



**Instituto Superior de Gestão e Administração de Santarém**

**Mestrado em Engenharia de Tecnologias e Sistemas Web**

**Dissertação**

Sistemas de Apoio Técnico Inteligentes

Carlos Silva, a22208573

Santarém

Ano 2023



**Instituto Superior de Gestão e Administração de Santarém**

**Mestrado em Engenharia de Tecnologias e Sistemas Web**

**Dissertação**

Sistemas de Apoio Técnico Inteligentes

Carlos Silva, a22208573

Dissertação submetida para satisfação parcial dos requisitos do grau de Mestre em Engenharia de Tecnologias e Sistemas Web sob a orientação do Prof. Doutor Pedro Sobreiro.

Santarém

Ano 2023

## RESUMO

**Introdução:** O relacionamento com os clientes é crucial para o sucesso dos negócios, com o e-mail, a funcionar com canal de comunicação principal. No entanto, o seu tratamento em tempo útil pode ser difícil e moroso. Este trabalho tem como objetivo propor um mecanismo de classificação de textos e análise de sentimentos, permitindo que as entidades façam o tratamento dos pedidos de forma a priorizem os clientes com mensagens negativas para um atendimento mais imediato, de forma automática.

**Método:** Foi usado um conjunto de dados referente a solicitações de assistências de clientes. Estes foram anonimizados. Foi realizado o seu tratamento, através do pré-processamento, extração de atributos e balanceamento por categoria. De seguida foram implementados os algoritmos de DT, RF, NB, KNN e LSTM, quer para a classificação, quer para análise de sentimento. Por fim, foi realizada a avaliação dos resultados obtidos.

**Resultado:** De todos os algoritmos implementados, o que teve melhor desempenho foi o RF e AB, com valores de *accuracy*, *recall* e *f1-score*, de 100% e 99.38%, para classificação e sentimento, respetivamente. As técnicas de balanceamento usadas foram ROS e SMOTE e a extração de atributos através de TF-IDF, sem a aplicação de CV.

**Discussão:** Os resultados obtidos para o processo de classificação são melhores do que os analisados na revisão de literatura, como o algoritmo RF a obter 100% de *accuracy*. Em relação à análise de sentimento, os resultados estão de acordo com os analisados, com o melhor a ser RF com 99,38% de *accuracy*.

**Conclusão:** Os métodos e algoritmos implementados, têm uma eficácia bastante grande, com recuso à implementação de técnicas de balanceamento devido a estes estarem enviesados. Os resultados obtidos, mostram que a implementação de um sistema que faça uso destas ferramentas, pode tratar de analisar os emails de forma automática com uma fiabilidade elevada.

*Palavras-chave: algoritmos; análise de sentimento; classificação texto; helpdesk; pós-venda.*

## ABSTRACT

**Introduction:** Relationships with customers are crucial to business success, with email serving as the main communication channel. However, timely treatment can be difficult and time-consuming. This work aims to propose a text classification and sentiment analysis mechanism, allowing entities to process requests to prioritize customers with negative messages for more immediate service, automatically.

**Method:** A set of data relating to customer assistance requests was used. These have been anonymized. Its treatment was carried out through pre-processing, attribute extraction and balancing by category. The DT, RF, NB, KNN and LSTM algorithms were then implemented, both for classification and sentiment analysis. Finally, the results obtained were evaluated.

**Result:** Of all the algorithms implemented, the one that performed best was RF and AB, with accuracy, recall and f1-score values of 100% and 99.38%, for classification and sentiment, respectively. The balancing techniques used were ROS and SMOTE and attribute extraction through TF-IDF, without applying CV.

**Discussion:** The results obtained for the classification process are better than those analyzed in the literature review, with the RF algorithm obtaining 100% accuracy. Regarding sentiment analysis, the results are in line with those analyzed, with the best being RF with 99.38% accuracy.

**Conclusion:** The implemented methods and algorithms are very effective, with the refusal to implement balancing techniques due to them being biased. The results obtained show that the implementation of a system that makes use of these tools can analyze emails automatically with high reliability.

*Keywords: algorithms; sentiment analysis; text classification; helpdesk; after sales.*

## ÍNDICE

<b>RESUMO</b> .....	<b>I</b>
<b>ABSTRACT</b> .....	<b>I</b>
<b>ÍNDICE</b> .....	<b>III</b>
<b>ÍNDICE DE FIGURAS</b> .....	<b>VI</b>
<b>ÍNDICE DE TABELAS</b> .....	<b>VII</b>
<b>ABREVIATURAS E SÍMBOLOS</b> .....	<b>VIII</b>
<b>INTRODUÇÃO</b> .....	<b>1</b>
PRÉ-PROCESSAMENTO .....	2
EXTRAÇÃO DE ATRIBUTOS .....	3
CLASSIFICAÇÃO DE TEXTO .....	4
ANÁLISE DE SENTIMENTO .....	5
PROBLEMA E QUESTÕES DE INVESTIGAÇÃO .....	5
ESTRUTURA DO RELATÓRIO .....	7
<b>REVISÃO DA LITERATURA</b> .....	<b>8</b>
QUAL É O ESTADO DA PESQUISA ATUAL SOBRE SISTEMAS DE APOIO TÉCNICO INTELIGENTES? .....	10
QUAIS SÃO OS ALGORITMOS QUE SÃO APLICADOS EM SISTEMAS DE SUPORTE TÉCNICO PARA CLASSIFICAÇÃO DE TEXTO E ANÁLISE DE SENTIMENTO? .....	10
DE QUE FORMA SÃO UTILIZADOS OS ALGORITMOS PARA CLASSIFICAR TEXTO E PARA ANÁLISE DE SENTIMENTO? .....	11
<i>Classificação de Texto</i> .....	11
<i>Análise de Sentimento</i> .....	14
COMO SÃO INTEGRADOS ESTES ALGORITMOS NOS SISTEMAS DE <i>HELPDESK</i> ? .....	16
COMO É MEDIDA A SUA EFICÁCIA E EFICIÊNCIA? .....	16
CONCLUSÃO .....	18
<b>METODOLOGIA</b> .....	<b>18</b>
TECNOLOGIAS A USAR .....	18
ARQUITETURA .....	19
PLANEAMENTO .....	19
<i>Análise de Risco</i> .....	20

CONJUNTO DE DADOS .....	21
PRÉ-PROCESSAMENTO .....	23
EXTRAÇÃO DE ATRIBUTOS.....	23
<i>TF-IDF</i> .....	24
<i>Word2Vec</i> .....	24
<i>Doc2Vec</i> .....	24
ALGORITMOS USADOS .....	24
<i>Gaussian NB</i> .....	25
<i>DT</i> .....	25
<i>RF</i> .....	27
<i>AB</i> .....	27
<i>KNN</i> .....	27
<i>LSTM</i> .....	28
AVALIAÇÃO DESEMPENHO .....	30
BALANCEAMENTO .....	31
<b>RESULTADOS .....</b>	<b>33</b>
POR EXTRATOR .....	33
<i>TF-IDF</i> .....	33
<i>Word2Vec</i> .....	35
<i>Doc2Vec</i> .....	36
ALGORITMO .....	38
<i>NB</i> .....	38
<i>DT</i> .....	40
<i>RF</i> .....	41
<i>AB</i> .....	43
<i>KNN</i> .....	44
<i>LSTM</i> .....	46
GLOBAL .....	48
<b>DISCUSSÃO DOS RESULTADOS .....</b>	<b>50</b>
DADOS .....	51
PRÉ-PROCESSAMENTO E EXTRAÇÃO DE ATRIBUTOS.....	51
CLASSIFICAÇÃO DE TEXTO .....	52
ANÁLISE DE SENTIMENTO .....	55

SISTEMAS PROPOSTOS .....	57
LIMITAÇÕES .....	58
<b>CONCLUSÕES .....</b>	<b>59</b>
TRABALHOS FUTUROS .....	60
<b>BIBLIOGRAFIA .....</b>	<b>61</b>
<b>APÊNDICES .....</b>	<b>65</b>

## ÍNDICE DE FIGURAS

<b>Figura 1.</b> <i>Pré-processamento de Texto</i> .....	2
<b>Figura 2.</b> <i>Processo de Classificação de Texto</i> .....	4
<b>Figura 3.</b> <i>Arquitetura</i> .....	19
<b>Figura 4.</b> <i>Distribuição dos dados pelas classes e sentimento</i> .....	22
<b>Figura 5.</b> <i>Exemplo de DT</i> .....	26
<b>Figura 6.</b> <i>Fluxograma KNN</i> .....	27
<b>Figura 7.</b> <i>Modelo LSTM</i> .....	29
<b>Figura 8.</b> <i>Loss e accuracy - classificação</i> .....	46
<b>Figura 9.</b> <i>Loss e accuracy - sentimento</i> .....	47

## ÍNDICE DE TABELAS

<b>Tabela 1.</b> <i>Critérios PICOC</i> .....	8
<b>Tabela 2.</b> <i>Resumo Métricas</i> .....	16
<b>Tabela 3.</b> <i>Tabela de Riscos</i> .....	20
<b>Tabela 4.</b> <i>Descrição do conjunto de dados</i> .....	21
<b>Tabela 5.</b> <i>Comparação de Algoritmos de DT</i> .....	26
<b>Tabela 6.</b> <i>Matriz de Confusão</i> .....	30
<b>Tabela 7.</b> <i>10 melhores resultados para classificação - TF IDF</i> .....	33
<b>Tabela 8.</b> <i>10 melhores resultados para sentimento - TF IDF</i> .....	34
<b>Tabela 9.</b> <i>10 melhores resultados para classificação - Word2Vec</i> .....	35
<b>Tabela 10.</b> <i>10 melhores resultados para sentimento - Word2Vec</i> .....	36
<b>Tabela 11.</b> <i>10 melhores resultados para classificação - Doc2Vec</i> .....	36
<b>Tabela 12.</b> <i>10 melhores resultados para sentimento – Doc2Vec</i> .....	37
<b>Tabela 13.</b> <i>10 melhores resultados para classificação – Naive Bayes</i> .....	38
<b>Tabela 14.</b> <i>10 melhores resultados para sentimento – Naive Bayes</i> .....	39
<b>Tabela 15.</b> <i>10 melhores resultados para classificação – Decision Tree</i> .....	40
<b>Tabela 16.</b> <i>10 melhores resultados para sentimento – Decision Tree</i> .....	41
<b>Tabela 17.</b> <i>10 melhores resultados para classificação – Random Forest</i> .....	41
<b>Tabela 18.</b> <i>10 melhores resultados para sentimento – Random Forest</i> .....	42
<b>Tabela 19.</b> <i>10 melhores resultados para classificação – Adaboost</i> .....	43
<b>Tabela 20.</b> <i>10 melhores resultados para sentimento – Adaboost</i> .....	43
<b>Tabela 21.</b> <i>10 melhores resultados para classificação – KNN</i> .....	44
<b>Tabela 22.</b> <i>10 melhores resultados para sentimento – KNN</i> .....	45
<b>Tabela 23.</b> <i>Melhores resultados para classificação – LSTM</i> .....	46
<b>Tabela 24.</b> <i>Melhores resultados para sentimento – LSTM</i> .....	47
<b>Tabela 25.</b> <i>10 melhores resultados para classificação – GLOBAL</i> .....	48
<b>Tabela 26.</b> <i>10 melhores resultados para sentimento – GLOBAL</i> .....	49
<b>Tabela 27.</b> <i>Algoritmos usados, artigos vs tese</i> .....	50
<b>Tabela 28.</b> <i>Resumo outras técnicas de tratamento de texto</i> .....	52
<b>Tabela 29.</b> <i>Resumo melhores resultados - classificação</i> .....	54
<b>Tabela 30.</b> <i>Resumo melhores resultados - sentimento</i> .....	56
<b>Tabela 31.</b> <i>Matriz avaliação novos pedidos</i> .....	57

## ABREVIATURAS E SÍMBOLOS

AB	Adaboost
ADASYN	Adaptive Synthetic
AI	Artificial Intelligence
AMT	Amazon Mechanical Turk
ANN	Artificial Neural Network
API	Application Programming Interface
BERT	Bidirectional Encoder Representations For Transformers
BOA	Bag-Of-Words
C4.5	Successor Of Id3
CART	Classification And Regression Tree
CBOW	Continuous Bag-Of-Words
CHAID	Chi-Squared Automatic Interaction Detector
CNN	Convolutional Neural Network
COTA	Customer Obsession Ticket Assistant
CRM	Customer Relationship Management
CRUISE	Classification Rule With Unbiased Interaction Selection And Estimation
CSA	Customer Support Automation
CTREE	Conditional Inference Trees
CV	Cross-Validation
DBOW	Distributed Bag Of Words
DL	Deep Learning
DM	Distributed Memory
DT	Decision Tree
ETL	Extract-Transform-Load
ERP	Enterprise Resource Planning
FS	FunctionSampler
GBT	Gradient-Boosted Trees
GUIDE	Generalized, Unbiased, Interaction Detection And Estimation
IA	Inteligência Artificial
ID3	Iterative Dichotomies 3
IVA	Intelligent Virtual Agents

KNN	K-Nearest Neighbors
LIME	Local Interpretable Model-Agnostic Explanations
LSBD	LingPipe Sentence Boundary Detector
LPC	Ling Pipe Classifier
LR	Logistic Regression
LSA	Latent Semantic Analysis
LSTM	Long Short Term Memory
MARS	Multivariate Adaptive Regression Splines
ML	Machine Learning
MLP	Multilayer Perceptron
MNB	Multinomial Naïve Bayes
NB	Naive Bayes
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
PICOC	Problem, Intervention, Comparison, Outcomes And Context
POS	Partofspeech Tagging
QUEST	Quick, Unbiased And Efficient Statistical Tree
RF	Random Forest
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
ROS	RandomOverSampler
SHAP	Shapley Additive Explanations
SMO	Sequential Minimal Optimization
SMOTE	Synthetic Minority Oversampling Technique
SPAM	Sending And Posting Advertisement In Mass
SVD	Singular Value Decomposition
SVM	Support Vector Machines
TF-IDF	Term Frequency-Inverse Document Frequency
TI	Tecnologias De Informação
XGB	Extreme Gradient Boosting

## INTRODUÇÃO

Nos dias de hoje, as empresas ou entidades têm uma especial atenção à sua relação com os seus clientes, quer no processo de contato inicial, quer no processo de pós-venda. Com uma forte presença no nosso quotidiano, a correspondência eletrónica, por exemplo, os emails, e, com a WEB 2.0, as redes sociais, revestem-se de uma importância relevante para as empresas (McGuire & Kampf, 2015).

Tratar estes contatos ou opiniões, torna-se assim, uma das prioridades dessas entidades, para que lhes seja possível manter os seus clientes ou até mesmo parceiros, satisfeitos (Senthurvelautham & Hettiarachchi, 2018; Xin et al., 2015).

No entanto, para poder tratar todos estes dados, na sua maioria não estruturados, estas têm de ter recursos humanos disponíveis e capazes de analisar e tirar conclusões de toda essa informação. Como tal, estes processos têm o potencial para se tornarem dispendiosos, colocando nas fiances das empresas um peso e pressão extra (Molino et al., 2018; Xin et al., 2015).

Para colmatar ou minimizar esse custo, muitas entidades recorrem à mecanismos de ML e AI para que lhes seja possível automatizar estes processos e reduzir, assim, o seu gasto (Ebadi et al., 2019).

É aqui que entra o uso de técnicas de NLP, um campo da AI que é responsável pelo estudo da compreensão de linguagem humana por sistemas digitais, cujo objetivo, é a criação de um sistema eletrónico que permita a análise e geração de textos em linguagem humana (Ohata et al., 2022; Padmavathy & Pakkir Mohideen, 2020; Sunagar et al., 2021).

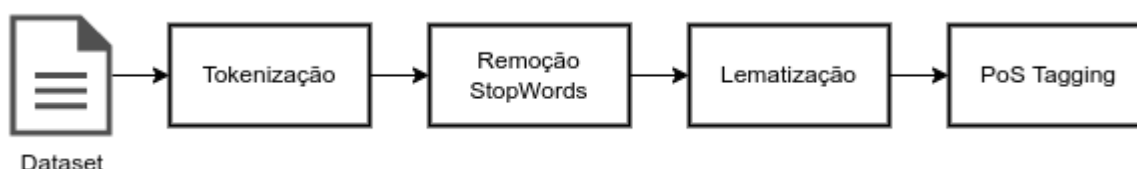
Estas ferramentas são atualmente usadas em diversas áreas e empresas, tais como sistemas de decisão medica, como auxílio no diagnóstico e tratamento de pacientes, em sistemas de CSA para registo/criação de uma base de conhecimento de iterações anteriores com os clientes para que seja possível prever a solução de uma iteração futura. São ainda usados por empresas da área legal para análises complexas e automatizadas de documentos com o objetivo de reduzir o tempo despendido e os seus respetivos custos, por empresas de telecomunicações para classificar os emails de contato com os clientes de forma a melhorar o suporte dado a estes e até por entidades publicas para gerir o contato com os seus respetivos cidadãos (Chen, 2012; Gaut et al., 2018; Parmar et al., 2018).

No caso particular das empresas que prestam serviços de apoio/suporte técnico, estas técnicas podem ser usadas para tratar automaticamente os pedidos de assistência que recebem, quer sejam problemas ou dúvidas, abertura e elevação de tickets, resolução automatizada de problemas, encaminhamento de pedidos para o seu destino, entre outras (Gupta, 2020; Ricciardi Celsi et al., 2021).

O NPL, tradicionalmente, serve para realizar uma série de passos que permitem tratar os dados de forma que estes possam ser interpretados por sistemas digitais. Estes são a detecção do idioma, correção de palavras com erros ortográficos, pré-processamento dos dados e extração de atributos. No final destes passos, passa a ser possível, com recurso a algoritmos de IA, fazer a classificação dos textos, de forma a categorizá-los, assim como, realizar uma análise do sentido e emoção presente no texto, vulgo Análise de Sentimento, para que seja possível determinar o tom da mensagem, por exemplo, positivo, negativo ou neutro (Moharil et al., 2020; Ohata et al., 2022; PhD Research Student, MUIT, Lucknow. et al., 2019; Ranganathan & Tzacheva, 2019).

### Pré-Processamento

O pré-processamento é um passo que é realizado para tratar os dados antes de estes poderem ser analisados pelos algoritmos de AI. É aqui que vão ser analisadas as entradas inválidas ou nulas, a formatação dos campos, por exemplo as datas ou números de telefone ou do código postal, para que possam estar todos com o mesmo formato.



**Figura 1.** Pré-processamento de Texto

Semelhante ao processo de ETL, conforme a figura 1, pode englobar vários passos:

- **Tokenização:** Processo que transforma as frases dos textos em análise em listas de palavras, de forma a facilitar a sua classificação (Moharil et al., 2020).
- **Remoção de Stopwords:** Neste passo são removidas as palavras de ligação, tais como, a, e, de, se, e as palavras sem valor semântico (Sunagar et al., 2021).
- **Stemming:** Processo que remove os últimos caracteres de uma palavra para obter para que a palavra fique numa forma mais pequena, por exemplo, carros fica

carro. É usado em casos em que há uma limitação da capacidade e tempo de processamento, devido a ser mais leve do que lematização (Ahmad & Laroche, 2017).

- Lematização: Este termo refere-se ao processo que determina a forma canônica das palavras. Permite que as palavras possam ser agrupadas pelo seu lema, a palavra que está presente nos dicionários ou vocabulários. É uma evolução do passo anterior, na medida em que tem em conta o contexto para obter a forma base da palavra. O seu processo é mais demorado que o anterior (Moharil et al., 2020).
- POS: Neste passo, é determinado a classe gramatical das palavras, ou seja, é determinado se uma palavra é um nome, pronome, verbo, adjetivos, entre outros (Pugsee et al., 2019).

Findo este processo, os dados em análise estão prontos para serem utilizados na fase de extração de atributos.

### **Extração de Atributos**

A extração de atributos é o processo que permite a redução de dimensão da amostra em análise para possa ser mais fácil o seu processamento. Um método para reduzir a dimensão em NPL é o BOA. Este processo extrai as palavras (atributos) das frases e classifica-as pela sua frequência nos textos (Parmar et al., 2018). A classificação das palavras pode ser realizada de diversas formas:

- TF-IDF: Aqui é atribuído um valor às palavras de acordo com a sua frequência nos textos em análise (Nadaf & Patil, 2015; Parmar et al., 2018). Como anteriormente foram removidos os termos sem valor semântico, este processo acaba por usar apenas as palavras mais relevantes, em vez de ter de analisar todo o texto.
  - TF mostra a importância de um termo, tendo em conta as vezes que este aparece num texto (Kaewnoo & Senivongse, 2019).

$$TF(m) = \frac{\text{Frequência do termo } m \text{ no texto}}{\text{Total de termos no texto}}$$

- IDF apresenta a relevância de um termo no texto, ignorando as *stopwords* e penalizando as palavras que mais vezes estão presentes em todos os textos em análise (Kaewnoo & Senivongse, 2019).

$$IDF(m) = \ln \left( \frac{\text{Total de documentos}}{\text{N}^\circ \text{ de documentos com } m} \right)$$

- *Word Embeddings* ou *Word Vectors*: É uma técnica que pretende extrair a identificação de informação semântica e contextual de um texto. Esta é feita com recurso a criação de vetores numéricos, em que as palavras com um significado aproximado têm um vetor semelhante.

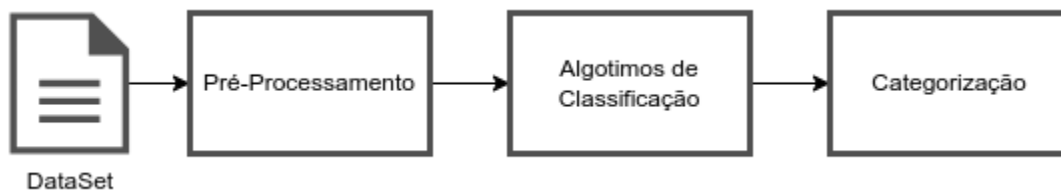
### Classificação de texto

A classificação textual refere-se ao processo pelo qual grandes quantidades de texto, tradicionalmente não estruturados, são organizados em diversas categorias pré-definidas. Por outras palavras, é o método de atribuir rótulos à informação em análise. Tem uma quantidade razoável de aplicabilidade, nomeadamente, a deteção de correio eletrónico não solicitado, vulgo SPAM, nas plataformas de comércio eletrónico ou de redes sociais para classificar comentários, em sistemas de CRM para analisar e distribuir tarefas, entre outros (Etaiwi & Naymat, 2017; Parmar et al., 2018).

Estes processos podem ser realizados por pessoas em vez de máquinas, mas a sua eficácia e objetividade em que os textos são analisados, estão dependentes da pessoa que os analisa e podem variar de pessoa para pessoa (Ebadi et al., 2019).

Como cada vez mais se privilegia o contato via canais digitais, a quantidade de informação que as empresas/entidades dispõe para ser analisada é enorme. Como tal, executar o processo manual pode não ser eficaz, na medida em que se torna dispendioso e pode acabar por não resultar em informação importante em tempo útil (Ebadi et al., 2019; McGuire & Kampf, 2015; Nadaf & Patil, 2015).

Desta forma, a utilização de mecanismos automatizados poder ser uma vantagem importante e por vezes, essencial para as empresas. Quanto mais rápido for possível obter informação relevante para o processo de tomada de decisão, melhor as empresas podem planear o seu dia-a-dia e resolver os seus problemas e ultrapassar os obstáculos (Ebadi et al., 2019).



**Figura 2.** Processo de Classificação de Texto

A classificação de texto automatizada por técnicas de AI, conforme apresentada na figura 2, tem várias fases, determinar os dados iniciais, pré-processar os dados, classificar os dados com recurso a um ou mais algoritmos e por fim, fazer a sua categorização (Ohata et al., 2022; Parmar et al., 2018; Sunagar et al., 2021).

### **Análise de sentimento**

A análise de sentimento é uma formas em que se pode aplicar a classificação de texto. Esta análise recorre a técnicas de NLP para determinar as emoções ou sentimentos do autor do documento em relação a um determinado assunto ou em relação ao texto completo (PhD Research Student, MUIT, Lucknow. et al., 2019; Ranganathan & Tzacheva, 2019).

É maioritariamente usada em análise de comentários nas redes sociais relacionados com o lançamento de um produto, por exemplo, em questionário de satisfação de clientes para medir o desempenho do relacionamento da empresa com o este, pesquisa de mercado para saber a sua ou outra reputação, entre outras (PhD Research Student, MUIT, Lucknow. et al., 2019; Pugsee et al., 2019).

Dá às entidades uma noção de como os seus clientes, atuais ou potenciais, a vêem. Permite que se tomem medidas de acordo com os resultados, para melhorar ou reparar a perceção destes de forma a melhorar o seu desempenho.

### **Problema e questões de investigação**

Atualmente os sistemas de apoio/assistência técnica funcionam maioritariamente de forma manual, com a gestão dos pedidos que chegam às empresas/entidades a serem tratados pelos funcionários.

Isto é um processo algo moroso e dispendioso pode ser propenso a erros, como a catalogação de um determinado problema de forma incorreta ou o esquecimento de registar algum dos pedidos que chegam.

Com a enorme quantidade de dados de comunicações eletrónicas, nomeadamente, e-mail, geradas pelas organizações, assim como, a evolução de sistema de AI, torna-se atrativo a implementação de sistemas que possam analisar de forma automatizada estes dados, de forma a poder otimizar o processo de comunicação com os clientes, quer na categorização dos textos, quer na análise do seu sentimento.

Este trabalho visa investigar se as comunicações eletrônicas, vulgo emails, recebido por alguma entidade com contexto do relacionamento com os seus clientes, podem ser classificados automaticamente e avaliados de acordo com o seu sentimento. Estes processos serão realizados de acordo com dados históricos dessa mesma entidade.

Considerando estes pressupostos procurou-se compreender o estado da arte relativamente, a como é que a investigação tem sido desenvolvida nomeadamente:

- Qual é o estado da pesquisa atual sobre sistemas de apoio técnico inteligentes? Esta questão tem como objetivo identificar os estudos publicados nesta temática. Desta forma, será possível perceber quais os sistemas que já foram estudados e perceber o seu funcionamento
- Quais são os algoritmos que são aplicados em sistemas de suporte técnico para classificação de texto e análise de sentimento? Para analisarmos os algoritmos usados em sistemas de apoio técnico para perceber como estes foram usados e configurados e obter uma perspetiva da evolução dos próprios algoritmos e de quais são os mais usados neste campo
- De que forma são utilizados os algoritmos para classificar texto e para análise de sentimento? Para avaliarmos como é que os algoritmos estão a ser implementados e configurados.
- Como são integrados estes algoritmos nos sistemas de *helpdesk*? Onde pretendemos perceber como é que estão a ser implementados.
- Como é medida a sua eficácia e eficiência no desempenho dos vários algoritmos utilizados? Para ser possível compreendermos métricas utilizadas na sua avaliação

### **Objetivos, geral e específicos**

Este trabalho tem como objetivo estudar a classificação de texto e análise de sentimento usada em contexto de sistemas de assistência técnica inteligentes. Com base neste estudo, pretende-se criar uma abordagem que possibilite a classificação de texto de forma automática e a realização de análise de sentimento de forma a ser possível criara uma fila de prioridades. Esta fila será criada com o tipo de classificação obtida e através do grau de sentimento obtido.

## **Estrutura do relatório**

Este trabalho está organizado da seguinte forma: no capítulo seguinte está presente a revisão bibliográfica da temática em análise, onde são dadas respostas às questões de pesquisa identificadas e está dividida em vários subcapítulos nos quais são descritos a metodologia de pesquisa e as respostas às perguntas de pesquisas identificadas anteriormente.

## REVISÃO DA LITERATURA

Para responder às questões de investigação levantadas sobre o estado da arte relativamente a sistemas de apoio técnico inteligentes. Realizamos uma revisão sistemática para responder às questões de investigação e problemas identificadas. Nesta seção descrevemos qual é o estado atual da pesquisa sobre os sistemas de assistência técnica inteligentes depois abordamos quais são os algoritmos usados para a classificação e determinação do sentimento, a sua utilização e integração. A sua avaliação será realizada através da matriz de confusão.

Esta revisão foi desenvolvida em três fases, a primeira define a questão principal a ser respondida por este trabalho. A segunda define o processo pelo qual, serão procurados e selecionados os trabalhos a fazerem parte da revisão bibliográfica. O terceiro e último passo, passa por dar resposta às perguntas de pesquisa identificadas anteriormente, com base nos trabalhos selecionados no ponto anterior.

A estratégia de pesquisa foi iniciada através da técnica PICOC (Petticrew & Roberts, 2006), população, intervenção, comparação, resultados (*outcomes*) e contexto, para ser possível definir as *keywords* de pesquisa.

**Tabela 1.** *Critérios PICOC*

PICOC	Descrição
População	Publicações referentes a sistemas inteligentes de classificação de emails com análise de sentimento
Intervenção	Algoritmos de aprendizagem máquina para classificação de texto e análise de sentimento
Comparação	Diversos algoritmos para classificação de texto e análise de sentimento
Resultado	Procura de publicações que estudem a temática dos sistemas e mecanismos de apoio técnico/helpdesk, de forma a identificar os algoritmos de inteligência máquina usados como complemento a estes sistemas, possibilidades de melhoria e melhores práticas
Contexto	Publicações académicas

As palavras-chave identificadas foram *helpdesk*, *ticketing*, *customer care*, *Technical Support*, *Machine Learning*, *artificial intelligence* e *sentiment analysis*.

De acordo com os dados anteriores, o texto da pesquisa ficou o seguinte, (helpdesk OR ticketing OR customer care OR Technical Support) AND (“Machine Learning” OR

“artificial intelligence” OR “sentiment analysis”), o qual foi aplicado à pesquisa de títulos, resumo e keywords. O período de pesquisa foi definido de 2010 a 2022.

A pesquisa foi realizada a base de dados SCOPUS no dia 07/12/2022, e foram identificados 147 artigos.

Os filtros de inclusões e exclusões aplicados na pesquisa foram os seguintes:

1. Inclusão:

- a) De 2010 a 2022
- b) Apenas artigos e relatórios de conferências
- c) Apenas as áreas de COMPUTER SCIENCE e ENGINEERING

2. Exclusão:

- a) Conference Review
- b) Book Chapter
- c) Review
- d) Book
- e) Short Survey
- f) NON-ENGLISH ARTICLES

Após a obtenção dos artigos a analisar, um foi removido devido a não estar escrito em inglês e outro por não ter sido possível obtê-lo.

Para auxiliar na filtragem dos artigos selecionados, foi usada uma ferramenta de inteligência artificial que consegue estimar a probabilidade de um documento poder ser incluído a realizar a ordenação dos artigos do mais para o menos relevante. Desta forma, o autor pode identificar antecipadamente os artigos a incluir. Desta forma, foi usado o software *AsReview* para realizar esta tarefa. Foi selecionado o classificador NB e a extração de atributos foi feita com recurso TD-IDF. Os passos realizados por este foram os seguintes:

1. Seleção de cinco artigos relevantes de cinco irrelevantes.
2. Como a apresentação dos artigos vai do mais relevante para o menos relevantes, torna o processo de análise dos resumos mais fácil e célere.
3. Após a revisão dos resumos dos artigos, foram excluídos 119 artigos.

No fim deste processo de seleção de artigos, ficaram 27 para analisar e usar neste trabalho.

### **Qual é o estado da pesquisa atual sobre sistemas de apoio técnico inteligentes?**

Esta pergunta de pesquisa pretende perceber em qual o estado da arte em relação a sistemas de apoio técnico inteligentes. De todos os artigos analisados, estes sistemas são maioritariamente propostos para empresas de tecnologias de informação, para auxiliar a resolução de pedidos de assistência (tickets) (Jordán et al., 2013; Kaewnoo & Senivongse, 2019; Molino et al., 2018; Ohata et al., 2022), fazer a sua classificação automaticamente e prever se ou quando estes serão escalados (Gupta, 2020). São também usados assistentes inteligentes (*chatbots*) (Nayak et al., 2021) que, com base em interações com clientes passadas, possam resolver problemas ou dar sugestões para a resolução.

Foi ainda documentado, duas soluções automatizadas de sistemas de CRM que usam técnicas de AI para auxiliar o relacionamento com os clientes, IVA e *Pegasystems*. (Chen, 2012) O primeiro é um assistente virtual, que oferece soluções personalizadas na interação com o cliente. O segundo é um sistema semelhante ao anterior, mas em vez de devolver respostas em tempo real, usa a análise dos clientes para facilitar o funcionário na sua interação com o cliente, fornecendo-lhe informação e sugestões.

Noutra vertente, existem soluções que capturam as comunicações de voz entre os utilizadores e os técnicos de forma a ser possível fazer o reconhecimento do texto e registá-lo para uso futuro. (Lohse et al., 2015) Desta maneira, na próxima interação com o cliente, serão apresentadas soluções de acordo com o problema reportado.

A análise de métricas sobre todo o processo de interação com o cliente também foi documentada. Esta análise permite que se possa otimizar os recursos usados para a interação com os clientes. (Xin et al., 2015)

Por fim, é ainda analisado um sistema de governamental que permite a automatização do relacionamento digital dos cidadãos com o governo regional. (Gaut et al., 2018)

### **Quais são os algoritmos que são aplicados em sistemas de suporte técnico para classificação de texto e análise de sentimento?**

Após analisar os trabalhos referentes a sistemas inteligentes de apoio ao cliente, foram analisados quais os algoritmos usados para a classificação de texto e análise de sentimento.

Foram identificados treze diferentes tipos de algoritmos. O mais popular, aplicado em cinco artigos, foi RF (Gaut et al., 2018; Kaewnoo & Senivongse, 2019; Lohse et al., 2015; Molino et al., 2018; Ohata et al., 2022), seguido de LR (Gaut et al., 2018; Kaewnoo & Senivongse, 2019), NB (Kaewnoo & Senivongse, 2019; Nayak et al., 2021; Ohata et al., 2022) e SVM (Nayak et al., 2021; Ohata et al., 2022), todos usados em dois trabalhos. Usado apenas uma vez, temos *LASSO Regression* (Xin et al., 2015), *probabilistic topic modeling* (Lohse et al., 2015), *Encoder-Combiner-Decoder* (Molino et al., 2018), Linear SVC (Kaewnoo & Senivongse, 2019), ensemble decision tree models - bagging/boosting (Nayak et al., 2021), MLP, XGB e KNN (Ohata et al., 2022).

### **De que forma são utilizados os algoritmos para classificar texto e para análise de sentimento?**

O objetivo desta pergunta prende-se com o uso dos algoritmos no processo de classificação de texto e análise de sentimento de uma forma genérica. Mais em concreto, os dados que são usados e a forma como estes são tratados, como são configurados os algoritmos e como estes usam os dados.

A revisão está dividida em duas partes, a primeira analisa os artigos referentes a classificação de texto. A segunda faz o mesmo para os artigos de análise de sentimento. Os artigos analisados estão ordenados do mais antigo para o mais recente.

#### ***Classificação de Texto***

Lohse et al. (2015) faz uma análise de como a interação entre os técnicos de *helpdesk* e os clientes pode ser capturada, transformada em texto e registada para uso futuro. Com recurso a técnicas de NPL, as chamadas de voz podem ser transformadas em texto para serem posteriormente analisadas por classificadores. Neste caso o classificador usado foi RF, o qual, depois de ser analisado o pedido do cliente, é sugerida uma possível solução ao técnico.

Nadaf & Patil (2015) analisam a classificação de texto, tendo como base um *dataset* composto por pedidos de assistência técnica de forma a facilitar a resolução destes. O *dataset* foi obtido do site data.gov e é constituído por dez variáveis e tem 2, 64 e 545 observações. Foram usados três algoritmos de *clustering*, *K-Means*, *Dirichlet* e *Fuzzy K-means*, com valores de k diferentes, 10, 20, 30, 40 e 50. O algoritmo *Dirichlet* foi o que mais tempo demorou a formar os clusters, com o *Fuzzy K-means* a demorar menos tempo. A vectorização

do *dataset* é feito através de TF-IDF e é usada a métrica da soma do erro quadrado para avaliar os clusters, com a conclusão de que quanto maior o valor de *k*, maior o erro.

Etaiwi & Naymat (2017) analisam a detecção automática de opiniões no site *TripAdvisor* para os vinte hotéis mais populares em Chicago, de forma a poder diferenciar opiniões falsas. A recolha de dados foi realizada através da AMT. Foram recolhidas 40 opiniões reais e 40 falsas por cada um dos hotéis, num total de 800 reais e 800 falsas. O pré-processamento passou pelo uso de BOW com POS para criar os vetores de palavras. Foram realizados diversos testes, vectorização sem mais nenhum processamento, com *stemming*, com remoção de pontuação, remoção de *stopwords* e com todos os anteriores. O resultado para os algoritmos usados, GBT, NB, RF, DT, SVM, foi para a *precision*, o RF teve melhores resultados para qualquer um dos métodos. Para a *accuracy* e *recall* foi NB que teve melhor desempenho para todos os métodos exceto o último, em que ficou em igualdade com SVM.

Molino et al. (2018) apresenta uma solução para a melhoria do apoio aos clientes, COTA, nas versões 1 e 2. Os dados eram compostos por pedidos de suporte da Uber, com aproximadamente 3 milhões de observações. Ambas realizam o pré-processamento da mesma forma, tokenização, transformação para minúsculas, remoção de *stopwords* e *lemmatization*, seguido do uso de TF-IDF e LSA para atribuir peso às palavras. A v1 usa RF e a v2 usa tanto CNN, como RNN. Os resultados foram 45,5% e 53,12, respetivamente.

Gaut et al. (2018) tenta melhorar um sistema de comunicação com o público do governo mexicano. Os dados analisados referem-se aos pedidos submetidos na plataforma, num total de 69402 registos. A extração de atributos foi realizada através de TF-IDF ou tokenização em *unigrams* ou *unigrams+bigrams* ou através de *stemming*. Os dados foram analisados por RF e LR, com a primeira a ter melhor resultados.

Parmar et al. (2018) analisa a classificação de pedidos de assistência técnica de forma a poder ser feito o seu envio para o departamento correto. Os dados continham 5500 mensagens recebidas pelo suporte. A extração de atributos é realizada através de TF-IDF. O desempenho dos algoritmos usados foi o seguinte, MNB 41,69%, SVM 60,76%, DT 49,53%, RF 39,85% e KNN 52,19.

Ebadi et al. (2019) testa o uso de cinco ferramentas de ML, H2O AutoML, *auto\_ml*, TPOT, SAP Predictive Analytics software e Auto-sklearn, contra uma solução desenvolvida manualmente. Foram testados com sete problemas de ML, três de classificação e quatro de regressão. Os dados foram obtidos da plataforma Kaggle e têm entre 8 e 1933 atributos e

entre 506 e 1058297 observações. Os modelos gerados pela H2O tiveram melhor performance que todos os outros para seis dos sete problemas e `auto_ml` foi a única a conseguir gerar previsões para todos os problemas. No entanto, a solução manual, desenvolvida por um técnico, foi a que obteve melhores resultados. O autor refere ainda, que no caso das empresas que não têm meios para desenvolver soluções à medida, podem olhar para estas ferramentas com uma alternativa.

Kaewnoo & Senivongse (2019) analisa a classificação de texto para tentar otimizar a resolução de problemas com a manutenção de software, no contexto de pós-venda. Os dados usados provêm do sistema de gestão de pedidos de assistência e contêm 1884 observações, divididas por 11 categorias. O pré-processamento foi realizado através da remoção de caracteres especiais, tokenização, remoção de *stopwords* e *stemming*. A extração de atributos foi realizada através de TF-IDF. Para normalizar os dados, foi usada a técnica de SMOTE para preencher as classes com menos dados. Os algoritmos usados foram MNB, Linear SVC (SVM com *linear kernel*), RF e LR. Os resultados foram os seguintes, Linear SVC teve melhor desempenho em todos os testes, independente das técnicas usadas (*unigram + bigram*, *unigram + trigram*, entre outras), com uma percentagem entre 91,69% e 74,51%, para *f1-score* e *accuracy*, respetivamente.

Pandey et al. (2020) propõe um modelo híbrido para a deteção de e-mail de *phishing*. Esse modelo é composto pelos algoritmos RF e SVM. O primeiro treina o *dataset*, dividindo-o em várias unidades. Estas seriam depois reclassificadas pelo SVM e os resultados serão fundidos através do uso de *bagging*. Os dados usados provêm do repositório da Universidade da Califórnia, e contêm 10 atributos com 1353 observações. Os resultados obtidos mostram que o modelo proposto é mais eficaz que SVM e RF, com a média de *f1-score* a ser 91,75%, 70% e 88,5%, respetivamente para cada um dos algoritmos (modelo híbrido, SVM e RF).

Nayak et al. (2021) apresenta a um *chatbot* para auxiliar na resolução de problemas de assistência técnica e auxiliar no reencaminhamento de pedidos. Faz uso de base de dados de conhecimento de pedidos anteriores para sugerir ações futuras. Os dados antes de serem gravados na base de dados são tratados através de *n-gram*, TF-IDF, POS, entre outras. Os dados são posteriormente classificados através dos algoritmos NB, SVM, e DT com *bagging/boosting*. No caso de o resultado ser superior a 90%, é automaticamente gerado um registo no sistema, caso contrário, é marcada para análise mais detalhada.

Ricciardi Celsi et al. (2021) apresenta uma solução para a previsão de reabertura de pedidos de assistência técnica relacionado com redes de fibra ótica. Os dados utilizados têm 307 atributos com 600000 observações. Foram recolhidos durante 30 dias a 600000 utilizadores. Os algoritmos usados para realizar a análise foram DT, RF, boosting, LR, ANN e SVM. Os resultados para os vários algoritmos são 69% e 79%, 75% e 80%, 71% e 79%, 70% e 81%, 66% e 67% e 73% e 69%, relativamente á métrica *accuracy*, para o *dataset* original e o reduzido, respetivamente.

Sunagar et al. (2021) propõem uma metodologia para a classificação de artigos de notícias de acordo com rótulos pré-definidos. O *dataset* usado foi AG's News Topic Classification Dataset. Este tem quatro classes, com 30000 observações de treino e 1900 de teste, cada. Para o pré-processamento e extração de atributos, foi realizada a tokenização do texto, remoção de *stopwords*, *stemming* e TF-IDF, respetivamente. Os algoritmos usados, NB, SVM, KNN, Rocchio, Boosting e Bagging, tiveram a seguinte classificação para *f1-score*, 87%, 91%, 81%, 72%, 61%, 83%, respetivamente.

Ohata et al. (2022) pretende desenvolver um assistente para ser usado como complemento numa empresa de fornece serviços de assistência técnica, de forma que os problemas mais comuns possam ser resolvidos com o envio de mensagens automáticas. Os dados usados referem-se a 1049 problemas reportados, labelados, aos quais foram atribuídos 9 *templates* de resposta. O pré-processamento foi realizado através da deteção da língua (o estudo apenas considera textos em Inglês), tokenização, conversão para letras minúsculas, *lemming*, correção ortográfica, remoção da pontuação e de caracteres especiais, espaços em branco e *stopwords*. Para a extração de atributos foi usado a técnica TF-IDF com a limitação a *unigrams*, *bigrams* ou *trigrams*, assim como, Word2Vec, Doc2Vec e BERT. Além disso, palavras que aparecem em menos de 5 documentos foram removidas. Os algoritmos usados foram SVM, RF, MLP, NB, XGB e KNN. O melhor desempenho foi observado com o algoritmo RF, com uma combinação de TF-IDF com meta dados adicionais, com um score de 69,2% para f1-score.

### ***Análise de Sentimento***

Laryea et al. (2015) analisa o uso de algoritmos para extrair emoções de opiniões de produtos eletrónicos presentes em Amazon.com. O sistema proposto, recolhe os dados do site referido, armazena numa base de dados, faz a extração das frases, executa POS, desenvolve o modelo, classifica-o e apresenta os dados numa plataforma Web. Os dados

recolhidos são compostos por 43227 observações repartidos por cinco classes. As frases foram extraídas através do LSBD. O classificador usado para construir o modelo foi LPC, o qual teve um resultado de 70%. Os resultados são apresentados ao utilizador através de um *dashboard*, visualizado através de um navegador WEB.

Ahmad & Laroche (2017) analisa como podem ser analisadas as opiniões sobre produtos online, no contexto da evolução da Web 2.0 e como esta veio alterar os hábitos de compra dos consumidores. Os dados analisados provêm da Amazon.com e referem-se a opiniões sobre eletrodomésticos para cozinhas. Contêm 15701 opiniões, no entanto, foram selecionadas aleatoriamente 147 opiniões negativas e 258 positivas para a análise. O tratamento dos dados passou pela criação de uma matriz com as palavras que compõem as frases e a sua respetiva frequência. O pré-processamento passou pela remoção de palavras únicas e *stopwords* através de *stemming*. A extração de atributos foi realizada através de TF-IDF. De seguida foi executado o SVD. Este processo de decomposição da amostra produz três matrizes, *term eigenvectors*, *document eigenvectors* e *square roots of eigenvalues (singular values)*. Por fim, foi executada uma regressão para prever a quais os fatores os clientes dão mais importância. Para as opiniões negativas, os fatores que têm mais importância são o processo de compra online, erros no envio da encomenda e portes de envio. Para as opiniões positivas, os fatores com maior importância são *branding*, afetivo, aspetos técnicos, online, estética e serviço.

Ceyhan et al. (2017) estuda a opinião de utilizadores de serviços de saúde na Turquia durante sete anos num hospital privado. Os dados totalizam 2018 comentários positivos e 1394 negativos. Estes foram divididos em quatro *datasets*, cada um com 450 comentários de treino e 50 de teste. Na fase do pré-processamento foi realizada uma análise morfológica dos dados e desambiguação dos dados. A extração de atributos foi realizada através do cálculo da sua frequência nos textos. Para a classificação foram usados os algoritmos NB, SMO e DT J48. Os resultados obtidos por cada um são 89,4%, 95,8% e 91,2%, respetivamente.

PhD Research Student, MUIT, Lucknow. et al. (2019), (Pugsee et al., 2019) e Ranganathan & Tzacheva (2019) analisam dados de *posts* no Twitter de forma a determinar o sentimento destes. Os dados utilizados nestes estudos provêm todos do *twitter* API e referem-se a *posts* sobre filmes, produtos de beleza e assuntos genéricos. Na fase de pré-processamento todos eles necessitaram de realizar a remoção do jargão, pontuação, símbolos e outros caracteres especiais. De seguida foi aplicado a um dos artigos a remoção de *stopwords* e a outro POS. No terceiro, foi aplicado o *stemming*. No segundo e terceiro artigo,

a polaridade das palavras é realizada através de SentiWordNet e do filtro TweetToSparseFeatureVector presente em Weka Affective Datasets. Os resultados obtidos para o primeiro artigo são 75% para KNN, 78,6% para SVM e 82% para o algoritmo proposto no artigo. No segundo, os resultados são 82% para NB e 86% para SVM. No terceiro e último, os resultados são 98% para o classificador SVM LibLinear.

### **Como são integrados estes algoritmos nos sistemas de *helpdesk*?**

Dos trabalhos analisados, os algoritmos de AI são integrados, na sua maioria, nos sistemas já existentes. Podem ser integrados entre o técnico que serve de primeiro contato com o cliente e a segunda linha de apoio (Lohse et al., 2015).

Outros são colocados na camada de negócio, de forma que possam analisar os pedidos que chegam diretamente, para proceder à sua análise e respetiva classificação ou encaminhamento (Molino et al., 2018; Nayak et al., 2021; Senthurvelautham & Hettiarachchi, 2018; Xin et al., 2015).

Um dos artigos refere a testagem do sistema proposto, antes de o colocar em produção (Gaut et al., 2018). Neste caso, à semelhança dos anteriores, a integração foi feita nos sistemas atuais, com a responsabilidade de fazer a análise dos pedidos, a sua aceitação e posterior encaminhamento.

Por último, um dos sistemas proposto passou pelo desenvolvimento de uma plataforma Web para a introdução de pedidos de assistência (Kaewnoo & Senivongse, 2019). O sistema ao receber uma solicitação, faz a classificação deste autenticamente. Noutra secção da página, é apresentado um gráfico com o resumo das categorizações dos pedidos.

### **Como é medida a sua eficácia e eficiência?**

Uma das mais importantes fases do estudo de algoritmos de AI, é a forma de avaliação de desempenho. São estas métricas que nos permitem avaliar a validade dos modelos criados e a sua eficácia.

Nos trabalhos analisados, foram aplicadas várias métricas de avaliação de desempenho Tabela 2, sendo as mais populares, *accuracy*, *precision*, *recall* e *f-measure (f1-score)*.

**Tabela 2.** *Resumo Métricas*

Artigo	Accuracy	Precision	Recall	Specificity	Sensitivity	F-Measure	ROC
--------	----------	-----------	--------	-------------	-------------	-----------	-----

(Ceyhan et al., 2017)	X						
(Etaiwi & Naymat, 2017)	X	X	X				
(Molino et al., 2018)	X					X	
(Gaut et al., 2018)		X	X				
(Parmar et al., 2018)	X	X	X			X	
(Senthurvelautham & Hettiarachchi, 2018)		X	X			X	
(Ranganathan & Tzacheva, 2019)		X	X			X	
(Kaewnoo & Senivongse, 2019)	X	X	X			X	
(Pugsee et al., 2019)	X	X	X				
(PhD Research Student, MUIT, Lucknow. et al., 2019)	X	X	X			X	
(Padmavathy & Pakkir Mohideen, 2020)		X	X			X	
(Pandey et al., 2020)	X	X		X	X	X	
(Sunagar et al., 2021)	X	X	X			X	X
(Ricciardi Celsi et al., 2021)	X						
(Ohata et al., 2022)	X					X	

Outras métricas usadas para a avaliação de desempenho foram as seguintes, Sum of Squared Error (Nadaf & Patil, 2015), Classification error, Mean absolute error e Mean square error (Pandey et al., 2020), Gini Coefficient e Youden Index (Ricciardi Celsi et al., 2021), Friedman teste sobre *accuracy*, com  $\alpha = 0.05$  e Nemenyi test (Ohata et al., 2022).

Foram ainda usadas duas ferramentas para explicar os resultados dos algoritmos, SHAP (Gupta, 2020) e LIME (Ohata et al., 2022).

## **Conclusão**

No decorrer da revisão bibliográfica realizada, foi detetada uma lacuna nos procedimentos realizados. Ao realizar a classificação de texto em contexto de serviços de apoio/assistência técnica, poderia ser interessante a realização da análise de sentimento.

Desta forma, as empresas/entidades poderiam determinar, em conjunção com a importância do cliente, uma fila de prioridades. Assim, um cliente que tivesse uma importância mais elevada, como por exemplo, um cliente que tivesse importância estratégica significativa, e com uma mensagem negativa, teria uma prioridade maior. Pelo contrário, um cliente com um peso mais baixo e com uma mensagem neutra, teria uma prioridade menor.

Como a literatura nos indica, a quantidade de correspondência eletrónica gerada e armazenada pelas diversas entidades é enorme, pelo que, dados para treinar os algoritmos e alimentar os modelos por eles gerados, é algo que existe em abundância. No entanto, para que estes possam ser analisados, têm de ser colocados num suporte que permita e facilite o acesso, tratamento, assim como a sua manipulação.

Só assim se pode planear o processo de AI, para a classificação e análise sentimental. Com o tratamento dos dados, o pré-processamento, a extração de atributos, a sua classificação e a atribuição de valores e respetiva avaliação.

## **METODOLOGIA**

A metodologia usada foi realizada em diversas fases. Principalmente foi realizada a definição das tecnologias usadas, a arquitetura e o planeamento do sistema. De seguida, foi obtido e tratado o conjunto de dados, o seu tratamento, através do processo de pré-processamento e a extração de atributos. Após este processo foram aplicados diversos algoritmos de ML, em conjunto de várias técnicas de *oversampling* e por fim, foi realizada a avaliação dos resultados dos modelos implementados.

### **Tecnologias a usar**

As tecnologias a usar neste processo de testes de avaliação de algoritmos de AI para a classificação de texto e análise de sentimento são os seguintes:

- Linguagem de programação: Python v3.10.
- Bibliotecas de IA: scikit-learn, pandas, plotly, tensorflow, nltk, keras, gesim.

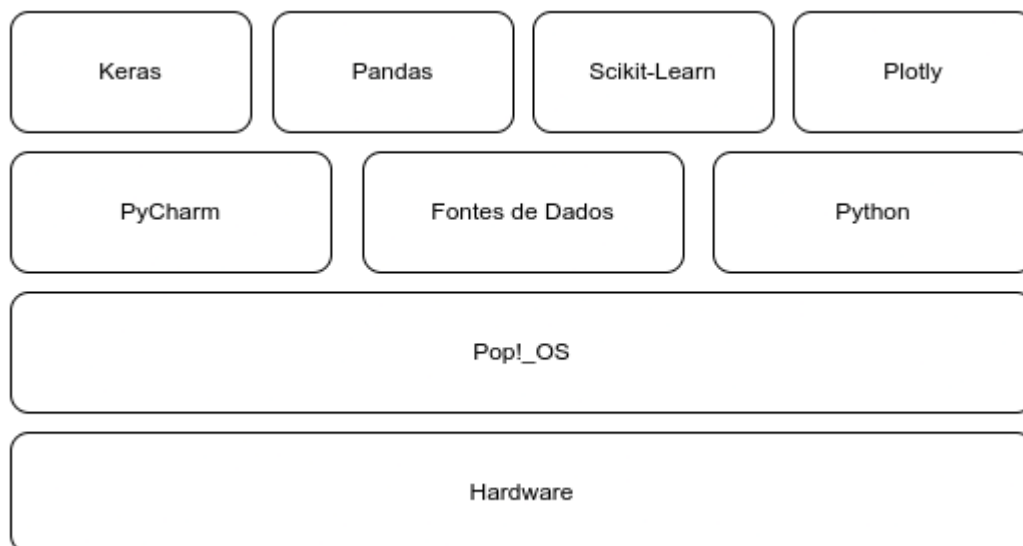
- IDE: JetBrains PyCharm.

Os teste serão realizados num computador com o sistema operativo Pop!\_OS versão 22.04 LTS, baseado em Linux Ubuntu, que por sua vez, é baseado em Linux Debian. O hardware é o seguinte:

- CPU: Intel i7-7700HQ.
- GPU: Nvidia GeForce GTX 1050Ti Mobile.
- Memória: 32GB

### Arquitetura

A arquitetura a usar neste processo é a que é apresentada na imagem seguinte.



**Figura 3.** *Arquitetura*

### Planeamento

Para poder testar as aplicações dos algoritmos, pretende-se avaliar o seu desempenho através da aplicação de diversas técnicas de pré-processamento, nomeadamente, limpeza do texto através da remoção de caracteres especiais, espaços em branco e quebras de linha, assim como, a conversão para letras minúsculas. Por fim, será realizada a tokenização e remoção de *stopwords*.

Para a extração de atributos, pretende-se avaliar o desempenho de vários da geração de vários BOW distintos. As técnicas a usar serão as seguintes: TF-IDF e *Word Embeddings* ou *Word Vectors*, mais concretamente, *Word2Vec* e *Doc2Vec*.

De seguida, para a criação do modelo de classificação, serão considerados os algoritmos da família de NB, nomeadamente, GaussianNB, RF, DT e KNN. Relativamente a algoritmos de DL, pretende-se a implementação de LSTM.

Para a divisão dos dados, pretende-se testar o corte em 70/30, com e sem a utilização de K-Fold Cross-Validation, com o número de k igual a 5.

A validação dos resultados, será realizada através das medidas estatísticas, *accuracy*, *precision*, *recall* e *f-measure*, vulgo matriz de confusão.

### **Análise de Risco**

De forma a ser possível antecipar alguma eventualidade com o decorrer do desenvolvimento e implementação do projeto, executou-se uma análise de risco. Nela estão identificados os potenciais problemas, a importância de cada um deles, a sua probabilidade da ocorrência, o seu impacto e um plano de contingência.

**Tabela 3. Tabela de Riscos**

Risco	Categoria	Probabilidade	Impacto	Plano Contingência
Evolução tecnológica com impacto nos módulos desenvolvidos	DE	20%	4	Assegurar que o desenvolvimento é executado de forma genérica
Bibliotecas utilizadas não servem o propósito	DE	50%	3	Pesquisa sobre as bibliotecas a usar, ver exemplos
Tempo de desenvolvimento insuficiente	PS	50%	3	Otimização do calendário
Dificuldade para executar a implementação	ST	70%	2	Pesquisa sobre as tecnologias usadas para a implementação
Falta de conhecimentos tecnológicos	ST	30%	1	Recorrer a pessoas mais experientes para auxílio
Problemas com o hardware usado para o desenvolvimento	TE	10%	4	Manter os equipamentos em bom estado de funcionamento

## Conjunto de Dados

Os dados usados neste trabalho foram obtidos através de uma empresa de consultoria informática, a qual, dedica-se à comercialização e à prestação de assistência técnica a sistemas ERP e outros. A empresa é portuguesa e conta com cerca de 14 funcionários e está em atividade desde 2011.

A sua relação com os clientes é realizada, em relação aos pedidos de assistência técnica, na sua maioria, por email. Estes pedidos são posteriormente registados em sistema informático próprio e geridos no mesmo. Cada email válido, ou seja, com um assunto relevante e válido, dá origem a um pedido de assistência (*ticket*). Estes, por sua vez, podem ter associadas diversas

De forma a seleccionar alguns destes dados para uso neste trabalho, foi realizada a sua extração diretamente do sistema de base de dados relacionado usado pelo sistema informático referido anteriormente. Isto referente aos dados relativos a pedidos de assistência e a intervenções.

Já os dados referentes a outros assuntos, por exemplo, emails de publicidade, de relatórios internos, agendamento, assuntos comerciais, entre outros, foram recolhidos diretamente da caixa de email.

Os campos recolhidos foram os emails de origem e destino, o assunto do email e o corpo do mesmo. Posteriormente, foram acrescentadas duas novas colunas, uma para a classe e outra para o sentimento.

**Tabela 4.** Descrição do conjunto de dados

Variável	Tipo	Classificação	N.º Observações
indice	Inteiro	Quantitativo	214
mail_origem	Texto	Qualitativo	214
mail_destino	Texto	Qualitativo	214
assunto	Texto	Qualitativo	214
corpo	Texto	Qualitativo	214
label	Texto	Qualitativo	214
sentimento	Texto	Qualitativo	214

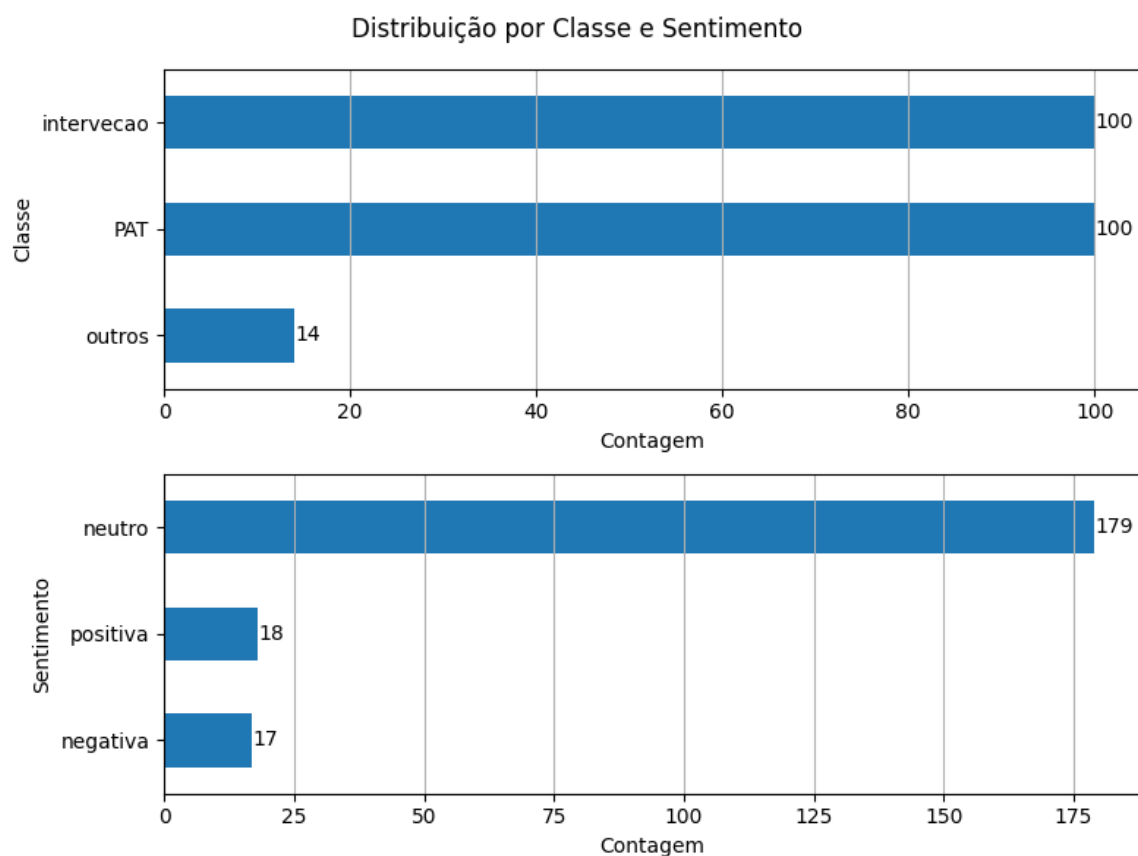
As variáveis apresentadas na tabela anterior são descritas de seguida:

- indice – Campo que mostra o índice do registo.
- mail\_origem – Campo que contem o email de onde foi enviado o email.

- mail\_destino – Campo com o ou os emails de destino.
- assunto – Campo com o assunto do email.
- corpo – Campo com o corpo do email recebido.
- label – Campo com a classe referente ao tipo de email recebido.
- sentimento – Campo com o nível de sentimento do email recebido.

Os valores apresentados nas variáveis mail\_origem, mail\_destino, assunto e corpo sofreram alteração de forma a anonimizar os mesmos. Foram substituídas todas as referências a emails, nomes, telefones, moradas e empresas, por textos genéricos.

intervenção.



**Figura 4.** Distribuição dos dados pelas classes e sentimento

Na figura 4, podemos observar que as classes com maior número de observações são intervenção e pat, com cem observações cada. A restante classe conta com quatorze observações. Já a distribuição dos dados pelo sentimento, também tem contagens de observações distintas por cada um dos sentimentos. Aqui, neutro conta com 179, positiva

tem 18 e negativa tem 17. Posteriormente, no processo de implementação dos algoritmos, serão usadas técnicas para balancear estes registros de acordo com a sua classe ou sentimento.

Os dados estão distribuídos por três classes:

- pat – Classe que rotula os emails com novos pedidos de assistência.
- intervencoes – Classe referente aos emails de intervenções realizadas pelos técnicos ou de emails de resposta a uma solicitação/duvida dos técnicos por parte dos clientes.
- outros – Classe que categoriza os emails que não se enquadram em nenhuma das anteriores.

Já o sentimento de cada um dos dados em análise pode ser positiva, negativa ou neutro.

### **Pré-Processamento**

Para que os dados em análise possam ser usados em algoritmos de classificação, foi necessário proceder ao seu tratamento. Este processo decorreu em três fases complementares, limpeza e homogeneização dos textos, tokenização e remoção de *stopwords*.

Para a primeira fase, foram limpos todos os caracteres especiais (@, #, “, %, entre outros), assim como, todos os espaços em branco. Foi ainda necessário remover os \_ e as quebras de linha (caractere \n). Por fim, todo o texto foi convertido para letras minúsculas

Na segunda fase, os textos foram separados em palavras, chamado de tokenização, com recurso a uma biblioteca de Python, denominada de nltk. Através desta biblioteca foi obtido um modelo já treinado com textos em português, o qual serviu para tratar os textos.

Na terceira e última fase, foram removidas todas as *stopwords* com recurso a um dicionário destas para a língua portuguesa.

Todas estas fases foram realizadas apenas na variável corpo, devido a ser com base nesta que vão ser gerados os modelos de classificação de texto e análise de sentimento.

### **Extração de Atributos**

Para fazer a extração de atributos de forma a analisar os seus atributos mails importantes e relevantes, foram usadas três técnicas, TF-IDF, Word2Vec e Doc2Vec. Estas são três técnicas populares, com metodologias diferentes. A primeira destaca a importâncias das

palavras nos textos enquanto a outras tentam obter o significado semântico das palavras nos textos em análise.

### ***TF-IDF***

Para esta técnica, apenas foi necessário definir os parâmetros a usar na sua implementação, mas especificamente, o número a atributos máximo e a amplitude dos ngram, respetivamente, 10000 e 1 e 2.

### ***Word2Vec***

Neste processo, o modelo gerado teve como parâmetros, o *dataset* de treino, o algoritmo de treino (*sg*), o tamanho dos vetores, distancia (*window*) e o número de palavras mínimo.

Para o parâmetro *sg*, foram usados os dois algoritmos disponíveis para a biblioteca gensim, CBOW e skip – gram. Os restantes atributos, com exceção do *dataset* de treino, foram os seguintes, 300, 5 e 2, respetivamente, tamanho dos vetores, distancia (*window*) e o número de palavras mínimo.

### ***Doc2Vec***

Para esta técnica, o processo é semelhante à anterior, na medida em que aqui também é treinado um modelo para extrair os atributos e este também permite dois algoritmos de treino.

Os parâmetros usados para a criação do modelo foram os algoritmos de treino (*dm*), o tamanho do vetor de palavras, o número de palavras mínimo e o número de *epochs*. Os valores usados foram DBOW e DM para o *dm*, 100, 2 e 40 para os restantes.

Depois de criado o modelo, procedeu se à construção do vocabulário e ao seu treino. Por fim, foram criados os vetores com base no modelo criado e nos dados em análise, quer para os dados de treino e teste.

## **Algoritmos Usados**

De forma a proceder à classificação do texto em análise e analisar o seu sentimento, foram testados diversos algoritmos. Estes são na sua maioria algoritmos de ML, com exceção do LSTM, que é uma rede neural. Desta forma é possível comparar os dois tipos de algoritmos no processamento de linguagem natural.

Todos os algoritmos descritos de seguida foram implementados com e sem CV. por forma a analisar/comparar, a sua robustez e eliminar o *underfitting* e *overfitting*. O CV é um dos métodos de reamostragem mais amplamente utilizado para estimar o verdadeiro erro de predição dos modelos de classificação e ajustar os respetivos parâmetros.

### ***Gaussian NB***

Os métodos NB, são um conjunto de algoritmos supervisionados, baseados no teorema de Bayes. O teorema de Bayes é uma fórmula matemática que é usada para o cálculo de probabilidades condicionais. Probabilidade condicional refere-se à medição de um evento acontecer dado que outro evento ocorreu, quer seja por suposição, presunção, afirmação ou por provas.

A fórmula do teorema de Bayes é a que se segue:

$$P(A|B) = \frac{P(B|A)}{P(B)}$$

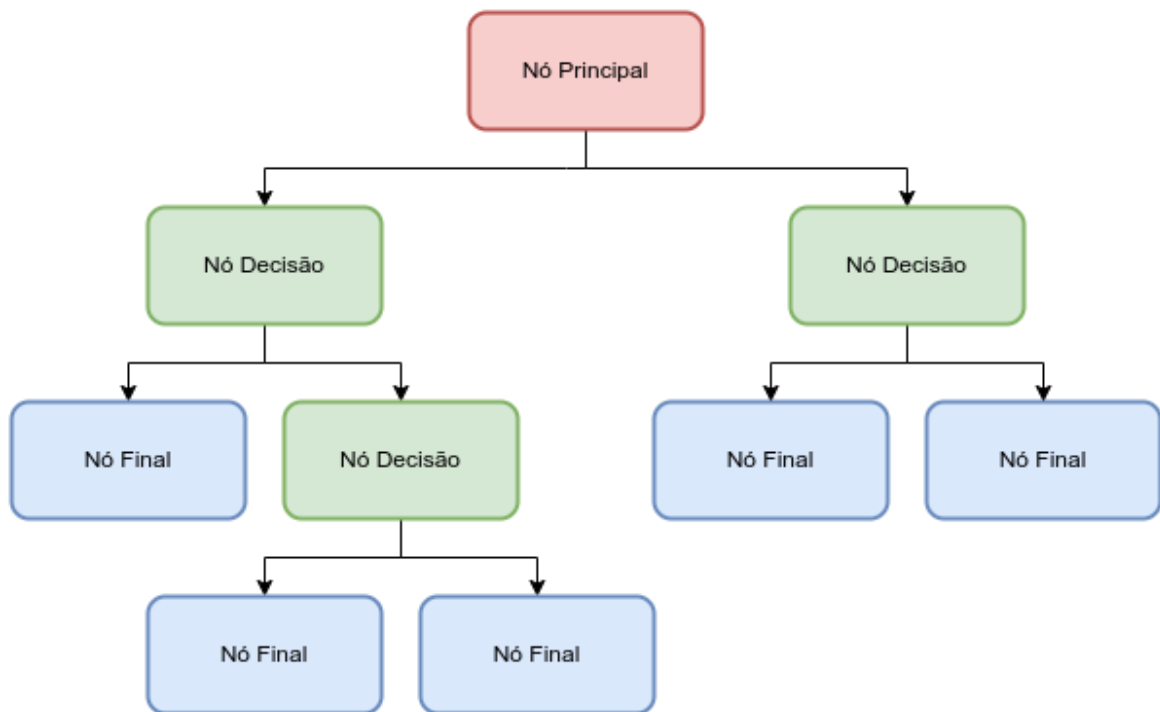
Na fórmula do teorema temos a probabilidade de A acontecer dado que B ocorreu ( $P(A|B)$ ) é igual à probabilidade de B acontecer dado que A ocorreu ( $P(B|A)$ ) vezes a probabilidade de A ( $P(A)$ ) e a dividir pela probabilidade de B ( $P(B)$ ).

Este algoritmo assume que cada um dos atributos é independente e tem a mesma contribuição para o resultado.

Este algoritmo foi implementado de seguida ao pré-processamento e extração de atributos e não teve nenhuma configuração adicional.

### ***DT***

Este algoritmo pertence à família de algoritmos supervisionados e pode ser usado para resolver problemas de regressão e classificação. O seu objetivo é a criação de um modelo de treino que é usado para prever a classe ou valor da variável em análise, de acordo com um conjunto de regras de decisão obtidas de dados anteriormente analisados.



**Figura 5.** Exemplo de DT

Na figura 5 é mostrado um exemplo de um DT. O processo começa no nó principal. Este é comparado com o valor em análise e segue para o próximo nó de acordo com a comparação. Este processo é repetido sucessivamente até chegar ao nó final.

Existem diversos algoritmos de DT, tais como, ID3, C4.5, CART, CHAID, MARS, GUIDE, CTREE, CRUISE e QUEST.

**Tabela 5.** Comparação de Algoritmos de DT

Método	CART	C4.5	CHAID	QUEST
Pruning	Antes, com passagem única do algoritmo	Antes, com passagem única do algoritmo	Antes, com recurso ao Chi-square	Posterior
Variáveis	Catégorica/Contínua	Catégorica/Contínua	Catégorica/Contínua	Catégorica/Contínua
Separação Nó	Binária	Múltipla	Múltipla	Múltipla

Na tabela 5, é apresentada uma comparação de vários algoritmos de DT.

A configuração desta implementação teve como parâmetros *criterion* e *random\_state*. Estes tiveram como valores *entropy* e 0, respetivamente.

## **RF**

Baseia-se em várias árvores de decisão, cujo principal objetivo consiste em agregar os votos de diferentes formações aleatórias das árvores de decisão para determinar a classe final do objeto de teste. É assim um modelo mais robusto e com um melhor desempenho de generalização, por forma a garantir que seja menos suscetível ao *overfitting*.

Para este classificador, foi usado o valor de 50 para o número de estimadores, auto para o número máximo de *features* e o *random\_state* teve como valor 44.

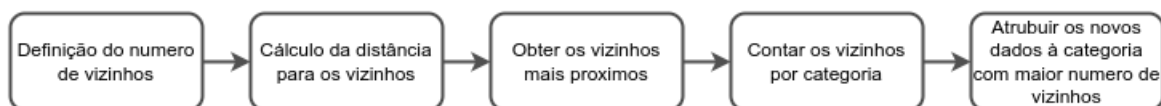
## **AB**

Combina múltiplos classificadores para aumentar a precisão dos mesmos. O classificador desenvolve um classificador forte combinando múltiplos classificadores que têm um desempenho fraco, de modo a obter um classificador forte de alta precisão. Procura estabelecer os pesos dos classificadores e treinar a amostra de dados em cada iteração de forma a garantir previsões precisas para observações incomuns.

Semelhante ao anterior, os parâmetros definidos aqui também foram o número de estimadores e o *random\_state*. Desta vez, os valores usados foram 100 e 0, respetivamente.

## **KNN**

O KNN é um algoritmo supervisionado que serve para analisar problemas de classificação e regressão. Este é um algoritmo não paramétrico, não são baseados num modelo matemático. Em vez disso, aprendem através dos dados fornecidos. Este algoritmo é do tipo *lazy*, o que significa que não necessita dos dados de treino para gerar um modelo. Utiliza os dados do *dataset* na fase de teste, o que torna o treino mais rápido, mas em contrapartida, o teste e a validação ficam mais lentos.



**Figura 6.** Fluxograma KNN

O seu funcionamento, exemplificado na figura 6, passa pela seleção do número de vizinhos a considerar (o valor de K) e pelo cálculo da distância, euclidiana, do novo dado para cada um dos vizinhos. De seguida são contabilizados os vizinhos mais próximos de acordo com a sua categoria. Por fim, de acordo com a métrica obtida no ponto anterior, é

determinada a categoria do novo dado, de acordo com o maior número de vizinhos de uma determinada categoria.

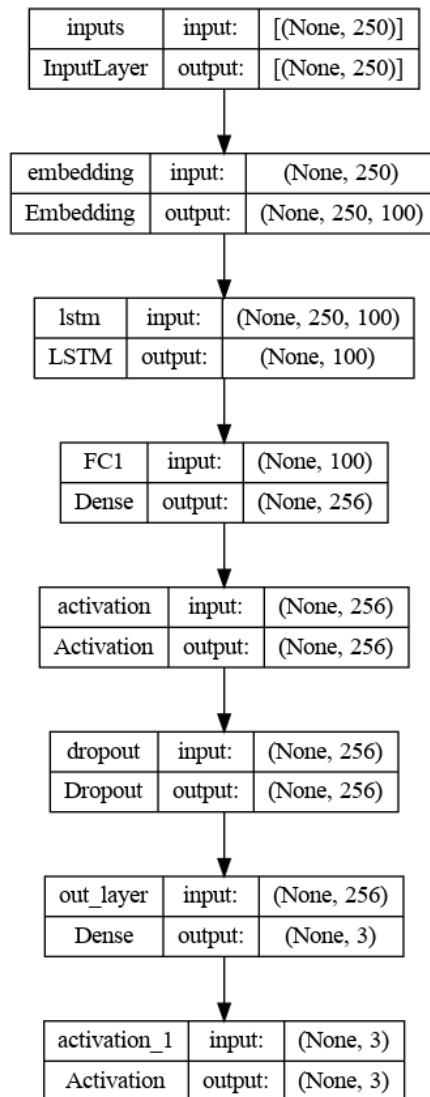
Este classificador foi implementado com duas variações, em que apenas se varia a métrica usada, *minkowski* e *euclidean*. Para o número de vizinhos foi usado o valor de 3.

### ***LSTM***

Este é um tipo de RNN, cuja característica principal passa pela captura e registro de informação em longas sequências (Gupta, 2020).

As suas principais características e componentes são os seguintes:

- *Memory Cells*: São células de memória especiais que podem armazenar e recuperar informações ao longo do tempo, o que permite que estes algoritmos tenham “memória”.
- *Gates*: Existem três tipos de portas: entrada, esquecimento e saída. São estas que controlam o fluxo de informações de entrada e saída das células de memória.
  - Forget Gate: Remove as informações que deixam de ser relevantes;
  - Input Gate: Decide quais as informações que devem ser mantidas;
  - Output Gate: Extrai as informações úteis para que estas possam ser apresentadas como saída.
- *Hidden State*: É uma função do estado oculto anterior, da entrada atual e das saídas dos portões.
- *Long-Term and Short-Term Memory*: Armazenam seletivamente informações nas células de memória por longos períodos, permitindo-lhes capturar dependências de dados de longo prazo. Simultaneamente, podem atualizar e adaptar-se rapidamente a novas informações, permitindo-lhes capturar também dependências de curto prazo.



**Figura 7. Modelo LSTM**

O modelo apresentado na imagem anterior mostra a forma da rede LSTM. É composta por sete camadas, as quais se descrevem de seguida:

- *Input Layer*: Camada que recebe os dados de entrada, com sequencias de dados com um tamanho máximo de 250.
- *Embedding Layer*: Camada que trata da otimização dos dados para que estes possam ser processados com maior eficiência. Tem como parâmetros *vocab\_size*, que representa o número de palavras únicas no *dataset*, dimensão de saída e *input\_length*. Os valores para cada uma destes parâmetros é 5000, 100, 200, respetivamente.

- *LSTM Layer*: Camada com o parâmetro de dimensão do espaço de saída com o valor 100.
- *Dense Layer (Fully Connected Layer)*: Camada densa com o parâmetro *units* com o valor de 256 e função de ativação *relu*.
- *Dropout Layer*: Nesta camada foi aplicada uma taxa de perda de 0.2.
- *Dense Output Layer*: Semelhante á outra camada densa, tem desta feita, 3 como o valor de *units* e a função de ativação é, desta vez, *sigmoid*.

Este algoritmo teve como otimizador o valor de adam, sem parâmetros adicionais definidos e teve o valor de 5 para os *epochs*.

### Avaliação desempenho

As avaliações feitas aos algoritmos têm como parâmetros de avaliação a *accuracy*, *recall* e *f-score*. A precisão é a medida dos valores relevantes verdadeiros determinados pelo modelo. O *recall* é a medida dos verdadeiros encontrados pelo modelo. Já a *f-score*, refere-se à combinação da *accuracy* e do *recall* em uma única métrica.

**Tabela 6.** Matriz de Confusão

	Previsto		
Atual	True Positive (TP)	False Negative (FN)	Recall $TP/(TP+FN)$
	False Positive (FP)	True Negative (TN)	Especificidade $TN/(TN+FP)$
	Precision $TP/(TP+FP)$	Negative Rate $TN/(TN+FN)$	Accuracy $TP+TN/(TP+TN+FP+FN)$

Conforme apresentado na tabela 6, a matriz de confusão é composta pelos resultados divididos por categorias, TP, FP, FN e TN.

A *precision* mostra quantos registros dos previstos corretamente, eram casos positivos.

O *negative rate* mostra quantos dos registros previstos como negativos, eram na verdade negativos.

O *recall* explica quantos dos casos positivos foram corretamente previstos pelo modelo.

A *specificity* é o contrário do *recall*, explica quantos dos casos negativos foram corretamente previstos.

O *f-score*, também conhecida com *f1 – score*, descrita anteriormente, permite a avaliação do desempenho do modelo. É usada para comparar dois ou mais classificadores.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Para o LSTM foi usada, para além da precisão, a perda.

A perda, também conhecida como custo ou função objetivo, quantifica o erro ou a dissimilaridade entre as previsões do modelo e os valores alvo reais no conjunto de dados de treino. O objetivo durante este é minimizar essa função de perda. Funções de perda comuns para tarefas de classificação incluem perda de entropia cruzada (binária ou categórica), erro quadrático médio (para tarefas de regressão) e outras. Durante o treino, os parâmetros do modelo (pesos e vieses) são ajustados usando técnicas de otimização (por exemplo, gradiente descendente) para minimizar a função de perda, o que significa fazer as previsões o mais próximo possível dos verdadeiros valores alvo.

A matriz de confusão foi implementada sem recurso a parâmetros adicionais, com exceção do cálculo da *accuracy*, *recall* e *f1 – score*, os quais tiveram como *average* o valor de *weighted*.

Para o algoritmo LSTM, a função de perda usada foi *sparse\_categorical\_crossentropy*.

## **Balanceamento**

Como os dados em análise não estavam balanceados pelas classes de tipo de email e pelo sentimento destes, foi necessário recorrer a técnicas de balanceamento sintético dos dados para tentar melhorar a performance dos modelos gerados.

Desta forma, foram usadas quatro técnicas de balanceamento de geração de novos dados para igualar o número de registos por classes. Foi usado, quer para o tipo de email, quer para o sentimento. As técnicas usadas foram as seguintes:

- FS: Técnica implementada sem definição de parâmetros adicionais.
- ROS: Técnica implementada com o valor de 0 para o parâmetro *random\_state*.
- SMOTE: Semelhante à anterior, com o valor de 0 para o parâmetro *random\_state*.
- ADASYN: Também aqui foi usado 0 para o parâmetro *random\_state*. Foi ainda usado o valor de *minority* para o parâmetro *sampling\_strategy*.

Além das técnicas de balanceamento enumeradas, foi ainda testado o *dataset* sem qualquer tipo de técnica aplicada.

## RESULTADOS

Os resultados serão apresentados por cada uma das técnicas de extração de atributos, pelas técnicas de balanceamento implementadas e por fim, será feito um resumo global. O método de avaliação usado para todas as implementações será a matriz de confusão, sendo que para o algoritmo LSTM, será apresentada também a métrica referente à perda (*loss*).

Serão listados apenas os dez melhores resultados por cada uma das subsecções, sendo que, os resultados integrais podem ser consultados nos apêndices 1, 2, 3 e 4.

### Por extrator

Nesta secção serão apresentados os resultados por cada uma das técnicas de extração de atributos usadas. Como o algoritmo LSTM foi implementado sem recuso a nenhuma destas técnicas, não figura nos dados apresentados nesta secção.

### *TF-IDF*

Para esta técnica de extração de atributos, os algoritmos com melhor resultados em relação á classificação de texto, foram o RF e o AB, com 100% para *accuracy*, *precision*, *recall* e *f1 – score*.

**Tabela 7.** 10 melhores resultados para classificação - TF IDF

Algoritmo	Equilibrador	CV	Accuracy	Precision	Recall	F1 - Score
RF	ROS	False	100	100	100	100
AB	ROS	False	100	100	100	100
AB	ADASYN	True	98,9	98,93	98,9	98,9
AB		False	98,46	98,51	98,46	98,22
AB	FS	False	98,46	98,51	98,46	98,22
DT	ADASYN	True	97,8	97,92	97,8	97,79
DT	ROS	False	97,78	97,92	97,78	97,78
RF	SMOTE	False	97,78	97,9	97,78	97,77
DT		False	96,92	96,92	96,92	96,92
DT		True	96,92	94,03	96,92	95,43

Na tabela 7 podemos verificar que estes resultados foram obtidos com recurso à técnica de balanceamento ROS e sem ter sido realizado o processo de CV. De seguida, com 98.9, 98.93, 98.9 e 98,9 para *accuracy*, *precision*, *recall* e *f1 – score*, respetivamente, temos

novamente o algoritmo AB. Desta feita, estes resultados foram obtidos com o equilibrador ADASYN e com o uso da técnica de CV.

Dos resultados apresentados, os resultados variam entre 98.46 e 96.92 para *accuracy*, 98.51 e 94.03 para *precision*, 98.46 e 96.92 para *recall* e para *f1 – score*, 98.22 e 95.43.

Os algoritmos mais representados neste top-10 são RF e AB com quatro aparições e RF com duas, sendo que a técnica de balanceamento mais usada foi ROS e sem técnica com três usos, ADASYN com dois e FS e SMOTE com um.

Já em relação ao uso de CV, apenas por três vezes esta técnica foi usada.

Em relação à análise de sentimento dos textos, o melhor desempenho foi obtido pelo algoritmo RF, com uma classificação de 99.38, 99.39, 99.38 e 99.38, para *accuracy*, *precision*, *recall* e *f1 – score*, respetivamente.

**Tabela 8.** 10 melhores resultados para sentimento - TF IDF

Algoritmo	Equilibrador	CV	Accuracy	Precision	Recall	F1 - Score
RF	ROS	False	99,38	99,39	99,38	99,38
RF	SMOTE	False	99,38	99,39	99,38	99,38
NB	ROS	False	97,53	97,62	97,53	97,51
NB	SMOTE	False	97,53	97,62	97,53	97,51
DT	SMOTE	False	95,68	95,69	95,68	95,68
RF	ADASYN	False	94,74	93,27	94,74	93,51
DT	ROS	False	94,44	94,93	94,44	94,33
RF	ROS	True	93,21	93,56	93,21	93,23
RF	SMOTE	True	91,98	92,88	91,98	92,06
NB	SMOTE	True	91,36	92,38	91,36	91,02

De acordo com a tabela 8, a técnica de balanceamento dos dados usada para o melhor resultado foi ROS em igualdade com SMOTE. Ambas sem recuso ao CV.

Os restantes resultados variam entre 97.53 e 91.36 para *accuracy*, 97.62 e 92.38 para *precision*, 97.53 e 91.36 para *recall* e 97.51 e 91.02 para *f1 – score*.

À semelhança da classificação, o algoritmo mais representado nestes resultados é o RF, desta feita com cinco entradas, seguido pelo NB e DT com duas entradas cada. A técnica de balanceamento mais usada foi SMOTE com cinco registos, seguida de ROS com quatro e ADASYN com uma.

A técnica de CV apenas foi usada por três vezes, estando posicionada nos últimos registros deste top 10.

### **Word2Vec**

A criação do modelo para esta técnica de extração de atributos fez uso de dois algoritmos distintos, *skip-gram* e CBOW.

Para a classificação, o melhor desempenho foi obtido pelo algoritmo NB, com uso de várias técnicas de balanceamento dos dados e com uso do algoritmo CBOW para criação do modelo. Aliás, em todas as entradas deste top 10, o algoritmo usado para a classificação foi o NB, com o algoritmo do modelo Word2Vec usado a ser o CBOW.

**Tabela 9.** 10 melhores resultados para classificação - Word2Vec

Algoritmo	Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
NB		CBOW	True	84,62	82,15	84,62	83,29
NB	FS	CBOW	True	84,62	82,15	84,62	83,29
NB	ROS	CBOW	True	84,62	82,15	84,62	83,29
NB	SMOTE	CBOW	True	84,62	82,15	84,62	83,29
NB	ADASYN	CBOW	True	84,62	82,15	84,62	83,29
NB		CBOW	False	83,08	83,85	83,08	83,4
NB	FS	CBOW	False	83,08	83,85	83,08	83,4
NB	ROS	CBOW	False	83,08	83,85	83,08	83,4
NB	SMOTE	CBOW	False	83,08	83,36	83,08	82,91
NB	ADASYN	CBOW	False	83,08	83,36	83,08	82,91

Na tabela 9 podemos verificar que os cinco melhores resultados têm todos a mesma avaliação, com 84.62, 82.15, 84.62 e 83.29, para *accuracy*, *precision*, *recall* e *f1 – score*, respectivamente.

Estes resultados foram obtidos com recursos a todos os métodos de balanceamento dos dados, assim como, sem o uso dessa técnica. De referir ainda, que em todos eles foi aplicado a CV.

Os restantes resultados são semelhantes aos primeiros cinco, com a diferença principal a ser o não uso de CV e a *accuracy* e *recall* a terem valores ligeiramente inferiores, como a *precision* e *f1 – score* a terem valores ligeiramente superiores.

Em relação ao sentimento, os melhores resultados foram obtidos pelo algoritmo RF, com o balanceamento a ser realizado pelas técnicas SMOTE e ADASYN, respectivamente, e o algoritmo do modelo de extração de atributos a ser *skip-gram* e CBOW.

**Tabela 10.** 10 melhores resultados para sentimento - Word2Vec

Algoritmo	Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
RF	SMOTE	skip - gram	False	89,23	89,63	89,23	88,38
RF	ADASYN	CBOW	False	89,23	90,45	89,23	86,51
AB		skip - gram	False	87,69	81,56	87,69	83,27
AB	FS	skip - gram	False	87,69	81,56	87,69	83,27
RF		CBOW	False	87,69	81,56	87,69	83,27
RF	FS	CBOW	False	87,69	81,56	87,69	83,27
RF	ROS	CBOW	False	87,69	81,56	87,69	83,27
RF		skip - gram	False	86,15	80,41	86,15	80,78
AB		skip - gram	True	86,15	78,83	86,15	81,95
RF	FS	skip - gram	False	86,15	80,41	86,15	80,78

Os resultados dos dois primeiros registos da tabela 10, são muito semelhantes, com os seus valores a serem 89.23, 89.63, 89.23 e 88.38 e 89.23, 90.45, 89.23 e 86.51, para *accuracy*, *precision*, *recall* e *f1 - score*, respetivamente.

Neste caso, a maior parte das entradas foi obtida sem recurso a técnicas de balanceamento, as restantes fizeram uso de FS por três vezes, ROS, SMOTE e ADASYN por uma vez, cada. Já os algoritmos usados foram RF por sete vezes e AB por três.

Para a geração do modelo de extração de atributos, foram usados os dois algoritmos disponíveis, sendo o *skip-gram* usado por seis vezes e o CBOW por quatro.

Apenas em uma entrada deste top 10, a penúltima, temos o uso de CV.

### **Doc2Vec**

Para a criação do modelo de extração de atributos, foram usados os dois algoritmos possíveis, DM e DBOW.

Relativamente ao processo de classificação, os melhores resultados foram obtidos pelo algoritmo RF, com e sem recuso a técnicas de balanceamento, com o modelo de extração de atributos usado a ser DM e sem uso de CV.

**Tabela 11.** 10 melhores resultados para classificação - Doc2Vec

Algoritmo	Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
-----------	--------------	----------	----	----------	-----------	--------	------------

RF	ROS	DM	False	76,92	77,41	76,92	76,86
RF		DM	False	75,38	77,96	75,38	75,79
RF	FS	DM	False	75,38	77,96	75,38	75,79
RF		DM	True	73,85	71,58	73,85	72,68
RF	FS	DM	True	73,85	71,58	73,85	72,68
RF	ROS	DM	True	73,85	71,58	73,85	72,68
RF	SMOTE	DM	False	73,85	75,95	73,85	74,43
RF	SMOTE	DM	True	73,85	71,58	73,85	72,68
DT	ADASYN	DM	False	73,85	75,21	73,85	74,08
RF	ADASYN	DM	False	73,85	75,95	73,85	74,43

Na tabela 11 podemos analisar os melhores resultados. Os três primeiros têm valores de avaliação muito aproximados, com ligeiras variações, 76.92, 77.41, 76.92 e 76.86 e 75.38, 77.96, 75.38 e 75.79, para *accuracy*, *precision*, *recall* e *f1 – score*, respetivamente.

Os algoritmos mais usados para esta técnica de extração foram o RF com 9 entradas e uma para o DT. Já em relação à técnica de balanceamento, temos dois registos para cada uma delas, incluindo, registos sem o seu uso. Todos os resultados apresentados fazem uso o algoritmo DM para a geração do modelo de extração de atributos.

Em relação ao processo de CV, temos 4 entradas que fazem uso deste, posições 3, 4, 5 e 7.

Para a análise de sentimento, o melhor resultado foi obtido pelo algoritmo RF, FS e sem uso de técnica de balanceamento, com o algoritmo de criação do modelo de extração de atributos usado a ser DBOW e sem CV.

**Tabela 12.** 10 melhores resultados para sentimento – Doc2Vec

Algoritmo	Variante	Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 – Score
RF			DBOW	False	86,15	78,83	86,15	81,95
RF		FS	DBOW	False	86,15	78,83	86,15	81,95
KNN	Mink.		DBOW	False	84,62	76,37	84,62	79,64
KNN	Euclidean		DBOW	False	84,62	76,37	84,62	79,64
KNN	Mink.	FS	DBOW	False	84,62	76,37	84,62	79,64
KNN	Euclidean	FS	DBOW	False	84,62	76,37	84,62	79,64
RF		ROS	DBOW	False	84,62	76,37	84,62	79,64
AB			DBOW	False	83,08	74,9	83,08	78,58
AB		FS	DBOW	False	83,08	74,9	83,08	78,58

RF	DM	False	83,08	76,41	83,08	79,6
----	----	-------	-------	-------	-------	------

Os valores das métricas dos melhores resultados, apresentados na tabela 12, são 86.15, 78.83, 86.15 e 81.95, para *accuracy*, *precision*, *recall* e *f1 – score*, respectivamente.

Nesta parte, os algoritmos RF e KNN, este com uso das distâncias *minkowski* e *euclidean*, com duas entradas para cada uma, são os que estão representados em maior número, com quatro entradas cada. As restantes são preenchidas pelo AB, com duas. Para o balanceamento, cinco das entradas não fizeram uso destas técnicas, quatro usaram FS e a restante usou ROS. Já para o algoritmo usado para criação do modelo de extração de atributos, existe uma preponderância do DBOW, com nove entradas e uma para DM.

No top 10 apresentado, não foi usado o procedimento de CV em nenhuma entrada.

## Algoritmo

Aqui podemos verificar os resultados por cada um dos algoritmos implementados.

### *NB*

Para este algoritmo, no processo de classificação, os melhores resultados foram obtidos com cada uma das técnicas de balanceamento usadas, assim como sem recurso a estas. O extrator usado foi o Word2Vec com o algoritmo implementado neste processo a ser o CBOW e com o recurso a CV.

**Tabela 13.** 10 melhores resultados para classificação – NB

Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
	Word2Vec - CBOW	True	84,62	82,15	84,62	83,29
FS	Word2Vec - CBOW	True	84,62	82,15	84,62	83,29
ROS	Word2Vec - CBOW	True	84,62	82,15	84,62	83,29
SMOTE	Word2Vec - CBOW	True	84,62	82,15	84,62	83,29
ADASYN	Word2Vec - CBOW	True	84,62	82,15	84,62	83,29
	Word2Vec - CBOW	False	83,08	83,85	83,08	83,4
FS	Word2Vec - CBOW	False	83,08	83,85	83,08	83,4

ROS	Word2Vec - CBOW	False	83,08	83,85	83,08	83,4
SMOTE	Word2Vec - CBOW	False	83,08	83,36	83,08	82,91
ADASYN	Word2Vec - CBOW	False	83,08	83,36	83,08	82,91

A tabela 13, as métricas para os melhores resultados são 84.62, 82.15, 84.62 e 83.29, para *accuracy*, *precision*, *recall* e *f1 – score*, respetivamente.

Todos os registos apresentados são semelhantes, com ligeiras variações nos valores das métricas e no uso de CV. Todos fazem uso da mesma técnica de extração de atributos, Word2Vec com CBOW, sendo que os cinco primeiros usam CV e os restantes não. As métricas dos cinco primeiros são exatamente as mesmas, mas nos últimos cinco, temos variações na *precision* e *f1 – score*.

Para a análise de sentimento, os melhores resultados obtiveram-se com recurso ao balanceador ROS e SMOTE, com a extração de atributos realizada através de TF-IDF e sem CV.

**Tabela 14.** 10 melhores resultados para sentimento – NB

Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
ROS	TF-IDF	False	97,53	97,62	97,53	97,51
SMOTE	TF-IDF	False	97,53	97,62	97,53	97,51
SMOTE	TF-IDF	True	91,36	92,38	91,36	91,02
ADASYN	TF-IDF	False	90,35	87,88	90,35	89,09
ROS	TF-IDF	True	85,8	86,17	85,8	85,59
	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
FS	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
ROS	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
SMOTE	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
ADASYN	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56

Conforme o apresentado na tabela 14, os primeiros cinco resultados fazem uso da extração de atributos através de TF-IDF e os restantes Word2Vec com o algoritmo CBOW.

Estes têm um maior uso de algoritmos sem CV, em contraste com os últimos, os quais usam esta técnica em todas as entradas.

Relativamente ao balanceador usado, temos três estradas com ROS e SMOTE, duas com ADSYN e uma com FS e sem balanceador, cada.

### ***DT***

Desta feita, para este algoritmo, os melhores resultados para a classificação foram obtidos com recursos à técnica de balanceamento ADASYN e ROS, ao extrator TF-IDF e com e sem CV.

**Tabela 15.** 10 melhores resultados para classificação – DT

Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
ADASYN	TF-IDF	True	97,8	97,92	97,8	97,79
ROS	TF-IDF	False	97,78	97,92	97,78	97,78
	TF-IDF	False	96,92	96,92	96,92	96,92
	TF-IDF	True	96,92	94,03	96,92	95,43
FS	TF-IDF	False	96,92	96,92	96,92	96,92
FS	TF-IDF	True	96,92	94,03	96,92	95,43
ADASYN	TF-IDF	False	96,7	96,97	96,7	96,68
SMOTE	TF-IDF	False	95,56	96,01	95,56	95,52
SMOTE	TF-IDF	True	93,33	93,73	93,33	93,28
ROS	TF-IDF	True	91,11	91,11	91,11	91,11

Para os melhores resultados, apresentados na tabela 15, os valores das métricas foram os seguintes 97.8, 97.92, 97.8 e 97.79 e 97.78, 97.92, 97.78 e 97.78, para *accuracy*, *precision*, *recall* e *f1 – score*, respetivamente.

Foram usadas todas as técnicas de balanceamento, como a ausência desta, tendo cada uma, duas entradas. O extrator implementado em todos os 10 primeiros resultados foi TF-IDF e existe um misto de uso de CV, cinco entradas a usar e as outras cinco, a não fazer uso desta.

Para o processo de análise de sentimento, os melhores resultados foram obtidos através da implementação das técnicas de balanceamento SMOTE e ROS, com o método de extração de atributos usado a ser TF-IDF e sem recurso a CV.

**Tabela 16.** 10 melhores resultados para sentimento – DT

Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
SMOTE	TF-IDF	False	95,68	95,69	95,68	95,68
ROS	TF-IDF	False	94,44	94,93	94,44	94,33
ROS	TF-IDF	True	88,27	88,76	88,27	88,07
ADASYN	TF-IDF	False	87,72	87,08	87,72	87,24
SMOTE	TF-IDF	True	85,19	85,16	85,19	85,16
ADASYN	TF-IDF	True	85,09	80,89	85,09	82,84
	TF-IDF	True	84,62	76,37	84,62	79,64
FS	TF-IDF	True	84,62	76,37	84,62	79,64
	TF-IDF	False	83,08	84,66	83,08	81,74
FS	TF-IDF	False	83,08	84,66	83,08	81,74

As métricas *accuracy*, *precision*, *recall* e *f1 - score*, para os dois primeiros resultados, têm os valores de 95.68, 95.69, 95.68 e 95.68 e 94.44, 94.93, 94.44 e 94.33, respetivamente, sendo os únicos a ter valores superiores a 90%, de acordo com a tabela 16.

À semelhança da classificação, aqui também existe uma distribuição homogenia das técnicas de balanceamento, com duas entradas para cada uma, assim como o uso do método de extração de atributos a também ser TF-IDF. Até o uso de CV tem o mesmo número de entradas, no entanto, com posições diferentes.

### **RF**

A implementação deste algoritmo, no processo de classificação, teve o melhor resultado, quando usado em conjunto com a técnica de balanceamento ROS, o método de extração de atributos TF-IDF e sem CV.

**Tabela 17.** 10 melhores resultados para classificação – RF

Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
ROS	TF-IDF	False	100	100	100	100
SMOTE	TF-IDF	False	97,78	97,9	97,78	97,77
ADASYN	TF-IDF	False	96,7	96,97	96,7	96,68
	TF-IDF	False	95,38	95,48	95,38	95,16
	TF-IDF	True	95,38	92,45	95,38	93,89
FS	TF-IDF	False	95,38	95,48	95,38	95,16
FS	TF-IDF	True	95,38	92,45	95,38	93,89
ADASYN	TF-IDF	True	91,21	92,88	91,21	91,34

ROS	TF-IDF	True	90	92,05	90	90,1
SMOTE	TF-IDF	True	86,67	90,07	86,67	86,76

Desta feita, o melhor resultado listado na tabela 17, tem 100% para todas as métricas usadas, *accuracy*, *precision*, *recall* e *f1 – score*.

Em todo semelhante ao algoritmo DT, também aqui, temos o uso todos os balanceadores e registo sem o seu uso, duas vezes cada. O extrator é na mesma TF-IDF, com a metade inferior dos registos a fazer uso de CV e a metade superior, não.

Já a análise de sentimento tem resultados diferentes. Os seus melhores registos fazem uso dos balanceadores ROS e SMOTE, do extrator TF-IDF e não implementam CV.

**Tabela 18.** 10 melhores resultados para sentimento – RF

Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
ROS	TF-IDF	False	99,38	99,39	99,38	99,38
SMOTE	TF-IDF	False	99,38	99,39	99,38	99,38
ADASYN	TF-IDF	False	94,74	93,27	94,74	93,51
ROS	TF-IDF	True	93,21	93,56	93,21	93,23
SMOTE	TF-IDF	True	91,98	92,88	91,98	92,06
SMOTE	Word2Vec - skip - gram	False	89,23	89,63	89,23	88,38
ADASYN	Word2Vec - CBOW	False	89,23	90,45	89,23	86,51
	Word2Vec - CBOW	False	87,69	81,56	87,69	83,27
FS	Word2Vec - CBOW	False	87,69	81,56	87,69	83,27
ROS	Word2Vec - CBOW	False	87,69	81,56	87,69	83,27

Na tabela 18 é possível verificar que a técnica de balanceamento de dados mais usada é ROS em igualdade com SMOTE, com três entradas. O ADASYN tem duas e o FS tem uma, assim como, o registo que não faz uso de nenhuma técnica. Os extratores de atributos usados são TF-IDF para os cinco primeiros resultados e Word2Vec para os restantes cinco. Este último fez uso dos dois algoritmos para a criação do seu modelo, *skip-gram* e CBOW.

Aqui, o processo de CV apenas foi usado duas vezes, nas entradas 4 e 5.

## AB

Este classificador teve como melhor resultado, para a classificação, 100% em todas as métricas. Para esta entrada, o balanceador de dados usado foi o ROS e o extrator o TF-IDF, sem a implementação de CV.

**Tabela 19.** 10 melhores resultados para classificação – AB

Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
RandomOverSampler	TF-IDF	False	100	100	100	100
ADASYN	TF-IDF	True	98,9	98,93	98,9	98,9
	TF-IDF	False	98,46	98,51	98,46	98,22
FunctionSampler	TF-IDF	False	98,46	98,51	98,46	98,22
	TF-IDF	True	96,92	94,03	96,92	95,43
FunctionSampler	TF-IDF	True	96,92	94,03	96,92	95,43
SMOTE	TF-IDF	True	96,67	96,93	96,67	96,65
ADASYN	TF-IDF	False	94,51	94,52	94,51	94,5
SMOTE	TF-IDF	False	94,44	94,67	94,44	94,41
RandomOverSampler	TF-IDF	True	93,33	93,41	93,33	93,31

Na tabela 19 e à semelhança dos algoritmos de RF e DT, existe uma distribuição por todos os métodos de balanceamento de dados, com todos os utilizadores a terem dois registos cada. Também a não implementação desta técnica, teve dois resultados. O extrator em todas as entradas é TF-IDF e o método de CV, mais uma vez, semelhante ao DT o RF, está dividido com cinco registos para a implementação e outros cinco para a não implementação.

No que diz respeito ao processo de análise de sentimento, os melhores resultados foram obtidos com e sem implementação de balanceador, com o uso do extrator TF-IDF e sem CV.

**Tabela 20.** 10 melhores resultados para sentimento – AB

Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
	TF-IDF	False	87,69	86,27	87,69	86,37
FS	TF-IDF	False	87,69	86,27	87,69	86,37
	Word2Vec - skip - gram	False	87,69	81,56	87,69	83,27
FS	Word2Vec - skip - gram	False	87,69	81,56	87,69	83,27
	Word2Vec - skip - gram	True	86,15	78,83	86,15	81,95
FS	Word2Vec - skip - gram	True	86,15	78,83	86,15	81,95

	skip - gram						
ROS	Word2Vec - skip - gram	False	86,15	80,41	86,15	80,78	
ROS	Word2Vec - skip - gram	True	86,15	78,83	86,15	81,95	
SMOTE	Word2Vec - skip - gram	True	86,15	78,83	86,15	81,95	
ADASYN	Word2Vec - skip - gram	True	86,15	78,83	86,15	81,95	

Os valores das métricas para os primeiros quatro resultados da tabela 20, apresentam valores semelhantes, com ligeiras diferenças para *precision* e *f1 - score*, sendo os valores apresentados 87.69, 86.27, 87.69 e 86.37, para os dois primeiros e 87.69, 81.56, 87.69 e 83.27 para os outros dois, sendo estas *accuracy*, *precision*, *recall* e *f1 - score*, respetivamente. Nos primeiros lugares existe uma maioria do não de balanceador e do uso de FS, com três registos cada. Os restantes são ROS com duas entradas e SMOTE e ADASYN com uma cada. Apesar de os melhores resultados fazerem uso do extrator TF-IDF, todos os restantes oito, usam Word2Vec com o algoritmo usado para o seu modelo a ser *skip-gram*.

### ***KNN***

A implementação deste algoritmo teve o mesmo resultado para os seus dez primeiros lugares, para o processo de classificação de texto.

**Tabela 21.** 10 melhores resultados para classificação – KNN

Variante	Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
Minkowski		Word2Vec - CBOW	True	80	77,54	80	78,75
Euclidean		Word2Vec - CBOW	True	80	77,54	80	78,75
Minkowski	FS	Word2Vec - CBOW	True	80	77,54	80	78,75
Euclidean	FS	Word2Vec - CBOW	True	80	77,54	80	78,75
Minkowski	ROS	Word2Vec - CBOW	True	80	77,54	80	78,75
Euclidean	ROS	Word2Vec - CBOW	True	80	77,54	80	78,75
Minkowski	SMOTE	Word2Vec - CBOW	True	80	77,54	80	78,75

Euclidean	SMOTE	Word2Vec - CBOW	True	80	77,54	80	78,75
Minkowski	ADASYN	Word2Vec - CBOW	True	80	77,54	80	78,75
Euclidean	ADASYN	Word2Vec - CBOW	True	80	77,54	80	78,75

Conforme apresentado na tabela 21, existe uma distribuição dos resultados por todas as técnicas usadas. Metade dos registos usa a distância *minkowski* e a outra, *euclidean*. Cada uma das técnicas de balanceamento usadas conta com dois registos, assim como, a não implementação desta. O extrator usado nos dez primeiros resultados é Word2Vec com o algoritmo do modelo a ser CBOW, e todos estes, implementam o método de CV.

Os valores das métricas para *accuracy*, *precision*, *recall* e *f1 – score* são 80, 77.54, 80, e 78.75, respetivamente.

A análise de sentimento o melhor desempenho foi obtido com recuso às distâncias *minkowski* e *euclidean*, à técnica de balanceamento ROS, ao extrator TF-IDF e sem implementar CV.

**Tabela 22.** 10 melhores resultados para sentimento – KNN

Variante	Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
Minkowski	ROS	TF-IDF	False	85,19	88,08	85,19	84,02
Euclidean	ROS	TF-IDF	False	85,19	88,08	85,19	84,02
Minkowski		Doc2Vec - DBOW	False	84,62	76,37	84,62	79,64
Euclidean		Doc2Vec - DBOW	False	84,62	76,37	84,62	79,64
Minkowski	FS	Doc2Vec - DBOW	False	84,62	76,37	84,62	79,64
Euclidean	FS	Doc2Vec - DBOW	False	84,62	76,37	84,62	79,64
Minkowski		Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Euclidean		Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Minkowski	FS	Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Euclidean	FS	Word2Vec - CBOW	True	84,62	81,21	84,62	81,72

De acordo com a tabela 22, os valores para os dois primeiros resultados são 85.19, 88.08, 85.19 e 84.02, para *accuracy*, *precision*, *recall* e *f1 – score*, respetivamente.

De resto, semelhante à classificação, cinco registos usam a distância *minkowski* e os outros cinco, *euclidean*. A não implementação de 4 técnica de balanceamento, tem quatro registos, o mesmo número que a técnica FS. O restantes dois, os melhores, implementam a técnica ROS.

O extrator usado nos primeiros é TF-IDF, seguido pelo Doc2Vec com DBOW para os próximos quatro e para os últimos quatro, foi usado Word2Vec com CBOW.

A CV apenas foi implementada para os últimos quatro registos.

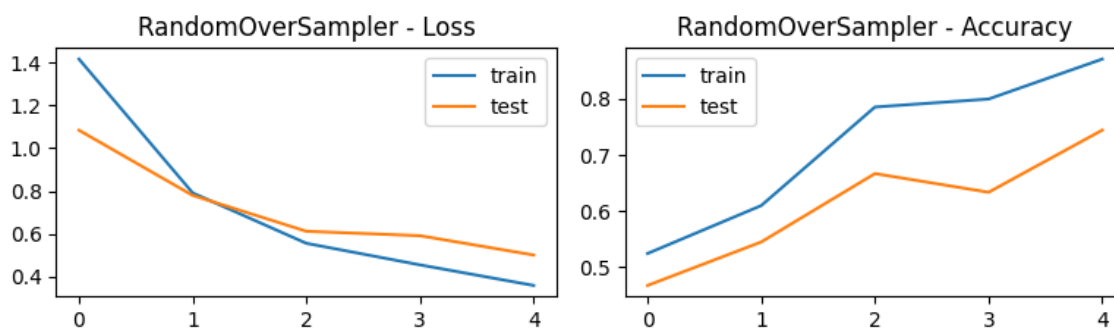
### **LSTM**

Este algoritmo foi implementado sem fazer uso de métodos de extração de atributos. Neste caso, o melhor resultado obtido para a classificação, ocorreu quando se fez uso do balanceador ROS.

**Tabela 23. Melhores resultados para classificação – LSTM**

Equilibrador	Accuracy	Precision	Recall	F1 - Score	Loss
ROS	74,44	74,74	74,44	74,44	0.5
ADASYN	64,44	67,08	64,44	61,83	1.29
FS	63,08	61,59	63,08	61,95	0.83
	58,46	61,35	58,46	53,98	0.76
SMOTE	55,56	55,83	55,56	55,22	1.43

Na tabela 23, o melhor resultado teve 74.44, 74.74, 74.44, 74.44 e 0.5, para *accuracy*, *precision*, *recall*, *f1 – score* e *loss*, respetivamente.



**Figura 8. Loss e accuracy - classificação**

O gráfico com as curvas de *loss* e *accuracy* pode ser consultado na figura 8. Nela pode ser vista a sua evolução, dos dados de treino e teste, por cada um dos *epochs*.

As restantes técnicas de balanceamento, obtiveram resultados entre 67.08 e 55.56 para *accuracy*, 67.08 e 55.83 para *precision*, 64.44 e 55.56 para *recall*, 61.83 e 55.22 para *f1 – score* e 1.29 e 1.43 para *loss*.

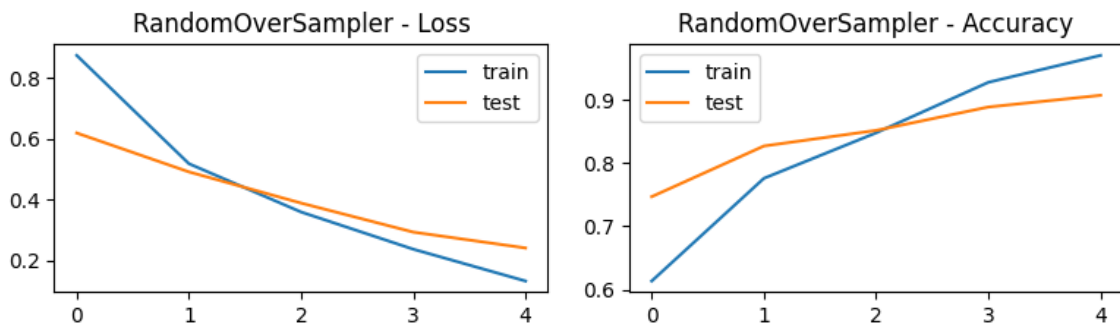
Os gráficos com as curvas de *accuracy* e *loss* para cada uma das implementações dos balanceadores de dados para a classificação, pode ser consultada no apêndice 5.

Também para a análise de sentimento, o melhor resultado foi obtido quando se fez uso da técnica ROS.

**Tabela 24.** Melhores resultados para sentimento – LSTM

Equilibrador	Accuracy	Precision	Recall	F1 - Score	Loss
ROS	90,74	92,01	90,74	90,38	0.24
	84,62	71,6	84,62	77,56	0.55
FS	84,62	72,72	84,62	78,22	0.61
ADASYN	72,97	71,72	72,97	71,67	0.64
SMOTE	56,17	56,02	56,17	55,85	1.07

Os valores das métricas presentes da tabela 24 para o melhor resultado são 90.74, 92.01, 90.74, 90.38 e 0.24, para *accuracy*, *precision*, *recall*, *f1 – score* e *loss*, respetivamente.



**Figura 9.** Loss e accuracy - sentimento

Na figura 9 são apresentados os gráficos de *loss* e *accuracy* por cada dos *epochs*, com os dados de treino e teste.

Todos os outros estão compreendidos entre 82.62 e 56.17 para *accuracy*, 71.6 e 56.02 para *precision*, 84.62 e 56.17 para *recall*, 77.56 e 55.85 para *f1 – score* e 0.24 e 1.07 para *loss*.

Os gráficos com as curvas de *accuracy* e *loss* para cada uma das implementações dos balanceadores de dados para a análise de sentimento, pode ser consultada no apêndice 6.

## Global

A nível global, o melhor resultado para a classificação foi obtido pelos algoritmos RF e AB, com o uso da técnica de balanceamento ROS, com a extração de atributos a ser realizada através de TF-IDF e sem CV.

**Tabela 25.** 10 melhores resultados para classificação – GLOBAL

Algoritmo	Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
RF	ROS	TF-IDF	False	100	100	100	100
AB	ROS	TF-IDF	False	100	100	100	100
AB	ADASYN	TF-IDF	True	98,9	98,93	98,9	98,9
AB		TF-IDF	False	98,46	98,51	98,46	98,22
AB	FS	TF-IDF	False	98,46	98,51	98,46	98,22
DT	ADASYN	TF-IDF	True	97,8	97,92	97,8	97,79
DT	ROS	TF-IDF	False	97,78	97,92	97,78	97,78
RF	SMOTE	TF-IDF	False	97,78	97,9	97,78	97,77
DT		TF-IDF	False	96,92	96,92	96,92	96,92
DT		TF-IDF	True	96,92	94,03	96,92	95,43

Na tabela 25, é possível constatar que os dois melhores resultados tiveram 100 para cada uma das métricas usadas.

Em termos de algoritmos, os mais representados são AB e DT, com quatro entradas e RF com duas. As técnicas de balanceamento dividem-se por ROS e sem implementação, com três entradas, ADASYN com duas e SMOTE com uma.

O método de extração de atributos usado foi sempre TF-IDF e o processo de CV foi usado em três registos apenas, posições 3, 6 e 10.

De referir, que todas as métricas usadas devolvem valores superiores a 95%, sendo que o valor mais baixo é obtido pela métrica f1 – score, para o algoritmo DF, sem balanceador, como extrator TF-IDF e com CV.

Para a análise de sentimento, os melhores resultados são obtidos pelo algoritmo RF, com o balanceamento a cargo de ROS e SMOTE, a extração de atributos feita através de TF-IDF e sem CV.

**Tabela 26.** 10 melhores resultados para sentimento – GLOBAL

Algoritmo	Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
RF	ROS	TF-IDF	False	99,38	99,39	99,38	99,38
RF	SMOTE	TF-IDF	False	99,38	99,39	99,38	99,38
NB	ROS	TF-IDF	False	97,53	97,62	97,53	97,51
NB	SMOTE	TF-IDF	False	97,53	97,62	97,53	97,51
DT	SMOTE	TF-IDF	False	95,68	95,69	95,68	95,68
RF	ADASYN	TF-IDF	False	94,74	93,27	94,74	93,51
DT	ROS	TF-IDF	False	94,44	94,93	94,44	94,33
RF	ROS	TF-IDF	True	93,21	93,56	93,21	93,23
RF	SMOTE	TF-IDF	True	91,98	92,88	91,98	92,06
NB	SMOTE	TF-IDF	True	91,36	92,38	91,36	91,02

De acordo com o apresentado na tabela 26, os valores das métricas para os dois melhores resultados são 99.38, 99.39, 99.38 e 99.38 para *accuracy*, *precision*, *recall* e *f1 – score*, respectivamente.

Para esta implementação, existe uma predominância do algoritmo RF com cinco entradas, sendo que as restantes estão divididas por NB, com três e DT com duas.

Relativamente à técnica de balanceamento de dados mais cotada, é a SMOTE que tem maior número de entradas, com cinco, seguida pelo ROS, com quatro e ADASYN com uma.

Semelhante ao processo de classificação, também aqui, o extrator usado nos melhores resultados é TF-IDF, assim como, o método de CV é usado em três entradas. No entanto, nestes resultados, as suas posições são diferentes, 8, 9 e 10.

Por fim, todas as métricas aqui apresentadas, têm valores superiores a 90%, com o valor mais baixo, mais uma vez, à semelhança da classificação, a ser a métrica *f1 – score*, com 91.02%.

## DISCUSSÃO DOS RESULTADOS

Os sistemas de apoio técnico inteligentes surgem da necessidade de otimizar e melhorar o processo de assistência dado aos clientes. Na sua maioria, são usados nas empresas de TI para permitir uma maior celeridade no processo de resposta aos pedidos dos clientes, vulgo tickets (Jordán et al., 2013; Kaewnoo & Senivongse, 2019; Molino et al., 2018; Ohata et al., 2022).

Outra finalidade dos sistemas inteligentes, passa pela implementação de assistentes virtuais, *chatbots*, que servem de primeira linha de contato com os clientes no processo de assistência (Nayak et al., 2021).

Por fim, temos ainda os sistemas já existentes de CRM, os quais, têm implementados sistemas de AI, cujo objetivo, é facilitar a interação com o cliente. Estes sistemas podem ainda, analisar o histórico das interações entre os técnicos e os clientes, de forma a fazerem sugestões com propostas de resolução dos problemas reportados (Chen, 2012). Outra forma de recolher e analisar os dados, passa pela captação de comunicações voz, para este poder ser analisado posteriormente (Lohse et al., 2015). Todo este processo, pode posteriormente, ser usado para a análise da qualidade do atendimento prestado (Xin et al., 2015).

Á semelhança dos trabalhos analisados, também aqui foram usados diversos algoritmos de AI, quer para o processo de classificação de textos, quer para o processo de análise de sentimento.

**Tabela 27.** Algoritmos usados, artigos vs tese

Artigo	Algoritmo	Usado Neste Trabalho
(Gaut et al., 2018; Kaewnoo & Senivongse, 2019; Lohse et al., 2015; Molino et al., 2018; Ohata et al., 2022)	RF	Sim
(Gaut et al., 2018; Kaewnoo & Senivongse, 2019)	LR	Não
(Kaewnoo & Senivongse, 2019; Nayak et al., 2021; Ohata et al., 2022)	NB	Sim, variante Gaussian
(Nayak et al., 2021; Ohata et al., 2022)	SVM	Não
(Xin et al., 2015)	LASSO Regression	Não
(Lohse et al., 2015)	probabilistic topic modeling	Não
(Molino et al., 2018)	Encoder-Combiner-Decoder	Não
(Kaewnoo & Senivongse, 2019)	Linear SVC	Não

(Nayak et al., 2021)	ensemble DT models - bagging/boosting	Não
(Ohata et al., 2022)	MLP	Não
(Ohata et al., 2022)	XGB	Não
(Ohata et al., 2022)	KNN	Sim

Além dos algoritmos assinalados na tabela 27, a qual compara os algoritmos usados pelos diversos trabalhos analisados, com os que foram usados neste trabalho, foram ainda implementados os algoritmos RT, AD e LSTM. Todos eles, foram implementados na classificação e análise de sentimento.

## Dados

Os dados usados para este trabalho, têm como grande lacuna, as poucas observações, assim como, uma distribuição enviesada. Estes foram obtidos através de uma empresa de consultoria informática e representam emails que foram recebidos e registados em base de dados, por essa empresa. Como estes dados, na sua forma original, contem informação que pode identificar pessoas e empresas, foi necessário fazer a sua anonimização.

Dos trabalhos analisados, a origem dos dados destes é diversa, sendo que na sua grande maioria, são provenientes, à semelhança dos usados neste trabalho, de sistemas de apoio técnico (Kaewnoo & Senivongse, 2019; Molino et al., 2018; Nadaf & Patil, 2015; Ohata et al., 2022; Parmar et al., 2018; Ricciardi Celsi et al., 2021). Os restantes, referem-se a dados de opiniões (Ceyhan et al., 2017; Etaiwi & Naymat, 2017; Laryea et al., 2015; McGuire & Kampf, 2015), email (Pandey et al., 2020), notícias (Sunagar et al., 2021) e posts de redes sociais (PhD Research Student, MUIT, Lucknow. et al., 2019).

Estes têm como origem diversas fontes, plataformas de dados (Ebadi et al., 2019; Nadaf & Patil, 2015; Pandey et al., 2020; Sunagar et al., 2021), plataformas de comercio (Etaiwi & Naymat, 2017; Laryea et al., 2015; McGuire & Kampf, 2015), sistemas próprios (Ceyhan et al., 2017; Gaut et al., 2018; Kaewnoo & Senivongse, 2019; Molino et al., 2018; Nayak et al., 2021; Ohata et al., 2022; Parmar et al., 2018; Ricciardi Celsi et al., 2021) e redes sociais (PhD Research Student, MUIT, Lucknow. et al., 2019; Pugsee et al., 2019; Ranganathan & Tzacheva, 2019).

## Pré-Processamento e Extração de Atributos

Para a primeira parte, o pré-processamento, os métodos usados vão de encontro aos descritos nos trabalhos analisados. O processo de limpeza do texto (Etaiwi & Naymat, 2017; Kaewnoo & Senivongse, 2019; McGuire & Kampf, 2015; Molino et al., 2018; Nayak et al., 2021; Ohata et al., 2022; PhD Research Student, MUIT, Lucknow. et al., 2019; Sunagar et al., 2021), através do qual se fez a limpeza de caracteres especiais, espaços, pontuação e conversão para minúsculas, tokenização e remoção de *stopwords*, é semelhante ao implementado neste trabalho.

**Tabela 28.** *Resumo outas técnicas de tratamento de texto*

Artigo	Stemming	Lemmatization	POS
(Etaiwi & Naymat, 2017)	Sim		Sim
(Molino et al., 2018)		Sim	
(Gaut et al., 2018)	Sim		
(Kaewnoo & Senivongse, 2019)	Sim		
(Nayak et al., 2021)			Sim
(Sunagar et al., 2021)	Sim		
(Ohata et al., 2022)		Sim	
(McGuire & Kampf, 2015)	Sim		
(PhD Research Student, MUIT, Lucknow. et al., 2019)			Sim

No entanto, além das técnicas descritas anteriormente, foram usadas outras que não foram implementadas neste trabalho, como é o caso do *stemming*, *lemmatization* e POS, conforme listado na tabela 28.

Relativamente à extração de atributos, dos métodos implementados, o que está mais representado nos trabalhos analisados é TF-IDF (Gaut et al., 2018; Kaewnoo & Senivongse, 2019; McGuire & Kampf, 2015; Molino et al., 2018; Nadaf & Patil, 2015; Nayak et al., 2021; Ohata et al., 2022; Parmar et al., 2018; Sunagar et al., 2021). Outros usam outras técnicas, como a frequência do termo (TF) (Ceyhan et al., 2017; Etaiwi & Naymat, 2017). As técnicas de extração de atributos do tipo *word vectors*, como Word2Vec ou Doc2Vec, estão presentes apenas estão representadas em um artigo (Ohata et al., 2022).

### **Classificação de Texto**

Nadaf & Patil (2015) fazem uso de vários algoritmos de *clustering*, testando cada um com diversos valores para o número de clusters, valor de k. No entanto, a definição do valor de k

deveria ter sido feita com números ímpares, para evitar que um registro possa ter valores iguais para mais que um cluster.

Etaiwi & Naymat (2017) implementam três algoritmos que são também usados neste trabalho, NB, RF e DT. Os seus resultados são significativamente mais baixos que os obtidos aqui, com a métrica *precision*, para cada uma das suas implantações a ser no máximo 52.8, 73.7 e 52, para NB, RF e DT, respetivamente. No entanto, o algoritmo NB teve valores mais elevados, quer para *accuracy*, quer para *recall*, com 92.6 e 88.7, respetivamente, semelhantes aos obtidos aqui. O conjunto de dados usado tem cerca de quatro vezes mais observações, com 800. Já o processo de tratamento de texto e extração de atributos realizado foi semelhante ao deste trabalho, com mais alguns passos.

Molino et al. (2018) faz o tratamento do *dataset* da mesma forma, com a adição da *lemma*. A extração de atributos é feita com recuso a TF-IDF com LSA. Os seus resultados, quer para RF, quer para as redes neuronais, é bastante inferior ao obtido, com 45.5 e 53.12, para cada uma.

Gaut et al. (2018) implementa algoritmos de RF e LR para análise de um *dataset* com 69402 observações. Aqui os parâmetros usados na definição do de RF são diferentes. Estes aplicam  $n\_estimators = 3000$  e  $max\_features = \log 2$ . O seu processo de tratamento dos dados é semelhante com a adição do processo de *stemming*, assim como a técnica de extração de atributos. O autor refere que o melhor desempenho foi obtido pela RF, à semelhança deste trabalho, no entanto não refere as métricas.

Parmar et al. (2018) utiliza vários algoritmos para realizar a classificação de texto. Estes são aplicados a um *dataset* com 5500 observações, em que não é referido se lhe é aplicado algum tratamento e os atributos são extraídos via TF-IDF. Os resultados obtidos são muitos inferiores aos deste trabalho, com o melhor desempenho a ser obtido pelo SVM, com 60.76 para f1 – score. Não é referido quais são os parâmetros usados em cada um dos algoritmos.

Kaewnoo & Senivongse (2019) usa como *dataset* um conjunto de pedidos de assistência com 1884 e com várias categorias. O seu tratamento teve apenas mais uma fase, o processo de *stemming* e a extração de atributos foi feita com TF-IDF. Este trabalho, á semelhança desta tese, também fez uso de técnicas de balanceamento de dados, nomeadamente, o uso de SMOTE e implementa CV com o valor de 10. Os resultados obtidos estão de acordo com os obtidos para NB, se bem que aqui foi usada a variante Multinomial. Para o RF, os resultados foram inferiores, com o valor de f1 – score a ser no máximo 74.19.

Pandey et al. (2020) propõe um modelo híbrido para a detecção de e-mail de *phishing*, composto por uma combinação de RF e SVM. O *dataset* é multilabel e contém 1353 observações e foi dividido em 60/40, treino e teste, respectivamente. O autor não refere se foi realizado algum tratamento dos dados, nem refere se foi usada alguma técnica de extração de atributos. Os resultados obtidos para RF, f1 – score, são inferiores. Já o modelo proposto, tem uma avaliação que fica no top 30 dos resultados obtido nesta tese.

Ricciardi Celsi et al. (2021) usam um *dataset* de pedidos de assistência, com 307 atributos com 600000 observações. O processo de tratamento de dados apenas refere que estes foram tratados, não especificando quais os passos utilizados. Também não é indicado qual o método de extração de atributos usado. Os resultados obtidos tiveram melhor performance com um *dataset* reduzido, mas ainda assim, com valores inferiores aos desta tese, com exceção das redes neurais, com ANN a ter 79% o que vai de encontro aos obtidos pela rede neuronal usada, LSTM.

Sunagar et al. (2021) analisa um *dataset* com 31900 observações, com vários *labels*. Estes são divididos em treino e teste e são tratados da mesma forma, com a adição do processo de *stemming*. A extração de atributos é realizada através de TF-IDF. Os resultados obtidos para NB foram melhores, com uma diferença de 15%, assim como o KNN, que obteve 81% de f1 – score.

Ohata et al. (2022) tem com base um *dataset* de pedidos de assistência resolvidos com 1049 observações com várias categorias. Aqui o processo de tratamento dos dados é semelhante, com a adição do passo de *lemming*. O processo de extração de atributos faz uso das mesmas técnicas implantadas nesta tese, com a adição do BERT. A divisão de dados em treino e teste, faz-se 80/20. O algoritmo NB foi usado na sua variante Multnomial e obteve valores inferiores, com a sua pontuação máxima a ser 51.18 para f1 – score, com meta dados + TF-IDF com extrator. Para KNN, houve a utilização de vários valores de k, todos eles ímpares, sem serem apresentados os resultados por cada um deles. Desta feita, o valor médio de f1 – score é 63.41, o que fica dentro da média, quando comparado com os valores obtidos nesta tese. Para o RF, é apresentado o valor para *n\_estimators*, que varia entre 20 e 2000. Aqui também não é apresentado o valor do resultado por cada um. O valor obtido é inferior em aproximadamente 30% com os desta tese.

**Tabela 29.** *Resumo melhores resultados - classificação*

Fonte	Extrator	Balanceamento	Algoritmo	Accuracy	F1-
-------	----------	---------------	-----------	----------	-----

					Score
Etaiwi & Naymat (2017)	BoW		NB	88,7%	
Molino et al. (2018)	TF-IDF			53,12%	
Gaut et al. (2018)	TF-IDF		SVM	63,02%	60,76%
Kaewnoo & Senivongse (2019)	TF-IDF	SMOTE	LinearSVM	91,79%	91,69%
Pandey et al. (2020)			Modelo Híbrido (RF + SVM)	94,74%	91,75%
Ricciardi Celsi et al. (2021)			LR	81%	
Sunagar et al. (2021)	TF-IDF		SVM	90,85%	91%
Ohata et al. (2022)	TF-IDF + Metadata		RF	72,67%	69,18%
Métodos Analisados	TF-IDF	ROS	RF e ADB	100%	100%
Métodos Analisados cont.	TF-IDF	ADASYN	ADB com CV	98,9%	98,90%

Na tabela 29, é apresentado um resumo dos principais resultados dos algoritmos analisados na revisão de literatura e dos algoritmos implementados neste trabalho, para o processo de classificação. De acordo com os dados apresentados, os métodos analisados por esta tese, têm um desempenho superior a todos os outros. Os dados usados, com recurso a técnicas de balanceamento, conseguiram uma pontuação de 100%, quer para *accuracy*, quer para *f1 - score*.

### Análise de Sentimento

Laryea et al. (2015) analisa o sentimento de opiniões sobre produtos num site de comercio digital. Estes têm 43227 observações divididas por 5 classes. Estes dados foram tratados de forma ligeiramente diferente dos desta tese, com a conversão para minúsculas e a utilização do POS. A extração de atributos também fez uso de uma técnica diferente, com recurso LPSBD, sendo depois usado LPC para fazer a análise de sentimento. O resultado obtido, cerca de 70% para *accuracy*, encaixa-se nos últimos dez resultados obtidos nesta tese.

Ahmad & Laroche (2017) faz a análise não do sentimento da frase diretamente, mas de quais são os termos mais importantes nas opiniões negativas e nas positivas. O *dataset* é composto por 405 observações, divididas em duas classes, positivo e negativo, menos uma que esta tese. O processo de tratamento de texto teve os mesmos passos, com a inclusão do

processo de *stemming*. A extração ficou a cargo de TF-IDF e foi ainda usado SVD para obter os termos mais importantes.

Ceyhan et al. (2017) tem como *dataset* um conjunto de opiniões com 2018 observações, dividido em positivo e negativo. Aqui o processo de tratamento dos dados foi diferente do usado neste trabalho, com este a ser realizado com a análise morfológica e a desambiguação das palavras. Também a extração de atributos foi realizada de forma distinta com recurso a Binary List, Frequency List, Binary Words e Words Frequencies. Os resultados aqui obtidos, em relação a *accuracy*, para NB são cerca de 10% inferiores e para DT são ligeiramente inferiores com uma diferença de aproximadamente 5%.

PhD Research Student, MUIT, Lucknow. et al, (2019), Pugsee et al., (2019) e Ranganathan & Tzacheva (2019) os dados são obtidos de tweets publicados no Twitter. Apenas no terceiro é referido o número de observações, 520000. Os métodos usados para tratamento dos textos estão de acordo com os usado nesta tese, com a adição de *stemming* e POS. Já a extração de atributos é realizada de forma distinta deste trabalho, com recurso a dois vetores, 8-feature e BOW, para o primeiro, SentiWordNet e do filtro TweetToSparseFeatureVector presente em Weka Affective Datasets, respetivamente, o segundo e terceiro. Os resultados do primeiro são inferiores para o KNN, com uma diferença de pouco menos de 10%, para o segundo, o NB também teve pior pontuação, com um défice de pouco mais de 10% e para o último o algoritmo implementado SVM, teve 98%. Este não foi implementado aqui, mas o seu resultado compara-se a RF ou NB, com o balanceador ROS e SMOTE, com a extração via TF-IDF, ficando com uma variação de menos de 1%.

**Tabela 30.** *Resumo melhores resultados - sentimento*

Fonte	Extrator	Balanceamento	Algoritmo	Accuracy	F1-Score
Laryea et al. (2015)	LSBD		LPC	70%	
Ceyhan et al. (2017)	Words Frequencies		SMO	95,8%	
PhD Research Student, MUIT, Lucknow. et al, (2019)	BOW		Proposto	82%	
Pugsee et al., (2019)	SentiWordNet		SVM com CV (10)	92%	
Ranganathan & Tzacheva (2019)	TweetToSparse FeatureVector (Weka Affective tweets)		LibLinear SVM	98%	

Métodos Analisados	TF-IDF	ROS e SMOTE	RF	99,38%	99,38%
Métodos Analisados cont.	TF-IDF	ROS e SMOTE	NB	97,53%	97,51%

Na tabela 30, é possível verificar que, à semelhança da classificação, também as metodologias analisadas por esta tese, tiveram melhores resultados. No entanto, houve alguns trabalhos analisados que tiveram valores aproximados aos desta, mas com métodos algoritmos distintos.

### Sistemas Propostos

Apesar de não ter sido implementado um sistema final, a sugestão de um sistema seria dependente da sua finalidade. Desta forma, as tecnologias usadas variam de acordo com a sua implementação. Se for uma plataforma independente (Jordán et al., 2013; Kaewnoo & Senivongse, 2019; Laryea et al., 2015; Moharil et al., 2020; Pugsee et al., 2019; Senthurvelautham & Hettiarachchi, 2018), pode ser usado qualquer tipo de tecnologias de forma a satisfazer essa necessidade. No entanto, este sistema também pode ser integrado nos softwares já existentes (Chen, 2012; Gaut et al., 2018; Lohse et al., 2015; Molino et al., 2018; Nayak et al., 2021; Xin et al., 2015) ou então, poderá ainda ser usado uma plataforma já desenvolvida que permite o uso de algoritmos de IA (Ebadi et al., 2019).

Aqui, o principal seria a gestão da prioridade dos novos pedidos. Complementar a esta, seria a análise do sentimento, quer dos novos pedidos, quer das comunicações consequentes.

**Tabela 31.** *Matriz avaliação novos pedidos*

Positivo	Neutro	Negativo	Importância Cliente
-1	0	+1	Valor numérico, por exemplo, de 1 a 5.

Na tabela 31, está exemplificada uma matriz que será usada para a avaliação de novos pedidos. Por exemplo, um cliente com uma importância alta, digamos 4, com uma mensagem negativa, teria uma prioridade maior,  $4 + 1 = 5$ . Já um cliente com uma baixa importância, com 3 e com uma mensagem neutra, já teria uma prioridade menor,  $3 + 0 = 3$ . Para complementar esta fórmula, por cada de passe sem ser possível tratar o pedido em aberto, é acrescentado um ponto.

Além da análise aos novos pedidos, também as comunicações subsequentes realizadas entre o técnico e o cliente, também seriam avaliadas de acordo com o seu sentimento. No caso de as mensagens serem conseqüentemente negativas, são adicionados novos pontos à métrica anterior. Esta parte, poderá ser utilizada para fazer uma avaliação à interação dos técnicos com os clientes (Senthurvelautham & Hettiarachchi, 2018).

### **Limitações**

Para este projeto, a grande limitação passou pela pouca quantidade de dados para analisar. Como são dados reais, com informação identificadora das empresas, pessoas e documentos, aumenta a dificuldade ao seu acesso. Como estes tiveram de ser anonimizados, processo que requer muito tempo, só foi possível obter uma pequena amostra do que seria desejável.

Outra grande limitação, passou pela impossibilidade em encontrar sistemas que apliquem análise de sentimento ao processo de assistência técnica. Desta forma, este estudo fica limitado e compara os resultados desta parte com análises realizadas sobre conteúdos de redes sociais ou opiniões sobre produtos.

## CONCLUSÕES

Atualmente a quantidade de correspondência eletrônica gerada é em grande quantidade, sobretudo, no relacionamento com os clientes. Isto faz com que a análise destes dados possa ser um processo moroso e com propensão a falhas ou esquecimentos.

É aqui que entram os sistemas de apoio inteligentes, os quais podem ter um papel preponderante na gestão desta correspondência. Ao analisar os dados de forma automática, podem no imediato, separar os dados que são relevantes ou importantes dos restantes. Desta forma, poupa-se tempo e reduz-se os erros cometidos.

Este trabalho teve como objetivo principal a análise dos sistemas atuais de forma a perceber quais as suas vantagens e o seu funcionamento. Existem já alguns mecanismos que podem ser usados para esta vertente, mas que ainda são ultrapassados em termos de resultados, pela implementação de técnicas personalizadas.

Os algoritmos usados neste estudo, são na sua maioria de aprendizagem supervisionada, com um destes, a ser do tipo ensemble e outro do tipo rede neuronal. Serviu sobretudo, para comparar a performance dos dois de forma a perceber qual o mais adequado para o problema em estudo. No entanto, este último sai prejudicado devido ao reduzido tamanho do *dataset*, contudo, mostra potencial na área de análise de sentimento, onde obteve melhores resultados do que no processo de classificação.

Os melhores resultados são obtidos com os algoritmos de RF e ADB, os quais superam a concorrência, quer na classificação de texto, quer na análise de sentimento. Neste caso, os dados em análise são balanceados com recurso a várias técnicas. A espaços, existem resultados positivos, em que não foi aplicada esta técnica. O mesmo se passa com a CV. Na sua maioria não foi implementada, mas a espaços aparecem resultados altos com esta técnica. Relativamente ao extrator de atributos, as posições cimeiras são dominadas pelo TF-IDF, sendo que o primeiro dos outros a ser contabilizado é Word2Vec com *skip-gram* e CBOW.

Alguns dos sistemas propostos integram estes processos através de uma plataforma independente. No entanto, ainda há algumas plataformas que permitem a sua integração diretamente com elas.

Estes mecanismos são avaliados na sua maioria através de uma matriz de confusão. Esta mostra a eficiência dos resultados dos modelos gerados, contabilizando os resultados corretamente classificados. Aqui as métricas mais importantes são o *accuracy*, que avalia os

dados que foram efetivamente bem classificados e o *fl – score*, que faz a avaliação global dos resultados gerados pelo algoritmo.

Ao proceder à classificação dos dados em conjunto com a análise de sentimento, pode ser possível às entidades implementarem uma ordem para a resolução dos pedidos de apoio que recebem. Dada a importância a uma cliente, em conjunto com o sentimento da sua mensagem, permite que calcule um peso, que será usado para determinar a sua posição na fila de prioridades. Desta forma, otimiza se o processo e atendimento de forma a dar resposta aos pedidos dos clientes mais importantes com as mensagens mais negativas.

### **Trabalhos Futuros**

Como proposta de trabalho futuro, propunha-se a continuação deste estudo, através a implementação em ambiente de produção, de uma combinação de algoritmos para a classificação e análise de sentimento. Pode ser o mesmo para as duas vertentes, um para cada ou ainda, vários para cada uma delas, de forma a determinar quais os mais adequados.

Por fim, dependendo os sistemas de informação disponíveis, caso estes permitissem a integração de novas funcionalidades, podia ser realizada a integração de um sistema para estes processos. No caso de não ser possível, realizava-se o desenvolvimento de uma plataforma independente para a gestão da correspondência e respetiva determinação da fila de prioridades.

## BIBLIOGRAFIA

- Ahmad, S. N., & Laroche, M. (2017). Analyzing electronic word of mouth: A social commerce construct. *International Journal of Information Management*, 37(3), Artigo 3. <https://doi.org/10.1016/j.ijinfomgt.2016.08.004>
- Ceyhan, M., Orhan, Z., & Domnori, E. (2017). Health service quality measurement from patient reviews in Turkish by opinion mining. Em A. Badnjevic (Ed.), *CMBEBIH 2017* (Vol. 62, pp. 649–653). Springer Singapore. [https://doi.org/10.1007/978-981-10-4166-2\\_97](https://doi.org/10.1007/978-981-10-4166-2_97)
- Chen, T. F. (2012). Applying Artificial Intelligence in CRM: Case Studies of Intelligent Virtual Agents and Pegasystems. *Applied Mechanics and Materials*, 182–183, 878–882. <https://doi.org/10.4028/www.scientific.net/AMM.182-183.878>
- Ebadi, A., Gauthier, Y., Tremblay, S., & Paul, P. (2019). How can Automated Machine Learning Help Business Data Science Teams? *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, 1186–1191. <https://doi.org/10.1109/ICMLA.2019.00196>
- Etaiwi, W., & Naymat, G. (2017). The Impact of applying Different Preprocessing Steps on Review Spam Detection. *Procedia Computer Science*, 113, 273–279. <https://doi.org/10.1016/j.procs.2017.08.368>
- Gaut, G., Navarrete, A., Wahedi, L., van der Boor, P., De Unánue, A., Díaz, J., Clark, E., & Ghani, R. (2018). Improving Government Response to Citizen Requests Online. *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, 1–10. <https://doi.org/10.1145/3209811.3209872>
- Gupta, S. (2020). A Hybrid Machine Learning Framework of Gradient Boosting Decision Tree and Sequence Model for Predicting Escalation in Customer Support. *2020 IEEE International Conference on Big Data (Big Data)*, 5511–5518. <https://doi.org/10.1109/BigData50022.2020.9377831>
- Jordán, J., Heras, S., Valero, S., & Julián, V. (2013). ArgCBR-CallCentre: A Call Centre Based on CBR Argumentative Agents. Em Y. Demazeau, T. Ishida, J. M. Corchado, & J. Bajo (Eds.), *Advances on Practical Applications of Agents and Multi-Agent Systems* (Vol. 7879, pp. 292–295). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-38073-0\\_30](https://doi.org/10.1007/978-3-642-38073-0_30)

- Kaewnoo, P., & Senivongse, T. (2019). Identification of Software Problem Report Types Using Multiclass Classification. *Proceedings of the 2019 3rd International Conference on Software and E-Business*, 104–109. <https://doi.org/10.1145/3374549.3374570>
- Laryea, B. N. L., Choi, C.-H., Jung, I.-S., Lee, K.-H., & Cho, W.-S. (2015). Web Application for Sentiment Analysis Using Supervised Machine Learning. *International Journal of Software Engineering and Its Applications*, 9(1), Artigo 1.
- Lohse, J., Sanati-Mehrizi, R., & Minaie, A. (2015). Data Mining in Call Centers: The Overlooked Interaction between Employees. *2015 ASEE Annual Conference and Exposition Proceedings*, 26.439.1-26.439.12. <https://doi.org/10.18260/p.23778>
- McGuire, M., & Kampf, C. (2015). Using social media sentiment analysis for interaction design choices: An exploratory framework. *Proceedings of the 33rd Annual International Conference on the Design of Communication*, 1–7. <https://doi.org/10.1145/2775441.2775472>
- Moharil, A., Singh, S., Dravid, Y., Dharap, H., & Bhanuse, V. (2020). Integrated Feedback Analysis And Moderation Platform Using Natural Language Processing. *2020 Fourth International Conference on Inventive Systems and Control (ICISC)*, 872–877. <https://doi.org/10.1109/ICISC47916.2020.9171202>
- Molino, P., Zheng, H., & Wang, Y.-C. (2018). COTA: Improving the Speed and Accuracy of Customer Support through Ranking and Deep Networks. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 586–595. <https://doi.org/10.1145/3219819.3219851>
- Nadaf, M. A., & Patil, S. S. (2015). Performance evaluation of categorizing technical support requests using advanced K-Means algorithm. *2015 IEEE International Advance Computing Conference (IACC)*, 409–414. <https://doi.org/10.1109/IADCC.2015.7154740>
- Nayak, S. P., Rai, A., Vankataramanappa, K., Pansuriya, J. A., & Singler, J. (2021). Adaptive MoD Chatbot: Toward Providing Contextual Corporate Summarized Document as Suggestions and Reported Issue Ticket Routing. Em N. N. Chiplunkar & T. Fukao (Eds.), *Advances in Artificial Intelligence and Data Engineering* (Vol. 1133, pp. 659–674). Springer Singapore. [https://doi.org/10.1007/978-981-15-3514-7\\_50](https://doi.org/10.1007/978-981-15-3514-7_50)

- Ohata, E. F., Mattos, C. L. C., Gomes, S. L., Reboucas, E. D. S., & Rego, P. A. L. (2022). A Text Classification Methodology to Assist a Large Technical Support System. *IEEE Access*, *10*, 108413–108421. <https://doi.org/10.1109/ACCESS.2022.3213033>
- Padmavathy, P., & Pakkir Mohideen, S. (2020). An efficient two-pass classifier system for patient opinion mining to analyze drugs satisfaction. *Biomedical Signal Processing and Control*, *57*, 101755. <https://doi.org/10.1016/j.bspc.2019.101755>
- Pandey, A., Gill, N., Sai Prasad Nadendla, K., & Thaseen, I. S. (2020). Identification of Phishing Attack in Websites Using Random Forest-SVM Hybrid Model. Em A. Abraham, A. K. Cherukuri, P. Melin, & N. Gandhi (Eds.), *Intelligent Systems Design and Applications* (Vol. 941, pp. 120–128). Springer International Publishing. [https://doi.org/10.1007/978-3-030-16660-1\\_12](https://doi.org/10.1007/978-3-030-16660-1_12)
- Parmar, P. S., Biju, P. K., Shankar, M., & Kadiresan, N. (2018). Multiclass Text Classification and Analytics for Improving Customer Support Response through different Classifiers. *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 538–542. <https://doi.org/10.1109/ICACCI.2018.8554881>
- Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide* (1.<sup>a</sup> ed.). Wiley. <https://doi.org/10.1002/9780470754887>
- PhD Research Student, MUIT, Lucknow., Raut, A., Pandey, Dr. R. K., & Research Guide, MUIT, Lucknow. (2019). Sentiment Analysis using Optimized Feature Sets in Different Twitter Dataset Domains. *International Journal of Innovative Technology and Exploring Engineering*, *8*(11), Artigo 11. <https://doi.org/10.35940/ijitee.K2195.0981119>
- Pugsee, P., Nussiri, V., & Kittirungruang, W. (2019). Opinion Mining for Skin Care Products on Twitter. Em B. W. Yap, A. H. Mohamed, & M. W. Berry (Eds.), *Soft Computing in Data Science* (Vol. 937, pp. 261–271). Springer Singapore. [https://doi.org/10.1007/978-981-13-3441-2\\_20](https://doi.org/10.1007/978-981-13-3441-2_20)
- Ranganathan, J., & Tzacheva, A. (2019). Emotion Mining in Social Media Data. *Procedia Computer Science*, *159*, 58–66. <https://doi.org/10.1016/j.procs.2019.09.160>
- Ricciardi Celsi, L., Caliciotti, A., D’Onorio, M., Scocchi, E., Sulieman, N. A., & Villari, M. (2021). On Predicting Ticket Reopening for Improving Customer Service in 5G Fiber

Optic Networks. *Future Internet*, 13(10), Artigo 10.  
<https://doi.org/10.3390/fi13100259>

Senthurvelautham, S., & Hettiarachchi, S. (2018). SentScore: Autonomous Text Sentiment scoring and Summarizing System related to Complaint Management. *2018 4th International Conference for Convergence in Technology (I2CT)*, 1–8.  
<https://doi.org/10.1109/I2CT42659.2018.9058199>

Sunagar, P., Kanavalli, A., Nayak, S. S., Mahan, S. R., Prasad, S., & Prasad, S. (2021). News Topic Classification Using Machine Learning Techniques. Em V. Bindhu, J. M. R. S. Tavares, A.-A. A. Boulogeorgos, & C. Vuppapapati (Eds.), *International Conference on Communication, Computing and Electronics Systems* (Vol. 733, pp. 461–474). Springer Singapore. [https://doi.org/10.1007/978-981-33-4909-4\\_35](https://doi.org/10.1007/978-981-33-4909-4_35)

Xin, Z., Li, F., Li, Q. C., & Sarkar, S. R. (2015). Applying Analytics to Improve Hardware and Software Maintenance Support Services. *2015 IEEE International Conference on Services Computing*, 355–362. <https://doi.org/10.1109/SCC.2015.56>

## APÊNDICES

## Apêndice 1 – Resultados Classificação

Tipo	Algoritmo	Variante	Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
Classificação	RF		ROS	TF-IDF	False	100	100	100	100
Classificação	AB		ROS	TF-IDF	False	100	100	100	100
Classificação	AB		ADASYN	TF-IDF	True	98,9	98,93	98,9	98,9
Classificação	AB			TF-IDF	False	98,46	98,51	98,46	98,22
Classificação	AB		FS	TF-IDF	False	98,46	98,51	98,46	98,22
Classificação	DT		ADASYN	TF-IDF	True	97,8	97,92	97,8	97,79
Classificação	DT		ROS	TF-IDF	False	97,78	97,92	97,78	97,78
Classificação	RF		SMOTE	TF-IDF	False	97,78	97,9	97,78	97,77
Classificação	DT			TF-IDF	False	96,92	96,92	96,92	96,92
Classificação	DT			TF-IDF	True	96,92	94,03	96,92	95,43
Classificação	AB			TF-IDF	True	96,92	94,03	96,92	95,43
Classificação	DT		FS	TF-IDF	False	96,92	96,92	96,92	96,92
Classificação	DT		FS	TF-IDF	True	96,92	94,03	96,92	95,43
Classificação	AB		FS	TF-IDF	True	96,92	94,03	96,92	95,43
Classificação	DT		ADASYN	TF-IDF	False	96,7	96,97	96,7	96,68
Classificação	RF		ADASYN	TF-IDF	False	96,7	96,97	96,7	96,68
Classificação	AB		SMOTE	TF-IDF	True	96,67	96,93	96,67	96,65
Classificação	DT		SMOTE	TF-IDF	False	95,56	96,01	95,56	95,52
Classificação	RF			TF-IDF	False	95,38	95,48	95,38	95,16
Classificação	RF			TF-IDF	True	95,38	92,45	95,38	93,89

Classificação	RF		FS	TF-IDF	False	95,38	95,48	95,38	95,16
Classificação	RF		FS	TF-IDF	True	95,38	92,45	95,38	93,89
Classificação	AB		ADASYN	TF-IDF	False	94,51	94,52	94,51	94,5
Classificação	AB		SMOTE	TF-IDF	False	94,44	94,67	94,44	94,41
Classificação	AB		ROS	TF-IDF	True	93,33	93,41	93,33	93,31
Classificação	DT		SMOTE	TF-IDF	True	93,33	93,73	93,33	93,28
Classificação	RF		ADASYN	TF-IDF	True	91,21	92,88	91,21	91,34
Classificação	DT		ROS	TF-IDF	True	91,11	91,11	91,11	91,11
Classificação	RF		ROS	TF-IDF	True	90	92,05	90	90,1
Classificação	RF		SMOTE	TF-IDF	True	86,67	90,07	86,67	86,76
Classificação	NB	Gaussian		Word2Vec - CBOW	True	84,62	82,15	84,62	83,29
Classificação	NB	Gaussian	FS	Word2Vec - CBOW	True	84,62	82,15	84,62	83,29
Classificação	NB	Gaussian	ROS	Word2Vec - CBOW	True	84,62	82,15	84,62	83,29
Classificação	NB	Gaussian	SMOTE	Word2Vec - CBOW	True	84,62	82,15	84,62	83,29
Classificação	NB	Gaussian	ADASYN	Word2Vec - CBOW	True	84,62	82,15	84,62	83,29
Classificação	NB	Gaussian		Word2Vec - CBOW	False	83,08	83,85	83,08	83,4
Classificação	NB	Gaussian	FS	Word2Vec - CBOW	False	83,08	83,85	83,08	83,4
Classificação	NB	Gaussian	ROS	Word2Vec - CBOW	False	83,08	83,85	83,08	83,4
Classificação	NB	Gaussian	SMOTE	Word2Vec - CBOW	False	83,08	83,36	83,08	82,91
Classificação	NB	Gaussian	ADASYN	Word2Vec - CBOW	False	83,08	83,36	83,08	82,91
Classificação	AB		ROS	Word2Vec - CBOW	False	81,54	79,11	81,54	80,29
Classificação	KNN	Minkowski		Word2Vec - CBOW	True	80	77,54	80	78,75
Classificação	KNN	Euclidean		Word2Vec - CBOW	True	80	77,54	80	78,75
Classificação	KNN	Minkowski	FS	Word2Vec - CBOW	True	80	77,54	80	78,75

Classificação	KNN	Euclidean	FS	Word2Vec - CBOW	True	80	77,54	80	78,75
Classificação	KNN	Minkowski	ROS	Word2Vec - CBOW	True	80	77,54	80	78,75
Classificação	KNN	Euclidean	ROS	Word2Vec - CBOW	True	80	77,54	80	78,75
Classificação	KNN	Minkowski	SMOTE	Word2Vec - CBOW	True	80	77,54	80	78,75
Classificação	KNN	Euclidean	SMOTE	Word2Vec - CBOW	True	80	77,54	80	78,75
Classificação	KNN	Minkowski	ADASYN	Word2Vec - CBOW	True	80	77,54	80	78,75
Classificação	KNN	Euclidean	ADASYN	Word2Vec - CBOW	True	80	77,54	80	78,75
Classificação	AB			Word2Vec - CBOW	False	78,46	76,13	78,46	77,21
Classificação	AB		FS	Word2Vec - CBOW	False	78,46	76,13	78,46	77,21
Classificação	RF		ADASYN	Word2Vec - CBOW	False	78,46	78,88	78,46	78,29
Classificação	RF		ROS	Doc2Vec - DM	False	76,92	77,41	76,92	76,86
Classificação	AB			Word2Vec - skip - gram	False	76,92	74,77	76,92	75,76
Classificação	AB		FS	Word2Vec - skip - gram	False	76,92	74,77	76,92	75,76
Classificação	RF		ROS	Word2Vec - skip - gram	False	76,92	78,51	76,92	76,39
Classificação	AB		ROS	Word2Vec - skip - gram	False	76,92	75,19	76,92	75,75
Classificação	RF		ADASYN	Word2Vec - skip - gram	False	76,92	77,89	76,92	76,44
Classificação	RF		SMOTE	Word2Vec - CBOW	False	76,92	77,29	76,92	76,8
Classificação	KNN	Minkowski	ROS	TF-IDF	False	75,56	79,67	75,56	74,4
Classificação	KNN	Euclidean	ROS	TF-IDF	False	75,56	79,67	75,56	74,4
Classificação	RF			Doc2Vec - DM	False	75,38	77,96	75,38	75,79
Classificação	RF		FS	Doc2Vec - DM	False	75,38	77,96	75,38	75,79
Classificação	RF			Word2Vec - skip - gram	False	75,38	74,56	75,38	73,48
Classificação	RF			Word2Vec - skip - gram	True	75,38	73,03	75,38	74,17
Classificação	RF		FS	Word2Vec - skip - gram	False	75,38	74,56	75,38	73,48

Classificação	RF		FS	Word2Vec - skip - gram	True	75,38	73,03	75,38	74,17
Classificação	RF		ROS	Word2Vec - skip - gram	True	75,38	73,03	75,38	74,17
Classificação	RF		SMOTE	Word2Vec - skip - gram	True	75,38	73,03	75,38	74,17
Classificação	RF		ADASYN	Word2Vec - skip - gram	True	75,38	73,03	75,38	74,17
Classificação	RF			Word2Vec - CBOW	False	75,38	73,1	75,38	74,07
Classificação	RF		FS	Word2Vec - CBOW	False	75,38	73,1	75,38	74,07
Classificação	AB		SMOTE	Word2Vec - CBOW	False	75,38	73,14	75,38	74,17
Classificação	AB		ADASYN	Word2Vec - CBOW	False	75,38	76,34	75,38	74,79
Classificação	KNN	Minkowski	ROS	TF-IDF	True	74,44	75,03	74,44	74,06
Classificação	KNN	Euclidean	ROS	TF-IDF	True	74,44	75,03	74,44	74,06
Classificação	RF			Doc2Vec - DM	True	73,85	71,58	73,85	72,68
Classificação	RF		FS	Doc2Vec - DM	True	73,85	71,58	73,85	72,68
Classificação	RF		ROS	Doc2Vec - DM	True	73,85	71,58	73,85	72,68
Classificação	RF		SMOTE	Doc2Vec - DM	False	73,85	75,95	73,85	74,43
Classificação	RF		SMOTE	Doc2Vec - DM	True	73,85	71,58	73,85	72,68
Classificação	DT		ADASYN	Doc2Vec - DM	False	73,85	75,21	73,85	74,08
Classificação	RF		ADASYN	Doc2Vec - DM	False	73,85	75,95	73,85	74,43
Classificação	RF		ADASYN	Doc2Vec - DM	True	73,85	71,58	73,85	72,68
Classificação	KNN	Minkowski		Word2Vec - skip - gram	False	73,85	74,28	73,85	73,8
Classificação	KNN	Euclidean		Word2Vec - skip - gram	False	73,85	74,28	73,85	73,8
Classificação	KNN	Minkowski	FS	Word2Vec - skip - gram	False	73,85	74,28	73,85	73,8
Classificação	KNN	Euclidean	FS	Word2Vec - skip - gram	False	73,85	74,28	73,85	73,8
Classificação	RF		SMOTE	Word2Vec - skip - gram	False	73,85	74,64	73,85	73,32
Classificação	AB		SMOTE	Word2Vec - skip - gram	False	73,85	72,16	73,85	72,72

Classificação	AB		ADASYN	Word2Vec - skip - gram	False	73,85	71,78	73,85	72,73
Classificação	RF			Word2Vec - CBOW	True	73,85	71,55	73,85	72,61
Classificação	RF		FS	Word2Vec - CBOW	True	73,85	71,55	73,85	72,61
Classificação	RF		ROS	Word2Vec - CBOW	True	73,85	71,55	73,85	72,61
Classificação	RF		SMOTE	Word2Vec - CBOW	True	73,85	71,55	73,85	72,61
Classificação	RF		ADASYN	Word2Vec - CBOW	True	73,85	71,55	73,85	72,61
Classificação	NB	Gaussian	ADASYN	TF-IDF	False	72,53	73,37	72,53	71,94
Classificação	DT			Doc2Vec - DM	False	72,31	72,43	72,31	72,3
Classificação	DT		FS	Doc2Vec - DM	False	72,31	72,43	72,31	72,3
Classificação	AB			Word2Vec - CBOW	True	72,31	70,91	72,31	71,17
Classificação	AB		FS	Word2Vec - CBOW	True	72,31	70,91	72,31	71,17
Classificação	AB		ROS	Word2Vec - CBOW	True	72,31	70,91	72,31	71,17
Classificação	AB		SMOTE	Word2Vec - CBOW	True	72,31	70,91	72,31	71,17
Classificação	AB		ADASYN	Word2Vec - CBOW	True	72,31	70,91	72,31	71,17
Classificação	NB	Gaussian	ADASYN	TF-IDF	True	71,43	70,6	71,43	70,24
Classificação	RF			Doc2Vec - DBOW	False	70,77	73,93	70,77	71,56
Classificação	RF		FS	Doc2Vec - DBOW	False	70,77	73,93	70,77	71,56
Classificação	KNN	Minkowski		Doc2Vec - DM	False	70,77	73,18	70,77	69,92
Classificação	KNN	Euclidean		Doc2Vec - DM	False	70,77	73,18	70,77	69,92
Classificação	KNN	Minkowski	FS	Doc2Vec - DM	False	70,77	73,18	70,77	69,92
Classificação	KNN	Euclidean	FS	Doc2Vec - DM	False	70,77	73,18	70,77	69,92
Classificação	DT		ROS	Doc2Vec - DM	False	70,77	74,58	70,77	72,26
Classificação	KNN	Minkowski	ROS	Doc2Vec - DM	False	70,77	77,35	70,77	72,01
Classificação	KNN	Euclidean	ROS	Doc2Vec - DM	False	70,77	77,35	70,77	72,01

Classificação	DT		SMOTE	Doc2Vec - DM	False	70,77	73,33	70,77	71,81
Classificação	NB	Gaussian		TF-IDF	True	70,77	71,28	70,77	70,57
Classificação	NB	Gaussian	FS	TF-IDF	True	70,77	71,28	70,77	70,57
Classificação	DT		ROS	Word2Vec - skip - gram	False	70,77	72,57	70,77	71,04
Classificação	DT		ADASYN	Word2Vec - skip - gram	False	70,77	72,62	70,77	71,23
Classificação	RF		ROS	Word2Vec - CBOW	False	70,77	68,79	70,77	69,7
Classificação	RF		ROS	Doc2Vec - DBOW	False	69,23	72,01	69,23	69,95
Classificação	RF		SMOTE	Doc2Vec - DBOW	False	69,23	72,01	69,23	69,95
Classificação	RF		ADASYN	Doc2Vec - DBOW	False	69,23	72,01	69,23	69,95
Classificação	KNN	Minkowski		TF-IDF	True	69,23	69,94	69,23	67,46
Classificação	KNN	Euclidean		TF-IDF	True	69,23	69,94	69,23	67,46
Classificação	KNN	Minkowski	FS	TF-IDF	True	69,23	69,94	69,23	67,46
Classificação	KNN	Euclidean	FS	TF-IDF	True	69,23	69,94	69,23	67,46
Classificação	NB	Gaussian		Word2Vec - skip - gram	True	69,23	67,57	69,23	67,58
Classificação	DT			Word2Vec - skip - gram	True	69,23	68,44	69,23	68,02
Classificação	NB	Gaussian	FS	Word2Vec - skip - gram	True	69,23	67,57	69,23	67,58
Classificação	DT		FS	Word2Vec - skip - gram	True	69,23	68,44	69,23	68,02
Classificação	NB	Gaussian	ROS	Word2Vec - skip - gram	True	69,23	67,57	69,23	67,58
Classificação	DT		ROS	Word2Vec - skip - gram	True	69,23	68,44	69,23	68,02
Classificação	NB	Gaussian	SMOTE	Word2Vec - skip - gram	True	69,23	67,57	69,23	67,58
Classificação	DT		SMOTE	Word2Vec - skip - gram	True	69,23	68,44	69,23	68,02
Classificação	NB	Gaussian	ADASYN	Word2Vec - skip - gram	True	69,23	67,57	69,23	67,58
Classificação	DT		ADASYN	Word2Vec - skip - gram	True	69,23	68,44	69,23	68,02
Classificação	NB	Gaussian	SMOTE	TF-IDF	True	68,89	67,86	68,89	67,25

Classificação	NB	Gaussian	ROS	TF-IDF	False	67,78	68,77	67,78	67,37
Classificação	NB	Gaussian	ROS	TF-IDF	True	67,78	66,85	67,78	66,62
Classificação	NB	Gaussian	SMOTE	TF-IDF	False	67,78	68,77	67,78	67,37
Classificação	DT			Doc2Vec - DBOW	False	67,69	70,61	67,69	68,38
Classificação	DT		FS	Doc2Vec - DBOW	False	67,69	70,61	67,69	68,38
Classificação	DT		ROS	Doc2Vec - DBOW	False	67,69	69,45	67,69	68,01
Classificação	DT		SMOTE	Doc2Vec - DBOW	False	67,69	69,45	67,69	68,01
Classificação	DT		ADASYN	Doc2Vec - DBOW	False	67,69	69,45	67,69	68,01
Classificação	AB			Word2Vec - skip - gram	True	67,69	67,22	67,69	66,41
Classificação	AB		FS	Word2Vec - skip - gram	True	67,69	67,22	67,69	66,41
Classificação	AB		ROS	Word2Vec - skip - gram	True	67,69	67,22	67,69	66,41
Classificação	AB		SMOTE	Word2Vec - skip - gram	True	67,69	67,22	67,69	66,41
Classificação	AB		ADASYN	Word2Vec - skip - gram	True	67,69	67,22	67,69	66,41
Classificação	AB			Doc2Vec - DBOW	False	66,15	67,65	66,15	66,42
Classificação	AB		FS	Doc2Vec - DBOW	False	66,15	67,65	66,15	66,42
Classificação	NB	Gaussian	SMOTE	Doc2Vec - DBOW	False	66,15	66,69	66,15	66,26
Classificação	AB			Doc2Vec - DM	True	66,15	66,68	66,15	66,06
Classificação	AB		FS	Doc2Vec - DM	True	66,15	66,68	66,15	66,06
Classificação	AB		ROS	Doc2Vec - DM	True	66,15	66,68	66,15	66,06
Classificação	KNN	Minkowski	SMOTE	Doc2Vec - DM	False	66,15	75,08	66,15	68,22
Classificação	KNN	Euclidean	SMOTE	Doc2Vec - DM	False	66,15	75,08	66,15	68,22
Classificação	AB		SMOTE	Doc2Vec - DM	True	66,15	66,68	66,15	66,06
Classificação	KNN	Minkowski	ADASYN	Doc2Vec - DM	False	66,15	76	66,15	68,67
Classificação	KNN	Euclidean	ADASYN	Doc2Vec - DM	False	66,15	76	66,15	68,67

Classificação	AB		ADASYN	Doc2Vec - DM	True	66,15	66,68	66,15	66,06
Classificação	DT			Word2Vec - skip - gram	False	66,15	69,85	66,15	67,13
Classificação	KNN	Minkowski		Word2Vec - skip - gram	True	66,15	64,21	66,15	65,1
Classificação	KNN	Euclidean		Word2Vec - skip - gram	True	66,15	64,21	66,15	65,1
Classificação	DT		FS	Word2Vec - skip - gram	False	66,15	69,85	66,15	67,13
Classificação	KNN	Minkowski	FS	Word2Vec - skip - gram	True	66,15	64,21	66,15	65,1
Classificação	KNN	Euclidean	FS	Word2Vec - skip - gram	True	66,15	64,21	66,15	65,1
Classificação	KNN	Minkowski	ROS	Word2Vec - skip - gram	True	66,15	64,21	66,15	65,1
Classificação	KNN	Euclidean	ROS	Word2Vec - skip - gram	True	66,15	64,21	66,15	65,1
Classificação	KNN	Minkowski	SMOTE	Word2Vec - skip - gram	True	66,15	64,21	66,15	65,1
Classificação	KNN	Euclidean	SMOTE	Word2Vec - skip - gram	True	66,15	64,21	66,15	65,1
Classificação	KNN	Minkowski	ADASYN	Word2Vec - skip - gram	True	66,15	64,21	66,15	65,1
Classificação	KNN	Euclidean	ADASYN	Word2Vec - skip - gram	True	66,15	64,21	66,15	65,1
Classificação	NB	Gaussian		Doc2Vec - DBOW	True	64,62	63,2	64,62	63,14
Classificação	NB	Gaussian	FS	Doc2Vec - DBOW	True	64,62	63,2	64,62	63,14
Classificação	NB	Gaussian	ROS	Doc2Vec - DBOW	False	64,62	65,09	64,62	64,76
Classificação	NB	Gaussian	ROS	Doc2Vec - DBOW	True	64,62	63,2	64,62	63,14
Classificação	NB	Gaussian	SMOTE	Doc2Vec - DBOW	True	64,62	63,2	64,62	63,14
Classificação	NB	Gaussian	ADASYN	Doc2Vec - DBOW	False	64,62	66,62	64,62	65,02
Classificação	NB	Gaussian	ADASYN	Doc2Vec - DBOW	True	64,62	63,2	64,62	63,14
Classificação	KNN	Minkowski		TF-IDF	False	64,62	64,89	64,62	63,96
Classificação	KNN	Euclidean		TF-IDF	False	64,62	64,89	64,62	63,96
Classificação	KNN	Minkowski	FS	TF-IDF	False	64,62	64,89	64,62	63,96
Classificação	KNN	Euclidean	FS	TF-IDF	False	64,62	64,89	64,62	63,96

Classificação	DT			Word2Vec - CBOW	False	64,62	67,69	64,62	65,76
Classificação	DT		FS	Word2Vec - CBOW	False	64,62	67,69	64,62	65,76
Classificação	NB	Gaussian		Doc2Vec - DBOW	False	63,08	63,54	63,08	63,15
Classificação	RF			Doc2Vec - DBOW	True	63,08	62,21	63,08	61,9
Classificação	NB	Gaussian	FS	Doc2Vec - DBOW	False	63,08	63,54	63,08	63,15
Classificação	RF		FS	Doc2Vec - DBOW	True	63,08	62,21	63,08	61,9
Classificação	KNN	Minkowski	ROS	Doc2Vec - DBOW	False	63,08	65,42	63,08	63,59
Classificação	KNN	Euclidean	ROS	Doc2Vec - DBOW	False	63,08	65,42	63,08	63,59
Classificação	RF		ROS	Doc2Vec - DBOW	True	63,08	62,21	63,08	61,9
Classificação	KNN	Minkowski	SMOTE	Doc2Vec - DBOW	False	63,08	66,58	63,08	63,98
Classificação	KNN	Euclidean	SMOTE	Doc2Vec - DBOW	False	63,08	66,58	63,08	63,98
Classificação	RF		SMOTE	Doc2Vec - DBOW	True	63,08	62,21	63,08	61,9
Classificação	KNN	Minkowski	ADASYN	Doc2Vec - DBOW	False	63,08	66,58	63,08	63,98
Classificação	KNN	Euclidean	ADASYN	Doc2Vec - DBOW	False	63,08	66,58	63,08	63,98
Classificação	RF		ADASYN	Doc2Vec - DBOW	True	63,08	62,21	63,08	61,9
Classificação	DT		SMOTE	Word2Vec - skip - gram	False	63,08	66,79	63,08	62,63
Classificação	KNN	Minkowski		Doc2Vec - DBOW	False	61,54	62,82	61,54	62,02
Classificação	KNN	Euclidean		Doc2Vec - DBOW	False	61,54	62,82	61,54	62,02
Classificação	KNN	Minkowski	FS	Doc2Vec - DBOW	False	61,54	62,82	61,54	62,02
Classificação	KNN	Euclidean	FS	Doc2Vec - DBOW	False	61,54	62,82	61,54	62,02
Classificação	DT			Doc2Vec - DM	True	61,54	62,88	61,54	61,32
Classificação	DT		FS	Doc2Vec - DM	True	61,54	62,88	61,54	61,32
Classificação	AB		ROS	Doc2Vec - DM	False	61,54	82,59	61,54	68,01
Classificação	DT		ROS	Doc2Vec - DM	True	61,54	62,88	61,54	61,32

Classificação	AB		SMOTE	Doc2Vec - DM	False	61,54	82,59	61,54	68,01
Classificação	DT		SMOTE	Doc2Vec - DM	True	61,54	62,88	61,54	61,32
Classificação	AB		ADASYN	Doc2Vec - DM	False	61,54	82,59	61,54	68,01
Classificação	DT		ADASYN	Doc2Vec - DM	True	61,54	62,88	61,54	61,32
Classificação	NB	Gaussian		TF-IDF	False	61,54	60,67	61,54	61,07
Classificação	NB	Gaussian	FS	TF-IDF	False	61,54	60,67	61,54	61,07
Classificação	NB	Gaussian	SMOTE	Doc2Vec - DM	False	58,46	61,32	58,46	59,42
Classificação	NB	Gaussian	ADASYN	Doc2Vec - DM	False	58,46	59,28	58,46	58,51
Classificação	NB	Gaussian	ROS	Word2Vec - skip - gram	False	58,46	69,65	58,46	58,17
Classificação	KNN	Minkowski	ROS	Word2Vec - skip - gram	False	58,46	70,95	58,46	62,67
Classificação	KNN	Euclidean	ROS	Word2Vec - skip - gram	False	58,46	70,95	58,46	62,67
Classificação	DT		SMOTE	Word2Vec - CBOW	False	58,46	61,83	58,46	59,94
Classificação	NB	Gaussian		Doc2Vec - DM	False	56,92	63,22	56,92	58,83
Classificação	NB	Gaussian	FS	Doc2Vec - DM	False	56,92	63,22	56,92	58,83
Classificação	NB	Gaussian	ROS	Doc2Vec - DM	False	56,92	63,22	56,92	58,83
Classificação	NB	Gaussian		Word2Vec - skip - gram	False	56,92	69,24	56,92	57,3
Classificação	NB	Gaussian	FS	Word2Vec - skip - gram	False	56,92	69,24	56,92	57,3
Classificação	DT		ADASYN	Word2Vec - CBOW	False	56,92	62,91	56,92	59,77
Classificação	AB			Doc2Vec - DBOW	True	55,38	55,81	55,38	54,59
Classificação	AB		FS	Doc2Vec - DBOW	True	55,38	55,81	55,38	54,59
Classificação	AB		ROS	Doc2Vec - DBOW	True	55,38	55,81	55,38	54,59
Classificação	AB		SMOTE	Doc2Vec - DBOW	True	55,38	55,81	55,38	54,59
Classificação	AB		ADASYN	Doc2Vec - DBOW	True	55,38	55,81	55,38	54,59
Classificação	NB	Gaussian		Doc2Vec - DM	True	55,38	53,78	55,38	53,93

Classificação	KNN	Minkowski		Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	KNN	Euclidean		Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	NB	Gaussian	FS	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	KNN	Minkowski	FS	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	KNN	Euclidean	FS	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	NB	Gaussian	ROS	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	KNN	Minkowski	ROS	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	KNN	Euclidean	ROS	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	NB	Gaussian	SMOTE	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	KNN	Minkowski	SMOTE	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	KNN	Euclidean	SMOTE	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	NB	Gaussian	ADASYN	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	KNN	Minkowski	ADASYN	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	KNN	Euclidean	ADASYN	Doc2Vec - DM	True	55,38	53,78	55,38	53,93
Classificação	NB	Gaussian	SMOTE	Word2Vec - skip - gram	False	55,38	75,94	55,38	57,34
Classificação	NB	Gaussian	ADASYN	Word2Vec - skip - gram	False	55,38	73,8	55,38	57,86
Classificação	KNN	Minkowski		Word2Vec - CBOW	False	55,38	65,21	55,38	59,24
Classificação	KNN	Euclidean		Word2Vec - CBOW	False	55,38	65,21	55,38	59,24
Classificação	KNN	Minkowski	FS	Word2Vec - CBOW	False	55,38	65,21	55,38	59,24
Classificação	KNN	Euclidean	FS	Word2Vec - CBOW	False	55,38	65,21	55,38	59,24
Classificação	DT		ROS	Word2Vec - CBOW	False	55,38	56,28	55,38	55,82
Classificação	DT			Doc2Vec - DBOW	True	53,85	54,11	53,85	53,92
Classificação	KNN	Minkowski		Doc2Vec - DBOW	True	53,85	51,88	53,85	52,23
Classificação	KNN	Euclidean		Doc2Vec - DBOW	True	53,85	51,88	53,85	52,23

Classificação	DT		FS	Doc2Vec - DBOW	True	53,85	54,11	53,85	53,92
Classificação	KNN	Minkowski	FS	Doc2Vec - DBOW	True	53,85	51,88	53,85	52,23
Classificação	KNN	Euclidean	FS	Doc2Vec - DBOW	True	53,85	51,88	53,85	52,23
Classificação	DT		ROS	Doc2Vec - DBOW	True	53,85	54,11	53,85	53,92
Classificação	KNN	Minkowski	ROS	Doc2Vec - DBOW	True	53,85	51,88	53,85	52,23
Classificação	KNN	Euclidean	ROS	Doc2Vec - DBOW	True	53,85	51,88	53,85	52,23
Classificação	DT		SMOTE	Doc2Vec - DBOW	True	53,85	54,11	53,85	53,92
Classificação	KNN	Minkowski	SMOTE	Doc2Vec - DBOW	True	53,85	51,88	53,85	52,23
Classificação	KNN	Euclidean	SMOTE	Doc2Vec - DBOW	True	53,85	51,88	53,85	52,23
Classificação	DT		ADASYN	Doc2Vec - DBOW	True	53,85	54,11	53,85	53,92
Classificação	KNN	Minkowski	ADASYN	Doc2Vec - DBOW	True	53,85	51,88	53,85	52,23
Classificação	KNN	Euclidean	ADASYN	Doc2Vec - DBOW	True	53,85	51,88	53,85	52,23
Classificação	AB			Doc2Vec - DM	False	50,77	79,31	50,77	58,38
Classificação	AB		FS	Doc2Vec - DM	False	50,77	79,31	50,77	58,38
Classificação	DT			Word2Vec - CBOW	True	50,77	50,54	50,77	50,6
Classificação	DT		FS	Word2Vec - CBOW	True	50,77	50,54	50,77	50,6
Classificação	DT		ROS	Word2Vec - CBOW	True	50,77	50,54	50,77	50,6
Classificação	DT		SMOTE	Word2Vec - CBOW	True	50,77	50,54	50,77	50,6
Classificação	DT		ADASYN	Word2Vec - CBOW	True	50,77	50,54	50,77	50,6
Classificação	KNN	Minkowski	ADASYN	Word2Vec - skip - gram	False	49,23	69,21	49,23	54,6
Classificação	KNN	Euclidean	ADASYN	Word2Vec - skip - gram	False	49,23	69,21	49,23	54,6
Classificação	KNN	Minkowski	ROS	Word2Vec - CBOW	False	49,23	65,64	49,23	55,22
Classificação	KNN	Euclidean	ROS	Word2Vec - CBOW	False	49,23	65,64	49,23	55,22
Classificação	KNN	Minkowski	SMOTE	Word2Vec - skip - gram	False	44,62	62,76	44,62	48,16

Classificação	KNN	Euclidean	SMOTE	Word2Vec - skip - gram	False	44,62	62,76	44,62	48,16
Classificação	KNN	Minkowski	SMOTE	Word2Vec - CBOW	False	44,62	70,98	44,62	51,55
Classificação	KNN	Euclidean	SMOTE	Word2Vec - CBOW	False	44,62	70,98	44,62	51,55
Classificação	KNN	Minkowski	ADASYN	Word2Vec - CBOW	False	44,62	68,14	44,62	51,38
Classificação	KNN	Euclidean	ADASYN	Word2Vec - CBOW	False	44,62	68,14	44,62	51,38
Classificação	AB		ROS	Doc2Vec - DBOW	False	41,54	64,6	41,54	46,26
Classificação	AB		SMOTE	Doc2Vec - DBOW	False	41,54	64,6	41,54	46,26
Classificação	AB		ADASYN	Doc2Vec - DBOW	False	41,54	64,6	41,54	46,26
Classificação	KNN	Minkowski	SMOTE	TF-IDF	False	37,78	41,82	37,78	25,46
Classificação	KNN	Euclidean	SMOTE	TF-IDF	False	37,78	41,82	37,78	25,46
Classificação	KNN	Minkowski	SMOTE	TF-IDF	True	36,67	40,79	36,67	23,22
Classificação	KNN	Euclidean	SMOTE	TF-IDF	True	36,67	40,79	36,67	23,22
Classificação	KNN	Minkowski	ADASYN	TF-IDF	True	36,26	47,99	36,26	23,8
Classificação	KNN	Euclidean	ADASYN	TF-IDF	True	36,26	47,99	36,26	23,8
Classificação	KNN	Minkowski	ADASYN	TF-IDF	False	35,16	35,9	35,16	22,25
Classificação	KNN	Euclidean	ADASYN	TF-IDF	False	35,16	35,9	35,16	22,25

**Apêndice 2 – Resultados Classificação – LSTM**

Tipo	Algoritmo	Loss	Equilibrador	CV	Accuracy	Precision	Recall	F1 - Score
Classificação	LSTM	0.5	ROS	False	74,44	74,74	74,44	74,44
Classificação	LSTM	1.29	ADASYN	False	64,44	67,08	64,44	61,83
Classificação	LSTM	0.83	FS	False	63,08	61,59	63,08	61,95
Classificação	LSTM	0.76		False	58,46	61,35	58,46	53,98
Classificação	LSTM	1.43	SMOTE	False	55,56	55,83	55,56	55,22

**Apêndice 3 – Resultados Sentimento**

Tipo	Algoritmo	Variante	Equilibrador	Extrator	CV	Accuracy	Precision	Recall	F1 - Score
Sentimento	RF		ROS	TF-IDF	False	99,38	99,39	99,38	99,38
Sentimento	RF		SMOTE	TF-IDF	False	99,38	99,39	99,38	99,38
Sentimento	NB	Gaussian	ROS	TF-IDF	False	97,53	97,62	97,53	97,51
Sentimento	NB	Gaussian	SMOTE	TF-IDF	False	97,53	97,62	97,53	97,51
Sentimento	DT		SMOTE	TF-IDF	False	95,68	95,69	95,68	95,68
Sentimento	RF		ADASYN	TF-IDF	False	94,74	93,27	94,74	93,51
Sentimento	DT		ROS	TF-IDF	False	94,44	94,93	94,44	94,33
Sentimento	RF		ROS	TF-IDF	True	93,21	93,56	93,21	93,23
Sentimento	RF		SMOTE	TF-IDF	True	91,98	92,88	91,98	92,06
Sentimento	NB	Gaussian	SMOTE	TF-IDF	True	91,36	92,38	91,36	91,02
Sentimento	NB	Gaussian	ADASYN	TF-IDF	False	90,35	87,88	90,35	89,09
Sentimento	RF		SMOTE	Word2Vec - skip - gram	False	89,23	89,63	89,23	88,38
Sentimento	RF		ADASYN	Word2Vec - CBOW	False	89,23	90,45	89,23	86,51
Sentimento	DT		ROS	TF-IDF	True	88,27	88,76	88,27	88,07
Sentimento	DT		ADASYN	TF-IDF	False	87,72	87,08	87,72	87,24
Sentimento	AB			TF-IDF	False	87,69	86,27	87,69	86,37
Sentimento	AB		FS	TF-IDF	False	87,69	86,27	87,69	86,37
Sentimento	AB			Word2Vec - skip - gram	False	87,69	81,56	87,69	83,27
Sentimento	AB		FS	Word2Vec - skip - gram	False	87,69	81,56	87,69	83,27
Sentimento	RF			Word2Vec - CBOW	False	87,69	81,56	87,69	83,27
Sentimento	RF		FS	Word2Vec - CBOW	False	87,69	81,56	87,69	83,27

Sentimento	RF		ROS	Word2Vec - CBOW	False	87,69	81,56	87,69	83,27
Sentimento	RF			Doc2Vec - DBOW	False	86,15	78,83	86,15	81,95
Sentimento	RF		FS	Doc2Vec - DBOW	False	86,15	78,83	86,15	81,95
Sentimento	RF			TF-IDF	False	86,15	80,41	86,15	80,78
Sentimento	RF		FS	TF-IDF	False	86,15	80,41	86,15	80,78
Sentimento	RF			Word2Vec - skip - gram	False	86,15	80,41	86,15	80,78
Sentimento	AB			Word2Vec - skip - gram	True	86,15	78,83	86,15	81,95
Sentimento	RF		FS	Word2Vec - skip - gram	False	86,15	80,41	86,15	80,78
Sentimento	AB		FS	Word2Vec - skip - gram	True	86,15	78,83	86,15	81,95
Sentimento	RF		ROS	Word2Vec - skip - gram	False	86,15	85,24	86,15	82,87
Sentimento	AB		ROS	Word2Vec - skip - gram	False	86,15	80,41	86,15	80,78
Sentimento	AB		ROS	Word2Vec - skip - gram	True	86,15	78,83	86,15	81,95
Sentimento	AB		SMOTE	Word2Vec - skip - gram	True	86,15	78,83	86,15	81,95
Sentimento	AB		ADASYN	Word2Vec - skip - gram	True	86,15	78,83	86,15	81,95
Sentimento	AB		ADASYN	TF-IDF	False	85,96	85,28	85,96	85,57
Sentimento	NB	Gaussian	ROS	TF-IDF	True	85,8	86,17	85,8	85,59
Sentimento	KNN	Minkowski	ROS	TF-IDF	False	85,19	88,08	85,19	84,02
Sentimento	KNN	Euclidean	ROS	TF-IDF	False	85,19	88,08	85,19	84,02
Sentimento	DT		SMOTE	TF-IDF	True	85,19	85,16	85,19	85,16
Sentimento	DT		ADASYN	TF-IDF	True	85,09	80,89	85,09	82,84
Sentimento	RF		ADASYN	TF-IDF	True	85,09	83,25	85,09	82,84
Sentimento	KNN	Minkowski		Doc2Vec - DBOW	False	84,62	76,37	84,62	79,64
Sentimento	KNN	Euclidean		Doc2Vec - DBOW	False	84,62	76,37	84,62	79,64
Sentimento	KNN	Minkowski	FS	Doc2Vec - DBOW	False	84,62	76,37	84,62	79,64

Sentimento	KNN	Euclidean	FS	Doc2Vec - DBOW	False	84,62	76,37	84,62	79,64
Sentimento	RF		ROS	Doc2Vec - DBOW	False	84,62	76,37	84,62	79,64
Sentimento	DT			TF-IDF	True	84,62	76,37	84,62	79,64
Sentimento	RF			TF-IDF	True	84,62	71,6	84,62	77,56
Sentimento	AB			TF-IDF	True	84,62	76,37	84,62	79,64
Sentimento	DT		FS	TF-IDF	True	84,62	76,37	84,62	79,64
Sentimento	RF		FS	TF-IDF	True	84,62	71,6	84,62	77,56
Sentimento	AB		FS	TF-IDF	True	84,62	76,37	84,62	79,64
Sentimento	NB	Gaussian		Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
Sentimento	RF			Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
Sentimento	KNN	Minkowski		Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Sentimento	KNN	Euclidean		Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Sentimento	NB	Gaussian	FS	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
Sentimento	RF		FS	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
Sentimento	KNN	Minkowski	FS	Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Sentimento	KNN	Euclidean	FS	Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Sentimento	NB	Gaussian	ROS	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
Sentimento	RF		ROS	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
Sentimento	KNN	Minkowski	ROS	Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Sentimento	KNN	Euclidean	ROS	Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Sentimento	NB	Gaussian	SMOTE	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
Sentimento	RF		SMOTE	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
Sentimento	KNN	Minkowski	SMOTE	Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Sentimento	KNN	Euclidean	SMOTE	Word2Vec - CBOW	True	84,62	81,21	84,62	81,72

Sentimento	NB	Gaussian	ADASYN	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
Sentimento	RF		ADASYN	Word2Vec - CBOW	True	84,62	71,6	84,62	77,56
Sentimento	KNN	Minkowski	ADASYN	Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Sentimento	KNN	Euclidean	ADASYN	Word2Vec - CBOW	True	84,62	81,21	84,62	81,72
Sentimento	NB	Gaussian	ADASYN	TF-IDF	True	83,33	82,69	83,33	82,37
Sentimento	AB			Doc2Vec - DBOW	False	83,08	74,9	83,08	78,58
Sentimento	AB		FS	Doc2Vec - DBOW	False	83,08	74,9	83,08	78,58
Sentimento	RF			Doc2Vec - DM	False	83,08	76,41	83,08	79,6
Sentimento	AB			Doc2Vec - DM	False	83,08	71,39	83,08	76,79
Sentimento	RF		FS	Doc2Vec - DM	False	83,08	76,41	83,08	79,6
Sentimento	AB		FS	Doc2Vec - DM	False	83,08	71,39	83,08	76,79
Sentimento	NB	Gaussian		TF-IDF	False	83,08	82,45	83,08	82,71
Sentimento	DT			TF-IDF	False	83,08	84,66	83,08	81,74
Sentimento	NB	Gaussian	FS	TF-IDF	False	83,08	82,45	83,08	82,71
Sentimento	DT		FS	TF-IDF	False	83,08	84,66	83,08	81,74
Sentimento	RF			Word2Vec - skip - gram	True	83,08	71,39	83,08	76,79
Sentimento	RF		FS	Word2Vec - skip - gram	True	83,08	71,39	83,08	76,79
Sentimento	RF		ROS	Word2Vec - skip - gram	True	83,08	71,39	83,08	76,79
Sentimento	RF		SMOTE	Word2Vec - skip - gram	True	83,08	71,39	83,08	76,79
Sentimento	AB		ADASYN	Word2Vec - skip - gram	False	83,08	78,46	83,08	80,37
Sentimento	RF		ADASYN	Word2Vec - skip - gram	True	83,08	71,39	83,08	76,79
Sentimento	NB	Gaussian	SMOTE	Word2Vec - CBOW	False	83,08	81,33	83,08	81,65
Sentimento	RF		SMOTE	Word2Vec - CBOW	False	83,08	74,9	83,08	78,58
Sentimento	AB		SMOTE	Word2Vec - CBOW	False	83,08	80,02	83,08	79,22

Sentimento	AB		ADASYN	TF-IDF	True	81,58	81,14	81,58	80,6
Sentimento	AB			Doc2Vec - DBOW	True	81,54	71,18	81,54	76,01
Sentimento	AB		FS	Doc2Vec - DBOW	True	81,54	71,18	81,54	76,01
Sentimento	AB		ROS	Doc2Vec - DBOW	True	81,54	71,18	81,54	76,01
Sentimento	AB		SMOTE	Doc2Vec - DBOW	True	81,54	71,18	81,54	76,01
Sentimento	AB		ADASYN	Doc2Vec - DBOW	True	81,54	71,18	81,54	76,01
Sentimento	KNN	Minkowski		Doc2Vec - DM	False	81,54	71,18	81,54	76,01
Sentimento	KNN	Euclidean		Doc2Vec - DM	False	81,54	71,18	81,54	76,01
Sentimento	AB			Doc2Vec - DM	True	81,54	71,18	81,54	76,01
Sentimento	KNN	Minkowski	FS	Doc2Vec - DM	False	81,54	71,18	81,54	76,01
Sentimento	KNN	Euclidean	FS	Doc2Vec - DM	False	81,54	71,18	81,54	76,01
Sentimento	AB		FS	Doc2Vec - DM	True	81,54	71,18	81,54	76,01
Sentimento	RF		ROS	Doc2Vec - DM	False	81,54	86,78	81,54	81,72
Sentimento	AB		ROS	Doc2Vec - DM	True	81,54	71,18	81,54	76,01
Sentimento	AB		SMOTE	Doc2Vec - DM	True	81,54	71,18	81,54	76,01
Sentimento	AB		ADASYN	Doc2Vec - DM	True	81,54	71,18	81,54	76,01
Sentimento	RF			Doc2Vec - DBOW	True	80	74,49	80	76,97
Sentimento	RF		FS	Doc2Vec - DBOW	True	80	74,49	80	76,97
Sentimento	RF		ROS	Doc2Vec - DBOW	True	80	74,49	80	76,97
Sentimento	RF		SMOTE	Doc2Vec - DBOW	True	80	74,49	80	76,97
Sentimento	RF		ADASYN	Doc2Vec - DBOW	True	80	74,49	80	76,97
Sentimento	RF			Doc2Vec - DM	True	80	70,97	80	75,21
Sentimento	RF		FS	Doc2Vec - DM	True	80	70,97	80	75,21
Sentimento	RF		ROS	Doc2Vec - DM	True	80	70,97	80	75,21

Sentimento	RF		SMOTE	Doc2Vec - DM	True	80	70,97	80	75,21
Sentimento	RF		ADASYN	Doc2Vec - DM	True	80	70,97	80	75,21
Sentimento	KNN	Minkowski		TF-IDF	False	80	75,71	80	77,63
Sentimento	KNN	Euclidean		TF-IDF	False	80	75,71	80	77,63
Sentimento	KNN	Minkowski	FS	TF-IDF	False	80	75,71	80	77,63
Sentimento	KNN	Euclidean	FS	TF-IDF	False	80	75,71	80	77,63
Sentimento	KNN	Minkowski		Word2Vec - skip - gram	True	80	73,85	80	76,76
Sentimento	KNN	Euclidean		Word2Vec - skip - gram	True	80	73,85	80	76,76
Sentimento	KNN	Minkowski	FS	Word2Vec - skip - gram	True	80	73,85	80	76,76
Sentimento	KNN	Euclidean	FS	Word2Vec - skip - gram	True	80	73,85	80	76,76
Sentimento	KNN	Minkowski	ROS	Word2Vec - skip - gram	True	80	73,85	80	76,76
Sentimento	KNN	Euclidean	ROS	Word2Vec - skip - gram	True	80	73,85	80	76,76
Sentimento	KNN	Minkowski	SMOTE	Word2Vec - skip - gram	True	80	73,85	80	76,76
Sentimento	KNN	Euclidean	SMOTE	Word2Vec - skip - gram	True	80	73,85	80	76,76
Sentimento	RF		ADASYN	Word2Vec - skip - gram	False	80	81,92	80	78,84
Sentimento	KNN	Minkowski	ADASYN	Word2Vec - skip - gram	True	80	73,85	80	76,76
Sentimento	KNN	Euclidean	ADASYN	Word2Vec - skip - gram	True	80	73,85	80	76,76
Sentimento	AB			Word2Vec - CBOW	True	80	72,13	80	75,86
Sentimento	AB		FS	Word2Vec - CBOW	True	80	72,13	80	75,86
Sentimento	AB		ROS	Word2Vec - CBOW	True	80	72,13	80	75,86
Sentimento	AB		SMOTE	Word2Vec - CBOW	True	80	72,13	80	75,86
Sentimento	AB		ADASYN	Word2Vec - CBOW	True	80	72,13	80	75,86
Sentimento	DT			Doc2Vec - DBOW	False	78,46	75,3	78,46	76,84
Sentimento	DT		FS	Doc2Vec - DBOW	False	78,46	75,3	78,46	76,84

Sentimento	AB		ROS	Doc2Vec - DM	False	78,46	70,74	78,46	74,4
Sentimento	NB	Gaussian		TF-IDF	True	78,46	75,55	78,46	76,42
Sentimento	NB	Gaussian	FS	TF-IDF	True	78,46	75,55	78,46	76,42
Sentimento	KNN	Minkowski		Word2Vec - skip - gram	False	78,46	77,88	78,46	78,03
Sentimento	KNN	Euclidean		Word2Vec - skip - gram	False	78,46	77,88	78,46	78,03
Sentimento	KNN	Minkowski	FS	Word2Vec - skip - gram	False	78,46	77,88	78,46	78,03
Sentimento	KNN	Euclidean	FS	Word2Vec - skip - gram	False	78,46	77,88	78,46	78,03
Sentimento	DT		ADASYN	Word2Vec - skip - gram	False	78,46	80,81	78,46	79,43
Sentimento	NB	Gaussian	ADASYN	Word2Vec - CBOW	False	78,46	81,32	78,46	79,25
Sentimento	KNN	Minkowski		Doc2Vec - DBOW	True	76,92	70,51	76,92	73,58
Sentimento	KNN	Euclidean		Doc2Vec - DBOW	True	76,92	70,51	76,92	73,58
Sentimento	KNN	Minkowski	FS	Doc2Vec - DBOW	True	76,92	70,51	76,92	73,58
Sentimento	KNN	Euclidean	FS	Doc2Vec - DBOW	True	76,92	70,51	76,92	73,58
Sentimento	AB		ROS	Doc2Vec - DBOW	False	76,92	76,21	76,92	76,27
Sentimento	KNN	Minkowski	ROS	Doc2Vec - DBOW	True	76,92	70,51	76,92	73,58
Sentimento	KNN	Euclidean	ROS	Doc2Vec - DBOW	True	76,92	70,51	76,92	73,58
Sentimento	KNN	Minkowski	SMOTE	Doc2Vec - DBOW	True	76,92	70,51	76,92	73,58
Sentimento	KNN	Euclidean	SMOTE	Doc2Vec - DBOW	True	76,92	70,51	76,92	73,58
Sentimento	KNN	Minkowski	ADASYN	Doc2Vec - DBOW	True	76,92	70,51	76,92	73,58
Sentimento	KNN	Euclidean	ADASYN	Doc2Vec - DBOW	True	76,92	70,51	76,92	73,58
Sentimento	KNN	Minkowski		Doc2Vec - DM	True	76,92	73,41	76,92	75,09
Sentimento	KNN	Euclidean		Doc2Vec - DM	True	76,92	73,41	76,92	75,09
Sentimento	KNN	Minkowski	FS	Doc2Vec - DM	True	76,92	73,41	76,92	75,09
Sentimento	KNN	Euclidean	FS	Doc2Vec - DM	True	76,92	73,41	76,92	75,09

Sentimento	KNN	Minkowski	ROS	Doc2Vec - DM	True	76,92	73,41	76,92	75,09
Sentimento	KNN	Euclidean	ROS	Doc2Vec - DM	True	76,92	73,41	76,92	75,09
Sentimento	KNN	Minkowski	SMOTE	Doc2Vec - DM	True	76,92	73,41	76,92	75,09
Sentimento	KNN	Euclidean	SMOTE	Doc2Vec - DM	True	76,92	73,41	76,92	75,09
Sentimento	KNN	Minkowski	ADASYN	Doc2Vec - DM	True	76,92	73,41	76,92	75,09
Sentimento	KNN	Euclidean	ADASYN	Doc2Vec - DM	True	76,92	73,41	76,92	75,09
Sentimento	DT			Word2Vec - skip - gram	True	76,92	75,6	76,92	76,26
Sentimento	DT		FS	Word2Vec - skip - gram	True	76,92	75,6	76,92	76,26
Sentimento	DT		ROS	Word2Vec - skip - gram	True	76,92	75,6	76,92	76,26
Sentimento	DT		SMOTE	Word2Vec - skip - gram	True	76,92	75,6	76,92	76,26
Sentimento	DT		ADASYN	Word2Vec - skip - gram	True	76,92	75,6	76,92	76,26
Sentimento	DT			Word2Vec - CBOW	False	76,92	75,6	76,92	76,26
Sentimento	AB			Word2Vec - CBOW	False	76,92	78,59	76,92	77,67
Sentimento	KNN	Minkowski		Word2Vec - CBOW	False	76,92	76,37	76,92	76,5
Sentimento	KNN	Euclidean		Word2Vec - CBOW	False	76,92	76,37	76,92	76,5
Sentimento	DT		FS	Word2Vec - CBOW	False	76,92	75,6	76,92	76,26
Sentimento	AB		FS	Word2Vec - CBOW	False	76,92	78,59	76,92	77,67
Sentimento	KNN	Minkowski	FS	Word2Vec - CBOW	False	76,92	76,37	76,92	76,5
Sentimento	KNN	Euclidean	FS	Word2Vec - CBOW	False	76,92	76,37	76,92	76,5
Sentimento	AB		ROS	TF-IDF	False	75,93	77,87	75,93	76,46
Sentimento	KNN	Minkowski	ROS	TF-IDF	True	75,93	76,54	75,93	74,93
Sentimento	KNN	Euclidean	ROS	TF-IDF	True	75,93	76,54	75,93	74,93
Sentimento	DT			Doc2Vec - DM	False	75,38	72,54	75,38	73,93
Sentimento	DT		FS	Doc2Vec - DM	False	75,38	72,54	75,38	73,93

Sentimento	DT		ROS	Word2Vec - skip - gram	False	75,38	75,51	75,38	75,42
Sentimento	AB		SMOTE	Word2Vec - skip - gram	False	75,38	75,51	75,38	75,42
Sentimento	NB	Gaussian		Word2Vec - CBOW	False	75,38	86,47	75,38	78,62
Sentimento	NB	Gaussian	FS	Word2Vec - CBOW	False	75,38	86,47	75,38	78,62
Sentimento	NB	Gaussian	ROS	Word2Vec - CBOW	False	75,38	86,72	75,38	78,77
Sentimento	DT		ADASYN	Word2Vec - CBOW	False	75,38	77,17	75,38	76,08
Sentimento	AB		ADASYN	Word2Vec - CBOW	False	75,38	80,75	75,38	77,63
Sentimento	AB		SMOTE	TF-IDF	False	75,31	78,02	75,31	75,87
Sentimento	DT		ROS	Doc2Vec - DBOW	False	73,85	73,59	73,85	73,71
Sentimento	KNN	Minkowski		TF-IDF	True	73,85	72,94	73,85	73,37
Sentimento	KNN	Euclidean		TF-IDF	True	73,85	72,94	73,85	73,37
Sentimento	KNN	Minkowski	FS	TF-IDF	True	73,85	72,94	73,85	73,37
Sentimento	KNN	Euclidean	FS	TF-IDF	True	73,85	72,94	73,85	73,37
Sentimento	DT			Word2Vec - skip - gram	False	73,85	80,49	73,85	76,65
Sentimento	DT		FS	Word2Vec - skip - gram	False	73,85	80,49	73,85	76,65
Sentimento	DT		ROS	Word2Vec - CBOW	False	73,85	77,26	73,85	75,43
Sentimento	AB		ADASYN	Doc2Vec - DM	False	72,31	79,12	72,31	74,02
Sentimento	DT			Doc2Vec - DBOW	True	70,77	74,16	70,77	72,39
Sentimento	DT		FS	Doc2Vec - DBOW	True	70,77	74,16	70,77	72,39
Sentimento	DT		ROS	Doc2Vec - DBOW	True	70,77	74,16	70,77	72,39
Sentimento	DT		SMOTE	Doc2Vec - DBOW	True	70,77	74,16	70,77	72,39
Sentimento	DT		ADASYN	Doc2Vec - DBOW	True	70,77	74,16	70,77	72,39
Sentimento	DT			Doc2Vec - DM	True	70,77	73,08	70,77	71,79
Sentimento	DT		FS	Doc2Vec - DM	True	70,77	73,08	70,77	71,79

Sentimento	DT		ROS	Doc2Vec - DM	False	70,77	80,94	70,77	74,02
Sentimento	DT		ROS	Doc2Vec - DM	True	70,77	73,08	70,77	71,79
Sentimento	DT		SMOTE	Doc2Vec - DM	False	70,77	87,94	70,77	75,93
Sentimento	DT		SMOTE	Doc2Vec - DM	True	70,77	73,08	70,77	71,79
Sentimento	DT		ADASYN	Doc2Vec - DM	True	70,77	73,08	70,77	71,79
Sentimento	DT			Word2Vec - CBOW	True	70,77	69,51	70,77	70,13
Sentimento	DT		FS	Word2Vec - CBOW	True	70,77	69,51	70,77	70,13
Sentimento	DT		ROS	Word2Vec - CBOW	True	70,77	69,51	70,77	70,13
Sentimento	DT		SMOTE	Word2Vec - CBOW	True	70,77	69,51	70,77	70,13
Sentimento	DT		ADASYN	Word2Vec - CBOW	True	70,77	69,51	70,77	70,13
Sentimento	KNN	Minkowski	ROS	Doc2Vec - DBOW	False	69,23	72,56	69,23	70,77
Sentimento	KNN	Euclidean	ROS	Doc2Vec - DBOW	False	69,23	72,56	69,23	70,77
Sentimento	AB		SMOTE	Doc2Vec - DM	False	69,23	79,1	69,23	72,42
Sentimento	DT		SMOTE	Word2Vec - CBOW	False	69,23	78,58	69,23	72,87
Sentimento	RF		ADASYN	Doc2Vec - DBOW	False	67,69	77,95	67,69	72,02
Sentimento	DT		ADASYN	Doc2Vec - DM	False	67,69	76,84	67,69	71,35
Sentimento	RF		ADASYN	Doc2Vec - DM	False	67,69	85,9	67,69	72,28
Sentimento	KNN	Minkowski	SMOTE	TF-IDF	True	66,67	78,11	66,67	55,13
Sentimento	KNN	Euclidean	SMOTE	TF-IDF	True	66,67	78,11	66,67	55,13
Sentimento	KNN	Minkowski	ADASYN	Doc2Vec - DBOW	False	66,15	80,16	66,15	70,28
Sentimento	KNN	Euclidean	ADASYN	Doc2Vec - DBOW	False	66,15	80,16	66,15	70,28
Sentimento	KNN	Minkowski	SMOTE	TF-IDF	False	66,05	77,84	66,05	53,8
Sentimento	KNN	Euclidean	SMOTE	TF-IDF	False	66,05	77,84	66,05	53,8
Sentimento	AB		ADASYN	Doc2Vec - DBOW	False	64,62	76,28	64,62	69,16

Sentimento	KNN	Minkowski	ROS	Doc2Vec - DM	False	64,62	78,42	64,62	69,93
Sentimento	KNN	Euclidean	ROS	Doc2Vec - DM	False	64,62	78,42	64,62	69,93
Sentimento	DT		SMOTE	Doc2Vec - DBOW	False	63,08	77,49	63,08	68,95
Sentimento	RF		SMOTE	Doc2Vec - DBOW	False	63,08	78,21	63,08	68,73
Sentimento	RF		SMOTE	Doc2Vec - DM	False	63,08	81,34	63,08	68,53
Sentimento	KNN	Minkowski	ROS	Word2Vec - skip - gram	False	63,08	79	63,08	68,9
Sentimento	KNN	Euclidean	ROS	Word2Vec - skip - gram	False	63,08	79	63,08	68,9
Sentimento	DT		SMOTE	Word2Vec - skip - gram	False	63,08	76,87	63,08	68,5
Sentimento	KNN	Minkowski	ROS	Word2Vec - CBOW	False	63,08	77,5	63,08	67,99
Sentimento	KNN	Euclidean	ROS	Word2Vec - CBOW	False	63,08	77,5	63,08	67,99
Sentimento	AB		SMOTE	TF-IDF	True	62,35	64,84	62,35	63,11
Sentimento	DT		ADASYN	Doc2Vec - DBOW	False	61,54	77,22	61,54	67,85
Sentimento	NB	Gaussian		Doc2Vec - DBOW	True	60	74,8	60	65,32
Sentimento	NB	Gaussian	FS	Doc2Vec - DBOW	True	60	74,8	60	65,32
Sentimento	NB	Gaussian	ROS	Doc2Vec - DBOW	True	60	74,8	60	65,32
Sentimento	NB	Gaussian	SMOTE	Doc2Vec - DBOW	True	60	74,8	60	65,32
Sentimento	NB	Gaussian	ADASYN	Doc2Vec - DBOW	True	60	74,8	60	65,32
Sentimento	AB		ROS	TF-IDF	True	56,17	63,46	56,17	54,35
Sentimento	KNN	Minkowski	SMOTE	Doc2Vec - DBOW	False	50,77	76,43	50,77	58,87
Sentimento	KNN	Euclidean	SMOTE	Doc2Vec - DBOW	False	50,77	76,43	50,77	58,87
Sentimento	KNN	Minkowski	ADASYN	Doc2Vec - DM	False	50,77	80,07	50,77	57,67
Sentimento	KNN	Euclidean	ADASYN	Doc2Vec - DM	False	50,77	80,07	50,77	57,67
Sentimento	AB		SMOTE	Doc2Vec - DBOW	False	49,23	69,43	49,23	57,61
Sentimento	NB	Gaussian		Word2Vec - skip - gram	True	49,23	69,74	49,23	56,24

Sentimento	NB	Gaussian	FS	Word2Vec - skip - gram	True	49,23	69,74	49,23	56,24
Sentimento	NB	Gaussian	ROS	Word2Vec - skip - gram	True	49,23	69,74	49,23	56,24
Sentimento	NB	Gaussian	SMOTE	Word2Vec - skip - gram	True	49,23	69,74	49,23	56,24
Sentimento	NB	Gaussian	ADASYN	Word2Vec - skip - gram	True	49,23	69,74	49,23	56,24
Sentimento	KNN	Minkowski	ADASYN	Word2Vec - skip - gram	False	47,69	83,3	47,69	56,75
Sentimento	KNN	Euclidean	ADASYN	Word2Vec - skip - gram	False	47,69	83,3	47,69	56,75
Sentimento	KNN	Minkowski	ADASYN	Word2Vec - CBOW	False	47,69	80,49	47,69	55,09
Sentimento	KNN	Euclidean	ADASYN	Word2Vec - CBOW	False	47,69	80,49	47,69	55,09
Sentimento	KNN	Minkowski	ADASYN	TF-IDF	False	47,37	70,55	47,37	33,17
Sentimento	KNN	Euclidean	ADASYN	TF-IDF	False	47,37	70,55	47,37	33,17
Sentimento	KNN	Minkowski	ADASYN	TF-IDF	True	47,37	70,55	47,37	33,17
Sentimento	KNN	Euclidean	ADASYN	TF-IDF	True	47,37	70,55	47,37	33,17
Sentimento	KNN	Minkowski	SMOTE	Word2Vec - skip - gram	False	46,15	84,04	46,15	52,56
Sentimento	KNN	Euclidean	SMOTE	Word2Vec - skip - gram	False	46,15	84,04	46,15	52,56
Sentimento	KNN	Minkowski	SMOTE	Doc2Vec - DM	False	44,62	83,81	44,62	51,95
Sentimento	KNN	Euclidean	SMOTE	Doc2Vec - DM	False	44,62	83,81	44,62	51,95
Sentimento	KNN	Minkowski	SMOTE	Word2Vec - CBOW	False	43,08	79,79	43,08	48,56
Sentimento	KNN	Euclidean	SMOTE	Word2Vec - CBOW	False	43,08	79,79	43,08	48,56
Sentimento	NB	Gaussian		Doc2Vec - DM	True	38,46	67,71	38,46	46,79
Sentimento	NB	Gaussian	FS	Doc2Vec - DM	True	38,46	67,71	38,46	46,79
Sentimento	NB	Gaussian	ROS	Doc2Vec - DM	True	38,46	67,71	38,46	46,79
Sentimento	NB	Gaussian	SMOTE	Doc2Vec - DM	True	38,46	67,71	38,46	46,79
Sentimento	NB	Gaussian	ADASYN	Doc2Vec - DM	True	38,46	67,71	38,46	46,79
Sentimento	NB	Gaussian	ADASYN	Word2Vec - skip - gram	False	35,38	87,08	35,38	40,37

Sentimento	NB	Gaussian	SMOTE	Word2Vec - skip - gram	False	33,85	81,52	33,85	39,4
Sentimento	NB	Gaussian		Word2Vec - skip - gram	False	30,77	86,96	30,77	34,24
Sentimento	NB	Gaussian	FS	Word2Vec - skip - gram	False	30,77	86,96	30,77	34,24
Sentimento	NB	Gaussian	ROS	Word2Vec - skip - gram	False	30,77	86,96	30,77	34,24
Sentimento	NB	Gaussian	ADASYN	Doc2Vec - DM	False	27,69	79,42	27,69	29,74
Sentimento	NB	Gaussian	ADASYN	Doc2Vec - DBOW	False	24,62	71,28	24,62	26,52
Sentimento	NB	Gaussian	SMOTE	Doc2Vec - DBOW	False	23,08	77,22	23,08	24,52
Sentimento	NB	Gaussian	SMOTE	Doc2Vec - DM	False	23,08	87,26	23,08	20,99
Sentimento	NB	Gaussian		Doc2Vec - DBOW	False	21,54	75,92	21,54	22
Sentimento	NB	Gaussian	FS	Doc2Vec - DBOW	False	21,54	75,92	21,54	22
Sentimento	NB	Gaussian	ROS	Doc2Vec - DBOW	False	21,54	75,92	21,54	22
Sentimento	NB	Gaussian		Doc2Vec - DM	False	18,46	86,76	18,46	15,1
Sentimento	NB	Gaussian	FS	Doc2Vec - DM	False	18,46	86,76	18,46	15,1
Sentimento	NB	Gaussian	ROS	Doc2Vec - DM	False	18,46	86,76	18,46	15,1
Sentimento	AB		ROS	Word2Vec - CBOW	False	10,77	51,5	10,77	9,77

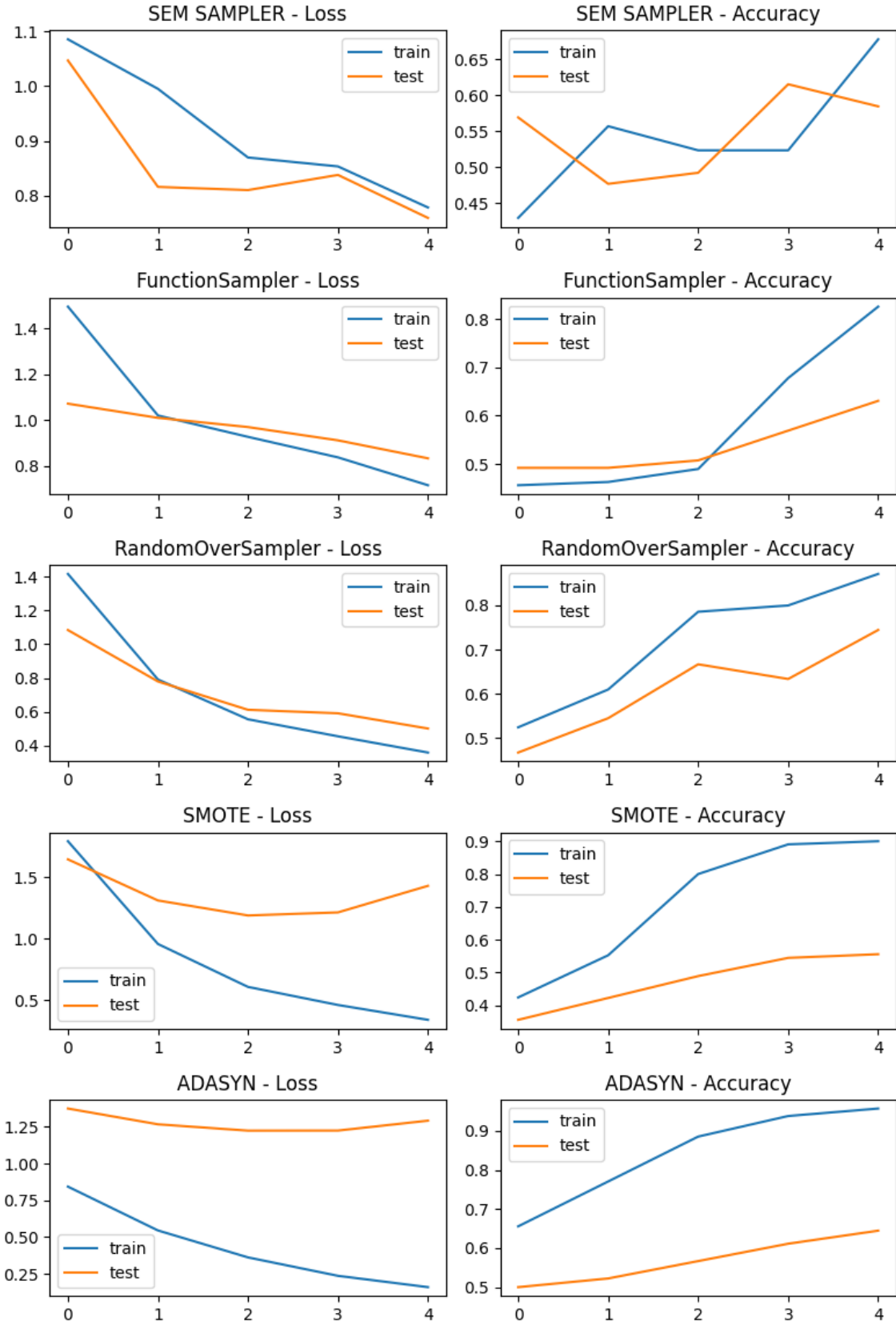
**Apêndice 4 – Resultados Sentimento – LSTM**

Tipo	Algoritmo	Loss	Equilibrador	CV	Accuracy	Precision	Recall	F1 - Score
Sentimento	LSTM	0.24	ROS	False	90,74	92,01	90,74	90,38
Sentimento	LSTM	0.55		False	84,62	71,6	84,62	77,56
Sentimento	LSTM	0.61	FS	False	84,62	72,72	84,62	78,22
Sentimento	LSTM	0.64	ADASYN	False	72,97	71,72	72,97	71,67
Sentimento	LSTM	1.07	SMOTE	False	56,17	56,02	56,17	55,85

**Apêndice 5 – Gráficos de accuracy e loss para cada um dos testes de LSTM –**

*Classificação*

Classificação



**Apêndice 6** – Gráficos de accuracy e loss para cada um dos testes de LSTM – Sentimento

