



UNIVERSIDADE ESTADUAL PAULISTA
“Júlio de Mesquita Filho”
Instituto de Geociências e Ciências Exatas
Câmpus de Rio Claro
Programa de Pós-Graduação em Ciência da Computação

KETERLY GEOVANA GOUVEIA SILVA

**PLN APLICADO À SAÚDE MENTAL:
EXPLORANDO ABORDAGENS PARA AUXÍLIO
NA CLASSIFICAÇÃO AUTOMÁTICA DE
SUICÍDIO E DEPRESSÃO EM TEXTOS**

Rio Claro - SP
2026

UNIVERSIDADE ESTADUAL PAULISTA
“Júlio de Mesquita Filho”
Instituto de Geociências e Ciências Exatas
Câmpus de Rio Claro

KETERLY GEOVANA GOUVEIA SILVA

**PLN APLICADO À SAÚDE MENTAL: EXPLORANDO ABORDAGENS PARA
AUXÍLIO NA CLASSIFICAÇÃO AUTOMÁTICA DE SUICÍDIO E DEPRESSÃO EM
TEXTOS**

Dissertação de Mestrado apresentada ao
Instituto de Geociências e Ciências Exatas
do Câmpus de Rio Claro, da Universidade
Estadual Paulista “Júlio de Mesquita Filho”,
como parte dos requisitos para obtenção do
título de Mestra em Ciência da Computação.

Orientador: Prof. Dr. Fabricio Aparecido
Breve

Rio Claro - SP
2026

S586p

Silva, Keterly Geovana Gouveia

PLN aplicado à saúde mental: explorando abordagens para auxílio na classificação automática de suicídio e depressão em textos / Keterly Geovana Gouveia Silva. -- Rio Claro, 2026

145 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP), Instituto de Geociências e Ciências Exatas, Rio Claro

Orientador: Fabricio Aparecido Breve

1. Processamento de linguagem natural. 2. Aprendizado de máquina. 3. Competição e cooperação entre partículas. 4. Depressão. 5. Suicídio. I. Título.

IMPACTO POTENCIAL DESTA PESQUISA

Ao explorar abordagens para a detecção automática de indícios de depressão e comportamento suicida em textos, esta pesquisa contribui com estratégias complementares que buscam auxiliar na prevenção, identificação e intervenção precoce em transtornos mentais, além do monitoramento da saúde mental em redes sociais.

POTENTIAL IMPACT OF THIS RESEARCH

By exploring approaches for the automatic detection of indicators of depression and suicidal behavior in texts, this research contributes complementary strategies aimed at supporting prevention, identification, and early intervention in mental disorders, as well as monitoring mental health on social networks.

UNIVERSIDADE ESTADUAL PAULISTA
“Júlio de Mesquita Filho”
Instituto de Geociências e Ciências Exatas
Câmpus de Rio Claro

KETERLY GEOVANA GOUVEIA SILVA

PLN APLICADO À SAÚDE MENTAL: EXPLORANDO ABORDAGENS PARA AUXÍLIO
NA CLASSIFICAÇÃO AUTOMÁTICA DE SUICÍDIO E DEPRESSÃO EM TEXTOS

Dissertação de Mestrado apresentada ao
Instituto de Geociências e Ciências Exatas
do Câmpus de Rio Claro, da Universidade
Estadual Paulista “Júlio de Mesquita
Filho”, como parte dos requisitos para
obtenção do título de Mestra em Ciência da
Computação.

Comissão Examinadora

Prof. Dr. Fabricio Aparecido Breve
IGCE/UNESP/Rio Claro (SP)

Profa. Dra. Ana Estela Antunes da Silva
FT/UNICAMP UNICAMP/Limeira (SP)

Prof. Dr. João Paulo Papa
FC/UNESP/Bauru (SP)

Conceito: Aprovado.

Rio Claro (SP), 8 de dezembro de 2025.

AGRADECIMENTOS

Agradeço à minha família, ao meu namorado e aos meus amigos pelo apoio constante ao longo desta jornada, por meio de gestos e palavras que fortaleceram minha confiança e me motivaram a seguir em frente.

Agradeço, também, ao meu orientador, Prof. Dr. Fabrício Aparecido Breve, pela parceria, pelo conhecimento compartilhado e pelas revisões e sugestões que contribuíram significativamente para o avanço desta pesquisa.

RESUMO

A depressão é um transtorno mental que afeta o estado emocional e comportamental do indivíduo, impactando em questões como a sua capacidade de realizar tarefas do dia a dia, se relacionar com o outro e, até mesmo, alterando seus processos fisiológicos. Entre os sintomas observados em pacientes diagnosticados com depressão, a perda de peso, do prazer e humor deprimido são alguns que podem ser listados. Já o suicídio, que muitas vezes é associado a depressão, é um ato individual e intencional executado por aquele que deseja tirar sua própria vida, quando, para ele, é a única alternativa para acabar com um sofrimento constante. A identificação de indícios de um comportamento suicida é fundamental para a diminuição do número de indivíduos que de fato realizam tal ato, da mesma forma que a identificação precoce da depressão pode não apenas proporcionar um tratamento efetivo nos primeiros momentos da condição, mas também evitar que os diagnosticados encontrem o suicídio como o último estágio de sua perturbação emocional. Diante disso, formas de auxiliar na detecção automática dos transtornos mentais são cada vez mais exploradas, se beneficiando principalmente da quantidade de dados disponíveis em redes sociais, local onde muitos usuários se expressam e depositam suas emoções. Este trabalho busca aplicar técnicas de Processamento de Linguagem Natural e Aprendizado de Máquina para avaliar modelos supervisionados e semi supervisionados na detecção de potencial depressão e risco de suicídio em textos extraídos de redes sociais. Foram analisados três conjuntos de dados, um para depressão e dois para suicídio. A extração de características se deu com o modelo BERT, sendo feita em dois cenários diferentes: com pré-processamento e sem pré-processamento dos dados. Posteriormente, a classificação ocorreu com os modelos Random Forest, SVM, Logistic Regression, Label Propagation, Competição e Cooperação entre Partículas, Self Training Classifier e Label Spreading. Além disso, os modelos foram testados com a aplicação de redução de dimensionalidade utilizando o algoritmo PCA. Os experimentos foram realizados com 80% dos dados destinados para treinamento e 20% destinados para teste, onde o modelo LR sem redução de dimensionalidade atingiu 92,91% de acurácia no cenário sem pré-processamento, enquanto a adição do pré-processamento resultou em 90,05%.

Palavras-chave: Processamento de Linguagem Natural; Aprendizado de Máquina; Competição e Cooperação entre Partículas; Depressão; Suicídio.

ABSTRACT

Depression is a mental disorder that affects an individual's emotional and behavioral state, impacting aspects such as their ability to perform daily tasks, relate to others, and even altering physiological processes. Among the symptoms observed in patients diagnosed with depression, weight loss, loss of pleasure, and depressed mood are some that can be listed. Suicide, which is often associated with depression, is an individual and intentional act carried out by someone who wishes to take their own life when, for them, it is the only alternative to end constant suffering. The identification of indications of suicidal behavior is essential to reduce the number of individuals who actually commit such an act, just as the early identification of depression can not only provide effective treatment in the initial stages of the condition, but also prevent those diagnosed from encountering suicide as the final stage of their emotional disturbance. Therefore, methods to assist in the automatic detection of mental disorders have been increasingly explored, mainly benefiting from the large amount of data available on social networks, where many users express themselves and deposit their emotions. This work aims to apply Natural Language Processing and Machine Learning techniques to evaluate supervised and semi-supervised models in the detection of potential depression and suicide risk in texts extracted from social networks. Three datasets were analyzed, one for depression and two for suicide. Feature extraction was performed using the BERT model, considering two different scenarios: with preprocessing and without preprocessing of the data. Subsequently, classification was carried out using the Random Forest, SVM, Logistic Regression, Label Propagation, Particle Competition and Cooperation, Self-Training Classifier, and Label Spreading models. In addition, the models were tested with the application of dimensionality reduction using the PCA algorithm. The experiments were conducted with 80% of the data allocated for training and 20% for testing, in which the LR model without dimensionality reduction achieved 92.91% accuracy in the scenario without preprocessing, while the addition of preprocessing resulted in 90.05%.

Keywords: Natural Language Processing; Machine Learning; Particle Competition and Cooperation; Depression; Suicide.

LISTA DE ILUSTRAÇÕES

Figura 1 – Agrupamento progressivo de unidades no nível sintático	24
Figura 2 – Processo de tokenização de uma frase	25
Figura 3 – Formação de <i>embeddings</i> na primeira camada do modelo BERT . .	30
Figura 4 – Etapa de pré-processamento do modelo BERT	30
Figura 5 – Etapa de ajuste fino do modelo BERT	31
Figura 6 – Abordagem da aprendizagem supervisionada	36
Figura 7 – Classes linearmente separáveis – SVM	37
Figura 8 – Representação de uma curva logística	38
Figura 9 – Aprendizado supervisionado X aprendizado semi-supervisionado . .	40
Figura 10 – Procedimento de duas etapas do GSSL	43
Figura 11 – Detecção de comunidades através do algoritmo de competição de partículas	45
Figura 12 – Geração de partículas para nós rotulados no grafo	46
Figura 13 – Níveis de domínio para nó rotulado	46
Figura 14 – Níveis de domínio para nó não rotulado	47
Figura 15 – Atualização da força de uma partícula	48
Figura 16 – Probabilidades da movimentação de partículas de forma aleatória e gulosa	49
Figura 17 – Fluxograma geral do trabalho proposto	56
Figura 18 – Nuvem de palavras – Tweets potencialmente depressivos do conjunto Mental Health Twitter	58
Figura 19 – Nuvem de palavras – Tweets potencialmente não depressivos do conjunto Mental Health Twitter	58
Figura 20 – Nuvem de palavras – Postagens classificadas como suicídio no con- junto Suicide Detection	60
Figura 21 – Nuvem de palavras – Postagens não classificadas como suicídio no conjunto Suicide Detection	60
Figura 22 – Nuvem de palavras – Postagens classificadas como ideação suicida no conjunto Suicidal Data	62
Figura 23 – Nuvem de palavras – Postagens não classificadas como ideação suicida no conjunto Suicidal Data	62
Figura 24 – Geração de diferentes <i>Word Embeddings</i> para um conjunto de dados	63
Figura 25 – Etapas realizadas no pré-processamento dos textos	64
Figura 26 – Exemplo de conversão de emoji	64
Figura 27 – Divisão dos conjuntos de dados	65
Figura 28 – Mental Health Twitter pré-processado - PCA com modelos RL, SVM e RF	75
Figura 29 – Mental Health Twitter pré-processado - PCA com modelo RF	75

Figura 30 – Mental Health Twitter não processado - PCA com modelos RL, SVM e RF	76
Figura 31 – Mental Health Twitter não processado - PCA com modelo RF	77
Figura 32 – Suicidal Data pré-processado - PCA com modelos RL, SVM e RF .	78
Figura 33 – Suicidal Data pré-processado - PCA com modelo RF	78
Figura 34 – Suicidal Data não processado - PCA com modelos RL, SVM e RF .	79
Figura 35 – Suicidal Data não processado - PCA com modelo RF	80
Figura 36 – Suicide Detection pré-processado - PCA com modelos RL, SVM e RF	81
Figura 37 – Suicide Detection pré-processado - PCA com modelo RF	81
Figura 38 – Suicide Detection não processado - PCA com modelos RL, SVM e RF	82
Figura 39 – Suicide Detection não processado - PCA com modelo RF	83
Figura 40 – Mental Health Twitter pré-processado - PCA com 60% de dados mascarados	91
Figura 41 – Mental Health Twitter pré-processado - PCA com 80% de dados mascarados	92
Figura 42 – Mental Health Twitter não processado - PCA com 60% de dados mascarados	92
Figura 43 – Mental Health Twitter não processado - PCA com 80% de dados mascarados	93
Figura 44 – Suicidal Data pré-processado - PCA com 60% de dados mascarados	94
Figura 45 – Suicidal Data pré-processado - PCA com 80% de dados mascarados	94
Figura 46 – Suicidal Data não processado - PCA com 60% de dados mascarados	95
Figura 47 – Suicidal Data não processado - PCA com 80% de dados mascarados	95
Figura 48 – Suicide Detection pré-processado - PCA com 60% de dados mascarados	96
Figura 49 – Suicide Detection pré-processado - PCA com 80% de dados mascarados	96
Figura 50 – Suicide Detection não processado - PCA com 60% de dados mascarados	97
Figura 51 – Suicide Detection não processado - PCA com 80% de dados mascarados	97
Figura 52 – Mental Health Twitter pré-processado - Autotreinamento com 60% de dados mascarados	99
Figura 53 – Mental Health Twitter pré-processado - Autotreinamento com 80% de dados mascarados	99
Figura 54 – Mental Health Twitter não processado - Autotreinamento com 60% de dados mascarados	100
Figura 55 – Mental Health Twitter não processado - Autotreinamento com 80% de dados mascarados	100

Figura 56 – Suicidal Data pré-processado - Autotreinamento com 60% de dados mascarados	101
Figura 57 – Suicidal Data pré-processado - Autotreinamento com 80% de dados mascarados	102
Figura 58 – Suicidal Data não processado - Autotreinamento com 60% de dados mascarados	103
Figura 59 – Suicidal Data não processado - Autotreinamento com 80% de dados mascarados	103
Figura 60 – Suicide Detection pré-processado - Autotreinamento com 60% de dados mascarados	104
Figura 61 – Suicide Detection pré-processado - Autotreinamento com 80% de dados mascarados	105
Figura 62 – Suicide Detection não processado - Autotreinamento com 60% de dados mascarados	106
Figura 63 – Suicide Detection não processado - Autotreinamento com 80% de dados mascarados	106

LISTA DE TABELAS

Tabela 1 – Sintomas da depressão - DSM IV	20
Tabela 2 – Aplicação dos métodos PorterStemmer e LancasterStemmer para termos em inglês	26
Tabela 3 – Aplicação dos métodos RSLPStemmer e PortugueseStemmer para termos em português	27
Tabela 4 – NLTK - Exemplos de <i>stopwords</i> em português	28
Tabela 5 – NLTK - Exemplos de <i>stopwords</i> em inglês	28
Tabela 6 – Formulário de avaliação dos modelos aplicados em um diálogo . .	50
Tabela 7 – Resultados e comparação entre trabalhos relacionados	51
Tabela 8 – Parâmetros de configuração da RNA	53
Tabela 9 – Valores de Acurácia da RNA e SVM	53
Tabela 10 – Resultados utilizando BERT-RF	54
Tabela 11 – 3 primeiras linhas e 5 colunas do conjunto de dados Mental Health Twitter	57
Tabela 12 – 3 primeiras linhas do conjunto de dados Suicide Detection	59
Tabela 13 – 3 primeiras linhas do conjunto de dados Suicidal Data	61
Tabela 14 – Resultados do conjunto Mental Health Twitter com classificação supervisionada	69
Tabela 15 – Resultados do conjunto Suicidal Data com classificação supervisionada	72
Tabela 16 – Resultados do conjunto Suicide Detection com classificação supervisionada	73
Tabela 17 – Resultados da combinação de modelos para o conjunto Mental Health Twitter	84
Tabela 18 – Resultados da combinação de modelos para o conjunto Suicidal Data	85
Tabela 19 – Resultados da combinação de modelos para o conjunto Suicide Detection	85
Tabela 20 – Melhores resultados obtidos para o conjunto Mental Health Twitter com aprendizado semi-supervisionado	87
Tabela 21 – Melhores resultados obtidos para o conjunto Suicidal Data com aprendizado semi-supervisionado	88
Tabela 22 – Melhores resultados obtidos para o conjunto Suicide Detection com aprendizado semi-supervisionado	89
Tabela 23 – Síntese Geral de Resultados - Classificação Supervisionada	107
Tabela 24 – Síntese Geral de Resultados - Classificação Semi-Supervisionada com mascaramento de 60%	108
Tabela 25 – Síntese Geral de Resultados - Classificação Semi-Supervisionada com mascaramento de 80%	109

Tabela 26 – PCA - Resultados da classificação supervisionada do conjunto Mental Health Twitter pré-processado	120
Tabela 27 – PCA - Resultados da classificação supervisionada do Suicidal Data pré-processado	121
Tabela 28 – PCA - Resultados da classificação supervisionada do Suicide Detection pré-processado	122
Tabela 29 – PCA - Resultados da classificação supervisionada do Mental Health Twitter não processados	123
Tabela 30 – PCA - Resultados da classificação supervisionada do Suicidal Data não processados	124
Tabela 31 – PCA - Resultados da classificação supervisionada do Suicide Detection não processados	125
Tabela 32 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Mental Health Twitter pré-processado	126
Tabela 33 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Mental Health Twitter não processado	126
Tabela 34 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Suicidal Data pré-processado	126
Tabela 35 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Suicidal Data não processado	127
Tabela 36 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Suicide Detection pré-processado	127
Tabela 37 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Suicide Detection não processado	127
Tabela 38 – Resultados obtidos para o conjunto Mental Health Twitter com aprendizado semi supervisionado	128
Tabela 39 – Resultados obtidos para o conjunto Suicidal Data com aprendizado semi supervisionado	129
Tabela 40 – Resultados obtidos para o conjunto Suicide Detection com aprendizado semi supervisionado	130
Tabela 41 – PCA Mental Health Twitter pré-processado - Classificação com 80% dos dados mascarados	131
Tabela 42 – PCA Mental Health Twitter não processado - Classificação com 80% dos dados mascarados	132
Tabela 43 – PCA Mental Health Twitter pré-processado - Classificação com 60% dos dados mascarados	133
Tabela 44 – PCA Mental Health Twitter não processado - Classificação com 60% dos dados mascarados	134

Tabela 45 – PCA Suicidal Data pré-processado - Classificação com 80% dos dados mascarados	135
Tabela 46 – PCA Suicidal Data não processado - Classificação com 80% dos dados mascarados	136
Tabela 47 – PCA Suicidal Data pré-processado - Classificação com 60% dos dados mascarados	137
Tabela 48 – PCA Suicidal Data não processado - Classificação com 60% dos dados mascarados	138
Tabela 49 – PCA Suicide Detection pré-processado - Classificação com 80% dos dados mascarados	139
Tabela 50 – PCA Suicide Detection não processado - Classificação com 80% dos dados mascarados	140
Tabela 51 – PCA Suicide Detection pré-processado - Classificação com 60% dos dados mascarados	141
Tabela 52 – PCA Suicide Detection não processado - Classificação com 60% dos dados mascarados	142
Tabela 53 – Resultados do Autotreinamento para o conjunto Mental Health Twitter com limiar de 75%	142
Tabela 54 – Resultados do Autotreinamento para o conjunto Mental Health Twitter com limiar de 50%	143
Tabela 55 – Resultados do Autotreinamento para o conjunto Suicidal Data com limiar de 75%	143
Tabela 56 – Resultados do Autotreinamento para o conjunto Suicidal Data com limiar de 50%	143
Tabela 57 – Resultados do Autotreinamento para o conjunto Suicide Detection com limiar de 75%	144
Tabela 58 – Resultados do Autotreinamento para o conjunto Suicide Detection com limiar de 50%	144

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Desafios de pesquisa	17
1.2	Objetivos	18
1.3	Hipóteses	18
1.4	Organização do documento	19
2	DEPRESSÃO E SUICÍDIO	20
3	PROCESSAMENTO DE LINGUAGEM NATURAL	22
3.1	Níveis da Língua	22
3.1.1	Nível Morfológico	23
3.1.2	Nível Sintático	23
3.1.3	Nível Semântico	24
3.1.4	Nível Pragmático	24
3.2	Técnicas aplicadas ao processamento de textos	25
3.2.1	Tokenização	25
3.2.2	Stemização	25
3.2.3	Stopwords	27
3.2.4	Word Embeddings com BERT Model	28
4	APRENDIZADO DE MÁQUINA	33
4.1	Aprendizado Supervisionado	35
4.1.1	SVM	36
4.1.2	Logistic Regression	37
4.1.3	Random Forest	38
4.2	Aprendizado Semi-supervisionado	39
4.2.1	TSVM	40
4.2.2	Auto-Treinamento	41
4.2.3	Co-Treinamento	42
4.2.4	Modelos generativos	42
4.2.5	Métodos baseados em grafos	43
4.2.5.1	Competição e Cooperação entre Partículas	44
5	TRABALHOS RELACIONADOS	50
5.1	Modelo de linguagem extenso pré-treinado aprimorado com conhecimento para diagnóstico e tratamento da depressão	50
5.2	Rede LSTM para detecção de suicídio	51
5.3	Detecção de ideação suicida nas redes sociais: uma abordagem de aprendizado de máquina	52

5.4	Detecção de ideação suicida no Twitter usando RNA	53
5.5	Incorporação contextualizada e recursos probabilísticos baseados em novos transformadores para detecção de depressão nas mídias sociais	54
6	METODOLOGIA	56
6.1	Proposta	56
6.2	Conjunto de Dados	57
6.2.1	Mental Health Twitter	57
6.2.2	Suicide Detection	59
6.2.3	Suicidal Data	61
6.3	Pré-processamento e geração de word embeddings	63
6.4	Divisão do conjunto e mascaramento de rótulos	65
6.5	Classificadores supervisionados	66
6.6	Classificadores semi-supervisionados	66
6.7	Redução de dimensionalidade	66
6.8	Combinação de modelos	67
7	RESULTADOS	68
7.1	Classificação Supervisionada	68
7.1.1	Classificação com dados pré-processados e não processados	68
7.1.2	Classificação com redução de dimensionalidade	74
7.1.3	Classificação com agrupamento de modelos	83
7.2	Classificação Semi-Supervisionada	85
7.2.1	Classificação com dados pré-processados e não processados	86
7.2.2	Classificação com redução de dimensionalidade	90
7.2.3	Classificação com Self Training	98
7.3	Discussão	107
8	CONCLUSÃO	111
	APÊNDICE A – TABELAS DE RESULTADOS	120
	DADOS CURRICULARES	145

1 INTRODUÇÃO

A saúde mental pode ser compreendida como um estado de bem-estar onde se é capaz de trabalhar de forma produtiva, lidar com os estresses da vida e perceber as próprias habilidades, sendo um componente essencial da promoção da saúde como um todo. Quando se fala em saúde mental, também se fala em transtornos mentais. A depressão, por exemplo, é um dos transtornos mais diagnosticados (World Health Organization, 2004).

A depressão é um transtorno mental que afeta a mente, altera como o indivíduo manifesta suas emoções e o compromete na totalidade, no âmbito físico e social. Com a manifestação da depressão, os relacionamentos interpessoais são afetados, conflitos surgem e sentimentos de inutilidade se tornam mais intensos devido à diminuição da autoestima (VIEIRA; COUTINHO, 2008).

Com a ocorrência da pandemia da COVID-19, foi observado um aumento no número de diagnósticos psiquiátricos que atestam transtornos mentais como a depressão e ansiedade. No final de 2020, nos Estados Unidos, as taxas de transtorno de ansiedade e depressão atingiram 37% e 30%, respectivamente. Em contrapartida, em 2019, no cenário chamado de “pré-pandemia”, as taxas eram de 8,1% para ansiedade e 6,5% para depressão. Além de um primeiro diagnóstico neurológico ou psiquiátrico estar associado à COVID-19 em pacientes sem histórico prévio, a pandemia também contribuiu para a recaída em pessoas com transtornos pré-existent (TAUSCH et al., 2022).

Em uma entrevista realizada com 1.210 pessoas durante a pandemia da COVID-19, 53% apresentaram sequelas psicológicas classificadas como moderadas ou severas, sendo que mulheres, estudantes, indivíduos que avaliaram sua saúde como ruim ou que apresentavam sintomas da doença foram os mais impactados. No cenário pós-pandemia, em entrevista com 52 mil habitantes chineses, as mulheres permaneceram no grupo dos mais impactados, incluindo também pessoas com mais de 60 anos, bem como indivíduos com maior nível de escolaridade e migrantes (FARO et al., 2020).

Com dados reportados em 2023, a Organização Mundial da Saúde (OMS) estima que 3,8% da população sofra de depressão, incluindo 5% dos adultos e 5,5% dos adultos com mais de 60 anos. A depressão é 50% mais comum entre mulheres, e, mais de 10% das gestantes e puérperas sofrem com esse transtorno (World Health Organization, 2023).

No mundo todo, estima-se que 280 milhões de pessoas sofrem com depressão, e mais de 700.000 pessoas cometem suicídio todos os anos, sendo esta a quarta principal causa de morte entre jovens de 15 a 29 anos (World Health Organization, 2023). O suicídio é um ato intencional onde o indivíduo causa sua própria morte, sendo associado ao fato de que quem o comete não é capaz de identificar outra forma, além da

morte, para a resolução de seus conflitos e angústias (BARBOSA; MACEDO; SILVEIRA, 2011).

Os dados estatísticos da Organização Mundial da Saúde reforçam a importância de um olhar atencioso para a saúde mental, assim como a pandemia da COVID-19 evidenciou ainda mais como as mudanças na vida dos indivíduos afetam seu estado emocional. É importante que o tratamento da depressão, por exemplo, seja realizado de forma correta e se torne efetivo, bem como a identificação do risco de suicídio pode contribuir para a prevenção.

Pesquisas já se direcionam para a busca de abordagens que possam auxiliar na identificação de transtornos mentais. Com o uso crescente das redes sociais, mais dados ficam disponíveis todos os dias e este é, no momento, uma das fontes predominantes nas pesquisas que buscam promover a detecção automática de indícios de ansiedade, depressão e outros transtornos mentais, além do suicídio.

Usuários utilizam as redes sociais para exporem suas experiências, opiniões, emoções, discutirem entre si e podem demonstrar indícios de depressão e risco de suicídio. Os textos são capazes de indicar o sentimento daquele que o escreve, como é possível compreender pela análise de Rios (2021) nas postagens de uma jovem no Instagram que cometeu suicídio. A análise indica que o conteúdo das publicações possibilitaram a identificação de sentimentos como tristeza e culpa.

Este processo foi realizado por uma profissional de psicologia, sem o uso de técnicas de classificação automática e permitiu a compreensão de que as publicações apresentam palavras possíveis de serem utilizadas para a determinação de padrões textuais e na detecção automática de textos que sigam estes padrões.

Para a detecção automática de textos, técnicas de Processamento de Linguagem Natural (PLN) e Aprendizado de Máquina podem ser aplicadas. Várias pesquisas já foram e vem sendo desenvolvidas com esta proposta, mas, por se tratar de saúde mental, encontrar um conjunto de dados rotulados por profissionais especializados na área ou um expressivo número de exemplos mais representativos é um dos principais desafios.

Diante disso, a pesquisa busca explorar diferentes abordagens para o auxílio na classificação automática de potencial depressão e risco de suicídio em textos representados por *Word Embeddings*. Para isso, técnicas de redução de dimensionalidade, aprendizado supervisionado e semi-supervisionado serão avaliadas, além da análise de como as etapas de pré-processamento ou a ausência delas influenciam nos resultados obtidos.

1.1 Desafios de pesquisa

O presente trabalho busca responder às seguintes perguntas de pesquisa:

- A utilização do modelo BERT para a geração de *embeddings* garante um bom

desempenho dos classificadores?

- O pré-processamento aplicado aos conjuntos impacta no resultado obtido pela classificação de embeddings BERT?
- A redução de dimensionalidade das embeddings geradas gera impacto positivo na classificação final?
- Algoritmos semi-supervisionados atingem bom desempenho no rastreio de sintomas de potencial depressão e risco de suicídio?
- Existem técnicas de aprendizagem supervisionada que superam ou se igualam aos modelos semi-supervisionados quando a mesma quantidade de dados rotulados esta disponível?

1.2 Objetivos

Esta pesquisa tem como objetivo explorar diferentes abordagens para o auxílio na classificação automática de potencial depressão e risco de suicídio em textos representados por *Word Embeddings*. Para isso, o presente trabalho tem como objetivos específicos:

- Implementar os algoritmos Máquina de Vetores de Suporte, Floresta Aleatória e Regressão Logística para à detecção de depressão e risco de suicídio em uma abordagem supervisionada;
- Implementar os algoritmos de Propagação de Rótulos, Espalhamento de Rótulos, Autotreinamento e Competição e Cooperação entre Partículas para à detecção de depressão e risco de suicídio em uma abordagem semi supervisionada;
- Comparar o impacto da aplicação de técnicas de pré-processamento na classificação dos textos;
- Aplicar redução de dimensionalidade utilizando o algoritmo de Análise de Componentes Principais e avaliar seu impacto;
- Realizar o comparativo entre aprendizagem supervisionada e semi supervisionada no contexto de saúde mental

1.3 Hipóteses

Com base nos objetivos e questões de pesquisa apresentados, este trabalho levanta as seguintes hipóteses:

1. As *embeddings* geradas pelo modelo BERT são contextuais e garantem um bom desempenho das classificações, mesmo com poucos dados disponíveis.

2. A aplicação de pré-processamento impacta no desempenho dos modelos, podendo gerar perda de informações relevante para uma classificação correta.
3. O impacto da redução de dimensionalidade depende da variância mantida e pode afetar negativamente o desempenho de alguns modelos.
4. Métodos de aprendizado semi-supervisionado são capazes de identificar, com desempenho satisfatório, indícios de sintomas de depressão e de comportamento suicida em textos, mesmo com uma quantidade reduzida de dados rotulados.
5. Algoritmos semi-supervisionados atingem bom desempenho na classificação dos textos, sendo o resultado afetado pelas configurações utilizadas.
6. Existem abordagens semi supervisionadas que podem gerar resultados próximos, mas não superiores ao desempenho obtido por modelos supervisionados treinados com a mesma quantidade de dados rotulados.

1.4 Organização do documento

Este trabalho apresenta 8 capítulos, organizados da seguinte forma:

Capítulo 2: Apresenta o conceito, sintomas e características da depressão e do suicídio;

Capítulo 3: Detalha o Processamento de Linguagem Natural, apresentando conceitos, técnicas para pré-processamento e geração de *embeddings* com o modelo BERT;

Capítulo 4: Explora o Aprendizado de Máquina, apresentando o aprendizado supervisionado e semi-supervisionado, além de alguns algoritmos de cada categoria;

Capítulo 5: Aponta pesquisas que também buscam detectar depressão e suicídio de forma automática;

Capítulo 6: Detalha a metodologia, apresentando a proposta da pesquisa, conjunto de dados e outras implementações;

Capítulo 7: Apresenta resultados obtidos na exploração das abordagens para classificação dos textos;

Capítulo 8: Conclui o documento, abordando considerações e apontando os trabalhos futuros;

Apêndice A: Apresenta tabelas com todos os resultados obtidos durante os experimentos.

2 DEPRESSÃO E SUICÍDIO

Depressão e suicídio são temas considerados sensíveis quando abordados em discussões. A depressão, transtorno mental do grupo de transtornos depressivos, é um dos mais diagnosticados atualmente. Compreende-se que transtornos mentais são quadros clínicos que ocasionam perturbação na cognição, regulação emocional ou no comportamento, alterando processos psicológicos, biológicos e outros relacionados ao funcionamento da mente (American Psychiatric Association, 2014).

Transtornos mentais, de modo geral, devem ser diagnosticados e tratados por profissionais especializados em saúde mental, podendo envolver o uso de medicamentos, internações e terapias. A depressão é diagnosticada considerando os sintomas manifestados, sua duração, frequência e intensidade, e tem o Manual Diagnóstico e Estatístico de Transtornos Mentais - DSM-IV como uma de suas bases (RODRIGUES, 2000). Para ser diagnosticado com o quadro clínico de depressão, os sintomas, como os exibidos na Tabela 1, devem ser manifestados na maior parte do dia e com permanência de pelo menos duas semanas (American Psychiatric Association, 2014).

Tabela 1 – Sintomas da depressão - DSM IV

Tristeza	Perda de peso
Ganho de peso	Diminuição do interesse
Perda de prazer	Insônia
Sensação de vazio	Hipersonia
Falta de esperança	Indecisão
Aumento do apetite	Fadiga
Sentimento de inutilidade	Culpa excessiva
Perda de energia	Falta de concentração
Ideação suicida	Tentativa de suicídio
Diminuição da capacidade de pensar	Redução do apetite

Fonte: Adaptado de American Psychiatric Association (2014)

Tais sintomas, quando frequentes, comprometem toda a vida do indivíduo, suas relações familiares, atividades profissionais e convivência social. Sentimentos de inferioridade, incapacidade, isolamento e falta de perspectiva futura, afetam drasticamente a saúde física e psíquica do sujeito acometido por esses sintomas.

A depressão também pode ser vista em uma associação com o suicídio, uma vez que indivíduos acometidos pelo transtorno apresentam maior risco de atentar contra a própria vida (CHACHAMOVICH et al., 2009). Entende-se por suicídio a ação individual executada visando causar a própria morte por meios considerados letais, como o envenenamento ou o uso de armas de fogo (CFM, 2014). Ele pode ser motivado por distorções cognitivas e desejos de finalizar o estado emocional de sofrimento, onde, ao pensar na morte, o indivíduo pode expressar a vontade e também um planejamento detalhado que inclui o lugar para realizar o suicídio (PORTO, 1999).

O suicídio pode ser planejado, incluindo a escrita de bilhetes e cartas que relatam o estado de sofrimento e desespero. Alguns podem ser impulsivos, mas, de modo geral, o comportamento suicida se inicia na ideação e termina no ato concluído. A ideação envolve os pensamentos e desejos de morte, partindo para a tentativa de por um fim a vida até que isso de fato ocorra (FUKUMITSU, 2014). O risco do ato consumado aumenta à medida em que o número de tentativas é observado com maior frequência (VIDAL; GONTIJO; LIMA, 2013). No entanto, mesmo que possa ser associado, o suicídio não está estritamente relacionado aos transtornos mentais. O ato pode ocorrer devido a momentos impulsivos, crises ou a sensação de incapacidade à frente de problemas (World Health Organization 2, 2023).

Pacientes diagnosticados com câncer, por exemplo, podem expressar o desejo de colocar um fim em suas vidas após um prognóstico, devido à dor e outros sintomas decorrentes da doença, assim como podem sofrer de um transtorno depressivo. Por conta disso, qualquer ideação suicida, independentemente do motivo aparente, deve ser avaliada criteriosamente para uma intervenção efetiva (MASSIE; GAGNON; HOLLAND, 1994).

O suicídio já é considerado um problema de saúde pública, sendo evidenciado que 700 mil pessoas morrem por este motivo todos os anos e que isso pode ser evitado. O Plano de Ação para Saúde Mental da OMS 2013-2030 visa reduzir a taxa de suicídio em um terço até 2030. Além disso, dificultar o acesso a meios letais, bem como a identificação precoce do comportamento suicida, são outras formas de contribuir para a diminuição do número de mortes decorrentes desta ação (World Health Organization 2, 2023).

A possibilidade de utilizar os inúmeros dados expostos em redes sociais como uma forma de rastrear indícios de um comportamento suicida também se enquadra na ação de prevenção. A rede social Facebook, por exemplo, já foi utilizada para um relato de despedida antes de um jovem encerrar sua vida, assim como outro jovem relatou o desejo de morrer antes de consumir o suicídio ao queimar seu corpo (RIOS, 2021). Dessa forma, o presente projeto aborda o auxílio na identificação de sintomas de depressão e risco de suicídio através de textos obtidos das redes sociais Reddit e Twitter, aplicando aprendizado de máquina para uma análise automatizada.

3 PROCESSAMENTO DE LINGUAGEM NATURAL

Processamento de linguagem natural (PLN) é uma área relacionada a inteligência artificial e a linguística computacional e busca dar às máquinas a habilidade de compreender a língua falada ou escrita pelos seres humanos, língua esta usada no dia a dia para a comunicação. No processamento da fala, diferente da escrita, o volume, sotaque e entonação são características que podem facilitar ou dificultar o processo de reconhecimento. De modo geral, o PLN pode ser compreendido como uma área que se divide entre a busca pela compreensão da linguagem natural e a geração de linguagem natural através da máquina (CASELI; NUNES, 2024).

Na compreensão da linguagem natural, as máquinas aprendem a analisar a língua, extraindo conceitos, identificando emoções, palavras-chave, entidades, entre outros aspectos. A compreensão de problemas enfrentados por clientes, quando expressos de forma escrita ou falada, é um exemplo de aplicação que utiliza a NLU (*Natural Language Understanding* - Compreensão da Linguagem Natural) voltada à área de atendimento. Por outro lado, a geração de linguagem natural permite que a máquina construa frases, sentenças e parágrafos a partir de uma representação interna. No entanto, não basta apenas gerar textos de forma aleatória. O objetivo é que as saídas obtidas sejam significativas (KHURANA et al., 2023).

Enfatizando o processamento de linguagem natural aplicado a língua escrita e compreendendo que existem diferentes domínios como a saúde e o direito, é possível citar, como um exemplo crescente, o uso do PLN aplicado às redes sociais. A rede social é um conjunto de atores e conexões onde são expressas opiniões, gírias e emojis são utilizados como forma de comunicação, e a linguagem é inclinada para a utilizada no cotidiano, não muito formal. Redes sociais como Facebook, Reddit e Twitter (atualmente X) são fontes para a obtenção de textos que podem ser utilizados em tarefas de PLN, sendo o Twitter a rede predominante nas pesquisas da área (CASELI; NUNES, 2024).

Além disso, o PLN contribuiu e contribui para a categorização automática de notícias, geração de resumos, *chatbots* e motores de busca mais precisos. Algumas técnicas para o processamento de textos serão apresentadas na seção seguinte, etapa crucial para que as aplicações citadas como exemplo possam ser de fato implementadas.

3.1 Níveis da Língua

Ao compreender sobre processamento de linguagem natural, também é importante compreender que a língua possui diferentes níveis e campos de estudo, uma vez que eles compõem o processamento nas abordagens que utilizam PLN (SONBOL; REBDAWI; GHNEIM, 2022). Os diferentes níveis da língua são conhecidos como: fonético ou fonológico; morfológico; sintático; semântico; pragmático e discursivo. A

seguir são apresentados, com mais detalhes, níveis que podem ser identificados na metodologia escolhida para esta pesquisa.

3.1.1 Nível Morfológico

A morfologia, ou nível morfológico, compreende a estrutura das palavras e de que forma é possível combinar morfemas. Os morfemas são partes de palavras que representam as menores unidades de significado da língua. Quando combinados, os morfemas permitem a criação de novas palavras. (KHURANA et al., 2023; PAJILA et al., 2023).

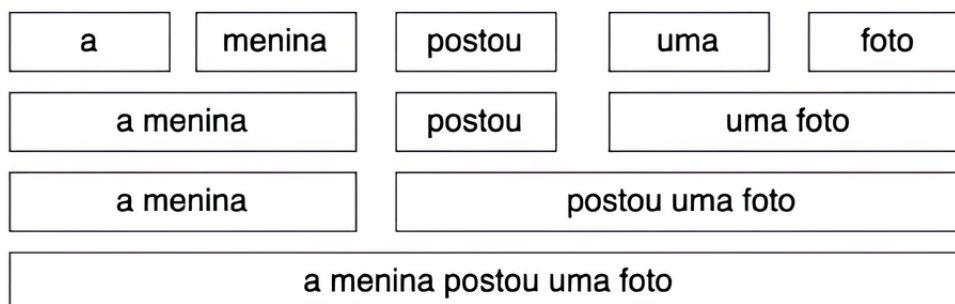
Palavras podem ser divididas em mais de um morfema para que seja analisada morfológicamente e tenha seu significado compreendido. Quando uma palavra não pode ser dividida em partes menores que contenham significado, ela é considerada um morfema lexical. Morfemas lexicais possuem significado por si mesmas como, por exemplo, as palavras mesa e casa (KHURANA et al., 2023).

Neste campo de estudo também é analisada a flexão, ação na qual palavras apresentam variação gramatical que podem indicar mudança de tempo (passado, presente, futuro) e gênero. Para indicar que uma ação ocorreu no passado, por exemplo, um verbo pode sofrer flexão e transformar “*I work*” (Eu trabalho) em “*I worked*” (Eu trabalhei). Também é possível formar palavras com a adição de prefixos e sufixos, derivando a palavra original e criando uma nova com um novo significado. O prefixo “*dis*” ou “*in*”, no português, é um exemplo que, ao ser atribuído à palavra “*satisfied*” (satisfeito), cria o termo “*dissatisfied*” (insatisfeito) e dá o significado oposto ao que era compreendido na palavra original (PAJILA et al., 2023).

3.1.2 Nível Sintático

A sintaxe é a área que estuda de que forma as palavras se organizam. Por meio dela entende-se as regras que determinam como as palavras se agrupam e quais agrupamentos são considerados corretos. Os agrupamentos funcionam como pequenos “tijolos” combinados para formar um significado. Eles se juntam em unidades maiores até que seja construído um significado relevante para determinada frase. A Figura 1 representa esta analogia, onde a junção progressiva de cada palavra resulta, no final, em uma frase. (CASELI; NUNES, 2024).

Por meio da sintaxe é possível transmitir significado, uma vez que, na maioria dos idiomas, a ordem e a dependência das palavras contribuem para isso. Duas frases com as mesmas palavras apresentam significados diferentes conforme a posição de cada uma delas, ou seja, a ordem. A frase “o cachorro perseguiu o gato” e a frase “o gato perseguiu o cachorro” exemplificam este conceito (LIDDY, 2001).

Figura 1 – Agrupamento progressivo de unidades no nível sintático

Fonte: Caseli e Nunes (2024)

3.1.3 Nível Semântico

A semântica é o campo que se dedica a análise do significado das palavras. No PLN, este é um nível onde a estrutura de uma frase é analisada para que seja possível identificar as palavras mais importantes presentes nela e seus relacionamentos. Neste ponto, palavras com diversos significados podem ser interpretadas e, para inferir seu significado, é levado em consideração de que forma ela está presente em uma determinada frase (PAJILA et al., 2023).

Para determinar os possíveis significados que uma frase pode apresentar, sua estrutura lógica é processada, de modo que palavras relevantes se tornam conhecidas e possibilitam compreender a relação entre palavras. Com isso, uma frase que se refere a um filme, mas não utiliza o termo em específico, ainda pode ser compreendida por meio do uso de palavras como “actress” (atriz) ou “actor” (ator) (KHURANA et al., 2023).

3.1.4 Nível Pragmático

A pragmática é o campo de estudo que se concentra na forma como a linguagem é utilizada em um contexto que pode ser linguístico ou situacional. “*Could you pass the salt?*” (Você poderia passar o sal?), por mais que seja gramaticamente uma pergunta, pode ser interpretada como um pedido em algumas situações, mesmo que não esteja no modo imperativo. Além disso, a pragmática permite determinar a entidade a que um pronome se refere e preencher informações ausentes (BRISCOE, 2011).

O contexto em que as palavras estão inseridas influencia na compreensão de um texto e, quando uma frase não é específica e seu contexto não contribui para sua especificação, a compreensão passa a ser ambígua. A ambiguidade pragmática ocorre quando um texto possui diferentes interpretações por pessoas diferentes, sendo estes significados dependentes do contexto e do conhecimento prévio que o leitor tem sobre o conteúdo (KHURANA et al., 2023).

3.2 Técnicas aplicadas ao processamento de textos

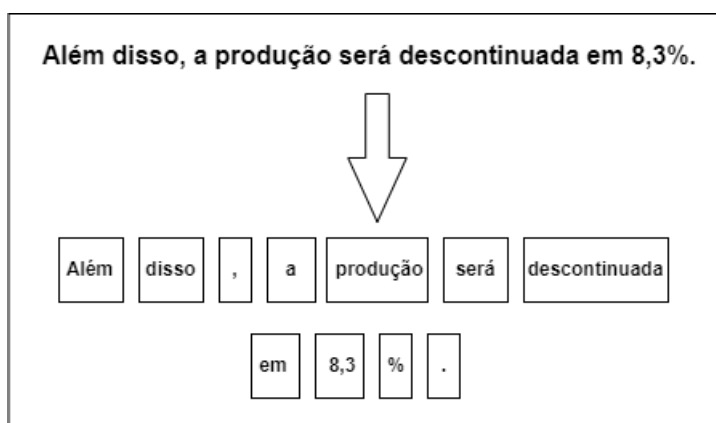
O pré-processamento dos dados textuais é uma tarefa importante no PLN. Pré-processar os dados pode incluir a remoção de caracteres especiais e emojis, principalmente quando tais textos são retirados de redes sociais, além da remoção de outras características indesejadas. Nas subseções seguintes serão detalhadas algumas técnicas: stemização, tokenização, remoção de *stopwords* e a representação de textos por meio de *word embeddings*.

3.2.1 Tokenização

A técnica de tokenização é utilizada para dividir uma sequência de caracteres em unidades menores, denominados tokens. Neste processo, alguns caracteres e pontuações podem ser filtrados. No inglês é comum dividir palavras pelos espaços, mas entidades nomeadas, por exemplo, podem ser adicionadas para isso (SUN; LUO; CHEN, 2017).

Em textos escritos na língua portuguesa, os tokens podem ser gerados através da separação de palavras por meio de delimitadores como espaços em branco, vírgula (,), dois-pontos (:), ponto e vírgula (;), ponto (.) e hífen (-). A frase “Além disso, a produção será descontinuada em 8,3%.” é um exemplo que pode ser tokenizado e obter 11 tokens como resultado, sendo 8,3 um pedaço da frase que não deve ser dividido, mesmo que uma vírgula esteja presente. A Figura 2 exibe todos os tokens desta frase (CASELI; NUNES, 2024).

Figura 2 – Processo de tokenização de uma frase



Fonte: Adaptado de Caseli e Nunes (2024)

3.2.2 Stemização

Stemming é a técnica que busca reduzir palavras à sua raiz, fazendo com que seu contexto na frase não seja considerado. Com este processo, diversas palavras que ocorrem em textos diferentes e que são semanticamente próximas podem ser

substituídas por seus radicais (KAUR; KAUR, 2023; GADRI; MOUSSAOUI, 2015). Para reduzir uma palavra, um algoritmo, conhecido como *stemmer*, remove os afixos que contém informações gramaticais ou lexicais desta palavra. O algoritmo Porter, por exemplo, realiza este processo se baseando em regras para a remoção do sufixo de uma palavra (SOUZA et al., 2021).

Existem casos onde a aplicação da técnica de stemming pode resultar em uma redução excessiva ou insuficiente de palavras. A redução excessiva ocorre quando duas palavras diferentes são reduzidas ao mesmo radical. A redução insuficiente, por outro lado, ocorre quando duas palavras iguais retornam resultados diferentes após a aplicação da técnica, resultando em reduções com radicais diferentes (KOIRALA; SHAKYA, 2020).

A Tabela 2 demonstra a utilização dos métodos *LancasterStemmer* (LS) e *PorterStemmer* (PS) para termos em inglês, e a Tabela 3 apresenta *RSLPStemmer* (RS) e *PortugueseStemmer* (PtS) aplicados ao português. Todos estão disponíveis na biblioteca NLTK¹.

Tabela 2 – Aplicação dos métodos PorterStemmer e LancasterStemmer para termos em inglês

Original	LP	PS	Original	LP	PS
love	lov	love	buyer	buy	buyer
loving	lov	love	buy	buy	buy
happy	happy	happi	bought	bought	bought
happiness	happy	happi	friend	friend	friend
happily	happy	happili	friends	friend	friend
unhappy	unhappy	unhappi	friendship	friend	friendship
marriage	marry	marriag	study	study	studi
married	marry	marri	student	stud	student
logic	log	logic	students	stud	student
logically	log	logic	work	work	work
kind	kind	kind	worked	work	work
kindness	kind	kind	employed	employ	employ
possible	poss	possibl	unemployment	unemploy	unemploy
impossible	imposs	imposs	unemployed	unemploy	unemploy
possibly	poss	possibl			

Fonte: Elaborado pela autora

Com base nas tabelas, é possível visualizar que algumas palavras apresentam resultados diferentes dependendo do algoritmo aplicado. Alguns métodos podem ser mais agressivos na redução das palavras, o que pode gerar maior perda de informação e um mesmo radical para termos com significados diferentes.

¹<https://www.nltk.org/>

Tabela 3 – Aplicação dos métodos RSLPStemmer e PortugueseStemmer para termos em português

Original	PtS	RS	Original	PtS	RS
amor	amor	am	amizade	amizad	amizad
amoroso	amor	amor	estudo	estud	estud
feliz	feliz	feliz	estudante	estud	estud
felicidade	felic	felic	estudantes	estud	estud
felizmente	feliz	feliz	trabalho	trabalh	trabalh
infeliz	infeliz	infeliz	trabalhado	trabalh	trabalh
casamento	casament	cas	empregado	empreg	empreg
casado	cas	cas	desemprego	desempreg	desempreg
logica	logic	logic	desempregado	desempreg	desempreg
logicamente	logic	logic	comprador	comprador	compr
gentil	gentil	gentil	comprar	compr	compr
gentileza	gentil	gentil	comprado	compr	compr
possivel	possivel	possi	amigo	amig	amig
impossivel	impossivel	impossi	amigos	amig	amig
possivelmente	possivel	possi			

Fonte: Elaborado pela autora

Além disso, durante a escolha é importante avaliar se o algoritmo é específico para determinada linguagem, ou se está bem ajustado ao idioma que se deseja trabalhar.

3.2.3 Stopwords

As *stopwords* são palavras que aparecem com muita frequência em um determinado documento e consideradas palavras com pouco peso semântico. Durante o pré-processamento, utilizando uma lista de termos, as *stopwords* são removidas. Apesar de ser uma técnica que pode trazer benefícios para a recuperação de informação, remover *stopwords* também pode ser vista como uma desvantagem quando dificulta a busca por textos que contenham termos considerados irrelevantes (JURAFSKY; MARTIN, 2023).

Uma lista de stopwords pode ser criada classificando um vocabulário com base na frequência e definindo como *stopwords* um número específico de palavras principais, assim como é possível utilizar uma lista padrão. A utilização de listas pré-definidas, como as fornecidas pela biblioteca NLTK, se tornou o método mais comum aplicado às pesquisas. No entanto, uma vez que a importância de cada palavra a nível semântico depende do problema a ser resolvido, listas genéricas podem resultar em perda de informação e afetar o desempenho de modelos (JURAFSKY; MARTIN, 2023; FAN; ARORA; TREUDE, 2023).

Na Tabela 4 estão algumas das *stopwords* retornadas pela biblioteca NLTK 3.8.1 para o idioma português, enquanto a Tabela 5 exibe alguns termos em inglês.

Nota-se que estes termos, se removidos sem uma análise, podem resultar

Tabela 4 – NLTK - Exemplos de *stopwords* em português

a	aos	aquela	depois	estávamos
estivemos	houveríamos	sem	ela	minha
for	ser	elas	e	meus
será	mais	isto	mas	tem
eu	era	ele	tu	meu
houvermos	não	tua	estejamos	estes
estive	você	nós	nossa	os
vocês	ou	para	pela	das

Fonte: Elaborado pela autora

Tabela 5 – NLTK - Exemplos de *stopwords* em inglês

a	of	that	after	were
we've	whom	he	she	my
myself	be	they	and	any
will	more	this	but	has
I	was	we	you	now
have	no	not	as	because
been	by	them	each	few
or	from	for	her	i'm

Fonte: Elaborado pela autora

na perda de informações importantes em um problema de análise de sentimentos, por exemplo. “*I’m not happy*” (Eu não estou feliz) e “*I’m happy*” (Eu estou feliz) possuem significados diferentes conforme a palavra “não” é inserida. Com base na lista de *stopwords* retornada pela biblioteca, esta palavra seria excluída. Por conta disso, é importante que o conjunto de *stopwords* seja personalizado de acordo com cada objetivo e domínio da pesquisa.

3.2.4 Word Embeddings com BERT Model

Transformar dados textuais em representações numéricas é um passo necessário para que o processo de classificação seja realizado com sucesso. *Word Embedding* é uma forma de capturar o significado de cada palavra e transformar textos em números reais, de modo que essa transformação é representada na forma de vetores n-dimensionais. Cada palavra é associada a um valor numérico e mapeada para uma posição espaço vetorial. Neste espaço, Palavras com significados e contextos semelhantes apresentam uma distância menor entre si, ou seja, possuem *embeddings* semelhantes. No entanto, palavras com significados diferentes tendem a ter *embeddings* diferentes (NEELIMA; MEHROTRA, 2023).

A criação de *embeddings* pode ser feita de modo estático, com frameworks como *word2vec*, ou de modo contextual com modelos como BERT (*Bidirectional Encoder Representations from Transformers*). Na abordagem estática, a representação de

cada palavra é independente do contexto, ou seja, cada palavra terá o mesmo valor para um vocabulário determinado, sem considerar o que está à sua volta (SARAKUL; RUTHERFORD, 2023).

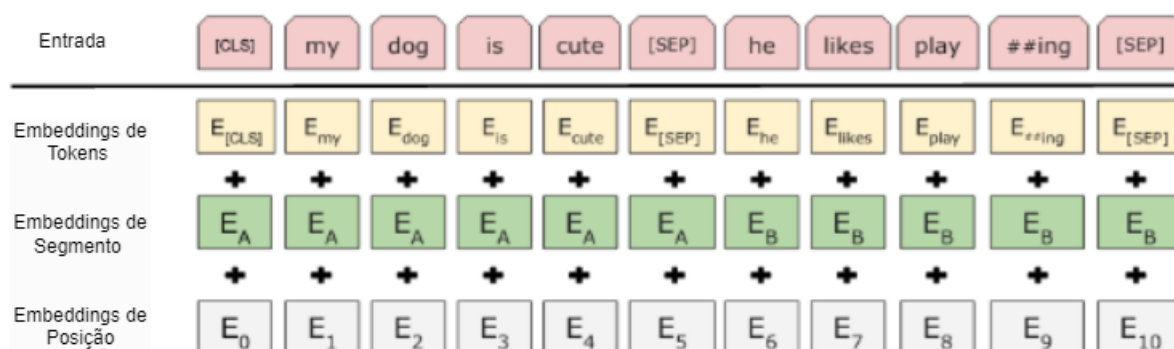
Na abordagem contextual, onde ocorre a criação de *embeddings* dinâmicos, a representação de cada palavra é modificada com base no contexto em que ela está inserida. A possibilidade de existir diferentes representações de uma palavra com base no contexto é considerada uma melhoria nas tarefas de PLN (PATIL et al., 2023). Para esta abordagem, como citado anteriormente, o modelo BERT é um dos modelos que podem ser aplicados para a geração de *embeddings* e é a escolha do trabalho atual.

O *Bidirectional Encoder Representations from Transformers* é baseado na arquitetura Transformer, arquitetura esta que utiliza mecanismos de atenção para capturar dependências globais entre entrada e saída, sem a necessidade de que tais mecanismos sejam combinados com uma rede neural recorrente. O Transformer segue a estrutura codificador-decodificador, utilizando camadas empilhadas de autoatenção. De modo geral, na estrutura codificador-decodificador, a camada de codificação é responsável por transformar uma determinada sequência de símbolos, como palavras de uma frase, em uma sequência de representações numéricas contínuas. Na camada de decodificação, com base na sequência gerada pelo codificador, uma nova sequência de saída de símbolos é então construída. Os símbolos gerados servem como entrada adicional para a geração do próximo símbolo, visto que o modelo é auto-regressivo. (VASWANI et al., 2023).

A arquitetura do modelo BERT permite resolver tarefas que envolvem dados sequenciais e é capaz de lidar com dependência de longo alcance, não exigindo que as sequências sejam processadas em ordem ou que estejam em posições próximas (GHANEM; ERBAY; BAKOUR, 2023).

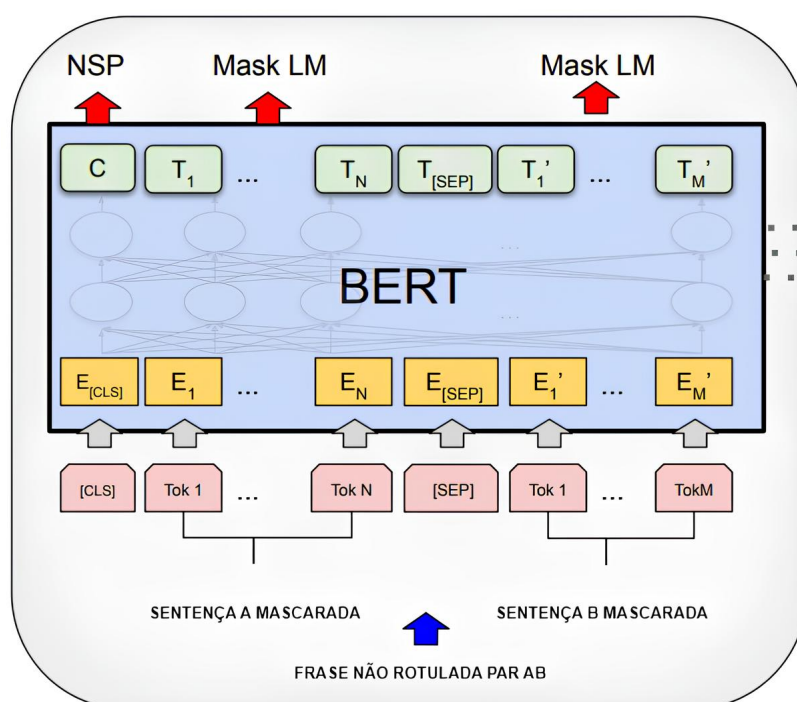
BERT, em sua configuração padrão, também conhecido como BERTBase, possui 12 camadas codificadoras do Transformer, 768 unidades ocultas e 12 cabeças de autoatenção. A Figura 3 demonstra de que forma os *embeddings* são gerados pelo modelo BERT. Para um determinado token, os *embeddings* de entrada são resultados da soma dos *embeddings* das outras três categorias: token, segmentação e posição. [CLS] é adicionado no início de todas as entradas, indicando que todo token que inicia uma sequência é um token de classificação especial. Já [SEP] é inserido para diferenciar sentenças em dois caminhos. Desta forma, ao adicionar uma *embedding* para cada token, é possível identificar se ele pertence à sentença A, finalizada no primeiro token [SEP], ou a sentença B (DEVLIN et al., 2019).

O funcionamento do modelo BERT consiste em duas etapas principais: o pré-treinamento e o ajuste fino. Na etapa de pré-treinamento, o modelo utiliza representações bidirecionais não rotuladas para capturar o significado das palavras. Isso quer dizer que ele considera, em todas as camadas, o contexto esquerdo e direito em que a

Figura 3 – Formação de *embeddings* na primeira camada do modelo BERT

Fonte: Adaptado de Devlin et al. (2019)

palavra está inserida. A Figura 4 demonstra visualmente a etapa de pré-processamento do modelo (DEVLIN et al., 2019).

Figura 4 – Etapa de pré-processamento do modelo BERT

Fonte: Adaptado de Devlin et al. (2019)

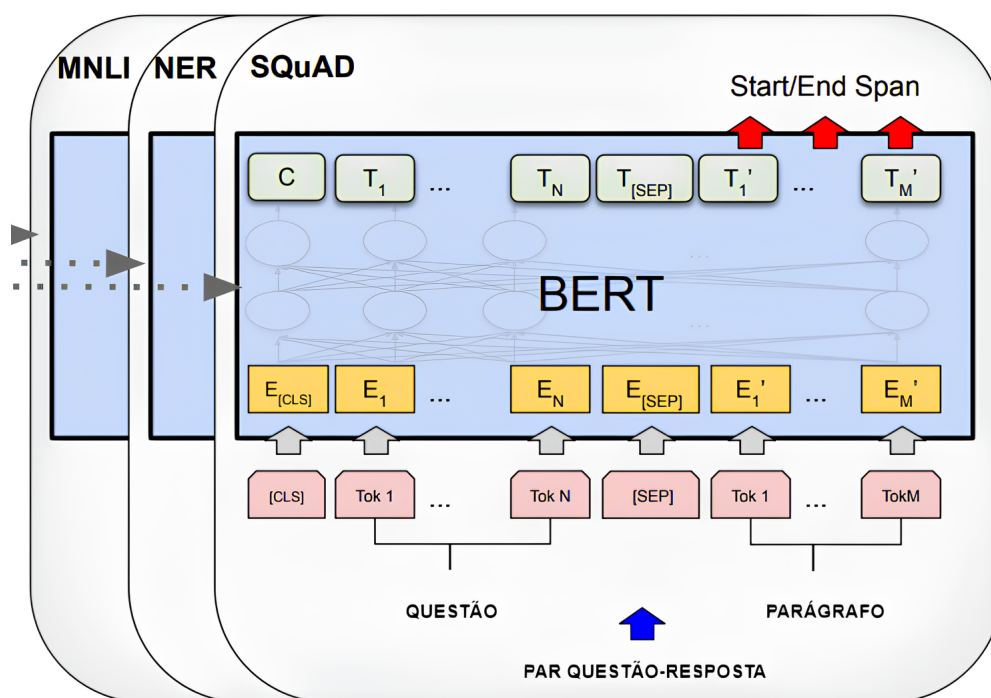
O pré-processamento do modelo utiliza duas estratégias: A Modelagem de Linguagem Mascarada (MLM) e a Predição da Próxima Frase (NSP), conforme mostra a Figura 4. A MLM é uma estratégia primária onde uma sequência, antes de ser inserida no modelo BERT, passa por um processo especial de tokenização. Além de ter a sentença separada em unidades menores (tokens), 15% dos tokens gerados são substituídos pelo token especial [MASK] e seus valores originais correspondentes precisam ser previstos. Nesta estratégia, o modelo aprende a prever as palavras

marcadas e completar corretamente a sequência com base no contexto construído pelos outros tokens não marcados (SEYYAR; YAVUZ; ÜNVER, 2022)

Na estratégia seguinte, NSP, frases são agrupadas em pares para serem utilizadas como entrada no treinamento do modelo BERT. Este processo busca ensinar o BERT a aprender o relacionamento entre as frases, prevendo se uma determinada frase B (segunda frase) é a continuação de uma frase A, ou seja, a próxima frase do documento. As entradas do treinamento são divididas com 50% delas contendo a segunda frase no documento, ou seja, a próxima frase do documento original, e 50% contendo a segunda frase selecionada de forma aleatória. Com um treinamento bem sucedido, o modelo se torna capaz de identificar que uma frase escolhida aleatoriamente em um corpus não está relacionada à primeira frase (SEYYAR; YAVUZ; ÜNVER, 2022).

A etapa de ajuste fino possibilita a execução de uma variedade de tarefas como, por exemplo, a de geração de respostas para uma determinada pergunta, conforme a Figura 5 ilustra. O ajuste do modelo se dá sem a necessidade de que sua arquitetura seja modificada para esta tarefa em específico, uma vez que é possível adaptá-lo por meio da inclusão de uma camada de saída adicional. No processo, o modelo é primeiro inicializado com parâmetros pré-treinados e depois, tais parâmetros são ajustados utilizando dados rotulados de tarefas seguintes (GHANEM; ERBAY; BAKOUR, 2023; DEVLIN et al., 2019).

Figura 5 – Etapa de ajuste fino do modelo BERT



Fonte: Adaptado de Devlin et al. (2019)

Por mais que seja visto como uma escolha para atingir altos desempenhos em tarefas de PLN, custo computacional e memória são pontos relevantes a serem

considerados para esta escolha. Diante disso, e também buscando melhores resultados, algumas variações do modelo BERT são propostas e utilizadas, como o modelo DistilBERT (GHANEM; ERBAY; BAKOUR, 2023)

DistilBERT é uma versão desenvolvida pela Hugging Face para ser menor, mais rápida e com custo computacional reduzido quando comparado ao modelo BERT. Para que essa melhoria seja obtida, DistilBERT utiliza a técnica de destilação de conhecimento. A destilação do conhecimento é uma técnica de compressão onde um modelo compacto é treinado e reproduz o comportamento de um modelo maior (ARDIMENTO, 2022).

4 Aprendizado de Máquina

À medida que a tecnologia avança e dados de diferentes formatos tornam-se acessíveis no dia a dia, a busca por aplicações com o uso da inteligência artificial também cresce.

Neste ponto, o aprendizado de máquina surge como uma sub-área da inteligência artificial, visando programar computadores de forma que aprendam com dados de exemplo ou experiências passadas. O aprendizado de máquina utiliza teorias da estatística para a criação de modelos preditivos, descritivos, ou ambos, podendo prever o futuro ou extrair conhecimento de uma amostra de dados (ALPAYDIN, 2004). Dentre as diferentes soluções que aplicam a técnica, é possível encontrar, por exemplo, programas para visão computacional e processamento de linguagem natural (JORDAN; MITCHELL, 2015).

De modo geral, o problema a ser resolvido pelo aprendizado de máquina busca melhorar o desempenho da execução de uma tarefa à medida que o modelo escolhido para isso passe por um processo de treinamento e aprenda por meio da experiência, sendo capaz de atribuir rótulos, no caso de uma tarefa de detecção. Para medir o desempenho na execução da tarefa, é comum a utilização de métricas como a acurácia, assim como a atribuição de penalidade para a rotulação incorreta de determinados rótulos também pode ser aplicada (JORDAN; MITCHELL, 2015). Além da acurácia, precisão, sensibilidade e F1 score também podem ser utilizadas para a análise do modelo

A acurácia é a métrica de avaliação que mede a proporção de rótulos previstos corretamente como positivos ou negativos, considerando todas as instâncias previstas por um modelo (DALIANIS, 2018). Esta medida é considerada a mais usada em problemas de classificação, mas que pode não ser a mais adequada em casos onde se tem um conjunto de dados desbalanceado. Sua fórmula pode ser encontrada abaixo (VAKILI; GHAMSARI; REZAEI, 2020):

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

onde VP representa o número de previsões verdadeiramente positivas, enquanto VN representa as previsões verdadeiramente negativas. FP representa as previsões positivas incorretas e FN, as previsões negativas incorretas. A soma das previsões corretas é dividida pelo total de previsões, considerando os resultados corretos e incorretos.

A precisão é a métrica que calcula a razão entre o número de rótulos previstos corretamente como positivos e o número total de rótulos previstos como positivos, considerando também os falsos positivos. Este cálculo pode ser compreendido pela fórmula abaixo (VAKILI; GHAMSARI; REZAEI, 2020):

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (2)$$

onde VP considera o número de instâncias classificadas corretamente como positivas, enquanto FP considera o número de instâncias classificadas incorretamente como positivas (os falsos positivos). Por outro lado, a sensibilidade é a métrica que considera todos os exemplos positivos do conjunto e busca calcular quantas delas o modelo foi capaz de prever. Para isso, o número de instâncias previstas corretamente como positivas é dividido pelo número de instâncias positivas no conjunto, considerando os falsos negativos. Seu cálculo pode ser feito aplicando a fórmula a seguir (VAKILI; GHAMSARI; REZAEI, 2020):

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (3)$$

onde VP considera o número de instâncias classificadas corretamente como positivas, enquanto FN considera instâncias classificadas incorretamente como negativas.

Por fim, a métrica F1-score pode ser calculada através da equação abaixo, retornando média harmônica entre outras duas métricas: a precisão e a sensibilidade (DALIANIS, 2018).

$$F1\text{-score} = 2 \times \frac{\text{Precisão} \times \text{Sensibilidade}}{\text{Precisão} + \text{Sensibilidade}} \quad (4)$$

Estudos têm mostrado que técnicas de aprendizado de máquina, especialmente em contextos supervisionados, são amplamente empregadas para auxiliar na detecção precoce de doenças, inclusive de patologias cardiovasculares, ao aprender padrões discriminativos em dados de exames clínicos e imagens médicas.

Técnicas de aprendizado de máquina têm sido utilizadas na área da saúde, como evidenciado por diferentes estudos publicados. Esses estudos demonstram sua aplicação na identificação de cardiopatias e do risco de recorrência de câncer, utilizando modelos treinados a partir de dados médicos, como prontuários, exames laboratoriais e exames de imagem. As previsões desses modelos devem apresentar alta precisão, uma vez que, dependendo do problema a ser resolvido, a identificação incorreta de exemplos positivos ou negativos pode resultar em consequências de diferentes níveis de gravidade. Na previsão do risco de recorrência em pacientes com câncer, por exemplo, a atribuição incorreta de um rótulo negativo a um paciente com alto risco de recorrência é mais perigosa do que a atribuição incorreta de rótulos positivos (ARAF; IDRI; CHAIRI, 2024).

No aprendizado de máquina existem diferentes tipos de aprendizagem, sendo a aprendizagem supervisionada uma das principais. Na aprendizagem supervisionada, o objetivo é encontrar modelos que apresentem bom desempenho na previsão de dados

não vistos durante seu treinamento. Para os dados de treino, uma rotulação prévia dos exemplos é feita (BARBER, 2012).

Na seção seguinte o aprendizado supervisionado será detalhado, assim como alguns algoritmos serão apresentados, incluindo o Random Forest. Este é um algoritmo que utiliza árvores de decisão durante seu treinamento para retornar a classe do exemplo visualizado. Cada árvore é responsável pela escolha de um rótulo, e a classe final corresponde ao rótulo com mais atribuições (SINGH; THAKUR; SHARMA, 2016).

4.1 Aprendizado Supervisionado

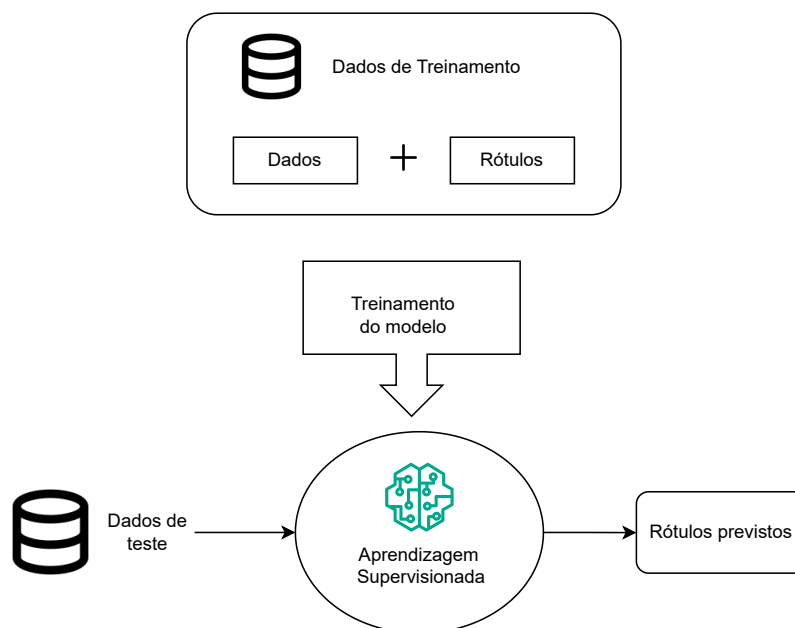
Entre as categorias de métodos mais conhecidas, o aprendizado supervisionado é o mais utilizado (JORDAN; MITCHELL, 2015). Nesta abordagem, os algoritmos são treinados com um conjunto de dados rotulados, onde os exemplos são compostos por pares de entrada e saída desejada. Após o treinamento, o modelo é testado e avaliado em um conjunto de testes também rotulado, mas que, embora tenha classes conhecidas, é utilizado como entrada para que o modelo demonstre a capacidade de prever seus rótulos com base no que foi aprendido durante o processo de treinamento. Um conjunto de dados a ser utilizado em uma tarefa contém amostras com base no problema proposto, podendo ser, por exemplo, registros de e-mails em um problema de detecção de spam. Nesta tarefa de detecção de spam, cada exemplo é representado por um conjunto de características que podem ser palavras e frases. Os algoritmos de aprendizado supervisionado utilizam tais características, também chamadas de preditores, para a previsão dos rótulos do conjunto de teste (FABRIS; MAGALHÃES; FREITAS, 2017).

O processo de rotulagem dos dados, como é possível compreender, é um passo essencial. Em problemas de domínios específicos, como saúde mental e física, construir um conjunto de exemplos rotulados se torna mais complexo. Profissionais especializados no problema proposto são necessários para garantir a precisão da rotulação, uma vez que seus conhecimentos permitem interpretar corretamente os dados.

A Figura 6 ilustra, de modelo geral, o processo da aprendizagem onde o próprio termo “supervisionado” expressa a ideia de que o conhecimento prévio sobre os resultados desejados guiam o método (MRABET; MAKKAOUJ; FAIZE, 2021). Os algoritmos desta categoria funcionam a partir da construção de uma função preditiva

$$F : X \rightarrow Y \quad (5)$$

A função realiza o mapeamento das características X para a variável alvo Y (rótulo) e, dessa forma, compreende-se que os modelos são treinados com dados em pares. Cada par contém um exemplo X e seu rótulo Y , formatado como uma tupla (X_i, Y_i) . O

Figura 6 – Abordagem da aprendizagem supervisionada

Fonte: Adaptado de Mrabet, Makkaoui e Faize (2021)

objetivo é que, após treinado, o algoritmo de aprendizado supervisionado seja capaz de prever valores de Y de acordo com as entradas X (FABRIS; MAGALHÃES; FREITAS, 2017; JORDAN; MITCHELL, 2015).

A aprendizagem supervisionada ainda pode ser dividida em duas categorias, sendo possível resolver problemas de regressão e de classificação. Modelos de regressão trabalham com dados contínuos e são aplicados em casos onde se deseja prever preços de casas, por exemplo. Em comparação a esta categoria, os modelos de classificação trabalham com dados binários ou multiclases, onde Y pode assumir um ou mais valores. Em um problema de classificação binária, o modelo pode prever as classes “aprovado” e “reprovado” em uma análise de perfil de crédito, e pode prever “esporte”, “cultura” e “política” como uma das categorias de um novo artigo em um problema de classificação multiclasse (GUPTA et al., 2022).

Existem diferentes algoritmos que podem ser utilizados em problemas resolvidos por meio da aprendizagem de máquina supervisionada. Máquina de Vetores de Suporte, Regressão Logística e Floresta Aleatória são alguns exemplos que podem ser citados. Os três algoritmos serão brevemente explicados nos tópicos a seguir, focando no objetivo central de cada um deles.

4.1.1 SVM

Máquina de Vetores de Suporte (*Support Vector Machine* — SVM) é um algoritmo de aprendizado de máquina para problemas de regressão e classificação.

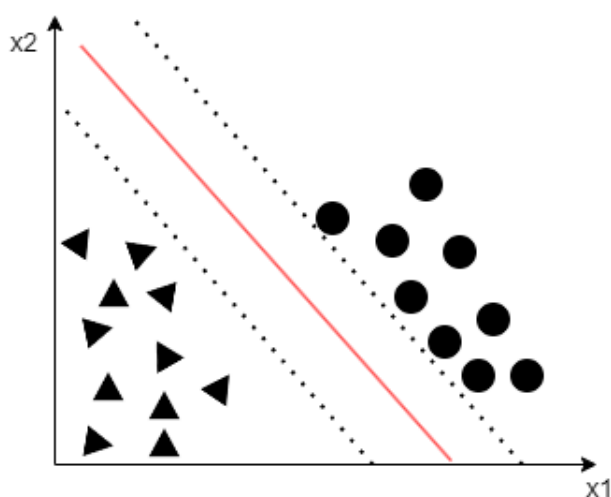
Em geral, o SVM busca encontrar um hiperplano, o qual representa uma superfície de separação entre classes diferentes. Para isso, os dados de entrada são transformados em um espaço onde as classes do conjunto, quando linearmente separáveis, podem ser divididas (RODRIGUEZ-GALIANO et al., 2015).

SVM, quando aplicado em dados linearmente separáveis, busca definir o hiperplano que melhor separa as classes de um conjunto. O objetivo é definir a superfície de separação que maximiza a margem do problema, ou seja, aquele que garante a maior distância entre as classes (CRISCI; GHATTAS; PERERA, 2012).

$$H^* = \frac{\operatorname{argmax} \gamma_H}{H} \quad (6)$$

H^* representa o hiperplano ótimo e único, encontrado por meio de uma otimização numérica. A função argmax representa a busca pelo valor que maximizará γ_H , a margem do hiperplano. A Figura 7 exemplifica um hiperplano no contexto da SVM para duas classes linearmente separáveis.

Figura 7 – Classes linearmente separáveis – SVM



Fonte: Adaptado de Crisci, Ghattas e Perera (2012)

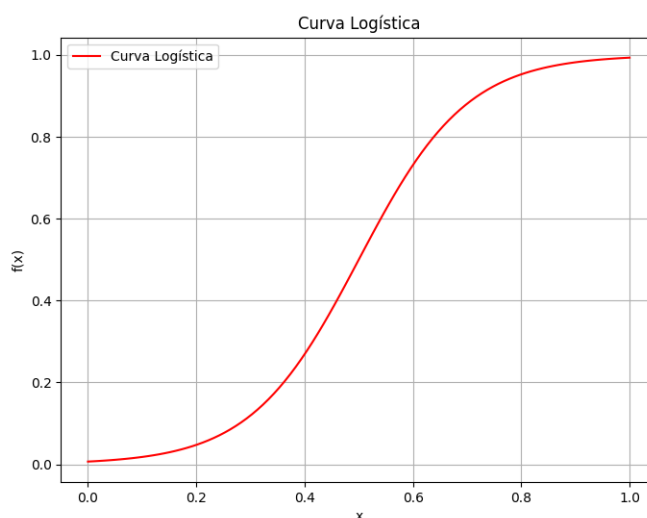
Por adotar a generalização, o SVM se ajusta aos padrões identificados nos dados fornecidos, reduzindo o risco de *overfitting*. O *overfitting* ocorre quando um modelo de aprendizado de máquina se ajusta excessivamente aos dados de treinamento, resultando em um desempenho ineficaz quando aplicado a dados não vistos anteriormente. No entanto, em conjuntos de dados com mais ruídos, o algoritmo não atinge sua melhor performance (RAY, 2019).

4.1.2 Logistic Regression

Regressão logística (*Logistic Regression*) é um algoritmo de aprendizado de máquina que utiliza conceitos estatísticos para determinar a classe de uma instância,

onde uma curva logística é ajustada aos dados e permite o cálculo da probabilidade de um exemplo pertencer a uma classe (LORENA et al., 2011). A curva logística é visualizada em forma de S e mapeia todos os números reais para números em um intervalo entre 0 e 1. Quando utilizado, o algoritmo busca prever o resultado da variável alvo Y por meio da variável X, que representam respectivamente a classe e o dado de entrada (CHOUDHARY; GIANEY, 2017). A Figura 8 exibe a exemplificação de uma curva logística.

Figura 8 – Representação de uma curva logística



Fonte: Elaborado pela autora

Ao realizar o mapeamento da probabilidade em valores entre 0 e 1, a regressão logística utiliza a função sigmoide. A fórmula matemática abaixo representa esta função, onde z é a soma ponderada das características dos dados de entrada (LIANG, 2023).

$$g(z) = \frac{1}{1 + e^{-z}} \quad (7)$$

Para transformar a probabilidade em um rótulo da classe, representando a decisão do modelo, é necessário definir um limiar. Este limiar, que na maioria das vezes é 0,5 indica que o exemplo é da classe positiva se a probabilidade for superior a ele. No entanto, em probabilidades menores do que 0,5 a classe negativa é escolhida (LIANG, 2023). Com o objetivo de reduzir o número de falsos positivos, como em aplicações na área da saúde, o limiar para a decisão das classes é maior do que 0,5. Em contrapartida, o aumento do valor do limiar pode elevar o número de falsos negativos.

4.1.3 Random Forest

Floresta Aleatória é a tradução literal do algoritmo *Random Forest*, que prevê a classe de um dado não rotulado por meio da utilização de árvores preditoras, onde

cada árvore indica a classe provável do exemplo e o valor final de Y corresponderá a categoria mais escolhida. (LORENA et al., 2011).

O processo de treinamento do algoritmo se dá pelo treinamento individual das árvores construídas. Para isso, os dados são selecionados aleatoriamente em um processo de amostragem e, posteriormente, cada árvore de decisão é treinada para que seja capaz de prever classes e produzir resultados independentes (DAI et al., 2018).

Uma árvore de decisão é uma estrutura de árvore que pode ser binária ou não, onde cada nó não folha representa um teste realizado em cada característica específica. Cada resultado possível é representado por uma ramificação, levando a diferentes partes da estrutura até a chegada ao nó folha, o qual armazena a classe final do exemplo (ZHU et al., 2019).

Random Forest é um algoritmo de aprendizado de máquina tradicional capaz de lidar com a ausência de valores em dados de entrada, o que é considerado uma grande vantagem (BOER et al., 2023). Ao estimar dados ausentes, o modelo permite que a precisão da previsão se mantenha (ZHU et al., 2019).

4.2 Aprendizado Semi-supervisionado

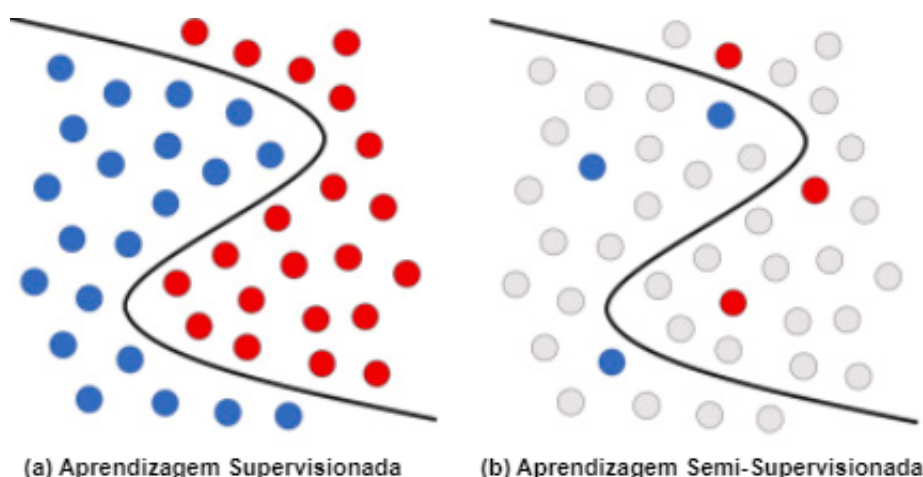
Em problemas resolvidos com o aprendizado de máquina, é possível que um dos desafios encontrados seja a disponibilidade de dados rotulados, sendo uma quantidade limitada em comparação com a grande quantidade de dados não rotulados que podem ser obtidos. Nesses cenários, o aprendizado semi-supervisionado surge como uma solução para lidar com a carência de dados rotulados, uma vez que busca aproveitar dados rotulados e não rotulados para a criação de modelos (BARBER, 2012). Por meio de dados rotulados, utilizados como uma espécie de referência, os modelos buscam identificar padrões e características que possam ser relevantes para a classificação de dados não rotulados (N; GUPTA, 2020). A Figura 9 compara visualmente o aprendizado supervisionado e semi-supervisionado.

Como mostrado na figura, o aprendizado supervisionado utiliza apenas dados rotulados para a construção de modelos. Os círculos vermelhos e azuis representam duas classes que estão sendo segmentadas. Na categoria do aprendizado semi-supervisionado, as duas classes são mantidas, mas a construção do modelo se dá com poucos dados rotulados.

Ao utilizar informações dos dados não rotulados, os algoritmos semi-supervisionados podem melhorar suas capacidades de classificação (BARBER, 2012). Esta categoria de aprendizado ainda pode ser dividida em duas configurações principais com base na disponibilidade dos dados de teste durante o processo de treinamento: transdutiva e indutiva (JIAO et al., 2024).

A aprendizagem transdutiva é uma configuração que assume que as amostras não rotuladas durante o treinamento são o conjunto de teste, os dados que devem

Figura 9 – Aprendizado supervisionado X aprendizado semi-supervisionado



Fonte: Adaptado de Jiao et al. (2024)

ter seus rótulos previstos. Nesta configuração, o objetivo é generalizar o modelo sobre essas amostras não rotuladas. Por outro lado, a aprendizagem indutiva aplica o modelo semi-supervisionado em novos dados não vistos anteriormente no processo de treinamento (JIAO et al., 2024).

O aprendizado semi-supervisionado é um aliado no desafio do treinamento de modelos com conjuntos de dados que apresentam rótulos ausentes e existem diferentes abordagens que podem ser empregadas. Modelos generativos e métodos baseados em grafos estão entre as abordagens que serão apresentadas a seguir.

4.2.1 TSVM

Máquina de Vetores de Suporte Transdutivo (*Transductive Support Vector Machine* — TSVM) é uma extensão da Máquina de Vetores de Suporte e utiliza dados não rotulados no processo de treinamento. Em uma Máquina de Vetores de Suporte padrão, SVM, apenas dados rotulados são utilizados para o treinamento e busca pela maximização da margem entre as classes do problema. O TSVM tem como objetivo maximizar a margem entre os dados rotulados e não rotulados em seu limite de decisão linear, definindo as categorias dos dados não rotulados (N; GUPTA, 2020; PISE; KULKARNI, 2008).

TSVM utiliza o método de pesquisa local para encontrar a solução ideal para o problema, sendo isso feito de forma iterativa. Um modelo SVM padrão é treinado com amostras rotuladas para que, posteriormente, categorias (pseudo-rótulos) sejam atribuídas às amostras não rotuladas. Assim que o hiperplano é calculado com a maximização da margem das amostras rotuladas e não rotuladas, os pseudo-rótulos atribuídos anteriormente são ajustados de acordo com este hiperplano. O processo é repetido até que um critério de parada seja alcançado (CHEN et al., 2023).

Diferentes problemas podem ser resolvidos utilizando Máquina de Vetores de Suporte Transdutivo. Singla, Ghosh e Shukla (2021) propuseram prever os pacientes infectados por COVID-19 por meio de radiografias de tórax, aplicando TSVM junto a uma função de perda para tornar o modelo mais resistente aos ruídos. pin-TSVM, como nomeado, foi o modelo desenvolvido para ser treinado em conjuntos com número significativo de dados não rotulados, conjuntos ruidosos e de diferentes tamanhos de amostras (SINGLA; GHOSH; SHUKLA, 2021).

O treinamento do TSVM apresenta uma desvantagem devido à velocidade lenta e a quantidade de iterações. Para lidar com esta desvantagem, o algoritmo de recozimento é frequentemente utilizado por ser uma heurística de otimização global (CHEN et al., 2023). Esta abordagem semi supervisionada, ao maximizar a margem com dados não rotulados, enfatiza mais a suposição de *cluster* do que a regularização do classificador, o que é apontado como um problema. Além disso, o problema de otimização deste algoritmo não é convexo com o modelo SVM, tornando difícil atingir uma solução ideal (ADANKON; CHERIET, 2007).

4.2.2 Auto-Treinamento

O auto-treinamento é uma abordagem semi supervisionada utilizada em cenários com poucos dados disponíveis, onde um classificador é inicialmente treinado com uma pequena quantidade de dados rotulados para que possa ser usado na classificação de dados não rotulados. Os exemplos não rotulados que demonstram classificação mais confiante são adicionados ao conjunto de dados de treinamento, juntamente com seus rótulos previstos pelo modelo. Após a inclusão, o classificador é novamente treinado com o novo conjunto de dados (N; GUPTA, 2020; PISE; KULKARNI, 2008).

Este ciclo de treinamento, previsão e atualização do conjunto de dados se repete até que o modelo atinja um critério de parada ou a sua convergência. Uma vez que o classificador utiliza suas próprias previsões para aprimorar seu desempenho, ensinando a si mesmo, o processo também é conhecido como *bootstrapping* ou auto-aprendizagem, sendo uma abordagem vista em várias tarefas de processamento de linguagem natural (N; GUPTA, 2020).

Um exemplo de aplicação de auto-treinamento em tarefas de processamento de linguagem natural é a classificação de sentimentos. Em um caso onde se deseja identificar conteúdos preconceituosos em postagens de redes sociais, rotular manualmente uma grande quantidade de dados pode ser inviável. Por meio do auto-treinamento, é possível utilizar um conjunto inicial de dados rotulados para treinar um modelo que posteriormente poderá ser aplicado na detecção de conteúdos não rotulados. Esta mesma prática pode ser vista em classificações de domínios mais específicos, como a identificação de sintomas do Transtorno de Ansiedade Generalizado em texto disponível online, uma vez que a rotulação destes exemplos exige a participação de profissionais

especializados.

É importante ressaltar que nesta abordagem semi supervisionada, métricas de seleção são necessárias para a escolha dos exemplos que foram previstos corretamente. Caso dados previstos incorretamente sejam incluídos no conjunto de treinamento, o desempenho final do modelo poderá ser comprometido (ROSENBERG; HEBERT; SCHNEIDERMAN, 2005).

4.2.3 Co-Treinamento

O co-treinamento foi proposto inicialmente por Blum e Mitchell (1998) e envolve o treinamento de dois modelos distintos em diferentes visões de características dos mesmos dados rotulados. Algumas suposições fundamentais podem ser feitas em relação ao co-treinamento (PISE; KULKARNI, 2008):

1. Os recursos podem ser divididos em dois conjuntos distintos, cada um contendo as características dos dados.
2. Cada conjunto de sub-características é capaz de treinar um classificador de forma eficaz e individual.
3. Os dois conjuntos de recursos são considerados independentes devido à classe dos dados, significando que as características presentes em um conjunto não estão relacionadas com as características presentes em outro.

Esta abordagem de aprendizado semi-supervisionado, que envolve dois classificadores distintos, necessita de dois conjuntos diferentes de recursos nos dados rotulados utilizados. Cada um dos classificadores é treinado com um conjunto específico de recursos para poder ser capaz de classificar dados não rotulados disponíveis. Cada classificador realiza a previsão da classe dos dados não rotulados e, aquelas que são consideradas as mais confiantes, são usadas de forma iterativa pelo outro classificador como se fossem dados de treinamento rotulados (SHEIKHPOUR et al., 2017).

No co-treinamento, os dois classificadores precisam concordar tanto na previsão dos dados não rotulados, que são mais numerosos, quanto na predição dos dados rotulados. Isso é necessário, uma vez que dois modelos estão sendo construídos a partir de diferentes visualizações dos dados e a concordância indica uma maior confiabilidade na classificação (PISE; KULKARNI, 2008).

4.2.4 Modelos generativos

Modelos generativos são considerados modelos probabilísticos baseados na regra de Bayes, que utilizam a probabilidade para calcular a similaridade entre dados onde se conhece a classe e dados onde se deseja prever o valor da variável alvo Y (TSABITA et al., 2018).

Sendo assim, os modelos generativos assumem a fórmula $p(x, y) = p(y)p(x|y)$, onde $p(x|y)$ é uma distribuição de mistura identificável, como a mistura Gaussiana. Para determinar a distribuição da mistura, mesmo em grandes conjuntos de dados não rotulados, é necessário que apenas um dado por componente esteja rotulado. Este método é visto como um dos métodos de aprendizado semi-supervisionado mais antigos, onde os exemplos não rotulados disponíveis são classificados utilizando componentes da mistura associados às classes do problema (ZHU, 2005; CHEBLI; DJEBBAR; MAROUANI, 2018).

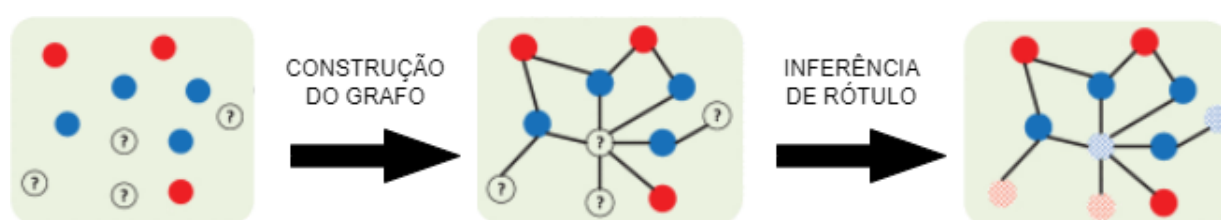
Considerando uma mistura Gaussiana, o algoritmo de Maximização de Expectativa (*Expectation-Maximization (EM)*) pode ser utilizado para encontrar os parâmetros correspondentes as classes. Sobre os modelos generativos, também é possível ressaltar que a abordagem de transformação de dados em uma representação de características determinadas pelo modelo generativo é uma aplicação viável. Estas representações podem então serem inseridas como entrada para um classificador discriminativo para categorização de imagens, por exemplo (ZHU, 2005; CHAPPELLE; SCHÖLKOPF; ZIEN, 2006).

4.2.5 Métodos baseados em grafos

A estrutura de um grafo é composta por conjunto de vértices V e arestas E , podendo ser definida como $G = (V, E)$. Um grafo é criado pelos seus nós e conexões entre eles, que são feitas através de arestas. Os nós podem representar padrões de dados e as arestas o indicativo da relação entre eles (N; GUPTA, 2020).

Métodos baseados em grafos trabalham com a construção de um grafo onde cada nó representa uma amostra no conjunto de dados, que podem ou não ser rotulada, e as arestas ponderadas geralmente representam a similaridade entre os exemplos. Nesta estrutura, entende-se que quando nós conectados por arestas com pesos significativos possuem maior probabilidade de pertencerem a mesma classe. De modo geral, o processo de construção destes modelos contém a etapa de construção do grafo de similaridade junto a ponderação das arestas, como citado, e a inferência de rótulos (SONG et al., 2023). A Figura 10 exemplifica estes dois passos.

Figura 10 – Procedimento de duas etapas do GSSL



Fonte: Adaptado de Song et al. (2023)

Conforme demonstra a figura, o processo se inicia com a construção do

grafo com base no conjunto de dados disponíveis, incluindo amostras rotuladas e não rotuladas. Cada cor representa uma classe de um dado rotulado e as amostras inseridas com o símbolo “?” indicam que são dados não rotulados. O principal desafio é fazer com que a relação entre as amostras originais sejam representadas de maneira correta dentro do grafo. Na segunda etapa, que consiste na inferência de rótulo, o objetivo é realizar a propagação das informações dos nós rotulados para os não rotulados. Isso é feito por meio das informações topológicas do grafo de similaridade construído (SONG et al., 2023). Como um exemplo da categoria de métodos baseado em grafo, o algoritmo de Competição e Cooperação entre Partículas terá seu funcionamento detalhado.

4.2.5.1 Competição e Cooperação entre Partículas

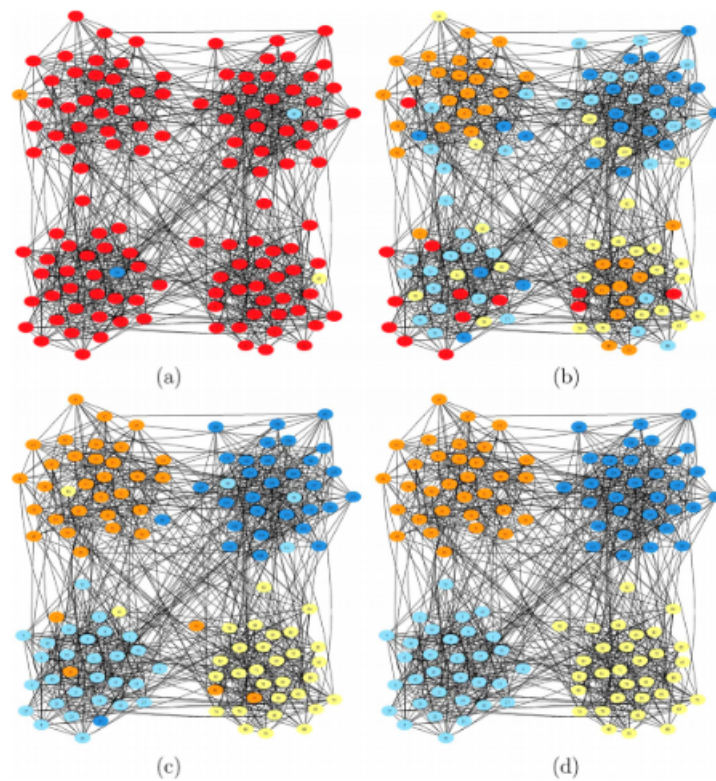
Da categoria de algoritmos de aprendizado semi-supervisionado, Competição e Cooperação entre Partículas é um modelo de classificação baseado em grafos e derivado de outro modelo, o de competição de partículas. O modelo de competição de partículas foi proposto para a detecção de comunidades em redes complexas, onde partículas caminham na rede de modo a competirem entre si e buscarem conter o menor número de nós, marcando seus territórios e rejeitando partículas consideradas intrusas. A Figura 11 demonstra visualmente o processo da detecção de comunidades utilizando o algoritmo (QUILES et al., 2008).

A Figura 11.a ilustra as configurações iniciais do modelo com 4 partículas dispostas aleatoriamente na rede, onde os nós vermelhos são representações não rotuladas. A Figura 11.b e 11.c representam, respectivamente, 250 e 3500 iterações na rede complexa e seu estado após isso. Por fim, a Figura 11.d representa a rede após 7000 iterações.

Tendo este conhecimento em vista, o modelo de competição e cooperação entre partículas pode ser melhor explicado. Diferente do algoritmo explicado anteriormente, o modelo inclui cooperação e competição de forma combinada e unificada. O primeiro passo para a implementação deste modelo se dá pela construção de uma rede não direcionada e não ponderada, a qual representa um conjunto de dados de entrada. Isso é feito com a transformação de cada dado do conjunto em um nó no grafo, sendo eles conectados posteriormente por meio de arestas (BREVE et al., 2012).

Os nós, representando dados rotulados ou não rotulados, podem ser conectados de diferentes maneiras, como a ligação entre nós com distância menor do que um valor pré-estabelecido ou pelo método do k -vizinhos mais próximos. A primeira abordagem citada utiliza o cálculo da distância Euclidiana, por exemplo, e determina que existirá uma aresta ligando dois nós v_i e v_j apenas se este par possuir uma distância menor do que a estabelecida por um limiar σ . Na segunda abordagem é determinado que cada nó estará conectado aos seus k -vizinhos mais próximos, de modo que x_j é conectado com x_i se pertencer aos seus k -vizinhos mais próximos. Esta condicional é

Figura 11 – Detecção de comunidades através do algoritmo de competição de partículas



Fonte: Quiles et al. (2008)

representada abaixo (BREVE et al., 2012).

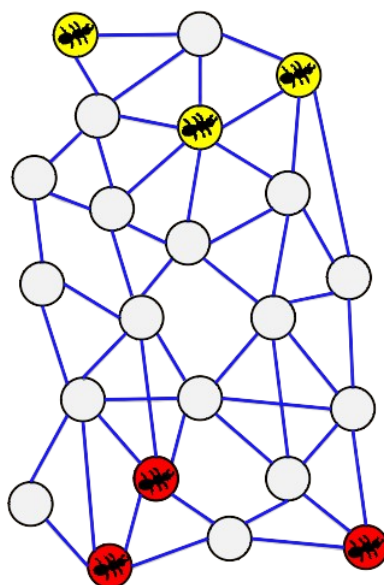
$$W_{ij} = \begin{cases} 1 & \text{se } x_j \text{ esta entre os } k\text{-vizinhos mais próximos } x_i \text{ ou vice-versa} \\ 0 & \text{caso contrário} \end{cases}$$

Após a construção do grafo e da ligação entre o nós, partículas são geradas para cada nó rotulado. Isso quer dizer que, para cada dado rotulado e representado no grafo, uma partícula p_i é gerada e sua posição inicial definida como v_i , seu rótulo na rede e também chamado de nó-casa. A Figura 12 ilustra a criação dessas partículas (BREVE, 2010).

A Figura 12 exibe nós pintados de vermelho e amarelo, os quais representam os nós rotulados no grafo. Por serem rotulados, a ilustração de uma formiga indica que partículas foram geradas para cada um deles. Já os nós cinzas representam dados não rotulados, significando que nenhuma partícula foi gerada para eles.

Cada nó representado no grafo possui um vetor de domínio, estrutura que é utilizada durante a dinâmica de nós e que indica a dominância de cada classe em relação a eles. Em um nó rotulado como classe A e 4 classes possíveis, por exemplo, o vetor de domínio seria configurado de forma que a classe A dominasse completamente

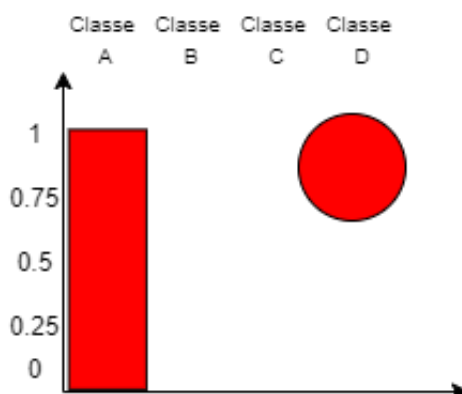
Figura 12 – Geração de partículas para nós rotulados no grafo



Fonte: Breve (2024)

tal nó. A Figura 13 é um exemplo de configuração inicial de níveis de domínio para um nó rotulado (BREVE, 2010).

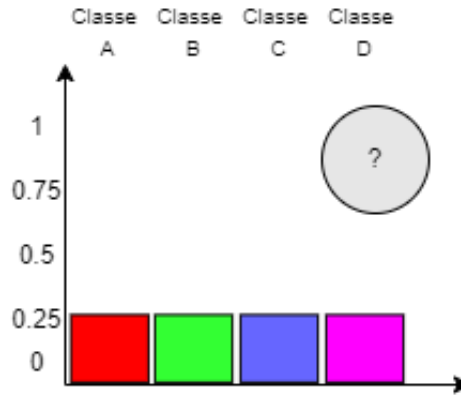
Figura 13 – Níveis de domínio para nó rotulado



Fonte: Elaborado pela autora

A Figura 13 ilustra um nó rotulado como classe A. Seu vetor de domínio seria configurado como $[1, 0, 0, 0]$, sendo a classe A 100% dominante em relação as outras 3 possíveis. Para nós não rotulados, a distribuição é diferente, conforme mostra a Figura 14.

O nó de cor cinza representa uma amostra não rotulada. Uma vez que sua classe é desconhecida, a probabilidade é dividida entre todas as possíveis, resultando em 25% de dominância para cada uma, ou seja, seu vetor é configurado como $[0,25, 0,25, 0,25, 0,25]$. Nesta dinâmica, cada partícula seleciona um nó para visitar a cada iteração, buscando aumentar o nível de domínio de sua classe em relação a este alvo.

Figura 14 – Níveis de domínio para nó não rotulado

Fonte: Elaborado pela autora

Em contraste a isso, o nível de domínio das outras classes diminuem para o mesmo nó, que deve ser não-rotulado, uma vez que nós rotulados possuem níveis de domínio fixos. A atualização dos vetores de domínio pode ser representada pela formulação abaixo, com um conjunto de rótulos L onde $y_i \in L$ representa uma amostra rotulada e $y_i = \emptyset$ uma amostra não rotulada (BREVE et al., 2012):

$$v_i^{\omega_\ell}(t+1) = \begin{cases} \max \left\{ 0, v_i^{\omega_\ell}(t) - \frac{\Delta v \rho_j^\omega(t)}{c-1} \right\} & \text{se } y_i = \emptyset \text{ e } \ell \neq \rho_j^f \\ v_i^{\omega_\ell}(t) + \sum_{q \neq \ell} v_i^{\omega_q}(t) - v_i^{\omega_q}(t+1) & \text{se } y_i = \emptyset \text{ e } \ell = \rho_j^f \\ v_i^{\omega_\ell}(t) & \text{se } y_i \in L \end{cases} \quad (8)$$

A taxa de alteração dos níveis de domínio é controlada por um parâmetro Δv , valor que deve ser maior que zero e menor ou igual a 1. Caso este parâmetro seja um valor alto, os níveis de domínio mudam de forma mais rápida, enquanto o contrário ocorre para um valor baixo. Após a atualização do vetor de domínio, a força da partícula também é atualizada. Isso é feito de forma que, caso o nó alvo seja dominado por sua classe, sua força é aumentada. No entanto, sua força é reduzida quando o nó alvo é dominado por uma classe oposta, conforme demonstra a equação abaixo (BREVE et al., 2012):

$$\rho_j^\omega(t+1) = v_i^{\omega_\ell}(t+1) \quad (9)$$

A cada iteração da partícula ρ_j , sua força ρ_j^ω é atualizada de acordo com o valor do nível de domínio de sua classe $v_i^{\omega_\ell}$ em relação a um nó alvo v_i . Com isso, cada vez que o nó alvo v_i é dominado por classes opostas, a força da partícula ρ_j^ω diminui. A Figura 15 representa esta atualização (BREVE et al., 2012).

A Figura 15.a representa a força de uma partícula onde seu nó alvo é dominado por sua própria classe, fazendo com que essa força aumente. No entanto, a Figura 15.b representa a redução da força de uma partícula quando seu nó alvo é dominado por outra classe. Para que este segundo cenário seja contornado, o modelo

Figura 15 – Atualização da força de uma partícula



Fonte: Breve et al. (2012)

propõe o uso da informação da distância entre a partícula e o seu nó-casa a cada iteração.

Ao saber o quão longe está de seu nó-casa, a partícula evita a redução de sua força ao não ter como alvo os nós dominados por outras classes e, consequentemente, essa abordagem também permite que partículas vizinhas sejam protegidas. Antes de se mover do nó atual para o nó alvo, a partícula obtém a informação de uma tabela de distâncias. Quando o nó alvo apresenta distância maior do que o nó atual, este valor é atualizado. Isso é feito através da soma da distância do nó atual e a constante 1, de modo que a partícula permanecerá sem se mover até a próxima iteração. Este cálculo pode ser visualizado na equação abaixo (BREVE et al., 2012):

$$\rho_j^{d_k}(t+1) = \begin{cases} \rho_j^{d_i}(t) + 1 & \text{se } \rho_j^{d_i}(t) + 1 < \rho_j^{d_k}(t) \\ \rho_j^{d_k}(t) & \text{caso contrário} \end{cases} \quad (10)$$

onde $\rho_j^{d_i}(t)$ representa a distância do nó-casa a partir do nó atual e $\rho_j^{d_k}(t)$ representa a distância a partir do nó alvo. Este funcionamento consiste no mecanismo de competição do algoritmo, fazendo com que a partícula visite apenas nós dominados por sua classe (BREVE et al., 2012).

A movimentação das partículas pelo grafo pode ocorrer de duas formas: escolhendo o próximo nó de forma aleatória ou gulosa. A movimentação aleatória corresponde de modo que a escolha do vizinho a ser visitado não é influenciada pelo nível de domínio ou distância do nó-casa. Na movimentação gulosa, no entanto, a partícula tende a escolher um nó dominado pela sua própria classe e com uma distância menor em relação ao seu nó-casa. As duas formas de movimentação ocorrem para que a partícula exiba o comportamento de explorar e defender de forma equilibrada, uma vez que a movimentação gulosa é importante para a defesa do território de uma classe e a aleatória é útil para a exploração de novos nós no grafo (BREVE et al., 2012).

Durante a movimentação aleatória, a partícula ρ_j escolhe um nó v_i para

visitar com probabilidade definida pela fórmula:

$$p(v_i|\rho_j) = \frac{W_{qi}}{\sum_{\mu=1}^n W_{q\mu}} \quad (11)$$

Quando existe uma aresta entre o nó atual da partícula ρ_j , sendo q o índice deste nó, e qualquer outro nó v_i , é atribuída a constante 1 para W_{qi} . Caso contrário, W_{qi} recebe 0.

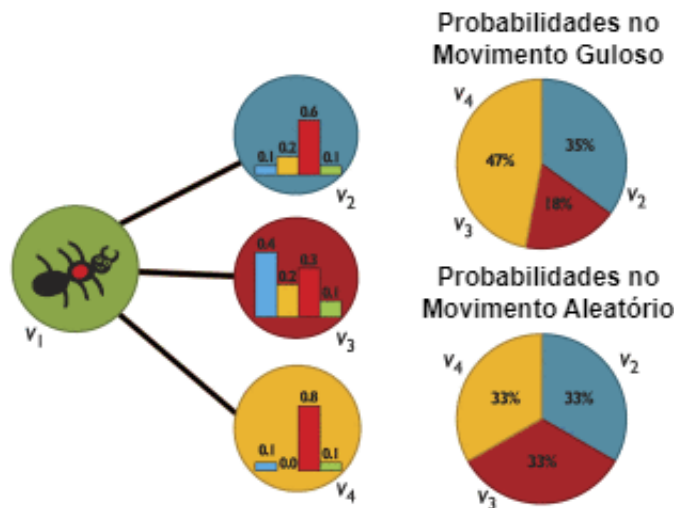
Já na movimentação gulosa, a escolha do nó se dá pela probabilidade definida pelo nível de domínio de sua classe, representado por $\rho_j^{w_l}$ na fórmula abaixo, e pelo inverso da distância ($\rho_j^{d_i}$) entre o nó v_i e o nó casa v_j da partícula

$$p(v_i|\rho_j) = \frac{W_{qi} v_i^{w_l} (1 + \rho_j^{d_i})^{-2}}{\sum_{\mu=1}^n W_{q\mu} v_\mu^{w_l} (1 + \rho_j^{d_\mu})^{-2}} \quad (12)$$

onde, novamente, q é o índice do nó atual da partícula ρ_j e $\ell = \rho_j^f$, onde ρ_j^f é o rótulo da classe da partícula ρ_j .

A cada iteração, cada partícula tem uma probabilidade p_{grd} de realizar a movimentação gulosa e probabilidade $1 - p_{grd}$ de realizar o movimento aleatório ($0 \leq p_{grd} \leq 1$). A Figura 16 ilustra a probabilidade de movimentação para estas duas formas.

Figura 16 – Probabilidades da movimentação de partículas de forma aleatória e gulosa



Fonte: Adaptado de Breve et al. (2012)

Além de uma dinâmica de movimentação, as partículas também possuem um critério de parada. É possível monitorar a rede de modo que, quando o algoritmo não apresentar aumento na média dos maiores níveis de domínio dos nós, as partículas deixem de caminhar pelo grafo. O monitoramento da média desses níveis de domínio é um critério considerado eficiente para o algoritmo (BREVE et al., 2012).

5 TRABALHOS RELACIONADOS

A busca pelo auxílio de um diagnóstico precoce de transtornos mentais específicos vem sendo cada dia mais abordada em pesquisas que buscam a interdisciplinariedade, unindo conceitos da psicologia e as grandes subáreas da Inteligência Artificial. Processamento de Linguagem Natural, Aprendizado de Máquina e Aprendizado Profundo são conceitos e áreas que sustentam as metodologias de cada artigo em busca de resultados nas aplicações de classificação textual.

5.1 Modelo de linguagem extenso pré-treinado aprimorado com conhecimento para diagnóstico e tratamento da depressão

Wang, Liu e Wang (2023) propõe uma abordagem incrementada de pré-treinamento em um conjunto de dados em chinês, utilizando um grafo de conhecimento sobre depressão para um conhecimento aprofundado do domínio e os modelos LLaMA-7B, ChatGLM-6B e Alpaca. O estudo buscou enfrentar os desafios do problema da alucinação, que ocorre quando modelos de linguagem geram resultados enganosos, e o da não compreensão abrangente sobre a depressão, justificado pela ausência de conjuntos de dados específicos.

Para o conjunto de dados de pré-treinamento, os autores utilizam treinamento não supervisionado, junto a dados gerais e específicos da depressão. Para mitigar o problema da escassez de dados da depressão, técnicas de aumento de dados foram empregadas através do método de Chataug, parafraseando sentenças e criando exemplos similares com variação semântica (WANG; LIU; WANG, 2023).

A pesquisa ajustou os modelos de linguagem com a utilização de instruções geradas por meio do grafo de conhecimento e *prompts* adaptados para melhorar a obtenção do conhecimento específico. Uma vez que a classificação da depressão requer cuidado para não ocasionar efeito prejudicial, os modelos foram desenvolvidos seguindo o aprendizado por reforço com *feedback* de especialistas, que se deu com base na Escala de Avaliação de Depressão de Hamilton (HAMD) no contexto de diálogos. A Tabela 6 demonstra os resultados obtidos pelos autores.

Tabela 6 – Formulário de avaliação dos modelos aplicados em um diálogo

Modelo	Segurança	Usabilidade	Fluência
LLaMA-7B	7,31	3,97	9,78
ChatGLM-6B	7,03	3,34	9,89
Alpaca	7,44	4,31	9,85
LLMs + Conhecimento	9,58	7,74	9,75

Fonte: Adaptado de Wang, Liu e Wang (2023)

Em uma avaliação de escala de 1 a 10, os autores compreenderam que a fluência do diálogo em modelos de grande escala não é impactada pela inclusão do

conhecimento. No quesito segurança, modelos sem aprimoramento do conhecimento específico de saúde mental apresentam mais riscos e menor usabilidade. Os modelos integrados ao grafo de conhecimento apresentaram diálogos mais direcionados para o tratamento do transtorno mental, com maior segurança em sua abordagem e, consequentemente, melhor usabilidade.

Embora tenha o objetivo de desenvolver modelos para diagnóstico e tratamento da depressão, a pesquisa ressalta a importância dos profissionais especializados na área da saúde mental para a realização desse papel, pontuando que os modelos de linguagem podem servir como auxiliares e não substitutos destes profissionais.

5.2 Rede LSTM para detecção de suicídio

A depressão é um transtorno que pode levar ao suicídio e, sabendo disso, Gupta e Pirzada (2023) buscaram desenvolver um sistema para identificar depressão utilizando uma rede neural LSTM e um conjunto de dados composto por postagens da rede social Reddit, mais especificamente, de exemplos extraídos dos *subreddits* “SuicideWatch” e “teenagers”. 232,074 mil linhas foram obtidas, sendo 116,037 mil exemplos de cada classe.

O pré-processamento dos dados é uma tarefa crucial em aplicações de processamento de linguagem natural. A pesquisa realizou a remoção de caracteres especiais, *stopwords*, tokenização e representação de palavras com *GloVe Embedding*. Vetores globais para representação de palavras geram *embedding* de palavras utilizando análise estatística para descobrir a relação entre as palavras.

Os autores realizaram os experimentos com a divisão do conjunto de dados com 20% para teste e 80% para treino em um cenário com classes balanceadas. A rede LSTM foi configurada com 30 épocas e lote de tamanho 256. A Tabela 7 apresenta a comparação realizada pelos próprios autores entre os resultados obtidos em seu estudo e os apresentados no trabalho de Nikhileswar et al. (2021).

Tabela 7 – Resultados e comparação entre trabalhos relacionados

Referência	Metodologia	Acurácia (%)
Nikhileswar et al.	Word2Vec + LSTM	92,4
Nikhileswar et al.	Word2Vec + LSTM-CNN	92,6
Nikhileswar et al.	Word2Vec + rede neural totalmente conectada	90,77
Gupta e Pirzada	<i>GloVe Embedding</i> + LSTM	93

Fonte: Adaptado de Gupta e Pirzada (2023)

O trabalho de Gupta e Pirzada alcançou acurácia de 93%, demonstrando

que, em comparação com pesquisas relacionadas elencadas pelos autores, a aplicação da rede LSTM com *GloVe Embedding* apresentou melhor desempenho.

A conclusão dos autores é que o *GloVe Embedding* é responsável pelo desempenho superior do método proposto. No entanto, a pesquisa ressalta que a falta de dados é um ponto negativo, além da ausência de rotulagem manual do conjunto utilizado.

Ao extrair os dados, o estudo pressupõe que as postagens do *subreddits* “SuicideWatch” contêm conteúdos suicidas, enquanto as do *subreddits* “teenagers” representam conteúdos não suicidas. Essa abordagem, que se baseia apenas no título do *subreddit* para definir a classe de um exemplo, pode enviesar o conjunto de dados e aumentar o número de falsos positivos para este domínio de pesquisa.

5.3 Detecção de ideação suicida nas redes sociais: uma abordagem de aprendizado de máquina

Chadha, Gupta e Kumar (2022) trabalham para diferenciar, usando aprendizado de máquina, postagens suicidas e não suicidas coletadas de redes sociais. A pesquisa executou diferentes algoritmos como SVM e Regressão Logística em um conjunto de dados com postagens extraídas do Reddit, sendo 10.000 coletadas do *subreddit* “SuicideWatch”.

Neste trabalho, palavras-chave como “ansiedade” e “diversão” foram utilizadas para retornar postagens que indicassem o estado emocional do usuário. Os dados coletados passaram pela etapa de pré-processamento, com a aplicação de técnicas como tokenização e TF-IDF, e foram divididos em duas classes: “suicida” e “não suicida”.

Os dados pré-processados foram utilizados para treinamento supervisionado dos modelos SVM, Regressão Logística e *AdaBoost*. Quando a acurácia é analisada, o modelo de Regressão Logística atinge 80,75%, considerado o melhor resultado. Porém, quando a precisão é analisada, o modelo Máquina de Vetores de Suporte apresenta o maior valor, 80,72%. Os autores ainda identificam a possibilidade de aplicar aprendizado profundo em dados extraídos de redes sociais para prever não apenas a ideação suicida, mas estresse e ansiedade também.

É importante observar que o processo de rotulação manual dos dados posteriormente testados neste trabalho, não especifica se critérios científicos foram utilizados ou se profissionais especializados em saúde mental auxiliaram na atribuição das classes. Este é um ponto a ser considerado nesta pesquisa, assim como em outros trabalhos que buscam aplicar a inteligência artificial na detecção de transtornos mentais, uma vez que emoções analisadas de forma isolada podem levar a classificação manual incorreta dos exemplos.

5.4 Detecção de ideação suicida no Twitter usando RNA

Yatapala e Kumara (2021) pesquisaram a possibilidade de detectar ideações suicidas por meio de postagens do Twitter, utilizando uma rede neural artificial.

O conjunto de dados escolhido pelos autores teve suas características extraídas através da técnica de *Word Embedding*, utilizando Word2Vec e Doc2Vec, e por meio da aplicação de *Term Frequency-Inverse Document Frequency* (TF-IDF), que leva em consideração a frequência e importância das palavras nos documentos.

Os autores optaram pelo aprendizado supervisionado para o sistema de classificação com o conjunto de dados dividido em 80% para treinamento e 20% para teste. A Tabela 8 lista todas as configurações da rede neural artificial proposta.

Tabela 8 – Parâmetros de configuração da RNA

Parâmetros	Valores
Tamanho do lote	10
Épocas	20
Perda	binary_crossentropy
Otimizador	adam
Função de ativação - camada de saída	sigmoid
Função de ativação - camadas ocultas	relu

Fonte: Adaptado de Yatapala e Kumara (2021)

A Rede Neural Artificial (RNA) foi construída com três camadas: uma camada de entrada com os vetores de características obtidos de Word2Vec, Doc2Vec e TF-IDF; uma camada oculta e uma camada de saída com função de ativação sigmoid. Para analisar seu desempenho, a pesquisa ainda aplicou, ainda em aprendizagem supervisionada, uma nova abordagem para a classificação.

O modelo Máquinas de Vetores de Suporte (SVM) foi utilizado e teve sua acurácia comparada com a primeira abordagem. A Tabela 9 demonstra os resultados obtidos pela RNA e pela SVM.

Tabela 9 – Valores de Acurácia da RNA e SVM

Vetor de recurso	RNA	SVM
Word2Vec	0,8568	0,7958
Doc2vec	0,8119	0,7399
TF-IDF	0,8676	0,8494

Fonte: Adaptado de Yatapala e Kumara (2021)

A rede neural artificial proposta teve melhor desempenho em comparação a aplicação da SVM, chegando ao valor máximo de 0,8676 com o uso de vetores construídos através do TF-IDF.

Os autores concluem ser possível reconhecer a ideação suicida no conteúdo publicado por usuários em redes sociais, e expressam o desejo da utilização de outros algoritmos de aprendizado de máquina e de uma rede neural profunda.

5.5 Incorporação contextualizada e recursos probabilísticos baseados em novos transformadores para detecção de depressão nas mídias sociais

Propondo o que considera ser um método inovador para engenharia de recursos, Abbas et al. (2024), intitulando tal método como BERT-RF, extrai *embeddings* contextuais e recursos probabilísticos para a detecção de depressão em postagens de redes sociais. O trabalho utiliza um conjunto de dados com 20.000 tweets classificados como depressivos e não depressivos, disponível publicamente na plataforma Kaggle.

Estes dados de entrada passam pelo processo de extração de recursos utilizando o modelo BERT e então, são inseridos no modelo Random Forest para a geração de recursos probabilísticos de cada classe. Com esta metodologia, Abbas et al. (2024) busca melhorar a identificação da depressão em conteúdos de mídias sociais e aplica diferentes classificadores para uma avaliação e comparação de seus desempenhos.

Para a classificação dos tweets, Random Forest, *Multilayer Perceptron* (MLP), *K-Neighbours Classifier* (KNC), Logistic Regression e LSTM foram escolhidos e treinados com 80% dos dados do conjunto e testados com os 20% restantes. A Tabela 10 exibe os resultados obtidos nesta pesquisa.

Tabela 10 – Resultados utilizando BERT-RF

Técnica	Acurácia	Precisão	sensibilidade	F1
RF	0,98	0,99	0,99	0,99
KNC	0,98	0,99	0,99	0,99
MLP	0,98	0,98	0,98	0,98
LSTM	0,94	0,95	0,94	0,94
LR	0,99	0,99	0,99	0,99

Fonte: Adaptado de Abbas et al. (2024)

Analisando as métricas aplicadas, a pesquisa expõe resultados promissores, atingindo taxa de acurácia de 99% e ressaltando o potencial uso da aprendizagem profunda para a detecção da depressão. Dentre todos os modelos testados, o modelo Logistic Regression (LR) apresentou melhor desempenho, mantendo taxa de 99% em todas as métricas. Já o modelo LSTM obteve o menor desempenho entre todos, mas ainda atingindo taxas maiores do que 90%.

É possível observar, analisando esses resultados, que os modelos com melhores desempenhos representam métodos mais tradicionais de aprendizado de máquina. A rede neural LSTM, que fica em último lugar no ranking, é uma arquitetura bastante utilizada em tarefas de classificação de texto, mas que foi superada por estruturas que

não utilizam neurônios. Isso tem relação direta com a forma como as características textuais foram obtidas, uma vez que o modelo BERT é uma arquitetura Transformer que considera o contexto no qual as palavras estão inseridas e melhora significativamente a classificação de textos.

Por fim, os autores concluem ser possível, em trabalhos futuros, incorporar os modelos em uma API para a detecção de depressão em tempo real. No entanto, não é exposto se o uso da API geraria algum tipo de mensagem para o usuário ou se esta informação será utilizada de outra forma.

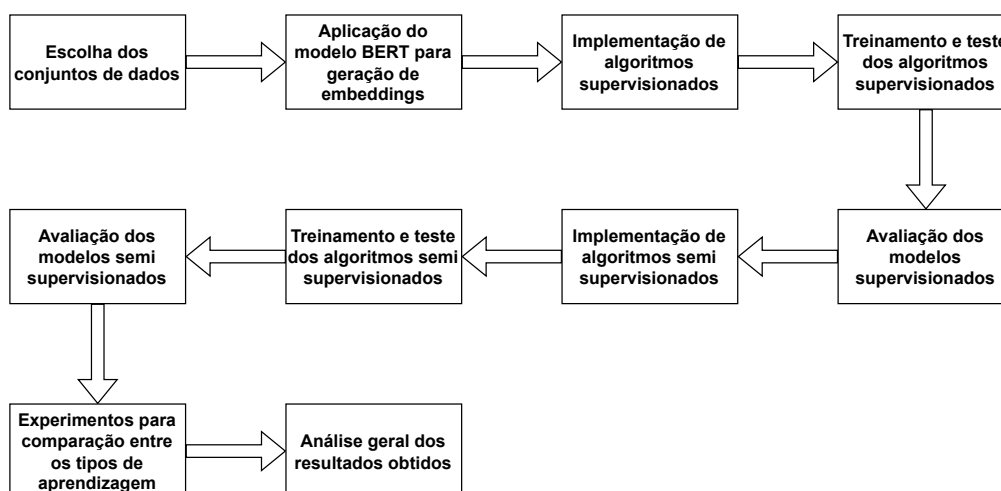
6 METODOLOGIA

Esta seção descreve o plano de trabalho executado durante o desenvolvimento do projeto. Serão descritas as tecnologias, os conjuntos de dados e cenários de experimentos realizados.

6.1 Proposta

A Figura 17 apresenta o fluxo geral do trabalho proposto, o qual tem o objetivo de explorar abordagens para o auxílio na classificação automática de potenciais postagens depressivas ou com indícios de ideação suicida em textos extraídos de redes sociais.

Figura 17 – Fluxograma geral do trabalho proposto



Fonte: Elaborado pela autora

A metodologia proposta se inicia pela escolha dos conjuntos de dados a serem classificados, dois para suicídio e um para depressão. Com esta definição, o processo posterior tem como responsabilidade a aplicação do modelo BERT para geração de *word embeddings*, criando uma nova coluna com tais resultados e salvando, em arquivos csv, para cada um dos conjuntos.

Após estas etapas, algoritmos supervisionados e semi-supervisionados foram aplicados a cada um dos conjuntos de dados, tendo seus desempenhos avaliados com a utilização de métricas como acurácia e F1. Experimentos foram realizados em diferentes cenários para identificar de que forma as abordagens escolhidas poderiam impactar positiva ou negativamente nos resultados finais.

A escolha do aprendizado semi-supervisionado em comparação ao supervisionado se deu pelo fato de que, com o aprendizado semi-supervisionado, é possível aproveitar conjuntos com um menor número de dados rotulados, uma vez que este

problema de domínio mais específico tende a possuir barreiras quanto a possibilidade de extração e rotulagem correta de tais dados.

6.2 Conjunto de Dados

O projeto buscou auxiliar na classificação de textos potencialmente depressivos e com risco de suicídio e, para isso, três conjuntos foram utilizados no treinamento e teste de modelos supervisionados e semi-supervisionados.

6.2.1 Mental Health Twitter

O conjunto de dados Mental Health Twitter, disponível na plataforma Kaggle², foi escolhido para auxiliar na detecção da depressão. Este é um conjunto também utilizado em trabalhos relacionados e publicados, como o trabalho de Abbas et al. (2024) que propõe a detecção de depressão utilizando modelos como Random Forest e LSTM.

20 mil postagens de usuários foram coletadas através da API do Twitter, atualmente X, sendo 10 mil marcadas como “0” (não depressivo) e 10 mil marcadas como “1” (depressivo). Na Tabela 11 é possível visualizar as 3 primeiras linhas do conjunto e 5 das 10 colunas totais. Além das exibidas, também é acessível às informações relativas ao id da postagem (*post_id*), data de criação (*post_created*), seguidores (*followers*), id do usuário (*user_id*) e número de *retweets* (*retweets*).

Tabela 11 – 3 primeiras linhas e 5 colunas do conjunto de dados Mental Health Twitter

post_text	friends	favourites	statuses	label
It's just over 2 years since I was...	211	251	837	1
It's Sunday, I need a break, so I'm...	211	251	837	1
Awake but tired. I need to sleep...	211	251	837	1

Fonte: Elaborado pela autora

Apenas as colunas referentes ao texto e sua classe foram utilizadas na pesquisa e, para explorar os exemplos de tweets extraídos, 2 nuvens de palavras foram criadas. A Figura 18 exibe a nuvem de palavras para tweets rotulados como depressivos, e a Figura 19, para tweets rotulados como não depressivos.

²<https://www.kaggle.com/datasets/infamouscoder/mental-health-social-media>

incorretos.

6.2.2 Suicide Detection

O segundo conjunto de dados baixado da plataforma Kaggle, Suicide Detection³, contém postagens extraídas da rede social Reddit e rotuladas em duas classes: “non-suicide” e “suicide”. Os subreddits “SuicideWatch” e “teenagers” foram alvo desta coleta e o processo de rotulagem não contou com o auxílio de profissionais da saúde. No entanto, o conjunto também foi utilizado no trabalho publicado por Gupta e Pirzada (2023).

A Tabela 12 mostra as primeiras 3 linhas do conjunto, que contém apenas duas colunas representando o texto e sua classe. Para este conjunto, a conversão dos valores da coluna classe foi necessária, transformando “non-suicide” em 0 e “suicide” em 1. Originalmente, o conjunto de dados foi obtido contendo 232.074 linhas, sendo o maior escolhido para este projeto.

Tabela 12 – 3 primeiras linhas do conjunto de dados Suicide Detection

text	class
Ex Wife Threatening SuicideRecently I left my wife...	suicide
Am I weird I don't get affected by compliments if it's coming...	non-suicide
Finally 2020 is almost over... So I can never hear...	non-suicide

Fonte: Elaborado pela autora

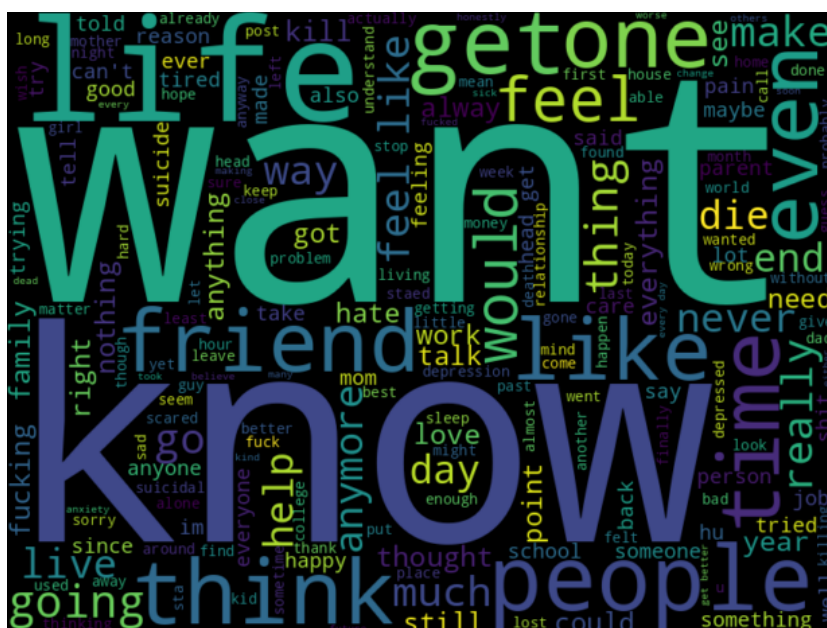
Devido a limitações computacionais e pela escolha do modelo BERT para a criação da representação dos dados, apenas 20 mil exemplos escolhidos aleatoriamente foram utilizados, sendo 10 mil exemplos positivos e 10 mil exemplos negativos. A decisão foi tomada por causa da exigência de memória RAM para o processamento dos textos.

Para este conjunto, nuvens de palavras também foram geradas para cada classe. A Figura 20 exhibe as postagens classificadas como suicídio, e a Figura 21, para as postagens não consideradas suicídio.

Por meio das imagens é possível identificar termos como “want” e “life” em destaque nos exemplos positivos. Todos os exemplos positivos foram extraídos do subreddit “SuicideWatch”, com uma rotulação automática que pressupõe que as postagens publicadas estão relacionadas ao suicídio. Este é um ponto que deve ser levado em consideração, uma vez que relatos de tentativas anteriores, por exemplo, podem ou não conter nova ideação suicida.

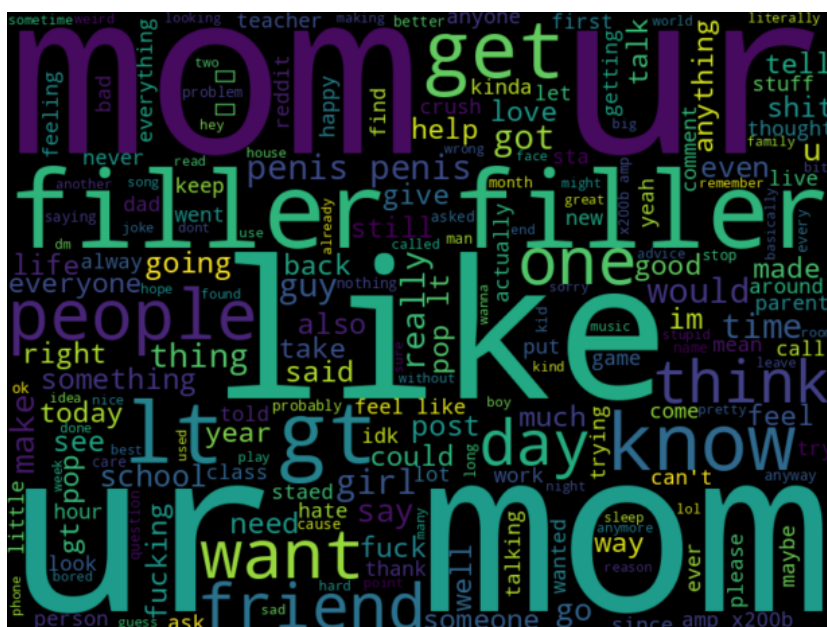
³<https://www.kaggle.com/datasets/nikhileswarkomati/suicide-watch>

Figura 20 – Nuvem de palavras – Postagens classificadas como suicídio no conjunto Suicide Detection



Fonte: Elaborado pela autora

Figura 21 – Nuvem de palavras – Postagens não classificadas como suicídio no conjunto Suicide Detection



Fonte: Elaborado pela autora

6.2.3 Suicidal Data

O último conjunto de dados, intitulado Suicidal Data⁴, é um conjunto disponibilizado pelo artigo científico de Yatapala e Kumara (2021) para a predição de ideação suicida em postagens extraídas do Twitter. O conjunto contém 5.121 exemplos rotulados como 0 (negativos) e 3.998 exemplos rotulados como 1 (positivos).

A Tabela 13 exibe os 3 primeiros exemplos e as duas únicas colunas do conjunto, que armazenam o tweet e sua respectiva classe. Para complementar a exploração do conjunto, 2 nuvens de palavras foram criadas.

Tabela 13 – 3 primeiras linhas do conjunto de dados Suicidal Data

label	tweet
1	my life is meaningless i just want to end my life so badly my life is...
1	muttering i wanna die to myself daily for a few months now i feel...
1	work slave i really feel like my only purpose in life is to make a...

Fonte: Elaborado pela autora

De acordo com a Figura 23, o termo “die” é encontrado com maior frequência em exemplos negativos do que em exemplos positivos, conforme mostra a Figura 22. O tweet *“i have a headache and i have to pee and my laptop is about to die fail”* (Eu tenho uma dor de cabeça e eu tenho que urinar e meu laptop está prestes a morrer falha) é um exemplo negativo que, mesmo utilizando o termo “die”, não se refere ao ato de cometer suicídio e sim, de que a bateria de seu laptop estaria acabando.

Para exemplos positivos, que são exibidos na Figura 22, o termo “life” é destacado como um uso frequente.

⁴<https://github.com/AminulIsrael/Predicting-Suicide-Ideation>

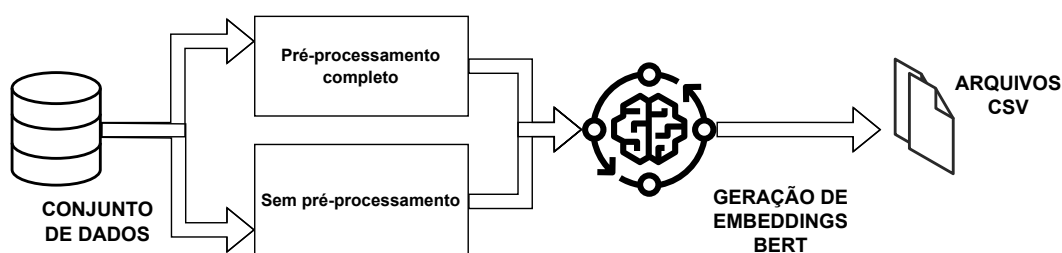
Figura 22 – Nuvem de palavras – Postagens classificadas como ideação suicida no conjunto Suicidal Data

Estes exemplos demonstram a importância de considerar o contexto neste domínio de pesquisa, além do cuidado no processo de rotulagem. No conjunto, no entanto, não existe indicativo de que as classes foram atribuídas com o auxílio de um profissional da saúde.

6.3 Pré-processamento e geração de word embeddings

Cada exemplo nos conjunto de dados específicos foi representado por *word embeddings* contextuais geradas pelo modelo BERT, com dimensão fixa de 768, permitindo a extração de características em um formato que possa ser compreendido pelo classificador. O modelo BERT (*Bidirectional Encoder Representations from Transformers*) foi escolhido por ser, atualmente, um dos principais utilizados para geração de vetores de *embeddings* dinâmicos, uma vez que constrói cada representação com base no contexto em que as palavras estão inseridas. Para fins comparativos, 2 arquivos de *embeddings* foram criados para cada um dos conjuntos utilizados, conforme mostra a Figura 24.

Figura 24 – Geração de diferentes *Word Embeddings* para um conjunto de dados

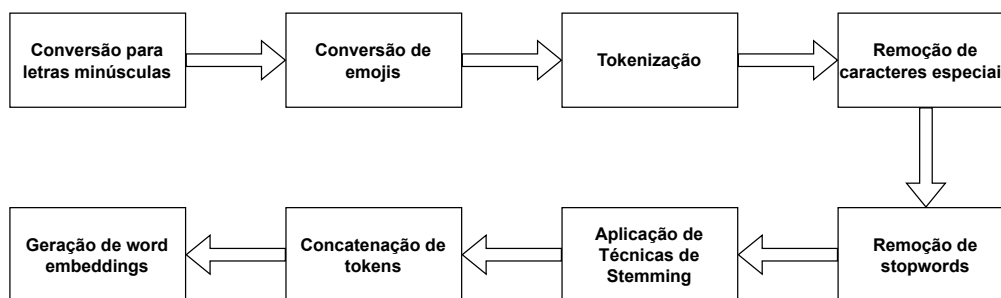


Fonte: Elaborado pela autora

Conforme mostra a Figura 24, cada conjunto de dados teve *embeddings* geradas em 2 cenários, o que resultou 2 arquivos diferentes: O primeiro arquivo foi pré-processado com a aplicação de diferentes técnicas, enquanto o segundo não contou com manipulação inicial.

Para a geração do primeiro arquivo, a Figura 25 exhibe o fluxograma de todas as tarefas que englobaram o pré-processamento aplicado, processo que antecedeu a utilização do modelo BERT para a geração das *embeddings*. Antes de treinar um classificador e posteriormente testá-lo, é comum que as amostras envolvidas passem pela etapa de pré-processamento. Esta etapa, além de realizar a remoção de dados nulos e outros aspectos não desejados, também permite que os textos tenham suas características representadas em vetores numéricos.

Iniciando o pré-processamento, a conversão de emojis é implementada utilizando a biblioteca emoji e o método demojize. Cada emoji encontrado em um exemplo dos conjuntos de dados é transformado em uma representação textual, como o exemplo da Figura 26. Por mais que, usualmente, os emojis sejam removidos durante

Figura 25 – Etapas realizadas no pré-processamento dos textos

Fonte: Elaborado pela autora

o pré-processamento, esta metodologia o mantém, com a hipótese de que os emojis podem auxiliar na compreensão da emoção dos usuários.

Figura 26 – Exemplo de conversão de emoji

```

[9] emoji.demojize("Hoje me sinto tão 😊")
    'Hoje me sinto tão :beaming_face_with_smiling_eyes:'

emoji.demojize("Estou me sentindo tão 😞")
    'Estou me sentindo tão :disappointed_face:'
  
```

A imagem mostra duas linhas de código executadas em um terminal. A primeira linha converte a frase 'Hoje me sinto tão 😊' para 'Hoje me sinto tão :beaming_face_with_smiling_eyes:'. A segunda linha converte 'Estou me sentindo tão 😞' para 'Estou me sentindo tão :disappointed_face:'.

Fonte: Elaborado pela autora

As postagens, antes de serem tokenizadas, foram submetidas a conversão de suas palavras, de modo que todas as letras se tornam minúsculas. Seguindo para o processo de tokenização, cada uma das sequências de strings presentes nos conjuntos de dados foram divididas em partes menores. No passo seguinte, a metodologia buscou remover todas as características indesejadas presentes nos exemplos.

Nesta etapa, ainda mais considerando que são postagens retiradas de redes sociais, URLs, @, # e números foram excluídos. A técnica da remoção de *stopwords* foi aplicada, após a remoção dos caracteres, com a utilização da biblioteca NLTK para a obtenção da lista de termos irrelevantes. A lista, com termos em inglês, foi modificada com a exclusão da palavra “não” do *array*, uma vez que, para este problema, é um termo relevante. “Não me sinto depressivo” e “me sinto depressivo”, por exemplo, são frases com significados diferentes, mas que se tornam iguais após a remoção da palavra “não”.

A técnica de *stemming*, utilizando o algoritmo SnowballStemmer, foi aplicada para reduzir cada palavra à sua raiz. Já a concatenação dos tokens foi utilizada para

retornar a sequência das strings pré-processadas que representam as postagens dos usuários. O passo final do pré-processamento, visando transformar cada palavra em uma representação numérica, consistiu na aplicação do modelo BERT para a criação de *embeddings* dinâmicos que consideram o contexto em que a palavra está inserida.

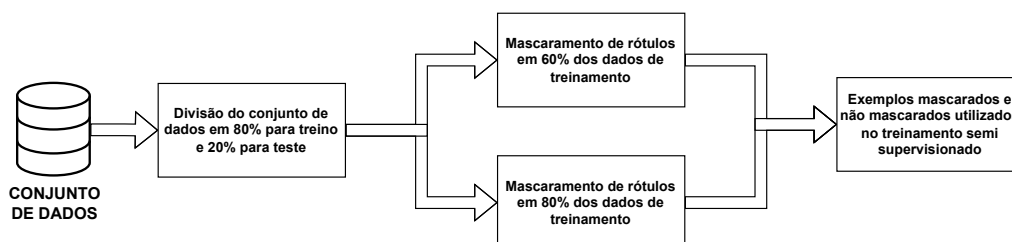
O segundo arquivo foi criado levando em consideração que, para a geração de embeddings BERT, algumas etapas de pré-processamento deixam de ser necessárias, podendo até mesmo atrapalhar o desempenho do modelo. Sendo assim, as *embeddings* foram geradas a partir de um conjunto de dados sem nenhuma manipulação inicial. Isso quer dizer que, nem mesmo o processo de remoção de caracteres especiais foi aplicado, assim como a conversão para letras minúsculas também foi excluído.

6.4 Divisão do conjunto e mascaramento de rótulos

Assim como definido em trabalhos relacionados e comparados nesta pesquisa, os conjuntos foram divididos em 80% dos dados para treino e 20% para teste, com a utilização de sementes para garantir que, ao reproduzir os experimentos, os subconjuntos fossem os mesmos. A distribuição dos dados nos subconjuntos de treinamento e teste foi escolhida com base nos trabalhos relacionados, os quais utilizaram a mesma porcentagem. Esta forma de divisão, conhecida como *holdout*, foi executada 50 vezes para cada um dos algoritmos e com N sementes diferentes. Desta forma, as métricas finais de avaliação foram obtidas através da média das métricas de cada uma das execuções.

Uma vez que esta pesquisa busca implementar modelos semi-supervisionados, uma função de mascaramento de rótulos foi criada e aplicada aos dados de treinamento. A Figura 27 é um esquema deste processo.

Figura 27 – Divisão dos conjuntos de dados



Fonte: Elaborado pela autora

Conforme demonstrado pelo fluxograma da Figura 27, cada conjunto de teste (20% do conjunto total) foi submetido ao mascaramento de rótulos. O mascaramento de rótulos atribui, a uma determinada porcentagem de exemplos, o rótulo -1 como indicativo de uma amostra não rotulada. Para os experimentos, 80 e 60 foram os

valores escolhidos, de forma aleatória, para corresponder a porcentagem de dados mascarados.

6.5 Classificadores supervisionados

Para a construção dos classificadores supervisionados, a biblioteca de código aberto scikit-learn⁵ foi escolhida. Esta é uma biblioteca de apoio para o desenvolvimento de tarefas de aprendizado de máquina utilizando a linguagem de programação Python, com diversos modelos e métodos auxiliares disponíveis. Para esta pesquisa, os modelos SVM, RandomForestClassifier e LogisticRegression foram escolhidos. A avaliação do desempenho dos modelos se deu pela aplicação das métricas acurácia, F1 score, sensibilidade e precisão.

6.6 Classificadores semi-supervisionados

Para a aplicação do algoritmo de Competição e Cooperação entre Partículas, o código disponibilizado pelo autor no GitHub⁶, escrito em MATLAB, foi utilizado. Os classificadores restantes foram construídos com a utilização da biblioteca de scikit-learn, em linguagem Python. LabelPropagation, LabelSpreading e SelfTrainingClassifier são os três modelos semi-supervisionados listados na documentação da biblioteca, sendo o último o classificador que promete permitir que classificadores supervisionados funcionem como um classificador semi-supervisionado.

Todos os modelos foram treinados e testados nos conjuntos de dados com exemplos rotulados e não rotulados, além da aplicação das métricas acurácia, F1 score, sensibilidade e precisão para avaliação.

6.7 Redução de dimensionalidade

Para a redução de dimensionalidade das *word embeddings* geradas, o algoritmo de Análise de Componentes Principais (*Principal Component Analysis* - PCA) foi implementado utilizando a biblioteca scikit-learn.

O PCA permite que a dimensionalidade de um conjunto seja reduzida, minimizando a perda de informações neste processo e aumentando a capacidade de interpretação. Sendo considerada uma técnica estatística, a Análise de Componentes Principais busca encontrar novas variáveis não correlacionadas que maximizam a variância dos dados (JOLLIFFE; CADIMA, 2016).

Para os experimentos supervisionados, o algoritmo foi avaliado utilizando os valores 150, 300, 450 e 600 para o parâmetro *n_components*, representando o números de componentes principais mantidos pelo PCA. Nos experimentos semi-supervisionados, o PCA foi configurado com 10, 20, 30, 40, 50, 100 e 150 componentes.

⁵<https://scikit-learn.org/stable/>

⁶<https://github.com/fbreve/Particle-Competition-and-Cooperation>

Em ambos os casos, os modelos foram treinados e testados nos cenários descritos, permitindo analisar e comparar o impacto da redução de dimensionalidade nos diferentes tipos de aprendizado.

6.8 Combinação de modelos

A abordagem de combinação de modelos foi definida como cenário para a classificação supervisionada dos conjuntos, aplicando a técnica *Stacking*.

Stacking é a técnica que combina diferentes modelos preditivos, chamados de modelos base, para que um metamodelo seja criado. Neste método, o novo modelo aprende a combinar as melhores previsões dos modelos utilizados como base. No entanto, o *Overfitting* é apontado como um possível problema a ser observado na aplicação da técnica (MOHAMMED; KORA, 2023).

A técnica foi implementada através do classificador Stacking Classifier, da biblioteca scikit-learn, e teve como metamodelo os classificadores que registraram os melhores resultados quando testados de forma individual.

7 RESULTADOS

Para a avaliação dos classificadores, nos três conjuntos de dados, foram aplicadas as métricas de sensibilidade, precisão, acurácia e F1 score. Para os experimentos, o método *holdout* foi executado 50 vezes, dividindo cada conjunto em subconjuntos de treino e teste.

Para garantir que ao repetir tais experimentos a mesma subdivisão fosse obtida, 50 sementes foram escolhidas aleatoriamente e salvas para serem informadas ao método `train_test_split` da biblioteca `sklearn`, responsável por dividir um conjunto em treino e teste com base em uma porcentagem escolhida. Todos os experimentos foram realizados com o conjunto de dados dividido em 80% para treino e 20% para teste.

7.1 Classificação Supervisionada

Os experimentos de classificação supervisionada foram realizados, nesta pesquisa, com a utilização dos modelos SVM, Logistic Regression e Random Forest. Os conjuntos Mental Health Twitter (MHT), Suicidal Data (SDa) e Suicide Detection (SDe) foram divididos em subconjuntos com 80% dos dados para treinamento e 20% dos dados para teste. O método, conhecido como *holdout*, foi executado 50 vezes e a média das métricas acurácia, precisão, sensibilidade e F1 score foram obtidas para todos eles.

Nesta seção, os melhores resultados obtidos nas abordagens escolhidas para a pesquisa serão apresentados em tabelas e gráficos. No entanto, os resultados completos, com todos os modelos e suas variações, estão incluídos no Apêndice deste trabalho.

7.1.1 Classificação com dados pré-processados e não processados

Para todos os conjuntos escolhidos para esta pesquisa, os experimentos foram realizados em dois cenários: com dados pré-processados e não processados. Para os experimentos no cenário não processado (NP), as *embeddings* foram geradas para os três conjuntos de dados, sem que nenhuma etapa de pré-processamento fosse aplicada. Para os experimentos no cenário pré-processado (PP), as *embeddings* foram geradas após diferentes etapas terem sido aplicadas para manipulação inicial dos dados.

Os resultados obtidos por esta pesquisa, em ambos os cenários, foram comparados com os artigos que disponibilizaram os conjuntos escolhidos. Mesmo utilizando apenas dados pré-processados, os trabalhos de Abbas et al. (2024) e Yatapala e Kumara (2021) foram comparados, auxiliando na análise do impacto da ausência de pré-processamento.

Mental Health Twitter

O conjunto MHT foi comparado com o trabalho de Abbas et al. (2024),

que utiliza *tweets* extraídos de redes sociais para auxiliar na detecção de depressão, aplicando o modelo BERT para a geração de *embeddings* contextuais. Embora ambas as pesquisas compartilhem técnicas semelhantes de pré-processamento, existem diferenças significativas em suas metodologias. Enquanto Abbas et al. (2024) realiza a remoção de emojis, a presente pesquisa aplica a conversão desses elementos em representações textuais. Dessa forma, as metodologias adotadas não são idênticas.

Na Tabela 14 os valores de acurácia, precisão, sensibilidade e F1 podem ser visualizados para os cenários PP e NP. Por executar a classificação 50 vezes, em subconjuntos diferentes, as métricas apresentadas representam a média dos valores obtidos em cada execução. Os resultados alcançados pelos modelos Logistic Regression (LR), SVM e Random Forest (RF) foram comparados com os resultados obtidos no trabalho de Abbas et al. (2024), que utilizou os classificadores LR, RF, Multi-Layer Perceptron (MLP) e K-Neighbors Classifier (KNC) em dados pré-processados.

Tabela 14 – Resultados do conjunto Mental Health Twitter com classificação supervisionada

Trabalho	Cenário	Técnica	Acurácia	Precisão	Sensibilidade	F1
Pesquisa atual	NP	SVM RBF	0,8115	0,8095	0,8144	0,8119
Pesquisa atual	NP	LR	0,8056	0,8087	0,8004	0,8045
Pesquisa atual	NP	RF 200 árvores	0,7931	0,7852	0,8068	0,7958
Pesquisa atual	NP	RF 100 árvores	0,7828	0,7806	0,7865	0,7835
Abbas et al. (2024)	PP	RF	0,71	0,72	0,72	0,72
Pesquisa atual	PP	SVM RBF	0,6876	0,6707	0,7364	0,7020
Pesquisa atual	PP	RF 200 árvores	0,6880	0,6740	0,7278	0,6998
Pesquisa atual	PP	SVM Linear	0,6810	0,6715	0,7081	0,6892
Pesquisa atual	PP	RF 100 árvores	0,6798	0,6709	0,7052	0,6875
Pesquisa atual	PP	LR	0,6809	0,6762	0,6935	0,6847
Abbas et al. (2024)	PP	KNC	0,61	0,61	0,61	0,61
Abbas et al. (2024)	PP	LR	0,56	0,56	0,56	0,56
Abbas et al. (2024)	PP	MLP	0,51	0,55	0,51	0,42

Fonte: Elaborado pela autora

Iniciando a análise pelo cenário pré-processado com a comparação entre

as técnicas LR, é possível observar que a pesquisa atual atingiu melhores resultados. Enquanto Abbas et al. (2024) obteve 0,56% em todas as métricas calculadas, esta pesquisa atingiu 0,6809 para acurácia, 0,6762 para precisão, 0,6935 para sensibilidade e 0,6847 para F1 score. Apesar da aplicação no mesmo conjunto de dados, com o mesmo modelo para classificação e com geração de *embeddings* utilizando BERT, diferenças na implementação podem explicar o desempenho superior obtido pela pesquisa atual.

Abbas et al. (2024) configurou o modelo LR com parâmetros `random_state=None`, `test_size=0.2`, `n_splits=10`, `shuffle=True`, enquanto a pesquisa atual utiliza a configuração padrão da biblioteca `sklearn`, onde os parâmetros `test_size`, `n_splits` e `shuffle` não estão listados na documentação como valores configuráveis. Além disso, Abbas et al. (2024) pré-processa os textos de forma diferente, realizando remoção de mais informações, como o descarte de emojis e nomes de usuários, bem como uma etapa adicional de normalização de dados.

Nesta pesquisa, Random Forest foi executado em cenários que variaram os valores para o parâmetro `n_estimators`, que indica o número de árvores na floresta. No trabalho proposto por Abbas et al. (2024), o modelo foi configurado com `n_estimators = 100` e `class_weight = None`, valores iguais ao configurado nesta pesquisa. As configurações dos trabalhos são diferentes para o parâmetro `max_depth` e `multi_class`, sendo o último uma informação não configurável no modelo implementado com `sklearn`.

Comparando com a técnica RF 100 árvores, Abbas et al. (2024) atingiram 71% para a métrica acurácia e 72% para precisão, sensibilidade e F1. A pesquisa atual, no entanto, obteve valores menores para as métricas calculadas. Com 200 árvores, a classificação com Random Forest apresentou melhora de desempenho em relação a 100 árvores, mas ainda manteve o resultado inferior aos 71% atingidos por (ABBAS et al., 2024).

Além das técnicas em comum, os modelos KNC e MLP implementados por Abbas et al. (2024) podem ser comparados com todos os resultados, assim como o modelo SVM. A aplicação da SVM RBF, por exemplo, atingiu acurácia de 0,6876 e registrou desempenho superior a média das métricas calculadas para LR, KNC e MLP, mas se manteve atrás dos resultados obtidos com o modelo Random Forest de Abbas et al. (2024).

De modo geral, o pior desempenho listado na Tabela 14 para o cenário PP foi observado no modelo MLP (com acurácia de 0,51), enquanto o melhor desempenho foi obtido pelo modelo RF (com acurácia de 0,71), ambos de Abbas et al. (2024). Nos modelos implementados na pesquisa atual, SVM RBF apresentou avaliação superior aos outros classificadores.

Para o cenário não processado, como exibiu a Tabela 14, os resultados obtidos para a classificação de textos potencialmente depressivos foram superiores

aos obtidos no cenário PP. O modelo SVM RBF atingiu o melhor desempenho em comparação a todos os classificadores testados, com acurácia de 0,8115 e F1 score de 0,8119, valores superiores a 0,6876 e 0,7020 que haviam sido obtidos com dados pré-processados.

Através da Tabela 14, também é possível determinar que a ausência de pré-processamento permitiu que a técnica RF atingisse resultados superiores ao obtido por Abbas et al. (2024) para o mesmo modelo, mas com dados pré-processados. Além disso, os resultados do modelo de Regressão Logística permaneceram maiores do que a pesquisa comparada, com uma diferença de 0,2456 para acurácia.

Em uma análise geral do impacto da abordagem, compreende-se que, ao não aplicar técnicas de pré-processamento neste conjunto, obtêm-se melhores resultados para o auxílio da detecção de potencial depressão. As classificações do conjunto pré-processado, como reforçam as métricas de ambos os trabalhos apresentados, não haviam conseguido obter resultados maiores do que 0,71 para acurácia e 0,72 para F1 score, observadas na técnica RF da pesquisa de Abbas et al. (2024). Na abordagem sem pré-processamento, 0,8115 e 0,8119 foram os maiores valores para acurácia e F1 score, exibidas nos resultados da técnica SVM RBF da pesquisa atual, com uma diferença de 0,1015 entre as acurácias citadas.

Suicidal Data

O conjunto SDa foi obtido do trabalho de Yatapala e Kumara (2021), que aplicou algumas técnicas para a representação dos textos e utilizou SVM como classificador. Escolhendo a melhor configuração para cada um dos modelos testados na pesquisa atual, LR, SVM e RF, a Tabela 15 exibe os resultados dos cenários pré-processado e não processado para comparação entre as duas pesquisas.

Embora utilizem SVM em um conjunto pré-processado, as pesquisas diferem em relação às técnicas utilizadas para a limpeza dos dados e extração de características. Para o pré-processamento do conjunto, Yatapala e Kumara (2021) aplicaram a remoção de identificadores do Twitter, números, caracteres especiais e pontuações. Palavras com menos de dois caracteres também foram removidas, mas os autores não deixam evidente se ocorreu a remoção de *stopwords* e aplicação da técnica de *stemming*. Por fim, Word2Vec, Doc2Vec e TF-IDF foram utilizados para a extração de características.

A pesquisa atual, por sua vez, utiliza o modelo BERT para a extração de características, gerando *embeddings* que incorporam informações semânticas com base no contexto em que a palavra está inserida. Esse tipo de representação é mais adequado para o domínio da pesquisa, quando comparado a *embeddings* estáticos, como os gerados pelo Word2Vec. Dessa forma, é possível observar que a presente pesquisa, no cenário PP, obteve melhores resultados para acurácia, única métrica que Yatapala e Kumara (2021) exibem nos resultados da pesquisa.

Com *embeddings* BERT classificadas com SVM RBF e SVM Linear, a acu-

Tabela 15 – Resultados do conjunto Suicidal Data com classificação supervisionada

Trabalho	Cenário	Técnica	Acurácia	Precisão	Sensibilidade	F1
Pesquisa atual	NP	LR	0,9267	0,9292	0,9019	0,9153
Pesquisa atual	NP	SVM RBF	0,9237	0,9516	0,8707	0,9093
Pesquisa atual	NP	RF 200 árvores	0,9056	0,9523	0,8267	0,8850
Pesquisa atual	PP	LR	0,9042	0,9125	0,8650	0,8880
Pesquisa atual	PP	SVM Linear	0,8987	0,9007	0,8649	0,8823
Pesquisa atual	PP	SVM RBF	0,8988	0,9439	0,8185	0,8767
Pesquisa atual	PP	RF 200 árvo- res	0,8906	0,9348	0,8075	0,8664
Yatapala e Kumara (2021)	PP	TF-IDF + SVM	0,8494	-	-	-
Yatapala e Kumara (2021)	PP	Word2vec + SVM	0,7958	-	-	-
Yatapala e Kumara (2021)	PP	Doc2vec + SVM	0,7399	-	-	-

Fonte: Elaborado pela autora

rácia atingida foi de 0,8988 e 0,8987, respectivamente. Para a metodologia proposta por Yatapala e Kumara (2021), a acurácia de maior valor foi obtida com a aplicação de TF-IDF junto a SVM, resultando em 0,8494.

A tabela também permite identificar que o modelo LR obteve o melhor desempenho em comparação a todos os modelos aplicados na pesquisa, assim como foi superior aos valores de acurácia obtidos por Yatapala e Kumara (2021) na execução do modelo SVM. O modelo Random Forest também se mostrou superior aos resultados de Yatapala e Kumara (2021), com acurácia de 0,8906.

Os resultados obtidos no cenário PP, de modo geral, indicaram que a escolha da técnica para extração de características impactou de forma significativa e positiva no desempenho dos modelos testados. Com o uso do modelo BERT, todos os modelos testados foram capazes de superar a acurácia de 0,8494 que foi obtida por Yatapala e Kumara (2021) na técnica TF-IDF + SVM.

Seguindo para a análise do cenário NP, visualizando as melhores configurações na Tabela 15, nota-se que o modelo de Regressão Logística atingiu o melhor desempenho entre todos os modelos testados com a ausência de pré-processamento, quando considerado que possui duas métricas com os maiores valores: acurácia e F1 score. O pior desempenho foi registrado com o modelo RF 200 árvores, com acurácia de 0,9056 e F1 score de 0,8850.

Enquanto a Regressão Logística manteve equilíbrio entre as métricas, com

valores entre 0,9019 e 0,9267, Random Forest apresentou um comportamento já observado em outros experimentos: a diferença significativa entre precisão e sensibilidade, justificando a métrica F1 mais baixa e indicando maiores chances de falsos negativos. Para este domínio de pesquisa, não identificar o desejo que precede a tentativa de suicídio se torna um problema impactante, uma vez que pode resultar na falta de intervenções que buscam reduzir o número de vítimas deste ato.

Além disso, assim como no cenário com dados pré-processados, todos os resultados da pesquisa atual foram superiores aos obtidos por Yatapala e Kumara (2021), com uma diferença ainda maior entre as métricas.

Suicide Detection

Diferente dos conjuntos anteriores, o conjunto SDe, que foi obtido no trabalho de Gupta e Pirzada (2023), não terá seus resultados comparados. Além de implementarem a rede LSTM para a detecção de potencial comportamento suicida, os autores utilizam o conjunto de dados completo, o que também não foi feito na metodologia desta pesquisa. A Tabela 16 apresenta os melhores resultados de acurácia, precisão, sensibilidade e F1 score obtidos para os modelos SVM, LR e RF nos cenários PP e NP.

Tabela 16 – Resultados do conjunto Suicide Detection com classificação supervisionada

Cenário	Técnica	Acurácia	Precisão	Sensibilidade	F1
NP	SVM RBF	0,9301	0,9324	0,9276	0,9300
NP	LR	0,9291	0,9295	0,9287	0,9291
PP	LR	0,9005	0,8971	0,9048	0,9009
PP	SVM Linear	0,8978	0,8952	0,9012	0,8982
NP	RF 200 árvores	0,8870	0,9091	0,8602	0,8839
PP	RF 200 árvores	0,8583	0,8712	0,8412	0,8559

Fonte: Elaborado pela autora

Para o cenário pré-processado, conforme pode ser observado na Tabela 16, o modelo LR atingiu maiores valores em todas as métricas calculadas, sendo considerado o melhor classificador entre todos os aplicados neste conjunto. O modelo RF, no entanto, apresentou o pior desempenho, mesmo com 200 árvores utilizadas, registrando acurácia de 0,8583 e F1 score de 0,8559.

No cenário não processado, a classificação com Regressão Logística apresentou acurácia e F1 score de 0,9291, com 0,0008 de diferença entre os valores das métricas precisão e sensibilidade, demonstrando capacidade significativa em identificar exemplos positivos e negativos. O modelo Random Forest, entretanto, apresentou acurácia e F1 score mais baixos, assim como uma diferença de 0,0489 entre precisão e sensibilidade.

O melhor desempenho geral foi registrado pelo modelo SVM com kernel RBF, que alcançou acurácia de 0,9301 e F1 score de 0,9300.

7.1.2 Classificação com redução de dimensionalidade

Para análise do impacto da redução de dimensionalidade no desempenho dos modelos de classificação em diferentes cenários de pré-processamento, o algoritmo de Análise de Componentes Principais (*Principal Component Analysis* - PCA) foi escolhido e executado nos conjuntos representados por *word embeddings*.

Com o auxílio da biblioteca sklearn, o algoritmo PCA foi implementado e aplicado aos conjuntos de dados, com diferentes valores para o número de componentes principais (k). Iniciando-se por 150, os conjuntos também foram testados com 300, 450 e 600 componentes.

Mental Health Twitter

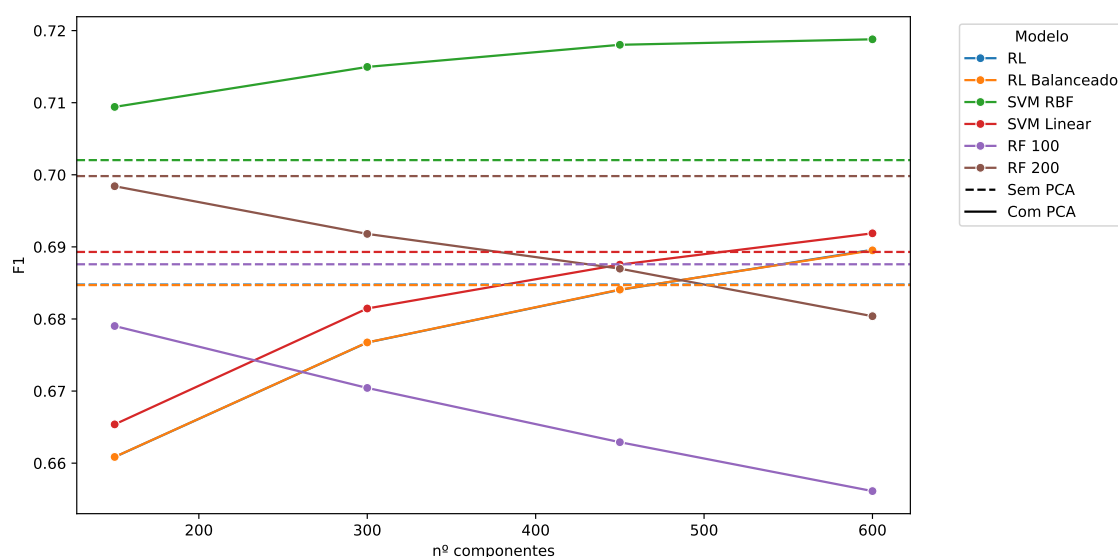
Nestes experimentos, o método *holdout* foi executado 50 vezes para cada configuração do algoritmo PCA e nos dois cenários já testados em seções anteriores: dados com pré-processamento completo e sem pré-processamento.

Dados pré-processados

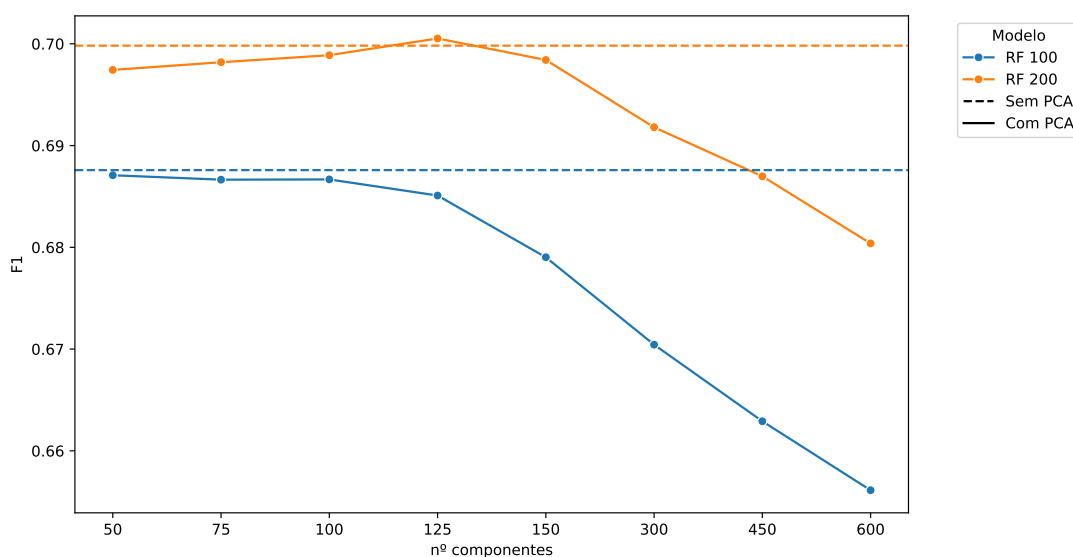
O conjunto MHT pré-processado teve o algoritmo PCA aplicado com 4 diferentes valores para o número de componentes principais. A variância total mantida após a execução foi de 0,8472 para 150 componentes, 0,9310 para 300, 0,9687 para 450 e 0,9889 para 600 componentes. Na Figura 28, é possível visualizar a variação da métrica F1 para diferentes números de componentes para os modelos RL, SVM e RF.

Conforme é visualizado no gráfico, à medida em que o número de componentes aumenta, a métrica F1 também evolui para os modelos SVM e RL. Para o modelo RF, no entanto, os valores de F1 tendem a diminuir.

Por conta desse comportamento, novos experimentos foram realizados exclusivamente para o modelo RF, buscando identificar se um número menor de componentes resultaria em resultados mais promissores. A Figura 29 exhibe estes resultados, com o valor da métrica F1 score para 50, 75, 100 e 125 componentes.

Figura 28 – Mental Health Twitter pré-processado - PCA com modelos RL, SVM e RF

Fonte: Elaborado pela autora

Figura 29 – Mental Health Twitter pré-processado - PCA com modelo RF

Fonte: Elaborado pela autora

Com a aplicação da redução de dimensionalidade, LR em sua versão padrão e balanceada obtiveram, respectivamente, 0,6831 e 0,6832 com 450 componentes. O modelo SVM RBF atingiu 0,7281 com 600 componentes, enquanto a versão linear alcançou 0,6900 com 450. Sendo esses os melhores resultados para cada classificador, é possível estabelecer uma relação com a variância mantida quando 450 e 600 componentes são utilizados na configuração do algoritmo PCA.

O modelo RF, que apresentou o pior desempenho com os mesmos componentes citados anteriormente, registrou o melhor desempenho com o uso de 50

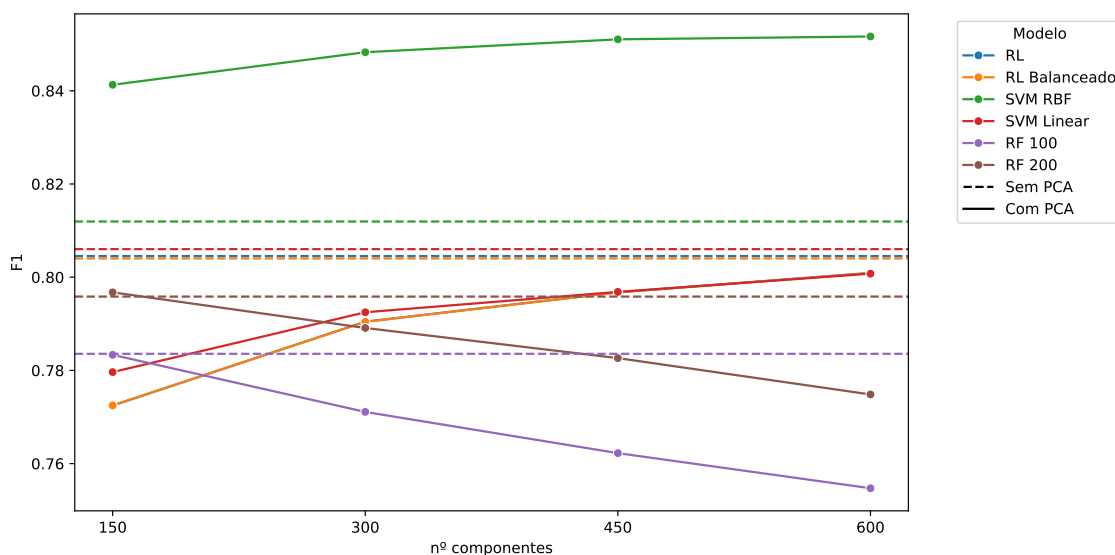
e 125 componentes, para a configuração com 100 e 200 árvores, respectivamente. Para RF 100, o melhor valor para F1 score foi de 0,6870, enquanto RF 200 obteve 0,7005, conforme mostra a figura 29. Os resultados das métricas acurácia, precisão e sensibilidade, que são exibidas no Apendice desta pesquisa, também reforçam a compreensão de que, mesmo com uma menor variância mantida pela aplicação do algoritmo PCA com baixo número de componentes, as classificações foram superiores.

Buscando identificar se este comportamento havia sido exibido em outras pesquisas, mesmo que não a variância mantida não fosse detalhada, o trabalho de Chen et al. (2022) foi encontrado. Na pesquisa, onde o modelo Random Forest foi utilizado junto ao PCA para classificação automática de textos jurídicos, o maior valor de F1 score foi obtido quando 400 componentes foram utilizados. Os experimentos foram realizados com números de 100 a 9 mil componentes e, à medida em que eles aumentaram, a métrica em questão sofreu oscilações, reduzindo seu valor em comparação ao uso de apenas 400 componentes.

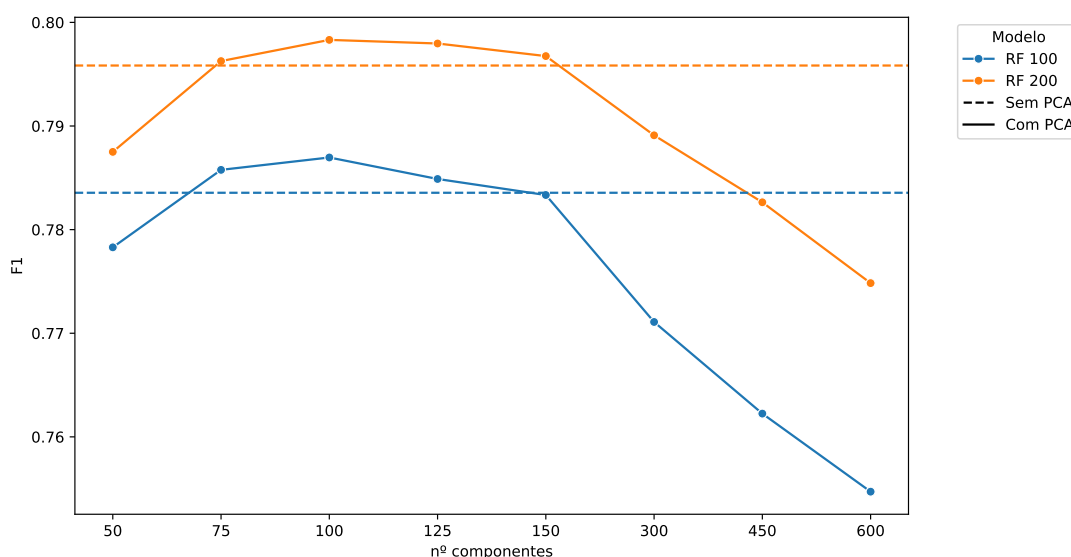
Dados não processados

O conjunto MHT sem pré-processamento também foi testado com redução de dimensionalidade. A variância total mantida após a execução do algoritmo PCA foi de 0,8478 para 150 componentes, 0,9324 para 300, 0,9698 para 450 e 0,9893 para 600 componentes.

Figura 30 – Mental Health Twitter não processado - PCA com modelos RL, SVM e RF



Fonte: Elaborado pela autora

Figura 31 – Mental Health Twitter não processado - PCA com modelo RF

Fonte: Elaborado pela autora

Reproduzindo o comportamento de melhoria observado nos experimentos com dados pré-processados, os modelos SVM e Regressão Logística apresentaram melhoria na métrica F1 score à medida que o número de componentes foi crescendo, conforme evidencia a Figura 30.

Com o modelo SVM de kernel linear, os resultados obtidos foram os melhores entre todos os registrados no experimento, independente do número de componentes utilizados. As maiores métricas do classificador foram obtidas com o uso de 600 componentes, com F1 de 0,8516.

Assim como no cenário de dados pré-processados, o modelo RF, que apresentou o pior desempenho considerando 150, 300, 450 e 600 componentes, foi testado com outras configurações do algoritmo PCA.

A Figura 31 exibe os resultados para o modelo RF com 100 e 200 árvores. Com 100 árvores, o classificador obteve o melhor valor com 100 componentes, registrando 0,7782. Com 200 árvores, para a mesma quantidade de componentes, o melhor resultado foi de 0,7983.

Suicidal Data

O conjunto Suicidal Data, para detecção de potencial risco de suicídio, foi inicialmente manipulado em diferentes cenários para que o impacto do pré-processamento fosse analisado. Agora, utilizando o algoritmo PCA, o impacto do uso da redução de dimensionalidade e os resultados obtidos pelos modelos LR, SVM E RF serão visualizados.

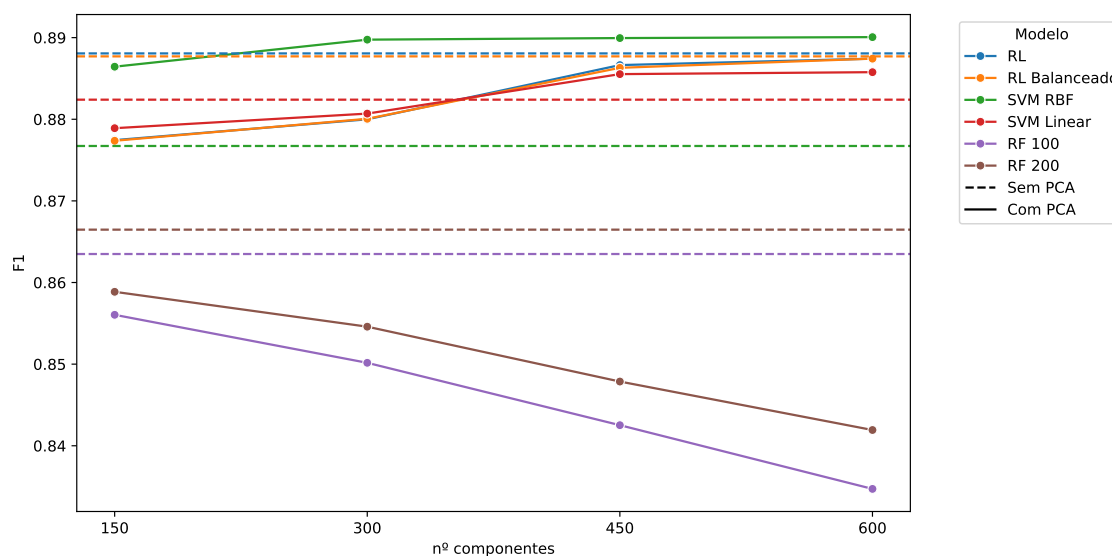
Dados pré-processados

Quando o conjunto pré-processado teve a dimensionalidade reduzida, a variância total mantida após a execução do PCA foi de 0,8856 para 150 componentes,

0,9502 para 300, 0,9782 para 450 e 0,9925 para 600 componentes.

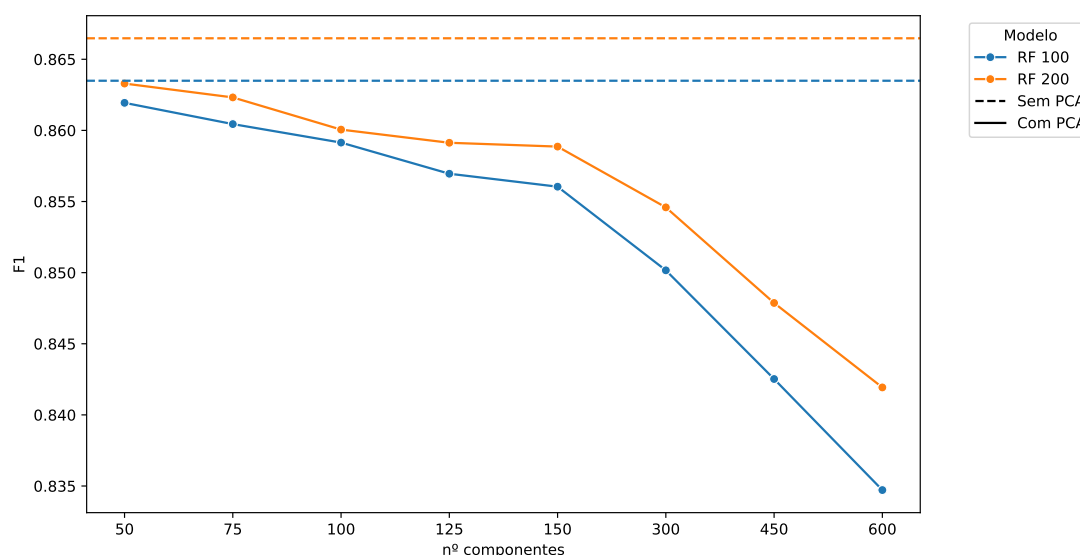
Conforme demonstra a Figura 32, os resultados obtidos no uso dos dados pré-processados variam consideravelmente à medida em que o número de componentes utilizadas no PCA são alterados.

Figura 32 – Suicidal Data pré-processado - PCA com modelos RL, SVM e RF



Fonte: Elaborado pela autora

Figura 33 – Suicidal Data pré-processado - PCA com modelo RF



Fonte: Elaborado pela autora

A Figura 32 evidencia que os modelos de Regressão Logística e SVM apresentaram os melhores desempenhos com o uso de 600 componentes, quando 0,9925 da variância do conjunto foi mantida, enquanto os piores resultados foram registrados

com 150 componentes. Para o modelo SVM RBF, por exemplo, foi registrado F1 score de 0,8900 com 600 componentes.

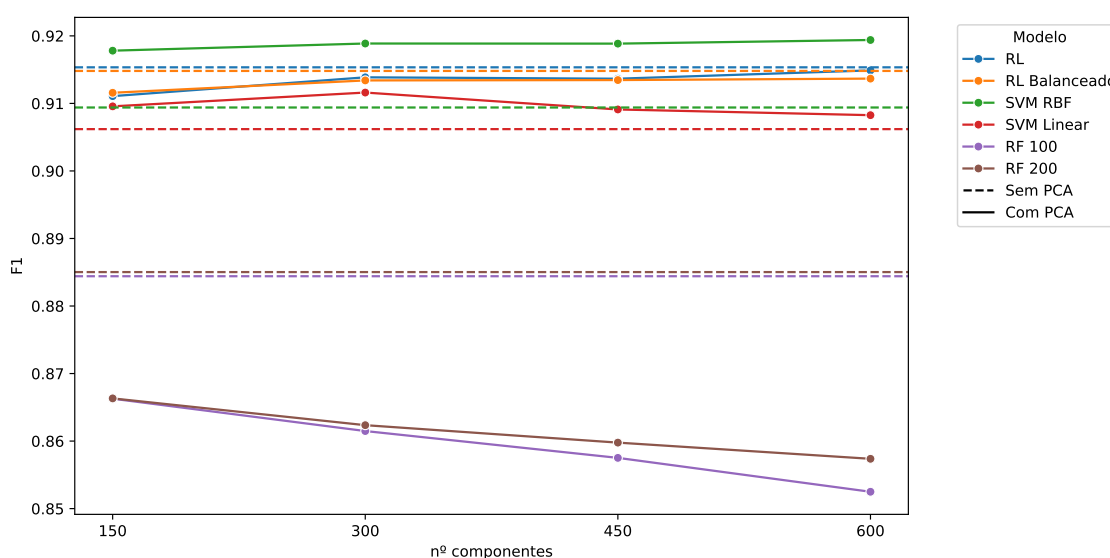
As duas configurações do modelo Random Forest, no entanto, apresentaram desempenho reduzido conforme o número de componentes aumentou, comportamento já observado nos experimentos para detecção de potencial depressão. Devido a este cenário, o classificador foi novamente testado com um número menor de componentes

A Figura 33 exibe as mudanças de valores para a métrica F1 durante a utilização de 50, 75, 100, 125, 150, 300, 450 e 600 componentes junto ao modelo Random Forest. Com 100 árvores, em seu melhor cenário com 50 componentes, o classificador registrou F1 score de 0,8619. Já para o uso de 200 árvores e o mesmo número de componentes, o valor foi de 0,8632, uma diferença de 0,0013

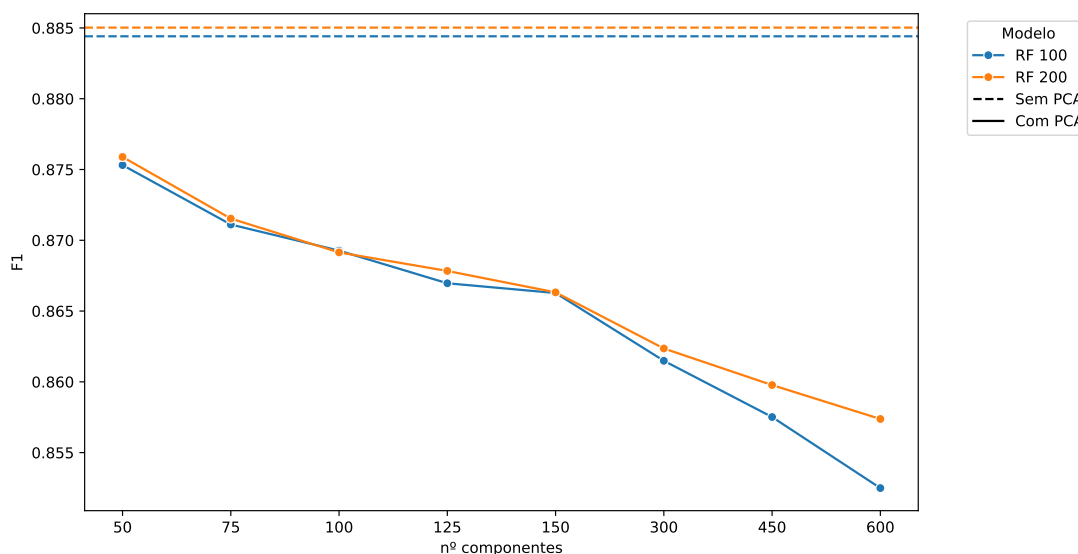
Dados não processados

Continuando os experimentos, agora com dados não processados, a Figura 34 apresenta os resultados obtidos para os modelos RL, SVM e RF. A redução de dimensionalidade manteve variância de 0,8789 para 150 componentes, 0,9452 para 300, 0,9754 para 450 e 0,9914 para 600 componentes.

Figura 34 – Suicidal Data não processado - PCA com modelos RL, SVM e RF



Fonte: Elaborado pela autora

Figura 35 – Suicidal Data não processado - PCA com modelo RF

Fonte: Elaborado pela autora

Assim como na versão pré-processada dos experimentos para o conjunto Suicidal Data, a Figura 34 também evidencia que o aumento no número de componentes permanece impactando de forma positiva o desempenho dos modelos, sendo 600 componentes o número do melhor cenário para os modelos de Regressão Logística e SVM RBF.

Para a configuração linear, diferente da versão RBF, SVM obteve os melhores resultados no cenário de 300 componentes, com F1 score de 0,9116. Com 450 e 600 componentes, o modelo apresentou queda em seus resultados, chegando ao seu pior desempenho no cenário de 600 componentes, com F1 score de 0,9082.

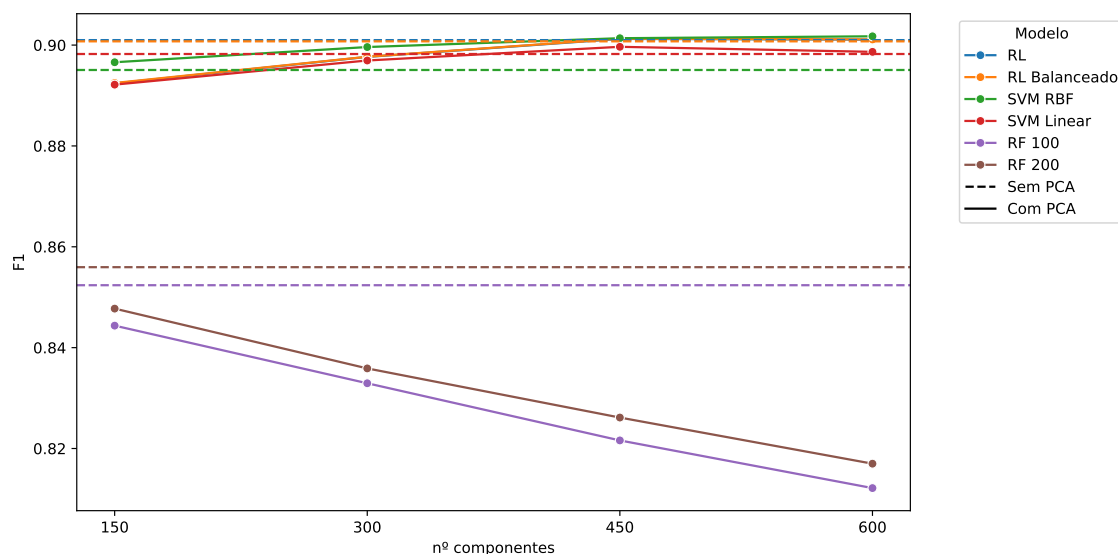
Para o modelo RF em suas duas configurações, a Figura 35 exibe os resultados de diferentes números de componentes utilizados para a redução da dimensionalidade do conjunto. O melhor desempenho para 100 árvores, assim como no conjunto pré-processado, foi obtido com 50 componentes, registrando 0,8753. Com 200 árvores o modelo atingiu 0,8758.

Suicide Detection

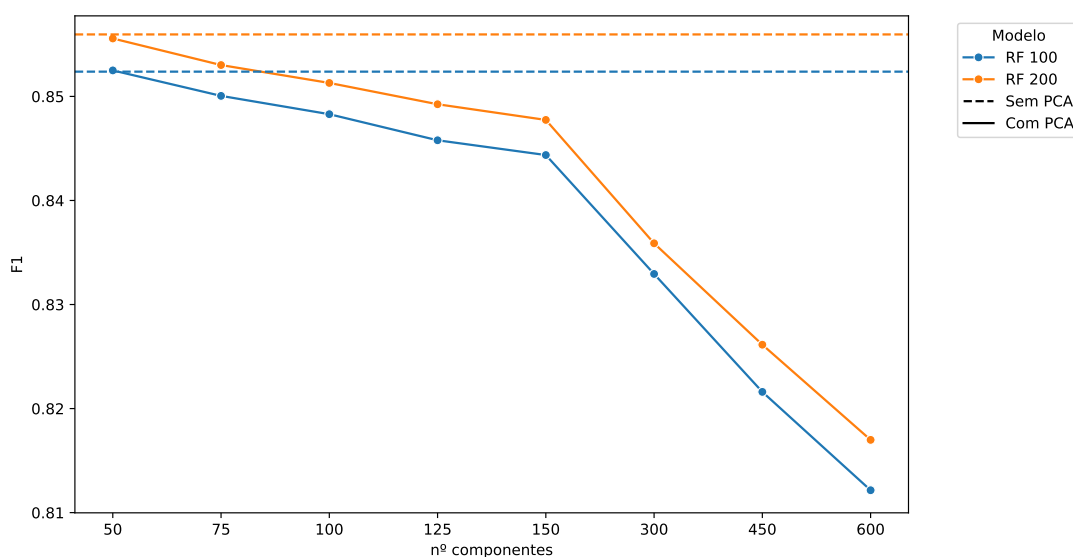
O segundo conjunto para detecção de potencial risco de suicídio, Suicide Detection, foi testado e avaliado após a aplicação da redução de dimensionalidade com o algoritmo PCA. Após 50 execuções de cada cenário, a média do desempenho de cada classificador foi obtida.

Dados pré-processados

Para os dados pré-processados, a Figura 36 exibe os resultados dos modelos LR, SVM e Random Forest. Após a aplicação do algoritmo PCA, foi possível calcular que a variância mantida foi de 0,8726 para 150 componentes, 0,9441 para 300, 0,9753 para 450 e 0,9913 para 600 componentes.

Figura 36 – Suicide Detection pré-processado - PCA com modelos RL, SVM e RF

Fonte: Elaborado pela autora

Figura 37 – Suicide Detection pré-processado - PCA com modelo RF

Fonte: Elaborado pela autora

Conforme mostra a Figura 36, os modelos de Regressão Logística e SVM Linear apresentaram crescimento nos valores de suas métricas à medida em que o número de componentes aumentou, atingindo melhores desempenhos com o uso de 450 componentes e pior desempenho quando 150 componentes principais foram utilizados. O modelo SVM RBF, no entanto, registrou o melhor desempenho com o uso de 600 componentes. À medida que o número de componentes aumentou, houve uma queda gradual nas métricas dos modelos Random Forest, sendo o pior desempenho obtido com o uso de 600 componentes.

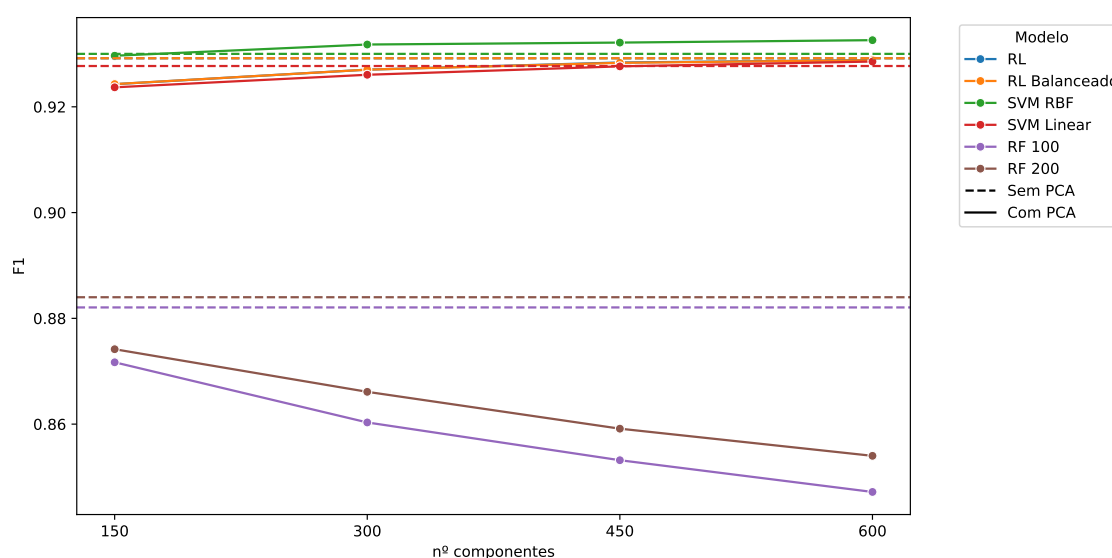
Aplicando 50, 75, 100 e 125 componentes exclusivamente para o modelo Random Forest, conforme mostra a Figura 37, foi possível obter resultados maiores do que 0,85 nos melhores cenários. Com apenas 50 componentes, o modelo com 100 árvores registrou 0,8524 para a métrica F1 score. Com 200 árvores, no entanto, foi possível obter 0,8555 utilizando o mesmo número de componentes.

Dados não processados

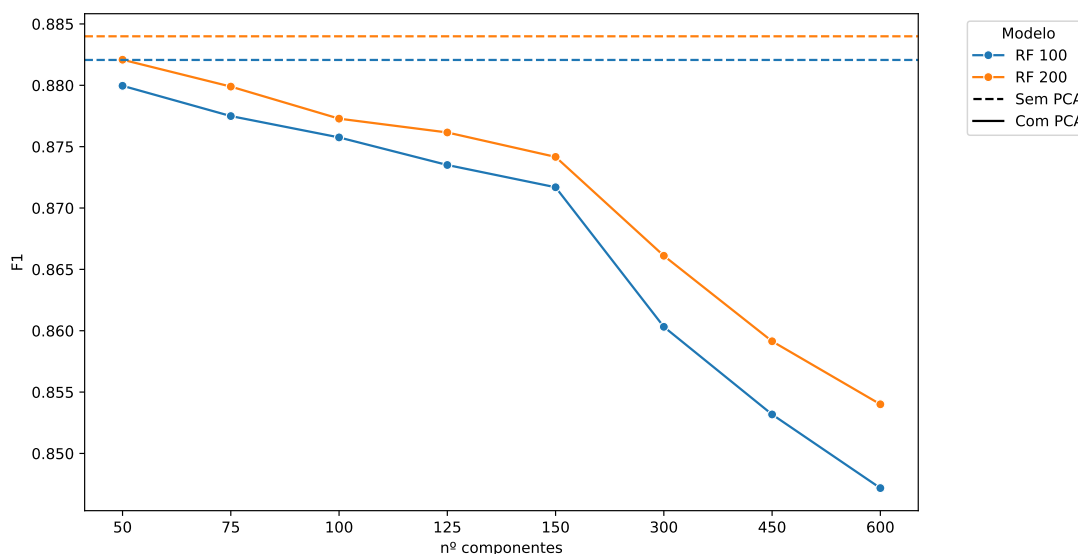
Finalizando os experimentos para a abordagem de redução de dimensionalidade, a Figura 38 apresenta os resultados obtidos durante a classificação do conjunto Suicide Detection não processado, com o objetivo de detectar potencial ideação suicida. Com a aplicação da redução de dimensionalidade, foi possível manter variância de 0,8650 para 150 componentes, 0,9375 para 300, 0,9714 para 450 e 0,9897 para 600 componentes.

O modelo de Regressão Logística, primeiro a ser testado, alcançou seu melhor desempenho com 600 componentes. Nesse cenário, LR obteve F1 score de 0,9288, mas com 150 componentes, seu pior cenário, o modelo atingiu F1 score de 0,9242.

Figura 38 – Suicide Detection não processado - PCA com modelos RL, SVM e RF



Fonte: Elaborado pela autora

Figura 39 – Suicide Detection não processado - PCA com modelo RF

Fonte: Elaborado pela autora

O comportamento, que já foi identificado nos experimentos de outros conjuntos, é o mesmo para os modelos LR balanceado, SVM RBF e SVM linear, todos tendo o melhor resultado registrado com a utilização de 600 componentes.

A Figura 38 ainda demonstra o modelo Random Forest apresentou queda em suas métricas conforme o número de componentes aumentou. A configuração de 150 componentes, para as duas versões do classificador, demonstrou ser a melhor. Com 100 árvores, o modelo obteve F1 score de 0,8716, enquanto o aumento para 200 árvores resultou em F1 score de 0,8741.

No entanto, assim como nos outros conjuntos testados, o algoritmo PCA foi configurado com 50, 75, 100 e 125 componentes para que o classificador fosse novamente testado, assim como evidencia a Figura 39. Com 50 componentes, bem como ocorreu com o conjunto pré-processado, RF 100 e RF 200 registraram os melhores valores para F1. Com 100 árvores, o valor foi de 0,8799, enquanto o uso de 200 árvores resultou em 0,8820.

De modo geral, em todos os conjuntos pré-processados e não processados, seja para auxílio na detecção de depressão ou suicídio, o modelo Random Forest apresentou melhores resultados com um menor número de componentes. Para os dois conjuntos de suicídio, 50 componentes permitiram que as maiores métricas fossem alcançadas. Para LR e SVM, no entanto, 450 e 600 componentes são suficientes para resultados significativos.

7.1.3 Classificação com agrupamento de modelos

Ensemble Methods, que constitui a aprendizagem por conjunto, é uma abordagem escolhida para ser testada nos conjuntos desta pesquisa. Os modelos LR, SVM

e RF foram testados individualmente e tiveram suas métricas calculadas nos cenários com pré-processamento, sem pré-processamento e com redução de dimensionalidade. No entanto, buscando identificar a melhor abordagem para o auxílio na detecção de potencial depressão e risco de suicídio, esta pesquisa também busca avaliar o desempenho desses classificadores em conjunto.

O método de empilhamento, em inglês, *Stacking*, foi o método escolhido para os experimentos realizados. A implementação da técnica foi realizada com o auxílio da biblioteca scikit-learn, utilizando o classificador Stacking Classifier (SC). Para os experimentos, os modelos RF com 100 árvores, SVM RBF e SVM Linear foram utilizados como modelos base, e suas previsões serviram para treinar e ensinar o metamodelo LR, uma vez que se apresentou como melhor classificador em alguns resultados anteriores.

Mental Health Twitter

A Tabela 17 exibe os resultados obtidos pela técnica SC em dois cenários: com dados pré-processados (PP) e dados não processados (NP). Para comparação, o modelo LR testado individualmente em cada cenário também foi adicionado a Tabela.

Tabela 17 – Resultados da combinação de modelos para o conjunto Mental Health Twitter

Cenário	Técnica	Acurácia	Precisão	Sensibilidade	F1
PP	SC	0,7007	0,6894	0,7300	0,7091
PP	LR	0,6809	0,6762	0,6935	0,6847
NP	SC	0,8165	0,8110	0,8251	0,8180
NP	LR	0,8056	0,8087	0,8004	0,8045

Fonte: Elaborado pela autora

As métricas obtidas no cenário PP demonstram que, para esse conjunto, a técnica SC apresenta melhor desempenho na classificação dos dados. Enquanto LR obteve acurácia de 0,6809, o empilhamento dos modelos permitiu atingir 0,7007. Para o cenário NP, o comportamento se repetiu, tendo a técnica SC com o melhor desempenho em comparação ao modelo LR testado individualmente, com uma diferença de 0,0109 para acurácia e 0,0135 para F1 score.

De modo geral, considerando todos os experimentos já realizados, a utilização da técnica SC demonstrou ser uma escolha viável para a melhora significativa no desempenho das classificações do conjunto MHT.

Suicidal Data

Para o conjunto Suicidal Data, a Tabela 18 exibe as métricas calculadas. Diferente do observado no conjunto MHT, algumas métricas são melhores na técnica SC e outras são melhores na técnica LR.

No cenário PP, LR obteve sensibilidade e F1 score maiores do que a técnica

Tabela 18 – Resultados da combinação de modelos para o conjunto Suicidal Data

Cenário	Técnica	Acurácia	Precisão	Sensibilidade	F1
PP	SC	0,9052	0,9275	0,8510	0,8875
PP	LR	0,9042	0,9125	0,8650	0,8880
NP	SC	0,9247	0,9365	0,8890	0,9121
NP	LR	0,9267	0,9292	0,9019	0,9153

Fonte: Elaborado pela autora

SC, que atingiu acurácia e precisão mais altos. Para o cenário NP, no entanto, apenas a métrica precisão foi superior quando a técnica SC foi aplicada.

Os resultados demonstram que, neste conjunto para detecção de suicídio, o empilhamento dos modelos escolhidos e com as configurações escolhidas, não trouxeram melhorias significativas que justifiquem o aumento do custo computacional.

Suicide Detection

Finalizando os experimentos, a Tabela 19 exibe os resultados para o conjunto Suicide Detection e evidencia que, para estes dados, a técnica SC não superou os resultados obtidos com a classificação individual do modelo LR.

Tabela 19 – Resultados da combinação de modelos para o conjunto Suicide Detection

Cenário	Técnica	Acurácia	Precisão	Sensibilidade	F1
PP	SC	0,8983	0,8963	0,9010	0,8986
PP	LR	0,9005	0,8971	0,9048	0,9009
NP	SC	0,9288	0,9298	0,9278	0,9288
NP	LR	0,9291	0,9295	0,9287	0,9291

Fonte: Elaborado pela autora

No cenário PP, o modelo LR apresentou resultado superior em todas as métricas calculadas. A técnica SC obteve, por exemplo, acurácia de 0,8983, enquanto LR obteve 0,9005.

Com a ausência de pré-processamento, a técnica SC também não foi capaz de superar os resultados já obtidos para este conjunto. Com acurácia e F1 score de 0,9288, ela apresentou desempenho inferior a LR, que registrou 0,9291 para acurácia e F1 score, um pequeno aumento de 0,0003.

7.2 Classificação Semi-Supervisionada

Para avaliar o uso da aprendizagem semi-supervisionada neste domínio de pesquisa, os experimentos foram realizados com duas variações no percentual de dados mascarados: 80% e 60%. O método *holdout* foi executado 50 vezes, com 80% dos dados para treino e 20% dos dados para teste.

Enquanto as classificações supervisionadas utilizaram os modelos de Regressão Logística, SVM e Random Forest, as classificações semi-supervisionadas foram realizadas com a execução dos modelos baseados em grafos: Label Propagation (LP), Label Spreading (LS) Competição e Cooperação entre Partículas (CCP).

Para a realização dos testes, o algoritmo LP foi configurado com `kernel = 'knn'`, `n_neighbors = 10`, `max_iter = 100000` e `tol = 0.001`, assim como o modelo LS, mantendo o restante dos parâmetros com os valores padrões do `sklearn`. O modelo CCP utilizou o mesmo número de vizinhos em suas configurações, atribuindo `k = 10`, `disttype = 'euclidean'` e `mex = false`.

Posteriormente, para testar a técnica de Autotreinamento, o modelo Self Training Classifier foi aplicado, sendo configurado com `criterion = 'threshold'`, `max_iter = 10000` e variando os valores para o parâmetro *threshold*. O modelo Self Training Classifier, que tem o objetivo de fazer um classificador supervisionado funcionar como um classificador semi-supervisionado, utilizou o modelo de Regressão Logística (STLR) para as classificações. A escolha dos modelos se deu pelo fato terem registrado os melhores desempenhos na classificação supervisionada.

7.2.1 Classificação com dados pré-processados e não processados

Esta sub-seção exibe os resultados registrados durante os experimentos de classificação de dados pré-processados (PP) e não processados (NP), sem a utilização de redução de dimensionalidade. As Tabelas exibem os resultados obtidos pela melhor classificação de cada modelo: LP, LS e CCP.

Mental Health Twitter

Os resultados obtidos na classificação do conjunto Mental Health Twitter não foi comparado com o trabalho relacionado de Abbas et al. (2024), como ocorreu com os resultados supervisionados. Na Tabela 20, as métricas acurácia, precisão, sensibilidade e F1 score são exibidas para as classificações com dados pré-processados e não processados.

Os modelos LP e LS foram testados em duas configurações diferentes, alternando entre os valores KNN e RBF para o parâmetro *kernel*. Enquanto isso, o modelo CCP foi testado com uma única configuração. Para a discussões dos resultados, as melhores configurações são exibidas na Tabela, uma vez que os resultados completos podem ser visualizados no Apêndice da monografia.

Conforme mostra a Tabela 20, os melhores resultados foram obtidos com o mascaramento de 60% do conjunto de dados, mantendo 40% dos exemplos com seus rótulos originais. Isso ocorre porque a quantidade de rótulos conhecidos e confiáveis influencia diretamente a propagação das informações no grafo e, consequentemente, a atribuição de classes aos exemplos não rotulados.

Neste cenário, e com os dados pré-processados, o modelo LP com o uso do

Tabela 20 – Melhores resultados obtidos para o conjunto Mental Health Twitter com aprendizado semi-supervisionado

Técnica	Cenário	Masc.	Acurácia	Precisão	Sensibilidade	F1
CCP	NP	60%	0,7838	0,7863	0,7795	0,7828
LS KNN	NP	60%	0,7661	0,7965	0,7148	0,7532
LP KNN	NP	60%	0,7565	0,8046	0,6776	0,7354
CCP	PP	60%	0,6600	0,6567	0,6706	0,6635
LS KNN	PP	60%	0,6461	0,6429	0,6579	0,6499
LP KNN	PP	60%	0,6443	0,6417	0,6552	0,6475
CCP	NP	80%	0,7508	0,7534	0,7465	0,7497
LS KNN	NP	80%	0,7442	0,7767	0,6858	0,7280
LP KNN	NP	80%	0,7217	0,8021	0,5900	0,6787
CCP	PP	80%	0,6399	0,6410	0,6367	0,6387
LS KNN	PP	80%	0,6312	0,6318	0,6290	0,6299
LP KNN	PP	80%	0,6281	0,6310	0,6229	0,6237

Fonte: Elaborado pela autora

kernel KNN permitiu atingir acurácia de 0,6443 e F1 score de 0,6475. Para o modelo LS, utilizando o mesmo kernel, a métrica F1 registrada foi de 0,6499. O modelo CCP, testado com apenas uma configuração, apresentou o melhor desempenho entre todos os classificadores.

Com a aplicação do algoritmo de Competição e Cooperação entre Partículas, a acurácia registrada foi de 0,6600, com F1 score de 0,6635. Em comparação ao modelo LS, por exemplo, a diferença foi de 0,0136 para a métrica F1. Com 80% de dados mascarados, no entanto, todos os modelos registraram queda significativa de desempenho. O modelo CCP diminuiu 0,0248 do valor obtido para a métrica F1 score com a utilização de 60% dos dados mascarados.

Removendo a etapa de pré-processamento, as métricas das classificações se mostraram superiores às obtidas no cenário PP. Por meio da Tabela 20, nota-se que o modelo CCP se manteve como o melhor classificador para a detecção de potencial depressão, quando comparado com os outros dois modelos baseados em grafos.

No cenário NP, o algoritmo de Competição e Cooperação entre Partículas atingiu acurácia de 0,7838 e F1 score de 0,7828 quando 60% dos dados foram mascarados. Os modelos LP e LS, no entanto, obtiveram 0,7354 e 0,7532 para a métrica F1 score, respectivamente.

De modo geral, a classificação semi-supervisionada do conjunto MHT demonstrou ser possível utilizar modelos baseados em grafos para o domínio da pesquisa. Além disso, reforça que a ausência de pré-processamento contribui para a obtenção de métricas mais elevadas. Como era esperado, quanto maior o número de dados não

maskarados, melhor o desempenho dos modelos testados. No entanto, mesmo com apenas 20% dos rótulos originais disponíveis durante os experimentos, os resultados são promissores.

Suicidal Data

Os experimentos realizados com o conjunto Suicidal Data também não foram comparados com o trabalho de Yatapala e Kumara (2021), uma vez que não utiliza os mesmos modelos e tipo de aprendizado de máquina. A Tabela 21 apresenta os resultados obtidos com a classificação semi-supervisionada para os modelos LP KNN, LS KNN e CCP nos cenários PP e NP.

Tabela 21 – Melhores resultados obtidos para o conjunto Suicidal Data com aprendizado semi-supervisionado

Técnica	Cenário	Masc.	Acurácia	Precisão	Sensibilidade	F1
CCP	NP	60%	0,9030	0,9199	0,8533	0,8853
LS KNN	NP	60%	0,8961	0,8903	0,8710	0,8805
LP KNN	NP	60%	0,8951	0,8901	0,8688	0,8792
CCP	PP	60%	0,8839	0,9017	0,8257	0,8620
LS KNN	PP	60%	0,8693	0,8481	0,8560	0,8519
LP KNN	PP	60%	0,8669	0,8440	0,8554	0,8495
CCP	NP	80%	0,8980	0,9177	0,8432	0,8788
LP KNN	NP	80%	0,8861	0,8761	0,8636	0,8694
LS KNN	NP	80%	0,8893	0,8791	0,8678	0,8732
CCP	PP	80%	0,8786	0,8986	0,8154	0,8549
LS KNN	PP	80%	0,8589	0,8286	0,8567	0,8422
LP KNN	PP	80%	0,8505	0,8114	0,8608	0,8350

Fonte: Elaborado pela autora

Da mesma forma como ocorreu com o conjunto Mental Health Twitter, a utilização de 60% dos dados mascarados foi superior às avaliações obtidas com 80% de exemplos sem o rótulo original. O melhor modelo para a classificação de potencial suicídio, assim como para depressão, foi o modelo de Competição e Cooperação entre partículas.

Iniciando a análise pelo experimento com dados pré-processados e avaliando o seu melhor cenário, com 60% dos dados mascarados, é possível identificar que modelo LP KNN obteve desempenho inferior aos outros classificadores testados. Enquanto o uso do classificador LS permitiu F1 score de 0,8519 e a aplicação do algoritmo CCP registrou 0,8620 para a mesma métrica, LP KNN alcançou 0,8495.

Ao realizar novas execuções, agora com a ausência de pré-processamento, os resultados aumentaram de forma significativa, em todos os modelos e cenários de mascaramento. Com 60% dos dados mascarados, o modelo LP KNN, que antes

havia obtido F1 score de 0,8495, registrou 0,8792 no cenário NP, um aumento de 0,0297. Mesmo com a melhora do desempenho, seus resultados foram superados pelos modelos LS e CCP. Enquanto o modelo LS registrou F1 score de 0,8805, o modelo CCP obteve 0,8853, mantendo a posição de melhor classificador para este conjunto.

A classificação semi-supervisionada do conjunto Suicidal Data também apresentou resultados promissores, com métricas mais elevadas do que as obtidas no conjunto de depressão, repetindo o padrão observado na classificação supervisionada.

O algoritmo de Competição e Cooperação entre Partículas foi classificado como o melhor em ambos os conjuntos, evidenciando que este pode ser visto como uma alternativa para a aplicação de métodos baseados em grafos em problemas de classificação de textos.

Suicide Detection

Para o conjunto Suicide Detection, testado no cenário pré-processado e não processado, a Tabela 22 exhibe os resultados dos experimentos realizados com 80% e 60% dos dados mascarados.

Tabela 22 – Melhores resultados obtidos para o conjunto Suicide Detection com aprendizado semi-supervisionado

Técnica	Cenário	Masc.	Acurácia	Precisão	Sensibilidade	F1
CCP	NP	60%	0,8603	0,8502	0,8749	0,8623
LS KNN	NP	60%	0,8343	0,7840	0,9233	0,8479
LP KNN	NP	60%	0,8245	0,7714	0,9229	0,8403
CCP	PP	60%	0,8256	0,8121	0,8474	0,8294
LS KNN	PP	60%	0,7758	0,7099	0,9334	0,8064
LP KNN	PP	60%	0,7539	0,6846	0,9426	0,7931
CCP	NP	80%	0,8505	0,8413	0,8642	0,8525
LS KNN	NP	80%	0,8212	0,7678	0,9214	0,8376
LP KNN	NP	80%	0,7970	0,7382	0,9211	0,8195
CCP	PP	80%	0,8143	0,8024	0,8339	0,8178
LS KNN	PP	80%	0,7554	0,6879	0,9358	0,7929
LP KNN	PP	80%	0,6837	0,6186	0,9604	0,7524

Fonte: Elaborado pela autora

Mantendo o padrão dos outros experimentos realizados, o melhor cenário permaneceu sendo a configuração com 60% de dados mascarados e 40% de dados com a preservação dos rótulos originais.

Iniciando pelo cenário PP, o modelo LP KNN apresentou acurácia de 0,7539 e F1 score de 0,7931, os menores valores entre os três modelos testados na pesquisa. A aplicação do modelo LS no conjunto de detecção de suicídio obteve, para a métrica

acurácia, 0,7758, enquanto o modelo CCP apresentou 0,8256. A diferença entre a acurácia da classificação com LP e CCP foi de 0,0717.

Em uma análise geral, os modelos LS e LP apresentam uma diferença significativa entre as métricas de precisão e sensibilidade, justificando o valor da métrica F1 score. LS KNN com 80% dos dados mascarados, por exemplo, obteve F1 score de 0,7929, mas uma precisão de 0,6879 e sensibilidade de 0,9358, indicando uma quantidade significativa de falsos positivos.

Para o conjunto Suicide Detection com a ausência de pré-processamento, os modelos registraram desempenho superior ao cenário testado anteriormente. A técnica LP com kernel KNN obteve acurácia de 0,8245, precisão de 0,7714, sensibilidade de 0,9229 e F1 score de 0,8403. Para o modelo LS, a acurácia foi de 0,8343 e F1 score de 0,8479.

A classificação com o modelo de Competição e Cooperação entre Partículas, assim como no cenário PP, registrou as melhores métricas entre todos os modelos. Com 60% dos dados mascarados, a acurácia foi de 0,8603, precisão de 0,8502, sensibilidade de 0,8749 e F1 score de 0,8623. A diferença entre os valores da métrica F1 do algoritmo CCP e o algoritmo LP foi de 0,022.

As classificações do conjunto Suicide Detection concluem os experimentos que buscaram avaliar o impacto da utilização e da ausência do pré-processamento para a detecção de potencial depressão e risco de suicídio. Reforçando o que foi observado nos outros dois conjuntos, remover as diferentes técnicas aplicadas no pré-processamento dos textos resultou em um melhor desempenho dos modelos semi-supervisionados, assim como foi observado com os modelos supervisionados.

Além disso, este experimento também elenca o modelo CCP como o melhor método baseado em grafos para esta pesquisa, quando os modelos LP e LS também são levados em consideração.

7.2.2 Classificação com redução de dimensionalidade

Os experimentos com redução de dimensionalidade foram realizados apenas nas configurações que apresentaram os melhores resultados nos cenários anteriores: LP KNN, LS KNN e CCP. Todos os conjuntos foram testados com 60% e 80% dos dados mascarados, com exemplos pré-processados e, posteriormente, com a ausência de pré-processamento. Para esta abordagem, o algoritmo PCA foi configurado com 10, 20, 30, 40, 50, 100 e 150 componentes.

Mental Health Twitter

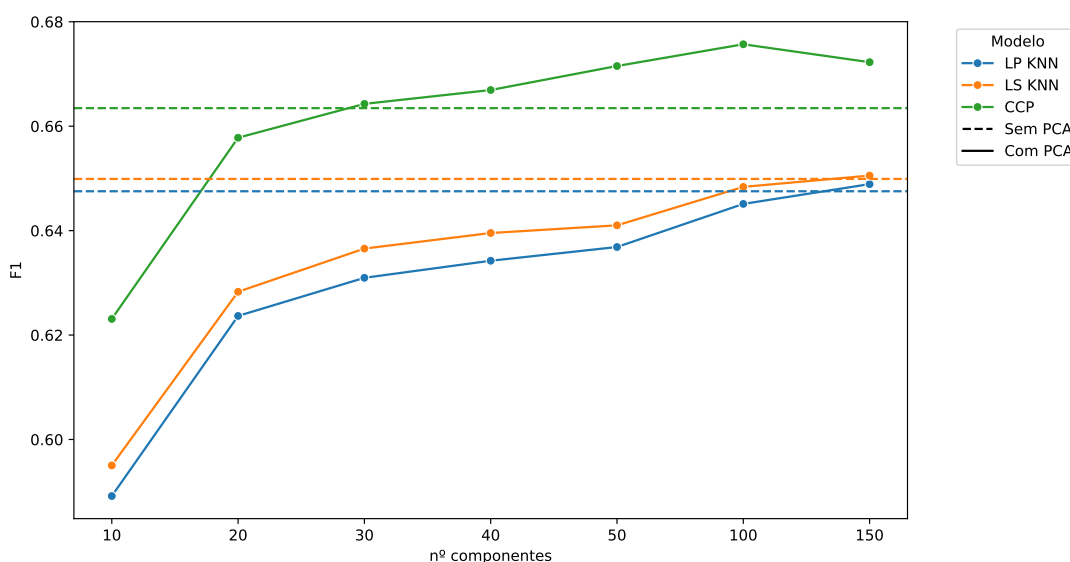
A Figura 40 exibe a variação nos valores de F1 score na classificação com 60% dos dados pré-processados e mascarados, à medida em que o número de componentes aumentava. O algoritmo CCP iniciou os experimentos com 10 componentes, registrando F1 score de 0,6230, que aumentou e diminuiu ao longo dos testes. Com

100 componentes, o resultado foi de 0,6756. Para LP o melhor resultado foi de 0,6489, com 150 componentes, o mesmo cenário de LS, que registrou 0,6505.

Para o uso de 80% dos dados, ainda no cenário pré-processado, a Figura 41 mostra que o desempenho do modelo CCP reduziu. Com 10 componentes o modelo obteve 0,61300, evoluindo até atingir a máxima de 0,6522, também com 100 componentes utilizados. Para os modelos LP e LS, o melhor resultado foi alcançado com 150 componentes, registrando F1 score de 0,6307 e 0,6334, respectivamente.

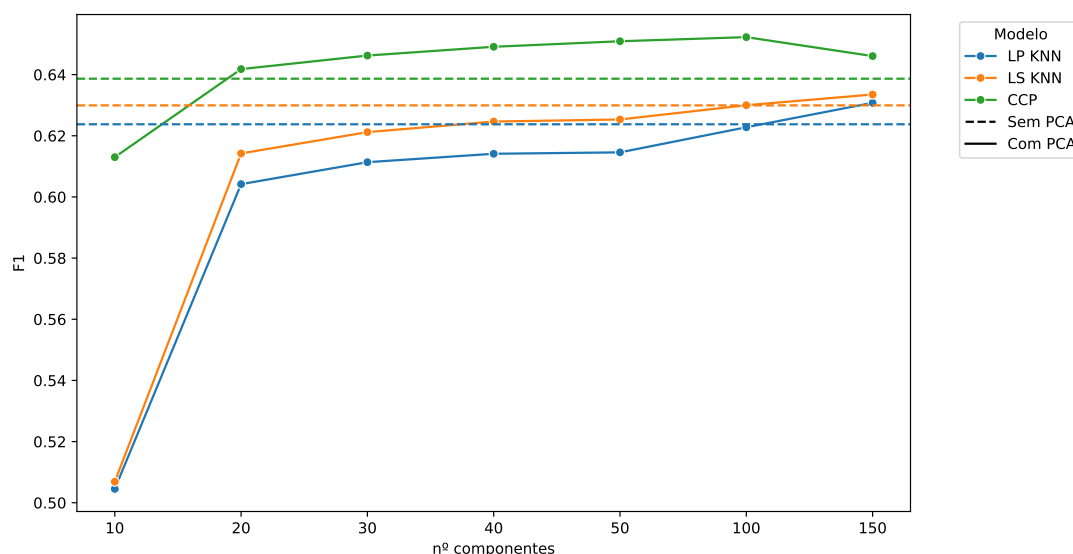
Comparando a métrica F1 score obtida sem a aplicação do algoritmo PCA, a classificação com o modelo de Competição e Cooperação entre Partículas teve um aumento de 0,0135 e 0,0121 no uso de 60% e 80% de dados mascarados, respectivamente, ao considerar o melhor cenário de 100 componentes.

Figura 40 – Mental Health Twitter pré-processado - PCA com 60% de dados mascarados



Fonte: Elaborado pela autora

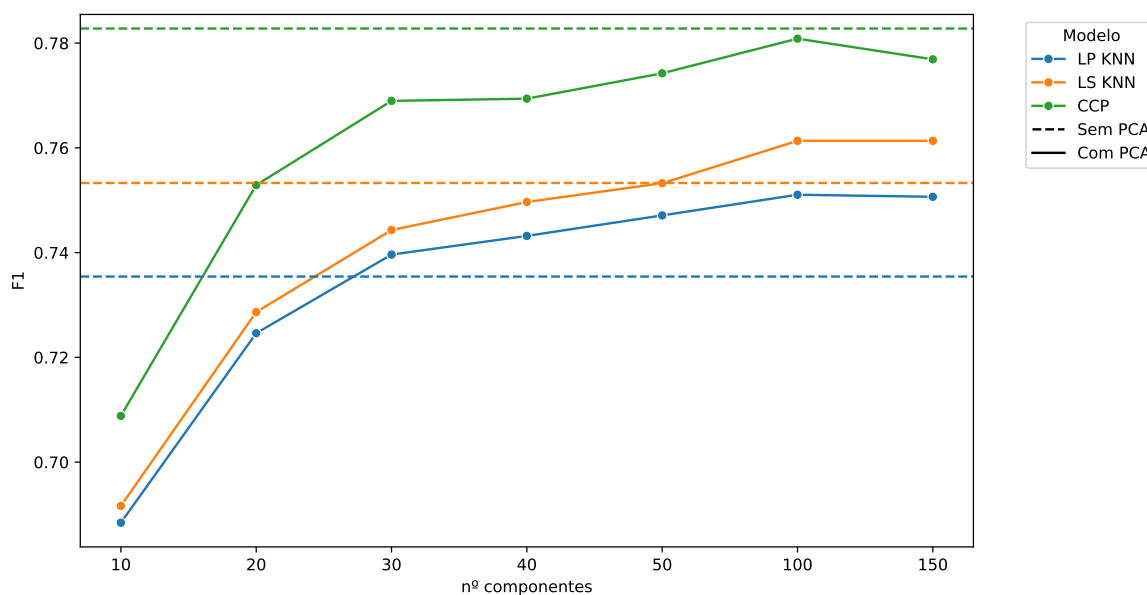
Figura 41 – Mental Health Twitter pré-processado - PCA com 80% de dados mascarados



Fonte: Elaborado pela autora

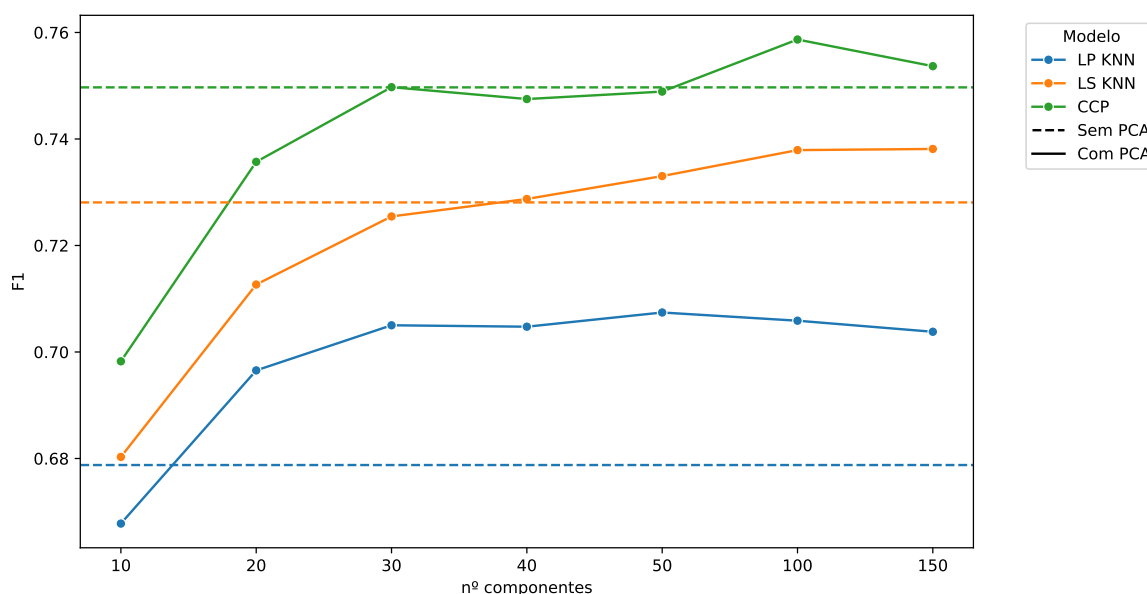
Seguindo para os experimentos com dados não processados, conforme mostra a Figura 42, o melhor cenário também se deu com o uso de 60% dos dados mascarados. Nesta configuração, o modelo CCP manteve 100 componentes como o número que permitiu seu melhor desempenho, registrando F1 score de 0,7808. Para LP, no entanto, 100 componentes resultaram em 0,7510, diferente do modelo LS que atingiu 0,7613 com 150 componentes.

Figura 42 – Mental Health Twitter não processado - PCA com 60% de dados mascarados



Fonte: Elaborado pela autora

Figura 43 – Mental Health Twitter não processado - PCA com 80% de dados mascarados



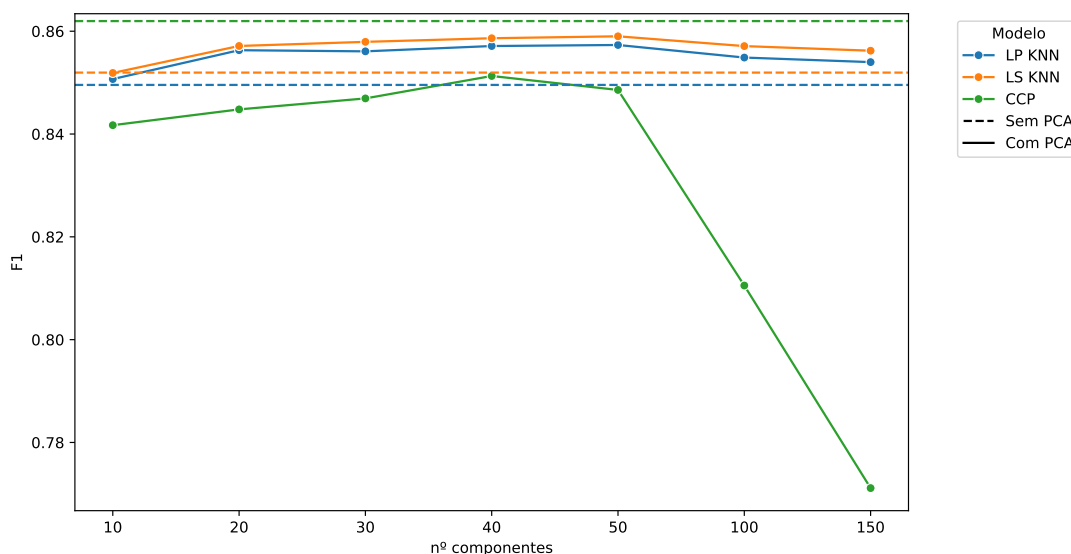
Fonte: Elaborado pela autora

Por fim, a Figura 43 exibe a diminuição das métricas ao utilizar 80% dos dados mascarados. Para o modelo CCP com 100 componentes utilizados, a F1 score foi de 0,7586, uma diferença de 0,0222. Com o uso do modelo LP, a melhor métrica, de 0,7074, foi registrada com 50 componentes, enquanto o modelo LS obteve 0,7381 com 150 componentes.

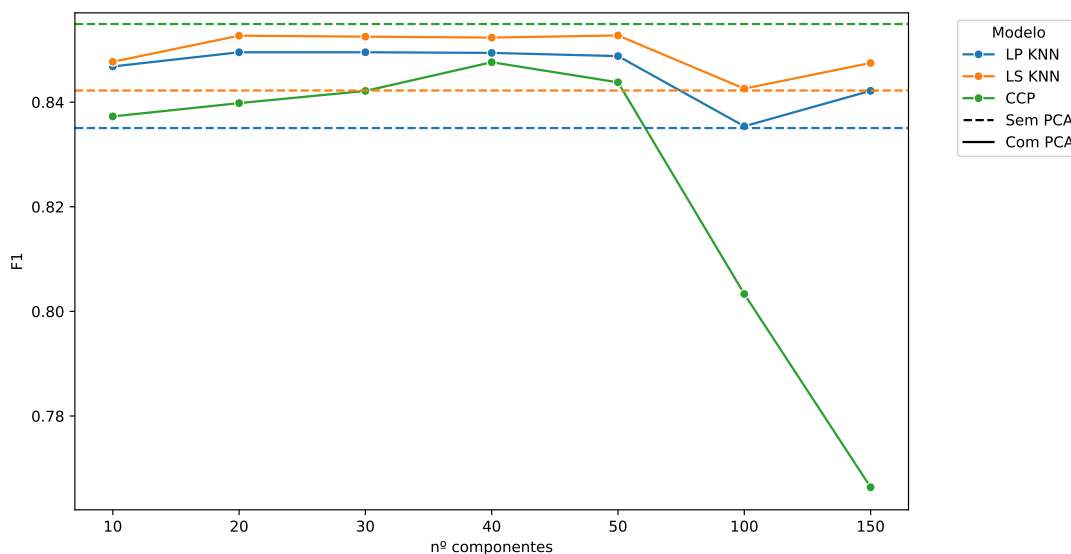
Suicidal Data

Para o conjunto Suicidal Data, a Figura 44 exibe os resultados obtidos com 60% dos dados rotulados e pré-processados. Para este conjunto, diferente do MHT, o modelo CCP não superou os resultados obtidos pelos dois outros modelos baseados em grafos. O valor máximo que o modelo CCP atingiu, com 40 componentes, foi de 0,8512, registrando queda a partir disso. Já os modelos LP e LS obtiveram, respectivamente, 0,8573 e 0,8590 com 50 componentes utilizados.

Com 80% dos dados, conforme demonstra a Figura 45, os valores máximos atingidos pelos modelos foram inferiores aos resultados obtidos no cenário anterior. Com 30 componentes, o classificador LP registrou F1 score de 0,8495, enquanto LS obteve 0,8527 com 50 componentes e o modelo CCP, com 40 componentes, alcançou 0,8476.

Figura 44 – Suicidal Data pré-processado - PCA com 60% de dados mascarados

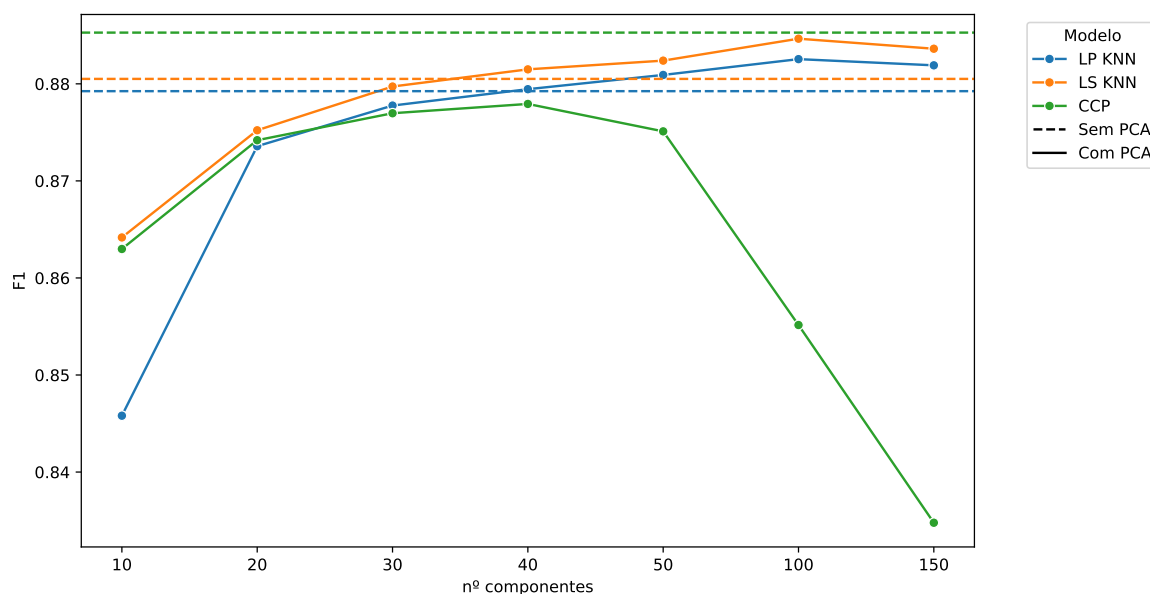
Fonte: Elaborado pela autora

Figura 45 – Suicidal Data pré-processado - PCA com 80% de dados mascarados

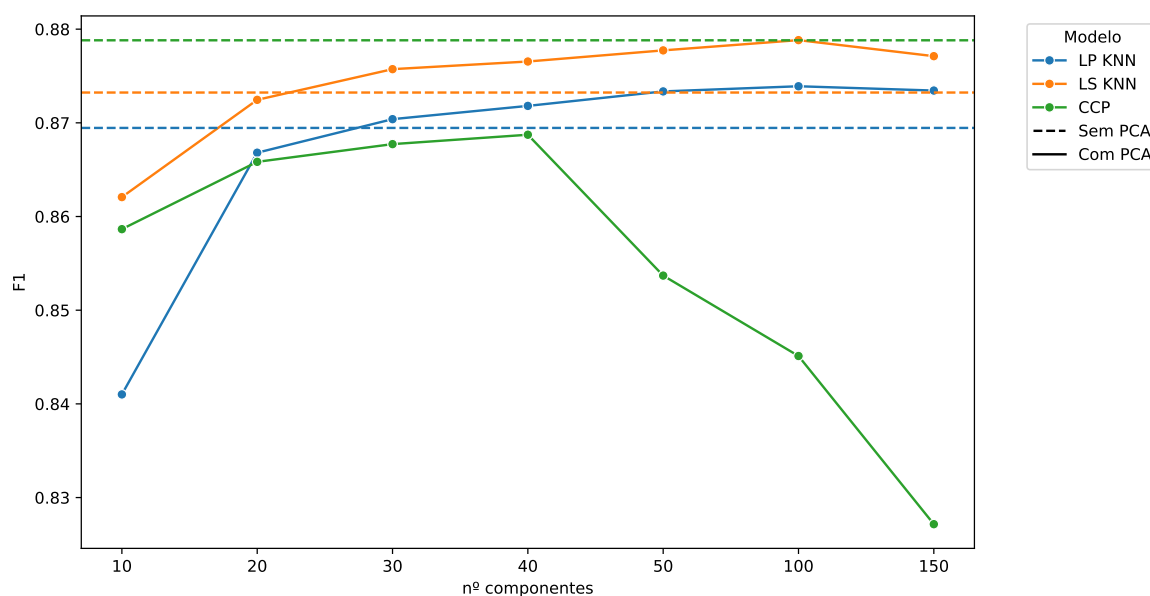
Fonte: Elaborado pela autora

Para a classificação com dados não processados, o comportamento se repetiu. O modelo CCP, mesmo atingindo resultado superior ao algoritmo LP em algumas configurações do algoritmo PCA, ainda se manteve inferior aos melhores resultados registrados pelos outros dois algoritmos baseados em grafos.

Com 60% dos dados mascarados, o melhor cenário, e com o uso de 100 componentes, os modelos LP e LS obtiveram, respectivamente, 0,8825 e 0,8846 para a métrica F1. O modelo CCP, no entanto, atingiu valor máximo de 0,8779, conforme exibe a Figura 46.

Figura 46 – Suicidal Data não processado - PCA com 60% de dados mascarados

Fonte: Elaborado pela autora

Figura 47 – Suicidal Data não processado - PCA com 80% de dados mascarados

Fonte: Elaborado pela autora

Além disso, com base na Figura 47, é possível perceber que o mascaramento de 80% dos dados não processados fez com que os valores diminuíssem, assim como ocorreu no cenário de dados pré-processados. O maior valor para a métrica F1, de 0,8788, foi obtido pelo modelo LS com o uso de 100 componentes.

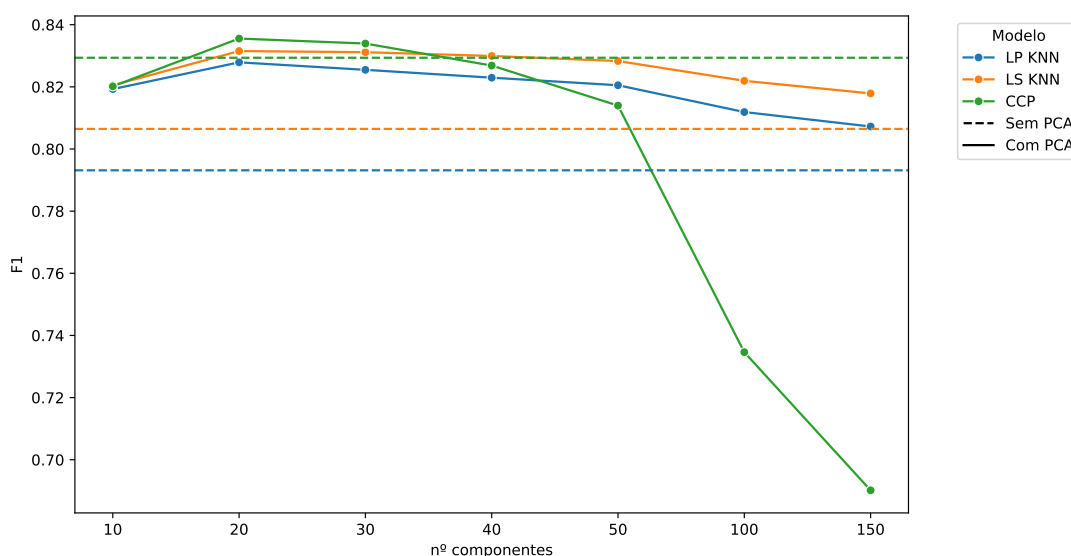
Suicide Detection

Finalizando os experimentos com redução de dimensionalidade, as Figuras 48 e 49 exibem os resultados obtidos na classificação de dados processados com 60%

e 80% de rótulos mascarados, respectivamente.

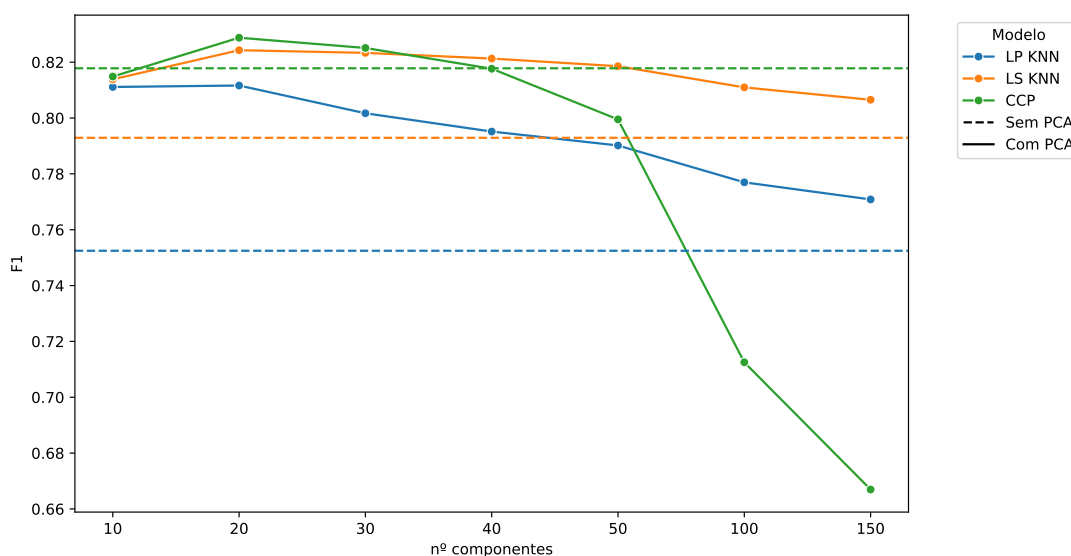
Diferente do observado no conjunto Suicidal Data, também para detecção de potencial risco de suicídio, o modelo CCP foi capaz de obter F1 score superior às máximas atingidas pelos modelos LP e LS. Com apenas 20 componentes, o algoritmo de Competição e Cooperação entre Partículas registrou 0,8355 com o uso de 60% de dados mascarados e 0,8287 com 80% de rótulos mascarados. Seu pior desempenho, no entanto, ocorreu com o uso de 150 componentes.

Figura 48 – Suicide Detection pré-processado - PCA com 60% de dados mascarados



Fonte: Elaborado pela autora

Figura 49 – Suicide Detection pré-processado - PCA com 80% de dados mascarados

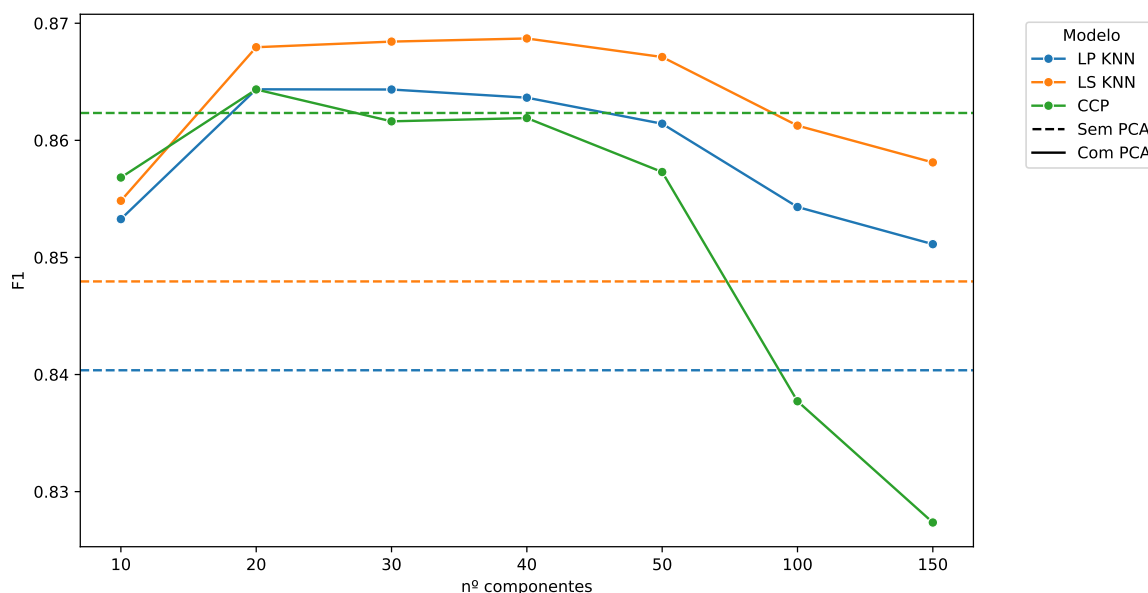


Fonte: Elaborado pela autora

Para o conjunto não processado, o modelo CCP permaneceu obtendo me-

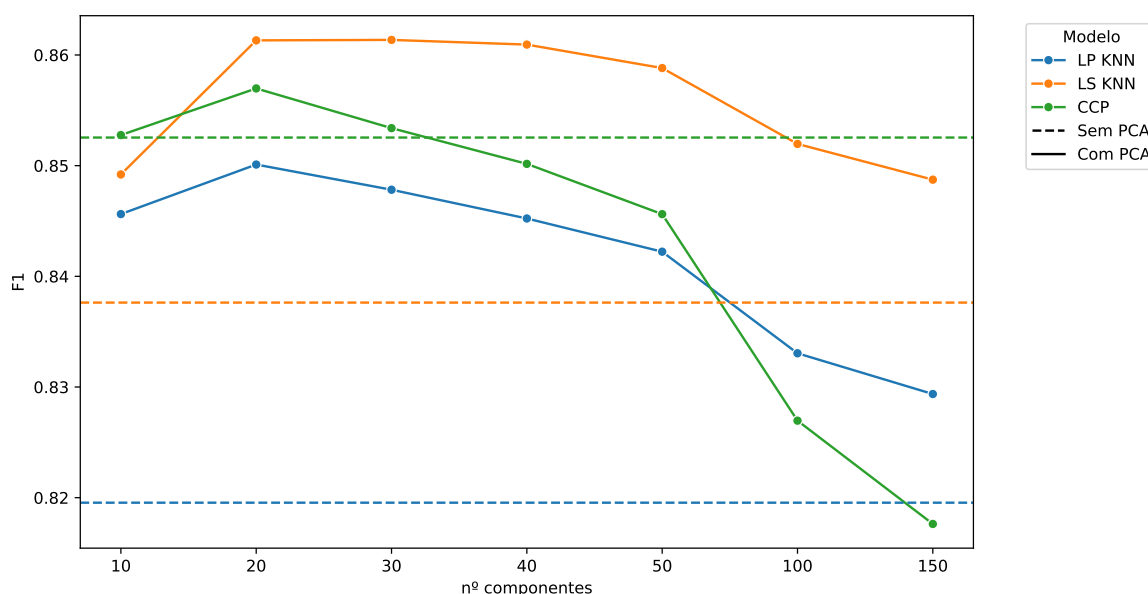
lhor desempenho com 20 componentes, mas não foi capaz de alcançar os melhores resultados em comparação aos outros modelos testados. Conforme mostra a Figura 50, o classificador LS apresentou os maiores valores para a métrica F1, obtendo a máxima de 0,8687 com 30 componentes e 60% dos dados mascarados.

Figura 50 – Suicide Detection não processado - PCA com 60% de dados mascarados



Fonte: Elaborado pela autora

Figura 51 – Suicide Detection não processado - PCA com 80% de dados mascarados



Fonte: Elaborado pela autora

Além disso, de acordo com a Figura 51, com 80% dos dados mascarados, o modelo LS também registrou o melhor desempenho entre todos os outros. Com 30

componentes, obteve 0,8613, enquanto LP alcançou 0,8501 com 20 componentes. O modelo CCP, que superou a máxima do algoritmo LP, obteve 0,8569 com o mesmo número de componentes utilizados.

7.2.3 Classificação com Self Training

O classificador Self Training (ST) foi aplicado para exploração da abordagem de autotreinamento. O algoritmo foi testado e configurado para que o critério *threshold* fosse utilizado na tomada de decisão das classificações confiáveis. Esta técnica foi comparada com o modelo LR padrão em uma classificação supervisionada.

Para o modelo ST, que utilizou LR como modelos base nas classificações, foram definidos 0,5 e 0,75 como *threshold* e 10 mil iterações máximas para o autotreinamento. O modelo base foi configurado com mil iterações máximas, considerando o custo computacional das execuções. A execução individual do modelo LR supervisionado também foi configurado com mil iterações máximas.

Uma vez que as classificações semi-supervisionadas utilizaram 80 e 60 como percentual para o mascaramento dos rótulos, os modelos supervisionados foram testados com 80% e 60% dos dados do conjunto, sendo o restante utilizado para o treinamento. Para a avaliação das classificações, as métricas de acurácia, precisão, sensibilidade e F1 score foram calculadas considerando 0,5 e 0,75 como valor mínimo para que uma predição fosse considerada confiável.

Todos os resultados podem ser visualizados nas tabelas adicionadas ao Apêndice deste trabalho, enquanto esta sub-seção apresenta as métricas F1 score e acurácia através de gráficos.

Mental Health Twitter

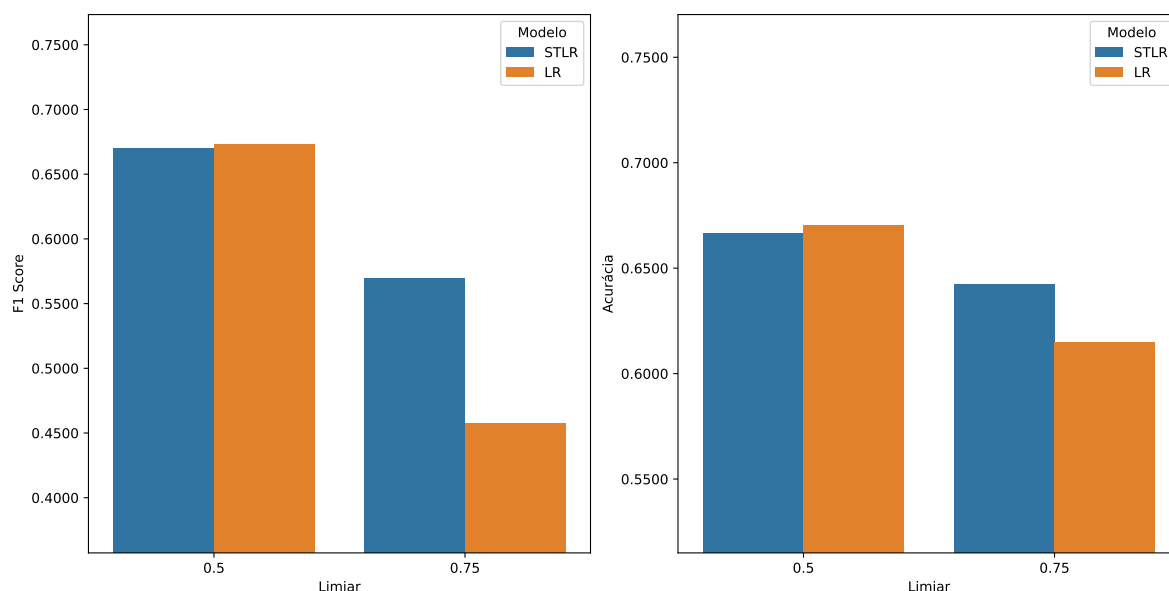
As Figuras 52 e 53 apresentam, respectivamente, os resultados para os experimentos com 60% e 80% dos dados mascarados. Esse mesmo percentual diz respeito ao número de exemplos pré-processados que foram utilizados para teste do modelo supervisionado.

Através dos gráficos é possível identificar que, com limiar de 0,5, o modelo STRL tem desempenho superior quando 60% dos rótulos é mascarado, assim como o modelo LR registra métricas mais elevadas ao utilizar 40% dos dados para seu treinamento. Além disso, fica evidente que a técnica de autotreinamento supera a classificação supervisionada quando o limiar para a determinação das classes sobe para 0,75.

O limite de 0,5 muitas vezes é utilizado como padrão para as classificações, influenciando nas métricas calculadas e, consequentemente, no desempenho do modelo ao prever novos dados. No entanto, alguns problemas exigem que este valor seja maior. No domínio desta pesquisa, que busca auxiliar na classificação de potencial depressão e risco de suicídio, ter classificações positivas com um maior grau

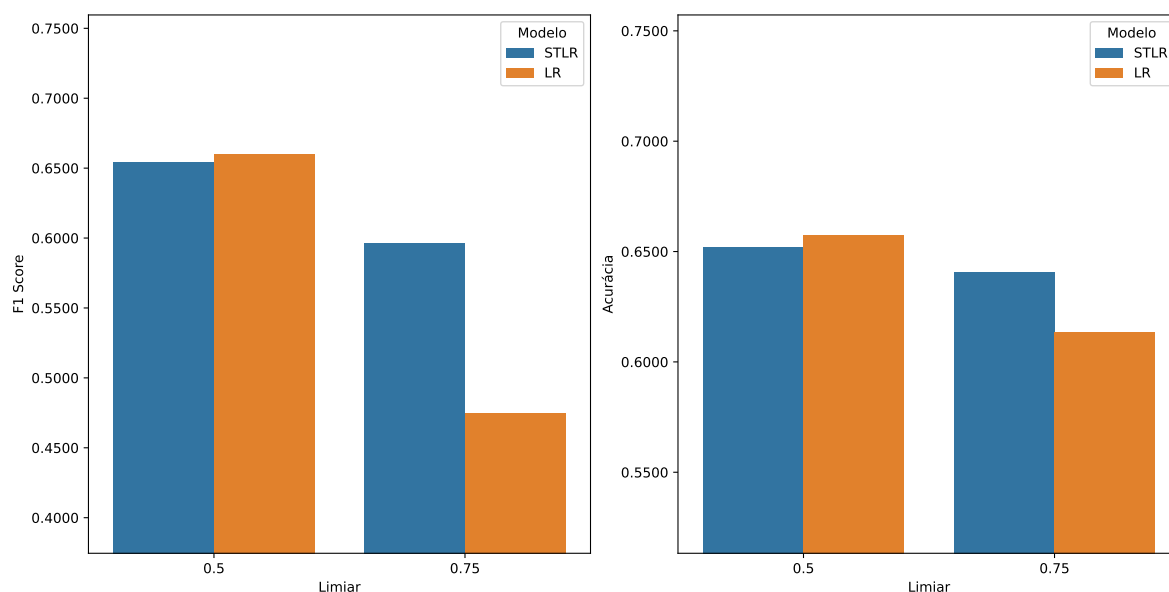
de confiança pode diminuir a geração de alertas incorretos. No entanto, casos incertos podem ser classificados como negativos e impedir a detecção precoce de ideias suicidas, por exemplo.

Figura 52 – Mental Health Twitter pré-processado - Autotreinamento com 60% de dados mascarados



Fonte: Elaborado pela autora

Figura 53 – Mental Health Twitter pré-processado - Autotreinamento com 80% de dados mascarados

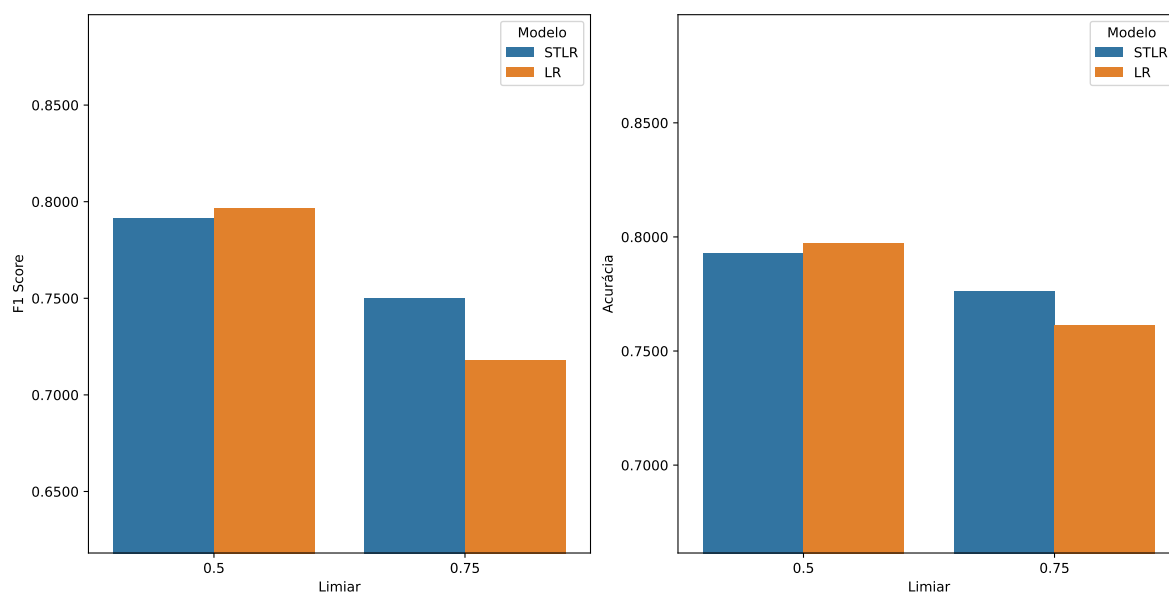


Fonte: Elaborado pela autora

Para o conjunto não processado, o comportamento se mantém. Com 80% dos dados mascarados, conforme mostra a Figura 55, os modelos apresentam queda

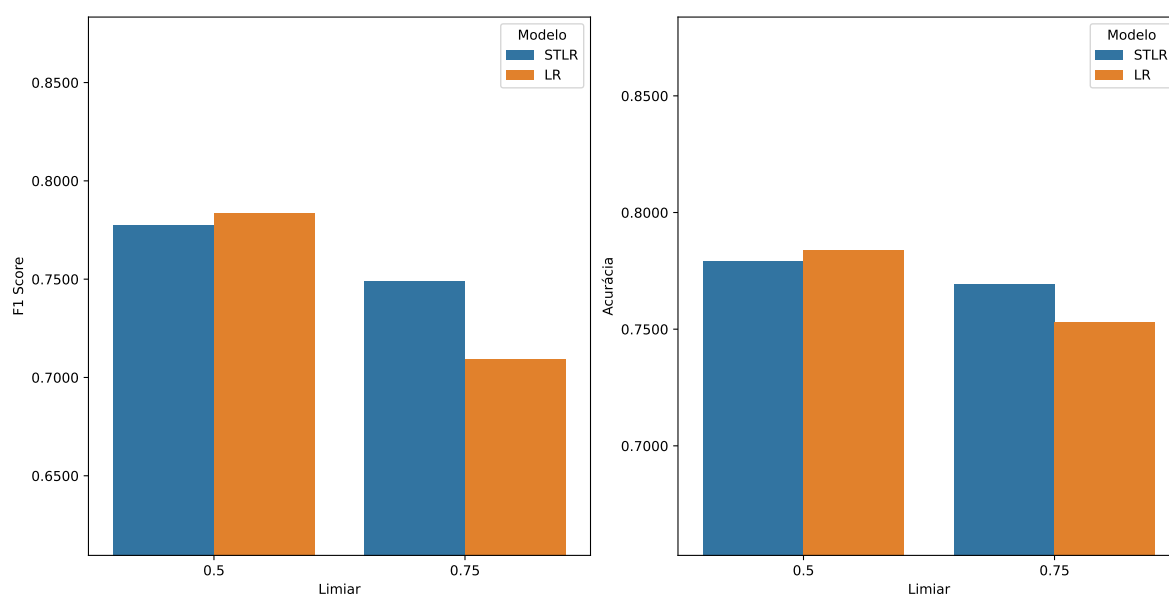
nos resultados, quando comparados com os obtidos com 60%, que são exibidos na Figura 52. Com este percentual e com limiar de 0,75, o modelo STLR registrou acurácia de 0,7692 e F1 score de 0,7489. O modelo LR, no entanto, obteve acurácia de 0,7530 e F1 de 0,7095.

Figura 54 – Mental Health Twitter não processado - Autotreinamento com 60% de dados mascarados



Fonte: Elaborado pela autora

Figura 55 – Mental Health Twitter não processado - Autotreinamento com 80% de dados mascarados



Fonte: Elaborado pela autora

De modo geral, para a detecção de potencial depressão, o conjunto não

processado com 60% de rótulos mascarados é a melhor configuração para o modelo STLR. Para o limiar padrão, de 0,5, a classificação supervisionada com 60% dos dados para teste ainda se mantém superior, mesmo com pequenas diferenças. No entanto, para classificações mais robustas e confiantes, a aprendizagem semi-supervisionado apresenta resultados mais promissores.

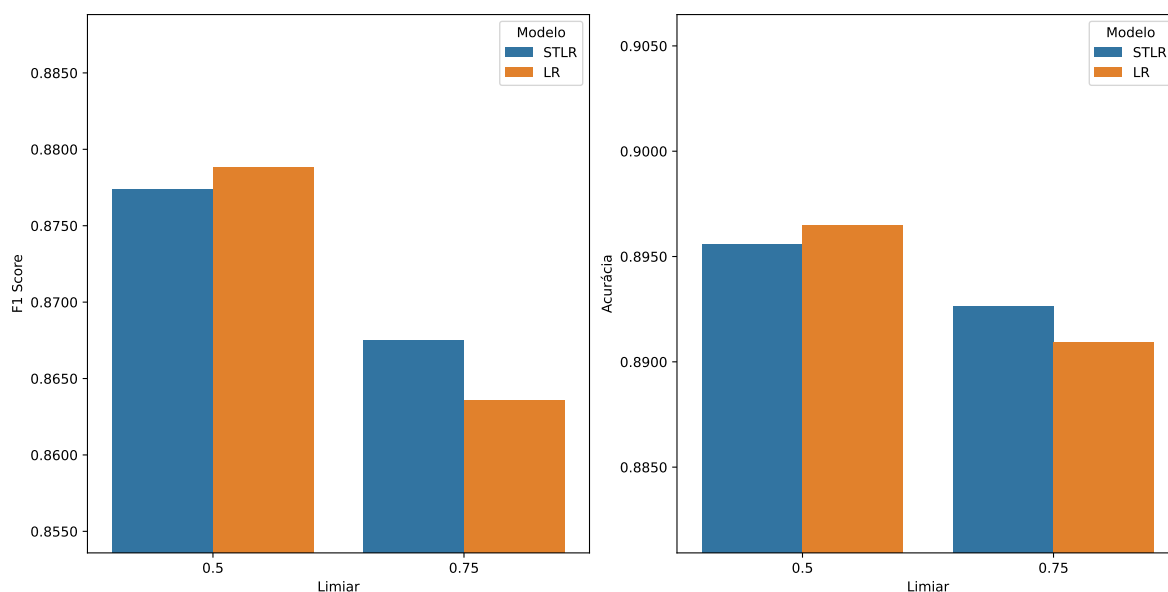
Suicidal Data

O primeiro conjunto para detecção de suicídio, Suicidal Data, foi avaliado inicialmente com os dados pré-processados. As Figuras 56 e 57 exibem, respectivamente, os cenários com 60% e 80% dos dados mascarados, sendo o primeiro considerado o melhor cenário.

O melhor cenário evidencia que, com o limiar padrão de 0,5, o modelo LR atingiu as maiores métricas, com acurácia de 0,8964 e F1 score de 0,8788. Com limiar de 0,75, os valores diminuem, chegando a 0,8909 para acurácia e 0,8635 para a métrica F1. Ao aplicar o autotreinamento e limiar de 0,5, o modelo obteve acurácia de 0,8955 e F1 score 0,8773, enquanto o aumento para 0,75 resultou em acurácia de 0,8926 e F1 score de 0,8675.

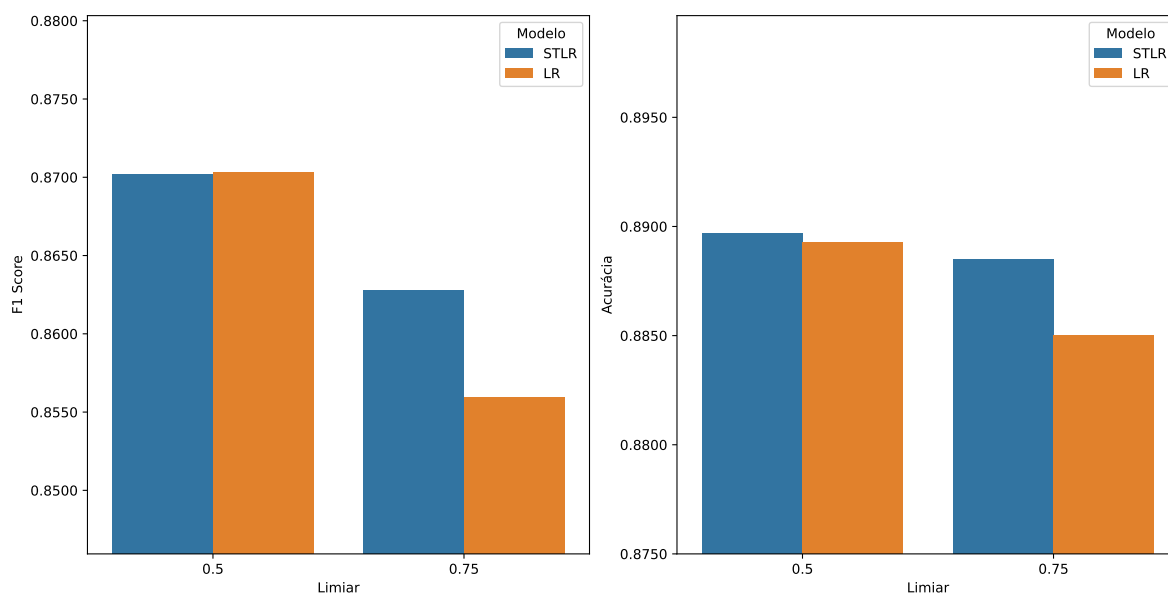
Com limiar de 0,75, a classificação semi-supervisionada foi superior, assim como ocorreu na detecção de potencial depressão. Além disso, a diferença entre a métrica F1 score durante o aumento do **threshold** foi menor com o modelo STLR. Enquanto LR diminuiu 0,0153, STLR reduziu 0,0098.

Figura 56 – Suicidal Data pré-processado - Autotreinamento com 60% de dados mascarados



Fonte: Elaborado pela autora

Figura 57 – Suicidal Data pré-processado - Autotreinamento com 80% de dados mascarados

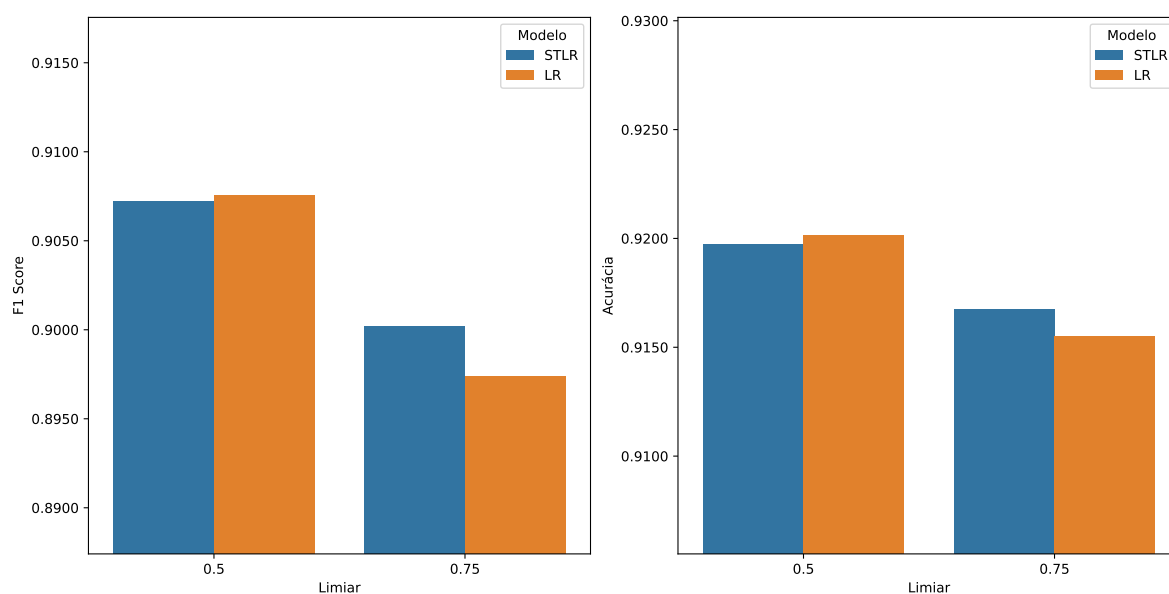


Fonte: Elaborado pela autora

Com a ausência de pré-processamento, o desempenho dos modelos aumentou, assim como ocorreu em outros experimentos para este mesmo conjunto. As Figuras 58 e 59 apresentam, respectivamente, os resultados obtidos com o uso de 60% e 80% de dados mascarados e, conseqüentemente, o mesmo percentual para o teste do modelo supervisionado.

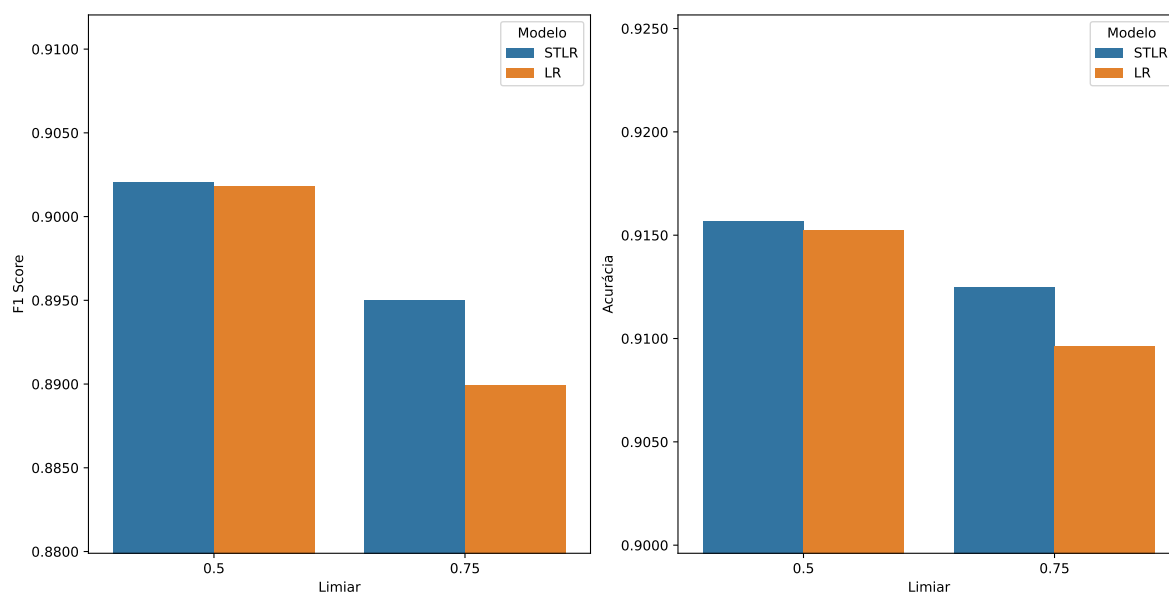
No melhor cenário, com 60% de dados mascarados, o modelo STLR atingiu F1 score de 0,9072 e 0,9001 para o limiar de 0,5 e 0,75, nesta ordem. O modelo LR, com limiar de 0,5, registrou F1 score de 0,9075, apenas 0,0003 a menos que a técnica supervisionada.

Figura 58 – Suicidal Data não processado - Autotreinamento com 60% de dados mascarados



Fonte: Elaborado pela autora

Figura 59 – Suicidal Data não processado - Autotreinamento com 80% de dados mascarados



Fonte: Elaborado pela autora

Suicide Detection

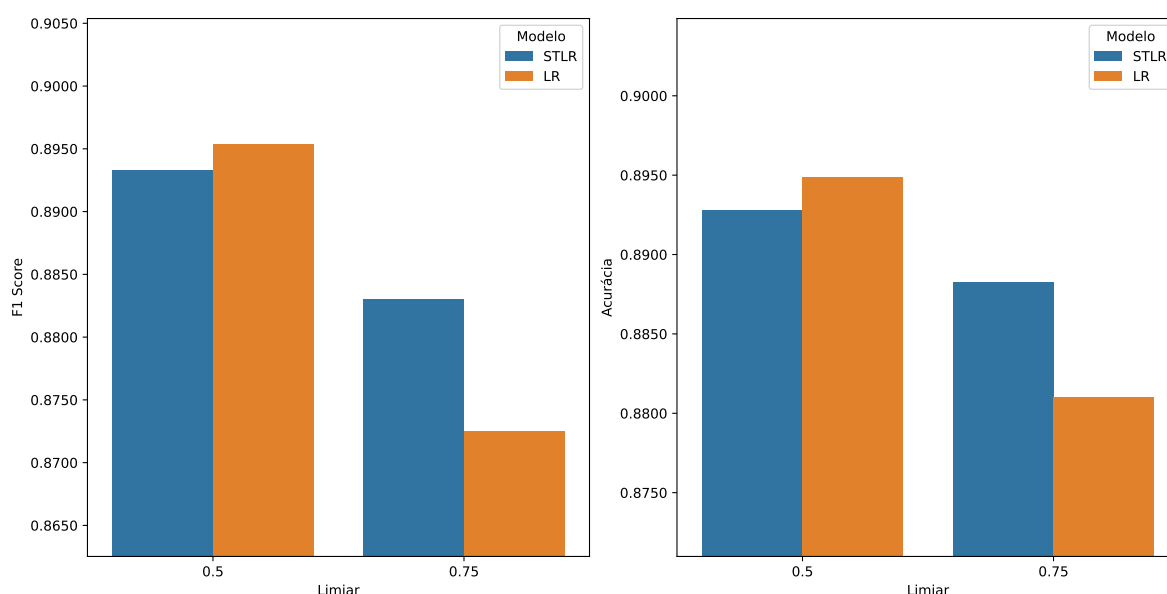
Para finalizar os experimentos do autotreinamento, o conjunto Suicide Detection foi avaliado. Com os dados pré-processados, as melhores métricas para a detecção de risco de suicídio foram obtidas no cenário exibido pela Figura 60, com 60% de rótulos mascarados para a classificação semi-supervisionada e o mesmo percentual de

exemplos para teste do modelo supervisionado.

O modelo STLR, neste cenário, obteve F1 score de 0,8932 com o uso de 0,5 como limite de classificação. Nesta mesma configuração, o modelo LR registrou desempenho superior, com 0,0021 a mais do que a técnica de autotreinamento. Ao aumentar o limiar da classificação, ocorreu queda nos valores das métricas calculadas, assim como observado nos experimentos dos conjuntos anteriores. Com a mudança da configuração, o autotreinamento demonstrou ser superior, chegando a um F1 score de 0,8830, enquanto a Regressão Logística atingiu 0,8725.

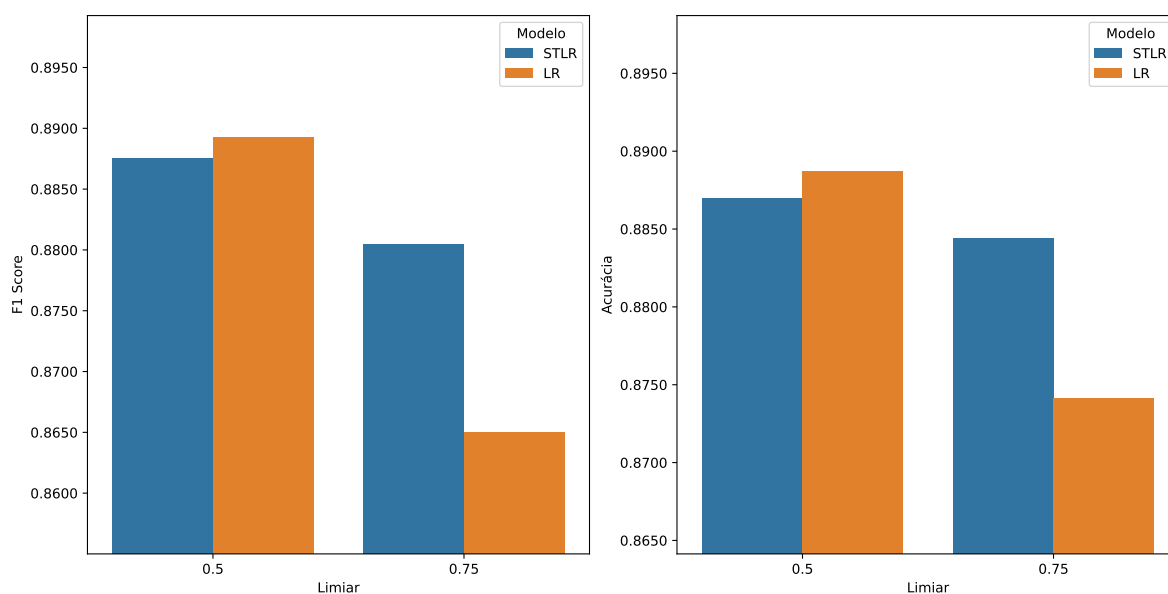
Com 80% dos dados, os resultados são exibidos na Figura 61 e demonstram o desempenho inferior dos modelos para este percentual de mascaramento e de exemplos utilizados no teste supervisionado.

Figura 60 – Suicide Detection pré-processado - Autotreinamento com 60% de dados mascarados



Fonte: Elaborado pela autora

Figura 61 – Suicide Detection pré-processado - Autotreinamento com 80% de dados mascarados

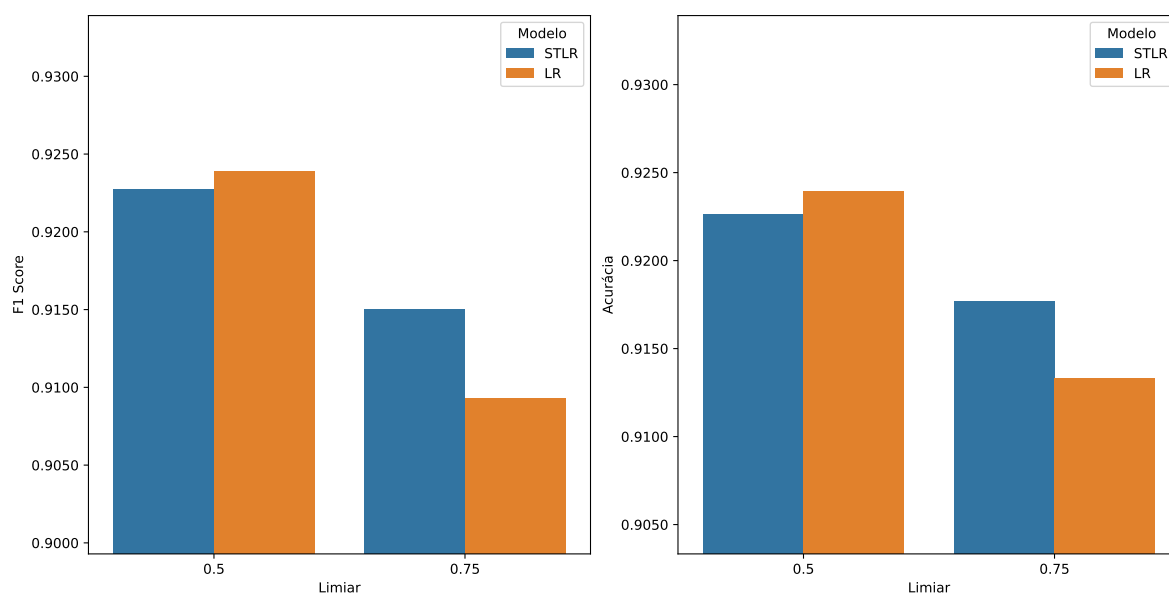


Fonte: Elaborado pela autora

Com a ausência do pré-processamento, os resultados foram superiores, conforme mostram as Figuras 62 e 63. Novamente, com 60% de dados mascarados e limiar de 0,75, a classificação semi-supervisionada apresentou os maiores resultados em comparação a classificação supervisionada com o mesmo limiar.

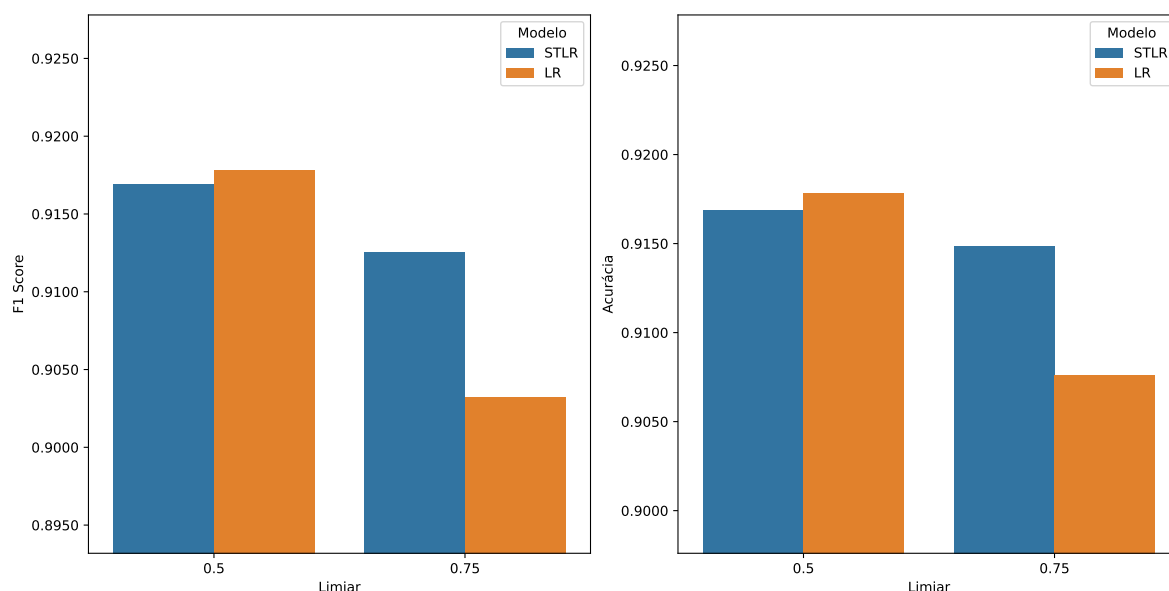
A Figura 62 exibe o cenário citado, onde STLR registrou acurácia de 0,9176 e F1 score de 0,9150, enquanto LR obteve acurácia de 0,9133 e F1 score de 0,9092. A escolha do *threshold* demonstrou ter impacto na classificação final, assim como as técnicas de pré-processamento e ausência delas.

Figura 62 – Suicide Detection não processado - Autotreinamento com 60% de dados mascarados



Fonte: Elaborado pela autora

Figura 63 – Suicide Detection não processado - Autotreinamento com 80% de dados mascarados



Fonte: Elaborado pela autora

De modo geral, o autotreinamento com o modelo Regressão Logística exibiu resultados que o tornam uma escolha possível para a classificação semi-supervisionada na área da saúde mental, podendo ser superior a classificação supervisionada em casos onde a atribuição de rótulos positivos exija maior grau de confiança.

Ter o modelo com menor taxa de falsos positivos e maior robustez pode

gerar aumento no número de falsos negativos. A escolha manual do limiar deve ser realizada considerando o problema a ser resolvido, critérios de aceite para a resolução e, sempre que possível, com a contribuição de um profissional especializado na área.

7.3 Discussão

Diferentes experimentos foram realizados durante a pesquisa, e esta seção busca sintetizar os resultados obtidos e definir, em uma análise geral, a melhor abordagem entre diferentes comparações realizadas, conforme mostra a Tabela 23 para o aprendizado supervisionado, e as tabelas 24 e 25 para o aprendizado semi-supervisionado.

Tabela 23 – Síntese Geral de Resultados - Classificação Supervisionada

Comparação	Cenário	Conjunto	Melhor abordagem	F1-Score
Cenário PP vs NP	PP/NP	MHT	SVM RBF - NP	0,8119
Cenário PP vs NP	PP/NP	SDa	LR - NP	0,9153
Cenário PP vs NP	PP/NP	SDe	SVM RBF - NP	0,9300
Com PCA vs Sem PCA	PP	MHT	SVM RBF - PCA k=600	0,7281
Com PCA vs Sem PCA	PP	SDa	SVM RBF - PCA k=600	0,8900
Com PCA vs Sem PCA	PP	SDe	SVM RBF - PCA k=600	0,9017
Com PCA vs Sem PCA	NP	MHT	SVM RBF - PCA k=600	0,8516
Com PCA vs Sem PCA	NP	SDa	SVM RBF - PCA k=600	0,9193
Com PCA vs Sem PCA	NP	SDe	SVM RBF - PCA k=600	0,9325
Com SC vs Sem SC	PP	MHT	Com SC	0,7091
Com SC vs Sem SC	PP	SDa	Sem SC	0,8880
Com SC vs Sem SC	PP	SDe	Sem SC	0,9009
Com SC vs Sem SC	NP	MHT	Com SC	0,8180
Com SC vs Sem SC	NP	SDa	Sem SC	0,9153
Com SC vs Sem SC	NP	SDe	Sem SC	0,9291

Fonte: Elaborado pela autora

A Tabela 23 permite compreender que a melhor abordagem depende do conjunto de dados analisado. Em todos os conjuntos utilizados nesta pesquisa, o cenário sem pré-processamento apresentou os melhores valores nas métricas avaliadas. O comportamento observado confirma a hipótese de que a aplicação de técnicas de pré-processamento impacta o desempenho dos modelos, principalmente na classificação de potencial depressão. Esse resultado está associado à representação textual adotada na metodologia, que utiliza o modelo BERT para a geração de *embeddings*. Modelos de linguagem pré-treinados podem ter seu desempenho afetado pelas técnicas de pré-processamento aplicadas, bem como pela combinação dessas técnicas, uma vez que tais procedimentos podem resultar na perda de informações relevantes e comprometer a

distinção entre palavras semanticamente diferentes, em decorrência do uso de *stemming* (SIINO; TINNIRELLO; La Cascia, 2024).

Além disso, também fica evidente que, no geral, o modelo SVM RBF é o melhor classificador para os experimentos da pesquisa e se beneficia da aplicação do algoritmo PCA para redução de dimensionalidade, atingindo melhores resultados com 600 componentes. Para os modelos não listados no resumo da tabela, por obterem resultado inferior ao SVM RBF, a aplicação de PCA não trouxe melhora no desempenho das classificações, como ocorreu com o modelo Random Forest. Esta análise permite confirmar a hipótese de que o impacto da redução de dimensionalidade depende da variância mantida e da natureza e configuração dos modelos, podendo afeta-los negativamente.

Por fim, a utilização do classificador Stacking Classifier para combinação de modelos exibiu melhora apenas no conjunto MHT, não sendo a melhor abordagem para os conjuntos de detecção de potencial suicídio. Para a classificação semi-supervisionada, 60% e 80% dos dados tiveram seus rótulos mascarados. Com um maior número de rótulos conhecidos, é esperado que o modelo tenha desempenho superior, o que foi evidenciado pelos resultados obtidos na pesquisa. No entanto, cada cenário também foi analisado de forma individual para que fosse possível definir a melhor abordagem.

Tabela 24 – Síntese Geral de Resultados - Classificação Semi-Supervisionada com mascaramento de 60%

Comparação	Cenário	Masc.	Conjunto	Melhor abordagem	F1-Score
Cenário PP vs NP	PP/NP	60%	MHT	CCP - NP	0,7828
Cenário PP vs NP	PP/NP	60%	SDa	CCP - NP	0,8853
Cenário PP vs NP	PP/NP	60%	SDe	CCP - NP	0,8623
Com PCA vs Sem PCA	PP	60%	MHT	CCP k=100	0,6756
Com PCA vs Sem PCA	PP	60%	SDa	Sem PCA	0,8620
Com PCA vs Sem PCA	PP	60%	SDe	CCP k=20	0,8355
Com PCA vs Sem PCA	NP	60%	MHT	Sem PCA	0,7828
Com PCA vs Sem PCA	NP	60%	SDa	Sem PCA	0,8853
Com PCA vs Sem PCA	NP	60%	SDe	LS k=40	0,8687
Com ST vs Sem ST	PP	60%	MHT	ST limiar=0,5	0,6704
Com ST vs Sem ST	PP	60%	SDa	ST limiar=0,5	0,8773
Com ST vs Sem ST	PP	60%	SDe	ST limiar=0,5	0,8932
Com ST vs Sem ST	NP	60%	MHT	ST limiar=0,5	0,7915
Com ST vs Sem ST	NP	60%	SDa	ST limiar=0,5	0,9072
Com ST vs Sem ST	NP	60%	SDe	ST limiar=0,5	0,9227

Fonte: Elaborado pela autora

Para o mascaramento de 60% dos dados, a Tabela 24 demonstra que a ausência de pré-processamento também resultou em melhor desempenho para classificações semi-supervisionadas, reafirmando a hipótese citada na análise das classificações supervisionadas. Para as classificações sem redução de dimensionalidade, a pesquisa identificou que o modelo CCP é a melhor escolha, com F1-Score superior ao que foi registrado pelos modelos LS e LP.

A redução de dimensionalidade não beneficiou todos os conjuntos de dados. Considerando dados não processados, apenas o conjunto SDe conseguiu obter resultado superior ao utilizar apenas 40 componentes, enquanto os conjuntos MHT e SDa mantiveram o melhor desempenho com a ausência do algoritmo PCA.

A Tabela 24 também permite concluir que a aplicação do Autotreinamento com o a utilização do modelo Self Training Classifier configurado com o modelo Logistic Regression e limiar de 0,5 superou os resultados obtidos pelos modelos semi-supervisionados LP, LS e CCP.

Tabela 25 – Síntese Geral de Resultados - Classificação Semi-Supervisionada com mascaramento de 80%

Comparação	Cenário	Masc.	Conjunto	Melhor abordagem	F1-Score
Cenário PP vs NP	PP/NP	80%	MHT	CCP - NP	0,7497
Cenário PP vs NP	PP/NP	80%	SDa	CCP - NP	0,8788
Cenário PP vs NP	PP/NP	80%	SDe	CCP - NP	0,8525
Com PCA vs Sem PCA	PP	80%	MHT	CCP k=100	0,6522
Com PCA vs Sem PCA	PP	80%	SDa	Sem PCA	0,8549
Com PCA vs Sem PCA	PP	80%	SDe	CCP k=20	0,8287
Com PCA vs Sem PCA	NP	80%	MHT	CCP k=100	0,7586
Com PCA vs Sem PCA	NP	80%	SDa	Sem PCA	0,8788
Com PCA vs Sem PCA	NP	80%	SDe	LS k=20	0,8613
Com ST vs Sem ST	PP	80%	MHT	ST limiar=0,5	0,6537
Com ST vs Sem ST	PP	80%	SDa	ST limiar=0,5	0,8702
Com ST vs Sem ST	PP	80%	SDe	ST limiar=0,5	0,8875
Com ST vs Sem ST	NP	80%	MHT	ST limiar=0,5	0,7774
Com ST vs Sem ST	NP	80%	SDa	ST limiar=0,5	0,9020
Com ST vs Sem ST	NP	80%	SDe	ST limiar=0,5	0,9168

Fonte: Elaborado pela autora

Com 80% dos dados mascarados, a ausência de pré-processamento também demonstrou ser a melhor escolha, conforme evidencia a Tabela 25. Diferente dos resultados comentados anteriormente para a aplicação de redução de dimensionalidade, os conjuntos MHT e SDe registraram melhores desempenhos com a utilização do algoritmo PCA, sendo necessário 100 componentes para o melhor resultado na detecção

de potencial depressão e 20 componentes para a detecção de potencial risco de suicídio. O conjunto SDa, no entanto, não se beneficiou da redução de dimensionalidade.

Por fim, a utilização do Self Training Classifier também apresentou melhor resultado para todos os conjuntos, com limiar de 0,5 e dados sem redução de dimensionalidade.

Embora os modelos semi-supervisionados tenham apresentado resultado inferior aos modelos supervisionados, os experimentos confirmaram a hipótese de que é possível utiliza-los para identificar, com desempenho satisfatório, indícios de sintomas de depressão e de comportamento suicida em textos. O modelo CCP é o melhor classificador para os conjuntos utilizados na pesquisa caso seja preferível trabalhar com métodos baseados em grafos. No entanto, o autotreinamento é a abordagem que traz resultados mais próximos aos obtidos com classificação supervisionada, mesmo que não os supere.

8 CONCLUSÃO

Esta pesquisa propôs a aplicação de técnicas de Processamento de Linguagem Natural e Aprendizado de Máquina para avaliar o desempenho de modelos supervisionados e semi-supervisionados na detecção de potencial depressão e risco de suicídio, explorando diferentes abordagens para o comparativo de seus impactos.

Os experimentos foram realizados com a aplicação dos modelos em três conjuntos de dados contendo postagens em língua inglesa, sendo eles divididos na proporção 80/20, com 80% dos dados para treino e 20% dos dados para teste. Estes conjuntos foram escolhidos por meio da busca de trabalhos relacionados, garantindo que tais dados fossem públicos e relevantes. No entanto, principalmente se tratando de postagens extraídas da rede social Reddit, foi importante ter em mente que a atribuição das classes, com base no *subreddit* em que o texto foi disponibilizado, poderia gerar mais rotulações incorretas por parte dos autores do conjunto.

Os resultados obtidos durante os diferentes cenários explorados, junto à comparação com trabalhos relacionados, confirmaram a hipótese de que as *embeddings* geradas pelo modelo BERT garantem um bom desempenho das classificações, uma vez que, mesmo com poucos dados não mascarados ou utilizados para treino, os modelos apresentaram resultados promissores. Também foi possível compreender o impacto da escolha das técnicas de pré-processamento nos textos disponíveis e reforçar que determinadas etapas podem gerar perda de informações relevantes para as classificações.

A aplicação de redução de dimensionalidade com o algoritmo PCA não trouxe benefício para todos os modelos escolhidos para a pesquisa. Em alguns casos, como o modelo Random Forest, as métricas tiveram queda significativa ao aplicar PCA, mesmo com diferentes números de componentes, confirmando a hipótese de que esta abordagem pode afetar negativamente o desempenho de alguns modelos.

Com a aplicação dos modelos semi-supervisionados foi possível classificar os textos e obter resultados satisfatórios, além de destacar o algoritmo de Competição e Cooperação entre Partículas como uma das melhores escolhas para os conjuntos utilizados.

Ao aplicar a técnica de auto-treinamento, a pesquisa também confirmou parcialmente a hipótese de que existem abordagens semi-supervisionadas que podem gerar resultados próximos, mas não superiores ao desempenho obtido por modelos supervisionados treinados com a mesma quantidade de dados rotulados. A confirmação foi parcial porque, dependendo do *threshold* escolhido para o auto-treinamento e a avaliação do modelo supervisionado, a abordagem semi-supervisionada é capaz de ser superior, como foi observado com o *threshold* de 0,75.

A pesquisa contribui para a área de PLN aplicada à saúde mental, ao demonstrar que é possível detectar potencial depressão e risco de suicídio em tex-

tos, alcançando desempenho satisfatório em classificações supervisionadas e semi-supervisionadas. Além disso, o trabalho explora o impacto das etapas de pré-processamento, analisando como a aplicação, combinação ou ausência dessas técnicas afeta o desempenho dos modelos. Por fim, a pesquisa também inova ao aplicar o modelo de Competição e Cooperação entre Partículas na classificação de textos extraídos de redes sociais, trazendo uma nova abordagem semi-supervisionada nesse contexto.

Para melhorar o que foi conquistado nesta pesquisa, é possível criar uma etapa para escolha do *threshold* de classificação e escolha do modelo com base na avaliação de um profissional especialista em saúde mental. Ter um profissional da área neste processo fará com que o modelo escolhido seja o mais assertivo e robusto, levando em consideração as métricas calculadas e os critérios adotados pela psicologia.

REFERÊNCIAS

- ABBAS, M. A.; MUNIR, K.; RAZA, A.; SAMEE, N. A.; JAMJOOM, M. M.; ULLAH, Z. Novel transformer based contextualized embedding and probabilistic features for depression detection from social media. *IEEE Access*, v. 12, p. 54087–54100, 2024.
- ADANKON, M. M.; CHERIET, M. Learning semi-supervised svm with genetic algorithm. In: *2007 International Joint Conference on Neural Networks*. [S.l.: s.n.], 2007. p. 1825–1830.
- ALPAYDIN, E. *Introduction to machine learning*. [S.l.]: MIT Press, 2004. ISBN 978-0-262-01211-9.
- American Psychiatric Association. *Manual Diagnóstico e Estatístico de Transtornos Mentais: DSM-5*. Porto Alegre, 2014.
- ARAF, I.; IDRI, A.; CHAIRI, I. Cost-sensitive learning for imbalanced medical data: a review. *Artificial Intelligence Review*, 2024.
- ARDIMENTO, P. Predicting bug-fixing time: Distilbert versus google bert. In: SPRINGER. *International Conference on Product-Focused Software Process Improvement*. [S.l.], 2022.
- BARBER, D. *Bayesian Reasoning and Machine Learning*. [S.l.]: Cambridge University Press, 2012.
- BARBOSA, F. d. O.; MACEDO, P. C. M.; SILVEIRA, R. M. C. d. Depressão e o suicídio. *Revista da SBPH*, scielopepsic, v. 14, p. 233 – 243, 06 2011. ISSN 1516-0858. Disponível em: <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1516-08582011000100013&nrm=iso>.
- BLUM, A.; MITCHELL, T. Combining labeled and unlabeled data with co-training. In: *Proceedings of the eleventh annual conference on Computational learning theory*. [S.l.: s.n.], 1998. p. 92–100.
- BOER, Y.; VALENCIA, L.; SETIADI, M. R.; SETIAWAN, K. E.; HASANI, M. F. Classification of heart disease: Comparative analysis using knn, random forest, gaussian naive bayes, xgboost, svm, decision tree, and logistic regression. In: *2023 5th International Conference on Cybernetics and Intelligent System (ICORIS)*. [S.l.: s.n.], 2023. p. 1–5.
- BREVE, F. *Computação inspirada pela natureza*. 2024. Material de aula (slides). Acesso em: 21 jun. 2024. Disponível em: <<https://fabriciobreve.com/class-materials/computacao-inspirada-pela-natureza/>>.
- BREVE, F.; ZHAO, L.; QUILES, M.; PEDRYCZ, W.; LIU, J. Particle competition and cooperation in networks for semi-supervised learning. *IEEE Transactions on Knowledge and Data Engineering*, v. 24, n. 9, p. 1686–1698, 2012.
- BREVE, F. A. *Aprendizado de máquina em redes complexas*. Tese (Doutorado) — Tese de Doutorado em Ciências de Computação e Matemática Computacional. Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2010.
- BRISCOE, T. Introduction to linguistics for natural language processing. *Computer Laboratory, Univ. Cambridge*, Citeseer, v. 4, p. 25, 2011.

CASELI, H. M.; NUNES, M. G. V. (Ed.). *Processamento de Linguagem Natural: Conceitos, Técnicas e Aplicações em Português*. 2. ed. BPLN, 2024. ISBN 978-65-00-95750-1. Disponível em: <<https://brasileiraspln.com/livro-pln/2a-edicao/>>.

CFM, C. F. de M. *Suicídio: informando para prevenir*. Brasília: CFM, 2014.

CHACHAMOVICH, E.; STEFANELLO, S.; BOTEAGA, N.; TURECKI, G. Quais são os recentes achados clínicos sobre a associação entre depressão e suicídio? *Brazilian Journal of Psychiatry*, Associação Brasileira de Psiquiatria, v. 31, p. S18–S25, May 2009. ISSN 1516-4446. Disponível em: <<https://doi.org/10.1590/S1516-44462009000500004>>.

CHADHA, A.; GUPTA, A.; KUMAR, Y. Suicidal ideation detection on social media: A machine learning approach. In: *2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS)*. [S.l.: s.n.], 2022. p. 685–688.

CHAPELLE, O.; SCHÖLKOPF, B.; ZIEN, A. *Semi-supervised Learning*. Cambridge, MA: The MIT Press, 2006. (Adaptive Computation and Machine Learning).

CHEBLI, A.; DJEBBAR, A.; MAROUANI, H. F. Semi-supervised learning for medical application: A survey. In: *2018 International Conference on Applied Smart Systems (ICASS)*. [S.l.: s.n.], 2018. p. 1–9.

CHEN, H.; WU, L.; CHEN, J.; LU, W.; DING, J. A comparative study of automated legal text classification using random forests and deep learning. *Information Processing Management*, v. 59, n. 2, p. 102798, 2022. ISSN 0306-4573. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0306457321002764>>.

CHEN, H.; YU, Y.; JIA, Y.; GU, B. Incremental learning for transductive support vector machine. *Pattern Recognition*, v. 133, p. 108982, 2023. ISSN 0031-3203.

CHOUDHARY, R.; GIANEY, H. K. Comprehensive review on supervised machine learning algorithms. In: *2017 International Conference on Machine Learning and Data Science (MLDS)*. [S.l.: s.n.], 2017. p. 37–43.

CRISCI, C.; GHATTAS, B.; PERERA, G. A review of supervised machine learning algorithms and their applications to ecological data. *Ecological Modelling*, v. 240, p. 113–122, 2012. ISSN 0304-3800.

DAI, B.; CHEN, R.-C.; ZHU, S.-Z.; ZHANG, W.-W. Using random forest algorithm for breast cancer diagnosis. In: *2018 International Symposium on Computer, Consumer and Control (IS3C)*. [S.l.: s.n.], 2018. p. 449–452.

DALIANIS, H. Evaluation metrics and evaluation. In: _____. *Clinical Text Mining: Secondary Use of Electronic Patient Records*. Cham: Springer International Publishing, 2018. p. 45–53. ISBN 978-3-319-78503-5. Disponível em: <https://doi.org/10.1007/978-3-319-78503-5_6>.

DEVLIN, J.; CHANG, M.-W.; LEE, K.; TOUTANOVA, K. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. 2019.

FABRIS, F.; MAGALHÃES, J. P. de; FREITAS, A. A. A review of supervised machine learning applied to ageing research. *Biogerontology* 18, 171–188, Springer, 2017.

FAN, Y.; ARORA, C.; TREUDE, C. *Stop Words for Processing Software Engineering Documents: Do they Matter?* 2023. Disponível em: <<https://arxiv.org/abs/2303.10439>>.

FARO, A.; BAHIANO, M. d. A.; NAKANO, T. d. C.; REIS, C.; SILVA, B. F. P. d.; VITTI, L. S. Covid-19 e saúde mental: a emergência do cuidado. *Estudos de Psicologia (Campinas)*, Programa de Pós-Graduação em Psicologia, Pontifícia Universidade Católica de Campinas, v. 37, p. e200074, 2020. ISSN 0103-166X. Disponível em: <<https://doi.org/10.1590/1982-0275202037e200074>>.

FUKUMITSU, K. O. O psicoterapeuta diante do comportamento suicida. *Psicologia USP*, Instituto de Psicologia da Universidade de São Paulo, v. 25, n. 3, p. 270–275, Sep 2014. ISSN 0103-6564. Disponível em: <<https://doi.org/10.1590/0103-6564D20140001>>.

GADRI, S.; MOUSSAOUI, A. Information retrieval: A new multilingual stemmer based on a statistical approach. In: *2015 3rd International Conference on Control, Engineering Information Technology (CEIT)*. [S.l.: s.n.], 2015. p. 1–6.

GHANEM, R.; ERBAY, H.; BAKOUR, K. Contents-based spam detection on social networks using roberta embedding and stacked blstm. *SN Computer Science*, Springer, v. 4, 2023.

GUPTA, A.; PIRZADA, U. S. M. Lstm network for suicide detection. In: *2023 5th Biennial International Conference on Nascent Technologies in Engineering (ICNTE)*. [S.l.: s.n.], 2023. p. 1–5.

GUPTA, V.; MISHRA, V. K.; SINGHAL, P.; KUMAR, A. An overview of supervised machine learning algorithm. In: *2022 11th International Conference on System Modeling Advancement in Research Trends (SMART)*. [S.l.: s.n.], 2022. p. 87–92.

JIAO, R.; ZHANG, Y.; DING, L.; XUE, B.; ZHANG, J.; CAI, R.; JIN, C. Learning with limited annotations: A survey on deep semi-supervised learning for medical image segmentation. *Computers in Biology and Medicine*, v. 169, p. 107840, 2024. ISSN 0010-4825.

JOLLIFFE, I. T.; CADIMA, J. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 374, n. 2065, p. 20150202, 2016. Disponível em: <<https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2015.0202>>.

JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. *Science*, v. 349, n. 6245, p. 255–260, 2015.

JURAFSKY, D.; MARTIN, J. H. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Third. [S.l.]: Stanford University and University of Colorado at Boulder, 2023.

KAUR, R.; KAUR, A. Text generator using natural language processing methods. In: *2023 International Conference on Computational Intelligence, Communication Technology and Networking (CICTN)*. [S.l.: s.n.], 2023. p. 421–425.

KHURANA, D.; KOLI, A.; KHATTER, K.; SINGH, S. Natural language processing: state of the art, current trends and challenges. *Multimedia Tools and Applications*, v. 82, n. 3, p. 3713–3744, 2023. Disponível em: <<https://doi.org/10.1007/s11042-022-13428-4>>.

KOIRALA, P.; SHAKYA, A. *A Nepali Rule Based Stemmer and its performance on different NLP applications*. 2020. Disponível em: <<https://arxiv.org/abs/2002.09901>>.

LIANG, Z. Effect of decision boundary for logistic regression classifiers on sleep apnea screening accuracy with wearable spo2 data. In: *2023 Fourteenth International Conference on Mobile Computing and Ubiquitous Network (ICMU)*. [S.l.: s.n.], 2023. p. 1–4.

LIDDY, E. D. *Natural language processing*. 2001.

LORENA, A. C.; JACINTHO, L. F.; SIQUEIRA, M. F.; GIOVANNI, R. D.; LOHMANN, L. G.; de Carvalho, A. C.; YAMAMOTO, M. Comparing machine learning classifiers in potential distribution modelling. *Expert Systems with Applications*, v. 38, n. 5, p. 5268–5275, 2011. ISSN 0957-4174.

MASSIE, M. J.; GAGNON, P.; HOLLAND, J. C. Depression and suicide in patients with cancer. *Journal of Pain and Symptom Management*, v. 9, n. 5, p. 325–340, 1994. ISSN 0885-3924. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0885392494901929>>.

MOHAMMED, A.; KORA, R. A comprehensive review on ensemble deep learning: Opportunities and challenges. *Journal of King Saud University - Computer and Information Sciences*, v. 35, n. 2, p. 757–774, 2023. ISSN 1319-1578. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1319157823000228>>.

MRABET, M. A. E.; MAKKAOUI, K. E.; FAIZE, A. Supervised machine learning: A survey. In: *2021 4th International Conference on Advanced Communication Technologies and Networking (CommNet)*. [S.l.: s.n.], 2021. p. 1–10.

N, T. R.; GUPTA, R. A survey on machine learning approaches and its techniques. In: *2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*. [S.l.: s.n.], 2020. p. 1–6.

NEELIMA, A.; MEHROTRA, S. A comprehensive review on word embedding techniques. In: *2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCOIS)*. [S.l.: s.n.], 2023. p. 538–543.

NIKHILESWAR, K.; VISHAL, D.; SPHOORTHY, L.; FATHIMABI, S. Suicide ideation detection in social media forums. In: *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*. [S.l.: s.n.], 2021. p. 1741–1747.

PAJILA, P. B.; SUDHA, K.; SELVI, D. K.; KUMAR, V. N.; GAYATHRI, S.; SUBRAMANIAN, R. S. A survey on natural language processing and its applications. In: *2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC)*. [S.l.: s.n.], 2023. p. 996–1001.

PATIL, R.; BOIT, S.; GUDIVADA, V.; NANDIGAM, J. A survey of text representation and embedding techniques in nlp. *IEEE Access*, v. 11, p. 36120–36146, 2023.

PISE, N. N.; KULKARNI, P. A survey of semi-supervised learning methods. In: *2008 International Conference on Computational Intelligence and Security*. [S.l.: s.n.], 2008. v. 2, p. 30–34.

PORTO, J. A. D. Conceito e diagnóstico. *Brazilian Journal of Psychiatry*, Associação Brasileira de Psiquiatria, v. 21, p. 06–11, May 1999. ISSN 1516-4446. Disponível em: <<https://doi.org/10.1590/S1516-44461999000500003>>.

QUILES, M. G.; ZHAO, L.; ALONSO, R. L.; ROMERO, R. A. F. Particle competition for complex network community detection. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, v. 18, n. 3, p. 033107, 2008.

RAY, S. A quick review of machine learning algorithms. In: *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*. [S.l.: s.n.], 2019. p. 35–39.

RIOS, A. D. *Suicídio e redes sociais: uma análise comportamental das postagens de uma jovem no Instagram*. Dissertação (Dissertação) — Universidade Católica de Brasília, 2021. Dissertação (Programa Stricto Sensu em Psicologia) - Universidade Católica de Brasília, Brasília, 2021. Disponível em: <<https://bdtd.ucb.br:8443/jspui/handle/tede/2841>>.

RODRIGUES, M. J. S. F. O diagnóstico de depressão. *Psicologia USP*, Instituto de Psicologia da Universidade de São Paulo, v. 11, n. 1, p. 155–187, 2000. ISSN 0103-6564. Disponível em: <<https://doi.org/10.1590/S0103-65642000000100010>>.

RODRIGUEZ-GALIANO, V.; SANCHEZ-CASTILLO, M.; CHICA-OLMO, M.; CHICA-RIVAS, M. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, v. 71, p. 804–818, 2015. ISSN 0169-1368. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169136815000037>>.

ROSENBERG, C.; HEBERT, M.; SCHNEIDERMAN, H. Semi-supervised self-training of object detection models. In: *2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05) - Volume 1*. [S.l.: s.n.], 2005. v. 1, p. 29–36.

SARAKUL, W.; RUTHERFORD, A. T. Contextualized vs. static word embeddings for word-based analysis of opposing opinions. In: *2023 20th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. [S.l.: s.n.], 2023. p. 95–100.

SEYYAR, Y. E.; YAVUZ, A. G.; ÜNVER, H. M. An attack detection framework based on bert and deep learning. *IEEE Access*, v. 10, p. 68633–68644, 2022.

SHEIKHPOUR, R.; SARRAM, M. A.; GHARAGHANI, S.; CHAHOOKI, M. A. Z. A survey on semi-supervised feature selection methods. *Pattern Recognition*, v. 64, p. 141–158, 2017. ISSN 0031-3203.

SIINO, M.; TINNIRELLO, I.; La Cascia, M. Is text preprocessing still worth the time? a comparative survey on the influence of popular preprocessing methods on transformers and traditional classifiers. *Information Systems*, v. 121, p. 102342, 2024. ISSN 0306-4379. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0306437923001783>>.

SINGH, A.; THAKUR, N.; SHARMA, A. A review of supervised machine learning algorithms. In: *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*. [S.l.: s.n.], 2016. p. 1310–1315.

SINGLA, M.; GHOSH, D.; SHUKLA, K. K. $\overline{\text{pin-tsvm}}$: A robust transductive support vector machine and its application to the detection of covid-19 infected patients. *Neural Processing Letters*, v. 53, n. 6, p. 3981–4010, 2021. ISSN 1573-773X.

SONBOL, R.; REBDAWI, G.; GHNEIM, N. The use of nlp-based text representation techniques to support requirement engineering tasks: A systematic mapping review. *IEEE Access*, v. 10, p. 62811–62830, 2022.

SONG, Z.; YANG, X.; XU, Z.; KING, I. Graph-based semi-supervised learning: A comprehensive review. *IEEE Transactions on Neural Networks and Learning Systems*, v. 34, n. 11, p. 8174–8194, 2023.

SOUZA, E.; MORIYAMA, G.; VITÓRIO, D.; CARVALHO, A. C. de; FÉLIX, N.; ALBUQUERQUE, H. O.; OLIVEIRA, A. L. Assessing the impact of stemming algorithms applied to brazilian legislative documents retrieval. In: SBC. *Anais do XIII Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana*. [S.l.], 2021. p. 227–236.

SUN, S.; LUO, C.; CHEN, J. A review of natural language processing techniques for opinion mining systems. *Information Fusion*, v. 36, p. 10–25, 2017. ISSN 1566-2535. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1566253516301117>>.

TAUSCH, A.; e Souza, R. O.; VICIANA, C. M.; CAYETANO, C.; BARBOSA, J.; HENNIS, A. J. Strengthening mental health responses to covid-19 in the americas: A health policy analysis and recommendations. *The Lancet Regional Health - Americas*, v. 5, p. 100118, 2022. ISSN 2667-193X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2667193X21001149>>.

TSABITA, F.; ROHMAN, W. N.; ROSMALIATI; LYSTIANINGRUM, B. V.; PURNOMO, M. H. Semi-supervised learning optimization based on generative models to identify type of electric load at low voltage. In: *2018 International Seminar on Intelligent Technology and Its Applications (ISITIA)*. [S.l.: s.n.], 2018. p. 209–214.

VAKILI, M.; GHAMSARI, M.; REZAEI, M. *Performance Analysis and Comparison of Machine and Deep Learning Algorithms for IoT Data Classification*. 2020. Disponível em: <<https://arxiv.org/abs/2001.09636>>.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, L.; POLOSUKHIN, I. *Attention Is All You Need*. 2023. Disponível em: <<https://arxiv.org/abs/1706.03762>>.

VIDAL, C. E. L.; GONTIJO, E. C. D. M.; LIMA, L. A. Tentativas de suicídio: fatores prognósticos e estimativa do excesso de mortalidade. *Cadernos de Saúde Pública*, Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz, v. 29, n. 1, p. 175–187, Jan 2013. ISSN 0102-311X. Disponível em: <<https://doi.org/10.1590/S0102-311X2013000100020>>.

VIEIRA, K. F. L.; COUTINHO, M. d. P. d. L. Representações sociais da depressão e do suicídio elaboradas por estudantes de psicologia. *Psicologia: Ciência e Profissão*, Conselho Federal de Psicologia, v. 28, n. 4, p. 714–727, 2008. ISSN 1414-9893. Disponível em: <<https://doi.org/10.1590/S1414-98932008000400005>>.

WANG, X.; LIU, K.; WANG, C. Knowledge-enhanced pre-training large language model for depression diagnosis and treatment. In: *2023 IEEE 9th International Conference on Cloud Computing and Intelligent Systems (CCIS)*. [S.l.: s.n.], 2023. p. 532–536.

World Health Organization. *Promoting mental health : concepts, emerging evidence, practice : summary report / a report from the World Health Organization, Department of Mental Health and Substance Abuse in collaboration with the Victorian Health Promotion Foundation and the University of Melbourne*. [S.l.]: World Health Organization, 2004. "An editorial summary of a report of the same title published by WHO" p.

World Health Organization. *Depressive disorder (depression)*. 2023. Acesso em: 15 mar 2025. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/depression>>.

World Health Organization 2. *Suicide Fact Sheet*. 2023. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/suicide>>.

YATAPALA, K.; KUMARA, B. Detection of suicide ideation in twitter using ann. In: *2021 6th International Conference on Information Technology Research (ICITR)*. [S.l.: s.n.], 2021. p. 1–5.

ZHU, L.; QIU, D.; ERGU, D.; YING, C.; LIU, K. A study on predicting loan default based on the random forest algorithm. *Procedia Computer Science*, v. 162, p. 503–513, 2019. ISSN 1877-0509. 7th International Conference on Information Technology and Quantitative Management (ITQM 2019): Information technology and quantitative management based on Artificial Intelligence.

ZHU, X. Semi-supervised learning literature survey. *Computer Sciences, University of Wisconsin-Madison*, 2005.

APÊNDICE A – TABELAS DE RESULTADOS

Tabela 26 – PCA - Resultados da classificação supervisionada do conjunto Mental Health Twitter pré-processado

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LR	150	0,6661	0,6633	0,6741	0,6686
	300	0,6704	0,6667	0,6809	0,6737
	450	0,6790	0,6742	0,6923	0,6831
	600	0,6793	0,6747	0,6916	0,6830
	*	0,6809	0,6762	0,6935	0,6847
LR Balanceado	150	0,6663	0,6635	0,6741	0,6687
	300	0,6705	0,6668	0,6808	0,6737
	450	0,6793	0,6745	0,6923	0,6832
	600	0,6793	0,6748	0,6915	0,6830
	*	0,6808	0,6760	0,6936	0,6847
SVM RBF	150	0,7079	0,6886	0,7586	0,7219
	300	0,7128	0,6941	0,7603	0,7257
	450	0,7150	0,6957	0,7637	0,7280
	600	0,7152	0,6962	0,7633	0,7281
	*	0,6876	0,6707	0,7364	0,7020
SVM Linear	150	0,6662	0,6560	0,6982	0,6764
	300	0,6751	0,6652	0,7044	0,6842
	450	0,6798	0,6684	0,7132	0,6900
	600	0,6790	0,6683	0,7102	0,6886
	*	0,6810	0,6715	0,7081	0,6892
RF 100	150	0,6771	0,6681	0,7034	0,6852
	300	0,6658	0,6570	0,6936	0,6747
	450	0,6589	0,65011	0,6880	0,6683
	600	0,6508	0,6418	0,6821	0,6612
	*	0,6798	0,6709	0,7052	0,6875
RF 200	150	0,6866	0,6721	0,7284	0,6990
	300	0,6799	0,6645	0,7265	0,6940
	450	0,6725	0,6554	0,7269	0,6892
	600	0,6674	0,6502	0,7245	0,6852
	*	0,6880	0,6740	0,7278	0,6998

Tabela 27 – PCA - Resultados da classificação supervisionada do Suicidal Data pré-processado

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LR	150	0,8962	0,9122	0,8453	0,8774
	300	0,8977	0,9077	0,8539	0,8799
	450	0,9031	0,9121	0,8626	0,8866
	600	0,9036	0,9113	0,8648	0,8874
	*	0,9042	0,9125	0,8650	0,8880
LR Balanceado	150	0,8941	0,8933	0,8620	0,8773
	300	0,8959	0,8910	0,8694	0,8800
	450	0,9012	0,8960	0,8768	0,8863
	600	0,9021	0,8968	0,8783	0,8874
	*	0,9024	0,8975	0,8782	0,8877
SVM RBF	150	0,9057	0,9412	0,8377	0,8864
	300	0,9084	0,9442	0,8413	0,8897
	450	0,9086	0,9449	0,8411	0,8899
	600	0,9087	0,9453	0,8410	0,8900
	*	0,8988	0,9439	0,8185	0,8767
SVM Linear	150	0,8984	0,9230	0,8388	0,8789
	300	0,8989	0,9146	0,8493	0,8806
	450	0,9021	0,9112	0,8613	0,8855
	600	0,9019	0,9071	0,8654	0,8857
	*	0,8987	0,9007	0,8649	0,8823
RF 100	150	0,8831	0,9332	0,79074	0,8560
	300	0,8798	0,9405	0,7757	0,8501
	450	0,8750	0,9444	0,7606	0,8425
	600	0,8703	0,9491	0,7450	0,8347
	*	0,8882	0,9318	0,8046	0,8634
RF 200	150	0,8852	0,9345	0,7946	0,8588
	300	0,8831	0,9430	0,7814	0,8545
	450	0,8791	0,9483	0,7668	0,8478
	600	0,8755	0,9529	0,7542	0,8419
	*	0,8906	0,9348	0,8075	0,8664

Tabela 28 – PCA - Resultados da classificação supervisionada do Suicide Detection pré-processado

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LR	150	0,8917	0,8874	0,8974	0,8923
	300	0,8970	0,8931	0,9022	0,8976
	450	0,9007	0,8972	0,9052	0,9012
	600	0,9005	0,8970	0,9051	0,9010
	*	0,9005	0,8971	0,9048	0,9009
LR Balanceado	150	0,8918	0,8875	0,8975	0,8924
	300	0,8971	0,8932	0,9022	0,8976
	450	0,9007	0,8974	0,9051	0,9012
	600	0,9006	0,8970	0,9052	0,9011
	*	0,9002	0,8966	0,9048	0,9007
SVM RBF	150	0,8971	0,9016	0,8916	0,8965
	300	0,9000	0,9044	0,8949	0,8996
	450	0,9018	0,9060	0,8967	0,9013
	600	0,9022	0,9066	0,8969	0,9017
	*	0,8948	0,8935	0,8966	0,8950
SVM Linear	150	0,8918	0,8902	0,8941	0,8921
	300	0,8965	0,8937	0,9001	0,8969
	450	0,8992	0,8964	0,9029	0,8996
	600	0,8982	0,8956	0,9016	0,8986
	*	0,8978	0,8952	0,9012	0,8982
RF 100	150	0,8514	0,8871	0,8055	0,8443
	300	0,8426	0,8884	0,7840	0,8329
	450	0,8338	0,8875	0,7648	0,8216
	600	0,8263	0,8849	0,7504	0,8121
	*	0,8551	0,8693	0,8361	0,8523
RF 200	150	0,8543	0,8887	0,8104	0,8477
	300	0,8452	0,8903	0,7878	0,8358
	450	0,8377	0,8900	0,7708	0,8261
	600	0,8305	0,8884	0,7562	0,8169
	*	0,8583	0,8712	0,8412	0,8559

Tabela 29 – PCA - Resultados da classificação supervisionada do Mental Health Twitter não processados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LR	150	0,7733	0,7750	0,7698	0,7723
	300	0,7926	0,7961	0,7863	0,7911
	450	0,7973	0,7997	0,7930	0,7963
	600	0,8021	0,8054	0,7965	0,8009
	*	0,8056	0,8087	0,8004	0,8045
LR Balanceado	150	0,7733	0,7750	0,7699	0,7724
	300	0,7928	0,7963	0,7865	0,7913
	450	0,7972	0,7997	0,7928	0,7962
	600	0,8021	0,8055	0,7963	0,8008
	*	0,8052	0,8084	0,7997	0,8040
SVM RBF	150	0,8397	0,8336	0,8485	0,8409
	300	0,8468	0,8420	0,8538	0,8478
	450	0,8501	0,8453	0,8569	0,8510
	600	0,8508	0,8463	0,8570	0,8516
	*	0,8115	0,8095	0,8144	0,8119
SVM Linear	150	0,7763	0,7775	0,7741	0,7757
	300	0,7942	0,7966	0,7899	0,7932
	450	0,7981	0,7948	0,7948	0,7973
	600	0,8018	0,8045	0,7970	0,8007
	*	0,8067	0,8084	0,8036	0,8060
RF 100	150	0,7799	0,7738	0,7910	0,7822
	300	0,7697	0,7632	0,7818	0,7723
	450	0,7595	0,7516	0,7748	0,7630
	600	0,7506	0,7421	0,7678	0,7547
	*	0,7828	0,7806	0,7865	0,7835
RF 200	150	0,7925	0,7810	0,8126	0,7964
	300	0,7845	0,7699	0,8114	0,7900
	450	0,7766	0,7604	0,8073	0,7831
	600	0,7669	0,7490	0,8027	0,7748
	*	0,7931	0,7852	0,8068	0,7958

Tabela 30 – PCA - Resultados da classificação supervisionada do Suicidal Data não processados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LR	150	0,9232	0,9282	0,8946	0,9110
	300	0,9254	0,9279	0,9003	0,9138
	450	0,9252	0,9270	0,9007	0,9136
	600	0,9263	0,9293	0,9009	0,9148
	*	0,9267	0,9292	0,9019	0,9153
LR Balanceado	150	0,9226	0,9157	0,9075	0,9115
	300	0,9241	0,9163	0,9105	0,9134
	450	0,9242	0,9166	0,9104	0,9134
	600	0,9245	0,9180	0,9093	0,9136
	*	0,9255	0,9196	0,9101	0,9148
SVM RBF	150	0,9300	0,9490	0,8886	0,9178
	300	0,9310	0,9513	0,8886	0,9188
	450	0,9310	0,9513	0,8885	0,9188
	600	0,9315	0,9521	0,8889	0,9193
	*	0,9237	0,9516	0,8707	0,9093
SVM Linear	150	0,9219	0,9260	0,8937	0,9095
	300	0,9234	0,9249	0,8987	0,9116
	450	0,9211	0,9208	0,8977	0,9091
	600	0,9202	0,9178	0,8990	0,9082
	*	0,9182	0,9137	0,8988	0,9061
RF 100	150	0,8916	0,9464	0,7987	0,8662
	300	0,8886	0,9497	0,7883	0,8614
	450	0,8861	0,9525	0,7798	0,8575
	600	0,8827	0,9536	0,7708	0,8524
	*	0,9050	0,9506	0,8269	0,8844
RF 200	150	0,8916	0,9460	0,7991	0,8663
	300	0,8892	0,9505	0,7892	0,8623
	450	0,8878	0,9543	0,7823	0,8597
	600	0,8864	0,9573	0,7764	0,8573
	*	0,9056	0,9523	0,8267	0,8850

Tabela 31 – PCA - Resultados da classificação supervisionada do Suicide Detection não processados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LR	150	0,9243	0,9258	0,9227	0,9242
	300	0,9270	0,9281	0,9259	0,9270
	450	0,9284	0,9290	0,9277	0,9283
	600	0,9288	0,9296	0,9280	0,9288
	*	0,9291	0,9295	0,9287	0,9291
LR Balanceado	150	0,9243	0,9258	0,9227	0,9242
	300	0,9270	0,9281	0,9259	0,9270
	450	0,9283	0,9289	0,9276	0,9283
	600	0,9288	0,9295	0,9280	0,9288
	*	0,9291	0,9293	0,9290	0,9291
SVM RBF	150	0,9298	0,9327	0,9266	0,9296
	300	0,9319	0,9346	0,9289	0,9317
	450	0,9323	0,9350	0,9293	0,9321
	600	0,9327	0,9353	0,9298	0,9325
	*	0,9301	0,9324	0,9276	0,9300
SVM Linear	150	0,9237	0,9242	0,9231	0,9236
	300	0,9260	0,9265	0,9256	0,9260
	450	0,9276	0,9274	0,9279	0,9276
	600	0,9284	0,9278	0,9293	0,9285
	*	0,9276	0,9267	0,9286	0,9277
RF 100	150	0,8767	0,9093	0,8371	0,8716
	300	0,8673	0,9088	0,8167	0,8603
	450	0,8613	0,9072	0,8052	0,8531
	600	0,8566	0,9078	0,7941	0,8471
	*	0,8853	0,9082	0,8573	0,8820
RF 200	150	0,8790	0,9113	0,8399	0,8741
	300	0,8725	0,9129	0,8239	0,8661
	450	0,8669	0,9128	0,8114	0,8591
	600	0,8628	0,9132	0,8020	0,8540
	*	0,8870	0,9091	0,8602	0,8839

Tabela 32 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Mental Health Twitter pré-processado

Técnica	Comp.	Acurácia	Precisão	Recall	F1
RF 100	50	0,6801	0,6721	0,7029	0,6870
	75	0,6788	0,6700	0,7042	0,6866
	100	0,6789	0,6701	0,7042	0,6866
	125	0,6771	0,6682	0,7029	0,6850
RF 200	50	0,6870	0,6747	0,7219	0,6974
	75	0,6864	0,6727	0,7258	0,6981
	100	0,6873	0,6736	0,7262	0,6988
	125	0,6885	0,6742	0,7292	0,7005

Tabela 33 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Mental Health Twitter não processado

Técnica	Comp.	Acurácia	Precisão	Recall	F1
RF 100	50	0,7767	0,7725	0,7842	0,7782
	75	0,7835	0,7772	0,7946	0,7857
	100	0,7850	0,7797	0,7944	0,7869
	125	0,7829	0,7774	0,7926	0,7848
RF 200	50	0,7841	0,7750	0,8005	0,7875
	75	0,7922	0,7806	0,8126	0,7962
	100	0,7945	0,7833	0,8139	0,7983
	125	0,7938	0,7820	0,8146	0,7979

Tabela 34 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Suicidal Data pré-processado

Técnica	Comp.	Acurácia	Precisão	Recall	F1
RF 100	50	0,8864	0,9254	0,8067	0,8619
	75	0,8857	0,9288	0,8015	0,8604
	100	0,8847	0,9280	0,7999	0,8591
	125	0,8835	0,9310	0,7939	0,8569
RF 200	50	0,8875	0,9268	0,8080	0,8632
	75	0,8872	0,9300	0,8039	0,8623
	100	0,8855	0,9296	0,8003	0,8600
	125	0,8851	0,9318	0,7970	0,8591

Tabela 35 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Suicidal Data não processado

Técnica	Comp.	Acurácia	Precisão	Recall	F1
RF 100	50	0,8977	0,9429	0,8167	0,8753
	75	0,8948	0,9441	0,8086	0,8711
	100	0,8935	0,9436	0,8058	0,8692
	125	0,8919	0,9445	0,8012	0,8669
RF 200	50	0,8982	0,9434	0,8174	0,8758
	75	0,8952	0,9446	0,8090	0,8715
	100	0,8935	0,9446	0,8049	0,8691
	125	0,8927	0,9459	0,8017	0,8678

Tabela 36 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Suicide Detection pré-processado

Técnica	Comp.	Acurácia	Precisão	Recall	F1
RF 100	50	0,8566	0,8787	0,8278	0,8524
	75	0,8553	0,8829	0,8196	0,8500
	100	0,8541	0,8843	0,8151	0,8482
	125	0,8522	0,8850	0,8098	0,8457
RF 200	50	0,8594	0,8801	0,8323	0,8555
	75	0,8579	0,8845	0,8237	0,8530
	100	0,8567	0,8853	0,8198	0,8512
	125	0,8552	0,8868	0,8147	0,8492

Tabela 37 – PCA - Resultados do modelo RF na classificação supervisionada do conjunto Suicide Detection não processado

Técnica	Comp.	Acurácia	Precisão	Recall	F1
RF 100	50	0,8834	0,9075	0,8540	0,8799
	75	0,8814	0,9086	0,8484	0,8774
	100	0,8800	0,9090	0,8448	0,8757
	125	0,8781	0,9088	0,8408	0,8735
RF 200	50	0,8854	0,9088	0,8568	0,8820
	75	0,8836	0,9098	0,8519	0,8798
	100	0,8814	0,9099	0,8469	0,8772
	125	0,8806	0,9104	0,8443	0,8761

Tabela 38 – Resultados obtidos para o conjunto Mental Health Twitter com aprendizado semi supervisionado

Técnica	Cenário	Masc.	Acurácia	Precisão	Recall	F1
LP KNN	PP	80%	0,6281	0,6310	0,6229	0,6237
	PP	60%	0,6443	0,6417	0,6552	0,6475
LP RBF	PP	80%	0,6106	0,6188	0,5753	0,5961
	PP	60%	0,6279	0,6360	0,5972	0,6159
LS KNN	PP	80%	0,6312	0,6318	0,6290	0,6299
	PP	60%	0,6461	0,6429	0,6579	0,6499
LS RBF	PP	80%	0,6095	0,6147	0,5860	0,5998
	PP	60%	0,6282	0,6343	0,6051	0,6192
CCP	PP	80%	0,6399	0,6410	0,6367	0,6387
	PP	60%	0,6600	0,6567	0,6706	0,6635
LP KNN	NP	80%	0,7217	0,8021	0,5900	0,6787
	NP	60%	0,7565	0,8046	0,6776	0,7354
LP RBF	NP	80%	0,7167	0,7528	0,6452	0,6947
	NP	60%	0,7386	0,7735	0,6746	0,7206
LS KNN	NP	80%	0,7442	0,7767	0,6858	0,7280
	NP	60%	0,7661	0,7965	0,7148	0,7532
LS RBF	NP	80%	0,7217	0,7483	0,6678	0,7057
	NP	60%	0,7407	0,7690	0,6880	0,7262
CCP	NP	80%	0,7508	0,7534	0,7465	0,7497
	NP	60%	0,7838	0,7863	0,7795	0,7828

Tabela 39 – Resultados obtidos para o conjunto Suicidal Data com aprendizado semi supervisionado

Técnica	Cenário	Masc.	Acurácia	Precisão	Recall	F1
LP KNN	PP	80%	0,8505	0,8114	0,8608	0,8350
	PP	60%	0,8669	0,8440	0,8554	0,8495
LP RBF	PP	80%	0,8255	0,7732	0,8538	0,8113
	PP	60%	0,8322	0,7811	0,8593	0,8182
LS KNN	PP	80%	0,8589	0,8286	0,8567	0,8422
	PP	60%	0,8693	0,8481	0,8560	0,8519
LS RBF	PP	80%	0,8275	0,7839	0,8393	0,8104
	PP	60%	0,8346	0,7899	0,8500	0,8186
CCP	PP	80%	0,8786	0,8986	0,8154	0,8549
	PP	60%	0,8839	0,9017	0,8257	0,8620
LP KNN	NP	80%	0,8861	0,8761	0,8636	0,8694
	NP	60%	0,8951	0,8901	0,8688	0,8792
LP RBF	NP	80%	0,8527	0,8481	0,8103	0,8286
	NP	60%	0,8604	0,8598	0,8156	0,8370
LS KNN	NP	80%	0,8893	0,8791	0,8678	0,8732
	NP	60%	0,8961	0,8903	0,8710	0,8805
LS RBF	NP	80%	0,8531	0,8540	0,8036	0,8278
	NP	60%	0,8597	0,8613	0,8117	0,8357
CCP	NP	80%	0,8980	0,9177	0,8432	0,8788
	NP	60%	0,9030	0,9199	0,8533	0,8853

Tabela 40 – Resultados obtidos para o conjunto Suicide Detection com aprendizado semi supervisionado

Técnica	Cenário	Masc.	Acurácia	Precisão	Recall	F1
LP KNN	PP	80%	0,6837	0,6186	0,9604	0,7524
	PP	60%	0,7539	0,6846	0,9426	0,7931
LP RBF	PP	80%	0,7393	0,6872	0,8792	0,7714
	PP	60%	0,7539	0,7032	0,8793	0,7814
LS KNN	PP	80%	0,7554	0,6879	0,9358	0,7929
	PP	60%	0,7758	0,7099	0,9334	0,8064
LS RBF	PP	80%	0,7496	0,7069	0,8537	0,7733
	PP	60%	0,7621	0,7198	0,8590	0,7832
CCP	PP	80%	0,8143	0,8024	0,8339	0,8178
	PP	60%	0,8256	0,8121	0,8474	0,8294
LP KNN	NP	80%	0,7970	0,7382	0,9211	0,8195
	NP	60%	0,8245	0,7714	0,9229	0,8403
LP RBF	NP	80%	0,8095	0,8113	0,8073	0,8092
	NP	60%	0,8154	0,8201	0,8084	0,8142
LS KNN	NP	80%	0,8212	0,7678	0,9214	0,8376
	NP	60%	0,8343	0,7840	0,9233	0,8479
LS RBF	NP	80%	0,8071	0,8134	0,7975	0,8053
	NP	60%	0,8150	0,8235	0,8024	0,8128
CCP	NP	80%	0,8505	0,8413	0,8642	0,8525
	NP	60%	0,8603	0,8502	0,8749	0,8623

Tabela 41 – PCA Mental Health Twitter pré-processado - Classificação com 80% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 80%	10	0,5110	0,5110	0,4984	0,5045
	20	0,6178	0,6262	0,5847	0,6041
	30	0,6229	0,6304	0,5951	0,6113
	40	0,6281	0,6379	0,5937	0,6141
	50	0,6280	0,6374	0,5952	0,6145
	100	0,6297	0,6345	0,6143	0,6227
	150	0,6318	0,6325	0,6320	0,6307
LS KNN 80%	10	0,5113	0,5112	0,5028	0,5068
	20	0,6196	0,6227	0,6064	0,6142
	30	0,6258	0,6287	0,6144	0,6211
	40	0,6298	0,6331	0,6168	0,6246
	50	0,6305	0,6340	0,6174	0,6253
	100	0,6328	0,6346	0,6260	0,6299
	150	0,6345	0,6350	0,6326	0,6334
CCP 80%	10	0,6113	0,6104	0,6157	0,6130
	20	0,6373	0,6340	0,6500	0,6418
	30	0,6412	0,6374	0,6555	0,6462
	40	0,6427	0,6378	0,6612	0,6491
	50	0,6421	0,6354	0,6675	0,6509
	100	0,6386	0,6287	0,6779	0,6522
	150	0,6290	0,6178	0,6773	0,6460

Tabela 42 – PCA Mental Health Twitter não processado - Classificação com 80% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 80%	10	0,6891	0,7166	0,6261	0,6677
	20	0,7157	0,7463	0,6538	0,6965
	30	0,7244	0,7579	0,6598	0,7050
	40	0,7252	0,7610	0,6571	0,7047
	50	0,7271	0,7621	0,6612	0,7074
	100	0,7309	0,7779	0,6473	0,7058
	150	0,7312	0,7833	0,6403	0,7037
LS KNN 80%	10	0,6908	0,7038	0,6587	0,6802
	20	0,7217	0,7362	0,6910	0,7126
	30	0,7334	0,7475	0,7051	0,7254
	40	0,7365	0,7504	0,7087	0,7287
	50	0,7402	0,7533	0,7143	0,7330
	100	0,7470	0,7650	0,7131	0,7379
	150	0,7480	0,7679	0,7110	0,7381
CCP 80%	10	0,7033	0,71062	0,6864	0,6982
	20	0,7362	0,7375	0,7341	0,7357
	30	0,7464	0,7402	0,7597	0,7497
	40	0,7435	0,7364	0,7593	0,7475
	50	0,7454	0,7392	0,7595	0,7488
	100	0,7564	0,7521	0,7661	0,7586
	150	0,7503	0,7445	0,7641	0,7536

Tabela 43 – PCA Mental Health Twitter pré-processado - Classificação com 60% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 60%	10	0,6037	0,6112	0,5688	0,5891
	20	0,6298	0,6339	0,6140	0,6236
	30	0,6351	0,6378	0,6246	0,6309
	40	0,6398	0,6439	0,6252	0,6342
	50	0,6418	0,6456	0,6289	0,6368
	100	0,6455	0,6458	0,6452	0,6451
	150	0,6461	0,6436	0,6552	0,6489
LS KNN 60%	10	0,6048	0,6097	0,5812	0,5950
	20	0,6312	0,6329	0,6239	0,6282
	30	0,6377	0,6383	0,6351	0,6365
	40	0,6414	0,6426	0,6367	0,6395
	50	0,6427	0,6438	0,6385	0,6410
	100	0,6472	0,6461	0,6511	0,6483
	150	0,6478	0,6451	0,6565	0,6505
CCP 60%	10	0,6200	0,6181	0,6282	0,6230
	20	0,6493	0,6423	0,6742	0,6577
	30	0,6558	0,6484	0,6811	0,6642
	40	0,6580	0,6500	0,6849	0,6669
	50	0,6603	0,6500	0,6946	0,6715
	100	0,6602	0,6462	0,7081	0,6756
	150	0,6540	0,6386	0,7099	0,6722

Tabela 44 – PCA Mental Health Twitter não processado - Classificação com 60% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 60%	10	0,7010	0,7183	0,6611	0,6884
	20	0,7335	0,7492	0,7017	0,7246
	30	0,7471	0,7618	0,7189	0,7396
	40	0,7506	0,7657	0,7221	0,7431
	50	0,7542	0,7690	0,7266	0,7471
	100	0,7611	0,7836	0,7213	0,7510
	150	0,7624	0,7895	0,7158	0,7506
LS KNN 60%	10	0,7015	0,7150	0,6699	0,6916
	20	0,7353	0,7470	0,7113	0,7286
	30	0,7495	0,7598	0,7296	0,7443
	40	0,7545	0,7645	0,7355	0,7496
	50	0,7578	0,7673	0,7398	0,7532
	100	0,7676	0,7821	0,7418	0,7613
	150	0,7684	0,7850	0,7392	0,7613
CCP 60%	10	0,7133	0,7202	0,6979	0,7088
	20	0,7529	0,7529	0,7529	0,7528
	30	0,7663	0,7605	0,7777	0,7689
	40	0,7669	0,7614	0,7777	0,7693
	50	0,7720	0,7668	0,7820	0,7742
	100	0,7810	0,7815	0,7804	0,7808
	150	0,7777	0,7798	0,7742	0,7769

Tabela 45 – PCA Suicidal Data pré-processado - Classificação com 80% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 80%	10	0,8733	0,9032	0,7973	0,8468
	20	0,8735	0,8898	0,8130	0,8495
	30	0,8719	0,8779	0,8232	0,8495
	40	0,8711	0,8726	0,8276	0,8494
	50	0,8696	0,8650	0,8334	0,8488
	100	0,8509	0,8120	0,8608	0,8353
	150	0,8602	0,8359	0,8489	0,8421
LS KNN 80%	10	0,8726	0,8929	0,8071	0,8477
	20	0,8759	0,8914	0,8175	0,8527
	30	0,8748	0,8833	0,8240	0,8525
	40	0,8741	0,8792	0,8273	0,8523
	50	0,8737	0,8737	0,8330	0,8527
	100	0,8594	0,8291	0,8569	0,8425
	150	0,8664	0,8502	0,8451	0,8474
CCP 80%	10	0,8634	0,8771	0,8010	0,8372
	20	0,8679	0,8975	0,7891	0,8398
	30	0,8695	0,8974	0,7933	0,8421
	40	0,8727	0,8927	0,8069	0,8476
	50	0,8687	0,8825	0,8083	0,8437
	100	0,8199	0,7713	0,8382	0,8033
	150	0,7796	0,7168	0,8237	0,7663

Tabela 46 – PCA Suicidal Data não processado - Classificação com 80% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 80%	10	0,8783	0,9149	0,7783	0,8410
	20	0,8909	0,9350	0,8081	0,8668
	30	0,8932	0,9327	0,8160	0,8703
	40	0,8936	0,9268	0,8233	0,8718
	50	0,8943	0,9221	0,8297	0,8733
	100	0,8926	0,9031	0,8469	0,8739
	150	0,8914	0,8953	0,8531	0,8734
LS KNN 80%	10	0,8856	0,9174	0,8133	0,8620
	20	0,8937	0,9228	0,8275	0,8724
	30	0,8962	0,9240	0,8324	0,8757
	40	0,8961	0,9172	0,8395	0,8765
	50	0,8968	0,9152	0,8434	0,8777
	100	0,8963	0,9031	0,8560	0,8788
	150	0,8942	0,8955	0,8596	0,8771
CCP 80%	10	0,8837	0,9197	0,8052	0,8586
	20	0,8904	0,9350	0,8063	0,8658
	30	0,8925	0,9427	0,8038	0,8677
	40	0,8921	0,9321	0,8134	0,8687
	50	0,8796	0,9141	0,8008	0,8536
	100	0,8723	0,9036	0,7938	0,8451
	150	0,8599	0,9021	0,7637	0,8271

Tabela 47 – PCA Suicidal Data pré-processado - Classificação com 60% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 60%	10	0,8761	0,9042	0,8034	0,8506
	20	0,8798	0,9018	0,8153	0,8563
	30	0,8786	0,8939	0,8215	0,8560
	40	0,8790	0,8909	0,8259	0,8571
	50	0,8785	0,8855	0,8310	0,8573
	100	0,8748	0,8710	0,8395	0,8548
	150	0,8731	0,8635	0,8448	0,8539
LS KNN 60%	10	0,8766	0,9017	0,8075	0,8518
	20	0,8803	0,9017	0,8169	0,8571
	30	0,8801	0,8950	0,8239	0,8579
	40	0,8802	0,8920	0,8278	0,8586
	50	0,8799	0,8873	0,8326	0,8590
	100	0,8768	0,8741	0,8409	0,8571
	150	0,8753	0,8675	0,8453	0,8562
CCP 60%	10	0,8672	0,8827	0,8044	0,8417
	20	0,8715	0,8990	0,7968	0,8447
	30	0,8727	0,8974	0,8019	0,8469
	40	0,8752	0,8924	0,8138	0,8512
	50	0,8719	0,8818	0,8177	0,8485
	100	0,8263	0,7775	0,8465	0,8105
	150	0,7820	0,7152	0,8366	0,7711

Tabela 48 – PCA Suicidal Data não processado - Classificação com 60% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 60%	10	0,8810	0,9094	0,7907	0,8458
	20	0,8953	0,9307	0,8231	0,8735
	30	0,8984	0,9319	0,8297	0,8777
	40	0,8991	0,9264	0,8371	0,8794
	50	0,8999	0,9234	0,8423	0,8809
	100	0,8999	0,9111	0,8558	0,8825
	150	0,8988	0,9048	0,8603	0,8819
LS KNN 60%	10	0,8877	0,9226	0,8128	0,8641
	20	0,8961	0,9267	0,8292	0,8752
	30	0,8996	0,9288	0,8356	0,8797
	40	0,9004	0,9241	0,8427	0,8814
	50	0,9008	0,9216	0,8465	0,8823
	100	0,9015	0,9110	0,8598	0,8846
	150	0,9001	0,9055	0,8629	0,8836
CCP 60%	10	0,8865	0,9182	0,8141	0,8629
	20	0,8960	0,9319	0,8233	0,8741
	30	0,8987	0,9393	0,8224	0,8769
	40	0,8984	0,9295	0,8318	0,8779
	50	0,8957	0,9230	0,8319	0,8751
	100	0,8790	0,9011	0,8136	0,8551
	150	0,8641	0,8953	0,7819	0,8347

Tabela 49 – PCA Suicide Detection pré-processado - Classificação com 80% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 80%	10	0,7965	0,7574	0,8731	0,8111
	20	0,7898	0,7357	0,9051	0,8116
	30	0,7715	0,7086	0,9231	0,8017
	40	0,7598	0,6936	0,9316	0,7951
	50	0,7513	0,6838	0,9357	0,7901
	100	0,7278	0,6584	0,9476	0,7769
	150	0,7168	0,6478	0,9517	0,7708
LS KNN 80%	10	0,8053	0,7801	0,8506	0,8138
	20	0,8124	0,7758	0,8792	0,8242
	30	0,8074	0,7608	0,8970	0,8233
	40	0,8029	0,7516	0,9053	0,8213
	50	0,7981	0,7437	0,9102	0,8185
	100	0,7845	0,7228	0,9236	0,8109
	150	0,7771	0,7130	0,9283	0,8065
CCP 80%	10	0,8134	0,8086	0,8212	0,8148
	20	0,8285	0,8278	0,8297	0,8287
	30	0,8240	0,8200	0,8301	0,8250
	40	0,8155	0,8083	0,8272	0,8176
	50	0,7946	0,7811	0,8189	0,7995
	100	0,7033	0,6911	0,7356	0,7125
	150	0,6608	0,6552	0,6796	0,6669

Tabela 50 – PCA Suicide Detection não processado - Classificação com 80% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 80%	10	0,8469	0,8535	0,8379	0,8456
	20	0,8458	0,8276	0,8739	0,8501
	30	0,8413	0,8151	0,8833	0,8478
	40	0,8361	0,8010	0,8946	0,8452
	50	0,8315	0,7923	0,8988	0,8422
	100	0,8174	0,7678	0,9103	0,8330
	150	0,8116	0,7585	0,9148	0,8293
LS KNN 80%	10	0,8512	0,8611	0,8377	0,8492
	20	0,8593	0,8496	0,8734	0,8613
	30	0,8579	0,8415	0,8822	0,8613
	40	0,8557	0,8311	0,8929	0,8609
	50	0,8525	0,8239	0,8968	0,8588
	100	0,8419	0,8017	0,9089	0,8519
	150	0,8371	0,7925	0,9135	0,8487
CCP 80%	10	0,8565	0,8760	0,8307	0,8527
	20	0,8583	0,8655	0,8486	0,8569
	30	0,8522	0,8469	0,8600	0,8533
	40	0,8473	0,8345	0,8663	0,8501
	50	0,8415	0,8245	0,8678	0,8456
	100	0,8171	0,7848	0,8738	0,8269
	150	0,8056	0,7699	0,8715	0,8176

Tabela 51 – PCA Suicide Detection pré-processado - Classificação com 60% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 60%	10	0,8109	0,7851	0,8565	0,8192
	20	0,8157	0,7770	0,8858	0,8278
	30	0,8090	0,7603	0,9028	0,8254
	40	0,8038	0,7504	0,9111	0,8229
	50	0,7994	0,7429	0,9162	0,8205
	100	0,7844	0,7206	0,9296	0,8118
	150	0,7767	0,7105	0,9345	0,8072
LS KNN 60%	10	0,8136	0,7918	0,8513	0,8205
	20	0,8218	0,7890	0,8788	0,8315
	30	0,8180	0,7756	0,8952	0,8311
	40	0,8149	0,7682	0,9024	0,8299
	50	0,8117	0,7616	0,9077	0,8283
	100	0,8004	0,7422	0,9208	0,8219
	150	0,7937	0,7326	0,9256	0,8178
CCP 60%	10	0,8188	0,8144	0,8260	0,8201
	20	0,8351	0,8335	0,8376	0,8355
	30	0,8325	0,8271	0,8409	0,8339
	40	0,8241	0,8147	0,8393	0,8268
	50	0,8083	0,7909	0,8385	0,8139
	100	0,7211	0,7009	0,7718	0,7346
	150	0,6760	0,6614	0,7217	0,6901

Tabela 52 – PCA Suicide Detection não processado - Classificação com 60% dos dados mascarados

Técnica	Comp.	Acurácia	Precisão	Recall	F1
LP KNN 60%	10	0,8552	0,8654	0,8414	0,8532
	20	0,8624	0,8528	0,8762	0,8643
	30	0,8609	0,8440	0,8856	0,8643
	40	0,8584	0,8334	0,8961	0,8636
	50	0,8551	0,8260	0,8999	0,8614
	100	0,8444	0,8038	0,9115	0,8543
	150	0,8396	0,7947	0,9161	0,8511
LS KNN 60%	10	0,8568	0,8676	0,8424	0,8548
	20	0,8664	0,8588	0,8773	0,8679
	30	0,8657	0,8519	0,8856	0,8684
	40	0,8645	0,8430	0,8959	0,8687
	50	0,8620	0,8369	0,8996	0,8671
	100	0,8531	0,8163	0,9114	0,8612
	150	0,8485	0,8073	0,9158	0,8581
CCP 60%	10	0,8604	0,8795	0,8352	0,8568
	20	0,8654	0,8716	0,8571	0,8643
	30	0,8607	0,8562	0,8671	0,8616
	40	0,8592	0,8463	0,8781	0,8619
	50	0,8536	0,8364	0,8792	0,8572
	100	0,8283	0,7946	0,8857	0,8377
	150	0,8153	0,7770	0,8847	0,8273

Tabela 53 – Resultados do Autotreinamento para o conjunto Mental Health Twitter com limiar de 75%

Técnica	Cenário	Masc.	Acurácia	Precisão	Recall	F1
STLR	PP	80%	0,6406	0,6795	0,5318	0,5964
	PP	60%	0,6423	0,7139	0,4746	0,5700
LR	PP	80%	0,6132	0,7403	0,3494	0,4745
	PP	60%	0,6149	0,7748	0,3245	0,4573
STLR	NP	80%	0,7692	0,8205	0,6890	0,7489
	NP	60%	0,7761	0,8485	0,6719	0,7499
LR	NP	80%	0,7530	0,8615	0,6031	0,7095
	NP	60%	0,7614	0,8777	0,6077	0,7181

Tabela 54 – Resultados do Autotreinamento para o conjunto Mental Health Twitter com limiar de 50%

Técnica	Cenário	Masc.	Acurácia	Precisão	Recall	F1
STLR	PP	80%	0,6519	0,6499	0,6578	0,6537
	PP	60%	0,6667	0,6627	0,6785	0,6704
LR	PP	80%	0,6572	0,6552	0,6641	0,6595
	PP	60%	0,67025	0,6672	0,6796	0,6733
STLR	NP	80%	0,7788	0,7819	0,7732	0,7774
	NP	60%	0,7927	0,7958	0,7874	0,7915
LR	NP	80%	0,7837	0,7851	0,7816	0,7833
	NP	60%	0,7973	0,7993	0,7943	0,7967

Tabela 55 – Resultados do Autotreinamento para o conjunto Suicidal Data com limiar de 75%

Técnica	Cenário	Masc.	Acurácia	Precisão	Recall	F1
STLR	PP	80%	0,8884	0,9388	0,7983	0,8627
	PP	60%	0,8926	0,9468	0,8007	0,8675
LR	PP	80%	0,8850	0,9499	0,7789	0,8559
	PP	60%	0,8909	0,9547	0,7884	0,8635
STLR	NP	80%	0,9124	0,9468	0,8486	0,8949
	NP	60%	0,9167	0,9513	0,8543	0,9001
LR	NP	80%	0,9096	0,9554	0,8328	0,8898
	NP	60%	0,9155	0,9583	0,8438	0,8974

Tabela 56 – Resultados do Autotreinamento para o conjunto Suicidal Data com limiar de 50%

Técnica	Cenário	Masc.	Acurácia	Precisão	Recall	F1
STLR	PP	80%	0,8896	0,9003	0,8422	0,8702
	PP	60%	0,8955	0,9057	0,8510	0,8773
LR	PP	80%	0,8892	0,8949	0,8470	0,8703
	PP	60%	0,8964	0,9018	0,8570	0,8788
STLR	NP	80%	0,9156	0,9210	0,8839	0,9020
	NP	60%	0,9197	0,9219	0,8931	0,9072
LR	NP	80%	0,9152	0,9167	0,8874	0,9018
	NP	60%	0,9201	0,9207	0,8947	0,9075

Tabela 57 – Resultados do Autotreinamento para o conjunto Suicide Detection com limiar de 75%

Técnica	Cenário	Masc.	Acurácia	Precisão	Recall	F1
STLR	PP	80%	0,8844	0,9118	0,8513	0,8804
	PP	60%	0,8882	0,9264	0,8435	0,8830
LR	PP	80%	0,8741	0,9324	0,8067	0,8650
	PP	60%	0,8809	0,9390	0,8148	0,8725
STLR	NP	80%	0,9148	0,9378	0,8886	0,9125
	NP	60%	0,9176	0,9455	0,8865	0,9150
LR	NP	80%	0,9076	0,9483	0,8622	0,9031
	NP	60%	0,9133	0,9535	0,8689	0,9092

Tabela 58 – Resultados do Autotreinamento para o conjunto Suicide Detection com limiar de 50%

Técnica	Cenário	Masc.	Acurácia	Precisão	Recall	F1
STLR	PP	80%	0,8869	0,8836	0,8915	0,8875
	PP	60%	0,8927	0,8892	0,8974	0,8932
LR	PP	80%	0,8887	0,8847	0,8939	0,8892
	PP	60%	0,8948	0,8908	0,8999	0,8953
STLR	NP	80%	0,9168	0,9172	0,9166	0,9168
	NP	60%	0,9226	0,9223	0,9230	0,9227
LR	NP	80%	0,9178	0,9181	0,9174	0,9177
	NP	60%	0,9239	0,9239	0,9238	0,92390

DADOS CURRICULARES

IDENTIFICAÇÃO	
Nome completo	Keterly Geovana Gouveia Silva
Data de nascimento	02/11/1999
Naturalidade	Bauru - SP
Nacionalidade	Brasileira
Filiação	Nilcelene Aparecida Gouveia Silva
Nome em citações bibliográficas	SILVA, K. G. G.
Currículo Lattes	http://lattes.cnpq.br/5632584740280536
ORCID	https://orcid.org/0000-0003-4306-9772
FORMAÇÃO ACADÊMICA	
2019–2022	Ciência da Computação, Bacharelado UNISAGRADO – Bauru (SP)
PARTICIPAÇÃO EM EVENTOS CIENTÍFICOS	
XIII Workshop do Programa de Pós-Graduação em Ciência da Computação da UNESP, 2024. Depressão e Suicídio nas Redes Sociais: Uma Abordagem Supervisionada Para Classificação de Textos, 2024. XII Workshop do Programa de Pós-Graduação em Ciência da Computação da UNESP (WPPGCC2023), 2023, Rio Claro, SP. Seleção de Features Para Predição da Síndrome do Ovário Policístico utilizando o algoritmo Particle Swarm Optimization, 2023.	