

Instituto Politécnico de Coimbra
Instituto Superior de Contabilidade
e Administração de Coimbra

Previsão de emoções a partir de mineração de texto no Twitter

Hallan Patrick de Oliveira Roça

Previsão de emoções a partir de
mineração de texto no Twitter

Hallan Patrick de Oliveira Roça

ISCAC | 2021

Coimbra, julho de 2021



Instituto Politécnico de Coimbra
Instituto Superior de Contabilidade
e Administração de Coimbra

Hallan Patrick de Oliveira Roça

**Previsão de emoções a partir de
mineração de texto no Twitter**

Dissertação submetida ao Instituto Superior de Contabilidade e Administração de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de **Mestre em Análise de Dados e Sistemas de Apoio à Decisão**, realizada sob a orientação do Doutor Fernando Paulo Belfo e coorientação de Doutor António Trigo.

Coimbra, julho de 2021

TERMO DE RESPONSABILIDADE

Declaro ser o autor desta dissertação, que constitui um trabalho original e inédito, que nunca foi submetido a outra Instituição de ensino superior para obtenção de um grau acadêmico ou outra habilitação. Atesto ainda que todas as citações estão devidamente identificadas e que tenho consciência de que o plágio constitui uma grave falta de ética, que poderá resultar na anulação da presente dissertação.

PENSAMENTO

Minha motivação pessoal é gostar da área de comunicação social e ser entusiasmado por compreender o mundo por várias perspectivas. Na graduação em comunicação social, publicidade e marketing, conheci teorias e técnicas que me motivaram a gostar dessas áreas, incluindo política.

A comunicação tem um laço estreito com a psicologia, ciência essa que também tenho paixão e me interesse por continuar a estudar essa área. Portanto na realização desta Dissertação, me vejo no caminho em direção os meus objetivos.

Acredito que esse trabalho pode cooperar um mínimo para a ciência, na área de previsão de emoções em texto. E poder cooperar para a ciência, é um grande privilégio, no qual tenho a oportunidade de retribuir aos meus professores do ISCAC, da Unigran e a todos que dedicam suas vidas ao ensino, gratidão.

RESUMO

As emoções tiveram um papel fundamental para a adaptação e sobrevivência humana ao longo de toda história da nossa civilização, tendo ganho mais destaque na comunicação social nos últimos tempos com a criação e avanço das redes sociais. Na área de inteligência artificial, robôs têm simulado emoções humanas e por meio da visão computacional as expressões faciais têm sido detetadas. No marketing elas se destacam ao proporcionar experiências marcantes conectando o cliente ao consumidor.

Com propósito de cooperar no estudo das emoções, ainda que minimamente, esse trabalho objetivou desenvolver uma forma de identificar as emoções nos textos do Twitter, no cenário político brasileiro. Além de buscar padrões semânticos que diferenciasses as emoções, também foram aplicadas técnicas de classificação por meio de aprendizado de máquina.

A metodologia adotada neste estudo foi a CRISP-DM. A investigação baseou-se nas seis emoções primárias: tristeza, raiva, surpresa, medo, nojo e alegria. O conjunto de dados constituído por 3556 *tweets* foi retirado do Twitter, utilizando 190 *Tred Topics* sobre assuntos políticos do dia 22 de junho de 2020 até o dia 24 de agosto 2020.

Uma das principais contribuições deste estudo foi a forma como se fez a identificação e classificação das emoções nos textos foi desenvolvida a partir dos temas e gatilhos inatos. Alguns padrões se mostraram relevantes, como por exemplo, para raiva e nojo-desprezo onde os usuários tendem a utilizar palavras com significado de afastamento de grupo, enquanto para alegria mais palavras com sentimento de pertencimento a grupo. A precisão média do classificador foi de 75,61%. O estudo evidenciou a polarização política num sentido onde as pessoas se opuseram completamente umas contra as outras, discordando, demonstrando emoções maioritariamente pesadas, desejando o mal e até a morte umas das outras. Uma limitação importante foi a falta de mais poder computacional, limitando assim a afinação dos classificadores por meio de *GridSearch*. Outra limitação foi a de que alguns *tweets* possuíam duas emoções, como raiva e alegria, dificultado a sua classificação em uma única emoção. Algumas propostas de trabalhos futuros foram também avançadas.

Palavras-chave: mineração de texto, detecção de emoções, gatilhos emocionais, classificação de emoções, Twitter, Python, PLN, Linear SVM, política.

ABSTRACT

Emotions have played a fundamental role in human adaptation and survival throughout the history of our civilization, having gained more prominence in social communication in recent times with the creation and advancement of social networks. In the field of artificial intelligence, robots have simulated human emotions and through computer vision, facial expressions have been detected. In marketing, they stand out by providing outstanding experiences connecting the customer to the consumer.

In propouse to cooperate in the study of emotions, albeit minimally, this work aimed to develop a way to identify emotions in Twitter texts, in the Brazilian political scenario. In addition to seeking semantic patterns that differentiated emotions, classification techniques were also applied using machine learning.

The methodology adopted in this study was CRISP-DM. The investigation was based on six primary emotions: sadness, anger, surprise, fear, disgust and joy. The dataset consisting of 3556 tweets was taken from Twitter, using 190 Tred Topics on political issues from June 22, 2020 to August 24, 2020.

One of the main contributions of this study was the way in which emotions were identified and classified in texts, based on innate themes and triggers. Some patterns were shown to be relevant, for example, for anger and disgust-contempt, where users tend to use words with a meaning of distancing from the group, while for joy more words with a feeling of belonging to a group. The average accuracy of the classifier was 75.61%. The study evidenced the political polarization in a sense where people were completely opposed to each other, disagreeing, showing mostly heavy emotions, wishing each other harm and even death. An important limitation was the lack of more computational power, thus limiting the tuning of the classifiers through GridSearch. Another limitation was that some tweets had two emotions, such as anger and joy, making it difficult to classify them into a single emotion. Some proposals for future work were also put forward.

Keywords: text mining, emotion detection, emotional triggers, emotion classification, Twitter, Python, PNL, Linear SVM, politics.

ÍNDICE GERAL

1	Introdução	1
1.1	Apresentação do tema	1
1.2	Problema de investigação	1
1.3	Objetivos.....	2
1.4	Estrutura da dissertação	3
2	Revisão da literatura	4
2.1	Emoção	4
2.2	Sentimento	5
2.3	Como ocorrem as emoções.....	7
2.4	Emoções primárias	8
2.5	Emoções secundárias.....	11
2.6	A polaridade das emoções	12
2.7	Temas e gatilhos emocionais	13
2.8	Emoções e a comunicação escrita.....	14
2.9	A compreensão de cada emoção	16
2.10	Processamento de Linguagem Natural e Linguística.....	40
2.11	Aprendizado de máquina	43
2.12	Mineração de dados na análise de emoções e sentimentos.....	46
3	Metodologia	48
3.1	Metodologia CRISP-DM.....	48
3.2	Softwares utilizados	48
3.3	Campo de pesquisa e extração dos dados.....	49
3.4	Método de extração	51
3.5	Remoção de repetidos	52

3.6	Metodologia da classificação manual das emoções	52
3.7	Classificação manual e análise prática	58
3.8	Análise exploratória	68
3.9	Análise do significado das palavras	71
3.10	Limpeza dos dados	85
3.11	Formatação dos dados.....	86
4	Modelação e validação.....	92
4.1	Classificadores rejeitados	92
4.2	LinearSVC	93
4.3	SGDClassifier	95
4.4	Validação	96
4.5	Hiperparâmetros e GridSearch.....	98
5	Discussão dos resultados.....	100
5.1	Comportamentos emocionais expressos nas palavras.....	100
5.2	Ausência de algumas emoções.....	101
5.3	Significado das palavras por emoção	102
5.4	Discussão dos modelos.....	105
6	Conclusão.....	107
6.1	Objetivos alcançados	107
6.2	Contributos e implicações	110
6.3	Limitações e desafios	111
6.4	Trabalhos futuros	112
	REFERÊNCIAS BIBLIOGRÁFICAS	114
	APÊNDICES.....	Erro! Indicador não definido.
	APÊNDICE 1. Código Python para extração dos tweets	Erro! Indicador não definido.

APÊNDICE 2. Estrutura de código Python para classificação manual **Erro! Indicador não definido.**

APÊNDICE 3. Estrutura de código Python para modelação e validação..... **Erro! Indicador não definido.**

ÍNDICE DE TABELAS

Tabela 3.1: Quantidades e classes de tweets etiquetados manualmente.	69
Tabela 3.2: Estatísticas descritivas da quantidade de tokens do conjunto de emoções. .	71
Tabela 3.3: Frequência das 80 palavras mais usadas no corpus, por significado.	75
Tabela 3.4: Percentagem por significados de cada grupo de emoções.	84
Tabela 3.5: Abreviações encontradas no corpus.	87
Tabela 3.6: Aumento da precisão média depois da redução de dimensionalidade RFE.	89
Tabela 3.7: Extrato do dataset utilizado para o aprendizado de máquina.	90
Tabela 4.1: Detalhes da precisão do classificador LinearSVC.	95
Tabela 4.2: Detalhes da precisão do classificador SGDClassifier.	96
Tabela 4.3: Comparações dos hiperparâmetros dos classificadores.	98
Tabela 5.1: A presença das palavras mais comuns no dataset separado por emoção. ...	102

ÍNDICE DE FIGURAS

Figura 2.1: Representação Central de Componentes	15
Figura 3.1:.....	70
Figura 3.2: As 80 palavras mais frequentes dos conjuntos de emoções utilizados no aprendizado de máquina.	73
Figura 3.3: As 40 palavras mais frequentes da família das emoções de raiva.....	77
Figura 3.4: As 40 palavras mais frequentes da família das emoções de nojo-desprezo..	79
Figura 3.5: As 40 palavras mais frequentes da família das emoções da alegria.....	81
Figura 3.6: As 40 palavras mais frequentes dos tweets tidos como neutros.....	83
Figura 4.1: Machine Learning Map.....	93
Figura 4.2: Árvore de possibilidades para a modelação do dataset para o classificador LinearSVC.	94
Figura 4.3: Árvore de possibilidades para a modelação do SGDClassifier.....	96
Figura 4.4: Representação da validação cruzada estratificada StratifiedKFold.	97
Figura 5.1: Percentagem da quantidade de palavras com significado de pertencimento contra distanciamento de grupo para as três emoções, raiva, nojo-desprezo e alegria.	104

LISTA DE ABREVIATURAS E SIGLAS

APA – *American Psychological Association*

CNB – *Complement Naive Bayes*

CRISP-DM – *Cross Industry Standard Process for Data Mining*

DM – *Data Mining*

ETL – *Extract Transform Load*

KDD – *Knowledge Discovery in Databases*

KPCA – *Kernel Principal Component Analysis*

LDA – *Linear Discriminant Analysis*

NER – *Named Entity Recognition*

NLTK – *Natural Language Toolkit*

PCA – *Principal Component Analysis*

PLN – *Processamento de Linguagem Natural*

RFE – *Recursive Feature Elimination*

SGD – *Stochastic Gradient Descent*

SGDC – *Stochastic Gradient Descent Classifier*

SparsePCA – *Sparse Principal Components Analysis*

TF-IDF – *Term Frequency Inverse Document Frequency*

1 INTRODUÇÃO

Temos assistido a um aumento das faculdades mentais emocionais. Essa é uma possível consequência do crescimento exponencial da cultura industrial, como filmes, séries, livros, serviços como *streaming* de vídeo, etc. A indústria da arte nunca esteve tão desenvolvida como hoje. A partir da primeira revolução industrial, quando tudo era a preto e cinza, as cores e a personalização têm aparecido cada vez mais nas roupas, casas, carros e na sociedade de maneira geral. Estabelecimentos comerciais não são competitivos o suficiente se não forem esteticamente projetados ou se não tiverem no mínimo uma boa aparência interna e externa. Os ciclos da arte têm-se acelerado no design da arquitetura, moda e propaganda como nunca antes.

1.1 Apresentação do tema

Com a chegada das redes sociais a comunicação evoluiu, observando-se uma incrível extensão das relações sociais projetada nas redes sociais, possibilitando ser mensurada através de mineração de dados. E é essa a essência da atividade proposta aqui, analisar emoções na rede social Twitter.

A análise das emoções é um tema em ascensão, no marketing, na comunicação, na inteligência artificial, entre outras áreas. Por isso, é necessário aprofundar esta análise nos tempos que vivemos. A área de aplicação que está na base desta dissertação é o ambiente político, dizendo respeito ao debate realizado no Twitter por influenciadores, pessoas públicas e pessoas comuns.

Essa pesquisa não visou objetivos políticos, no sentido de trazer algum tipo de vantagem a alguma pessoa pública, partido ou qualquer outra entidade envolvida na política. Este trabalho pretendeu ser um estudo o mais isento e objetivo possível sobre uma determinada realidade. O investigador não pretendeu assumir qualquer estatuto de crítico ou guardião moral, mas apenas o de revelar e analisar as motivações que possam ter guiado as pessoas a terem certa afinidade por qualquer uma das posições, muitas vezes antagônicas, os sentimentos que com elas se relacionam e que se revelam através desse meio que é o Twitter.

1.2 Problema de investigação

Algumas vezes os comunicadores, provocam emoções indesejadas no público. Seja por meio da comunicação ou por fatores que fogem ao controle deles. Isso já causou muitos

problemas para as corporações, pessoas públicas ou entidades. Até os mais altos níveis como a destruição por completo da reputação de marcas e conseqüentemente da empresa em si.

Investigar sobre a previsão de emoções usando o texto que as pessoas registam numa rede social como o Twitter, possibilita que os comunicadores e outras partes interessadas pudessem monitorar a sua comunicação com base numa perspectiva emocional e por sua vez, ter um controle maior da sua comunicação.

1.3 Objetivos

Desde pensadores do mundo antigo como Heráclito e Parmênides, as perguntas sempre foram poderosas ferramentas para reflexão na busca por respostas a fim de resolver problemas. Por isso, os objetivos aqui foram feitos de forma semelhante, a partir de perguntas.

1.3.1 Análise dos significados, emoções, motivações e ações

- Quais são as emoções mais frequentes nos seres humanos?
- O que motiva as pessoas a se emocionarem?
- Há significados que despertam as emoções? Como temas da atualidade, pessoas públicas, ou elementos culturais, como música, filmes ou literatura?
- Há significados que despertam mais do que uma emoção simultaneamente?
- Quais ações as emoções despertam?
- Há palavras ou frases que são conseqüências de emoções?
- Há significados que despertam as emoções na mente coletiva ou na opinião pública?
- Motivações pragmáticas influenciam as emoções?
- Como ter um panorama pragmático de quem se comunica pelo Twitter?

1.3.2 Análise dos padrões identificados no texto

- Foi possível identificar algum padrão na estrutura das frases ou mesmo através dos significados expressos em palavras?
- Qual classificador melhor deteta esses padrões?

1.4 Estrutura da dissertação

O trabalho segue uma estrutura em grandes tópicos. A partir deste primeiro capítulo, o da Introdução, é feita uma revisão da literatura no capítulo 2, passando por cada uma das etapas necessárias, a fim de adquirir os conhecimentos para o desenvolvimento do trabalho. No capítulo 3, Metodologia, é explicado com detalhes as ferramentas utilizadas, o campo de pesquisa, o método de classificação manual dos dados, análise exploratória, análise do significado das palavras, pré-processamentos como a limpeza e formatação dos dados. No capítulo 4, Modelação e validação, é apresentada a explicação da abordagem mais técnica, passando pela seleção dos classificadores e da validação. No capítulo 5, Discussão dos resultados, são analisados os resultados obtidos do trabalho e o seu significado á luz da teoria apresentada. Por fim, no capítulo 6, Conclusão, são apresentados os objetivos alcançados, os principais contributos, limitações e algumas sugestões para trabalhos futuros que surgiram durante o trabalho.

2 REVISÃO DA LITERATURA

O objetivo do referencial teórico é a compreensão fundamental das áreas do conhecimento envolvidas neste trabalho a fim de executá-lo da melhor maneira possível, dentro do tempo estipulado. Na área da psicologia e afins, pretende-se rever o que a literatura apresenta para compreender o que são emoções primárias e secundárias e como estas ocorrem. Ainda nesta área, pretende-se rever o que são sentimentos e humor e compreender para cada uma das consideradas emoções primárias, em que contextos ocorrem, quais os gatilhos e temas que as levam a acontecer, as relações de umas com as outras e citando exemplos adequados.

Na área de Processamento de Linguagem Natural (PLN), busca-se compreender a teoria de como é feito este processo quando aplicado à mineração de texto e qual o envolvimento da linguística com essa área. Pretende-se obter o suporte teórico que permita reconhecer padrões na linguagem que possam ser aplicados no aprendizado de máquina para a classificação de textos com traços relativos às famílias das emoções abordadas.

Na área de aprendizado de máquina, busca-se compreender o suporte teórico básico sobre o assunto. Por fim, nos trabalhos relacionados, busca-se compreender o que os investigadores têm feito na área de reconhecimento de emoções e sentimentos, com intuito de desenvolver novas abordagens práticas e eficazes nesta área.

2.1 Emoção

A definição de emoção, de acordo com o dicionário de psicologia da *American Psychological Association* (APA), é um padrão de reação complexo, envolvendo elementos experienciais, comportamentais e fisiológicos, pelos quais um indivíduo tenta lidar com um assunto ou evento pessoalmente significativo. A qualidade específica da emoção, como medo ou vergonha, é determinada pelo significado específico do evento. Por exemplo, se o significado envolve ameaça, é provável que o medo seja gerado. Se o significado envolver desaprovação de outro, poderá ser gerada vergonha (APA, 2020).

Acredita-se que o contexto dos eventos que proporcionaram a evolução das emoções, estava relacionado com a importância para a sobrevivência do indivíduo. Atualmente, as emoções são despertadas tanto nesses eventos que configuram contextos de sobrevivência como em contextos sociais. Por exemplo, qualquer evento que represente risco de dano físico ou psicológico, pode despertar o medo. Esse é o tema desta emoção. Ekman (2011)

sustenta que as emoções determinam a qualidade de vida dos indivíduos, estando presentes nos relacionamentos humanos, como local de trabalho, nas amizades, nas relações com os membros da família, e até nos relacionamentos mais íntimos. As emoções tanto podem fazer com que o indivíduo se sinta bem e cooperar para o desenvolvimento individual e coletivo, como podem causar danos. Como é o caso da raiva que pode levar as pessoas a agirem de forma agressiva (Arruda, 2014; Ekman, 2011; Stets, 2006).

Numa abordagem sociológica, Kemper (1987) sustenta algumas características das emoções. Ele destaca que elas são fundamentais porque cada uma tem valor evolutivo de sobrevivência. Por exemplo, o medo e raiva incentivam os indivíduos a responder rapidamente quando estão em perigo. Elas aparecem nos estágios iniciais do desenvolvimento humano, são universalmente reconhecidos na face, têm respostas autonómicas únicas e emergem nas relações sociais (Kemper, 1987; Stets, 2006).

Davidson e Ekman (1994) identificam três funcionalidades das emoções: (A) as emoções têm propriedades motivacionais; (B) elas organizam os padrões comportamentais e fisiológicos para lidar com eventos que evocam emoções, interrompendo atividades menos importantes que estejam a decorrer e (C) os sinais emocionais informam os outros sobre as nossas motivações e motivam os outros a agir, o que é fundamental, nas interações sociais, ao longo da vida (Arruda, 2014; Davidson & Ekman, 1994).

Magalhães (2013) define emoção como uma resposta automática, intensa e rápida, inconsciente ou consciente, perante um estímulo e um impulso neuronal que leva o organismo a produzir uma ação. As funções da emoção encontram-se ligadas à adaptação e à expressão, funcionando como catalisador entre a conduta e o meio (Arruda, 2014; A. F. Magalhães, 2013).

Para António Damásio (1994), a emoção consiste numa variação psíquica e física, desencadeada por um estímulo, subjetivamente experimentada e automática e que coloca o indivíduo num estado de resposta ao estímulo. Um meio natural de avaliar o ambiente que nos rodeia e de reagir de forma adaptativa (Arruda, 2014; Damásio, 1994).

2.2 Sentimento

O sentimento é definido como, uma experiência fenomenal independente, subjetiva, que requer uma avaliação e independente das sensações, pensamentos ou imagens que o desperta. São avaliados como agradáveis ou desagradáveis. Eles podem ter qualidades intrapsíquicas mais específicas que diferencia um sentimento de outro, por exemplo, o

sentimento experimentado a partir do medo é diferente do que a raiva proporciona. A emoção normalmente envolve sentimento. Algumas diferenças entre eles são que, os sentimentos são puramente mentais, enquanto as emoções são projetadas para se envolverem com o mundo, de forma explícita ou implícita (APA, 2020).

Os sentimentos podem ser classificados em muitas categorias. Uma delas, é relativa aos sentimentos que dependem das emoções primárias e tem um envolvimento direto com elas. Para António Damásio, quando uma emoção é despertada, um sentimento relativo a ela é originado. Isso acontece num processo dinâmico de monitoramento que possibilita sentir as sensações corporais, que nesse caso, são geradas de uma determinada emoção. Por isso sentir essas sensações é a essência do sentimento que se origina de uma emoção. A experiência completa destes sentimentos, envolve a imagem mental desses padrões de sensações corporais, somado com uma outra imagem mental, que pode ser uma imagem visual de um rosto, a auditiva de uma melodia, etc. (Damásio, 1994).

Outra categoria de sentimentos, são os que se baseiam nas emoções que são pequenas variações das emoções primárias. São sentimentos originados por emoções universais sutis. Como por exemplo, a euforia e o êxtase são variações da felicidade, a melancolia e a ansiedade são variações da tristeza, o pânico e a timidez são variações do medo. Esses sentimentos, concretizam-se na conexão entre um estado cognitivo mais sutil e uma variação mais sutil de um estado emocional do corpo (Damásio, 1994).

Damásio define uma terceira categoria de sentimentos que é relativa a sentimentos de fundo, devido a estes não serem demasiados positivos e nem demasiados negativos e não terem origem em estados emocionais, como os outros citados. Essa categoria possui sentimentos agradáveis e desagradáveis numa intensidade mais leve. Relaciona-se com o humor, contribuindo provavelmente para um humor bom, mau ou indiferente. Esses sentimentos originam-se dos estados corporais de “fundo”, como o sentimento da própria vida, a sensação de existir. O estado de fundo do corpo é continuamente monitorado e o sentimento de fundo é a imagem desta paisagem do corpo, em momentos em que o corpo não se encontra agitado pela emoção, ou seja, nos intervalos entre as emoções. Por isso considera-se que são esses sentimentos e não os emocionais, que ocorrem com mais frequência ao longo da vida. Os sentimentos de fundo não mudam ao longo de horas e dias. Não se alteram com o fluxo e o refluxo do conteúdo dos pensamentos, sem eles, o âmago da nossa representação do eu seria destruída (Damásio, 1994).

No dicionário da APA há duas definições de humor. A primeira é qualquer estado emocional de curta duração, geralmente de baixa intensidade, como um humor alegre, um humor irritável. A outra é uma disposição para responder emocionalmente de uma maneira específica que pode durar horas, dias ou até semanas, sem que a pessoa saiba o que motivou o estado. Os humores diferem das emoções por falta de um objeto, por exemplo, a emoção da raiva pode ser despertada por um insulto, mas o humor de raiva pode surgir quando não se sabe o motivo de estar zangado ou o que provocou a raiva (APA, 2020).

Sentimento pode ser definido como a experiência subjetiva interna de uma emoção que é exclusiva para cada pessoa. O humor é um estado afetivo sem um objeto específico. Comparado às emoções, o humor geralmente, dura mais, tem menos intensidade e é mais difuso. Outros termos aparecem na literatura, entre os mais gerais, está o termo afeto, que é qualquer orientação avaliativa, positiva ou negativa em relação a um objeto (Stets, 2006).

2.3 Como ocorrem as emoções

De alguma forma, as emoções contemplam o envolvimento com o mundo. Por exemplo, quando uma pessoa está de frente com um cão feroz, esta situação pode configurar perigo de vida. Possivelmente o indivíduo sente medo e isso pode gerar batimentos cardíacos elevados e o sangue fluindo para grandes grupos musculares das pernas, preparando-o para fugir ou enfrentar a ameaça. Então temos uma situação que se envolve com o mundo e a partir disso gera uma emoção, que por fim interage com sistema endócrino e fisiológico. Note que nesta situação, houve um momento em que o cão apareceu e que ocorreu uma avaliação mental que possibilitou ao indivíduo, perceber que o cão era feroz e que representava um certo grau de perigo (Damásio, 1994; Ekman, 2011; Rhudy & Meagher, 2000).

O ser humano é dotado de um sistema de emoções e programado para agir de modo pré-organizado, diante de determinados eventos. Para uma emoção ser ativada, é necessário ser detetada uma ou mais características para gerar um estímulo específico e assim desencadear um processo emocional. Elas são desencadeadas após uma avaliação e detecção de características no ambiente, que ocorrem de forma automática. Através de mecanismos chamados de autoavaliadores, que examinam continuamente o ambiente fora de nossa consciência, em busca de temas e variações relevantes para a sobrevivência. O

conhecimento de um evento relevante é definido pelas experiências passadas individuais e pelo passado evolutivo, no qual o indivíduo sente que algo importante para o próprio bem-estar, está acontecendo. Neste caso, as características identificadas poderiam ser o porte físico do cão, juntamente com a possível aparência de feroz que ele tem. O cão também poderia ter mostrado os dentes, rosnando, levantado as orelhas e olhando atentamente em direção à pessoa. E tudo isso é o conjunto de características que gerou o estímulo. Os seres humanos são programados para reagirem com uma determinada emoção de acordo com o que se vivencia no mundo ou nos próprios corpos. Ao detectar algumas características, individualmente ou em conjunto, elas são processadas e em seguida detetadas pela amígdala, que é um componente do sistema límbico do cérebro. E, por sua vez, é desencadeada a ativação de um estado do corpo que caracteriza uma determinada emoção e a mesma coisa acontece com processamento cognitivo, que é alterado a fim de que corresponda ao estado desta emoção. O sistema emocional é rápido e eficiente de tal forma que não é necessário que seja reconhecido (Damásio, 1994; Ekman, 2011).

2.4 Emoções primárias

A diferença entre as emoções primárias e secundárias é que, as primárias são um conjunto de número limitado, manifestadas e reconhecidas universalmente através das culturas. Elas não precisam ser aprendidas ou construídas culturalmente. A lista de emoções primárias varia consoante os diferentes autores. Ela geralmente inclui medo, raiva, alegria, tristeza, nojo, desprezo e surpresa. Alguns autores também incluem vergonha, timidez e culpa (APA, 2020). Essas emoções também são chamadas de emoções básicas ou emoções centrais (APA, 2020).

Na sociologia, Kemper (1987) identificou um conjunto de quatro emoções primárias: raiva, medo, depressão e satisfação. Nota-se que são paralelas ao conjunto de emoções básicas de Ekman (1975). Turner (2007) fornece uma visão sociológica contemporânea dessas emoções. Ele sugere a seguinte lista: asserção-raiva, aversão-medo, desapontamento-tristeza e satisfação-felicidade. Esta lista tem paralelos com a lista de Ekman e Kemper (Kemper, 1987; Stets, 2006; Turner, 2007).

Os significados das emoções são socialmente construídos, pois quando aprendemos um termo para designar uma emoção, aprendemos segundo o jogo linguístico no qual aquela emoção é classificada, conforme as regras e padrões de uma determinada cultura. Assim,

diferentes culturas implicam variações de sentidos de determinadas emoções, e quais emoções são fundamentais, importantes para cada cultura. Portanto, a linguagem consiste numa ferramenta necessária para os significados que as emoções tomam (Arruda, 2014; Santos, 2003).

Episódios emocionais podem-se refletir no corpo em vários aspetos, por exemplo nas expressões faciais, vocalizações, linguagem corporal; indicadores fisiológicos, como a respiração, frequência cardíaca, pressão sanguínea, atividade electro dérmica, tensão muscular e neurológicos como os potenciais evocados, que são potenciais elétricos do sistema nervoso (Arriaga, 2010).

Para Damásio, as emoções primárias são inatas, pré-organizadas e estão ligadas ao processo adaptativo do ser humano, em termos de resposta, como estratégia de proteção ampliada consciente (Damásio, 1994).

Para Magalhães (2007), as emoções primárias são inatas, geneticamente programadas e se mostram essenciais para a sobrevivência do organismo. O desencadear de uma emoção acontece sem permissão, ou seja, não existe controlo na vontade de sentir, devido ao facto delas serem predeterminadas pela psicofisiologia.

O conjunto de emoções primárias defendido por Ekman são (Ekman & Friesen, 1975): tristeza, raiva, surpresa, medo, nojo e alegria. Alguns estudos foram realizados nos últimos sessenta anos e trazem evidências de indivíduos que nasceram cegos e podem expressar esse conjunto de emoções da mesma forma que indivíduos dotados de visão, através de expressões faciais. Esses estudos aplicam-se na manifestação de emoções espontâneas. Além destas, há uma pesquisa de Paul Ekman (1992), feita com tribos isoladas da civilização moderna que mostra as mesmas emoções básicas existentes, além de outras pesquisas feitas em diferentes países. Todos esses experimentos sustentam que esse conjunto de emoções são universais e inatas nos seres humanos, ou seja, são naturais e não derivam de construções sociais influenciadas pelo meio. Outra característica das emoções é que, em situações que envolve ativação de mais de uma emoção, elas podem alternar entre si a ponto de parecer que se está vivendo duas emoções juntas. Porém não há evidências científicas que comprovem que os seres humanos são capazes de vivenciar duas emoções simultaneamente (Ekman, 1992, 2011; Fulcher, 1942; Galati et al., 1997; Thompson, 1941).

O seguinte conjunto de emoções primárias (tristeza, raiva, surpresa, medo, nojo e alegria), são considerados por Ekman (1992) como famílias de emoções. Elas dão origem a várias outras emoções não primárias. Elas também possuem níveis de intensidade desde os mais brandos até os mais extremos, capazes de dominar as reações humanas. Um exemplo de família, pode ser a da emoção raiva. Foram identificadas mais de 60 expressões faciais diferentes relativas à raiva, onde cada emoção partilha certos padrões musculares que diferem da família de expressões de medo, expressões de nojo, etc. Existem cinco características compartilhadas entre cada uma das famílias das emoções primárias e que as diferencia de outros fenômenos afetivos: (1) sinais universais distintos; (2) presença em outros primatas; (3) fisiologia distinta; (4) eventos antecedentes universais e distintos; e (5) coerência entre os sistemas de resposta. (1) Os sinais estão relacionados as expressões faciais, sinalização vocal entre outros. Para cada emoção, mais de uma expressão universal foi identificada e também variam de intensidade. Contudo, dentro da família das emoções da alegria, muitas tem o mesmo sinal do sorriso, diferenciando nas intensidades e sinais vocais. (3) A fisiologia distinta significa que, ao serem disparadas, elas geram padrões físicos no corpo, atuando no sistema nervoso de forma diferente entre cada família. (4) Eventos antecedentes estão relacionados ao histórico adaptativo evolutivo das emoções, que deram origem aos temas emocionais. Mais à frente será abordado esse assunto de uma forma mais específica. (5) A coerência entre os sistemas de resposta está relacionada com as respostas independentes do sistema nervoso central e autônomo e as expressões faciais. Ekman defende que há uma conexão entre esses sistemas de resposta durante episódios emocionais genuínos, embora haja grandes diferenças individuais baseadas na aprendizagem social (Ekman, 1992, 1999, 2011; Ekman & Friesen, 1975).

Além das características compartilhadas entre as famílias de emoções, há outras que também diferenciam as emoções primárias de outros fenômenos afetivos. Como a (A) avaliação automática que acontece sem que se tenha consciência. O mecanismo de avaliação é capaz de operar automaticamente, exceto quando a avaliação se estende ao longo do tempo. O (B) início rápido, o que permite começar a lidar rapidamente com as atividades fundamentais, sem muito planejamento elaborado. Essa característica permite uma grande vantagem de sobrevivência e reação, especialmente em casos que se está de frente com um evento que exige resposta rápida. Uma (C) breve duração, geralmente duram alguns segundos. Em casos que os indivíduos relataram ter experimentado uma emoção

por 15 ou 20 minutos, as pessoas resumem em seu relatório verbal o que na verdade era uma série de episódios emocionais repetidos, porém diferentes. A (D) ocorrência de uma emoção não é permitida, elas ocorrem sem uma escolha prévia e com pouca ou nenhuma consciência da ocorrência e com alterações involuntárias da resposta na expressão e na fisiologia. Vivencia-se às emoções na forma natural com que elas acontecem. (E) Pensamentos distintos, imagens de memórias está relacionado a regulação que emoções específicas exercem sobre os pensamentos. Ekman defende que as emoções se evidenciam nas lembranças, imagens e expectativas. E que a relação entre emoções e pensamentos não seja apenas uma função do aprendizado social, devido a restrições biológicas impostas ao sistema cognitivo e ao sistema emocional. A (F) experiência subjetiva distinta da emoção, é como cada pessoa se sente quando vivencia e ao lembrar da experiência da emoção. Isso está para alguns, no centro do que é uma emoção. Engloba sensações físicas e outros sentimentos que são consequência do monitoramento e detecção das várias mudanças de resposta que ocorrem exclusivamente para cada emoção (Ekman, 1992, 1999, 2011; Ekman & Friesen, 1975).

2.5 Emoções secundárias

Emoção secundária é uma emoção que não é reconhecida ou manifestada universalmente entre culturas ou que requer experiência social para sua construção. Para alguns autores, o orgulho representa uma emoção secundária, decorrente da conjunção de uma emoção primária, alegria e uma reação pública favorável. Outras emoções secundárias incluem a inveja, a frustração e os ciúmes (APA, 2020).

Kemper (1987) sugere que há uma segunda camada de emoções que deriva de emoções primárias, às quais ele chama de emoções secundárias. Essas emoções são construídas socialmente na sua maior parte, ao contrário das emoções primárias que são inatas. Elas são aprendidas através da socialização em conexão com as emoções primárias, ou seja, uma combinação das emoções primárias pode levar às emoções secundárias. Por exemplo, a culpa é aprendida com o medo emocional primário. O medo e raiva podem levar a emoções secundárias de ódio, ciúme e inveja. Turner (2007) adota a analogia de Kemper (1987) de misturar emoções primárias para chegar a uma lista de emoções secundárias. Além do reconhecimento de que elas são socialmente construídas ou despertadas das interações sociais, geralmente, elas necessitam de alguma avaliação feita sobre si ou outra pessoa que esteja em determinada situação, uma avaliação consciente.

Algumas dessas emoções mais comuns investigadas, incluem a culpa, vergonha, empatia e simpatia (Kemper, 1987; Stets, 2006; Turner, 2007).

Para Damásio (1994), as emoções secundárias são desencadeadas a partir de um processo avaliativo consciente que ocorrem em relação a um evento ou a uma pessoa. Pode ser desencadeado por apenas projetar na mente determinado acontecimento, como imaginar o reencontro com um grande amigo que não se via a muito tempo, ou mesmo imaginar a morte deste. Esse processo avaliativo pode ser complexo, envolvendo imagens mentais que pode ser a aparência de uma determinada pessoa num determinado lugar, palavras e frases relativas a atributos, atividades, nomes, normas morais etc. Caso o episódio emocional ocorra em relação a uma pessoa, pode envolver uma infinidade de aspetos da relação do indivíduo com esta pessoa, além de reflexões sobre a situação atual e as consequências para si e para outros. Assim como as emoções primárias, as secundárias provocam reações no corpo. As respostas ao processo avaliativo são dirigidas ao corpo propriamente dito, resultando num estado fisiológico do corpo, mas também dirigidas ao próprio cérebro. Damásio argumenta que o circuito das emoções secundárias utiliza disposições pré-frontais adquiridas, distintas das disposições inatas necessárias para as emoções primárias e influenciadas, em sua construção, pelas primárias. Esses trabalham em conjunto com parte do circuito utilizado para desencadear as emoções primárias, no caso a amígdala entre outros. Em outras palavras, as emoções secundárias utilizam a maquinaria das emoções primárias (Damásio, 1994).

2.6 A polaridade das emoções

Para muitas pessoas as emoções têm sido fonte de motivação e prazer. Contudo, também existem as pessoas que não as admitem e ainda há aqueles que não procuram vivenciá-las, mas que não consideram essa experiência ruim na maioria dos casos. De um modo geral, grande número de cientistas da área das emoções separam as emoções entre positivas e negativas. Porém, essa questão não é consensual e ainda há um debate sobre isso. Muitas pessoas não gostam de viver experiências de medo e tomam atitudes extremas para as evitar. Há também quem goste de viver um medo que não é derivado de uma situação que represente perigo real, como assistir a filmes de terror, ou ler livros de terror e praticar alguns desportos seguros. E ainda há um comportamento extremo de algumas pessoas que procuram realmente colocar-se em perigo de vida nos desportos menos seguros. É provável que haja pessoas que gostem e que odeiem emoções

consideradas negativas, como a tristeza e o medo. Assim como existem as que gostam de ser surpreendidas e as que não gostam. Isso nos faz pensar sobre a polaridade das emoções na perspectiva do indivíduo. Pois se as emoções consideradas negativas possam ser positivas para as pessoas que gostam de vivenciá-las em determinadas situações, então nessa perspectiva, elas são positivas. Contudo, neste trabalho, quando o termo “emoção positiva” for utilizado, será para diferenciar as emoções relacionadas a alegria das outras (Ekman, 2011; Ekman & Friesen, 1975).

2.7 Temas e gatilhos emocionais

Na história da humanidade existiu e existem, várias tarefas e eventos fundamentais que as pessoas têm que lidar na vida cotidiana. São situações humanas universais, como conquistas, perdas ou frustrações, as quais se repetem inúmeras vezes. Por isso, acredita-se que as emoções evoluíram por seu valor adaptativo ao lidar com essas tarefas fundamentais da vida. Estas tarefas possuem cada uma, um significado de grande relevância para a sobrevivência e adaptação dos seres humanos. Os significados podem ser classificados por um conjunto de temas, que definem os aspectos interpessoal e intrapessoal, danos e benefícios inerentes a cada relacionamento pessoa-ambiente. Por isso, os temas são assuntos amplos considerados universais, inatos, uma herança biológica mediante a seleção natural. Que foram estabelecidos de acordo com o histórico adaptativo das emoções. E assim, proporcionaram reações universais a esses eventos, de forma a lidar com eles. Por exemplo, a perda de algo significativo para o indivíduo, é um antecedente da tristeza em muitas ou talvez em todas as culturas. Mas, algo significativo pode ser diferente de cultura para cultura, isso representa as diversas variações dos temas, que foram culturalmente aprendidas e que refletem a experiência individual de cada cultura. As variações desses temas são produto de várias influências. Como diferenças individuais na constituição biológica, diferentes experiências na aprendizagem e na natureza da ocasião específica do evento que desencadeia a emoção (Ekman, 1992, 1999; Ekman & Friesen, 1975; Richard S Lazarus, 1991).

Acredita-se que são como preparações, que permitem que alguns gatilhos se estabeleçam rapidamente, como sugere a experiência de Ohman. Portanto, não está provado que os temas atuem como gatilhos antes que o indivíduo tenha uma experiência emocional com o tema. Uma emoção pode ser despertada por vários eventos diferentes, mas o motivo principal que despertou a emoção está dentro do campo de abrangência do tema. O que

ativa uma determinada emoção é o gatilho. Esse é uma conexão estabelecida anteriormente. São capazes de despertar uma emoção por apenas lembrar-se do evento, ouvir ou ler um texto que remeta algo relacionado com o gatilho. Existem gatilhos comuns à maioria dos seres humanos, que são os considerados gatilhos inatos, ou seja, já nascem com as pessoas. A diferença de gatilhos inatos para temas, é que esses gatilhos são específicos, enquanto os temas são mais genéricos. Por exemplo, o tema associado ao medo é o perigo de dano físico ou psicológico, enquanto um gatilho inato pode ser um tipo de movimento, como a forma com que os répteis se movimentam. Existem ainda os gatilhos aprendidos, que foram estabelecidos por meio de uma experiência passada, em que determinado evento acontece acompanhado de um estado de excitação emocional, que fica na memória como um gatilho para despertar determinada emoção. Existem incontáveis maneiras dos gatilhos serem estabelecidos. E, geralmente estão associados a um tema de uma emoção. Portanto, acredita-se que as emoções possuem um ou mais temas ou variações deles e que cada indivíduo possui muitos gatilhos pessoais aprendidos e inatos, associados a cada tema (APA, 2020; Damásio, 1994; Ekman, 2011; Ekman & Friesen, 1975; Richard S Lazarus, 1991; Ohman, 2008).

2.8 Emoções e a comunicação escrita

As palavras são uma das formas de comunicação humana e as emoções são processos psicológicos e fisiológicos, como se apresentou anteriormente. A intenção neste trabalho é identificar palavras e textos que foram comunicados por indivíduos supostamente motivados por emoções, em parte ou na totalidade. Devido ao fato de que as palavras são apenas representações, elas possuem uma relação de significância com as emoções vivenciadas pelo indivíduo. Dessa forma, o processo de vivenciar a emoção e transmiti-la através de textos, é influenciado por uma quantidade de variáveis desconhecidas e que não é possível ter controle. Envolve dois conjuntos de variáveis essenciais, os aspetos internos, como a personalidade e aspetos ambientais, que prepara o cenário para a interpretação pessoal do evento significativo. Por exemplo, a forma como cada um reage à emoção e a intensidade com que elas se manifestam são diferentes em cada pessoa. Elas dependem da capacidade de comunicação do indivíduo, da quantidade de palavras conhecidas por ele, do grau de intensidade da emoção em si, do significado que o evento tem para o indivíduo, as palavras comumente utilizadas na cultura específica em que o evento ocorre. A interpretação do indivíduo em relação ao ambiente, da cultura do país, região e localidade que o indivíduo está inserido, da personalidade e características

comportamentais gerais e como elas se manifestam socialmente pelo indivíduo quando faz uso das redes sociais, a forma com que os utilizadores daquela determinada rede social geralmente comunicam-se, a quem se direciona o texto que a pessoa emocionada escreveu, do atraso entre o tempo em que a pessoa vivenciou a emoção, criou o texto e publicou nas redes sociais, etc. Portanto as emoções não podem ser reduzidas a palavras e textos, pois estes são apenas uma maneira de lidar e demonstrar a experiência vivenciada (Ekman, 2011; Richard S Lazarus, 1991).

Vivenciar uma emoção é um processo não linear e complexo, quando levado em consideração os fatores psicológicos, comunicacionais e ambientais. A abordagem complexa é vista também a partir de aspetos neurológicos, em que diferentes estruturas e processos interagem durante os episódios emocionais. Na Figura 2.1 vemos o modelo chamado *central representation of components* ou representação central de componentes (Scherer, 2009).

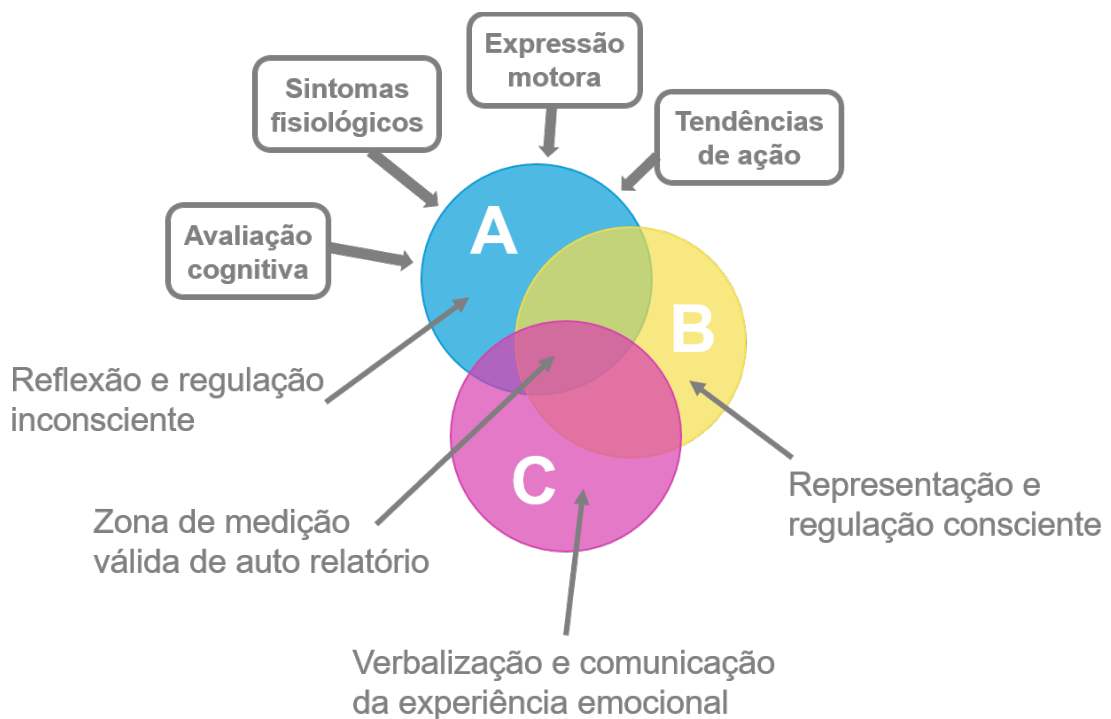


Figura 2.1: Representação Central de Componentes

Fonte: Scherer (2009)

A referida figura mostra os componentes responsáveis pela parte inconsciente, consciente até chegar na verbalização da experiência emocional não verbal. O primeiro círculo, (A) representa a reflexão pura dos processos emocionais. Esse componente é inconsciente, ele recebe uma grande quantidade de informação, de vários locais. Como por exemplo

numa situação de medo, as tendências de ação poderiam ser, fugir ou enfrentar, a expressão motora facial de medo, os sintomas fisiológicos como batimentos cardíacos acelerados e outros como discutidos anteriormente e o componente de autoavaliação cognitiva que deteta o perigo e continua ativo e sensível a qualquer mudança no ambiente. O componente (A) influencia os outros componentes que estão interligados. Como o círculo (B), que representa o surgimento da consciência no indivíduo sobre a emoção. Porém, no momento em que se vivencia uma emoção, nem sempre se têm consciência disso, especialmente no caso do medo, pois o foco cognitivo do indivíduo está concentrado na ameaça. O círculo (C) representa a categorização e verbalização da experiência emocional. Essa etapa de comunicação, é a que dedicamos os nossos esforços para realizar a presente pesquisa. Pois o indivíduo emocionado, comunica sua experiência emocional através de textos, sendo ele consciente ou não de suas emoções (Berkowitz, 1989; Ekman, 2011; Scherer, 2009).

2.9 A compreensão de cada emoção

Com objetivo de avaliar o estado sobre o campo de investigação das emoções, em 2016 Paul Ekman (2016) enviou um questionário a 248 cientistas que estudam as emoções por uma abordagem quantitativa, selecionados de acordo com determinados critérios. Uma parte do questionário pretendeu ajudar a verificar quais rótulos de emoção (de uma lista de 18) deveriam ser considerados empiricamente estabelecidos. De entre os participantes que responderam, houve alta concordância sobre cinco emoções, todas descritas por Darwin e Wundt (Ekman, 2016): raiva (91%), medo (90%), nojo (86%), tristeza (80%) e alegria (76%). Vergonha, surpresa e constrangimento foram aprovadas em 40% a 50%. Também houve alta concordância sobre outros assuntos, como o humor específico estar relacionado a uma emoção específica. Como a raiva estar relacionada à irritabilidade (88%), traços de personalidade específicos estarem relacionados à emoções específicas, como medo da timidez (82%) (Ekman, 2016).

Este trabalho limitou o conjunto das famílias das emoções primárias estudadas às seguintes: raiva, medo, nojo, tristeza, alegria. Sendo a alegria um conjunto de emoções distintas, porém da mesma família. É um desafio compreender a comunicação escrita numa abordagem afetiva devido à alta complexidade e inúmeras formas com que as emoções se manifestam nos seres humanos. Por isso, não temos a pretensão de concluir que, um indivíduo que escreveu determinado texto, estava definitivamente motivado por

uma emoção primária, pois isso não é possível ser feito apenas com análise linguística. Por isso, será considerado qualquer traço que represente as famílias das emoções citadas, independente da intensidade e se elas foram ou não influenciadas pelo humor, estados mentais, sentimentos, traços de personalidade, emoções, entre outros estados afetivos. Portanto nosso objetivo é compreender a comunicação através da ótica das famílias das emoções primárias e detectar quais textos que possuem relação com as emoções primárias.

Em seguida, este capítulo abordará as características de cada família das emoções, predominantemente pela abordagem de Paul Ekman (1992, 2011).

2.9.1 Tristeza

Considera-se a perda como o tema central que é capaz de gerar os estímulos que despertam a tristeza. Muitos gatilhos estão relacionados a tristeza como a rejeição de um amigo, namorada, cônjuge ou um familiar; e também a perda de uma pessoa próxima ou familiar que pode ser o distanciamento ou morte; a perda da autoestima devido a algum fracasso ou relacionado a aparência; a perda da admiração ou elogio de um superior ou as pessoas próximas ou mesmo de um público; a perda da saúde; a perda de alguma parte do corpo ou de função devido a um acidente ou doença; a perda de um objeto querido, etc. (Ekman, 2011; Ekman & Friesen, 1975).

A tristeza pode ser vivenciada mesmo quando não ocorre uma perda, mas esta não deixa de ser o tema que configura toda situação. Por exemplo, o indivíduo está aguardando uma notícia decisiva em relação à vitalidade de um ente querido, antecipando internamente essa perda, que representa algo devastador na vida desse indivíduo. Mas, após receber a notícia de que o seu ente querido está bem, há um momento de alívio e depois toda angústia que estava reprimida irrompe. Isso acontece quando o indivíduo se sente seguro para sentir essa emoção. A dor antecipada, mas contida, é agora expressa e a pessoa sente tanta dor como alívio, mesmo que a perda não se concretizou. Portanto, a interpretação e detecção da tristeza, toma uma dimensão mais desafiadora, pois agora além de ter que detectar o possível estímulo de perda, deve ser analisado se a tristeza demonstrada é derivada da antecipação dela (Ekman, 2011).

2.9.1.1 Diferenciação de angústia e tristeza

Alguns autores sugerem que a tristeza tem duas vertentes: tristeza e angústia. Em momentos de angústia, geralmente há protesto em relação ao que foi perdido. É como se fosse um estágio em que não se aceita a perda e que o indivíduo tenta lidar com a fonte da perda, mas é diferente de raiva. A tristeza se diferencia por ser uma fase de aceitação e desesperança. Um comportamento mais passivo onde não há luta para aceitar a perda. Acontece quando o indivíduo entende que nada pode ser feito para recuperar o que foi perdido. Quando abordamos a tristeza do luto, essas emoções podem se alternar e se repetir de forma cíclica durante o período do luto. Enquanto a angústia é protestante, a posterior tristeza é resignada. Em seguida, volta a angústia numa tentativa de recuperar a perda e depois a tristeza. É dessa forma que esse ciclo se pode continuar a repetir (Darwin, 1965; Ekman, 2011; Ekman & Friesen, 1975).

2.9.1.2 A mensagem e o compartilhamento da tristeza

Muitos seres humanos compartilham a tristeza com o objetivo de pedir ajuda a outros. Esse é um dos propósitos dessa emoção. Mostrar o seu sofrimento e obter a ajuda, sendo esta física, financeira ou psicológica. E quem ajuda, sente-se bem, pois ajudar geralmente traz alegria a quem ajuda e pode diminuir o sofrimento de quem é ajudado. Esse propósito também pode se estender ao coletivo, a fim de que a tristeza seja vivenciada em grupo, além de possibilitar ajuda coletiva. Alguns podem não vivenciar a angústia verdadeiramente se não estiverem na presença de pessoas que compartilhem da mesma tristeza. O significado do evento que entristece, pode-se tornar mais rico quando compartilhado, do ponto de vista da pessoa que compartilha. Observar as reações dos que sofrem junto deles, pode ser consolador. O compartilhamento da tristeza fortalece a justificativa de que as pessoas queiram expressar suas angústias e tristezas por meio da comunicação nas redes sociais. Há outras pessoas que não querem mostrar o próprio sofrimento. Muitos não querem ser ajudados quando estão tristes ou angustiados, preferindo a solidão a fim de não serem vistos nesta situação. Pois podem ficar envergonhados da fraqueza de tamanha dependência em relação a uma pessoa. Além disso, há uma relação social em algumas culturas, de que homens não devem sentir ou demonstrar angústia em público, o que pode reforçar o comportamento dos homens em não compartilhar suas angústias (Ekman, 2011; Ekman & Friesen, 1975).

2.9.1.3 Relação da angústia e da tristeza com a raiva

A angústia e a tristeza diferem da raiva. Num processo de luto intenso, há momentos em que a raiva pode ser ativada na angústia protestante. Raiva em relação à vida, a Deus, à pessoa ou à coisa que provocou a perda, à pessoa que morreu por ter morrido e especialmente se colocou a vida em risco de alguma forma. A pessoa que está de luto pode sentir raiva de si mesma por não ter feito nada para impedir, mesmo se racionalmente nada poderia ter sido feito para impedir a morte do ente querido. Pode sentir raiva também por não ter expressado algum sentimento importante ou por se sentir vulnerável. Algumas vezes a raiva também pode surgir como uma máscara para cobrir a tristeza em público, a fim de esconder a fraqueza e compactuar com o que determinadas culturas impõem. Especificamente neste caso, a raiva pode ser exagerada ou simulada para esconder a tristeza. A perda é o tema central da tristeza, mas em uma situação em que essa perda não seja permanente, como no caso da rejeição, a raiva pode surgir de maneira alternada com a tristeza. A raiva é dominante até ao momento em que o indivíduo sente a perda, a partir daí é ativada a tristeza (Ekman, 2011; Ekman & Friesen, 1975).

2.9.2 Raiva

Encontrar um tema central que desperte a raiva é um assunto que ainda está em debate entre os cientistas das emoções. Porém, um tema que se encaixa na maioria dos casos é a interferência. Quando alguém é interferido em algo que pretende fazer ou que estava fazendo. A raiva pode ser mais intensa se quem é interferido entender que houve intenção de quem interfere, ou seja, que a interferência foi feita de forma consciente e opcional. Alguns psicólogos do desenvolvimento constataram que para provocar raiva de maneira efetiva em bebês, é necessário segurar seus braços de modo que ele não consiga se libertar, isso serve como metáfora para a interferência. Diferentemente da tristeza, a qual possui apenas um tema, na raiva, é possível que haja algumas variações desse tema. De entre as variações do tema interferência, estão a frustração, a rejeição, a própria raiva de outra pessoa como causadora de raiva. Por exemplo, quando uma pessoa manifesta a sua raiva contra outra, a raiva pode ser ativada nesta outra pessoa ou em outras pessoas que estão próximas. Também pode ser ativada quando algo ou alguém ameaça causar dano, opinião ou conceito que é contrário ao que se acredita (Ekman, 2011; Ekman & Friesen, 1975; Sternberg & Campos, 1990).

2.9.2.1 *Variações do tema*

Há uma quantidade variada e desconhecida de causas da raiva. De entre as causas mais comuns está a frustração, que pode ser derivada de três fontes, em relação a alguém, a um objeto ou a si mesmo. Quando o objeto ou bem deixa de funcionar, não cumprindo as expectativas desejadas. Quando a frustração é derivada de alguém, podendo ser porque essa pessoa não fez o que se esperava ou fez ao contrário. A frustração em relação a si mesmo, que pode ser derivada de uma falha na capacidade, como uma falha de memória. Essas três fontes de frustração possuem em comum a interferência em alguma atividade ou na busca pelos objetivos do indivíduo. Como abordado anteriormente, a rejeição pode causar raiva além de tristeza. A raiva de outra pessoa pode gerar raiva no indivíduo. Quando alguém motivado pela raiva, direciona uma ação a outra pessoa, geralmente há uma reação furiosa por parte de quem sofreu a ação. A reação raivosa pode ser mais intensa quando o indivíduo se sente vítima de ações injustas ou injustificadas. Quando algo causa ou poderá causar dano é provável que ative a raiva. Quando essa ameaça de dano físico vem de uma pessoa, a raiva e o medo são reações prováveis. Se alguém ameaça causar dano psicológico, insultando, denegrindo a aparência ou desempenho, isso também pode despertar raiva e medo. Pode-se sentir raiva de ideias, ações ou crenças que se apresentem contrárias a opinião ou ofensivas ou que viole os valores morais do sujeito, a partir da interpretação dele mesmo. Apenas ler um texto, interpretar alguma forma de mensagem que é contrária, numa rede social, blogue, propaganda, jornal entre outros, pode gerar raiva. Ela também é sentida não somente pelo conteúdo em si, mas principalmente pela pessoa que falou ou agiu de forma contrária, mesmo que essa pessoa seja totalmente desconhecida. O desapontamento também pode gerar raiva, desaponta-se em relação ao comportamento de uma pessoa quando ela faz algo diferente ou contrário do que se esperava. Muitas vezes acontece em relação às pessoas de quem se tem maior proximidade ou intimidade. Nas primeiras fases do relacionamento romântico por exemplo, os cônjuges fantasiam o comportamento um do outro, de forma a criar uma expectativa e quando não é satisfeita, pode provocar raiva (Berkowitz, 1989; Ekman, 2011; Ekman & Friesen, 1975).

2.9.2.2 *Variações de tipos e intensidades da raiva*

A raiva possui muitos gatilhos inatos, que são motivações para ativá-la e baseiam-se em diferentes tipos de sentimentos de raiva, por isso possuem experiências diferentes com intensidades diferentes. Esses sentimentos variam desde o aborrecimento leve até a fúria.

A indignação é a raiva que está relacionada com a moral e o sujeito que se enfurece, geralmente, esta agindo de forma hipócrita. E muitas vezes o rótulo “hipócrita” é utilizado quando se discorda dos valores morais do sujeito enfurecido; o mau humor é a raiva passiva; a exasperação refere-se a ter a paciência provada em excesso. A vingança é um tipo de ação furiosa cometida, em geral, após um período de reflexão a respeito da ofensa, às vezes com maior intensidade que o ato que a provocou. Quando um ressentimento acontece por um período curto, é mais um integrante da família das emoções associadas à raiva. Quando há um ressentimento duradouro, ou seja, um rancor, pode surgir a raiva quando se vê ou pensa na pessoa que provocou o rancor. Os cientistas não sabem ao certo o quanto o meio e a genética influenciam na intensidade das emoções para cada pessoa. Sabe-se que alguns ficam com raiva mais rápido e com mais intensidade, alguns podem não ter a capacidade de ficar extremamente raivosos e talvez nunca tenham vivido um estado intenso de fúria (Ekman, 2011; Ekman & Friesen, 1975; Miles & Carey, 1997; Plomin et al., 1990).

2.9.2.3 A mensagem da raiva

Enquanto a mensagem que a tristeza comunica é um pedido de ajuda, na raiva não há uma frase que resuma completamente a comunicação dessa emoção. No entanto, a frase “Cai fora do meu caminho” parece capturar parte dos gatilhos da raiva. A raiva informa aos outros que existe um problema, é uma forma de intimidação a alguém que interfere. Porém não corresponde a raiva gerada da própria raiva de outra pessoa ou sentida por ler algo oposto ao que se acredita ser correto. Além disso a raiva nem sempre é o sentimento de querer que a pessoa ofensiva se distancie, mas um sentimento de querer ferir aquela pessoa (Ekman, 2011; Ekman & Friesen, 1975).

2.9.2.4 Relação da raiva com outras emoções

Muitas vezes o medo vem antes da raiva, como o medo de dano físico ou psicológico. Mas, se o ameaçado for mais poderoso do que quem ameaça, poderia ser mais provável sentir desprezo. Uma pessoa pode sentir raiva e alegria, mostrando desprezo ou presunção. Quando alguém que é considerado inferior, age com rejeição, pode provocar desprezo, surpresa e até raiva em outros. Existe o medo da própria raiva, por não saber como poderia agir, quando se está dominado por essa emoção, isso está relacionado a perder o controle e causar dano físico ou psicológico em outros. Em outras situações as pessoas podem vivenciar aversão e raiva, que gera repulsa do alvo que atacam. Essa

repugnância também pode ser sentida em relação a si próprio por ter ficado com raiva ou não ter conseguido se controlar. Culpa e vergonha também podem ser prováveis por ter sentido raiva (Ekman, 2011; Ekman & Friesen, 1975).

2.9.2.5 A força e a atitude da raiva

Aparentemente, uma atitude comum que as pessoas dominadas pela raiva têm é a de fazer o mal ao alvo ou até mesmo a um objeto. Agressão física ou psicológica através de palavras raivosas, gritadas ou proferidas agressivamente. Uma emoção capaz de dominar as atitudes dos seres humanos, evidentemente é uma emoção muito forte. Muitas pessoas se arrependem do que disseram em um momento de raiva. Pedem desculpas e explicam que estavam dominadas pela raiva e não queriam dizer o que foi dito. As suas atitudes e crenças verdadeiras foram distorcidas pelo poder dessa emoção. Um exemplo é a frase comum “Perdi a cabeça”. Assim como há pessoas que apreciam a tristeza, outras podem apreciar a raiva. Querem viver essa emoção intensa, procuram bons argumentos para justificar sua raiva. Trocas hostis e ataques verbais são sempre excitantes e gratificantes. Algumas pessoas apreciam até uma briga física (Ekman, 2011; Ekman & Friesen, 1975).

2.9.3 Surpresa

A forma como a surpresa é despertada está relacionada a um evento que acontece sem o conhecimento de que iria acontecer. Essa emoção só é vivenciada de forma verdadeira quando o evento que surpreende é imprevisível. Algo que não se esperava que acontecesse da forma como aconteceu ou algo que aconteceu de forma contrária ao que se esperava. Quando o evento se desenvolve lentamente, geralmente não há surpresa, pois a mente do indivíduo já previu o que irá acontecer. Muitas vezes a surpresa está relacionada a eventos que acontecem fora de contexto, como algo que ocorreu em um local que não deveria ter acontecido neste local ou que normalmente acontece em outros locais. Quase tudo pode ser surpreendente, desde que seja inesperado e relevante. Uma visão, som, cheiro, sabor ou toque. Não são apenas sensações físicas que causam surpresa, uma ideia, comentário, sugestão incorreta ou desconhecida. Um conhecimento que se adquire, por meio de alguém ou livro, vídeo, postagem em rede social, pela observação no mundo real, etc. O mesmo pode acontecer com um pensamento novo ou sentimento que não havia sentido (Ekman, 2011; Ekman & Friesen, 1975).

2.9.3.1 O tempo da surpresa

A surpresa é uma emoção que começa e termina de forma muito rápida, é a mais breve das emoções. Se o indivíduo tiver tempo para pensar no evento que o surpreende e considerar se está ou não surpreso, então já não estará mais, pois já terá passado. A surpresa passa assim que se entende o que está acontecendo, no momento em que a pessoa adquire consciência da natureza do evento surpreendente. Uma pessoa nunca ficará surpresa por muito tempo, a menos que o evento revele novos elementos que causem uma nova surpresa. Na maioria das vezes não há tempo de mobilizar esforços a fim de controlar o comportamento em caso de surpresa. Mas, se o indivíduo tiver tempo para antecipar o evento e o fizer corretamente, então é provável que ele não se surpreenda quando acontecer. As outras emoções podem ser tão breves quanto a surpresa, mas também podem ser muito mais longas do que ela. A surpresa não possui esta característica de ser variável na duração, o tempo dela é fixo, o que a diferencia de todas as outras emoções. Para alguns cientistas das emoções, esta característica levanta a questão de a surpresa ser sequer uma emoção. Há uma confusão entre surpresa e espanto. Contudo existem diferenças para sustentar que o espanto é um reflexo físico e não uma emoção derivada da surpresa. A fúria é um estado mais intenso da raiva, assim como terror é do medo e assim por diante. As expressões faciais desses estados intensos de cada emoção, se assemelham ao estado mais moderado delas, diferenciando na intensidade das contrações musculares. Porém o espanto não é um estado mais intenso de surpresa, uma vez que as expressões faciais do espanto são totalmente diferentes da surpresa moderada. Outras evidências como o tempo ainda mais reduzido do que a surpresa. A reação ao espanto é inevitável, mesmo sabendo que determinado evento iria acontecer, diferente da surpresa. O espanto está associado ao reflexo físico que se tem ao fechar os olhos quando um objeto está em curso, a caminho de se chocar com o rosto de forma inevitável ou quando rebenta uma bomba e a pessoa reage ao som num ato involuntário (Ekman, 2011; Ekman et al., 1985; Ekman & Friesen, 1975).

2.9.3.2 A relação da surpresa com outras emoções e a neutralidade

Quando o evento que causou a surpresa não é compreendido e resistir a interpretação, a surpresa pode permanecer, mas a pessoa pode ficar desorientada com medo e confusa. Por exemplo, suponhamos que uma mulher é recebida na porta pelo marido que ela acreditava ter sido morto na guerra. Ela ficará surpresa, com medo e confusa, mas a surpresa termina quando o marido dá uma explicação racional. A partir do momento que

se entende o que está acontecendo, a surpresa mistura-se ou dá lugar a outra emoção, como medo, diversão, alívio, raiva, aversão, entre outras, dependendo do significado que o evento surpreendente tem para o indivíduo. Mas também pode ser seguida de nenhuma emoção, se for determinado que o evento não tem consequências ou não é relevante. Há uma discussão entre os cientistas sobre a surpresa ser uma emoção neutra e somente a emoção posterior é que seria uma emoção positiva ou negativa. Paul Ekman acredita que não há polaridade nas emoções, uma vez que há muitas pessoas que gostam de vivenciar as consideradas emoções negativas, tristeza, raiva, medo, aversão e também as que não gostam, como discutido anteriormente. Ekman defende que a surpresa não é diferente dessas emoções, pois no momento antes de descobrir o que está acontecendo, a surpresa pode ser agradável ou não, dependendo se a pessoa gosta desta emoção. Algumas pessoas nunca querem ser surpreendidas, mesmo se o evento é positivo. Algumas pessoas podem até tentar evitar festas de aniversário surpresa, por não gostarem de ser surpreendidas. Essas planeiam o máximo que puderem para evitar o acaso. E há aquelas que gostam e deixam muitas coisas sem planejamento, para que possam vivenciar o inesperado. Podem até buscar experiências em que tendem a ser surpreendidas. Outra situação que gera o questionamento da surpresa como emoção, é que muitas pesquisas mostram uma grande dificuldade dos entrevistados em diferenciar as expressões faciais de medo e de surpresa, isso ocorreu entre vários países e uma tribo isolada da cultura contemporânea em Nova Guiné (Ekman, 2011; Ekman & Friesen, 1975).

2.9.4 Medo

O tema e os gatilhos do medo estão relacionados em torno do perigo de dano físico ou psicológico. Qualquer coisa compreendida e assimilada que possa fazer mal, é potencial para ativar o medo. Nesse processo de aprendizagem, não é necessário que o indivíduo sofra o dano, apenas ver outro a sofrer ou ter conhecimento que algo causou dano ou que possa causar dano. Seres, objetos ou eventos de que não se tem a certeza que causem dano e que possuem uma ou poucas características associadas com aprendizagem passadas, podem ativar o medo, mesmo que seja um medo sem fundamento. Muitos cientistas acreditam que há gatilhos desaprendidos, ou seja, a capacidade de ter medo de algo que não foi apreendido como causador de medo, um conhecimento que já existe nos seres humanos, como discutido anteriormente. Assim como a restrição física é um gatilho para raiva. O medo é responsável por diversos distúrbios emocionais, como as fobias que se caracterizam pelo medo de eventos ou situações interpessoais, como morte, ferimento,

doença, sangue, animais e situações como multidões ou espaços fechados (Ekman, 2011; Ekman & Friesen, 1975; Ohman, 2008).

2.9.4.1 Gatilhos do medo

Podem-se destacar aqui alguns gatilhos inatos ou desaprendidos, como por exemplo, um objeto arremessado através do espaço que rapidamente atingirá o indivíduo ou a perda inesperada de apoio, causando a sua queda. Ver animais assustadores como cobras, aranhas e répteis em geral, pode despertar o medo. O tamanho é um gatilho, como animais de grande porte, uma grande envergadura, como uma águia em voo, determinados sons, como um rugido. A própria ameaça de dor física é um gatilho desaprendido para o medo. No entanto, quando o indivíduo está a sentir a dor, não está necessariamente a sentir medo. É importante destacar que ao detetar essas características, o medo pode ser ativado, mesmo sem compreender qual animal está ameaçando (Damásio, 1994; Ekman, 2011; Ekman & Friesen, 1975; Ohman, 2008).

2.9.4.2 Presença e ausência de medos incomuns

Apesar de existirem gatilhos inatos, uma grande quantidade de pessoas parece não ter medo de cobras. Ao contrário, podem até gostar do contato físico, mesmo com cobras venenosas. Há sempre indivíduos diferentes em aspetos de comportamento da maioria das pessoas, assim também como em relação aos estímulos que despertam emoções ou em reações comuns. Por isso, existem indivíduos que temem coisas que não representam um risco real de dano físico ou psicológico, correspondendo a medos sem fundamento. Como o medo infantil do escuro, outro exemplo é que alguns adultos podem ter medo de ligar elétrodos no peito para medir a atividade cardíaca, num exame de eletrocardiograma, com medo de choque elétrico. Há muitos medos infundados, infantis, que são levados para a vida adulta, assim como há muitos medos desenvolvidos na vida adulta (Ekman, 2011; Ekman & Friesen, 1975).

2.9.4.3 Reações que o medo pode provocar

Para haver medo, geralmente há uma ou mais ameaças. Ameaças diferentes resultam em comportamentos diferentes. O medo gerado a partir de uma ameaça imediata, ou seja, que ocorre no tempo presente, de forma instantânea sem intervalo de tempo, conduz a dois tipos de ações de acordo com a ameaça, que são a paralisia ou a fuga. A reação a esta ameaça imediata é muitas vezes analgésica, reduzindo sensações de dor. Quando a

situação é de uma ameaça que está por vir, num tempo próximo, o indivíduo fica num estado de receio, uma preocupação extrema, que conduz a vigilância atenta na ameaça, onde é dedicado o foco total. Nesse momento há tensão muscular e um aumento na sensação da dor. Como todas emoções, o medo provoca reações no corpo, como batimento cardíaco acelerado e o sangue a correr para os grandes músculos das pernas (Damásio, 1994; Ekman, 2011; Ekman & Friesen, 1975; Rhudy & Meagher, 2000).

2.9.4.4 O envolvimento do medo com outras emoções

No momento que o indivíduo sente medo e está consciente disso, normalmente é difícil de sentir ou pensar em algo que não esteja relacionado com a ameaça, pois a mente e a atenção permanecem concentradas nela, até que seja eliminada ou abrandada. Por isso, na maioria das vezes, o medo se torna uma emoção dominante quando vivenciada. Ter raiva da ameaça é uma reação normal após sentir medo, como já foi discutido. Também, não é incomum sentir medo e raiva alternando entre si. Se é uma pessoa mais poderosa que ameaça, a tendência é sentir medo em vez de raiva. Porém, em algum momento, ou depois de escapar da ameaça é provável sentir raiva. O sujeito que vivenciou o medo pode sentir raiva de si mesmo, se ele acreditar que poderia lidar com a situação. Pelo mesmo motivo, também pode sentir aversão de si mesmo. Quando não há nada a se fazer diante do perigo, a não ser esperar o dano, a tendência é que sejam despertados níveis mais altos de medo, como o pavor. Mas, quando há algo para fazer a fim de evitar ou enfrentar o perigo, pode-se não sentir medo (Berkowitz, 1989; Ekman, 2011; Ekman & Friesen, 1975).

2.9.4.5 Medo como prazer

Para muitas pessoas, as emoções são experiências intensas e marcantes. Há pessoas que gostam de vivenciar as consideradas emoções negativas, como já foi visto. Algumas pessoas parecem gostar de sentir medo. A indústria cinematográfica explora esta atração pelo medo quando utiliza histórias como romances e filmes que assustam, e de fato são muito apreciados. No entanto, ler um livro ou assistir um filme, não é uma situação real, em que o indivíduo está em perigo de dano e ele sabe disso. Porém, existem pessoas que vão além, procuram experiências amedrontadoras e arriscam a vida com desportos extremos (Ekman, 2011; Ekman & Friesen, 1975).

2.9.5 Nojo-desprezo

Parte da literatura traz o termo aversão e nojo como sinónimos e dentro desta família há uma variação, que é o desprezo. Em seguida, serão discutidos os detalhes desta família de emoções.

2.9.5.1 Tema da aversão

O tema que desperta aversão está associado à incorporação oral, ou seja, ingerir algo desagradável ou ofensivo, que possa contaminar e causar dano. Geralmente o objetivo comum está relacionado a evitar, ou se afastar, ou mesmo expelir para fora do corpo. A aversão como as outras emoções, varia na intensidade. Quando a aversão é menos intensa, considera-se como a repugnância, que muitas vezes são exploradas em filmes de terror, piadas de *stand up comedy*, programas de televisão, programas infantis e brinquedos, pois são experiências que causam um certo fascínio em muitas crianças, jovens e em adultos. É importante notar que somente a partir dos quatro anos é que as crianças desenvolvem a noção de aversão. Porém, quando a aversão é intensa, as sensações são desagradáveis e podem conduzir à náusea. Em algumas vezes a raiva pode alternar com a aversão, por exemplo quando a pessoa enjoada sente raiva de algo ou alguém por ter provocado essa emoção ruim. Alguns tipos de raiva também podem preceder, suceder ou alternar com aversão, como é o caso da indignação (Ekman, 2011; Ekman & Friesen, 1975; Miller, 1997).

2.9.5.2 Diferentes gatilhos

O próprio gosto de algo desagradável, um cheiro que se quer impedir de entrar no nariz ou impedir de provocar a aversão. A visão de algo considerado desagradável, ou que lembre o gosto, pode provocar aversão. Assim com os sons, se estiver relacionado a um evento abominável. O tato, sentir através da pele uma sensação de algo desagradável, como um fluido viscoso também poderá provocar essa emoção. Outros vivenciam aversão quando veem uma pessoa deformada, aleijada ou considerada muito feia. Também ao ver uma pessoa machucada, com ferimentos expostos ou algum tipo doença na pele ou fluidos que assumam características desagradáveis. O relato de uma cirurgia ou acidente. Os gatilhos para aversão estendem-se além do tema de incorporação oral e da contaminação. As ações consideradas abomináveis de uma pessoa, podem causar aversão, assim como as ideias e ideologias dela. Alguém que maltrata um animal pode ser objeto de aversão. Alguém que cede ao desejo que se considera uma perversão sexual. Um modo humilhante

de tratar as pessoas, pode deixar quem observa enjoado (Ekman, 2011; Ekman & Friesen, 1975).

2.9.5.3 Aversão básica

Rozin e Ekman sugerem a divisão entre a aversão básica e a interpessoal, a primeira está relacionada à incorporação oral e contaminação. Para a aversão básica existem gatilhos universais mais potentes que estão ligados a produtos corporais. Como fezes, vômito, urina, muco e sangue. Cientistas da emoção sustentam que esses gatilhos universais são naturais, mas além desses, há desdobramentos de gatilhos que variam de acordo com a definição de cada indivíduo, que é influenciado pela cultura e as experiências pessoais. Existe uma grande variedade de produtos alimentícios entre as culturas humanas. Há pessoas que gostam de ostras frescas e outras consideram-nas repugnantes. Em determinadas regiões da China, cachorros, morcegos entre outros, são considerados iguarias suculentas, mas muitas pessoas, principalmente ocidentais, consideram essas repugnantes. Nesta forma de aversão, uma função importante é evitar a contaminação ou dano (Ekman, 2011; Ekman & Friesen, 1975; Miller, 1997; Rozin et al., 1999).

2.9.5.4 Aversão interpessoal

Para esta forma de aversão, considera-se quatro grupos de gatilhos interpessoais aprendidos: o estranho, o doente, o desafortunado e o moralmente corrompido. Estudos sugerem que essa classe de aversão pode ser mais repulsiva para adultos, em vez da aversão básica associada a incorporação oral. De entre os gatilhos interpessoais, o comportamento moralmente censurável mostrou-se com maior intensidade nos entrevistados. Como as atrocidades cometidas nos campos de concentração nazis e as ações sexualmente repugnantes, como ver alguém mantendo relações sexuais com uma criança. Algumas formas de aversão interpessoal variam de acordo com a cultura. Normalmente causa repugnância interpessoal nos japoneses, alguém que não se encaixa na ordem social ou que critica injustamente outras pessoas. Para os norte-americanos pode ser a brutalidade ou racismo. Rozin identificou que em diversas culturas, os políticos enjoam as pessoas e é possível que essa forma de aversão social não varie de acordo com a cultura. A função da aversão interpessoal está relacionada com o afastamento do que é considerado censurável, para evitar o dano social. Funciona como um julgamento moral onde não se pode conciliar com a pessoa ou com as ações repugnantes (Ekman, 2011; Ekman & Friesen, 1975; Rozin et al., 1999).

2.9.5.5 Superação e transformação dos gatilhos

Miller mostra que é mais fácil aprender novos gatilhos repugnantes do que fazer com que as pessoas os esqueçam. Contudo, existem situações que esses gatilhos são suspensos ou transformados, dando lugar a emoções agradáveis. Por exemplo, quando há intimidade sexual. Em um beijo, a língua de uma pessoa entrando na boca de outra, é agradável para elas. Houve uma transformação do gatilho em relação a beijar a pessoa amada. Porém se não houvesse essa intimidade sexual, a mesma situação poderia gerar um ataque repulsivo. O mesmo acontece no sexo, que além da transformação, significa a transgressão mútua dos limites defendidos pela aversão. Trocar fraldas do bebê é sempre repugnante, mesmo sendo filho. Mas, a aversão é superada ou suspensa, através do cuidado ou do amor ao bebê. Naturalmente quando o indivíduo tem a visão das entranhas de outra pessoa, a reação é ter náusea, evitar ou se afastar. No entanto, se a pessoa não é um estranho, mas um familiar, a aversão pode ser superada e o indivíduo é motivado a reduzir o sofrimento da pessoa íntima, em vez de fugir. Numa outra pesquisa, onde foram mostrados vídeos de cirurgias onde havia sangue e vídeos de exposições de feridas de outras pessoas, a maioria dos entrevistados dos Estados Unidos e do Japão demonstraram e relataram sentir aversão. Porém um grupo de 20% dos entrevistados reagiram com tristeza e dor, como se tivessem se identificado com as vítimas (Ekman, 2011; Ekman & Friesen, 1975; Miller, 1997).

2.9.5.6 Desprezo

O desprezo pode ser considerado uma emoção que está relacionada à aversão, ou um desdobramento dela, mas possui as suas próprias características. Essa emoção é vivenciada em relação a pessoas ou ações, não a sabores, cheiros, toques e sons repugnantes como na aversão. Pisar sobre fezes caninas pode causar aversão, mas não desprezo. No entanto, pode-se sentir desprezo por quem come coisas repulsivas. A emoção pode ser ativada quando o indivíduo não gosta de uma determinada pessoa ou das ações feitas por ela. No desprezo há um elemento de superioridade, arrogância por parte de quem despreza. O indivíduo sente uma sensação de superioridade em relação a quem ele despreza, geralmente uma superioridade moral. Isso pode provocar uma reação de desprezo por parte de quem é desprezado. Especialmente se há uma relação de subordinado e superior. As pessoas inferiorizadas, sabem que estão em posição inferior aos olhos dos outros. A ofensa feita pela pessoa desprezada é degradante, mas não se torna necessário afastar-se dela, como na aversão. O desprezo expressa poder e *status*

para os que estão ao redor, por isso aqueles incertos a respeito do próprio *status*, geralmente manifestam desprezo para afirmar sua superioridade sobre outros. Como todas emoções, o desprezo varia em intensidade. Essa emoção pode ser acompanhada pela raiva ou uma forma menos intensa de raiva, como o aborrecimento. Considera-se a função do desprezo como sinalizar a sensação de superioridade ou da falta de necessidade de se acomodar ou se envolver (Ekman, 2011; Ekman & Friesen, 1975).

2.9.6 Alegria

Atualmente há poucas pesquisas relacionadas as emoções agradáveis, pois a maioria das pesquisas dedicadas às emoções, são sobre emoções perturbadoras que causam a grande parte dos distúrbios emocionais (Ekman, 2011).

Outros motivos são que, as emoções positivas são menos distintas do que as negativas, e menos óbvias em termos de suas funções adaptativas. Por exemplo, enquanto emoções negativas básicas, como raiva ou repulsa, estão associadas a expressões faciais específicas, a experiência de algumas emoções positivas, não são frequentemente sinalizadas por movimentos faciais específicos. Geralmente são sinalizadas pelas mesmas expressões de alegria, representado pelo sorriso de “Duchenne”, “Duplay” entre outros. As emoções negativas parecem ter funções adaptativas óbvias, por exemplo, o medo promove a fuga perante a ameaça, enquanto as funções das emoções positivas são menos compreendidas em termos do vínculo com ações específicas (Beaumont, 2009; Ekman, 1992).

Cooperando no crescimento da Psicologia Positiva, Barbara Fredrickson propôs uma abordagem mais focada nas emoções positivas. Mostrando que as emoções positivas são únicas e ampliam os repertórios de pensamento e ação das pessoas, desfazem efeitos nocivos das emoções negativas, cria e fortalece resiliência psicológica e desencadeiam movimentos ascendentes em direção a o bem-estar emocional. Ao contrário das emoções negativas que direcionam as respostas às metas estreitas e específicas, raptando o foco cognitivo. Por exemplo, o interesse inicia a tendência de explorar e absorver informações, a alegria leva ao desejo de brincar e criar e o contentamento gera o desejo de saborear e integrar as circunstâncias (Beaumont, 2009; Fredrickson, 2001).

A palavra alegria pode ser usada para representar um estado mais intenso que satisfação ou felicidade, porém não é precisa para informar qual a emoção agradável vivenciada. Ekman sugere que exista mais de uma dúzia de emoções agradáveis universais e

diferentes. Elas estimulam a vida e motivam as pessoas a fazerem coisas que beneficiam a sobrevivência e desenvolvimento. Como relações sexuais e apoio ao crescimento dos filhos. O modelo que Ekman propõe de emoções positivas engloba duas categorias, os prazeres sensoriais, que estão ligados aos cinco sentidos humanos e quando são estimulados de determinadas formas e contextos, proporcionam sensações agradáveis. E algumas supostas emoções distintas, que são sugeridas como temas universais e estabelecidos a partir experiências individuais, que são: a diversão, contentamento, entusiasmo, alívio, assombro, êxtase, *fiero*, *naches*, elevação, gratidão e *schadenfreude*. Provavelmente haja inúmeras outras variações desses temas que são aprendidos e tornam-se fontes de emoções diferentes. Há dúvida se os prazeres sensoriais são de fato experiências diferentes dos temas que envolvem alegria (Ekman, 2011; Ekman & Friesen, 1975).

2.9.6.1 Prazeres sensoriais

Os prazeres sensoriais são despertados através dos cinco sentidos humanos: visual, tátil, olfativo, auditivo, gustativo e como visto anteriormente, esses mesmos sentidos também são utilizados para despertar as outras emoções além de alegria. Dentro de cada uma dessas fontes sensoriais, existem inúmeros gatilhos diferentes para o prazer, variando estes de acordo com as culturas. Quando os seres humanos tocam determinadas coisas, sentem sensações agradáveis. Assim como ser tocado por alguém em determinado contexto também gera essas sensações, principalmente quando o toque é de alguém por quem nos interessamos e feito de forma sensual ou carinhosa. Há visões agradáveis de contemplar, como um pôr do sol. Existem algumas teorias estéticas e conceitos artísticos de combinação de formas e cores que mostram como esses elementos devem ser organizados para proporcionar sensações agradáveis ao olhar dos seres humanos. Assim como há uma grande variedade de músicas agradáveis e sons prazerosos, como as ondas do mar, a água correndo nas pedras em um riacho e o som do vento mexendo as folhas das árvores. Gostos e cheiros são considerados em parte, pois para a maioria das pessoas gostos doces são naturalmente saborosos, mas a capacidade de apreciar os gostos azedos, amargos ou condimentados parece ser adquirida com o tempo. De modo semelhante, há inúmeros perfumes e cheiros de flores que são agradáveis, mas há também alguns tipos de queijos muito apreciados, que tem o que se considera um cheiro terrível. Devido a carência de pesquisas nessa área, não se sabe ao certo se os prazeres sensoriais são diferentes rotas para a mesma experiência emocional, considerando assim como a mesma

emoção ou se cada um é uma emoção diferente. Além disso, há uma discussão entre os cientistas das emoções a respeito dos prazeres sensoriais serem ou não emoções. Os prazeres sensoriais podem ir além destes citados. Ekman sustenta que a maioria dos prazeres sensoriais, são gatilhos aprendidos que envolvem uma avaliação estendida ou uma avaliação técnica, quando se está relacionado as artes. Isso pode ser aplicado para todos os sentidos, mas em menor grau pelo tato (Ekman, 2011; Ekman & Friesen, 1975).

2.9.6.2 Diversão

A diversão é uma das emoções agradáveis mais simples e está relacionada a algo cômico. Algo que faz com que as pessoas riem, como as piadas que despertam gargalhadas em um show de *stand-up comedy* ou um filme do gênero comédia. Muitas piadas baseiam-se em expectativas erradas para obter o humor. Nestes casos, a diversão depende de ser tomada pelo enredo da piada e surpreendida pelo final inesperado. Alguns gostam de fazer as pessoas ao seu redor rirem ou gostam de se entreterem com essa emoção, assim essas duas situações podem fazer essa emoção se tornar o ponto central da personalidade dessas pessoas. Grande parte da indústria do entretenimento dedica-se a produzir essa emoção. Ela pode variar de intensidade, com risadas estrondosas, risadas com lágrimas a até fora de controle (Ekman, 2011; Ekman & Friesen, 1975).

Piadas, vídeos ou conteúdo divertido ou jogos na forma de quebra-cabeças são a forma mental da diversão. Mas ela pode também ser experimentada durante uns jogos físicos, brincadeiras grosseiras como simulação de agressão, etc. Evidências em primatas e humanos bebês e adultos, sugerem que o jogo físico e o humor compartilham um sinal comportamental. Como sorrisos, risadas e uma mandíbula caída ou boca larga, que indica vontade de brincar. Esses sinais de diversão também foram vistos em contextos de assuntos que são tabus, na falta de jeito de outra pessoa que gera situações cômicas e nas brincadeiras grosseiras de primatas e de pai e filho (Shiota, 2009).

2.9.6.3 Contentamento

Esta emoção acontece em momentos em que o indivíduo entende que tudo parece certo à sua volta, que as coisas estão corretas, cada uma no seu lugar. Quando não há nada para fazer, nenhuma tarefa para cumprir, então a pessoa sente-se contente ou relaxada. É semelhante à sensação de satisfação. Há pessoas que buscam momentos de calma e tranquilidade para alcançar essa emoção. Ekman sugere que os músculos faciais possam ficar relaxados e que há sinais na voz (Ekman, 2011).

Nas experiências sexuais, pode-se sentir contentamento subsequente à queda da excitação, após o orgasmo (Ekman & Friesen, 1975).

Envolve tranquilidade e serenidade. É uma resposta emocional que tende a surgir sob condições que incluem alta segurança e baixo esforço. O contentamento produz uma tendência mais ampla de ação do pensamento do que estados neutros ou não emocionais. Além de gerar o desejo de saborear o momento presente e integrar-se às circunstâncias atuais de uma maneira que amplie a compreensão de si próprio. Levando ao potencial de uma mentalidade aprimorada em relação a si e ao mundo (Beaumont, 2009).

Numa experiência foi detetado que a experiência de contentamento provocada por um clipe de filme representando as ondas do mar quebrando na praia, reduziu a reatividade cardiovascular potencialmente prejudicial à saúde, que permanecia após às emoções negativas (Fredrickson et al., 2000).

2.9.6.4 *Entusiasmo*

O entusiasmo pode ser despertado quando há uma novidade ou um desafio, essa emoção surge como resposta a esses cenários. Alguns cientistas consideram o entusiasmo como uma forma intensa do interesse, pois de certo modo, quando o indivíduo vivencia a emoção, ele é dominado por uma potência ou vontade. Porém, para outros cientistas, o entusiasmo não se pode resumir num interesse intenso, pois o interesse é predominantemente cerebral, um estado de pensamento e não uma emoção. Contudo o entusiasmo possui uma relação com o interesse, pois assuntos que começam apenas interessantes podem se tornar excitantes, especialmente quando acontecem mudanças desafiadoras e rápidas, inesperadas ou originais. Ao vivenciar o entusiasmo, algumas vezes também pode-se vivenciar outras emoções que se alternem com ela de acordo com o contexto. Como um medo mais intenso, que é o pavor ou um medo indireto, que não representa um perigo real. Pode alternar com a raiva de forma mais intensa como a fúria e em outros contextos, pode se alternar com outras emoções agráveis (Ekman, 2011; Ekman & Friesen, 1975).

2.9.6.5 *Alívio*

Antes do alívio, geralmente vêm outras emoções, isso o torna incomum. Muitas vezes ele vem acompanhado de um suspiro e acontece quando algo que tinha despertado fortemente alguma emoção se aquieta ou se resolve de forma positiva. Muitas vezes o medo vem

antes do alívio, mas se a resolução da situação for negativa, em vez de alívio o indivíduo pode sentir tristeza. A tristeza pode ser sentida antes do alívio quando se espera uma notícia devastadora, da qual não se sabe o resultado, quando se recebe uma notícia boa em vez da ruim, o alívio pode ser sentido, como foi discutido anteriormente sobre a tristeza. Pode ser despertado quando uma pessoa angustiada é consolada ou tranquilizada com sucesso a respeito de uma perda. Depois de momentos de raiva intensa como a fúria também pode-se sentir alívio. Outros exemplos são, quando se descobre que o exame de cancro deu negativo, quando se acha o filho perdido no shopping ou quando se vai bem em uma avaliação difícil, que se pensava ter ido mal. Também se sente essa emoção depois de experiências positivamente valiosas, como o alívio da tensão sexual. Algumas vezes após o orgasmo, pode ser sentido entusiasmo misturado com o alívio a respeito do desempenho sexual. Além disso, os momentos de prazer intenso, também podem vir antes do alívio (Ekman, 2011; Ekman & Friesen, 1975).

2.9.6.6 Assombro

O assombro é uma emoção de ocorrência rara, a maioria das pessoas não estão acostumadas com essa emoção. Portanto, há uma certa dificuldade em conseguir que ela ocorra em um laboratório onde possa ser medida, por isso existem poucas experiências. Essa emoção desperta um sentimento predominante de um evento incompreensível, surreal, sobrenatural ou metafísico. A experiência do assombro é intensa e agradável. No entanto, ele pode alternar-se com o medo, nos casos em que a segurança do indivíduo é ameaçada por algo opressivo e difícil de compreender. Dacher Keltner e Jonathan Haidt defendem uma teoria sobre a estupefação, com um argumento de que ela é resultado da combinação de assombro e medo. Eles sustentam a dificuldade da mente em compreender o evento ou objeto causador de estupefação. Quase tudo que é incrível, incompreensível e fascinante pode ser fonte para despertar o assombro. Para Ekman, uma das sensações físicas mais intensas do assombro é o arrepiar da pele. Há mudanças respiratórias, suspiros ou o balançar da cabeça. Esse último comportamento está relacionado com a incredulidade. Pois, como o assombro está associado a um evento incompreensível, quem vivencia pode atribuir um significado sobrenatural, os que desacreditam nesse significado ou no evento, podem balançar a cabeça num gesto negativo. A possibilidade de poder atribuir o significado do evento a uma entidade sobrenatural é importante para desencadear dessa emoção. Numa situação em que o indivíduo conhece alguém e descobre algumas coincidências entre essa pessoa, o assombro pode ser vivenciado. Por

exemplo, ambos indivíduos cresceram na mesma cidade, frequentaram a mesma escola, mudaram-se para a mesma localidade, na mesma rua e mais tarde um viveu na mesma casa e no mesmo quarto que o outro viveu no passado, mas nunca se encontraram e não se conheciam durante tempo. Em relação a essas sequências de coincidências, note-se que pode haver uma interpretação no sentido de atribuir significado a uma entidade sobrenatural, como o destino ou algo semelhante como causador das coincidências. Portanto essa situação reúne três elementos para despertar o assombro, a sequência de acontecimentos incrivelmente coincidentes, pode ser tido como incompreensível e por último, permite atribuir a interpretação a algo sobrenatural como o destino (Ekman, 1992, 2011; Keltner & Haidt, 2003).

2.9.6.7 Êxtase

O êxtase é alcançado por meio de meditação, por experiências na natureza ou experiência sexual com a pessoa amada. Vivenciar o êxtase é uma experiência intensa e agradável, um estado de arrebatamento autotranscendente, similar ao entusiasmo e ao assombro, por isso não é algo que se vivencia levemente, em pequenas doses (Ekman, 2011; Ekman & Friesen, 1975).

A experiência de êxtase foi detetada em um estudo a partir da meditação. Foi medido a atividade cerebral de pessoas competentes em meditar. Os meditadores foram capazes de focar sua atenção para aumentar suas emoções positivas e alcançar o êxtase. Os pesquisadores detetaram um nível mais alto do que o normal de atividades elétricas nas células de seus lobos frontais (Holder, 2009).

2.9.6.8 Fiero

Essa emoção é sentida quando se enfrenta e supera um desafio. Geralmente um desafio difícil de acordo com a interpretação de quem enfrenta. Requer um sentimento favorável da pessoa a respeito de si mesma no momento da conquista. Envolve satisfação, mas é diferente, normalmente mais intenso. Semelhante a honra própria, triunfo, satisfazer-se com os próprios feitos, orgulho no sentido positivo e não no sentido de arrogância, onde há excesso de admiração por si próprio. A palavra *fiero* vem do italiano, segundo Ekman, é o termo que melhor traduz essa emoção, uma vez que não existe um termo semelhante em inglês. Geralmente é necessário que outras pessoas saibam da conquista, como uma plateia. Por isso o *fiero* pode ser visto em conquistas no mundo dos desportos. Também pode ser antecedido pelo entusiasmo, quando se inicia um desafio. Essa emoção também

pode ser sentida quando se descobre a solução para um problema intelectual difícil. Neste caso não há plateia, mas existe um grande desejo de que a conquista se torne pública. Por essa característica motivacional, acredita-se que o *fiero* foi essencial ao longo da evolução, motivando grandes conquistas (Ekman, 2011; M. Lewis, 2008).

2.9.6.9 *Naches*

Essa emoção também não possui uma palavra específica que a defina em inglês. *Naches* vem da língua iídiche, que é uma mistura de várias línguas, entre elas o hebraico, dialetos do alemão medieval e eslavismos. Essa emoção é semelhante a orgulho, mas têm suas próprias características e sensações corporais. *Naches* acontece quando se sente orgulho de um filho por uma conquista importante dele. Também pode acontecer quando um filho supera o pai em determinado assunto ou um aluno supera o professor, mentor ou mestre. É uma espécie de arrebatamento somando com orgulho, que o filho, aluno ou discípulo pode dar a seu pai ou mentor ao superar um grande obstáculo. Mas essa emoção é sentida somente pela pessoa que não realiza a conquista. Logo, vivenciar essa emoção positiva requer um desejo de ver a outra pessoa vencer ou realizar uma grande conquista. Há relatos de que, filhos, alunos ou discípulos também podem sentir essa emoção dos pais ou mentor. Portanto essa emoção pode ocorrer em duas direções. Ekman sugere que, quem não sente *naches*, provavelmente é competitivo com a pessoa do qual deveria se orgulhar e incentivar. Essa forma invejosa de sentimento, possivelmente seria um obstáculo psicológico, impossibilitando que alguns não sejam capazes de sentir *naches*. Assim como as outras emoções, *naches* está ligado a um fator que contribui para a espécie humana, neste caso ao desenvolvimento, uma vez que a emoção recompensa o investimento feito para o crescimento ou superação dos filhos, alunos ou discípulos (Ekman, 2011).

2.9.6.10 *Elevação*

Jonathan Haidt sugere que a elevação seja uma emoção. Ela está relacionada a atos altruístas. Paul Ekman não tem certeza se ela pode ser enquadrada como uma emoção. Pode sentir elevação, quem testemunha um ato moral exemplar realizado por outra pessoa. Pode ser descrita como uma sensação calorosa e edificante, que as pessoas experimentam ao ver atos inesperados de bondade humana. Portanto, essa emoção possui um papel moral nas relações humanas. É comumente provocada no contexto em que, o indivíduo ajudado carece de necessidades vitais, como pobreza, doença, quando a pessoa

está presa ou em outras situações consideradas difíceis. Pode ser despertada por outras ações que exemplificam atitudes morais e valores, especialmente aqueles que enfatizam a conexão social. Vivenciar a elevação, conduz o indivíduo a detetar virtudes morais no outro indivíduo que manifesta a ajuda. Como gentileza, compaixão, caridade, coragem e lealdade, que são virtudes potenciais para gerar um convívio benéfico. Por isso, essa experiência guia os julgamentos sobre a moralidade alheia. No sentido de identificar cooperadores exemplares, uma vez que pessoas que priorizam as necessidades de outras, são considerados bons para o relacionamento social. A elevação motiva o indivíduo a querer envolver-se com gestos altruístas e pode até conduzi-lo a unir forças, cooperando com quem iniciou o ato de ajuda. Portanto ao nível de comportamento do grupo, a elevação pode apoiar às normas comunitárias. Podendo provocar uma reação conjunta de outros membros do grupo a ajudar. Nesse sentido de motivação, a elevação pode ser considerada o oposto da aversão, em vez de conduzir o indivíduo a afastar-se de alguém doente, conduz a ajudar (Ekman, 2011; Haidt, 2000; Shiota et al., 2014).

2.9.6.11 Gratidão

O tema da gratidão é a apreciação de um presente altruísta. Pode vivenciar a gratidão quem recebe algo de bom de outra pessoa, como uma doação ou ajuda altruísta. Vivenciar a gratidão é semelhante a sentir apreço, estima, admiração ou respeito em relação ao doador. Há dois elementos necessários para sentir gratidão. O primeiro é o contexto interpessoal, ou seja, a necessidade de um remetente e um destinatário da boa ação. E o segundo é sobre a mente do destinatário, no ato de interpretar a intencionalidade do ajudador. Essa interpretação deve parecer não beneficiar quem doa, uma atitude que não caracterize troca, deve ser feita de forma gratuita. A implicação desses dois elementos, possibilita sentir gratidão independente do resultado do evento, por exemplo, quando alguém ajuda na busca frustrada de uma criança perdida. Numa perspectiva ampla, a moral está envolvida com o entendimento e o desencadear dessa emoção, tendo a empatia em sua base. São consideradas três variáveis que impactam na intensidade da gratidão, o grau em que ação é louvável, o desvio da ação do agente em relação às expectativas baseadas em papéis sociais. E o quanto a benfeitoria é oportuna e traz vantagens ao ajudado. Na maioria dos casos quando o ajudado é muito necessitado, a ajuda é bem-vinda, independente da intenção do ajudador. Em outros casos, a pessoa que é ajudada, pode sentir-se constrangida ou envergonhada, por ser diferenciada e receber tamanha atenção de quem ajuda. Normalmente, sente que está em dívida com o doador, devido a natureza

moral obrigatória da gratidão, salvo em algumas culturas que não levam em consideração esse fator. Para algumas pessoas, este sentimento de endividamento pode ser desconfortável de tal forma, que elas evitam receber presentes ou recursos. Quando a pessoa ajudada acredita que a doação foi feita por achar que ela estava muito necessitada ou carente, sendo que ela mesma não se considera carente, pode sentir ressentimento ou até raiva do doador. Isso faz da gratidão, uma emoção complicada, pois nunca se sabe quando surgirá, além de dificultar estudos empíricos em laboratório. A abordagem à gratidão não foi levada em consideração pela maior parte da comunidade científica que estuda as emoções. Mas recentemente tem ganhado espaço nos estudos de Psicologia Positiva. Na cultura popular e em muitas religiões, ela é amplamente difundida. A gratidão motiva às pessoas a retribuírem, com palavras ou gestos a quem doa ou ajuda. Funcionalmente, esse é um comportamento que recompensa a atitude pró-social do outro. Um gesto identificado por Ekman, é uma ligeira inclinação de cabeça, mas isso muda de acordo com a cultura. Ele supõe que há diferenças culturais importantes em relação a pessoa que ajuda. Por exemplo, alguns consideram apenas estar fazendo seu trabalho, dizem que não esperam retribuições por isso. A gratidão quando considerada um estado emocional ou humor, pode ser algo parecido com um sentimento de admiração ou apreço pela vida. Além das formas de expressão pessoais, ela pode ser expressa de forma impessoal, como gratidão em relação a natureza ou de outras formas em relação a deuses, animais, o universo etc. Algumas experiências mais profundas relatadas, estão relacionadas a uma admiração reverente religiosa. Numa pesquisa feita por Gallup com adolescentes e adultos americanos, mais de 90% dos entrevistados indicaram que, expressar gratidão os ajudou a se sentirem extremamente felizes ou um pouco mais felizes. Uma pessoa que se sente agradecida, pode estar mais inclinada a sentir-se amada e cuidada por outras pessoas (Ekman, 2011; Emmons & Shelton, 2002; Gallup, 1998; R. S. Lazarus & Lazarus, 1994; Shelton, 1990).

2.9.6.12 Schadenfreude

Esta palavra tem a sua origem no Alemão, pois é bem definida nesse idioma. O significado é parecido com alegria maliciosa. Há relação com a inveja, no sentido de ser um desejo sombrio de que o inimigo ou a pessoa invejada, perca a vantagem ou o objeto que provoca a inveja. *Schadenfreude* também é um prazer perverso que se sente ao saber que o inimigo ou a pessoa invejada sofre, falha ou fracassa. Pode ser considerada o oposto de sentir alegria ou orgulho na conquista de outra pessoa, como no caso do *naches* ou o

oposto de empatia, em vez de o indivíduo sentir vontade de ajudar, ele se satisfaz no sofrimento do outro. A empatia pelo sofrimento do outro implica na proximidade psicológica e convergência emocional, *schadenfreude* implica na distância psicológica e divergência emocional. É um tipo diferente de satisfação em relação às consideradas até aqui, pois vivenciá-la não é moralmente legítimo. Esse é um motivo de *schadenfreude* ser passiva, no sentido de ter o desejo ou esperar que algo ruim aconteça ao opositor, pois um comportamento ativo, implicaria em fazer o mal. Isso atribui a ela uma característica oportunista que tira a responsabilidade do mal que acontece, permitindo até que o indivíduo desfrute do infortúnio do outro disfarçando por meio de um viés altruísta. Algumas pessoas podem ter inveja de outras por incontáveis fatores. Por considerarem a pessoa invejada superior ou bem-sucedida, devido alguma qualidade ou habilidade específica, um bem de consumo, riquezas materiais, *status* adquirido, por parecer sempre de bom humor ou por ter um cargo superior. Essa emoção pode ser perigosa quando incorporada ao comportamento de grupo, pois possibilita que a agressão e a perseguição ao grupo oposto seja tolerada ou consentida. Há indícios de que o antissemitismo do holocausto, em parte foi sustentado pela inveja dos Alemães aos Judeus bem-sucedidos. Portanto, *schadenfreud* fazia parte do contexto criado naqueles tempos sombrios (Clanton, 2006; Ekman, 2011; Haidt & Keltner, 2004; Leach et al., 2003; Nietzsche, 2007; Spears & Leach, 2004).

2.10 Pensamento de grupo

Os indivíduos podem acreditar que estar num grupo é melhor, devido a falsa ideia de que a tomada de decisões em grupo é melhor do que por apenas uma pessoa. Geralmente, grupos tem objetivos explícitos e implícitos em comum que refletem as crenças do grupo. Algumas características do fenômeno *groupthink* serão apresentadas a seguir. A ilusão de vulnerabilidade, que está relacionado a assumir riscos, ignorando perigos, além de estar aliado a comportamentos extremamente otimistas. Racionalidade coletiva, tem haver com a postura dos membros do grupo em não se importar com alertas e rejeitar suas próprias hipóteses sobre os riscos. Crença na moralidade inerente do grupo, está relacionado as decisões dos membros serem moralmente corretas, ainda que isso viole a ética da maioria das pessoas de fora do grupo. Visões estereotipadas, está relacionado a uma construção, na maioria das vezes falsa, que os membros do grupo têm em relação aos opositores deste grupo. Pressão sobre dissidentes é uma pressão que o membro do grupo pode sofrer para não expressar argumentos contra os ideias ou crenças do grupo. Auto-censura é quando

o membro não se expressa, ele se censura, pois percebeu que suas ideias não são consenso grupo. Ilusão de unanimidade, acontece quando o membro acredita que todos concordam com a decisão do grupo, ainda que isso não seja verdade, como o silêncio pode parecer um consentimento coletivo. *Mindguards*, é a proteção do grupo e do líder desse grupo, em relação a informações que vão de encontro à coesão do grupo ou suas decisões. Por isso, pensamento de grupo nestas circunstâncias, pode ser considerado à deterioração da eficiência mental e do julgamento moral deste grupo (VOGT, 2017).

2.11 Processamento de Linguagem Natural e Linguística

O Processamento de Linguagem Natural (PLN) é uma área de estudo que se concentra em fazer com que computadores possam entender e gerar a comunicação natural dos seres humanos. A comunicação natural pode ser a escrita, fala, expressões faciais, etc. Essa área é estudada em dois sentidos, dos seres humanos para computadores e dos computadores para os seres humanos, esta última geralmente ocorre através da geração de textos e falas a partir de Inteligência Artificial. O PLN é conhecido como a ciência que estuda a interação entre computadores e humanos. Ela passou a existir a fim de facilitar o trabalho do utilizador em se comunicar com um computador através de linguagem natural (Khurana et al., 2017).

Essa área é largamente utilizada na escrita e na fala. Por isso, promove a necessidade da junção de duas áreas, a ciência da comunicação e a linguística, que é a base do PLN. Comumente utiliza técnicas de aprendizado de máquina ou *machine learning* (Roh et al., 2019).

O PLN abrange mais áreas do que somente a ciência da computação e a linguística, pois entender linguagem natural é um problema da Inteligência Artificial. Por isso, pode envolver outras ciências da linguagem ou também estatística, psicologia ou ciências cognitivas (Kocaleva et al., 2016).

Os seres humanos parecem aprender a partir de padrões, que são usados para diferenciar singular, plural, substantivo, verbos ou adjetivos. E nisso consiste a base para formar e diferenciar as frases. Teoricamente essa sequência de padrões que foram aprendidos pelos seres humanos na escola, poderiam também ser ensinados as máquinas de forma que possibilite que elas sejam capazes de aprender sobre a língua natural humana. E é isso que a maioria das técnicas de PLN tem feito. Treinar computadores para entender a

comunicação dos seres humanos e também gerar comunicação artificial e interagir (Kocaleva et al., 2016).

Portanto o PLN pode ser dividido nessas duas partes, a compreensão da linguagem natural e a geração de linguagem natural, que são basicamente compreender e gerar texto. Se a entrada de linguagem natural for falada, este conteúdo será convertido em texto para o processamento. E de modo semelhante, a linguagem gerada em texto pela máquina pode ser convertida em fala, para ser compreendida por seres humanos (Khurana et al., 2017).

A maioria dos sistemas de PLN utilizam níveis de compreensão que formam a arquitetura padrão desses sistemas (Kocaleva et al., 2016). Esse sistema é utilizado para compreender a linguagem escrita natural, vem da linguística, que é a ciência que envolve o significado da linguagem, o contexto e várias outras formas de estudo (Khurana et al., 2017). Em seguida serão apresentados os principais níveis desse sistema usado em PLN.

2.11.1 Fonologia

Estuda o sistema sonoro de um idioma, ocupa-se com os sons pertencentes a um sistema de linguagem e produzidos para a comunicação entre os seres humanos. O termo vem do Grego antigo *phono* e significa voz ou som e o sufixo *logia* refere-se a palavra ou discurso. Outras definições explicam que são os sons da língua e está preocupado com a função, comportamento e a organização dos sons como itens linguísticos (Khurana et al., 2017). A fonologia é muito importante para a linguagem falada e na compreensão desta, porém pouco importa para textos somente escritos (Kocaleva et al., 2016).

2.11.2 Morfologia

A morfologia trata do estudo da forma das palavras, as diferentes partes que a compõem. E a menor parte de uma palavra é chamada de morfema. Um exemplo pode ser a palavra “incerteza”. Morfologicamente ela está dividida em três morfemas, um prefixo “in”, a raiz ou radical “cert” e o sufixo “eza”. Uma determinada palavra pode ser quebrada até chegar na sua raiz, a fim de compreender apenas o significado, pois é imutável este significado. Logo, palavras desconhecidas também podem ser quebradas para chegar na raiz. No entanto, para compreender o significado da palavra no contexto deve-se analisar os outros morfemas, como prefixo e sufixo. Apenas um sufixo é suficiente para inverter o significado da palavra (Khurana et al., 2017).

2.11.3 Sintático

Esse é um tipo de análise que utiliza as palavras de uma frase como abordagem a fim de descobrir a estrutura gramatical desta frase, ou seja, as dependências estruturais entre as palavras de acordo com a gramática (Khurana et al., 2017).

A análise sintática inclui dividir a frase em componentes menores responsáveis pela origem dela, e assim compreender os regulamentos e condições definidas pela gramática que foram projetados para manter a frase junta. Para que ocorra a análise sintática, é necessário que exista um verbo na frase analisada. Os sistemas de PLN armazenam uma representação de todas as frases, armazenam o verbo e o tipo dele (Kocaleva et al., 2016).

Essa técnica nem sempre é tão eficaz, existindo várias formas gramaticais que podem impedir um analisador de captar a estrutura correta. É por isso que ainda existem muitos desafios nesta área. Na maioria dos idiomas, esta análise considera a ordem e dependência dos elementos, o que faz muita diferença ao nível de significado, ou seja, semântico. Por exemplo nas frases: “O gato perseguiu o rato” e “O rato perseguiu o gato” diferem em termos sintáticos, mas transmitem significados bastante diferentes (Khurana et al., 2017).

2.11.4 Semântico

Esse tipo de análise está ligado ao significado das frases, como mencionado anteriormente. Considera a forma como são construídas as estruturas sintáticas, por isso as análises em nível sintático e semântico são inseparáveis e complementares (Kocaleva et al., 2016).

O processamento semântico pode determinar os possíveis significados de uma frase, de acordo com as interações entre os significados das palavras (Khurana et al., 2017).

A partir da semântica, verifica-se que existem palavras com múltiplos significados e de acordo com a ocasião, algumas palavras podem adquirir múltiplos contextos. Isso resulta num problema de ambiguidade no significado das frases (Kocaleva et al., 2016).

E o nível de processamento semântico pode incorporar a desambiguação de palavras com múltiplos sentidos, o que coopera para compreender o significado destas frases. Essa forma de desambiguação opera no nível das palavras, baseando-se nos significados delas (Khurana et al., 2017).

Os significados são derivados do dicionário, mas também podem considerar o contexto que a palavra está inserida (Khurana et al., 2017). Ou seja, identifica o significado

apropriado olhando para o restante da frase ou da dependência do contexto (Kocaleva et al., 2016).

2.11.5 Discurso

A análise a nível do discurso, tem uma abordagem entre duas ou mais frases em um texto ou em um parágrafo, examinando as dependências dos significados entre elas. É utilizado em PLN para entender o papel de cada informação em um documento (Kocaleva et al., 2016).

O discurso considera unidades de texto maiores que uma frase, como o conjunto de frases que formam um parágrafo ou um conjunto de parágrafos que formam um texto. Essa análise se concentra nas propriedades do texto como um todo que transmitem significado. Analisa as conexões entre as frases considerando essas como pequenas componentes de um conjunto maior (Khurana et al., 2017).

2.11.6 Pragmático

Esse último nível tem como foco a análise do uso intencional da linguagem numa situação de contexto que vai além do conteúdo do texto, ou seja, os significados são extralinguísticos, vão além do que é dito objetivamente nas frases. É por isso que envolve conhecimento de mundo, das intenções, planos e objetivos, uma vez que o conteúdo não é todo expresso na escrita em análise (Khurana et al., 2017).

Esta abordagem trata da análise de sentenças e como elas são usadas em diferentes situações e como seus significados mudam de acordo com a situação (Kocaleva et al., 2016).

De acordo com a Enciclopédia Britânica, é o estudo de como os aspetos literais e não literais do significado comunicado, são determinados por princípios que se referem ao contexto físico ou social em que a linguagem é usada. Não o podemos confundir com Pragmatismo, que é uma escola filosófica (Britannica, 2020).

2.12 Aprendizado de máquina

O aprendizado de máquina é um ramo da inteligência artificial. São sistemas de análise capazes de aprender com dados de treinamento, em vez de seguir uma ordem exata de execução de instruções já programadas (Kocaleva et al., 2016).

A atual capacidade computacional disponível, aliada a técnicas e algoritmos de aprendizado de máquina (ML), permitem processar enormes quantidades de informação disponível em variados conjuntos de dados de diversas áreas. O aprendizado de máquina tem sido cada vez mais utilizado nos mais diversos domínios, seja no setor público, como nas áreas fiscal (António Seíça et al., 2019), educação (Pimenta, Ribeiro, Sá, & Belfo, 2018) e medicina (Brandão et al., 2021) ou, no sector privado, como no marketing (Cui et al., 2006), na indústria de *media* e entretenimento (Sereday & Cui, 2017), na indústria de eventos (Loureiro et al., 2014), no turismo (Pimenta, Belfo, & Trigo, 2009) e em muitas outras áreas, contribuindo para a criação de novos conhecimentos e ajudando as organizações a definir estratégias que lhes permitam aumentar seu desempenho.

Essa técnica é utilizada em PLN. Máquinas de PLN podem aprender a partir de informações linguísticas de cada palavra, sentença ou documento e analisa novas frases ou documentos de entrada com base nas informações aprendidas. E pode retornar ainda mais informações linguísticas dos textos analisados do que somente os significados (Roh et al., 2019).

Há uma tendência dentro das técnicas de aprendizado de máquina, que é a construção de sistemas artificiais autoajustáveis, que podem se desenvolver sozinhos e mudar com o tempo, a partir de experiências adquiridas pela coleta de novos conjuntos de dados (Kocaleva et al., 2016).

Existem dois tipos de aprendizado de máquina, o aprendizado supervisionado que geralmente envolve atividades relacionadas a previsão e aprendizado não supervisionado que geralmente envolve atividades descritivas (Kocaleva et al., 2016).

2.12.1 Aprendizado de máquina supervisionado

Aprender com supervisão significa que essa atividade é coordenada por seres humanos ao desenrolar das etapas. E pode ser uma atividade de classificação, previsões numéricas entre outras (Kocaleva et al., 2016).

A classificação é uma das possíveis atividades do aprendizado supervisionado, existem muitos métodos disponíveis para classificação. Nesta abordagem utilizam-se dois conjuntos de dados. Os dados de treino e os dados de teste fazem parte de um conjunto rotulado, ou seja, previamente classificado. A máquina aprende a relação entre os dados do conjunto de treino. E para testar a precisão do modelo aprendido, usa-se o conjunto de

dados de teste. Dessa forma é possível aplicar o modelo a fim de classificar novos dados (Hang & Banks, 2019).

2.12.2 Aprendizado de máquina não supervisionado

Nesta forma de aprendizado não há dependência de dados rotulados para o algoritmo aprender. Pode ser aplicado para identificar padrões ou regularidades nos dados. E também pode ser usado para calcular a saída correta para qualquer nova instância de dados. O aprendizado não supervisionado requer do cientista, a capacidade de identificar padrões no conjunto de dados (Kocaleva et al., 2016).

2.12.3 Aprendizado por reforço

Essa forma de aprendizado diferencia-se das outras por ser composta por um circuito que possui vários elementos que imitam a interação de um animal com o meio. A ideia básica é que esse circuito seja capaz de captar as características mais importantes do problema e interaja com o ambiente a fim de atingir uma meta. É denominado agente, aquele que age no ambiente para alterá-lo, este pode ser um robô físico, baseado em inteligência artificial ou puramente virtual e simulado, assim como o ambiente ao ser redor. O aprendizado por reforço é um sistema projetado para compreender o estado ambiente, planejar as próprias ações, executá-las e verificar a taxa de sucesso, baseado em estímulos de recompensa/punição. Assim, é possível aprender e aprimorar novas tarefas. É um modelo computacional poderoso para o aprendizado de tarefas orientadas à objetivos por exploração e recompensa. O estado refere-se à situação atual como percebida pelo agente. A recompensa é definida como o sinal positivo imediato (encorajamento) ou negativo (punição) determinando o que deve ser procurado pelo agente. Esse sistema é capaz de resolver um problema sequencial de tomada de decisão, aprendendo valores de ação para sequências de ações com base nas recompensas. Um valor de ação representa o retorno futuro, imediato e esperado da ação que é definido de acordo da situação. Na prática, os agentes aprendem valores de ação por exploração e quando certos o suficiente ou após uma determinada quantidade de esforço de exploração, o agente usa o conhecimento armazenado destes valores no seu componente de seleção de ações, para selecionar as ações apropriadas de acordo com determinadas situações. Como os valores das ações são aprendidos de acordo com o estado, em cada estado, um agente pode aprender a selecionar ações diferentes, desde que haja alguns recursos discriminativos desse estado que o torne suficientemente diferente dos outros (Broekens, 2018; Moerland et al., 2018).

2.13 Mineração de dados na análise de emoções e sentimentos

A análise de emoções através da mineração de dados tem sido um tema que tem merecido a atenção da comunidade académica (Batbaatar et al., 2019; Devika et al., 2016; Ducange et al., 2019; Sailunaz & Alhadjj, 2019; Seiça et al., 2019).

Num trabalho de Ducange, Fazzolari, Petrocchi e Vecchio (2019), foi desenvolvido um sistema do tipo *Decision Support System* (DSS), em português Sistema de Apoio a Decisão, que apresenta uma precisão significativa na área de análise de sentimento. Este sistema monitora continuamente várias mídias sociais, no qual coleta comentários a fim de mensurar a reputação de uma determinada empresa através de classificação de sentimento, que são positivo, negativo e neutro. Foram analisados um total de 100.000 comentários de três mídias sociais, nomeadamente Facebook, Twitter e Instagram. Para treinar o modelo, foram utilizados dois conjuntos de dados formados a partir de dois cenários diferentes, um voltado para restaurantes na Itália, coletado no TripAdvisor, que foi construído a partir de 1000 comentários positivos classificados com 5 estrelas pelos utilizadores e 1000 negativos classificados com 1 estrela. O segundo cenário voltado para lojas online de equipamentos eletrónicos, foi construído de modo semelhante e a partir de dados da Amazon, com 1000 comentários positivos com 5 estrelas e 1000 negativos com 1 estrela. Para o conjunto de textos neutros, foram usados 1000 comentários para ambos os cenários, mas dessa vez, extraídos de diferentes fontes, como discussões políticas, notícias e resenhas sobre tópicos distantes dos cenários selecionados, esses foram inspecionados manualmente para validar a neutralidade. Para selecionar o melhor método de classificação, em ambos os cenários foram treinados quatro modelos e comparados entre si os resultados a partir de uma validação cruzada, com os algoritmos *Support Vector Machine* (SVM), *Naive Bayes* (NB), C4.5 que é um algoritmo de árvore de decisão e *K-Nearest Neighbors* (K-NN). A precisão média, chamada de *accuracy* representa a eficácia geral do classificador e corresponde ao número de amostras classificadas corretamente sobre o número total de amostras. O SVM obteve os melhores resultados, com a precisão média de 91,95% para o cenário de restaurantes e 93,01% para o segundo cenário (Ducange et al., 2019).

Outro trabalho mostrou uma forma alternativa de classificação de sentimento dos *tweets* coletados desta rede social, relativo a autoridades fiscais portuguesas, em que ocorre duas etapas de classificação. A primeira é uma espécie de filtro, onde se usa a biblioteca Lexicon-PT, que possui dois léxicos em português etiquetados e polarizados para

classificação de sentimentos. Então, são identificadas as palavras contidas na amostra e também em ambos léxicos e assim, cria-se um novo conjunto de dados filtrados a partir desses léxicos. Nessa etapa, o léxico foi capaz de identificar na amostra, 86,7% de *tweets* com opiniões positivas, negativas ou neutras, equivalente a 880 de um total de 1015. Na segunda etapa foi aplicado o algoritmo *Naive Bayes* no conjunto de dados de 880 *tweets*, sendo 66% dados de treino e 34% conjunto de teste. E obteve a estatística *accuracy* de 74,83% (António Seïça et al., 2019).

Batbaatar, Li e Ryu (2019) propuseram o modelo *Semantic-Emotion Neural Network* (SENN) que faz uma análise em nível semântico, sintático e emocional para detecção de emoções. É baseado em *Deep Learning* (aprendizagem profunda), pois utiliza conjuntos de dados de treino maiores do que o aprendizado de máquina tradicionais, utilizando 10 conjunto de dados (*datasets*), com o maior deles a chegar a 24.361 ocorrências. Dessa forma, foi aliada a técnica de redes neurais, em que há duas sub-redes, a primeira usa memória bidirecional, de longo e curto prazo. No longo prazo analisou informações numa dimensão maior em que está envolvido o contexto ou assunto das frases analisadas. E no curto prazo está relacionado a uma dimensão curta, uma análise ao nível do significado das palavras, ou seja, ao nível semântico. A segunda sub-rede é uma rede neural convolucional, que se ocupa em extrair características emocionais e no relacionamento emocional entre as palavras do texto. Na experiência, foram utilizadas várias bases de dados, com diferentes formas de classificação e diferentes emoções. Essas bases de dados são constituídas por diálogos, *tweets*, contos de fadas, blogs e notícias. O modelo proposto obteve a melhor precisão comparado a outros modelos de ponta em 9 dos 10 conjuntos de dados testados: *F1-Score* de 84,8%, 51,1%, 61,3%, 74,6%, 91,0%, 56,3%, 59,3%, 98,8% e 70,8%. Os autores afirmaram que o vetor de palavras *FastText*, que foi usado para treinar o modelo em aprendizagem profunda, é mais sofisticado e devido a ele, se obteve um resultado melhor. Eles também afirmaram que provaram que o aprendizado profundo teve maior capacidade para previsão das emoções quando comparado aos modelos tradicionais (Batbaatar et al., 2019).

3 METODOLOGIA

Neste capítulo será abordado a metodologia utilizada para extrair, trabalhar e tratar os dados, a sua preparação, assim como também a sua compreensão através da sua exploração.

3.1 Metodologia CRISP-DM

Quando surgiu a ideia de KDD (Descoberta de conhecimento em bases de dados), diferentes modelos de processos e metodologias surgiram para realizar a tarefa de DM (Mineração de Dados). Estes evoluíram até chegar a uma metodologia considerada padrão, a CRISP-DM (do inglês *Cross-Industry Standard Process for Data Mining*), que emprestou ideias dos mais importantes modelos anteriores ao ano 2000, como o 5 A's, *Human-Centered*, SEMMA, Two Crows, entre outros. Concebida em 1966, esta metodologia consiste num modelo de processos hierárquicos, com diversas tarefas, distribuídas por vários níveis de abstração. O seu modelo de referência é composto por 6 fases principais, respetivamente, o entendimento do negócio, o entendimento dos dados, a preparação dos dados, a modelação, a avaliação do modelo e a implementação do modelo (Chapman et al., 2000). A pertinência e atualidade da metodologia CRISP-DM continuam a ser evidentes na área da mineração de dados e do *Business Intelligence* (Belfo & Andreica, 2018; Huber et al., 2019).

O CRISP-DM declara quais tarefas devem ser realizadas para concluir com êxito um projeto de DM. Por isso pode ser considerado um modelo de processo. É também um ciclo de vida em cascata. E também possui um componente metodológico, pois fornece recomendações sobre como executar algumas tarefas. A presente dissertação segue a metodologia CRISP-DM para realizar o trabalho de DM. Esta metodologia satisfaz as necessidades investigativas, além de acreditarmos que ela é a metodologia mais utilizada para o desenvolvimento de projetos de DM (Óscar et al., 2016).

3.2 Softwares utilizados

A ferramenta principal utilizada foi a linguagem de programação Python na versão 3.7, com bibliotecas padrão da linguagem e outras não padrão (Van Rossum & Drake, 2009). O ambiente de desenvolvimento foi o Google Colaboratory gratuito, mais conhecido como “Colab”, que é um projeto de pesquisa para criar modelos de aprendizado de

máquina. Ele fornece o ambiente de Notebook Jupyter para desenvolvimento interativo (Bisong, 2019).

Com o suporte do Python foi feito todo o processo de ETL, do inglês *Extract Transform Load* (Extrair Transformar Carregar), assim como a modelação e a validação. Para a extração dos *tweets* foi utilizado a biblioteca Tweepy, que faz conexão com a API da rede social Twitter (Roesslein, 2020). Para tratamento do conjunto de textos, em parte foram criadas estruturas na própria linguagem Python e em parte foi utilizada a biblioteca spaCy (Honnibal et al., 2020).

Os dados foram armazenados em uma base de dados sqlite3 (Hipp, 2020). Foram usadas outras bibliotecas para processamento, exploração, testes e visualização de gráficos como pandas, numpy, nltk, matplotlib, wordcloud e tqdm (Costa-Luis, 2019; Edward et al., 2009; Hunter, 2007; McKinney & others, 2010; Oesper et al., 2011; Oliphant, 2006).

Na etapa de modelação e validação, foi utilizada a biblioteca padrão do Python Scikit-learn (Pedregosa et al., 2011).

3.3 Campo de pesquisa e extração dos dados

A área delimitada para este estudo foi o debate no cenário político brasileiro no Twitter. O debate nesta rede social envolve pessoal e órgãos públicos, partidos políticos, poderes que constituem a governança brasileira, como os poderes Executivo, Legislativo e Judiciário. O debate desenvolve-se dia após dia, de acordo com as notícias que relatam eventos ocorridos com esses agentes. Ainda se estende a eventos isolados que ocorrem com cidadãos comuns, geralmente filmados ou transmitidos, que chocam a população e de alguma forma se relaciona num aspeto moral ou que envolve ideologias de partidos ou grupos políticos. Para nortear e selecionar a amostra coletada, foram identificadas as *hashtags* mais utilizadas em cada dia, que totaliza cerca de 50 *hashtags* diárias mais populares no cenário nacional brasileiro. As *hashtags* mais populares são chamadas de *Tred Topics* e podem ser selecionadas dentro da própria rede social abrangendo todo o território brasileiro, sem estarem filtradas pelo interesse pessoal do utilizador. As *Tred Topics* representam o assunto que está sendo debatido no momento e, por isso, envolvem muitos assuntos como futebol, programas televisivos, notícias, música e outras formas de cultura e o debate político que nos interessa. Portanto, destas 50 *Tred Topics* coletadas todos os dias, apenas as que são sobre o debate político variam desde duas a seis por dia.

Em média são quatro *Tred Topics* sobre assuntos políticos que foram coletadas do dia 22 de junho de 2020 até o dia 24 de agosto 2020, totalizando 190 *hashtags*.

As hashtags separadas pelos respectivos dias em que se tornaram tendências são: 22-jun: #FechadoComBolsonaroAte2016, #CassaAChapaTSEJa, #IncendioJaragua; 23-jun: #100MilhõesSemEsgoto, #MilitanciaPetistaPresente, #PL2360Nao, Angela Davis; 24-jun: #aguanaoemercadoria, Cármen Lúcia, #PSOL, #SenadoArquivaPL2630; 25-jun: #PTSempreAoLadoDoPovo, #PL2630NÃO, Censura; 26-jun: #auxilioemergencial, Gozando, Rio São Francisco, Parabéns Presidente, Oswaldo Eustáquio; 01-jul: #AlertaNacional, #CiroNaCasaDoSaber, #ParabensPresidenteBolsonaro, #BrequeDosApps, Coreia do Norte, #DevolvamNossaBandeira; 02-jul: #LavaJatoTraiuAPatria, #ParlerBrasil; 03-jul: #EaiMourão, Dória, Serra, Leblon, Bia Doria; 04-jul: Luciano Huck, #EraMelhorComPT; 06-jul: Sírio Libanês, #HaddadNoRodaViva, #CiroeSuplicy, Frida Kahlo; 07-jul: #forcacorona, Cloroquina, Bozo, Hélio Schwartsman, #URGENTE, Mandetta; 08-jul: #coisadeviado, Ricardo Eletro, #BolsonaroTemRazao, #VetaBolsonaro; 09-jul: #MexeuComACAixaMexeuComOBrasil, Dona Débora, João Otávio de Noronha, LGBTfobia, #GuedesRoubaOBrasil; 10-jul: O MBL, Leandro Narloch, Nego Ney, Ayan, #ForaBolsonaro, #DerreteMBL; 11-jul: #LavaJatoSempre, Suplicy, #DerreteMBL, Hélio Negão; 12-jul: #BicudaNoBozo, #Dia12ComCiro, #DelataAYAN, Janáina, Dra Nise, Micheque; 13-jul: #estudioi, #ConexaoComPT, #VidasNegrasImportam, Alphaville, Inpe, Gilmar Mendes, #Covid_19; 14-jul: #AprovaPL735ContraFome, #OrgulhoNaoEBinario, #verificanb, #MilitaresGenocidas, Átila, #impeachmentJa; 15-jul: Severino Cavalcanti, Geddel, Guerra Fria; 16-jul: #DerrubaVetosPL1142, Bob Jeff, Roberto Jefferson, Alexandre Garcia, #BrasilPedeSocorro; 17-jul: #HaiaParaBolsonaro, POSSO FALAR COMO CONVIDADA; 18-jul: nelson mandela, Marielle, #DerreteFelipeNeto; 19-jul: #PatriotasComBolsonaro; 20-jul: #PL735ContraFome, O Fundeb, Atila, Alexandre Garcia, #AprovaFundeb, Oxford, #DerreteFelipeNeto; 21-jul: #justiçaPorJosue, #NoBrasilEuMudaria, Qualicorp, Alcolumbre, Polícia Federal, #AprovaFundebJa, Serra, #FamiliasContraFelipeNeto; 22-jul: #HaddadAbsolvido, #TchauGuedes, #forcacovid, consulado Houston, Florestan Fernandes, Pfizer e BioNTech, Cloroquina, EUA e China; 23-jul: #DespejoZero, #MexeucomACAixaMexeuComoBrasil, Bia Kicis, Mandetta; 24-jul: #VozParaBernardoKuster, #GovernoDeFraudes, #BrequeDosApps, Paraná Pesquisas; 25-jul: #impeachmentalexandredemoraes, #DiadaMulherNegra, #25deJulho, Quariterê, BNDES, Toquinho, Barroso, #BolsonaroOrgulhoDoBrasil, #STFNaoVaiNosCalar, #BrequeDosApps; 26-jul: #LiveElzaPorMarielle, Haia; 27-jul: #HaiaCondeneBolsonaro, DEM e MDB, #DeltanNaLavajato, Alexandre Garcia, #NaturaNao, #TodosContraFelipeNeto; 28-jul: Silas Malafaia, Natura, Thammy, Altino, Bocardi, Chris Flores, #SemVacinaSemAulas; 29-jul: Casagrande, Inflação, Natura, ELE BATEU NELA, #BolsonaroAntiLavaJato, Marcelinho; 30-jul: #CemiterioBrasilRumoA100Mil, #NordesteComBolsonaro; 31-jul: #GloboApoiouFakesContraPT, Carlos Jordy, Allan,

Greta, #CiroNaCNN, *Fugiu*, *Weintraub*, #STFVergonhaMundial, *Allan dos Santos*, #FreeAllan; 15-ago: #HaddadFariaDiferente, #ACREDITENOBASIL, #EstamosComDeltan, #RachadinhaDaGlobo; 16-ago: #FolhaLixo, #ACREDITENOBASIL, *Sara Winter*, #RachadinhaNaGlobo; 21-ago: *Carla Zambelli*, #MandettaGenocida; 22-ago: #LavaJatoMaiorQueBolsonaro, #BrasilDesgovernado; 23-ago: #StopBolsonaroMundial, *Lewandowski*; 24-ago: #StopBolsonaroMundial, *Lewandowski*, #GloboLixo, *bundões*, #BolsonaroOrgulhoDoBrasil, *Flordelis*.

3.4 Método de extração

Dentro da biblioteca Tweepy foi usado o método `api.search()` para coletar cerca de 300 id de *tweets* mistos de cada *hashtag* acima. A configuração mista é dada pelo parâmetro `result_type='mixed'`, que retorna ids de *tweets* mistos, isto é, os que são populares com muitos compartilhamentos, interações e outros não populares. Outro parâmetro foi `lang='pt'` que retorna apenas *tweets* na língua portuguesa. A data da busca foi limitada para o dia em que a *hashtag* se tornou tendência. Essa coleta era feita no dia posterior do qual a *hashtag* tornou-se tendência.

A partir desses ids foi feito a coleta dos *tweets* específicos, pois o método anterior não retorna o *tweet* completo. Para isso foi utilizado o método `api.get_status()` configurado com parâmetro `tweet_mode="extended"`, a fim de obter o texto completo e outras informações adicionais da postagem. Dessa forma foi coletado e armazenado no base de dados SQLite o total de 42.610 *tweets* completos no seguinte formato:

Data completa da postagem, *hashtag*, id do utilizador, id do *tweet*, nome do utilizador, se é ou não *retweet*, url citado no *tweet*, *screen name* do utilizador, texto da postagem na íntegra.

Exemplo:

```
('2020-07-22      16:04:27',      '#HaddadAbsolvido',      1285968961285152768,
1260150566405144577, 'Avanci1983', 'RT @EdithMagalhaes', 'sem url', '@Avanci1983
#HaddadAbsolvido \n\nNão havia crime nenhum, era pura perseguição!!')
```

No campo “se é ou não retuite”, possui o “RT” indicando que é um *retweet* (foi compartilhado) seguido do *screen name* do utilizador do qual a postagem original foi compartilhada e se caso não for, retorna a palavra “não”.

3.5 Remoção de repetidos

Dos 42.610 *tweets* extraídos, muitos são repetidos pois são *retweets* de *tweets* originais. Uma vez que a busca é mista, em muitos casos ela retorna os *tweets* originais e vários *retweets* dele. Para contornar isso, foi criado um filtro, um código em Python para remover os *retweets* que forem completamente iguais aos *tweets* originais. Portanto alguns *retweets* que os utilizadores adicionaram texto ou emoji junto à publicação original não foram detetados. Depois deste filtro sobraram 19.610 *tweets* que representam 46,02% dos 42.610 totais de *tweets* extraídos. Nota-se que mais da metade dos *tweets* coletados eram completamente repetidos 53,97%.

3.6 Metodologia da classificação manual das emoções

3.6.1 Análise na perspectiva do emissor

Para fazer uma boa classificação, devemos considerar a relação emissor e recetor. Emissor é quem transmite a comunicação, ou seja, o utilizador que faz efetivamente a postagem no Twitter. E o recetor é quem recebe a mensagem ou lê a postagem (Wolf, 1989). A intenção deste trabalho está em interpretar a mensagem na perspectiva do emissor, ou seja, quais evidências estão presentes no texto de uma postagem que possibilita deduzir que a mensagem foi escrita num momento em que o emissor estava emocionado ou que a mensagem foi influenciada pela emoção e em qual das famílias das emoções as evidências se enquadraram.

Não existe a pretensão de analisar as postagens na perspectiva do recetor, no sentido de tentar identificar quais emoções o texto poderá causar. Pois pode haver muitos leitores diferentes e como foi apresentado na secção 2.7 Temas e gatilhos emocionais, podem existir gatilhos aprendidos estabelecidos por uma experiência passada singular, que são diferentes em cada pessoa. Logo, um pedaço de texto pode despertar emoções diferentes em recetores diferentes. Para melhor compreender isso, vamos supor que um apoiante do político A insulta o político B numa postagem. Apoiantes do político A geralmente irão concordar com o texto e podem cooperar ao sentir desprezo e zombar do político B nos comentários. Mas, apoiantes do político B, quando leem esse texto podem sentir raiva, ódio do emissor, dos apoiantes e até mesmo do político A. Por isso a tarefa de compreender as emoções que um texto pode causar é tão complexa e em grande parte fica no âmbito da pragmática, exigindo conhecimento além dos explícitos no texto, como

conhecimento de mundo sobre o assunto discutido, sobre como os grupos agem e sobre os recetores e os gatilhos aprendidos deles.

3.6.2 Gatilhos emocionais para classificar textos

No trabalho de etiquetagem, se quem fará a etiquetar apenas lesse as postagens e etiquetasse com a emoção que sente ao ler, provavelmente este usaria seu conhecimento prévio sobre os assuntos políticos sem qualquer pesquisa, assim como gostos pessoais, políticos que simpatiza e apoia ou que desaprova, usaria sentimentos e gatilhos prévios de experiências individuais passadas e isso tudo influenciaria na interpretação do que realmente o emissor pode ter sentido quando escreveu.

Por isso, foi desenvolvido neste trabalho uma forma de etiquetagem que segue os padrões identificados pelos cientistas das emoções. Cada família de emoção possui um tema e vários gatilhos inatos que podem ser extrapolados como gatilhos universais dos seres humanos, conforme visto na secção 2.7, Temas e gatilhos emocionais, e visto detalhadamente para cada emoção na secção 2.9, A compreensão de cada emoção. Abaixo veremos o resumo desses padrões para a identificação e classificação das emoções manualmente, baseado na bibliografia reunida na secção 2.9.

3.6.3 Identificação da Tristeza

Alguns autores sugerem uma divisão entre tristeza e angústia, porém são da mesma família, para este trabalho iremos considerar essas duas vertentes apenas como tristeza a fim de tornar a análise mais concisa.

Tema: a perda de algo ou alguém que representa muito valor e está ligado a quem perde. O propósito da tristeza compartilhada é pedir ajuda a outros.

Gatilhos e variações:

- Rejeição de um amigo, namorada, cônjuge ou um familiar;
- Perda de uma pessoa próxima ou familiar que pode ser o distanciamento ou morte;
- Perda da autoestima devido a algum fracasso ou relacionado a aparência, etc.;
- Perda da admiração ou elogio de um superior ou as pessoas próximas ou mesmo de um público;
- Perda da saúde ou de alguma parte do corpo ou de função devido a um acidente ou doença;
- Perda de um objeto querido.

3.6.4 Identificação da Raiva

Todas as variações da emoção raiva, serão consideradas apenas como raiva para este trabalho, de forma a facilitar a classificação.

Tema: interferência, quando alguém é interferido em algo que pretende fazer ou que estava fazendo. A frase “Cai fora do meu caminho” captura parte dos gatilhos da raiva, mas nem todos. Há também um sentimento de querer causar dano físico ou psicológico a pessoa que causou a raiva.

Gatilhos e variações:

- Frustração, em relação a alguém, a um objeto ou a si mesmo;
- A rejeição, assim como pode gerar tristeza, também pode gerar raiva;
- A própria raiva pode ser causadora, por exemplo, quando uma pessoa manifesta raiva contra outra, a raiva pode ser ativada nesta outra pessoa ou em outras pessoas que estão próximas;
- A ameaça de dano físico ou dano psicológico, como uma ofensa;
- Conceito contrário, ofensivo ou que viole os valores morais do sujeito. Ao ler um texto, interpretar alguma forma de mensagem que é contrária e pode vir também em forma de discurso, ação, crença, opinião, etc.;
- Desapontamento, que pode acontecer com pessoas que se tem maior proximidade ou intimidade;
- Indignação é a raiva moral, muitas vezes hipócrita;
- Mau humor como uma raiva passiva;
- Exasperação significa ter a paciência provada em excesso;
- A vingança, que é uma ação furiosa após um período de reflexão a respeito da ofensa;
- Ressentimento por um período curto, pois por um período longo deixa de ser uma emoção;
- No rancor, a raiva pode surgir quando se vê ou se lembra da pessoa que provocou o rancor.

3.6.5 Identificação da Surpresa

Tema: um evento que seja relevante e inesperado ou ao contrário do que se esperava.

- Sensações físicas como visão, som, cheiro, sabor ou toque surpreendentes;

- Outros aspetos como uma ideia, sentimento, comentário ou interpretar uma forma de comunicação, sugestão incorreta ou desconhecida, um conhecimento que se adquire, etc.

3.6.6 Identificação do Medo

Tema: está relacionado em torno do perigo de dano físico ou psicológico. É importante destacar que num ambiente de discussão política, as tentativas de dano psicológico por meio de insultos são comuns. O medo de uma ameaça imediata pode causar paralisia ou fuga. Para ameaça iminente o indivíduo fica num estado de receio, uma preocupação extrema, onde é dedicado o foco total.

Gatilhos e variações:

- Ameaças potenciais podem causar medo, ou seja, sem que o dano ocorra. Ameaças que não se tem certeza que cause dano e que possuem uma ou poucas características associadas com aprendizados passados, podem ativar o medo, mesmo que seja um medo sem fundamento como medo do escuro;
- Um objeto arremessado através do espaço que rapidamente atingirá o indivíduo, perder o apoio, como cair de uma cadeira, animais de grande porte, a movimentação dos répteis, sons de predadores, ver uma cobra ou aranha e etc.

3.6.7 Identificação do Nojo-desprezo

Como aversão e nojo são sinónimos e desprezo é uma variação da mesma família, com intuito de simplificar a classificação, foram considerados os comportamentos relacionados a aversão e seus desdobramentos como apenas “nojo-desprezo”, uma vez que são da mesma família.

Aversão básica: o tema está associado a incorporação oral, ou seja, ingerir algo desagradável ou ofensivo, que possa contaminar e causar dano. O objetivo do nojo em geral está relacionado a evitar a contato e contaminação, se afastar do que é considerado censurável ou mesmo expelir para fora do corpo.

Gatilhos e variações:

- Os gatilhos universais mais potentes estão ligados a produtos corporais, como fezes, vômito, urina, muco e sangue. Além destes, o gosto, cheiro, visão, tato ou sons de algo considerado desagradável ou que lembre, pode provocar aversão. Ver

uma pessoa deformada, aleijada ou considerada muito feia, machucada, com ferimentos expostos ou algum tipo de doença na pele ou fluidos que assumam características desagradáveis. Há variação cultural em relação a alimentos exóticos.

Aversão interpessoal: essa forma vai além do tema da incorporação oral, abrangendo conceitos considerados abomináveis. Mas a função é semelhante e está relacionada a afastar-se do que é considerado censurável, para evitar o dano social. Funciona como um julgamento moral onde não se pode conciliar com a pessoa ou com as ações repugnantes.

Gatilhos e variações:

- O estranho, o doente, o desafortunado, o moralmente corrompido e comportamento moralmente censurável. Ações, ideias, ideologias ou qualquer mensagem que seja considerada abominável, por exemplo maltratar um animal, perversão sexual, tratar de forma humilhante outras pessoas e etc. Há variação cultural, como para os japoneses alguém que não se encaixa na ordem social ou que critica injustamente outras pessoas. Para os americanos a brutalidade ou racismo. E em diversas culturas, os políticos enojam as pessoas.

Desprezo: é um desdobramento de aversão e o tema está relacionado a sensação de superioridade em relação a quem se despreza, geralmente uma superioridade moral.

Gatilhos e variações:

- Essa emoção é vivenciada em relação a pessoas ou ações, não a sabores, cheiros, toques e sons repugnantes como na aversão. A emoção pode ser ativada quando o indivíduo não gosta de uma determinada pessoa ou das ações feitas por ela. O indivíduo sente uma sensação de superioridade em relação a quem ele despreza, geralmente uma superioridade moral. Algumas vezes relação de subordinado e superior se faz presente. Pessoas incertas a respeito do próprio *status* podem manifestar desprezo.

3.6.8 Identificação da Alegria

As emoções da família da alegria, geralmente estimulam a vida e motivam as pessoas a fazerem coisas que beneficiam a sobrevivência e desenvolvimento.

Prazeres sensoriais:

- São despertados através dos cinco sentidos humanos: visual, tátil, olfativo, auditivo, gustativo.

Diversão:

- Algo que faz com que as pessoas deem risada, relacionado ao cômico.

Contentamento:

- Tende a surgir sob condições que incluem alta segurança e baixo esforço. São momentos de calma, tranquilidade ou relaxamento.

Entusiasmo:

- Pode ser despertado quando há uma novidade ou um desafio, essa emoção surge como resposta a esses cenários.

Alívio:

- Geralmente precede outras emoções e acontece quando algo que tinha despertado fortemente alguma emoção, se aquieta ou se resolve de forma positiva.

Assombro:

- Quase tudo que é incrível, incompreensível e fascinante pode ser fonte para despertar o assombro.

Êxtase:

- É alcançado por meio de meditação, por experiências na natureza ou experiência sexual com a pessoa amada.

Fiero:

- Essa emoção é sentida quando se enfrenta e supera um desafio. Semelhante a honra própria, triunfo, satisfazer-se com os próprios feitos, orgulho no sentido positivo.

Naches:

- *Naches* acontece quando se sente orgulho de um filho, aluno ou discípulo por uma conquista importante dele. Também pode ser sentida de forma inversa, por isso é bidirecional.

Elevação:

- Pode sentir elevação, quem testemunha um ato moral exemplar realizado por outra pessoa. É comumente provocada no contexto em que, o indivíduo ajudado carece de necessidades vitais.

Gratidão:

- É a apreciação de um presente altruísta, vivenciada por quem recebe algo de bom de outra pessoa, como uma doação ou ajuda altruísta. A gratidão motiva às pessoas a retribuírem, com palavras ou gestos a quem doa ou ajuda.

Schadenfreude:

- Alegria maliciosa, invejosa, satisfazer-se no sofrimento do outro.

3.7 Classificação manual e análise prática

Depois de removido os *tweets* repetidos, sobrou uma base de dados com 19.610, como foi discutido em 3.5 Remoção de repetidos. Através de um código desenvolvido em Python foi feito a randomização dos *tweets* com o método `randint()` da biblioteca `random` nativa do Python, para uma seleção aleatória a fim de serem lidos, interpretados e classificados manualmente.

A partir dos gatilhos emocionais apresentados na seção 2.9 e abordados anteriormente na seção 3.6 é possível interpretar os *tweets* para identificar a emoção presente. Para que aja uma boa interpretação é necessário conhecimento em comunicação, especificamente na forma como os utilizadores do Twitter se comunicam, neologismos, gírias, etc. Também requer conhecimento do assunto, que envolve o ambiente político Brasileiro, dos grupos políticos estabelecidos, das culturas que envolvem esses grupos, conhecimento dos

poderes que regem o país, conhecimento das intenções dos órgãos e pessoas públicas, estar atento nas notícias desse meio, etc.

Agora será mostrado alguns tipos de *tweets* mais comuns presentes na base de dados e classificados manualmente por emoção, seguido de uma breve explicação do contexto e motivo da atribuição de determinado rótulo.

Exemplo:

2 - “Mensagem do utilizador”

2 - Explicação do contexto e o porquê do rótulo

3.7.1 Classificação da Tristeza

Pouquíssimos *tweets* foram encontrados com a emoção tristeza, a maioria relacionados a decepção, no sentido de perda, afastamento. Apresentam-se em seguida alguns desses *tweets*:

1 - *“Eu assisti um pedaço da entrevista da Bianca Boca Rosa no programa pânico, confesso q o Emilio mudou negativamente e ele não era assim antes de 2016. Agora ele virou bozonarista e piorou de vez q eh uma pena. POSSO FALAR COMO CONVIDADA”*

1 - *O utilizador relata perda no sentido de afastamento de caráter, no caso Emilio, apresentador do programa pânico, que mudou na perspectiva do utilizador, ele agora possui um posicionamento Bolsonaroista, o que está longe do posicionamento do utilizador. Nesta entrevista e equipe do programa e a convidada Bianca, debatem sobre machismo. Uma frase dela num momento quente do debate ficou famosa e foi usada como hashtag que virou tendência no Twitter, que é “Posso falar como convidada”. No calor do debate os homens não estavam deixando ela falar, então ela disse essa frase. O sentido subjetivo aqui está relacionado justamente a não deixarem ela terminar o raciocínio que começou. E isso deu a impressão de que, a fala dela não é importante por ser uma mulher, o que é um traço característico da cultura machista. Então para Bianca, como ela não poderia falar como mulher, então poderia falar como convidada do programa.*

22 - *“cada vez que eu leio sobre o caso do Josué eu fico mais com o coração na mão e uma vontade de chorar absurda. o menino foi morto por um policial pq ele estava incomodando, fazendo barulho demais não aguento mais ler sobre. minha garganta deu um nó mt grande. #JusticaPorJosue”*

22 - *O utilizador relata sensações corporais e um sentimento de perda, característico da tristeza. Parece haver uma demonstração sutil de revolta na segunda frase, porém a tristeza é predominante. Ele refere-se ao caso de um menino de 16 anos que foi morto com um tiro*

na nuca por um policial penal. Essa notícia foi usada por grupos políticos no debate sobre racismo e causou muita revolta e um pouco de tristeza nas redes sociais.

82 - *“Trabalhei junto com Gilmar Mendes. Ele era assessor jurídico da Casa Civil no governo Collor. Duas vezes por semana nos reuníamos para traçar estratégias jurídicas na CPI do PC, ou no meu gabinete ou no dele. Sua mudança de caráter foi radical. Ele era simples e honesto.”*

82 - *O utilizador relata a perda de uma pessoa próxima, no sentido de afastamento. Como se Gilmar Mendes, o velho amigo honesto, tivesse mudado tanto o próprio caráter, que já não o considera mais como amigo. Gilmar Mendes é ministro do Supremo Tribunal Federal, e grupos de direita pró Bolsonaro tem acusado ele de imparcialidade e de favorecer partidos e pessoas de esquerda.*

41 - *“Tristeza e indignação, juntas e misturadas. Engasgando, apertando a garganta, dificultando a respiração. Esta é a sensação ao olhar esta imagem. #ForaBolsonaroGenocida #HaiaCondeneBolsonaro”*

41 - *O utilizador relata suas emoções e as sensações corporais sentidas ao olhar uma foto de muitas lápides, supostamente para os mortos de coronavírus no Brasil. Através do texto o utilizador expressa tristeza pela perda de muitas vidas. Apesar falar sobre indignação, isso não é evidenciado no texto, apenas nas hashtags. #HaiaCondeneBolsonaro representa um protesto em que a oposição fez um apelo ao Tribunal Penal Internacional (TPI), situado em Haia, Países Baixos, para que considere as ações de Bolsonaro perante a pandemia como genocidas.*

3.7.2 Classificação da Raiva

Apresentam-se em seguida alguns desses tweets:

390 - *“Expulsar as pessoas de suas casas em plena pandemia é o maior dos absurdos. #DespejoZero já!”*

390 - *Neste tweet o utilizador está expressando indignação, pelo fato de famílias estarem sendo despejadas por atrasarem aluguéis. #DespejoZero foi a hashtag usada na campanha pela esquerda a fim de se oporem a ação de despejo.*

940 - *“Aproveitem e denunciem o tweet da Sara Winter que expôs o endereço da menina, colocou a vida dela em risco. Denunciem”*

940 - *O utilizador está indignado com a atitude de Sara Winter em revelar a morada de uma criança vítima de estupro, que iria abortar. E assim convoca outros utilizadores a punirem a ativista Sara Winter através da denúncia do tweet feito pela ativista como forma de justiça. Sara Winter liderou um movimento anti abortista pró Bolsonaro, no qual defendiam que essa criança não deveria abortar. O aborto divide opiniões, em geral os grupos de esquerda são favoráveis e de direita contrários.*

13 - “a gente se lembra que esse país é uma merda, quando automaticamente sabe que hoje vai ter muito fdp fazendo pedido de propósito, só pra ir de contra os trabalhadores que estão organizando o #BrequeDosApps em busca de melhores condições e direitos.”

13 - O utilizador supõe que muitas pessoas irão contra a greve dos motoqueiros entregadores que trabalham para aplicativos de delivery e faz uma ofensa como meio de causar dano psicológico e elas. Na perspectiva do utilizador, as pessoas farão pedidos pelos aplicativos como forma de sabotar a greve. #BrequeDosApps representa essa paralisação e foi organizada por movimentos de esquerda principalmente na região da cidade de São Paulo.

242 - “Acho engraçado muita gente falando da thammy, sobre a relação de dia dos pais. Conheço tanto “MACHO” pagando de boy, fazendo filho e caindo fora, se achanado ganhão, e nem sequer está presente na vida dos seus filhos. Pra mim PAI/MÃE é quem cria!”

242 - Indignado o utilizador faz uma comparação entre Thammy Miranda, que é uma pessoa transexual e homens que engravidam as parceiras e não assumem a paternidade. Thammy Miranda anunciou que participaria de uma campanha publicitária da Natura do dia dos pais e isso dividiu opiniões nas redes, em que grupos de direita desaprovaram a ideia de uma pessoa transexual fazer o papel de pai e grupos de esquerda que foram contra a opinião dos de direita, seguindo a linha de ser a favor da diversidade. Pessoas públicas e alguns influenciadores envolvidos com o governo Bolsonaro representaram a opinião da direita.

1100 - “Esse Desgraçado @jairbolsonaro Usou a Lava-jato para se Eleger e hj é o MAIOR Genocida que o Brasil já Teve #LavaJatoSempre”

1100 - O utilizador expressa ressentimento e agressividade ao ofender Bolsonaro e ainda mencionar o perfil dele através do @jairbolsonaro, demonstrando um enfrentamento direto ao político. Este utilizador também manifesta apreço pela Operação Lava Jato, reforçando a ideia de ressentimento. Muitos acreditam que a Operação Lava Jato gerou revolta na população brasileira, possibilitando a entrada de Bolsonaro na presidência.

3.7.3 Classificação da Surpresa

Detetar *tweets* com a emoção surpresa não teve resultados satisfatórios. Alguns *tweets* com textos escritos com todas letras maiúsculas ou com pontos de exclamação, tentavam despertar essa emoção, mas as pessoas reagiam com raiva, orgulho, desprezo, etc, em vez de expressarem surpresa. Não se sabe o porquê de realmente as pessoas não relatarem surpresa ou se relataram em *tweets* que não foram analisados. Poderia deduzir que tenha alguma relação com a surpresa ser muito breve e geralmente antecede emoções que tem um potencial maior para fazer as pessoas reagirem, como é o caso da raiva. Em uma circunstância que isso fosse verdade, o utilizador poderia sentir surpresa no primeiro momento e logo em seguida sentir raiva e escrever o *tweet* motivado pela raiva em vez

de surpresa. Os *tweets* com títulos muitos polêmicos ou sensacionalistas, pareciam ter potencial para causar surpresa, mas também, potencial para causar raiva, desprezo ou orgulho, o que geralmente prevalecia nas respostas. Além disso, surpresa é uma reação mais física, que coloca a pessoa em estado de alerta momentâneo, por isso há uma certa dificuldade em comunicá-la através de palavras.

3.7.4 Classificação do Medo

Semelhante a tristeza, no medo houve pouquíssimos *tweets* identificados e ainda com poucas evidências. Poderia deduzir que as poucas ocorrências, podem ter alguma relação com o medo também possuir reações mais físicas, como paralisar ou fugir, por isso seria mais difícil de expressá-lo em palavras. Além disso o medo faz com que a pessoa fique num estado de total foco na ameaça, limitando assim algumas capacidades cognitivas.

Apresentam-se em seguida alguns desses *tweets*:

6 - *“Difícil acreditar nesse pesadelo que estamos vivendo! #GovernoDeFraudes”*

6 - *O utilizador compara o governo Bolsonaro a um pesadelo, algo que causa medo e terror. Neste caso ele estava falando sobre uma matéria sobre três milhões de pessoas que tiveram o auxílio emergencial bloqueado.*

19 - *“@felipeneto Só vai ter fim esse pesadelo, quando Bolsonaro sair da cadeira que ocupa. #HaiaCondeneBolsonaro”*

19 - *Através da interpretação do utilizador, o pesadelo representa o dano psicológico que ele sofre e pode estar ligado ao sofrimento da população brasileira pela má gestão do governo Bolsonaro, sendo este por dano físico ou psicológico, como mortes por covid-19, desemprego, poucos recursos para os mais desfavorecidos e etc. O tweet é uma resposta a outro utilizador mencionado no início, por isso há pouca informação sobre o contexto dessa resposta.*

22 - *“A segurança das crianças de sua família depende de você. #TodosContraFelipeNeto A mensagem subliminar da fábula de Chapeuzinho Vermelho é sobre o MAL em sexualizar as crianças. Não deixe o Lobo Mau com seus filhos. Use o material do MEC feito pelo professor @CarlosNadalim.”*

22 - *Esse o utilizador parece ter medo e utiliza-o como argumento, para provocar medo. No sentido de que o influenciador Felipe Neto, tem conteúdo que pode “sexualizar as crianças”, pois o público desse influenciador são crianças e adolescentes.*

23 - *“@Calladoproject Essa ideia é assustadora! #GovernoDeFraudes”*

23 - *O utilizador declara ser assustador um projeto de lei da direita chamado Escola sem Partido. Esse projeto reivindica a imparcialidade e a objetividade do professor em sala de*

aula alegando que, caso contrário, será negado ao aluno o acesso a outras explicações e abordagens alternativas para os fenômenos estudados. No âmbito do debate, de modo geral, grupos de direita acusam grupos de esquerda de doutrinar os alunos com ideologias de esquerda, em resposta a isso surgiu esse projeto, para buscar a imparcialidade dos professores. Na visão do utilizador, esse projeto estaria limitando o conhecimento dos alunos ao impedir professores de darem maiores explicações para os fenômenos estudados, por isso é uma ideia “assustadora”.

3.7.5 Classificação do Nojo-desprezo

Apresentam-se em seguida alguns desses tipos de tweets:

464 - “Malala levou um tiro e foi perseguida pelo Talibã. Felipe Neto e Greta, no máximo, foram bloqueados pelo Bolsonaro.”

464 - O utilizador despreza Felipe Neto e Greta ao comparar eles com Malala Yousafzai, atribuindo um título de superioridade a Malala comparando o que ela já passou. O assunto que esse tweet aborda, está relacionado com matéria e a capa publicada na revista INSTOÉ, que trouxe a foto de Greta Thunberg, Felipe Neto e Malala Yousafzai um ao lado do outro com o seguinte título “Geração Transformadora”.

104 - “Se 2022 for Huck, Mandetta e Doria, amigo... Não vai precisar nem de campanha. Bonoro 2022 já é realidade demais.”

104 - Semelhante ao tweet anterior, o utilizador faz uma comparação aos possíveis candidatos a presidência nas eleições de 2022, Huck, Mandetta e Doria comparados ao Bolsonaro, “Bonoro”. E despreza eles, no sentido de que são candidatos com pouquíssimas chances de ganhar se estiverem disputando contra o Bolsonaro, atribuindo o título de superior ao Bolsonaro.

24 - “@AndradeRNegro2 Huck teve uma chance de ouro de demonstrar consciência e discernimento. Falhou. A opinião dele passou a valer o mesmo que um papel higiênico usado. E é o que eu faço com a opinião do Luciano Huck, direto para a lixeira. Sem chance de reciclagem. Adubo sanitário ou incineração.”

24 - O utilizador desaprova totalmente a opinião de Luciano Huck, comparando a algo que normalmente despertaria aversão básica, um “papel higiênico usado”, caracterizando uma relação com aversão. O utilizador desaprova a opinião de Huck porque, julga que ele não agiu da forma moralmente correta. Em 2020 Huck começou articulações e mostrou a própria opinião, a fim de se candidatar para as eleições presidenciais de 2022.

704 - “Nojo é o que sinto assistindo ao vídeo da Bia Doria e uma amiga, dentro de uma mansão luxuosa, falando sobre moradores de rua gostarem de viver nessa situação, e deslegitimando as ações de solidariedade como a distribuição de alimentos. Elite nojenta e asquerosa!”

704 - *Aqui a aversão interpessoal fica clara pelo utilizador, que desaprova moralmente as ações de Bia Doria e destaca adjetivos relacionados a aversão básica em relação a elite, como “nojenta e asquerosa”. Bia Doria gerou polémica depois de dizer que não é correto dar comida as pessoas de rua. Bia Doria é casada com o governador do estado de São Paulo, por isso é a primeira-dama, ela também é presidente do Fundo Social de São Paulo.*

494 - *“O pior do Brasil, por mais incrível que possa parecer, não é seu Jair, é Paulo Guedes. Esse se alastra silenciosamente feito radiação, contamina com 100% de eficiência e é 100 % mortal. A dosagem nem precisava ser grande, mas além de ser mortal, ele é mal. #GuedesRoubaOBrasil”*

494 - *O utilizador mostra aversão interpessoal a Bolsonaro e Paulo Guedes, ministro da Economia. O principal alvo é Guedes, que é destacado por uma metáfora ligada a aversão básica, com as palavras “alastra”, “radiação”, “contamina” e por fim pode causar dano ou letalidade, “além de ser mortal”.*

3.7.6 Classificação do Alegria

Fiero não é uma emoção tão presente quanto as outras da família da alegria. Seguem alguns exemplos:

358 - *“Quem disse q a mobilização aqui nas redes sociais não funciona né? Foi uma vitória para #fundebaprovado Não conseguimos lotar as ruas por conta da pandemia, mas pegamos pesado aqui. Vamos continuar Lutando. #HaddadAbsolvido”*

358 - *O utilizador relata uma vitória pela aprovação do Fundeb (Fundo de Manutenção e Desenvolvimento da Educação Básica), o que caracteriza a emoção Fiero, vencer um grande obstáculo, mas em vez de ser uma conquista individual, ela é coletiva, pois neste tweet, o utilizador de esquerda supõe ser uma vitória apenas do grupo de esquerda.*

742 - *“@novaeramidia @Bezerrapqd1 Bora #PatriotasComBolsonaro é na rua que vamos vencer”*

742 - *O utilizador parece sentir fiero antecipadamente e também em grupo. Ele convoca mais dois outros utilizadores que fazem parte do mesmo grupo de apoiantes de Bolsonaro, para que vão à rua participar de uma manifestação, pois através disso serão vencedores.*

630 - *“#HaddadAbsolvido #Haddad2022 Cabeça erguida e vamos vencendo todos obstáculos 🇧🇷🇺🇸🇵🇷🇨🇦🇩🇪”*

630 - *O utilizador sente-se vencedor por fazer parte do grupo que apoia Haddad. A luta foi o julgamento desse político na chamada “Indústria da Multa” e a vitória é representada pela absolvição dele diante do tribunal em São Paulo.*

Naches está entre as emoções mais presentes coletadas dentro da família da alegria. Ela normalmente existe no sentido do filho, aluno ou discípulo sentir orgulho do professor,

mentor ou mestre e no âmbito da discussão política ela se manifesta de uma forma adaptada. As pessoas sentem orgulho dos políticos que apoiam, por isso essa emoção é vista muitas vezes caracterizada pela relação do utilizador que tem orgulho do político ou líder. Como se o político representasse para o utilizador alguém grandioso, digno de se ter orgulho. Segue alguns exemplos de *naches*:

67 - “@LiloVLOG Estou junto com o Presidente Jair Messias Bolsonaro. Junto com o BRASIL!!!”

67 - O utilizador expressa orgulho em apoiar Bolsonaro e parece gostar de fazer parte de um grupo que também apoia o presidente, sendo todo o “BRASIL” na visão do utilizador. As características se assemelham a *naches*, por ter orgulho do que Bolsonaro representa para ele, de acordo com avaliações anteriores, que podem ser sobre os valores morais de Bolsonaro, ações do passado ou presente, etc. Nesse caso Bolsonaro representa o papel de professor, mentor ou mestre e o utilizador o papel de filho, aluno ou discípulo.

433 - “@brasil247 Cada um dá o que tem. A conduta política do Suplicy é exemplo de nobreza para o povo brasileiro. #ForaBolsonaro”

433 - Semelhante ao tweet anterior, esse expressa orgulho do Suplicy, com a mesma relação, no sentido discípulo mestre.

968 - “Melhor presidente do Brasil... #GoBolsonaroMundial”

968 - O padrão de orgulho *naches* se repete muito dentro dos tweets classificados como alegria.

Schadenfreude também é visto excessivamente dentro das emoções classificadas como alegria. Os políticos possuem oposições, pessoas que são contra os interesses deles, portanto muitos mantêm uma relação de rivalidade. Os apoiantes desses políticos também mantêm uma relação de rivalidade com os políticos opositores e com todo o grupo desses opositores. Isso causa um grande movimento de rivalidade entre os grupos, semelhante a adeptos de dois clubes de futebol quando disputam uma partida. Esse cenário cria um ambiente propício para a alegria maliciosa *schadenfreude*, que surge ao verem os opositores tentarem algum tipo de perda, dano ou algo que cause mal a eles ou a reputação deles. Segue alguns exemplos:

226 - “@AmendoncaMJSP @danielPMERJ @policiafederal E daí que o Bolsonaro está com Covid-19 ? Eu não sou coveiro. Todo mundo tem que morrer um dia. Essa questão do coronavírus é muito mais uma fantasia que a grande mídia propala ou propaga. É uma gripezinha, é histeria. 🤔👉👉👉👉 #forcacovid #ForaBolsonaro #forcacورونا”

226 - O utilizador sente prazer ao saber da notícia de que Bolsonaro contraiu coronavírus, utilizando de ironia ele escreveu todo tweet com frases que Bolsonaro pronunciou em discursos variados, exceto a primeira frase onde houve uma adaptação. O utilizador expressa essa emoção também através das hashtags que ficaram populares na época em que essa notícia veio à tona, “#forcacovid” e “#forcacorona”. Quando alguém usa essas hashtags, está fazendo uma conotação para que o coronavírus tenha força e cause dano ou a morte de Bolsonaro, isso se encaixa com a alegria maliciosa de Schadenfreude.

493 - “Como é bom começar o dia sabendo que as máscaras dos grandes moralistas tbm caem. #DerreteMBL Sr. Kataguirí negando até que conhece os caras kkkkkkk”

493 - O utilizador sente prazer em saber que empresários ligados ao partido MBL são presos por um esquema de lavagem de dinheiro, caracterizando Schadenfreude. Kataguirí é integrante desse partido.

683 - “A Natura tá fazendo d vai continuar fazendo rios de dinheiro e esse povo tá ai espumando kkkk #EuNaoComproNatura”

683 - O utilizador sente alegria porque um grupo está com raiva, por não aprovarem a ideia da Natura sobre a campanha de dia dos pais ser feita com uma pessoa transexual, Thammy Miranda. Este grupo que manifesta a opinião contrária à do utilizador e a campanha, representa os opositores ou inimigos deste utilizador. A palavra “espumando”, refere-se à uma gíria, “espumando pela boca”, que significa que estão com raiva. #EuNaoComproNatura foi a hashtag usada por quem ia contra a campanha.

3.7.7 Classificação Neutro

Apresentam-se em seguida alguns desses tipos de tweets:

337 - “Hoje as 16h todos ligados na TV Câmara! Audiência de copa mundo. #AprovaFundeb #VotaFundeb”

337 - O padrão de tweet informativo é muito comum, são jornais, instituições, blogues, etc, que fazem publicações muitas vezes com um link redirecionando os utilizadores para o website da notícia.

369 - “@ancapball_br não confiem na mídia só acredito no Átila”

369 - É uma resposta a um tweet, em que o utilizador não demonstra emoção, apenas um voto de confiança ao Átila.

45 - “O Secretário de Educação de Porto Alegre, Adriano Naves de Brito, e também o próprio @fernandoschuler já se manifestaram mais profundamente sobre os problemas do projeto. Da mesma forma, o @novonacamara tem se manifestado em defesa de um FUNDEB que permita melhorar a educação.”

45 - *Esse é uma resposta de caráter informativo e sem sinais de emoção, traz uma informação parcial que soa como uma defesa aos políticos citados, no sentido de que eles se preocupam com a qualidade do projeto FUNDEB.*

187 - *“Bia Kicis cai da vice-liderança no governo @JairBolsonaro Os motivos da queda de @Biakicis seria o filho”*

187 - *Outro tweet informativo que é acompanhando de uma imagem de um print de uma matéria jornalística com o mesmo tema.*

135 - *“@Metropoles @jairbolsonaro @BolsonaroSP @ThammyReal Dia dos pais... Dia das mães... Devemos respeitar os dois lados, o que a pessoa escolheu e a data comemorativa de fato... A Natureza é livre pra propagar o que quiser e eu de aceitar se quiser. Não vale a pena discutir nada hoje em dia.”*

135 - *Algumas opiniões argumentativas parciais ou imparciais não demonstram padrões emocionais. Em vez disso têm uma abordagem mais lógica, portanto são neutras, ausentes de emoções. Talvez em um nível mais profundo, pragmático, possa ser motivado por uma emoção, mas isso não é visto no texto.*

3.7.8 Não classificados

Muitos *tweets* analisados não se encaixaram com os padrões citados na bibliografia das famílias das emoções por vários motivos, além de não se encaixarem como neutros. Outros passaram pelo filtro que removia os *tweets* repetidos, pois eram compartilhamentos em que o utilizador acrescentou apenas uma palavra no fim do *tweet* ou algo semelhante, fazendo com que o texto passasse despercebido pelo filtro. Esses receberam um rótulo de “não classificados” e agora será observado os principais.

Apresentam-se em seguida alguns desses tipos de *tweets*:

742 - *“Verdade! #ConexaoComPT”*

742 - *Houve muitos tweets que são respostas a outro tweet. Algumas vezes mostrando aprovação a um tweet, mas sem padrões de emoções claros ou totalmente sem padrões. Outras somente com uma ou mais hashtags, caracterizando-se com pouca informação.*

1541 - *“Estão esperando inflação? Qual o sentido de uma nota de R\$ 200,00?”*

1541 - *Outros tweets são perguntas diretas, sem ironia, sem emoções. Essa além de ser uma pergunta, também foge do assunto político, focando-se apenas em economia.*

374 - *“@pamcosta21 Boa noite. #DerreteMBL”*

374 - *Nesse caso o utilizador não tem nenhuma argumentação ou crítica, ele simplesmente compartilha uma hashtag com um amigo “@pamcosta21”, acompanhado de uma saudação de boa noite. Por isso não demonstra emoções.*

1414 - “Minha filhinha de 7 anos sabe que não prestam, nem ele e nem seu irmão. Ela não conhece a Globo, nem novelas, nem funk. Vira o rosto quando vê um beijo no filme. Pede para tirar quando vê violência. Isso é proteção! Tudo ao seu tempo. #FamíliasContraFelipeNeto”

1414 - Outros tweets que utilizam uma hashtag que foi criada no ambiente de discussão política, mas o utilizador aborda um assunto pessoal ou não focado em política.

695 - “Assista a Flávio Leandro - Chuva de Honestidade - legendado no YouTube #NordesteComBolsonaro”

695 - Esse é mais um exemplo de compartilhamento externo. Há um link para o YouTube e leva a uma música que fala sobre a corrupção no Nordeste do Brasil, mas dessa vez é um utilizador comum compartilhando uma música. Portanto não há relação direta com política no texto, ele faz uso apenas de uma hashtag criada no ambiente político.

1977 - “@EliezerMolivei1 @joicehasselman @RzLuana CARLA ZAMBELLI é uma mulher amarga, asquerosa, ... Vive nas trevas e carece da Graça de Deus. @EliezerMolivei1, GRATIDÃO por me SEGUIR! Estamos juntos por um Brasil melhor, em defesa Lava-Jato e Sérgio Moro 2022.”

1977 – Alguns tweets evidenciam duas emoções, o que complica a análise de máquina, porém não são tão comuns. No início há uma relação com nojo-desprezo, pela palavra “asquerosa” e por que Carla Zambelli não tem os mesmos valores morais religiosos que o utilizador, de acordo com a visão dele, ela “Vive nas trevas”. E a segunda parte ele evidencia alegria, pela gratidão porque o utilizador @EliezerMolivei1 o segue no Twitter e também demonstra orgulho em defender “um Brasil melhor” e fazer parte do grupo de apoiantes de Sérgio Moro, que de acordo com o utilizador, será candidato a Presidência da República em 2022.

3.8 Análise exploratória

Aqui veremos algumas contagens gerais dos dados coletados. No geral foram coletados 42.610 *tweets* completos. Depois de filtrar os repetidos sobrou 19.610. Então foi feita uma análise manual, onde foram lidos e etiquetados 3.672 *tweets*, em sete classes. Que são tristeza, raiva, medo, nojo-desprezo, alegria, neutro e não classificados, para a emoção surpresa não foi possível identificar um padrão, como discutido anteriormente na secção 3.7.3, Classificação da Surpresa . Veja na Tabela 3.1 abaixo um resumo dos *tweets* coletados e etiquetados.

Tabela 3.1: Quantidades e classes de tweets etiquetados manualmente.

Classes	Quantidade	Porcentagem
Total coletado	42.610	
Não repetidos	19.610	
Total etiquetados	3.672	100%
Tristeza	85	2,31%
Raiva	1335	36,36%
Surpresa	0	0%
Medo	31	0,84%
Nojo-desprezo	764	20,81%
Alegria	1060	28,87%
Neutro	397	10,81%

Aplicando a *tokenização* nos *tweets* com o método padrão do Python `.split()`, é possível fazer a contagem para visualizar o tamanho desses *tweets*. *Tokens* são basicamente as palavras como no exemplo abaixo:

Frase antes de ser tokenizada:

“Nesta frase há alguns tokens, vírgulas e pontos ficam juntas as palavras.”

A mesma frase depois de ser tokenizada:

“Nesta”, “frase”, “há”, “alguns”, “tokens,”, “vírgulas”, “e”, “pontos”, “ficam”, “juntas”, “as”, “palavras.”

Os elementos entre aspas são os *tokens*, note que essa simples técnica de *tokenização* inclui a pontuação junto das palavras, por ora é suficiente para termos uma noção do comprimento dos *tweets*. Abaixo na Figura 3.1 podemos ver a distribuição da quantidade de tokens no histograma.

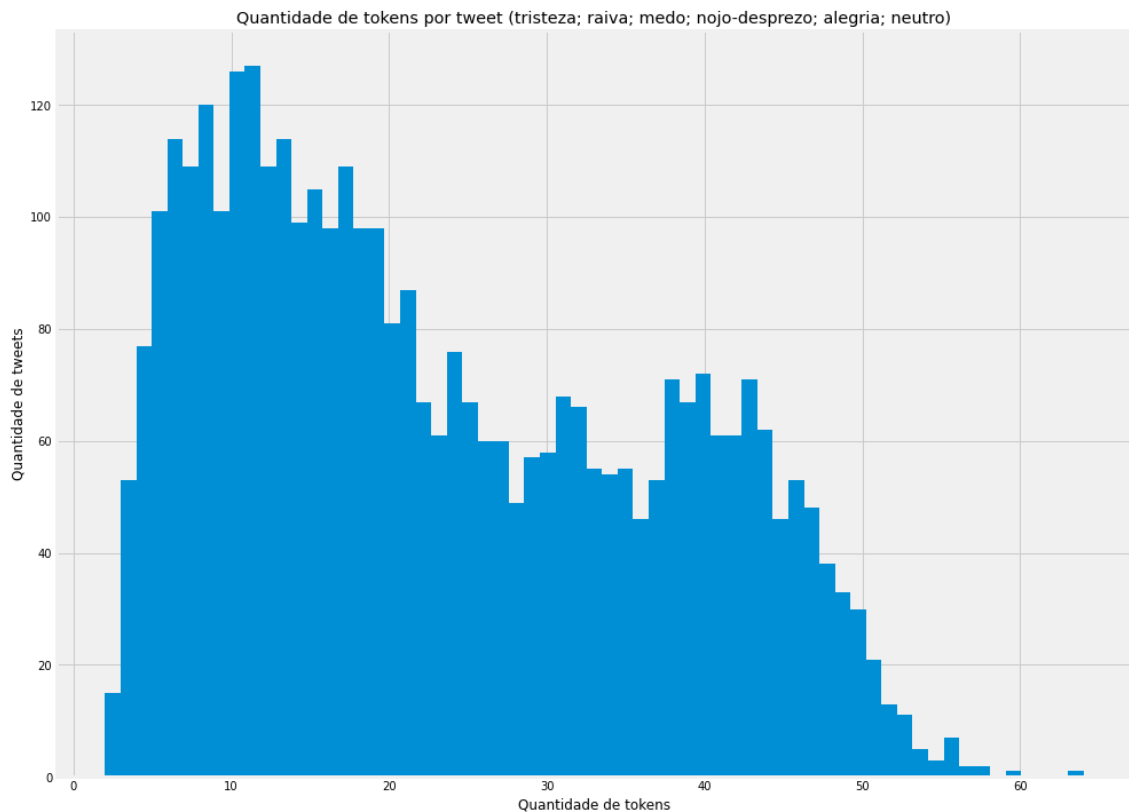


Figura 3.1: Histograma da quantidade de tokens.

No histograma feito a partir dessas quantidades de *tokens*, vemos que há muitas ocorrências no intervalo entre 4 e 23 *tokens*, em que esse volume passa de 60, diferenciando do restante da amostra. Possui um pico nos *tweets* com 10 e 11 *tokens* que passar se 120 ocorrências para cada um. Existem pouquíssimos *tweets* no intervalo entre 53 até 64 *tokens*.

Abaixo na Tabela 3.2 é possível ver a estatística descritiva da contagem desses *tokens*, feito com a biblioteca Pandas com o método `pandas.DataFrame.describe()` para as mesmas 5 emoções mais a classe neutro.

Tabela 3.2: Estatísticas descritivas da quantidade de tokens do conjunto de emoções.

Estatísticas	Quantidade
Contagem de tweets	3672
Média de tokens	23,20
Desvio padrão	13,83
Mínimo	2
25%	11
50% (mediana)	20
75%	35
Máximo	64

Nos extremos percebe-se que há *tweets* muito longos, com 64 palavras e muitos curtos com 2 palavras. Esses extremos dificultam a análise de máquina, por terem demasiado ou pouquíssimas palavras, forçando o algoritmo a extrapolar no momento da previsão.

3.9 Análise do significado das palavras

Baseado no significado e contexto dessas palavras mais frequentes podemos separá-las em grupos, como palavras com conotação negativa, positiva, imediatismo, sentimento de grupo, distanciamento de um grupo, generalização, pessoas públicas, grupos ou organizações e assuntos em pauta. Aqui as *hashtags* serão consideradas como palavras, pois possuem significado dentro do contexto e algumas vezes são a junção de mais de uma palavra adquirindo assim um novo significado.

Tomemos esse *tweet* como exemplo. Nele é utilizado a *hashtag* #forcacorona, que é a junção de duas palavras, “força” e “corona”, referindo-se ao coronavírus.

“#forcacorona contamos com você”

Essa *hashtag* surgiu depois que foi anunciado que o Bolsonaro, contraiu o vírus. Portanto o utilizador escreve como se tivesse falando com o vírus, no sentido de estar esperando uma ação de adoecimento ou morte do Bolsonaro causada pelo vírus. Aqui “forcacorona” e também uma variante “forcacovid” assumem um significado de desejar o mal de uma pessoa pública. Muitas outras frases foram postadas, dessa mesma forma, assumindo o significado implícito na *hashtag*, como:

“#forcacorona estamos com você!!”

O corpus é constituído por quatro emoções, correspondendo àquelas que mais tiveram ocorrências no momento da etiquetagem: raiva, nojo-desprezo, alegria e neutro. As emoções tristeza, surpresa e medo tiveram poucas ocorrências, por isso não foram levadas em consideração, uma vez que isso iria confundir demasiado os algoritmos de aprendizado da máquina pela quantidade muito pequena de ocorrências, como é o caso do medo que teve 31. Logo, quando explorarmos as emoções separadas estaremos a falar dessas quatro.

3.9.1 Palavras mais frequentes no corpus

Nesta seção de frequências foi utilizado o pacote nativo do Python NLTK para *tokenizar* e fazer a contagem das palavras com o método `FreqDist()`. A Figura 3.2 mostra as 80 palavras mais frequentes das emoções trabalhadas no aprendizado de máquina, que foram alegria, raiva, nojo e neutro. Essas palavras foram extraídas de 3556 *tweets*. O corpus em que foi feita a contagem, passou por uma eliminação de *stopwords*. O critério para seleção dessas 80 palavras foi, de entre as palavras com menos ocorrências, permanecer na faixa das 50 ocorrências por palavra. Se seleccionasse as 70 primeiras palavras, as com menos ocorrências ficariam com 52 e se seleccionasse 90, essas, ficariam com 46, o que já está um pouco longe da faixa de 50. Portanto com 80 palavras, a que tem menos ocorrências é “thammy” que representa 4,2% da soma dessas 80 mais frequentes, possuindo 48 ocorrências, portanto as 80 satisfazem o critério estipulado.

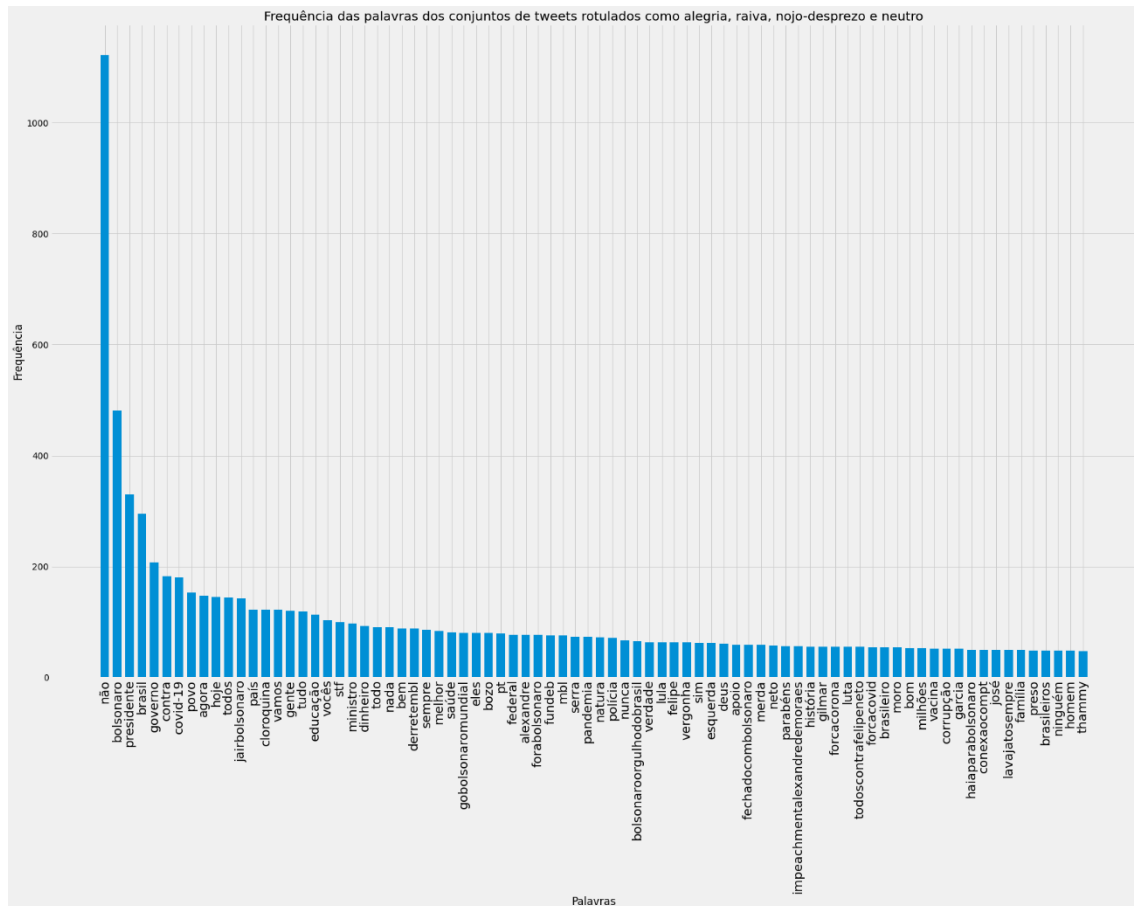


Figura 3.2: As 80 palavras mais frequentes dos conjuntos de emoções utilizados no aprendizado de máquina.

Há palavras que representam conotação negativa, ainda que empregadas em frases com significado positivo elas trazem um peso negativo. Como a palavra “não” que teve mais de mil ocorrências, demonstrando assim uma comunicação com peso mais negativo. Para compreender melhor esse peso negativo podemos usar a própria palavra “não” como exemplo. Observe a diferença entre as seguintes duas frases, uma empregando um sentido mais direto e usando a palavra “não” e a outra utilizando um sentido mais positivo.

“Eu não posso te ajudar, estou sobrecarregado.”

“Eu gostaria de ajudar, mas estou sobrecarregado.”

A formatação aqui será da seguinte forma, primeiro a palavra entre aspas, depois uma numeração correspondente a frequência que ela aparece no corpus ou dentro do grupo de cada emoção, que será visto posteriormente.

Das 80 palavras mais frequentes, as que se destacaram com sentido negativo somam 1527 que é equivalente a 18,11% do total de 8430, sendo que a palavra “esquerda” se repete

em dois grupos, logo é contato duas vezes de 62. As palavras com sentido negativo foram: “não”, 1121, “contra”, 183, “vergonha”, 63, “preso”, 49, “merda”, 59, “corrupção”, 52.

De forma semelhante, há a palavras que em vez de terem uma conotação negativa, possuem uma conotação positiva, como: “bem”, 89, “verdade”, 64, “sim”, 62, “apoio”, 59, “melhor”, 84, “parabéns”, 57, “bom”, 53. Essas, apesar de serem mais palavras, somam 468, o que é três vezes menos do que as palavras com conotação negativa e representam 5,55%.

Algumas palavras que possuem um significado de algo imediato, foram também empregadas dentro desse contexto, como: “agora”, 147, “hoje”, 145, totalizando 292 e representando 3,46%.

Há palavras e *hashtags* que definem um sentimento de pertença a um grupo, união, sendo verbos na primeira pessoa do plural, adjetivos que definem grupos ou *hashtags* utilizadas a favor de um grupo político específico. Esses totalizam 1010 e representam 11,98% são: “vamos”, 122, “povo”, 153, “todos”, 144, “gente”, 120, “brasileiros”, 49, “conexaocompt”, 50, “esquerda”, 62, “lavajatosempre”, 50, “gobolsonaromundial”, 81, “bolsonaroorgulhodobrasil”, 66, “fechadocombolsonaro”, 59, “brasileiro”, 54.

Ao contrário das anteriores, há palavras que mostram um sentido de distanciamento de outro grupo ou pessoa pública, ou mesmo uma forma de ridicularizar, como é o caso de “bozo”, que faz referência a Bolsonaro como sendo o palhaço Bozo. Outras, mostram um afastamento do locutor ou do grupo, como “eles” ou “vocês”. Essas somam 762, o que representa 9,04%. Veja: “vocês”, 103, “eles”, 81, “forabolsonaro”, 77, “esquerda”, 62, “derretembl”, 88, “bozo”, 80, “impeachmentalexandredemoraes”, 57, “haiaparabolsonaro”, 50, “todoscontrafelipeneto”, 55, “forcacorona”, 55, “forcacovid”, 54. Aqui, vale a pena destacar que a palavra “esquerda” foi utilizada tanto no contexto de grupo pelos seus apoiantes, como num contexto de distanciamento pelos contrários.

É comum encontrar as pessoas generalizando nas redes sociais e algumas palavras se destacam, como: “tudo”, 119, “todo”, 91, “nada”, 91, “nunca”, 67, “ninguém”, 49, “sempre”, 86, somam 503 equivalente a 5,97%.

A maior parte do debate gira em torno de pessoas públicas, grupos ou organizações, dos quais destacaram-se dentro das 80 mais frequentes: “bolsonaro”, 481, “presidente”, 330, “brasil”, 295, “governo”, 208, “jairbolsonaro”, 143, “stf”, 100, “ministro”, 97, “pt”, 79, “federal”, 77, “alexandre”, 77, “mbl”, 76, “serra”, 74, “natura”, 73, “polícia”, 71, “lula”,

64, “felipe”, 64, “neto”, 58, “gilmar”, 56, “moro”, 54, “garcia”, 52, “josé”, 50, “thammy”, 48, totalizam 2627, representa a maior fatia de 32,16%. Isso acontece devido ao ambiente de discussão geralmente ter uma entidade que fez algo, deixou de fazer ou esteve envolvida numa ação de forma passiva. Por exemplo, um político que diz ou faz algo, ou é alvo de alguma investigação ou acusação.

Outros são assuntos que não são uma entidade, mas estão em pauta de forma objetiva ou subjetiva. Isto é, são assuntos específicos trazidos pelos veículos de comunicação, influenciadores de opinião ou pessoas públicas, estão relacionados ao agendamento ou agenda-setting. É importante destacar que assuntos influenciados pelo agendamento fazem parte de todo debate e não exclusivamente desse grupo. Como “covid-19”, ou o projeto de lei “fundeb”, o setor de saúde pública como referenciada apenas por “saúde”. Esses totalizam 701 equivalente a 8,32%. São: “covid-19”, 180, “cloroquina”, 123, “educação”, 114, “saúde”, 82, “fundeb”, 76, “pandemia”, 74, “vacina”, 52.

Além dessas, há palavras que são auxiliares, substantivos que estão envolvidos subjetivamente no contexto do debate ou mesmo nos valores dos grupos, como as palavras “família” e “Deus”. Esses totalizam 540 equivalente a 6,41% e são: “país”, 123, “dinheiro”, 93, “deus”, 61, “história”, 56, “luta”, 55, “milhões”, 53, “família”, 50, “homem”, 49.

Na Tabela 3.3 abaixo, podemos ver um resumo dessas frequências de acordo com os significados discutidos.

Tabela 3.3: Frequência das 80 palavras mais usadas no corpus, por significado.

Significados	Quantidade	Porcentagem
Negativo	1527	18,11%
Positivo	468	5,55%
Imediato	292	3,46%
Sentimento de grupo	1010	11,98%
Distanciamento de outro grupo	503	5,97%
Pessoas, grupos ou organizações	2627	31,16%
Outros assuntos	701	8,32%
Auxiliares	540	6,41%
Total	8430	100%

As entidades como pessoas, grupos e organizações, palavras com conotação negativa e assuntos em pauta se destacam no grupo, totalizando juntas 64%.

3.9.2 Palavras mais frequentes separadas por emoções

Nesta secção serão discutidas as frequências das palavras em cada uma das principais emoções encontradas e etiquetadas. Dessa forma, é possível destacar as palavras específicas utilizadas no contexto das emoções sem serem enviesadas pelas emoções com maior quantidade, quando analisadas juntas. Portanto, em vez de 80, serão destacadas as 40 palavras mais frequentes dentro de cada emoção. Apesar da quantidade de *tweets* de cada emoção não serem semelhantes, o objetivo em selecionar as 40 palavras mais frequentes foi obter uma boa quantidade de palavras para enriquecer a análise dos significados sem selecionar muitas palavras que possuem poucas ocorrências, ou seja, as que estão na cauda do gráfico. Como seria impossível não selecionar parte dessa cauda dos grupos de emoções com poucos *tweets* sem ter a mesma quantidade de palavra para cada uma, optou-se por selecionar as 40 mais frequentes de todas as emoções. Portanto, para alegria, a palavra com menos ocorrências teve 23 ocorrências e foi “convidada”. A emoção raiva abrangeu 1335 *tweets* e as suas palavras com menos ocorrências tiveram 31 e foi a palavra “puta”. Nojo que teve poucos *tweets* e a palavra de menor ocorrência foi “nunca” com 21 ocorrências. Como neutro foi o menor entre todos, selecionou a maior parte da cauda, com a palavra “câmara” tendo apenas 9 ocorrências.

Começaremos com a emoções raiva. Abaixo na Figura 3.3 estão as 40 palavras mais frequentes desta emoção.

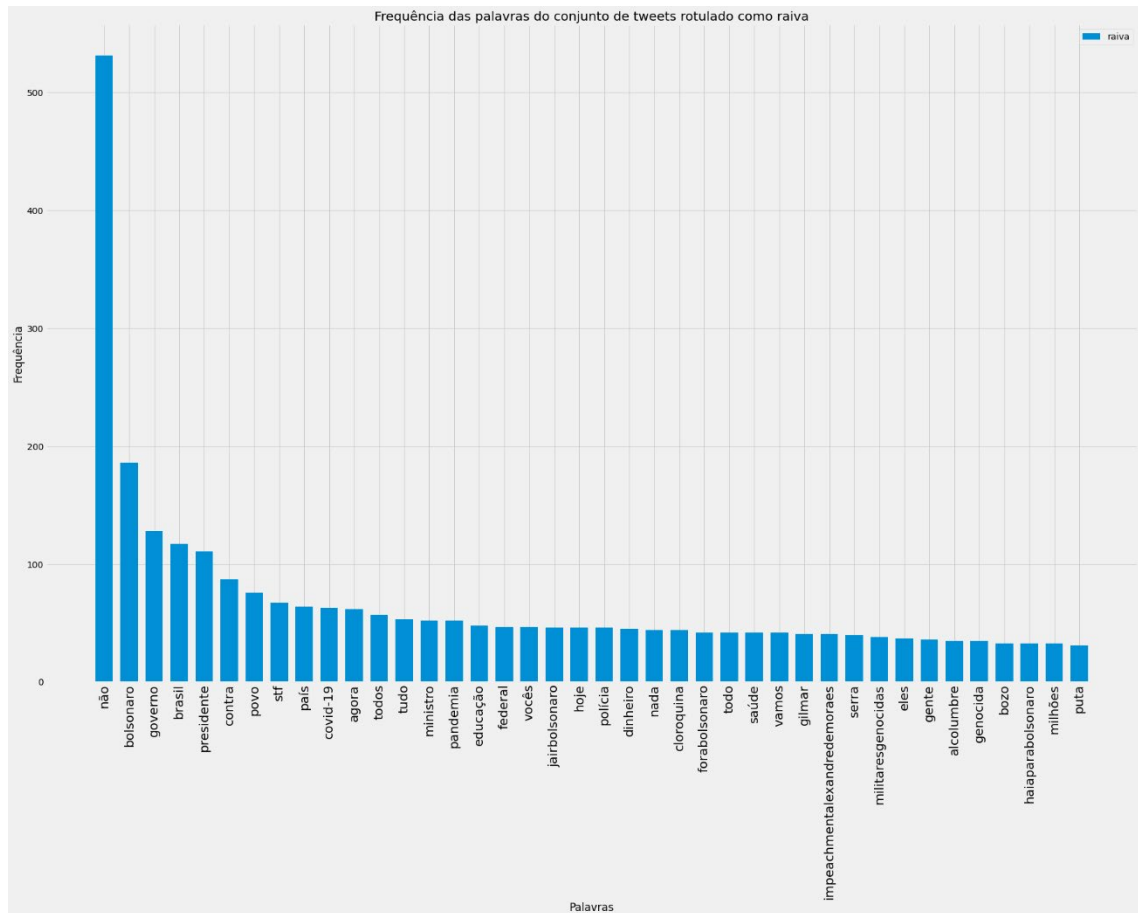


Figura 3.3: As 40 palavras mais frequentes da família das emoções de raiva

Com conotação negativa: “não”, 531, “contra”, 87, “puta”, 31 essas representam 23,86% do total de *tokens* da emoção raiva.

Com conotação positiva não houve ocorrências. Isso demonstra que parte do conteúdo da emoção raiva é negativo, com a ausência de palavras que possuem significado positivo nas 40 mais usadas.

Como urgência, surgiram as palavras “agora”, com 62 ocorrências e “hoje”, com 46 ocorrências, representando 3,97% do total.

Como sentimento de grupo, surgiram as palavras “povo”, com 76 ocorrências, “todos”, com 57 ocorrências, “vamos”, com 42 ocorrências, “gente”, com 36 ocorrências, representando um total de 7,76% do total.

Como distanciamento de grupo, surgiram as palavras “vocês”, com 47 ocorrências, “forabolsonaro”, 42 ocorrências, “impeachmentalexandredemoraes”, 41 ocorrências, “eles”, 37 ocorrências, “bozo”, 33 ocorrências, “haiaparabolsonaro”, 33 ocorrências,

“militaresgenocidas”, 38 ocorrências, “genocida”, 35 ocorrências, o que representa 11,25%.

Como generalização “tudo”, 53 ocorrências, “nada”, 44 ocorrências, “todo”, 42 ocorrências, isso é 5,11%.

Como pessoas públicas, grupos e organização, surgiram as palavras “bolsonaro”, 186, “governo”, 128, “brasil”, 117, “presidente”, 111, “stf”, 67, “ministro”, 52, “federal”, 47, “jairbolsonaro”, 46, “polícia”, 46, “gilmar”, 41, “serra”, 40, “alcolumbre”, 35, representando uma grande quantidade de 33,68%.

Como outros assuntos, surgiram as palavras “covid-19”, 63 ocorrências, “pandemia”, 52 ocorrências, “educação”, 48 ocorrências, “cloroquina”, 44 ocorrências, “saúde”, 42 ocorrências, o que representa 9,15%.

Como palavras auxiliares, surgiram as palavras “país”, com 64 ocorrências, “dinheiro”, 45 ocorrências, “milhões”, 33 ocorrências, isso é 5,22%.

Agora as 40 mais frequentes da emoção nojo-desprezo na Figura 3.4.

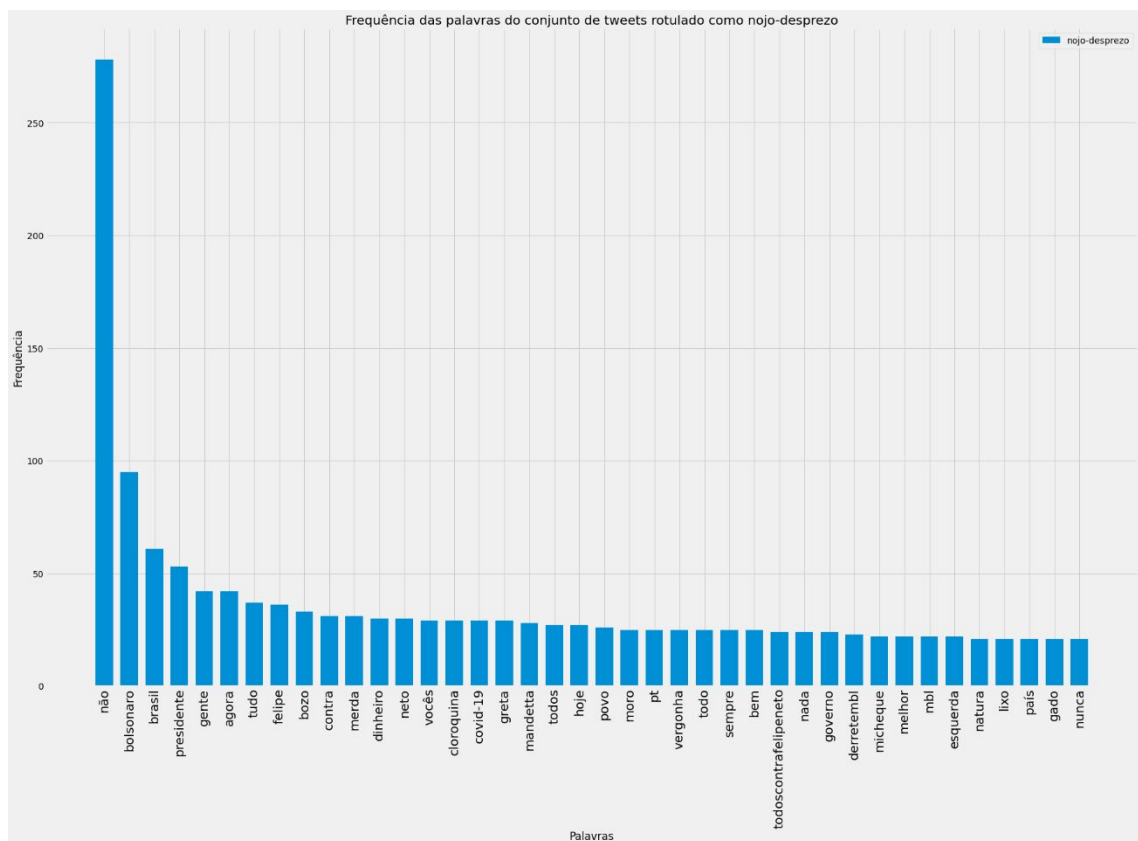


Figura 3.4: As 40 palavras mais frequentes da família das emoções de nojo-desprezo

Com conotação negativa, surgiram as palavras “não”, com 278 ocorrências, “contra”, 31 ocorrências, “merda”, 31 ocorrências, “vergonha”, 25 ocorrências, representado quase um quarto da quantidade de *tokens*, pelo peso do “não”, equivalente a uma grande parte de 24,61%.

Como conotação positiva, surgiram as palavras “bem”, com 25 ocorrências, “melhor”, 22, isso é apenas 3,17%.

Como urgência, surgiram as palavras “agora”, com 42 ocorrências, “hoje”, 27, representando 4,65%.

Como sentimento grupo, surgiram as palavras “gente”, com 42 ocorrências, “todos”, 27, “povo”, 26, “esquerda”, 22, isto é 7,89%.

Como distanciamento de grupo, surgiram as palavras “bozo”, com 33 ocorrências, “vocês”, 29, “todoscontrafelipeneto”, 24, “derretembl”, 23, “esquerda”, 22, “lixo”, 21, “gado”, 21, “micheque”, 22, o que representa 13,15%.

Como generalização, surgiram as palavras “tudo”, com 37 ocorrências, “todo”, 25, “sempre”, 25, “nada”, 24, “nunca”, 21, isso é 8,90%.

Como pessoas públicas, grupos e organização, surgiram as palavras “bolsonaro”, com 95 ocorrências, “brasil”, 61, “presidente”, 53, “felipe”, 36, “neto”, 30, “moro”, 25, “pt”, 25, “governo”, 24, “mbl”, 22, “natura”, 21, “greta”, 29, “mandetta”, 28, é equivalente a maior parte 30,28%.

Como outros assuntos, surgiram as palavras “cloroquina”, com 29 ocorrências, “covid-19”, 29, isso é 3,91%.

Como palavras auxiliares, surgiram as palavras “dinheiro”, com 30 ocorrências, “país”, 21, que representam 3,44%

Segue a Figura 3.5 com as 40 palavras mais frequentes de alegria.

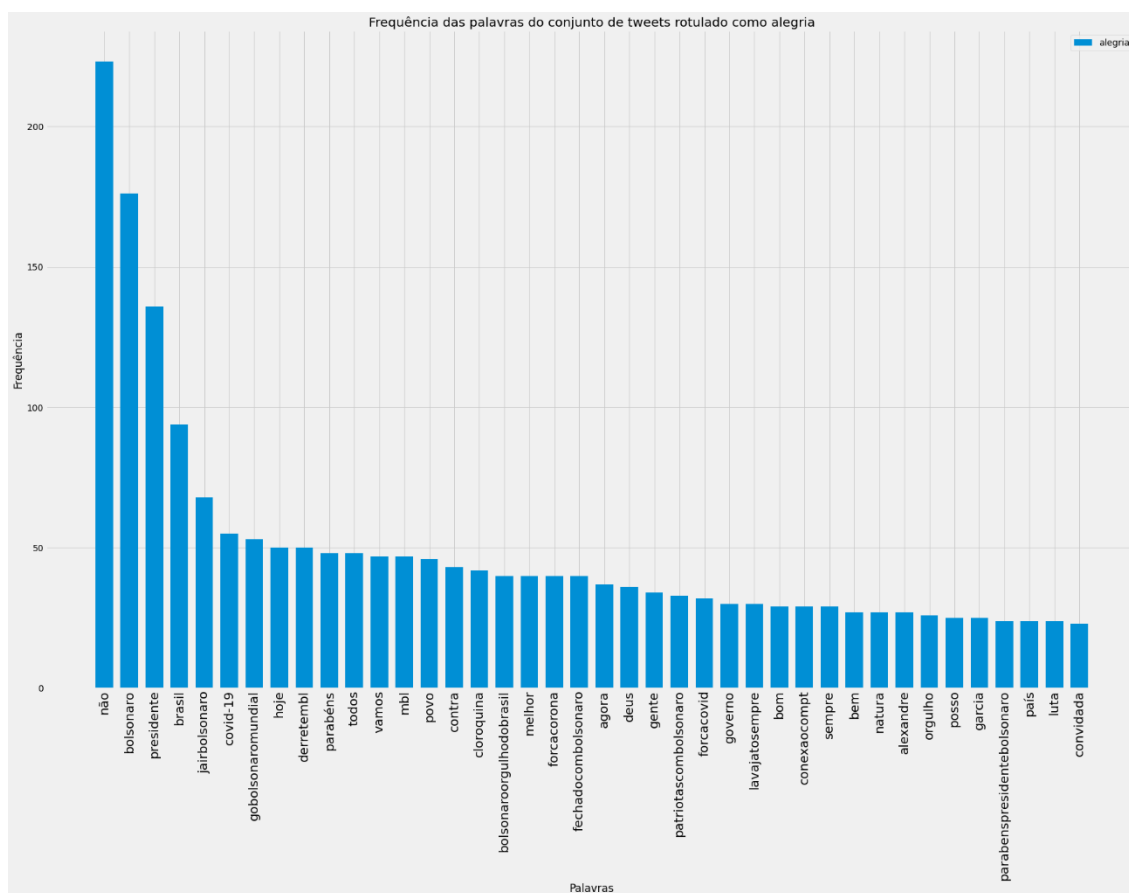


Figura 3.5: As 40 palavras mais frequentes da família das emoções da alegria

Com conotação negativa, “não”, 223, “contra”, com 43 ocorrências, o que representa 13,59%.

Com conotação positiva, surgiram as palavras “parabéns”, com 48 ocorrências, “melhor”, 40, “bom”, 29, “bem”, 27, “orgulho”, 26, isso é 8,69%.

Como urgência, surgiram as palavras “hoje”, 50, “agora”, 37, representando 4,45%.

Como sentimento de grupo, surgiram as palavras “gobolsonaromundial”, com 53 ocorrências, “todos”, 48, “vamos”, 47, “povo”, 46, “bolsonaroorgulhodobrasil”, 40, “fechadocombolsonaro”, 40, “gente”, 34, “lavajatosempre”, 30, “conexaocompt”, 29, “patriotascombolsonaro”, 33, “parabenspresidentebolsonaro”, 24, representando uma grande porção de 21,67%.

Como distanciamento de grupo, surgiram as palavras “derretemb”, com 50 ocorrências, “foracorona”, 40, “foracovid”, 32, isso é 6,23%.

Como generalização, surgiu a palavra “sempre”, com 29 ocorrências. Essa palavra representa apenas 1,48%.

Como pessoas públicas, grupos e organização, surgiram as palavras “bolsonaro”, com 176 ocorrências, “presidente”, 136, “brasil”, 94, “jairbolsonaro”, 68, “mbl”, 47, “governo”, 30, “natura”, 27, “alexandre”, 27, “garcia”, 25, aqui temos a maior porção, equivalente a 32,19%.

Como outros assuntos, surgiram as palavras “covid-19”, com 55 ocorrências, “cloroquina”, 42, “posso”, 25, “convidada”, 23, representando 7,41%.

Como palavras auxiliares, surgiram as palavras “deus”, com 36 ocorrências, “país”, 24, “luta”, 24, representando 4,29%.

Abaixo na Figura 3.6 as 40 palavras mais frequentes para neutro.

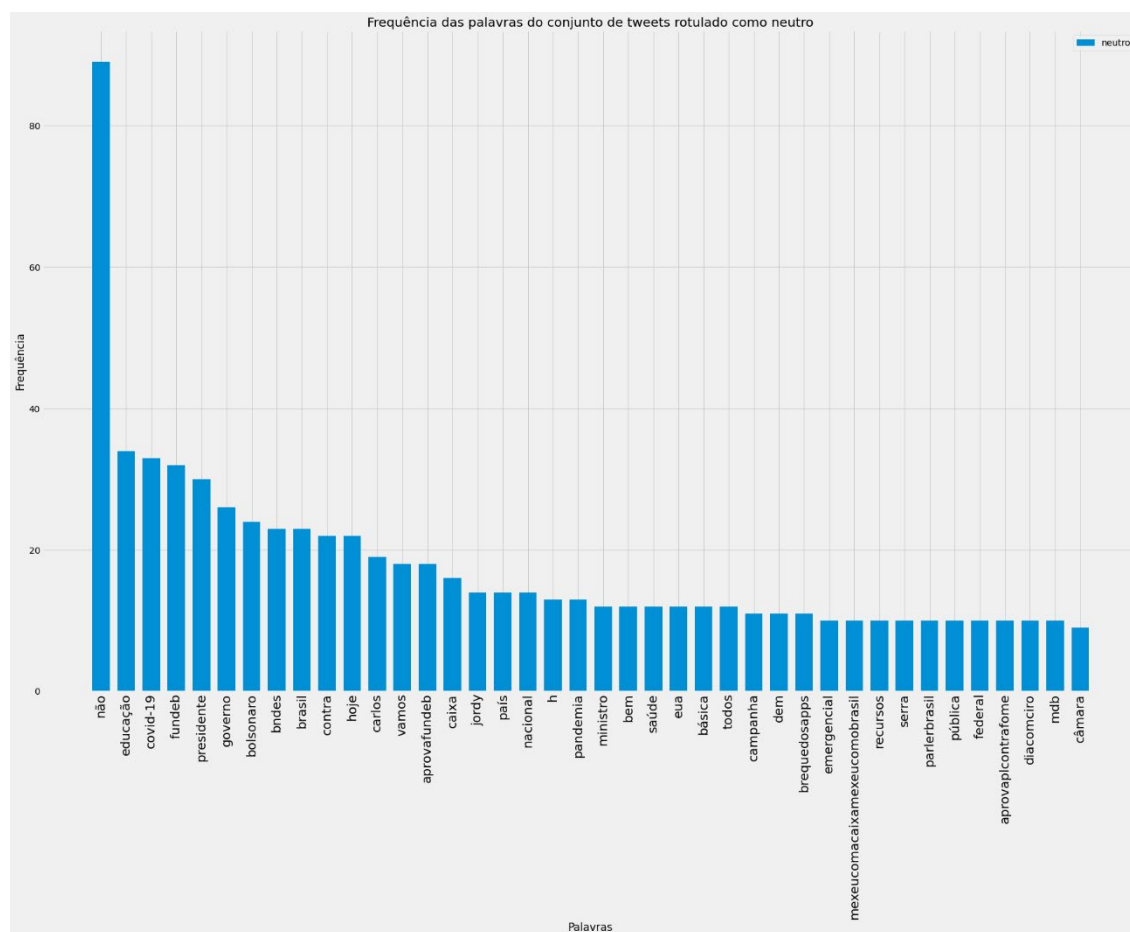


Figura 3.6: As 40 palavras mais frequentes dos tweets tidos como neutros

Com conotação negativa, surgiram as palavras “não”, com 89 ocorrências, “contra”, 22, representando 15,61%.

Com conotação positiva, surgiu a palavra “bem”, com 12 ocorrências, representando a menor porção com apenas 1,69%.

Como urgência, surgiu a palavra “hoje”, com 22 ocorrências, isto é, 3,09%.

Como sentimento grupo, surgiram as palavras “vamos”, com 18 ocorrências, “todos”, 12, “mexeucomacaixamexeucomobrasil”, 10, “diacomciro”, 10, representando 7,03%.

Para distanciamento de grupo e generalização não houve ocorrências.

Como pessoas públicas, grupos e organização, surgiram as palavras “presidente”, com 30 ocorrências, “governo”, 26, “bolsonaro”, 24, “brasil”, 23, “ministro”, 12, “serra”, 10, “federal”, 10, “jordy”, 14, “carlos”, 19, “eua”, 12, “dem”, 11, “caixa”, 16, “mdb”, 10, “bndes”, 23, “parlerbrasil”, 10, aqui tivemos a maior porção, de 35,16%.

Como outros assuntos, surgiram as palavras “educação”, com 34 ocorrências, “covid-19”, 33, “fundeb”, 32, “pandemia”, 13, “saúde”, 12, “emergencial”, 10, “aprovaplcontrafome”, 10, “brequedosapps”, 11, “aprovafundeb”, 18, que representam uma grande porção de 24,33%.

Como palavras auxiliares, surgiram as palavras “país”, com 14 ocorrências, “câmara”, 9 “nacional”, 14, “h”, 13, “pública”, 10, “recursos”, 10, “básica”, 12, “campanha”, 11, que representam 13,08%. Especificamente o “h” muito presente em neutro, geralmente aparece referindo-se as horas de um evento, como por exemplo um evento online “19h30”.

Na Tabela 3.4 é possível ver um resumo das percentagens de cada significado de acordo com as emoções abordadas.

Tabela 3.4: Percentagem por significados de cada grupo de emoções.

Significados	Raiva	Alegria	Nojo desprezo	Neutro
Negativo	23,86% ▲●	13,59% ▲●	24,61% ▲●	15,61% ▲●
Positiva	0%	8,69%	3,17%	1,69%
Urgência	3,97%	4,45%	4,65%	3,09%
Sentimento de grupo	7,76%	21,67% ▲●	7,89%	7,03%
Distanciamento de outro grupo	11,25% ▲●	6,23%	13,15% ▲●	0%
Generalistas	5,11%	1,48%	8,9%	0%
Pessoas, grupos ou organizações	33,68% ▲	32,19% ▲	30,28% ▲	35,16% ▲
Outros assuntos	9,15%	7,41%	3,91%	24,33% ▲
Palavras auxiliares	5,22%	4,29%	3,44%	13,08% ●
Soma 3 maiores (▲)	68,79%	67,45%	68,04%	75,11%
Soma 2 características (●)	35,11%	35,26%	37,76%	28,69%

No campo “Soma 3 maiores” é representada pela soma dos 3 grupos de palavras que mais se destacam dentro de cada emoção e são sinalizadas pelo triangulo (▲). De forma semelhante “Soma 2 características” é a soma de grupos de palavras específicos que de alguma forma faz com que determinada emoção se destaque. Essas são representadas pelo círculo (●). A análise completa dessa tabela será feita no capítulo 5.

3.10 Limpeza dos dados

Foram testadas algumas técnicas de limpeza dos textos, a fim de decidir qual delas era a melhor. Foram feitos testes com aprendizado de máquina e comparada a precisão (*accuracy*) antes e depois de aplicar a técnica. Esses resultados completos, a validação e os métodos utilizados são explicados com maior detalhe no capítulo 4, Modelação e validação.

3.10.1 Limpeza geral

A primeira limpeza nos dados aconteceu na arquitetura criada para extração dos *tweets*, onde foi identificado o *link* criado pela rede social referente ao *link* compartilhado pelo utilizador, que é algo semelhante a isso “<https://t.co/IxVfIL5xbD>”. Esse *link* é separado e guardado no base de dados junto com outras informações do *tweet*. Caso o utilizador não tenha compartilhado nenhum *link*, o código identifica isso e atribui o valor “sem url” na coluna onde haveria um *link*. Logo, sobra o texto original do *tweet* com caracteres de quebra de linha como “\n” e emojis. As quebras de linhas são removidas no momento de eliminação dos *tweets* repetidos com o método padrão do Python `.replace('\n', ' ')`, a eliminação dos *tweets* repetidos foi explicada na seção 3.5, Remoção de repetidos. Os *emojis* não são removidos, pois possuem parte do significado expressado pelos utilizadores. Outro tratamento necessário é remover qualquer tipo de aspas que os *tweets* possuem, pois dão erro ao inseri-los na base de dados.

3.10.2 Emojis

Foi feito um teste em que a amostra tinha todos os *emojis* e outro teste sem eles, com todos os parâmetros iguais, exceto pelos emojis. Para esse teste foi utilizado um classificador linear do pacote Scikit-learn `SGDClassifier`. O teste em que a amostra estava com os emojis obteve uma melhora na precisão média de 1,5%. Aqui verificamos que os emojis são necessários para o aprendizado de máquina nos dados deste trabalho.

3.10.3 Pontuação

Outro teste foi realizado removendo toda a pontuação, onde sobraram apenas as palavras em sequência e entre espaços. Foi utilizado o classificador do teste anterior: o `SGDClassifier`. O teste em que a amostra estava com a pontuação obteve uma melhoria na precisão média de 2,2%. Portanto, a pontuação cooperou para o aprendizado de máquina.

3.10.4 Stopwords

A amostra foi testada com e sem *stopwords* também. Outro classificador linear que teve bons resultados para os dados, foi o LinearSVC, também ele do pacote Scikit-learn. Os dados com *stopwords* obtiveram uma melhora na precisão média de 1,3%. Aqui verifica-se a vantagem em utilizar *stopwords* para os dados desse trabalho.

3.11 Formatação dos dados

Para a modelação foram feitas algumas formatações nos dados, verificando-se que essas contribuíram para melhorar a precisão média dos modelos. As emoções utilizadas para previsão foram raiva, com 1335 ocorrências, nojo-desprezo com 764 ocorrências, alegria com 1060 ocorrências e neutro com 397 ocorrências. As que foram descartadas por terem poucas ocorrências foram tristeza com 85 ocorrências, Surpresa com nenhuma e Medo com 31 ocorrências.

3.11.1 Correção de abreviações

Foi criada uma estrutura em Python para substituir abreviações por palavras completas, a fim de padronizar as muitas palavras que foram escritas de diversas formas. Como por exemplo o “não”, que foi escrito com algumas variações: “ñ”, “n”, “nao” e “nã”. O objetivo aqui foi fazer com que o texto ficasse o mais homogêneo possível, para que na etapa da modelação, o algoritmo de aprendizado de máquina não precisasse de diferenciar entre “ñ”, “n”, “nao” e “nã”. sendo que, como é a mesma palavra, se tal procedimento não se fizesse, isso poderia dificultar a aprendizagem.

Para validar essa técnica foi feito um teste com a amostra com todo o pré-processamento e formatação final dos dados, mas sem corrigir essas abreviações, com o classificador LinearSVC() do pacote Scikit-learn em modo padrão. Depois foi feito outro teste com os exatos mesmos parâmetros e procedimentos. A única alteração dessa vez foi a aplicação da correção de abreviações, tendo o classificador um desempenho 0,75% melhor do que da primeira vez. Foram identificadas manualmente 57 abreviações para serem corrigidas, elas são mostradas na Tabela 3.5 abaixo.

Tabela 3.5: Abreviações encontradas no corpus.

Abreviação	Palavra	Abreviação	Palavra
"n", "ñ", "nao"	"não"	"tá", "ta"	"está"
"pod"	"pode"	"fdp"	"filho da puta"
"q"	"que"	"sdv"	"segue de novo"
"d"	"de"	"agr"	"agora"
"pq"	"porque"	"pqp"	"puta que pariu"
"vc", "cê", "voce"	você	"nvo"	"novo"
"voces", "ces", "vcs"	"vocês"	"oto"	"outro"
"tt"	"twitter"	"bbs"	"bebês"
"rt"	"retweeta"	"ces", "vcs"	"vocês"
"hj"	"hoje"	"qm"	"quem"
"tb", "tbn", "tmb"	"também"	"vm"	"vem"
"td"	"tudo"	"c/"	"com"
"mt"	"muito"	"eh"	"é"
"kd"	"cadê"	"covid19", "covid_19", "covid"	"covid-19"
"s"	"sim"	"coronavirus", "coronávirus"	"covid-19"
"dps"	"depois"	"vsf"	"vai se foder"
"c", "cm"	"com"	"kralho"	"caralho"
"p/", "p", "pra"	"para"	"eua"	"estados unidos"
"pr"	"presidente"	"qrer"	"querer"
"pf", "policia federal"	"Polícia Federal"	"pfvr"	"por favor"

Outro recurso utilizado para padronizar, aplicado junto com a correção de abreviações, foi a função lower(), que deixa todas as letras em minúsculo. Veja em seguida a comparação de um exemplo de um *tweet* antes e depois do método de correção (as palavras sublinhadas são as que foram substituídas):

“Ai agr fui cancelado por n qrer o bem do bozo pfvr ne só me deixa zuar a condição dele, sendo ou n jogada política pra vender cloroquina ou fortalecer o bozonarismo”

“ai agora fui cancelado por não querer o bem do bozo por favor ne só me deixa zuar a condição dele, sendo ou não jogada política para vender cloroquina ou fortalecer o bozonarismo”.

3.11.2 Matriz esparsa de contagem de *tokens* Scikit-learn

O pacote Scikit-learn oferece uma forma de preparar os dados em forma de texto para o aprendizado de máquina, que consiste em transformá-los numa matriz esparsa, feita a partir da contagem de *tokens*. A maioria dos classificadores necessita de vetores numéricos com o mesmo tamanho, o que não acontece em um conjunto de textos, que são palavras no qual as frases possuem comprimento variável umas das outras.

Fazendo uso do pacote `scipy.sparse`, o Scikit-learn cria uma matriz que utiliza pouco espaço na memória além de ser otimizada para acelerar as operações algébricas feitas pela maioria dos classificadores. Então é realizada uma contagem de todos os *tokens* únicos do corpus. Em seguida, é atribuído um valor inteiro às palavras. Os dados foram convertidos de categóricos para numéricos para possibilitar os cálculos pelos classificadores. Nesse caso, foi utilizada uma função específica, `TfidfVectorizer()`, que para além de criar a matriz de contagem de *tokens*, faz algo semelhante a um escalonamento nos dados, deixando os valores em decimais e muito próximos uns dos outros, centralizando-os em torno de uma média, pois se apenas fossem inteiros distantes, eles teriam pesos diferentes sem nenhum critério, influenciando no aprendizado de máquina.

Além disso, foi possível aplicar a técnica de TF-IDF do inglês *Term Frequency Inverse Document Frequency*, que busca n-gramas dentro do corpus. Foram feitos muitos testes com e sem TF-IDF, mas os resultados não foram melhores com essa técnica. Para o `SGDClassifier()` em modo padrão, a melhor configuração foi com bi-gramas obtendo um aumento de precisão média de 1,57% em relação as dados sem n-gramas e com as mesmas configurações para todo resto. No entanto n-gramas fez cair a precisão depois de aplicar a última redução de dimensionalidade, ou seja, os classificadores selecionados preformaram melhor com os dados sem n-gramas e com a última redução de dimensionalidade selecionada e parametrizada, isso será abordado na secção seguinte. Além disso, a aplicação dos n-gramas faz com que aumente drasticamente o tamanho da matriz na quantidade de colunas, *features*. informações detalhadas da parametrização dos algoritmos podem ser encontradas no capítulo 4, Modelação e validação.

3.11.3 Reduções de dimensionalidade rejeitadas

Várias reduções de dimensionalidade foram testadas, como a *Principal Component Analysis* (PCA), não tendo essa redução se mostrado eficiente, selecionando as primeiras

features em ordem, deixando de lado as últimas. Poderia deduzir que o algoritmo não conseguiu analisar corretamente os dados para este caso, além de causar sobrecarga de memória quando utilizado uma quantidade alta *features*, portanto não foi possível prosseguir com essa redução.

Foi testada a redução *Sparse Principal Components Analysis* (SparsePCA) e teve o mesmo problema de sobrecarga de memória, por isso foi descartada. Outra que causou problema de sobrecarga na memória foi o *Linear Discriminant Analysis* (LDA), por isso também foi descartado. Além desses, também foram testados e descartados por sobrecarga de memória ou por baixar a performance dos classificadores, o *Randomized Singular Value Decomposition* (RandomizedSVD), *Truncated Singular Value Decomposition* (TruncatedSVD) e *Kernel Principal Components Analysis* (KernelPCA).

3.11.4 RFE redução de dimensionalidade escolhida

A redução de dimensionalidade que melhorou a precisão média dos modelos foi a *Recursive Feature Elimination* (RFE). Essa técnica reconfigura o peso das *features* de acordo com um estimador externo, junto com a classe com cada rótulo dos *tweets*. Nesse caso foi utilizado como estimador, os classificadores de maior performance para os dados, *LinearSVC* e o *SGDClassifier()*, um de cada vez e com configurações diferentes. A desvantagem dessa redução é que ela é demorada, cerca de dez a quinze minutos para cada teste, varia de acordo com as configurações dos parâmetros. Portanto, aqui foi um dos pontos exaustivos, pois para atingir a configuração ideal foi feito mais do que uma centena de testes para cada classificador. O resultado foi um aumento de 13,55% da precisão média do *LinearSVC*, quando os dados são comparados sem a aplicação do RFE e 17,15% para o *SGDClassifier()*. Abaixo a Tabela 3.6 mostra esse resumo.

Tabela 3.6: Aumento da precisão média depois da redução de dimensionalidade RFE.

Classificador e técnicas	Aumento na precisão média (%)
<i>LinearSVC()</i> : <i>TfidfVectorizer()</i> + RFE()	+ 13,55%
<i>SGDClassifier()</i> : <i>TfidfVectorizer()</i> + RFE()	+ 17,15%

3.11.5 Formato final

A sequência de formatação dos dados ficou da seguinte maneira, primeiro houve uma correção das abreviações para diminuir as palavras diferentes com mesmos significados. Depois as *features* foram extraídas a partir de *TfidfVectorizer*, o texto ficou em minúsculo

e sem a utilização do n-gramas. Por fim, as *features* foram treinadas em dois processos separados com o RFE, utilizando os estimadores selecionados, LinearSVC e SGDClassifier.

A formatação final dos previsores, antes de ser feito o aprendizado de máquina no classificador LinearSVC, ficaram na seguinte disposição: 3556 linhas, que são cada um dos *tweets* das 3 emoções mais neutro, por 1540 colunas, que eram as palavras de cada *tweet* e depois foram transformadas em dados numéricos. Para o SGDClassifier a disposição foi 3556 linhas por 2180 colunas. Abaixo, na Tabela 3.7, há um exemplo de como ficaram parte desses dados.

Tabela 3.7: Extrato do dataset utilizado para o aprendizado de máquina.

Início do dataset		Fim do dataset	
(0, 1995)	0.3842806104063306	(3553, 1475)	0.15029107093658758
(0, 920)	0.46975768771847676	(3553, 1676)	0.10948701169142216
(0, 2074)	0.5099686516196323	(3553, 558)	0.11595537264509419
(0, 1857)	0.6095802667905041	(3554, 531)	0.195590991628613
(1, 1644)	0.09231737943017701	(3554, 50)	0.16643965818918252
(1, 1481)	0.1854648025711329	(3554, 1295)	0.13680885130216797
(1, 1563)	0.1542582719746431	(3554, 1446)	0.16643965818918252
(1, 1979)	0.11504997832520249	(3554, 2086)	0.14965675037722675
(1, 2083)	0.2094057404438046	(3554, 483)	0.09713828198632055
(1, 545)	0.14573436110830806	(3554, 877)	0.12948984006149206
(1, 558)	0.15782564893313358	(3554, 814)	0.132873842565271
(1, 1432)	0.09239005298104822	(3554, 827)	0.09010366562838558
(1, 1695)	0.1854648025711329	(3554, 1340)	0.12265625317038106
(1, 892)	0.18370908958722132	(3554, 734)	0.06941163108934313
(1, 916)	0.3709296051422658	(3554, 1475)	0.11401602539427565
(1, 1067)	0.1536485569318646	(3554, 1676)	0.08306064909617698
(2, 1066)	0.27116396850183266	(3555, 475)	0.24960185722741846
(2, 1889)	0.1640559567485431	(3555, 502)	0.202004052782528

Entende-se a matriz final mostrada na Tabela 3.7 da seguinte forma, cada linha representa uma palavra, o primeiro número de uma linha da esquerda para a direita, é sobre o *tweet*, no sentido de que todos os zeros são as palavras do primeiro *tweet*, sendo o segundo *tweet* representado pelo número 1 e assim por diante. Logo, vai de zero até 3555, somando o total de ocorrências dos *tweets* etiquetados de 3556. Ao lado, o inteiro é o token da

respectiva palavra e por fim o peso que foi atribuído pela redução de dimensionalidade RFE.

A classe ou rótulos de cada *tweet*, é apenas uma coluna que possui o mesmo tamanho das linhas, 3556, essa está na seguinte disposição:

['alegria', 'alegria', ..., 'raiva', 'raiva', ..., 'nojo-desprezo', 'nojo-
desprezo', ..., 'neutro', 'neutro']

Na classe foi aplicado o método `LabelEncoder()`, para transformá-la de nominal para numérica e assim obteve o resultado final na seguinte disposição:

[0, 0, ..., 3, 3, ..., 2, 2, ..., 1, 1]

Logo 0: 'alegria', 3: 'raiva', 2: 'nojo-desprezo', 1: 'neutro', neste caso, os inteiros não influenciam no aprendizado de máquina.

4 MODELAÇÃO E VALIDAÇÃO

Foram testados exaustivamente muitos classificadores, probabilísticos, SVMs, Redes Neurais e etc. Principalmente do pacote Scikit-learn e um probabilístico do NLTK, *Naive Bayes*. A prioridade foi para aos algoritmos que melhor desempenharam levando em conta três critérios, a precisão média geral, o equilíbrio da precisão entre as classes e o tempo que o algoritmo leva para ser validado. A partir disso, foram selecionados dois deles como comentado anteriormente, o LinearSVC e o SGDClassifier. Portanto, destacaram-se os modelos baseados em *Linear Support Vector Machine*.

4.1 Classificadores rejeitados

Houve muitos classificadores testados que foram encontrados por meio de várias fontes, como através do ML Map mostrado na Figura 4.1, escolhidos baseando-se na documentação do Scikit-learn, por meio de artigos científicos e também em discussões em torno da comunidade de *Data Science* presente nas redes sociais.

Um classificador que obteve a precisão média significativa, mais ainda longe dos classificadores selecionados, foi uma rede neural multicamada MLPClassifier, mas demorou muito tempo de aprendizagem, então foi descartada. Assim como o *Naive Bayes* do NLTK, chamado por `nlk.NaiveBayesClassifier`, mas não teve um bom desempenho, na precisão, na preparação dos dados e validação. As duas últimas etapas foram extremamente demoradas.

Outros classificadores também foram escolhidos baseados no ML Map, como foi os classificadores baseados no Teorema de Bayes. Desses o que melhor preformou foi o ComplementNB, que é um classificador probabilístico projetado para conjuntos de dados desequilibrados, ideal para classificação de textos. Ele é uma adaptação MNB do inglês *Multinomial Naive Baye*. Ele não aceita dados negativos, uma vez que os cálculos são multiplicações, por isso RFE não pode ser aplicada. Ele também obteve uma boa precisão, mas ainda ficou mais de 10% a menos que os classificadores escolhidos.

Outros baseados no ML Map também foram rejeitados como KNeighbors Classifier e SVC por baixa precisão. Os classificadores baseados em Ensemble também foram testados, especificamente o RandomForestClassifier, GradientBoostingClassifier HistGradientBoostingClassifier mas não obtiveram resultados significativos.

Os classificadores rejeitados também foram testados em modo *default* e tiveram alguns dos hiperparâmetros configurados, mas não é possível afirmar que eles não devam ser usados para esse *dataset*. Afirmo aqui que, apenas que seria necessário mais tempo para verificar profundamente se eles seriam uma melhor escolha. Pois devido a grande quantidade de possibilidades, mais o acréscimo de tempo ao aplicar a redução de dimensionalidade para cada teste e a falta de poder computacional para realizar um teste de GridSearch, então foram rejeitados e optou-se pelos mais rápidos e mais precisos no primeiro momento.

4.2 LinearSVC

O *Linear Support Vector Classification* (LinearSVC) é um classificador linear capaz de realizar uma classificação de duas classes ou mais. Ele foi escolhido a partir do Machine Learning Map, fornecido pela documentação do Scikit-learn, esse é um fluxograma que auxilia na etapa de escolha de um classificador. O fluxograma é mostrado abaixo na Figura 4.1.

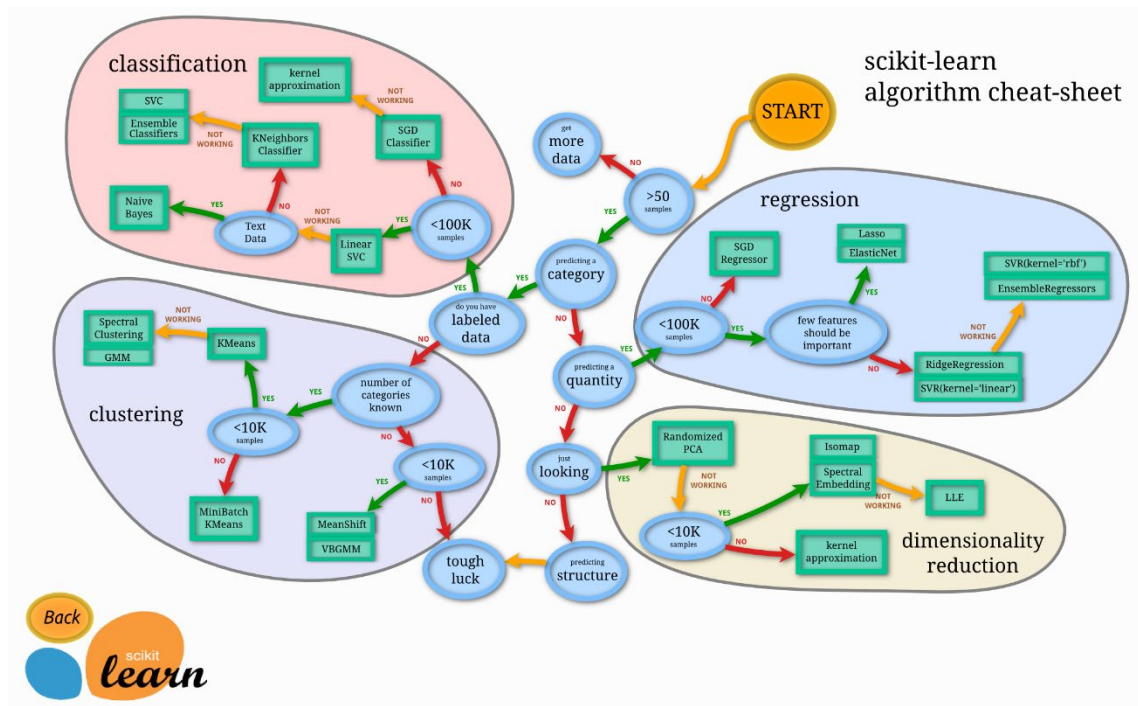


Figura 4.1: Machine Learning Map

Fonte: Pedregosa et al. (2011)

O LinearSVC é um classificador da família dos *Support Vector Machines* (SVMs), por isso classifica de forma semelhante a um SVM tradicional. Ele é uma implementação mais rápida do *Support Vector Classification* (SVC) com parâmetro `kernel='linear'` e

Agora, observemos os detalhes da precisão que o classificador LinearSVC alcançou na Tabela 4.1, a explicação de como foi gerado essas estatísticas é feita a seguir na seção 4.4, Validação.

Tabela 4.1: Detalhes da precisão do classificador LinearSVC.

Classificador	Est. Descritiva da precisão		Matriz de confusão	Precisão por classe
LinearSVC	Count	12.000000	[[818 35 60 147]	Alegria: 0.7717 Total: 1060
	Mean	0.731440	[78 176 27 116]	Neutro: 0.4433 Total: 397
	Std	0.028415	[107 6 467 184]	Noj-desp: 0.6113 Total: 764
	Min	0.672297	[102 23 70 1140]]	Raiva: 0.8539 Total: 1335
	25%	0.716933		
	50%	0.731874		
	75%	0.746826		
	Max	0.777027		

4.3 SGDClassifier

O SGDClassifier foi escolhido baseando-se no ML Map, mostrado na Figura 4.1, apesar que ser recomendado para amostras acima de 100 mil, ele entrou na lista dos vários classificadores testados. SGDClassifier é um método que possui vários classificadores, todos baseados na implementação de modelos lineares regularizados com aprendizagem de *Stochastic Gradient Descent* (SGD). De acordo com os hiperparâmetros configurados ele assume um tipo perda, como SVM, regressão logística, perceptron, etc. (Pedregosa et al., 2011).

O SGDClassifier, também faz parte do pacote científico do Python Scikit-learn, e é chamado pelo método `sklearn.linear_model.SGDClassifier()`. Em modo *default* ele possui os seguintes parâmetros: `SGDClassifier(loss='hinge', penalty='l2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=1000, tol=0.001, shuffle=True, verbose=0, epsilon=0.1, n_jobs=None, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5, early_stopping=False, validation_fraction=0.1, n_iter_no_change=5, class_weight=None, warm_start=False, average=False)`. O parâmetro *default* de perda, `loss='hinge'`, fornece um SVM Linear (Pedregosa et al., 2011).

Assim como foi para modelar o dataset para o LinearSVC, com o SGDClassifier também foi necessário correr muitas possibilidades para alcançar o resultado. É possível ver a árvore de possibilidades da modelação dos dados abaixo, na Figura 4.3.

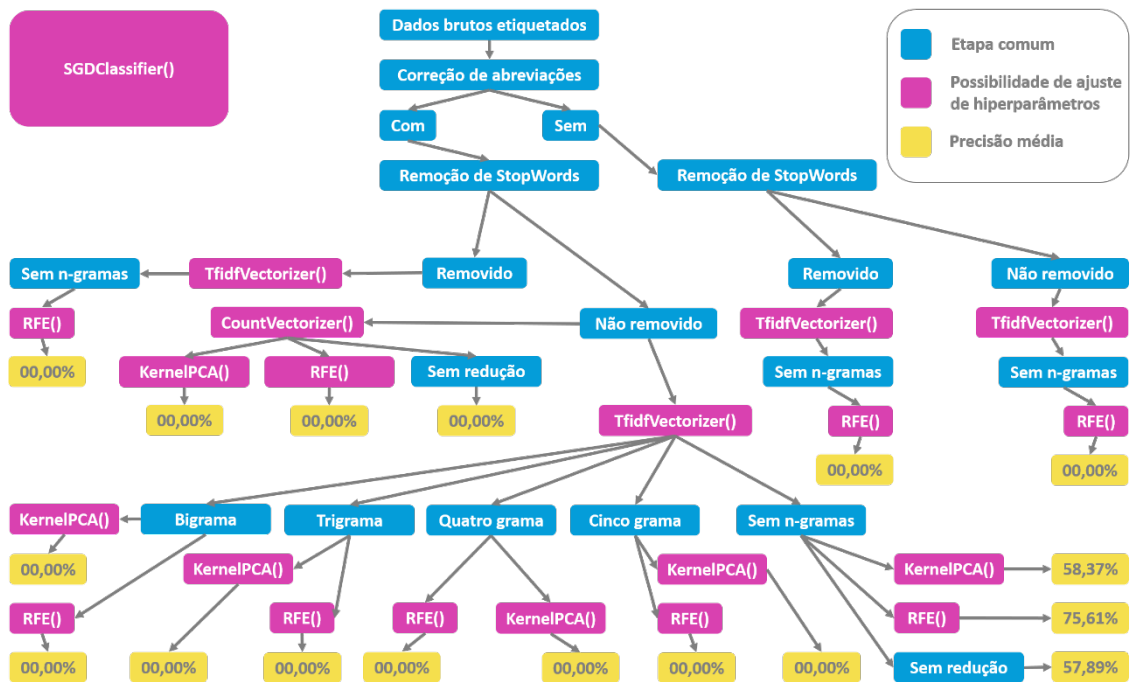


Figura 4.3: Árvore de possibilidades para a modelação do SGDClassifier.

A seguir temos a Tabela 4.2, que mostra o resultado da precisão em detalhes, do classificador SGDClassifier. Assim como a tabela de precisão da seção anterior, essa, será explicada a seguir na seção 4.4 Validação.

Tabela 4.2: Detalhes da precisão do classificador SGDClassifier.

Classificador	Est. Descritiva da precisão	Matriz de confusão	Precisão por classe
SGDClassifier	Count	12.000000	[[834 32 62 132]
	Mean	0.756187	[73 204 22 98]
	Std	0.022506	[89 9 499 167]
	Min	0.733108	[89 22 72 1152]]
	25%	0.737152	
	50%	0.752110	
	75%	0.766892	
	Max	0.801347	
			Alegria: 0.7868 total: 1060 Neutro: 0.5139 total: 397 Noj-desp: 0.6531 total: 764 Raiva: 0.8629 total: 1335

4.4 Validação

A forma de validação utilizada foi a *Stratified Cross Validator*, ou seja, Validação Cruzada Estratificada, a qual faz parte do pacote científico do Python Scikit-learn, e é representada pela função `StratifiedKFold()`, ela é chamada da seguinte forma `sklearn.model_selection.StratifiedKFold`.

Essa forma de validação divide a amostra em k partes, em que a parte maior corresponde ao conjunto de treino e a menor ao conjunto de teste. Então é realizado uma rodada de treino e teste, na próxima rodada a outra dobra será o novo conjunto de teste e o restante ficará para o treino e assim sucessivamente até rodar a quantidade de divisões

especificada. Cada parte preserva o percentual de cada classe presente na amostra, isso é fundamental para que a variação da precisão do teste de cada dobra seja a mínima possível e os resultados sejam mais consistentes. Na Figura 4.4, a baixo, é possível ver a representação da validação com o exemplo em cinco divisões (Pedregosa et al., 2011).



Figura 4.4: Representação da validação cruzada estratificada StratifiedKFold.

A quantidade de divisões é a quantidade de dobras, se for especificado 5 divisões, o algoritmo fará 5 dobras para o conjunto de teste, o que resulta em aproximadamente 20% do *dataset*, se forem selecionadas 10 divisões, 10%. No entanto, o StratifiedKFold faz apenas a divisão do conjunto, para medir a precisão após os testes, tendo sido utilizadas duas funções, `accuracy_score()`, `confusion_matrix()`, chamadas do mesmo módulo do pacote científico do Python Scikit-learn, `sklearn.metrics`. O *Accuracy Score* faz o cálculo da precisão, representado em fração, retornando a porcentagem de acertos de acordo com a classe, como visto nas tabelas de precisão Tabela 4.1 e Tabela 4.2 (Pedregosa et al., 2011).

Ao realizar os 12 testes feitos para cada dobra, os resultados são coletados a partir por `accuracy_score()`, colocados numa lista, assim são utilizados para criar um DataFrame, do pacote Pandas a fim de aplicar a função `describe()` nestas precisões, isso retorna a estatística descritiva delas como visto anteriormente.

Na função `confusion_matrix()`, é gerado a matriz de confusão com os resultados dos dados testados. Assim de modo semelhante, temos 12 matrizes feitas com cada dobra. Elas são colocadas numa lista e somadas com o método `np.sum()` do pacote NumPy (Oliphant, 2006). As matrizes são corretamente somadas pela função de acordo com a posição de cada item, o objetivo aqui é ter uma matriz com todos os dados do *dataset*.

A estatística “Precisão por classe” é gerada por um cálculo simples feito com a matriz de confusão final, com todo o *dataset*, onde foi dividido os acertos pelo total de cada classe. Assim obtém-se a percentagem de acerto por classe. Essas percentagens servem apenas para compreender os acertos de cada classe, diferente da estatística de precisão geral que foi feita baseado nos acertos gerais do modelo.

4.5 Hiperparâmetros e GridSearch

Depois de afinar o pré-processamento de acordo com cada classificador, rodando as possibilidades pertinentes, chegou a hora de afinar os classificadores e a técnica utilizada foi o GridSearch do Scikit-learn, chamado pelo método `sklearn.model_selection.GridSearchCV`, que é uma pesquisa exaustiva sobre valores de parâmetros especificados para um estimador, por isso é muito demorado, podendo demorar dias, isso requer um grande poder computacional. Inicialmente essa técnica foi aplicada utilizando um Notebook no Google Colab, mas a sessão expirava, perdendo o trabalho feito. Outra tentativa foi feita a partir do computador pessoal, mas a falta de poder computacional e o tempo limite para finalizar essa etapa, fizeram com que essa aplicação fosse cancelada. Por isso os classificadores seguiram sem que houvesse feito o ajuste necessário para obter um melhor resultado, ou seja, em modo *default*.

Os principais hiperparâmetros configurados para o pré-processamento, para ambos classificadores e para a validação são mostrados na Tabela 4.3.

Tabela 4.3: Comparações dos hiperparâmetros dos classificadores.

Pré-processamento	Classificador	Validação	Precisão
1ª Com correção de abreviações; 2ª Não removido as StopWords; 3ª TfidfVectorizer(lowercase=True, stop_words=None) 4ª RFE(n_features_to_select=1540)	LinearSVC(random_state=0)	StratifiedKFold(n_splits=12, shuffle=True, random_state=19)	count 12.000000 mean 0.731440 std 0.028415 min 0.672297 25% 0.716933 50% 0.731874 75% 0.746826 max 0.777027
1ª Com correção de abreviações; 2ª Não removido as StopWords; 3ª TfidfVectorizer(lowercase=True, stop_words=None) 4ª RFE(n_features_to_select=2180)	SGDClassifier(random_state=0)	StratifiedKFold(n_splits=12, shuffle=True, random_state=13)	count 12.000000 mean 0.756187 std 0.022506 min 0.733108 25% 0.737152 50% 0.752110 75% 0.766892 max 0.801347

Logo, a configuração do LinearSVC e do SGDClassifier não foram alteradas. Apenas foi atribuído o `random_state` como zero, para haver a possibilidade da experiência se replicar.

Já para a validação cruzada, utilizaram-se os seguintes hiperparâmetros: StratifiedKFold(n_splits=12, shuffle=True) e random state teve configurações diferentes em cada classificador.

Para este trabalho foram utilizadas 12 divisões, “n_splits=12”, pois o *dataset* tem um tamanho considerável para isso, neste caso o conjunto de teste ficou com aproximadamente 8,33% ou 296 *tweets* de 3556. Foi ativado o modo baralhar as amostras de cada classe antes de dividir em lotes por “shuffle=True”, pois o *dataset* foi agrupado sem aleatoriedade. Ou seja, os *tweets* e as classes estavam na ordem com que foram sendo adicionadas na variável, primeiro alegria, depois raiva, nojo-desprezo e neutro no final. O parâmetro *Random State* afeta a ordem dos índices, ele controla a semente aleatória de cada dobra para cada classe. A princípio, foi definido como zero para todos os testes, a fim de obter a mesma aleatoriedade para testar todos algoritmos e perceber a real performance de cada algoritmo sem ser influenciado pela aleatoriedade da amostra. Outra vantagem, é que dessa forma os experimentos serão reproduzíveis.

5 DISCUSSÃO DOS RESULTADOS

Alguns pontos importantes, que já foram apresentados durante o trabalho, agora serão devidamente desenvolvidos.

5.1 Comportamentos emocionais expressos nas palavras

Nas etapas iniciais, durante o desenvolvimento da bibliografia, optou-se por procurar a presença das emoções em palavras, principalmente, através dos gatilhos inatos que cada uma das famílias de emoções possui e estão presentes na maioria dos seres humanos, como visto detalhadamente na seção 2, Revisão da literatura. Mas não se sabia ao certo se isso seria possível, devido a não se ter encontrado referências de outros estudos que tivessem a mesma abordagem a partir desses gatilhos. Porém, houve uma relação precisa entre muitos gatilhos emocionais identificados pelos cientistas, com os textos coletados na presente pesquisa, como discutido em 3.7, Classificação manual e análise prática.

A tristeza claramente relacionada ao tema, que é a perda. Embora houvesse poucas destas ocorrências, nessas, os utilizadores relatavam decepção com políticos, morte de pessoas por Covid-19 entre outras. Algumas palavras desses *tweets* ilustram isso, como “morre”, “pandemia”, “decepção”, “mortos”, “mortes”, “doença”, “dor”, “chorar” e etc.

Para raiva, a reação de querer causar dano físico e ou psicológico predominou. Os utilizadores manifestaram muitos níveis de raiva, dos mais baixos até ódio e também raiva junto com aversão. Destacaram-se alguns adjetivos pejorativos e muitos palavrões, como “genocida”, “corrupto”, “canalhas”, “bandidos”, “vagabundo”, “puta”, “vai tomar no cu”, “vai se fuder”, “porra”, “caralho” e “bicuda”, que é um chuto em uma pessoa.

Também, com pouquíssimas ocorrências os *tweets* que expressaram medo, foram bem claros. Geralmente relacionados com a suposta má gestão do governo Bolsonaro, algumas vezes até comparado a um pesadelo. Também relacionado ao Covid-19, sobre as mortes e a volta as aulas no sentido de que os alunos poderiam se infectar.

As palavras que ilustram isso foram “risco”, “covid”, “pesadelo”, “apocalipse”, “ameaças”, “pandemia”, “morte” e “assustador”.

Outros gatilhos se apresentaram de uma forma mais sofisticada na ligação entre o passado evolutivo das emoções e as palavras, como foi o caso do nojo-desprezo. Onde houve muitas palavras que remetem a aversão básica, que é a incorporação oral de algo que possa causar dano ou intoxicação, relacionado com outras pessoas que causaram nojo-

desprezo no utilizador que escreveu. Nestes casos, foram manifestados adjetivos como “lixo”, “merda”, “bosta”, “asquerosa”, “papel higiênico usado”, “contamina”, “sujo”, “verme”, “esgoto” e muitas outras. Algumas remetem a aversão interpessoal, relacionado a moralidade, como “hipocrisia”, “corruptos”, “miliciano”, “ladrão” e etc. Essas palavras poderiam ser usadas em contexto de raiva, no intuito de causar dano psicológico e muitas realmente o foram. Algumas vezes, quando a raiva é muito intensa, ela vem acompanhada de aversão. Por isso em alguns textos, o limite entre raiva e aversão se mostrou difuso.

Dentro da família da alegria, alguns gatilhos apresentaram-se de uma forma adaptada, porém sem perder a essência do gatilho, há o caso da emoção *naches*. Em vez da relação ser no sentido filho, aluno ou discípulo, sentir orgulho do professor, mentor ou mestre, foi caracterizada pela relação do utilizador sentir orgulho do político ou líder. *Fiero* se mostrou numa adaptação mais voltada para a conquista de desafio em grupo. E *schadenfreude* não teve adaptações e também foi uma emoção muito presente nos *tweets* classificados com a família da alegria. Algumas palavras se destacaram para ilustrar isso, foram “parabéns”, “melhor”, “bem”, “bom”, “orgulho”, “verdade”, “feliz” e etc.

No geral, as palavras destacadas aqui, servem para ilustrar algumas características das emoções e não serem usadas como meio definitivo de classificação, pois as frases dos *tweets* são complexas e por muitas vezes utilizam as mesmas palavras em vários contextos que remetem a emoções diferentes. Além disso, essas palavras cobrem apenas uma parte dos *tweets*.

Os *tweets* classificados como neutro, foram surgindo naturalmente na análise, sempre voltados para contextos do debate político, mas sem expressarem evidências emocionais e muitas vezes de forma argumentativa lógica. Eles também servem como ponto de partida para comparar com uma frase envolvida em emoções.

5.2 Ausência de algumas emoções

Na fase de etiquetagem manual dos dados, pode-se perceber que haviam quantidades desequilibradas das seis emoções, tristeza, raiva, surpresa, medo, nojo-desprezo e alegria. Porém, antes, acreditava-se que seriam equilibradas, devido à grande variedade de pessoas, culturas, formas diferentes de se expressarem e a riqueza de assuntos que as pessoas abordam nas redes sociais. No entanto, a realidade no ambiente político na época da coleta dos dados, mostrou-se mais pesada, negativa, carregada de raiva, nojo-desprezo e zombaria.

A Tristeza teve uma presença quase insignificante no *dataset*, totalizando 85 *tweets*. Ela é geralmente comunicada a fim de obter ajuda, como visto na secção 2.9.1, Tristeza, mas nesse cenário não se fez presente de forma significativa. Portanto, foi tida como uma emoção com pouca presença quando comparada com as ocorrências das outras emoções. Por isso, não foi utilizada no *dataset* final para o aprendizado de máquina.

De forma semelhante, o medo com apenas 31 ocorrências foi algo raro no ambiente de discussão política, dentro das circunstâncias conhecidas, e por isso, também não entrou para o *dataset* final.

Para a emoção surpresa, não houve como detetá-la, como discutido na secção 3.7.3, Classificação da Surpresa . Portanto foi outra que também ficou fora do *dataset* final.

5.3 Significado das palavras por emoção

Para compreender o peso que as palavras mais frequentes têm dentro do *dataset*, mostradas em 3.9 Análise do significado das palavras, foi feita uma estrutura em Python para saber a percentagem que ao menos uma dessas palavras aparecem dentro de cada *tweet* de cada emoção e obteve-se o resultado mostrado na Tabela 5.1.

Tabela 5.1: A presença das palavras mais comuns no dataset separado por emoção.

Classes	Quantidade	Percentagem
Raiva	1174	87,94%
Nojo-desprezo	672	87,96%
Alegria	942	88,86%
Neutro	353	88,91%

As percentagens ficaram todas próximas de 88% e isso significa que as palavras mais frequentes abordadas e divididas por significados, aparecem na maior parte do *dataset*, mostrando-se com um peso relevante para explicar a composição dos dados.

Agora, discutiremos com base na Tabela 3.4, onde foi dividido os grupos de palavras por significado. O assunto predominante em todas as emoções abordadas está relacionado com as palavras que definem Pessoas, grupos ou organizações. Sendo que a percentagem desse grupo passa de 30% em todas as emoções. Logo, os substantivos ocupam a maior frequência entre as 40 palavras mais usadas dentro de cada grupo de emoção.

As palavras que possuem um significado negativo também se destacam nas emoções, sendo menos frequentes para o grupo das emoções de alegria e neutro. Podemos notar

como emoções que são consideradas negativas possuem mais palavras com sentido negativo chegando em 23,86% para Raiva e 24,61% para Nojo-desprezo.

Buscando uma diferenciação entre os grupos de emoções, podemos destacar dentro de Raiva e Nojo-desprezo, que ambas possuem mais palavras com conotação de distanciamento de grupo, sendo 11,25% para Raiva e 13,15% para Nojo-desprezo. Aqui poderia sugerir, que os significados de afastamento de grupo se destacam em Nojo-desprezo devido a função dessa emoção, pois como visto em 2.9.5 Nojo-desprezo, para os cientistas das emoções, a função dessa emoção é evitar contaminar-se, tendo o objetivo de afastar-se do que causa nojo. E também está relacionado a algo moralmente repulsivo pelo locutor, que nesse caso, seria os valores do grupo político oposto e por consequência, o grupo inteiro propriamente dito. Logo, nesses casos, encaixa-se a função da aversão interpessoal que é afastar-se de algo considerado censurável para evitar o dano social. Dentro de uma abordagem pela teoria do groupthink, poderia sugerir-se que o afastamento aqui seria de algo que vai contra os princípios do grupo político do qual se pertence.

Diferente das outras emoção, a alegria destaca-se justamente pelo oposto, o segundo grupo de palavras com maior peso para essa emoção, são as palavras que possuem um significado de Sentimento de pertencimento a um grupo, sendo 21,67%, enquanto para todas as outras emoções essa percentagem fica na faixa de 7%. Em seguida, na Figura 5.1, podemos ver um gráfico de dispersão comparando as percentagens de palavras com conotação de sentimento de pertencimento a um grupo, contra as que possuem conotação de distanciamento de um grupo, para as três emoções trabalhadas, raiva, alegria e nojo-desprezo.

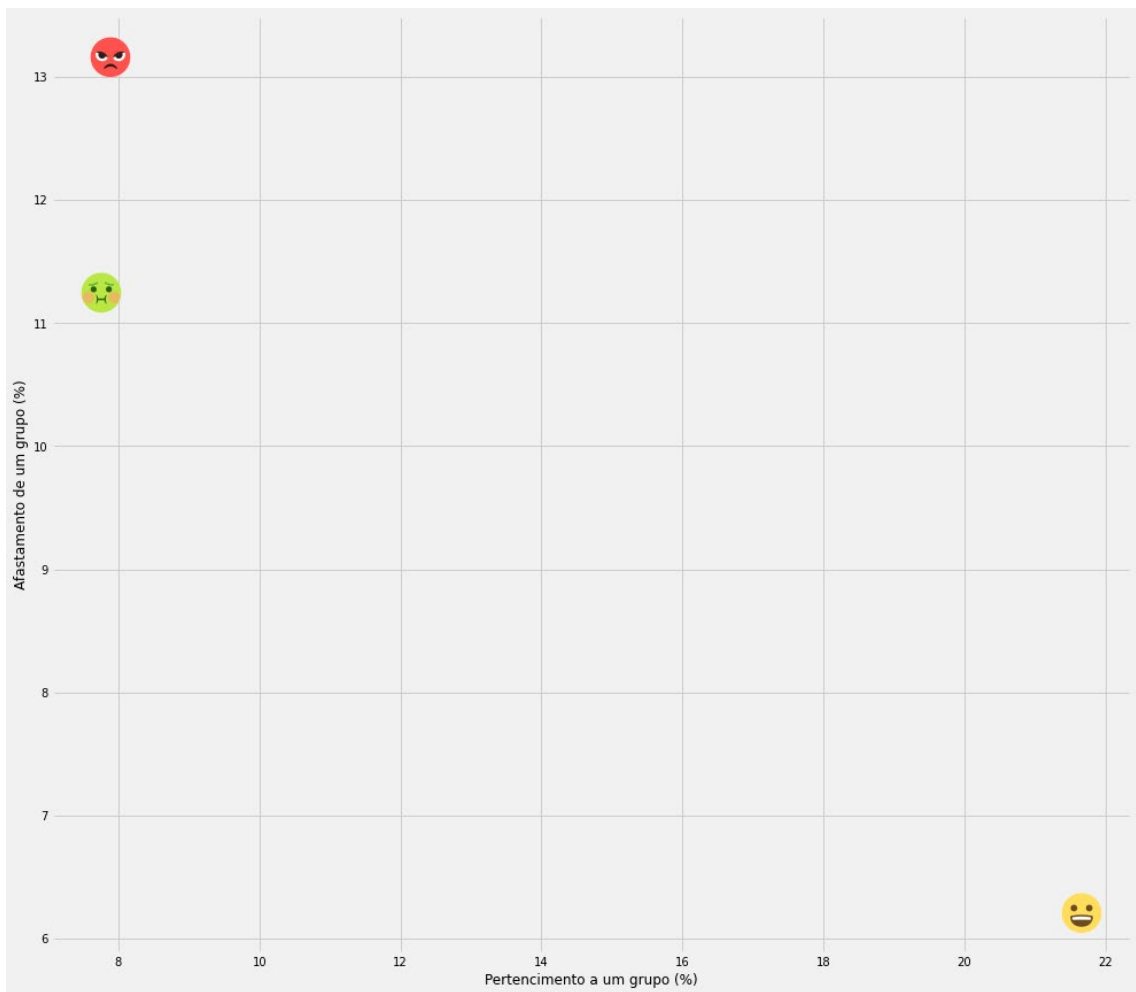


Figura 5.1: Percentagem da quantidade de palavras com significado de pertencimento contra distanciamento de grupo para as três emoções, raiva, nojo-desprezo e alegria.

É possível verificar na Figura 5.1 que as emoções consideradas negativas (raiva e nojo-desprezo) estão próximas uma da outra e com a percentagem mais alta no eixo do distanciamento de grupo. A emoção alegria fica no lado oposto, com palavras mais ricas em significados de pertencimento de grupo.

Para alegria, é interessante destacar o peso das palavras com conotação positiva, que chegam a 8,69% sendo a emoção com maior presença de palavras positivas. Depois de alegria, Nojo-desprezo possui a segunda maior quantidade de palavras positivas com apenas 3,17%, provavelmente devido a zombaria se fazer presente nesta emoção. E raiva não possui palavras com conotação positiva, isso evidencia o caráter da raiva.

Outro grupo de grande peso é o que aborda “Outros assuntos”, como já explicado anteriormente na secção 3.9.1, Palavras mais frequentes no corpus, esses assuntos não se encaixam com os nomes de entidades, como as palavras do maior grupo e são assuntos

relacionados principalmente as notícias. Exclusivamente esses dois grupos: Pessoas, grupos ou organizações e “Outros assuntos”, para neutro a percentagem é maior, sendo 35,16% e 24,33% respetivamente, somando 59,49% do total das 40 palavras mais frequentes de neutro.

5.4 Comportamento de grupo observado

Poderia sugerir que o ambiente das redes sociais pode funcionar como uma simulação do ambiente real, onde os indivíduos se sentem unidos com o grupo. Ilusão de invulnerabilidade foi frequente vista quando utilizadores agrediam verbalmente e até faziam citação ao perfil das vítimas como uma forma de enfrentamento direto. A racionalidade coletiva foi vista quando indivíduos desacreditaram em argumentos plausíveis, muitas vezes científicos, por serem contra os ideias do grupo. Crença na moralidade inerente do grupo é muitas vezes representada pelo termo “passar pano” entre os utilizadores, que significa, “independente do que um líder ou qualquer outro do grupo diga, eu encubro isso”, “passo pano para o político A”. Visões estereotipadas dos membros de grupos opostos estão presentes demasiadamente nas emoções negativas, onde os utilizadores fazem isso por meio de palavrão e adjetivos pejorativos, atribuindo qualidades negativas de pessoas das quais eles não conhecem. Nesse ponto é visível notar uma série de qualidades negativas que são sempre taxadas a membros de grupos opostos, como por exemplo “gado” que seria muitos indivíduos que não possuem senso crítico por pertencerem a um determinado grupo. Ilusão de unanimidade é representada por palavras com significados generalistas, como “todos sabem”, “todo o brasil” e etc.

5.5 Discussão dos modelos

A previsão obteve o melhor desempenho, no quesito precisão e velocidade de classificação, através de modelos de SVM lineares diferentes.

O classificador SGDClassifier obteve o maior desempenho em precisão média e também foi tão rápido na classificação quanto o LinearSVC. Em modo *default*, a função de perda, *loss='hinge'*, vem com “hinge”, por isso o classificador fornece um SVM Linear também.

De acordo com as tabelas, Tabela 4.1 Tabela 4.2, o LinearSVC teve uma precisão mínima de 67,22% e máxima de 77,70% cada um em uma das dobras de teste e desvio padrão de 2,84%. Enquanto SGDClassifier teve o mínimo de 73,31% e máximo de 80,13%, mas o

desvio padrão ficou próximo, com 2,25%. Logo a precisão média do SGDClassifier, foi 2,47% superior.

Pela matriz de confusão, a precisão das emoções alegria e raiva, ficaram próximas uma da outra, pois possuem mais dados, ultrapassando as mil ocorrências cada. A diferença maior foi para neutro e nojo-desprezo que tem menos ocorrências, dificultado o classificador de compreender melhor os dados. Para essas emoções, o SGDClassifier foi maior 7,06% e 4,18% respectivamente, com destaque para neutro, pois apesar de ter a menor precisão entre as emoções, terminou com 51,39%.

6 CONCLUSÃO

No desenvolvimento desta pesquisa, houve grandes desafios e limitações, mas também resultados e contribuições. Objetivos foram alcançados e novas perguntas e ideias surgiram. Portanto discutiremos sobre isso e por fim deixaremos o máximo de ideias possíveis, para qualquer um que se interesse por continuá-las.

6.1 Objetivos alcançados

Os objetivos desse trabalho foram alcançados, apesar de não saber se isso seria possível no início. Esta jornada também gerou muitas outras dúvidas sobre o assunto, algumas serão expostas na secção 6.4, Trabalhos futuros. Os objetivos principais atingidos foram divididos em dois grupos e serão apresentados em seguida.

6.1.1 Compreensão de como identificar as emoções nos textos

- Quais emoções presentes na maioria dos seres humanos?

Foi possível perceber que há emoções primárias, secundárias e sentimentos de fundo. Foi decidido trabalhar com as emoções primárias, pois possuem mais estudos e são menos variáveis do que as emoções secundárias. Percebeu-se também que há diferentes abordagens de acordo com os cientistas envolvidos na busca pelas emoções. Há investigadores que não consideram a surpresa uma emoção e outros investigadores que incluem vergonha, timidez e culpa nas emoções básicas. Portanto, optou-se pelo modelo de Ekman, devido ao vasto conteúdo disponível e pela quantidade de cientistas que aprovam essa abordagem. O modelo de Ekman (Ekman, 1992, 2011) inclui seis emoções, tristeza, raiva, surpresa, medo, nojo-desprezo e alegria.

- O que motiva as pessoas a se emocionarem?

Essa pergunta foi decisiva e orientou uma parte importante do trabalho. O tema e gatilhos inatos das emoções. Essas características definem como as emoções ativadas. Devido a esses gatilhos terem sido validados por muitos cientistas, neste trabalho foram utilizados como forma para abordar as emoções na interpretação dos textos. E a partir disso, foi norteado a classificação manual para a criação do *dataset*.

- Há significados que despertam as emoções? Como assuntos ou temas, pessoas públicas, elementos culturais como música, filmes ou literatura?

Essa pergunta avança num campo difuso, onde cada pessoa cria uma relação específica com os significados, sendo isso consciente ou não. A partir daqui foi explorado os significados das palavras por grupos. Na etapa da classificação manual, criou-se a impressão de que algumas pessoas nutrem um ódio tão grande contra certas pessoas públicas que apenas o nome de tal pessoa parecia despertar raiva entre outras emoções. Mas, isso não é possível afirmar, pois não foi encontrado uma forma de fazer uma experiência a partir apenas do texto publicado no Twitter. Um trabalho feito por Rozin (Rozin et al., 1999), citado anteriormente, sugere que políticos enojam as pessoas de muitas culturas.

- Há significados que despertam mais de uma emoção simultaneamente?

Vimos que as emoções se relacionam entre si, se alternam ou substituem uma a outra. Um desses exemplos é a raiva quando muito intensa, a qual pode provocar aversão e ambas serem vivenciadas juntas. Mas, quando falamos de significados, ficamos na mesma situação da pergunta anterior.

- Quais ações as emoções despertam?

Aqui foi norteado a parte das reações que as emoções causam, assim como o objetivo de delas, como a tristeza quando compartilhada que pode significar um pedido de ajuda.

- Há palavras ou frases que são consequências de emoções?

Identificamos muitos adjetivos, substantivos e palavras no ambiente do debate, que remetem a emoções específicas. Porém, não se podem diferenciar as classificações das emoções apenas por essas palavras, pois a comunicação por meio de texto, possui significados mais complexos e por vezes algumas palavras que ilustram bem uma emoção, podem ser usadas em outro contexto, evidenciando um gatilho para outra emoção.

- Há significados que despertam as emoções na mente coletiva ou na opinião pública?

Sobre significados a pergunta acaba sendo um pouco redundante, mas no aspeto do comportamento de grupo, pode-se perceber que no ambiente de discussão, as pessoas têm uma tendência em agir e reagir como grupos. Um exemplo foi evidenciado no gatilho da emoção *fiero*, onde grupos pareciam saborear uma conquista de todo o grupo.

- Motivações pragmáticas influenciam as emoções?

Essa pergunta não pôde ser respondida, pois, para entender as motivações que levaram as pessoas a se comunicarem da forma e com os objetivos com que fizeram, seria necessária uma pesquisa mais complexa e profunda que essa. No entanto, pode fazer-se uma sugestão com base nalguns textos argumentativos que se mostraram imparciais. Esses defendiam um lado político com argumentos lógicos e sem evidências emocionais, no entanto poderiam ser motivados por emoções, ou seja, essa pergunta de forma invertida. Em vez de haver motivações que influenciaram alguma emoção, as emoções podem ter sido o motivo pragmático.

- Como ter um panorama pragmático de quem se comunica pelo Twitter?

Visando esse objetivo, foram coletadas informações do perfil de cada utilizador que teve um tweet presente nos dados de estudo. Informações como data de criação do perfil, nome, *screen name*, *user id*, local onde reside, quantidade de seguidores, quantidade de perfis que segue, quantidade de postagens e comentários curtidos, quantidade de *tweets* e respostas publicadas, link colocado no perfil, em um local específico para isso e a descrição do perfil. No total foram coletados 26419 perfis não repetidos. Esses dados podem ser úteis para desenvolver alguma pesquisa no âmbito pragmático. Em particular, a descrição do perfil é muito interessante, pois nela a maioria das pessoas colocam o posicionamento político, orientação sexual, escolaridade, frases de escritores por vezes do universo político, gostos pessoais e muitas outras informações relacionadas a personalidade. No entanto, por falta de tempo, a presente pesquisa não contemplou desenvolvimento nesse sentido.

6.1.2 Identificação dos padrões analisados

- Foi possível identificar algum padrão na estrutura das frases ou mesmo através dos significados expressos em palavras?

Foram identificados alguns padrões a nível semântico, nas quantidades de palavras de acordo com os significados, como já discutido. Como por exemplo, as emoções consideradas negativas, raiva e nojo-desprezo que possuem palavras com mais significado de distanciamento de grupo, enquanto alegria possui palavras com mais sentimento de pertencimento a um grupo. Os gatilhos foram padrões que nortearam a classificação manual, mas eles não seguem um padrão sintático claro.

- Qual classificador melhor deteta esses padrões?

A princípio acreditava-se que classificadores baseados em agrupamentos resultariam melhor para os padrões identificados. Este é o caso do *KNeighborsClassifier*, mas, o qual não obteve um bom desempenho. Classificadores considerados bons para classificação de texto, como os probabilísticos, além de outros, como discutido anteriormente. No entanto, foram os modelos baseados em SVM lineares que prevaleceram.

6.2 Contributos e implicações

Um importante contributo desse trabalho, é a forma adotada de identificação das emoções básicas em textos, baseando-se nos temas e gatilhos inatos. É possível que a precisão obtida reflita uma boa qualidade de classificação manual, uma vez que o *dataset* é pequeno. A quantidade de classes também influencia na dificuldade com que o classificador faz o trabalho, pois, conforme aumenta a quantidade de classes, mais difícil é de se obter uma precisão mais alta. Além disso, os classificadores selecionados são métodos simples, quando comparado a técnicas de *Deep Learning* ou mesmo os Ensembles. Portanto, aqui há uma nova abordagem para a qualidade da classificação e a forma de como os textos envolvidos em emoções são vistos.

No trabalho de PLN principalmente nas redes sociais, há muitos neologismos, abreviações e gírias. Isso dificulta a padronização dos textos, pois sem ser corrigido, o classificador identifica cada palavras como diferentes, ainda que tenham o mesmo significado. Por isso, outro contributo desse trabalho está resumido na Tabela 3.5: Abreviações encontradas no corpus. Foram pesquisados manualmente e identificados 57 neologismos, abreviações e gírias. Portanto, essas palavras podem ser utilizadas por outros pesquisadores desta área.

Ao analisar o significado das palavras mais usadas, obtiveram-se algumas informações que diferenciaram uma emoção das outras, especificamente na seção 3.9.2 Palavras mais frequentes separadas por emoções. Aí, pode-se perceber que algumas emoções têm diferenças nas quantidades com que grupos de palavras com determinados significados aparecem. O que foi mais notável, é que os significados estão devidamente alinhados com as características relatadas pelos cientistas das emoções.

O conjunto de dados utilizado foi retirado do Twitter, utilizando em média, quatro *Tred Topics* sobre assuntos políticos que foram coletadas do dia 22 de junho de 2020 até o dia 24 de agosto 2020, totalizando 190 *hashtags*. Foi uma época politicamente conturbada e polarizada, agravada pela difícil situação económica e sanitária do Brasil. Com as

dificuldades provocadas pelo Covid-19, houve uma grande dificuldade de gestão por parte dos governos mundiais, incluindo o do Brasil, entre a opção de maior confinamento e menos consequências imediatas para a economia e a opção inversa. Isso contribuiu para que as pessoas ficassem mais engajadas na política. O estudo evidenciou a polarização política num sentido onde as pessoas se opuseram completamente umas contra as outras, discordando, demonstrando emoções maioritariamente pesadas, desejando o mal e até a morte umas das outras.

6.3 Limitações e desafios

O trabalho engloba aspetos de psicologia, sociologia, ciências políticas, comunicação social, linguística e ciência de dados. Portanto, envolve alguma complexidade na gestão e aplicação do conhecimento. Isso acabou por se tornar uma clara limitação.

A linguagem nas redes sociais é confusa. Muitas vezes as pessoas escrevem errado ou entendem algo errado, deturpando a sua compreensão da mensagem. Além disso, o uso de imagens, links externos, neologismos, gírias e abreviações, torna o trabalho de interpretação cansativo, onde, em alguns casos, envolve muita análise para compreender apenas um *tweet*.

Além disso, o trabalho de classificação manual, já é, por si só, demorado e requer um grande esforço. Calcula-se que para criar as estruturas em Python, extrair e classificar os dados manualmente se gastaram cerca de 645 horas, sendo a parte de classificação manual a mais dispendiosa delas.

Para fazer a modelação, combinando as possibilidades de pré-processamento, parâmetros da redução de dimensionalidade e de validação, houve muitas possibilidades e a técnica de redução demora cerca de 10 minutos para ser ajustada. Então, para afinar essa parte, foram feitas mais de 300 execuções.

Uma limitação importante, foi a falta de mais poder computacional, limitando assim a afinação dos classificadores por meio de GridSearch e impossibilitando de extrair o máximo de cada um. O tempo que cada execução demorou foi demasiado com os recursos computacionais utilizados.

Outra limitação foi a de que alguns *tweets* possuíam duas emoções, como raiva e alegria. Onde o utilizador utilizava adjetivos pejorativos no intuito de causar dano psicológico

nos opositores e em seguida fazia uma afirmação orgulhosa sobre o grupo político que ele se identifica. Como foram poucos, esses *tweets* ficaram de fora do *dataset*.

6.4 Trabalhos futuros

A fim de padronizar o texto foi aplicado lematização com auxílio do pacote NLTK. No entanto, essa técnica para a língua portuguesa não é desenvolvida. Por isso, houve muitos erros em muitos verbos e esses não puderam ser padronizados. Tal como foi o aplicado na padronização realizada com as abreviações, que obteve resultados melhores. Portanto, como trabalho futuro poderia sugerir-se a aplicação de uma lematização melhor traria alguma melhoria na classificação.

Na busca por diferenciar as emoções, chegou-se até aos significados dos grupos de palavras, que foram discutidos na secção 3.9.2, Palavras mais frequentes separadas por emoções. Num trabalho futuro, esses dados podem ser mais bem explorados numa etapa inicial, ao atribuir pesos diferentes de acordo com o enquadramento dos significados das palavras. Isso resultaria em acentuar as diferenças dos grupos de palavras característicos de cada emoção.

Outra ideia semelhante é a de atribuir pesos diferentes aos *emojis*, pois os de raiva, alegria e nojo-desprezo, mostraram-se com um padrão claro com *emojis* específicos de cada emoção.

A afinação dos classificadores com técnicas como o GridSearch, de forma a utilizar o máximo potencial de cada algoritmo, também se poderá revelar uma mais-valia em trabalhos futuros idênticos.

Houve uma suspeita comentada no desenvolvimento do trabalho, que está relacionada a gatilhos próprios, desenvolvidos no âmbito político, como nomes de algumas pessoas públicas, de partidos políticos ou de assuntos. Esses parecem despertar emoções por apenas aparecerem em alguma postagem, onde os utilizadores, dos mesmos grupos políticos, reagiram de forma parecidas. Um trabalho futuro poderia também tentar investigar esse possível efeito.

Poderia igualmente ser desenvolvido uma investigação que utilize as informações pessoais dos perfis dos utilizadores e também nas postagens deles, de forma a identificar padrões que se repetem na maneira em como determinado utilizador comunica. Essa investigação poderia tentar responder à questão de saber até onde os utilizadores com as

mesmas ideologias, valores, pertencentes a mesmos grupos políticos reagem a determinados assuntos de forma padronizada.

REFERÊNCIAS BIBLIOGRÁFICAS

- APA. (2020). *APA Dictionary of Psychology* (G. R. VandenBos (ed.)). American Psychological Association. <https://dictionary.apa.org/>
- Arriaga, P. & A. G. (2010). *Fabrica de emoções eficácia excerto filmes na inducao emocoos*. 8(1), 63–80.
- Arruda, M. de J. F. C. (2014). O ABC das emoções básicas. *Universidade Dos Açores*, 24–40.
- Batbaatar, E., Li, M., & Ryu, K. H. (2019). Semantic-Emotion Neural Network for Emotion Recognition From Text. *IEEE Access*, 7, 111866–111878. <https://doi.org/10.1109/access.2019.2934529>
- Beaumont, S. L. (2009). Contentment. In S. J. Lopez (Ed.), *The Encyclopedia of Positive Psychology*. Blackwell Publishing Ltd.
- Belfo, F. P., & Andreica, A. B. (2018). A Comprehensive Methodology to Implement Business Intelligence and Analytics Through Knowledge Discovery in Databases. *Mining Intelligence and Knowledge Exploration. MIKE 2018. Lecture Notes in Computer Science, 11308*, 102–111. https://doi.org/https://doi.org/10.1007/978-3-030-05918-7_10
- Berkowitz, L. (1989). Frustration-aggression hypothesis: Examination and reformulation. In *Psychological Bulletin* (Vol. 106, Issue 1, pp. 59–73). American Psychological Association. <https://doi.org/10.1037/0033-2909.106.1.59>
- Bisong, E. (2019). Google Colaboratory. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners* (pp. 59–64). Apress. https://doi.org/10.1007/978-1-4842-4470-8_7
- Brandão, L., Belfo, F. P., & Silva, A. (2021). Wavelet-based cancer drug recommender system. *Procedia Computer Science, Communications in Computer and Information Science, 181*, 487–494. <https://doi.org/https://doi.org/10.1016/j.procs.2021.01.194>
- Britannica, T. E. of E. (2020). Pragmatics. In *Encyclopedia Britannica*. <https://www.britannica.com/science/pragmatics>
- Broekens, J. (2018). *A Temporal Difference Reinforcement Learning Theory of Emotion: unifying emotion, cognition and adaptive behavior*.

- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-By-Step Data Mining Guide*. SPSS, CRISP-DM Consortium.
- Clanton, G. (2006). Jealousy and Envy. In J. E. Stets & J. H. Turner (Eds.), *Handbook of the sociology of emotions*. Springer Science+Business Media.
- Costa-Luis, C. O. da. (2019). tqdm: A Fast, Extensible Progress Meter for Python and CLI. *The Open Journal*, 4, 1277. <https://doi.org/10.21105/joss.01277>
- Cui, G., Wong, M. L., & Lui, H.-K. (2006). Machine learning for direct marketing response models: Bayesian networks with evolutionary programming. *Management Science*, 52(4), 597–612.
- Damásio, A. R. (1994). *O erro de Descartes: emoção, razão e o cérebro humano*. SCHWARCZ S.A.
- Darwin, C. (1965). *The expression of the emotions in man and animals*. University of Chicago Press.
- Davidson, R. J., & Ekman, P. (1994). Affective Science: A research agenda. In R. J. Davidson & P. Ekman (Eds.), *The Nature of Emotion: Fundamental Questions* (pp. 411–431). Oxford University Press.
- Devika, M. D., Sunitha, C., & Ganesh, A. (2016). Sentiment Analysis: A Comparative Study on Different Approaches. *Procedia Computer Science*, 87, 44–49. <https://doi.org/10.1016/j.procs.2016.05.124>
- Ducange, P., Fazzolari, M., Petrocchi, M., & Vecchio, M. (2019). An effective Decision Support System for social media listening based on cross-source sentiment analysis models. *Engineering Applications of Artificial Intelligence*, 78(November 2018), 71–85. <https://doi.org/10.1016/j.engappai.2018.10.014>
- Edward, L., Ewan, K., & Steven, B. (2009). *Natural Language Processing with Python*. O'Reilly Media Inc. <https://www.nltk.org/book/>
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>
- Ekman, P. (1999). Basic Emotions. In T. Dalgleish & M. J. Power (Eds.), *Handbook of Cognition and Emotion*. John Wiley & Sons.

- Ekman, P. (2011). A Linguagem das Emoções. In *Lua de Papel* (Vol. 1). Lua de Papel. <https://doi.org/10.1017/CBO9781107415324.004>
- Ekman, P. (2016). What Scientists Who Study Emotion Agree About. *Perspectives on Psychological Science*, *11*(1), 31–34. <https://doi.org/10.1177/1745691615596992>
- Ekman, P., & Friesen, W. V. (1975). Unmasking the face: A guide to recognizing emotions from facial clues. In *Unmasking the face: A guide to recognizing emotions from facial clues*. Prentice-Hall.
- Ekman, P., Friesen, W. V., & Simons, R. C. (1985). Is the startle reaction an emotion? In *Journal of Personality and Social Psychology* (Vol. 49, Issue 5, pp. 1416–1426). American Psychological Association. <https://doi.org/10.1037/0022-3514.49.5.1416>
- Emmons, R. A., & Shelton, C. M. (2002). Gratitude and the Science of Positive Psychology. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of Positive Psychology, 3rd Edition*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199396511.001.0001>
- Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, *56*(3), 218–226. <https://doi.org/10.1037/0003-066X.56.3.218>
- Fredrickson, B. L., Mancuso, R. A., Branigan, C., & Tugade, M. M. (2000). The undoing effect of positive emotions. *Motivation and Emotion*, *24*(4), 237–258. <https://doi.org/10.1023/A:1010796329158>
- Fulcher, J. S. (1942). “Voluntary” facial expression in blind and seeing children. *Archives of Psychology (Columbia University)*, *272*, 49.
- Galati, D., Scherer, K. R., & Ricci-Bitti, P. E. (1997). Voluntary facial expression of emotion: Comparing congenitally blind with normally sighted encoders. *Journal of Personality and Social Psychology*, *73*(6), 1363–1379. <https://doi.org/10.1037/0022-3514.73.6.1363>
- Gallup, G. H. J. (1998). *Thankfulness: America’s saving grace*.
- Haidt, J. (2000). The Positive emotion of elevation. *Prevention & Treatment*, *3*(1), No Pagination Specified-No Pagination Specified. <https://doi.org/10.1037/1522-3736.3.1.33c>

- Haidt, J., & Keltner, D. (2004). Appreciation of Beauty and Excellence. In C. Peterson & M. E. P. Seligman (Eds.), *Character Strengths and Virtues: A Handbook and Classification*. American Psychological Association & Oxford University Press USA.
- Hang, W., & Banks, T. (2019). Machine learning applied to pack classification. *International Journal of Market Research*, 61(6), 601–620. <https://doi.org/10.1177/1470785319841217>
- Hipp, R. D. (2020). *SQLite*. <https://www.sqlite.org/index.html>
- Holder, M. D. (2009). Neurobiology. In S. J. Lopez (Ed.), *The Encyclopedia of Positive Psychology*. Blackwell Publishing Ltd.
- Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2020). *spaCy: Industrial-strength Natural Language Processing in Python*. Zenodo. <https://doi.org/10.5281/zenodo.1212303>
- Huber, S., Wiemer, H., Schneider, D., & Ihlenfeldt, S. (2019). DMME: Data mining methodology for engineering applications—a holistic extension to the CRISP-DM model. *Procedia Cirp*, 79, 403–408. <https://doi.org/10.1016/j.procir.2019.02.106>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95.
- Keltner, D., & Haidt, J. (2003). Approaching awe: a moral, aesthetic, and spiritual emotion. *Cognition and Emotion*.
- Kemper, T. D. (1987). How Many Emotions Are There? Wedding the Social and the Autonomic Components. *American Journal of Sociology*, 93(2), 263–289. <https://doi.org/10.1086/228745>
- Khurana, D., Koli, A., Khatter, K., & Singh, S. (2017). *Natural Language Processing: State of The Art, Current Trends and Challenges. Figure 1*. <http://arxiv.org/abs/1708.05148>
- Kocaleva, M., Stojanov, D., Stojanovik, I., & Zdravev, Z. (2016). Pattern Recognition and Natural Language Processing: State of the Art. *Tem Journal*, 5(2), 236–240. <https://doi.org/10.18421/TEM52-18>
- Lazarus, R. S., & Lazarus, B. N. (1994). *Passion and reason: Making sense of our*

- emotions*. Oxford University Press.
- Lazarus, Richard S. (1991). *Emotion and Adaptation*. Oxford University Press.
- Leach, C. W., Spears, R., Branscombe, N. R., & Doosje, B. (2003). Malicious Pleasure: Schadenfreude at the Suffering of Another Group. *Journal of Personality and Social Psychology*, 84(5), 932–943. <https://doi.org/10.1037/0022-3514.84.5.932>
- Lewis, M. (2008). Self-conscious emotions. In M. Lewis & J. Haviland-Jones (Eds.), *The Handbook of Emotions*. Guilford Press.
- Loureiro, A., Lourenço, J., Costa, E., & Belfo, F. (2014). Indução de Árvores de Decisão na Descoberta de Conhecimento: Caso de Empresa de Organização de Eventos. In *VI Congresso Internacional de Casos Docentes em Marketing Público e Não Lucrativo*.
- Magalhães, A. (2007). *A Psicologia das emoções: o Fascínio do Rosto Humano*. Edições Universidade Fernando Pessoa.
- Magalhães, A. F. (2013). *O Código de Ekman - O Cérebro, a Face e a Emoção*. Feelab Science Books.
- McKinney, W., & others. (2010). Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference*, 445, 51–56.
- Miles, D. R., & Carey, G. (1997). Genetic and environmental architecture on human aggression. *Journal of Personality and Social Psychology*, 72(1), 207–217. <https://doi.org/10.1037//0022-3514.72.1.207>
- Miller, W. I. (1997). *The Anatomy of Disgust*. Harvard University Press.
- Moerland, T. M., Broekens, J., & Jonker, C. M. (2018). Emotion in reinforcement learning agents and robots: a survey. In *Machine Learning* (Vol. 107, Issue 2). Springer US. <https://doi.org/10.1007/s10994-017-5666-0>
- Nietzsche, F. (2007). *Genealogia da Moral: Uma Polêmica* (P. C. Souza (ed.)). Companhia das Letras.
- Oesper, L., Merico, D., Isserlin, R., & Bader, G. D. (2011). WordCloud: a Cytoscape plugin to create a visual semantic summary of networks. *Source Code for Biology and Medicine*, 6(1), 7.
- Ohman, A. (2008). Fear and anxiety: evolutionary, cognitive and clinical perspectives. In

- Micheal Lewis & J. Haviland-Jones (Eds.), *The Handbook of Emotions* (The Guilfo).
Oliphant, T. E. (2006). *A guide to NumPy* (Vol. 1). Trelgol Publishing USA.
- Óscar, M., Gonzalo, M., & Javier, S. (2016). A Data Mining & Knowledge Discovery Process Model. *Intech*. <https://doi.org/http://dx.doi.org/10.5772/57353>
- Pedregosa, F., Varoquaux, Gaël, Gramfort, A., Michel, V., Thirion, B., & Grisel, O. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*. <https://scikit-learn.org/stable/>
- Pimenta, C., Ribeiro, R., Sá, V., & Belfo, F. P. (2018). Fatores que Influenciam o Sucesso Escolar das Licenciaturas numa Instituição de Ensino Superior Portuguesa. In *Atas da 18ª Conferência da Associação Portuguesa de Sistemas de Informação (CAPSI 2018) Associação Portuguesa de Sistemas de Informação*. Associação Portuguesa de Sistemas de Informação.
- Pimenta, P., Belfo, F., & Trigo, A. (2009). Study the impact of Booking. com user scores and reviews in hotel management. *Book of Abstracts of the CENTERIS 2011–Conference on Enterprise Information Systems*, 30, 8.
- Plomin, R., Nitz, K., & Rowe, D. C. (1990). Behavioral genetics and aggressive behavior in childhood. In *Handbook of developmental psychopathology*. (pp. 119–133). Plenum Press. https://doi.org/10.1007/978-1-4615-7142-1_10
- Rhudy, J. L., & Meagher, M. W. (2000). Fear and anxiety: Divergent effects on human pain thresholds. *Pain*, 84(1), 65–75. [https://doi.org/10.1016/S0304-3959\(99\)00183-9](https://doi.org/10.1016/S0304-3959(99)00183-9)
- Roesslein, J. (2020). *Tweepy: Twitter for Python!* <https://github.com/Tweepy/Tweep>
- Roh, T., Jeong, Y., Jang, H., & Yoon, B. (2019). Technology opportunity discovery by structuring user needs based on natural language processing and machine learning. *PLoS ONE*, 14(10), 1–28. <https://doi.org/10.1371/journal.pone.0223404>
- Rozin, P., Haidt, J., & McCauley, C. R. (1999). Disgust: the body and soul emotion. In T. Dalgleish & M. J. Power (Eds.), *Handbook of Cognition and Emotion*. John Wiley & Sons.
- Sailunaz, K., & Alhaji, R. (2019). Emotion and Sentiment Analysis from Twitter Text. *Journal of Computational Science*, 36, 101003.

<https://doi.org/10.1016/j.jocs.2019.05.009>

- Santos, L. O. dos. (2003). O medo contemporâneo: abordando suas diferentes dimensões. *Psicologia: Ciência e Profissão*, 23(2), 48–49. <https://doi.org/10.1590/s1414-98932003000200008>
- Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition & Emotion*, 23(7), 1307–1351. <https://doi.org/10.1080/02699930902928969>
- Seiça, Alcides, Trigo, A., & Belfo, F. P. (2019). LexiNB - Uma abordagem bietápica de classificação de sentimentos em tweets relacionados com as autoridades fiscais portuguesas. *19.ª Conferência Da Associação Portuguesa de Sistemas de Informação (CAPSI'2019)*.
- Seiça, António, Trigo, A., & Belfo, F. P. (2019). LexiNB - Uma Abordagem Bietápica de Classificação de Sentimentos em Tweets Relacionados com as Autoridades Fiscais Portuguesas. *Proceedings of the 19.ª Conferência Da Associação Portuguesa de Sistemas de Informação (CAPSI'2019) Held in Lisboa, Portugal, 11-12 October 2019. Paper 5*.
- Sereday, S., & Cui, J. (2017). Using machine learning to predict future tv ratings. *Data Science, Nielsen*, 1(3), 3–12.
- Shelton, C. M. (1990). *Morality of the heart: A psychology for the Christian moral life*. New York: Crossroad.
- Shiota, M. N. (2009). Amusement. In S. J. Lopez (Ed.), *The Encyclopedia of Positive Psychology*. Blackwell Publishing Ltd.
- Shiota, M. N., Thrash, T. M., Danvers, A. F., & Dombrowski, J. T. (2014). Transcending the Self: Awe, Elevation and Inspiration. *Handbook of Positive Emotions*, c, 362–377.
- Spears, R., & Leach, C. W. (2004). Intergroup Schadenfreude: Conditions and Consequences. *The Social Life of Emotions*.
- Sternberg, C. R., & Campos, J. J. (1990). The development of anger expressions in infancy. In *Psychological and biological approaches to emotion*. (pp. 247–282). Lawrence Erlbaum Associates, Inc.

- Stets, J. E. (2006). Emotions and Sentiments. In J. Delamater (Ed.), *Handbook of social psychology*. Springer Science+Business Media.
- Thompson, J. (1941). Development of facial expression of emotion in blind and seeing children. *Archives of Psychology (Columbia University)*, 264, 47.
- Turner, J. H. (2007). *Human Emotions A Sociological Theory*. Routledge.
- Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.
<https://docs.python.org/3/reference/>
- VOGT, S. (2017). *O groupthink como aspecto que influencia a tomada de decisão em grupo*.
- Wolf, M. (1989). *TEORIAS DA COMUNICAÇÃO Mass media: contextos e paradigmas Novas tendências Efeitos a longo prazo O newsmaking*. Gruppo Editoriale Fabbri, Bompiani, Sonzogno, Etas S.p.A.