



INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Classificação de Emoções em Redes Sociais

Soraia Alexandra Cardoso Filipe

Mestrado em Engenharia Informática

Orientadores:

Doutor Fernando Manuel Marques Batista, Professor Associado,

Iscte – Instituto Universitário de Lisboa

Doutor Ricardo Daniel Santos Faro Marques Ribeiro, Professor Associado,

Iscte – Instituto Universitário de Lisboa

Dezembro, 2020



TECNOLOGIAS
E ARQUITETURA

Classificação de Emoções em Redes Sociais

Soraia Alexandra Cardoso Filipe

Mestrado em Engenharia Informática

Orientadores:

Doutor Fernando Manuel Marques Batista, Professor Associado,

Iscte – Instituto Universitário de Lisboa

Doutor Ricardo Daniel Santos Faro Marques Ribeiro, Professor Associado,

Iscte – Instituto Universitário de Lisboa

Dezembro, 2020

Resumo

Este trabalho foca-se na deteção automática de emoções em tweets escritos em língua portuguesa. Com o objetivo de classificar cada um dos tweets com presença ou ausência de cada uma de oito possíveis emoções, são utilizadas duas abordagens. Com vista a avaliar as abordagens propostas, foi criado um conjunto de 1000 tweets, manualmente anotado com a presença ou não das emoções consideradas.

Na primeira abordagem, é aplicado um léxico existente e diferentes estratégias para refinar e melhorá-lo, por meio de tradução automática e incorporação de palavras alinhadas com as existentes. Os resultados sugerem que se pode obter um melhor desempenho tanto através da melhoria de um léxico, como pela tradução direta de tweets portugueses para inglês e depois aplicando um léxico inglês existente.

Relativamente à abordagem supervisionada, pretende-se criar modelos que generalizem melhor, com base em grandes quantidades de informação. Inicialmente é feita a anotação dos tweets disponíveis com base no léxico de emoções. Seguidamente, os primeiros 5 milhões de tweets são usados para treinar um modelo para cada uma das emoções. Esses modelos são então utilizados para anotar 9 milhões de tweets, que por sua vez são filtrados com base na confiança do modelo. Os tweets com uma confiança acima de um dado threshold são usados para treinar modelos finais. Todos os modelos são avaliados com base nos dados de referência, revelando que o segundo modelo mostra em geral um maior sucesso na previsão de emoções do que o primeiro modelo. Esta abordagem apresenta melhores resultados face à primeira.

Palavras chave

Twitter, Língua Portuguesa, Deteção de emoções, Léxicos de emoções, Aprendizagem automática, Regressão Logística

Abstract

This work focuses on the automatic detection of emotions in tweets written in Portuguese. In order to classify each one of the tweets with the presence or absence of each one of eight possible emotions, two approaches are used: the first is based on a emotion lexicon; the second is supervised and is based on logistic regression models. In order to evaluate the two proposed approaches, a set of 1000 tweets was created, manually annotated with the presence or absence of each of the emotions.

Regarding the first approach, an existing and widely used lexicon is applied, as well as different strategies to refine and improve it, by means of machine translation and incorporating additional words. The results obtained suggest that a better performance can be achieved both by improving a lexicon and by directly translating Portuguese tweets into English and then applying an existing English lexicon.

Regarding the supervised approach, it was intended to create models that generalize better, based on large amounts of information. Initially the available tweets are annotated based on the emotion lexicon. Then, the first 5 million tweets are used to train a model for each of the emotions. These models are then used to annotate another 9 million tweets, which in turn were filtered based on the model's confidence. Tweets with a confidence above a given threshold are used to train final models. All models were evaluated based on the reference data, revealing that the first model generally shows greater success in predicting emotions. This approach presents better results than the previous one.

Keywords

Twitter, Portuguese Language, Emotion detection, Emotion lexicon, Machine learning, Logistic Regression

Agradecimentos

Enquanto percorria a dissertação, fui assaltada por um turbilhão de emoções... e, quase que vi o falecido Steve Jobs a acenar o dedo bem no meio da minha cara: “I’m convinced that about half of what separates successful entrepreneurs from the non-successful ones is pure perseverance.”. Ri-me, mas sem saber se havia de rir ou chorar. Sim, estou certa que a classificação emocional das próximas páginas estão carregadas de emoção. Mas não foi de mim que me lembrei...

A grande fonte de perseverança foram os meus professores, Fernando Batista e Ricardo Ribeiro — verdadeiros pilares. Devo-lhe estas páginas, o conhecimento pessoal e profissional que me passaram e o crescimento que daí adveio. Deram-me confiança quando a mesma faltou (que não foram poucas vezes...). Motivaram-me sempre. Incentivaram e animaram-me com palavras e gestos. Um muito obrigada não chega, pois estou convicta que em nenhuma altura o fizeram por obrigação! Sinto-me sobretudo agradecida pela infinita paciência, que nunca lhes faltou! Fizeram com certeza que esta experiência fosse valiosa e inigualável, de modo a que nunca poderia ser trocada por outra qualquer..

Não posso também deixar de agradecer àqueles que estiveram sempre ao meu lado: à minha família e aos meus amigos, pela paciência, apoio e compreensão. Obrigada por nunca se queixarem e estarem sempre lá!

Um sentido agradecimento aos que se disponibilizaram, perdendo o seu precioso tempo, para me ouvir e ajudar na anotação dos tweets.

E, por fim, agradeço, também, à minha segunda casa: ISCTE - Instituto Universitário de Lisboa. Não podendo esquecer, também, a minha terceira casa: INESC-ID (Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento de Lisboa).

Lisboa, 4 de Dezembro de 2020

Soraia Filipe

Conteúdo

1	Introdução	1
1.1	Enquadramento	1
1.2	Motivação e Relevância	2
1.3	Questões de Investigação	2
1.4	Objetivos	3
1.5	Abordagem Metodológica	3
1.5.1	Abordagens	3
1.5.2	Métricas de Avaliação	4
1.5.3	Coeficientes de Fiabilidade	4
1.6	Contribuições	5
1.7	Estrutura e Organização da Dissertação	5
2	Revisão da Literatura	7
2.1	Sobre as Emoções	7
2.2	Complexidade da Expressão de Emoções em Texto	8
2.3	Deteção de Emoções	8
3	Dados	15
3.1	Modelo e Descrição de Emoções	15
3.2	Descrição dos Dados	15
3.3	Anotação Manual de <i>Tweets</i>	17
3.4	Pré-processamento	19
4	Abordagem Baseada em Léxicos	23
4.1	Arquitetura do Classificador	23
4.2	Léxicos	24
4.2.1	Léxico NRC Original	24
4.2.2	LEX-I — Google Tradutor e DeepL	25
4.2.3	LEX-II — MUSE	25
4.3	Resultados	27
5	Abordagem Baseada em Aprendizagem Automática	35
5.1	Classificação Automática do <i>Dataset</i>	35
5.2	Classificador Baseado em Aprendizagem Supervisionada	37

5.3 Regressão Logística	39
5.4 Análise do Impacto do Tamanho do Conjunto de Treino e do Número de <i>Features</i>	40
5.5 Resultados	40
6 Conclusões e Trabalho Futuro	43
Bibliografia	47

Lista de Figuras

3.1	Roda de Emoções de Plutchik.	16
3.2	Distribuição de <i>tweets</i> por número de palavras.	17
3.3	Exemplos de <i>tweets</i> anotados.	17
4.1	Esquema geral do Classificador.	23
5.1	Número de <i>tweets</i> classificados por emoção.	36
5.2	Número de <i>tweets</i> por emoção e por número de emoções.	36
5.3	Porção de <i>tweets</i> classificados automaticamente.	37
5.4	Esquema das fases de treino da aprendizagem supervisionada.	38

Lista de Tabelas

1.1	Matriz de Confusão.	4
3.1	Descrição das emoções utilizadas.	16
3.2	Coeficientes de concordância entre as anotações manuais.	18
3.3	Número de <i>tweets</i> classificados, por emoção, pelos anotadores.	18
3.4	Exemplos de bigramas.	19
3.5	Exemplos de <i>stemming</i> e lematização, para a língua portuguesa.	20
3.6	Exemplos de <i>stemming</i> e lematização, para a língua inglesa	20
4.1	Exemplo do léxico NRC Word-Emotion com tradução em português.	24
4.2	Exemplo do léxico melhorado.	25
4.3	Excerto do LEX-II, utilizando o MUSE como recurso de tradução.	26
4.4	Métricas de avaliação, considerando a anotação manual 1.	28
4.5	Métricas de avaliação, considerando a anotação manual 2.	29
4.6	Métricas de avaliação das experiências, utilizando a primeira referência. . . .	30
4.7	Métricas de avaliação das experiências, utilizando a segunda referência. . . .	31
4.8	Métricas de avaliação das experiências, utilizando a terceira referência. . . .	32
4.9	Coeficientes de fiabilidade para a classificação.	33
5.1	Combinações de diferentes quantidades de <i>tweets</i> e número de <i>features</i>	40
5.2	Métricas de avaliação do Modelo 1.	41
5.3	Métricas de avaliação do Modelo 2.	41

Capítulo 1

Introdução

A sobrevivência do ser humano encontra-se diretamente interligada às emoções. Por vezes, as emoções humanas são pouco valorizadas, tendo o pensamento lógico e a razão maior peso. As emoções são exteriorizadas através das expressões faciais, da voz, da escrita e dos sinais fisiológicos, como a respiração, transpiração e aumento do ritmo cardíaco.

Este Capítulo explica o tema trabalhado e a sua importância, as questões e os objetivos que o guiam, como é abordada e avaliada a tarefa desenvolvida, e como é que o restante documento está delineado.

1.1 Enquadramento

A rede social *Twitter* é mundialmente conhecida e utilizada, sendo gerados todos os dias uma grande quantidade de publicações. As suas publicações, conhecidas como *tweets*, são particularmente interessantes para a comunidade científica por abrangerem todos os assuntos imagináveis, desde a política à partilha do dia-a-dia ou de desabafos. Maioritariamente, são partilhas menos pensadas e de momento, fator de genuinidade precioso quando o que se pretende é espelhar emoções reais. Estes elementos e a possível aquisição gratuita dos dados fez com que fossem os eleitos para esta dissertação. No entanto, os *tweets* consistem em texto não estruturado, com limite de caracteres e recorrência a, por exemplo: abreviaturas; *hashtags*; *emojis*; gíria; *emoticons*; erros de ortografia, sintaxe e semântica.

Neste trabalho, procura-se identificar e analisar as emoções básicas implícitas num determinado *tweet*, recorrendo a abordagens de Processamento de Língua Natural (PLN), para compreender e retirar significado do texto. Existem diversas escolhas no que diz respeito ao processamento e pré-processamento de dados, muitas das quais têm origem em PLN. Em alternativa, é possível adotar algoritmos já criados na área da Inteligência Artificial para o mesmo propósito ou, até mesmo, combinar diferentes métodos para obter melhores resultados.

1.2 Motivação e Relevância

A popularização da Internet facilitou a comunicação entre a população. A sua evolução, ao longo dos anos e a crescente necessidade dos indivíduos se manterem cada vez mais conectados e com maior rapidez, deu origem às redes sociais.

As redes sociais apresentam um rápido desenvolvimento e inovação, satisfazendo as necessidades de socialização dos seres humanos e criando novas necessidades, o que levou a uma adesão por um grande número de pessoas que as usa diariamente. Quer por uma questão de moda, conveniência ou, simplesmente, curiosidade, a aderência a estas redes continua a aumentar cada vez mais.

Atualmente, este fenómeno é de extrema importância uma vez que é eleito pelos seus utilizadores como forma primária de gerar e procurar conhecimento, explorar interesses, e canal de marketing e publicidade, entre outros. Torna-se assim uma das maiores ferramentas para influenciar e ganhar visibilidade entre a população, com impacto em todas as áreas da sociedade, como a política e a educação.

Estas plataformas são palco para exprimir opiniões, hábitos, costumes e gostos. Representam uma enorme amostra e diversificação da população, bem como uma grande fonte de dados, o que demonstra o imenso potencial para o estudo da população.

A deteção e análise de emoções sobre este conteúdo expõe o parecer global dos indivíduos em relação a diferentes aspetos da vida em sociedade.

Esta informação é relevante na medida em que permite perceber como chegar ao público alvo, traçar perfis psicológicos, detetar preferências, tendências (suicidas, *bullying*, moda, etc), entre muitas outras.

1.3 Questões de Investigação

A deteção de emoções é uma tarefa bastante atual que pode ser realizada com recurso a várias abordagens, tais como o uso de variadíssimos algoritmos de aprendizagem automática.

As questões de investigação que este trabalho pretende ver respondidas são as seguintes:

- Como extrair as emoções associadas à informação não estruturada presente nos *tweets*?
- É possível fazer esta extração utilizando léxicos de emoções?
- Qual o desempenho das abordagens baseadas em léxicos quando aplicadas à língua portuguesa?

- Qual o desempenho de uma abordagem baseada em aprendizagem automática tradicional no mesmo contexto?
- Quais são as emoções mais difíceis de identificar? E as mais fáceis?

1.4 Objetivos

Os principais objetivos deste trabalho são:

- Criar uma base de dados de *tweets* escritos em língua portuguesa e organizada por país;
- Desenvolver um classificador que, com base num léxico de emoções, identifique os tipos de emoções presentes em dado *tweet*;
- Treinar um novo classificador supervisionado, baseado num algoritmo de aprendizagem automática e com dados previamente anotados automaticamente;
- Anotar manualmente um conjunto de *tweets* que permita fazer uma avaliação dos classificadores desenvolvidos;
- Perceber qual abordagem favorece a identificação de emoções.

1.5 Abordagem Metodológica

A estratégia seguida para atingir com sucesso os objetivos centra-se em diversos métodos que são estudados, trabalhados, discutidos e necessários para responder às questões de investigação deste trabalho.

Os dados que constituem o *dataset* utilizado são retirados e filtrados, para que se encontrem escritos em língua portuguesa e com o país de origem associado, através da API (*Application Programming Interface*) do *Twitter*.

1.5.1 Abordagens

A tarefa principal por detrás dos objetivos mencionados na Secção 1.4 é a classificação de *tweets* a qual é, após o pré-processamento, abordada de duas formas diferentes.

Numa primeira abordagem é utilizado um léxico de emoções para perceber a relação entre as palavras dos *tweets* e as emoções estudadas.

Na abordagem baseada em aprendizagem automática o foco é a tarefa de classificação de texto, pretende-se saber se um texto tem ou não uma determinada emoção, o que se traduz em oito tarefas de acordo com o modelo de emoções usado.

		Valores Previstos	
		Positivo	Negativo
Valores Reais	Negativo	<i>Verdadeiros Positivos</i>	<i>Falsos Negativos</i>
	Positivo	<i>Falsos Positivos</i>	<i>Verdadeiros Negativos</i>

Tabela 1.1: Matriz de Confusão.

De modo a realizar este estudo, foi anotado manualmente, após o devido enquadramento, conhecimento e percepção do tema e da tarefa pretendida, um *dataset* de *tweets* de acordo com as emoções do modelo para conseguir validar os classificadores criados.

1.5.2 Métricas de Avaliação

Os classificadores elaborados nos Capítulos 4 e 5 refletem as abordagens mencionadas na Subsecção anterior e são avaliados através de quatro métricas: taxa de acerto, reflete quantas classificações (de ambas as classes) acertou o modelo — desempenho geral do classificador; precisão, demonstra quantos *tweets* foram classificados com emoção erradamente; cobertura, expõe quantos *tweets* foram classificados com emoção corretamente; e medida f1, que apresenta a média harmónica das duas anteriores. Estas são calculadas, utilizando a matriz de confusão apresentada na Tabela 1.1, através das seguintes formulas, em que *VP* corresponde a *verdadeiros positivos*, *VN* corresponde a *verdadeiros negativos*, *FP* corresponde a *falsos positivos* e *FN* corresponde a *falsos negativos*:

$$Taxa\ de\ acerto = \frac{VP + VN}{VP + VN + FP + FN} \quad (1.1)$$

$$Precisão = \frac{VP}{VP + VN} \quad (1.2)$$

$$Cobertura = \frac{VP}{VP + FN} \quad (1.3)$$

$$Medida\ f1 = 2 \times \frac{Precisão \times Cobertura}{Precisão + Cobertura} \quad (1.4)$$

f1

1.5.3 Coeficientes de Fiabilidade

Como anteriormente mencionado, foram anotados manualmente por duas pessoas 1000 *tweets* com o propósito de avaliar os trabalhos desenvolvidos no decorrer desta dissertação. De forma a perceber e analisar a concordância entre os anotadores e, tenho em conta que a combinação destas anotações é a base da referência para validar as abordagens,

retirar algumas conclusões adicionais sobre os resultados dos classificadores desenvolvidos, utilizam-se três coeficientes de concordância para dados nominais: percentagem média; Kappa de Cohen e Alpha de Krippendorff, calculadas através da ferramenta ReCal2¹. Tendo em conta que a probabilidade de fiabilidade observada (PF_O), a probabilidade de fiabilidade expectável (PF_E), o desacordo observada (D_O) e o desacordo expectável (D_E), as formulas concetuais dos últimos dois coeficientes são [25]:

$$Kappa\ de\ Cohen = \frac{PF_O - PF_E}{1 - PF_E} \quad (1.5)$$

$$Alpha\ de\ Krippendorff = 1 - \frac{D_O}{D_E} \quad (1.6)$$

Também são utilizados coeficientes de fiabilidade na Capítulo 4, com o objetivo de comparar várias experiências, a sua concordância e fazer uma análise de conteúdo. Uma vez que as experiências são mais do que duas, apenas se pode utilizar o coeficientes Alpha de Krippendorff. Os outros coeficientes utilizados são a média dos anteriores: percentagem média de acordo entre pares e média de Kappa de Cohen entre pares. Os três coeficientes são calculados com recurso à ferramenta ReCal3².

1.6 Contribuições

Esta dissertação fornece à sociedade científica possíveis soluções para a resolução da tarefa de deteção de emoções em textos portugueses característicos de redes sociais. Abordam-se dois métodos de extração de conhecimento diferentes e possíveis técnicas de melhoria dos mesmos ou de outros recursos semelhantes. É, também, anotado um *dataset* de *tweet* para o português - recurso raro e de valor. Em suma, representa uma boa base para trabalhos semelhantes e, mais importante, para uma possível construção de um léxico de emoções associado à língua portuguesa.

É de salientar a publicação do artigo *Different Lexicon-Based Approaches to Emotion Identification in Portuguese Tweets* no SLATE 2020³, referente ao trabalho elaborado no Capítulo 4.

1.7 Estrutura e Organização da Dissertação

Esta dissertação encontra-se organizada em seis Capítulos:

¹ReCal2: <http://dfreelon.org/utis/recalfront/recal2/>

²ReCal3: <http://dfreelon.org/utis/recalfront/recal3/>

³SLATE 2020: <https://drops.dagstuhl.de/opus/volltexte/2020/13012/pdf/oasics-vol083-slate2020-complete.pdf>

- Introdução;
- Revisão da Literatura;
- Dados;
- Abordagens Baseadas em Léxicos;
- Abordagem Baseada em Aprendizagem Automática;
- Conclusões e Trabalho Futuro.

Após a introdução do tema, visadas as questões de investigação, os objetivos a alcançar e a estratégia pensada, é apresentada no segundo Capítulo uma revisão da literatura sobre as emoções, definição e modelos; a complexidade e dificuldade dos indivíduos em expressarem e reconhecerem emoções em texto; e alguns trabalhos no âmbito deste tema ou relevantes para o mesmo, deste a criação de léxicos às técnicas de PLN e aos diferentes algoritmos utilizados.

O Capítulo 3 apresenta o modelo de emoções, a descrição do *dataset* utilizado para desenvolver este estudo, o processo e as dificuldades das anotações de *tweets* realizadas manualmente e a descrição do pré-processamento dos dados.

O Capítulo 4 explica o funcionamento do classificador com base em léxico, as diversas experiências feitas com diferentes léxicos e os resultados obtidos.

O Capítulo 5 ilustra a classificação automática de *tweets* no total dos dados utilizando a abordagem do Capítulo anterior, descreve o classificador com base num algoritmo de aprendizagem automática, a razão por detrás da escolha e funcionalidade do mesmo, e os resultados da classificação obtida.

Por último, são apresentadas as conclusões destes trabalhos e das questões de investigação pretendidas, as curiosidades e aspetos promissores que se traduzem num possível trabalho futuro.

Capítulo 2

Revisão da Literatura

A classificação de emoções em *tweets* é uma tarefa bastante interessante para a comunidade científica, no campo da informática, que se dedica à computação. Pode ser aplicada a todas as áreas existentes, fazendo uma ponte de ligação com outros possíveis interesses dos autores, como o comportamento humano, e ainda beneficiar a comunidade tanto a nível de mercado, como a nível pessoal.

Na literatura científica diversos autores exploram e expõem como abordam este tema, quais as técnicas que utilizam para resolver a tarefa de classificação, os problemas presentes na sua execução e os resultados obtidos.

As duas estratégias de classificação utilizadas são as seguintes: a construção ou aplicação de léxicos e a aplicação de diversos algoritmos de aprendizagem automática supervisionada.

Este Capítulo pretende elucidar o que são emoções, as dificuldades do estudo das mesmas em texto e apresentar brevemente alguns trabalhos relacionados, que serviram quer como inspiração quer como fonte de conhecimento.

2.1 Sobre as Emoções

Para um melhor entendimento deste trabalho apresentamos uma breve introdução sobre a definição e alguns modelos de emoções em psicologia.

As emoções são definidas, em psicologia, como estados que refletem a avaliação de julgamentos dos agentes sociais, incluindo o próprio, e do ambiente, tendo em consideração os objetivos (estados desejáveis) e crenças (conhecimento atual) dos organismos, que incentivam e coordenam a adaptação do comportamento. Encontram-se geralmente categorizadas como emoções básicas ou fundamentais, e emoções não básicas ou complexas. Estas últimas são difíceis de classificar e incluem emoções complexas como: orgulho, vergonha, culpa, amor, empatia, humilhação e desprezo, entre outras [14].

Não existe nenhum modelo de emoções universalmente aceite. No entanto, uma grande parte dos artigos referentes a detecção de emoções, que utiliza uma abordagem de classificação de múltiplas classes, baseia-se no modelo de Ekman que contabiliza seis emoções básicas: alegria, tristeza, raiva, medo, nojo e surpresa [27]. Outro modelo bastante usado é o de Plutchik, que combina pares adjacentes de emoções básicas e considera: alegria, confiança, medo, surpresa, tristeza, nojo, raiva e antecipação [30].

Esta dissertação só irá considerar as emoções básicas de Plutchik.

2.2 Complexidade da Expressão de Emoções em Texto

Identificar emoções sem ambiguidade é uma tarefa bastante difícil, mesmo na vida quotidiana, devido à sua complexidade.

Ben-Ze'ev [6] defende que esta complexidade advém da alta sensibilidade das emoções a fatores contextuais e pessoais; do fato das emoções aparecerem em conjuntos de atitudes emocionais, em vez de isoladas; e também porque os termos emocionais são usados linguisticamente de forma confusa.

Bazzanella [5] alega que a complexidade das emoções é visível a diversos níveis como: relação integrada com a mente, língua, comportamento e cultura; problemas lexicais e semânticos; número de correlações fisiológicas e particularidades neurológicas; universalidade ou relatividade das emoções.

De acordo com Lee [16], as emoções explícitas detetam-se no texto por meio de palavras-chave, podendo ser as próprias emoções ou palavras relacionadas com estas. As emoções implícitas estão presentes no texto através de informação inerente e indireta, sendo subentendidas.

Devido à dificuldade em transmitir o conhecimento necessário à máquina para identificar as pistas das emoções implícitas, a grande maioria das pesquisas de técnicas de PLN foca-se em detetar emoções explícitas no texto procurando verbos, advérbios e adjetivos das emoções. No entanto, a maior parte da população expressa as suas emoções de forma implícita. Na tentativa de ultrapassar este problema, Cambria et al. [10] propôs o Sentic Computing que retira partido do senso comum e trata a categorização afetiva, assentando na ideia de que a mente humana consiste em quatro esferas emocionais independentes com diferentes níveis de ativação.

2.3 Detecção de Emoções

Qualquer método de detecção de emoções em texto requer enormes quantidades de dados previamente anotados, tanto para desenvolvimento como para treino dos algoritmos em

aprendizagem automática. Isto consiste num problema, visto tais recursos serem raros. No entanto, é possível recorrer a abordagens de PLN tratando a deteção de emoções como um problema de classificação de múltiplas classes. Mohammad [19] comprova que *features* com base em léxicos de emoções melhoram e superam o desempenho de um classificador supervisionado comparativamente ao uso de unigramas.

Um léxico bastante conhecido e utilizado nesta área é o NRC Word-Emotion Association Lexion. Este foi desenvolvido manualmente por Mohammad e Turney [22], através de *crowdsourcing*, e associa cerca de 14 mil palavras em inglês a oito emoções básicas (alegria, tristeza, raiva, medo, nojo, surpresa, antecipação e confiança) e dois sentimentos (positivo e negativo). Também conhecido por EmoLex, encontra-se disponibilizado gratuitamente e em mais de 100 idiomas¹[22, 21].

O *Linguistic Inquiry and Word Count* (LIWC) é um analisador de texto que foi melhorado, pelos autores, ao longo dos anos [17, 29, 28]. A versão mais recente, LIWC2015², apresenta uma taxa de acerto superior às anteriores e detém uma gama maior de perceções sociais e psicológicas. A análise ao texto informa quantas palavras revelam os pensamentos do autor, sentimentos, personalidades e motivações, detetando também a percentagem de diferentes emoções, preocupações sociais e *part-of-speech* (POS). É composto por três sistemas de dicionários internos, o principal conta com cerca de 6400 palavras, os seus *stemmers* e as emoções selecionadas.

Soares et al. [35] desenvolveram uma adaptação do *Affective Norms for English Words* (ANEW) [8] para o português europeu com base na avaliação de normas afetivas recorrendo a 958 alunos do ensino superior nativos desta língua. Este estudo apresenta resultados bastante semelhantes à versão original e à adaptação espanhola [32]. Concluíram que existem especificidades socioculturais no modo como as palavras do recurso original são percebidas em todas as línguas e culturas. No entanto, contrariamente às duas versões mencionadas, obtiveram diferenças significativas para a valência das palavras e para as dimensões de excitação afetivas, entre o sexo feminino e masculino.

Alguns trabalhos nesta área focam-se em perceber qual algoritmo de aprendizagem automática supervisionada produz, em geral, melhores resultados para a tarefa de deteção de emoções, se varia consoante o tipo de dados, de emoção para emoção ou mesmo de métrica para métrica. Também são apresentados alguns trabalhos que seguem estas abordagens para a análise de sentimentos.

Regra geral, a análise de sentimentos é utilizada para a classificação dos sentimentos: positivo, negativo e neutro. No entanto, esta tarefa pode ser reforçada para fornecer informações mais ricas e significativas se, também, abranger-se os extremos: muito positivo e muito negativo. Para tal, Almatarneh e Gamallo [1] constroem um léxico com valores diferentes para positivo e muito positivo, negativo e muito negativo. De seguida, e através de

¹NRC Word-Emotion Association Lexion disponível em: <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

²LIWC disponível em: <http://liwc.wpengine.com/>

uma abordagem não-supervisionada são classificados textos com base neste léxico. Os resultados obtidos pelo método de classificação são baixos porque a diferença entre extremo e não extremo é muito subjetiva, aumentando a dificuldade da tarefa. Contudo, esta estratégia tem melhores resultados na identificação dos extremos do que os léxicos utilizados para a sua avaliação: SO-CAL e SentiWords.

Musto, Semeraro e Polignano [24] implementam quatro versões (básico, normalizado, enfatizado e enfatizado-normalizado) de uma abordagem baseada em léxicos (SentiWordNet [4], WordNet-Affect [36], MPQA [40] e SenticNet [9]) para a classificação de sentimentos em *tweets*. Cada *tweet* é dividido em micro-frases, utilizando para tal as pontuações, advérbios e conjunções encontradas nos mesmos. O sentimento é obtido através da soma das polaridades das micro-frases de cada *tweet*, apenas invertendo a polaridade em caso de negação. A avaliação foi feita com dois *datasets*. No primeiro *dataset* de treino os melhores resultados, tendo em conta a taxa de acerto, usaram os léxicos SentiWordNet (58,99%) e MPQA (58,25) e a implementação com as versões normalizado e enfatizado. No segundo *dataset* de treino o melhor desempenho foi de 74,65% de taxa de acerto e teve por base a aplicação do léxico SenticNet, não havendo grande diferença nos resultados com as quatro versões do classificador.

Anisa, Winarsih e Supriyanto [2] utilizaram *Naive Bayes*, *J48*, *K-Vizinhos Mais Próximos* e *Máquina de Vetores de Suporte-Otimização Mínima Sequencial* para identificar seis emoções (raiva, repugnância/nojo, medo, alegria, tristeza e surpresa) em textos da Indonésia. O algoritmo com melhor desempenho é o *Máquina de Vetores de Suporte-Otimização Mínima Sequencial*, com uma média de 80,78% de medida *f*.

Asghar et al. [3] compararam algoritmos de aprendizagem supervisionada - *Naive Bayes*, *Árvore de Decisão*, *Máquina de Vetores de Suporte*, *Regressão Logística* e *K-Vizinhos Mais Próximos*, com o objetivo de perceber qual obtém melhores resultados na classificação de cinco emoções (alegria, medo, tristeza, vergonha e culpa) em texto online. O algoritmo recomendado é a *Regressão Logística* com uma média *f1* de 66%, *K-Vizinhos Mais Próximos* é o que tem o pior *f1* com 57%.

Existem diversos trabalhos que combinam, de diferentes maneiras, as duas técnicas de forma a retirar melhor partido de ambas e obter melhores resultados.

BrainT utiliza o *Perceptrão* [13], com uma abordagem de múltiplas classes, de forma a classificar as emoções implícitas em *tweets*. Este trabalho comprova os benefícios de relacionar as emoções com conjuntos de palavras e do uso de técnicas de PLN como bigramas, trigramas e *POS*.

Os textos *online* por serem informais apresentam características peculiares, como *emojis*, *emoticons* (":)", ";)", ":b") e *hashtags* ("#sarcasmo" (ironia), "#entusiasmo" (alegria)), para um melhor entendimento do que se pretende expressar. Estas podem ser utilizadas como *features* individuais ou complementares na identificação de emoções ou sentimentos.

Wang et al. [38] produziram um *dataset* (cerca de 2,5 milhões de *tweets*) de dados

anotados, relacionando os *hashtags* às emoções. Esta relação é feita usando o trabalho de Shaver et al. [33], onde os autores organizam as emoções em duas camadas. A primeira contabiliza seis emoções básicas (raiva, tristeza, amor, alegria, surpresa e medo) e a segunda vinte e cinco emoções secundárias, que são subcategorias das emoções da primeira camada. Cada emoção tem uma lista de palavras associadas à mesma. Wang et al. [38] expandem estas listas de palavras, conseguindo associar 131 palavras usadas em *hashtags* às 7 emoções básicas consideradas – as 6 mencionadas atrás mais gratidão. De forma a aumentar a qualidade dos dados apenas mantém os *tweets* considerados relevantes, ou seja, *tweets* com *hashtags* no final, mais de cinco palavras, não contendo URL ou citações e com menos de quatro *hashtags*. Exploraram uma variedade de recursos para encontrar características, como por exemplo: n-gramas, léxicos de emoções, POS, entre outros, utilizando dois algoritmos, *LIBLINEAR* e *Multinomial Naive Bayes*. Os melhores resultados obtidos combinam o uso de unigramas, bigramas, léxico e características POS. A maior percentagem de precisão obtida é de 65,57%, utilizando cerca de 2 milhões de *tweets* para treino.

Wang e Zhou [37] propuseram um modelo para prever a intensidade das emoções em *tweets*. Para tal, começaram por "limpar" os *tweets* retirando *hashtags* e referências a *usernames*, entre outros. De seguida construíram um léxico de características e procuraram combinações ótimas, para produzir um vetor final de modo a representar um *tweet*. Por último, treinaram um modelo de regressão recorrendo a Redes Neurais.

Mohammad e Kiritchenko [20] construíram um léxico que associa palavras com *hashtags* a emoções, através de um *dataset* de *tweets* etiquetados com emoções, e um classificador baseado no algoritmo Máquina de Vetores de Suporte-Otimização Mínima Sequencial, que demonstra resultados positivos na deteção de emoções básicas e na melhoria da taxa de acerto de classificadores de emoções.

Emoji2emotion é um método proposto por Rakhmetullina, Trautmann e Groh [31] que estabelece a correspondência entre *emojis* e possíveis classes relacionadas, com o objetivo de detetar sentimentos ou emoções, e está disponível online³. As emoções básicas consideradas são as oito de Plutchik [30]. Neste trabalho foi selecionada uma lista dos *emojis* mais recorrentes no *Twitter* e no texto online em geral, recorrendo ao projeto Novak et al. [26] e ao Emojitracker⁴, um site que mostra os *emojis* utilizados em tempo real no *Twitter*. Para obterem um *dataset* de qualidade, recolheram 500 *tweets* com as seguintes características: sem URL, *tags* ou *hashtags*; de cinco a quinze palavras; não mais que duas palavras em maiúsculas; com palavras lematizáveis; e sem certas palavras chave. As anotações foram feitas manualmente por três pessoas, através de uma interface web desenvolvida pelos autores, onde para cada *tweet* definiam a emoção ou emoções correspondentes (incluindo nenhuma) e o valor de intensidade de cada emoção entre 1-3. Metade dos *tweets* anotados revelam apenas uma emoção e a outra metade até quatro emoções. Utilizando os dados

³Emoji2emotion disponibilizado em: <https://github.com/Aisulu/emoji2emotion>.

⁴Emojitracker: <http://www.emojitracker.com/>

anotados, analisou-se a percentagem de ocorrência de *emojis* por emoção e vice-versa. O mapeamento emoji2emotion resulta de testes condicionais para todas as combinações possíveis de emoções e *emojis*. Para testar o modelo utilizaram os algoritmos: *Stochastic Gradient Descent*, *Naive Bayes*, *Random Forest* e K-Vizinhos Mais Próximos. Devido ao desequilíbrio de dados entre emoções, apenas foram testadas quatro emoções (raiva, alegria, tristeza e surpresa). K-Vizinhos Mais Próximos é o algoritmo que apresenta melhor cobertura (86,73%), mas a nível de medida f1 o melhor resultado é através do *Stochastic Gradient Descent* (23,56%).

Relativamente à classificação de emoções em textos escritos em português, Silva Fernando, Carvalho Ariadne e Roman Norton [34] descrevem a metodologia e as duas ferramentas que usaram para anotar manualmente emoções em *tweets* escritos em português do Brasil, relacionados com a bolsa de valores. Cada *tweet* é anotado de acordo com a Teoria da Roda de Plutchik [30] e pode ter até quatro emoções ou ser neutro, isto porque as emoções são consideradas em pares opostos (alegria ou tristeza, raiva ou medo, nojo ou surpresa, antecipação ou confiança).

O EmoSpell é uma extensão do analisador morfológico português Jspell que utiliza como dicionário o do Jspell e as bases lexicais EMOTAIX.PT e SentiLex-PT, com o propósito de aumentar o poder de reconhecimento do EMOTAIX.PT [18]. Os resultados mostram que o EmoSpell encontra mais palavras emocionais nos textos do que o EMOTAIX.PT e deteta a polaridade dos mesmos. O SentiLex-PT, em alguns casos, tem um desempenho superior ao EmoSpell. No entanto, o EmoSpell deteta tantas palavras emocionais, em média, como o SentiLex-PT, fornece mais informação sobre estas e deteta a emoção dominante presente nos textos.

Moraes et al. [23] compararam abordagens baseadas em léxico e aprendizagem automática para perceber qual define melhor a subjetividade dos *tweets* portugueses e apresenta melhores resultados na classificação da polaridade. As experiências foram realizadas utilizando os léxicos SentiLex e WordAffectBR, e os algoritmos *Sequential Machine Optimization* e *Naive Bayes*. O algoritmo de *Sequential Machine Optimization* apresenta a melhor medida f1 de 71,3%, enquanto que os léxicos conseguem obter uma medida f1 de 64%. Assim sendo, os autores defendem que uma abordagem de aprendizagem automática é a melhor opção.

Para prever a personalidade dos utilizadores portugueses do *Twitter*, Jusupova, Batista e Ribeiro [15] construíram uma plataforma baseada em léxico focada no modelo dos Cinco Grandes Traços de Personalidade (extroversão, amabilidade, conscienciosidade, neuroticismo e abertura). Esta plataforma analisa o perfil do autor do *tweet* e, com base em informações públicas, fornece o tipo de personalidade do mesmo, mostra a análise de sentimentos do *tweet*, a visualização temporal de atividades do utilizador, entre outras.

Finalmente, Duarte, Macedo e Oliveira [12] reportam um estudo onde exploram *emojis* para o reconhecimento de emoções em textos portugueses.

Conclui-se que, para a tarefa de detecção de emoções, as abordagens com base em léxicos de emoções e com base em aprendizagem supervisionada permitem resultados bastante positivos. O pré-processamento é utilizado, pelos autores, para clarificar a informação e facilitar a extração de *features* do texto. Alguns exemplos mais recorrentes da limpeza de dados, com base nos artigos mencionado são: remover o símbolo *hashtags*, as pontuações (exceto "!", "?") e os caracteres numéricos; substituir as *tags* de utilizador por *username*; separar as abreviações; substituir todos os caracteres em letra minúscula. Outras decisões de pré-processamento mais utilizadas são tokenização, remover *stop words*, *stemming*, n-gramas e POS. Pode afirmar-se que os léxicos e analisadores apresentam uma boa sugestão em conjunto com algumas destas técnicas de PLN como recurso para adicionar contexto e extrair características importantes dos textos ou até mesma em substituição de algumas delas (exemplo: n-gramas), aplicando posteriormente um classificador supervisionado.

No âmbito de textos em redes sociais e de forma a melhorar os resultados da abordagem ou abordagens escolhidas, os autores para além do pré-processamento também mencionam a possibilidade de explorar outras características, para além das palavras: *hashtags*, *emojis* e *emoticons*, bastante comuns e com contexto emocional nestes textos informais e não estruturados.

Tendo em conta esta revisão da literatura, o valor mais elevado obtido através de um modelo baseado em léxico é 74% de medida de f1, no entanto este léxico é de sentimentos e remete para uma tarefa de detecção de polaridades. No que diz respeito aos diversos algoritmos utilizados, os que apresentam melhores resultados segundo a medida f1 são: o modelo de Redes Neurais com base na biblioteca *LIBLINEAR* (72,1%), *Back Propagation Neural* (67% em média), Regressão Logística (66,5% em média), e *Sequential Machine Optimization* (71,3%).

Capítulo 3

Dados

Este Capítulo descreve os dados utilizados nesta tese: o modelo de emoções e uma breve definição das mesmas, a descrição dos dados estudados, as dificuldades sentidas nas anotações manuais, e as decisões tomadas no pré-processamento do *dataset* analisado.

3.1 Modelo e Descrição de Emoções

Plutchik [30] propôs a roda de emoções observada na Figura 3.1 [21], onde representa as oito emoções básicas que defende como pares adjacentes (alegria-tristeza, raiva-medo, confiança-nojo, antecipação-surpresa) e que quanto mais próximo do centro da roda mais intensas são as emoções. Afirma ainda existirem emoções, encontradas nos espaços em branco, que são a combinação de ambas as emoções básicas ao seu redor.

O modelo de emoções adotado é o de Plutchik, visto ser dos modelos mais utilizados e, também, por ser aquele em que se baseia o léxico que vamos explorar. Consideram-se as oito emoções básicas e dois sentimentos (positivo e negativo). Na Tabela 3.1 observa-se a tradução destas emoções, bem como uma breve descrição do considerado na anotação de *tweets* de forma a facilitar o trabalho dos anotadores e para que considerem o mesmo significado das emoções.

3.2 Descrição dos Dados

Os dados foram retirados diretamente do *Twitter*, correspondendo a cerca de 16 milhões de *tweets*, escritos em português e geolocalizados, tendo origem em cerca de 500 países, incluindo Portugal e Brasil.

Este trabalho centra-se na comunidade portuguesa, sendo direcionado para os *tweets* escritos em português e de origem na Comunidade dos Países de Língua Portuguesa (CPLP), que inclui Portugal, Brasil, Timor Leste, e os Países Africanos de Língua Oficial

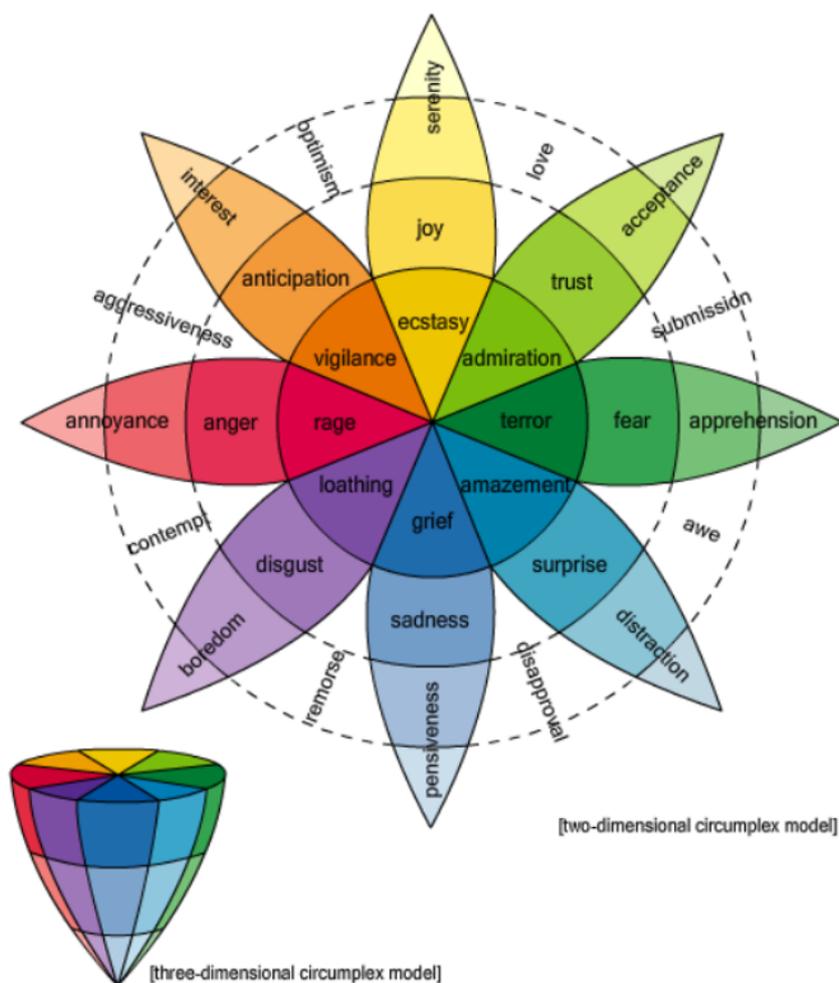


Figura 3.1: Roda de Emoções de Plutchik.

Emoção	Tradução	Breve descrição
Anger	Raiva	Fúria ou irritação, perante algo ou alguém.
Anticipation	Antecipação	Previsão de algo sem conhecimento.
Disgust	Nojo	Repugnância, aversão ou desaprovação perante um indivíduo ou acontecimento.
Fear	Medo	Receio de algo ou alguém.
Joy	Alegria	Felicidade, alegria momentânea ou divertimento perante alguém ou algum acontecimento. Não abrange o gozo.
Sadness	Tristeza	Infelicidade, desânimo, mágoa ou desgosto perante algo ou algum acontecimento.
Surprise	Surpresa	Acontecimento não esperado.
Trust	Confiança	Confiança ou fé, em algo ou alguém que não o próprio.

Tabela 3.1: Descrição das emoções utilizadas.

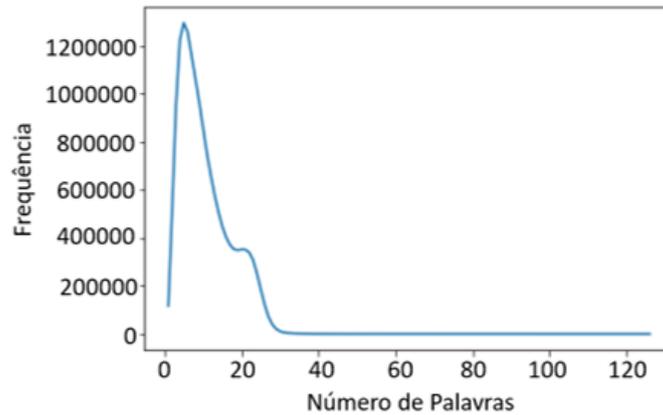


Figura 3.2: Distribuição de *tweets* por número de palavras.

Estava tudo tranquilo aqui no reino do leste até descobrir q há uma igreja evangélica aqui perto. Imagem a zuada!
 eU sOu mACHo eU DirlJO CaRROs
 odeio ficar bld cr, saio descontando em quem não tem nada a ver, uoor isso. 😏 😏
 eu: a talita: muito piranha minha amiga
 Susseguei So Vou Sair Na Micareta Agora Guarda Muita Energiaaaaa
 sp, ba, rj, pe, es, mg, e varios outros 😊
 Por qué me da risa esto en LDJDLKDLKSK

Figura 3.3: Exemplos de *tweets* anotados.

Portuguesa (PALOP), compreendendo Angola, Cabo Verde, Guiné-Bissau, Guiné Equatorial, Moçambique, e São Tomé e Príncipe. O Brasil tem a maioria dos *tweets*, com 14995078 *tweets*, seguido por Portugal, com 383874 *tweets*, Timor Leste e países PALOP com 15369 *tweets*, 1632 *tweets* não têm nenhum país associado e os restantes países têm um total de 549874 *tweets*.

O comprimento médio de cada *tweet* é de aproximadamente 10 palavras e varia de 1 a 125. A Figura 3.2 apresenta a distribuição de palavras por *tweet*, revelando que a maioria contém menos de 25 palavras. Os dados consistem em texto não estruturado contendo: abreviaturas, *hashtags*, *emojis*, calão e *emoticons*. Apresentam, assim, um conjunto de fenómenos que tornam a sua análise mais difícil. Por exemplo, alguns *tweets* são incorretamente identificados como escritos em português e podem conter palavras escritas noutra língua; erros ortográficos, de sintaxe e semânticos; abreviaturas desconhecidas; gírias; uso de *emojis* sem ter em conta o seu suposto significado; entre outras questões.

3.3 Anotação Manual de *Tweets*

A identificação de emoções é uma tarefa por si só difícil: pela complexidade das emoções; por ser necessário retirar significado das frases, subentendendo o que os autores pretendem transmitir; e pela subjetividade da tarefa, ou seja, na mesma frase diferentes indivíduos podem perceber emoções distintas. O que leva a uma escassez de *datasets* ano-

	Raiva	Antec.	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Percentagem média de acordo entre pares	90,2%	94,3%	82,9%	98,9%	78,4%	82,5%	89%	88,3%
Kappa de Cohen	0,431	0,191	0,273	0,347	0,478	0,507	0,271	0,270
Alpha de Krippendorff	0,430	0,190	0,261	0,348	0,478	0,505	0,182	0,182

Tabela 3.2: Coeficientes de concordância entre as anotações manuais.

	Raiva	Antec.	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
AM-1	109	45	89	8	275	201	71	75
AM-2	81	28	178	9	309	258	93	79

Tabela 3.3: Número de *tweets* classificados, por emoção, pelos anotadores.

tados, principalmente para a língua portuguesa. Assim sendo e de forma a conseguirmos validar os nossos resultados, foram anotados manualmente, por duas pessoas, 1000 *tweets* da amostra inicial, de 0 a 8 emoções, podendo estes conter mais do que uma emoção. Foi perceptível um acréscimo na dificuldade desta tarefa resultante do tipo de serviço que os microblogs oferecem: mensagens curtas, de escrita informal, em tempo real para qualquer pessoa e parte do mundo, como é o caso do *Twitter*. Observou-se o uso recorrente à gíria e ao calão; erros ortográficos e de semântica; abreviaturas; expressões próprias de outras línguas; escrita confusa; mau uso de pontuação; e alguns *tweets* escritos em espanhol e inglês. Na Figura 3.3 observa-se exemplos destes *tweets*. A ironia e o gozo também são um ponto a ter em conta, uma vez que estão bastante presentes nos *tweets* e, por vezes, não são completamente claros, levantando a dúvida de se o sujeito estará ou não a ser irónico ou a expressar gozo, e se não, quais as verdadeiras emoções que está a encobrir. Por este motivo, a maioria dos *tweets* com estes indícios não foram classificados.

A Tabela 3.2 mostra o acordo, por emoção, entre as pessoas que classificaram manualmente os *tweets*. A baixa concordância observada, na maioria das emoções, deve-se à falta de exemplos dessa emoção na amostra, como é possível observar na Tabela 3.3, e à subjetividade das emoções associadas principalmente à complexidade e dificuldade da tarefa. Tal como referido, a Tabela 3.1 pretende clarificar e remover alguma subjetividade das emoções. Contudo, a compreensão dos *tweets*, assumindo que são decifráveis, varia de pessoa para pessoa e é influenciada por numerosos fatores simples como a idade e o sexo, e outros mais imprevisíveis como a experiência pessoal, onde contextos iguais adquirem experiências e verdades diferentes e nos levam a experimentar emoções diferentes e intensidades e a enfrentá-las e considerá-las de forma diferente.

<i>Token 1</i>	[O meu] vestido preferido é azul.
<i>Token 2</i>	O [meu vestido] preferido é azul.
<i>Token 3</i>	O meu [vestido preferido] é azul.
<i>Token 4</i>	O meu vestido [preferido é] azul.
<i>Token 5</i>	O meu vestido preferido [é azul].

Tabela 3.4: Exemplos de bigramas.

3.4 Pré-processamento

O pré-processamento de dados prepara os mesmos para a análise desejada e pode incluir quatro etapas: unitização e tokenização; padronização e limpeza de dados de texto; remoção de *stop words*; e *stemming* ou lematização, que removem informação desnecessária do texto e diminuem a complexidade das escolhas necessárias para a tarefa a realizar, contribuindo assim para o seu sucesso.

O passo "unitização e tokenização" consiste na escolha da unidade de texto a analisar, isto é, o *token*, e na separação do texto de acordo com o mesmo. O *token* deve ser a unidade mais proveitosa do ponto de visto do projeto, podendo ser uma única palavra, grupos de palavras ou frases. Em alternativa a esta forma mais simples, existem os n-gramas que guardam informação sobre a coocorrência de palavras pois o *token* representa palavras sequenciais de comprimento "n", como mostra a Tabela 3.4 com "n" igual a dois. Na segunda etapa a limpeza de dados passa pela eliminação de caracteres que não acrescentam valor, por exemplo: pontuações, espaços extra, *emojis* e números. A padronização serve para uniformizar representações. Por exemplo, para que "Cão" e "cão" sejam consideradas a mesma palavra é necessário passar todos os caracteres para letras minúsculas.

A remoção de *stop words* tem como propósito excluir palavras como "e", "que" e "de" que são utilizadas para fins gramaticais, ou seja, como ligação entre as palavras e que, portanto, adicionam pouco conteúdo.

Antes de realizar a última etapa, é necessário compreender-se que a sintaxe diz respeito à estrutura da frase e a semântica ao significado desta. Importa referir que POS são categorias gramaticais ou classes de palavras, como substantivo, verbo e adjetivo, que se bem identificados nos *tokens* facilitam e conduzem ao seu significado.

O tratamento dos dados facilita a extração das características necessárias para o sistema identificar as emoções presentes nos *tweets*. Este processo, tendo em consideração a abordagem baseada em léxico, engloba tokenização, limpeza, e *stemming* ou *lematização*, todas realizadas de forma automatizada. A tokenização garante a decomposição dos *tweets* em palavras, pontuações, *emoticons* e *emojis*, permitindo uma análise palavra a palavra por *tweet*. A limpeza pretende diminuir o ruído nas palavras para facilitar a comparação das mesmas, sendo substituídas as letras maiúsculas por minúsculas e todas as letras iguais

<i>Tweet</i>	@mais Um emprego novo, amém.	@rayssaster n reparou que eu te mandei um beijo??
<i>Tokens</i>	'@mais', 'Um', 'emprego', 'novo', ',', 'amém', ','	'@rayssaster', 'n', 'reparou', 'que', 'eu', 'te', 'mandei', 'um', 'beijo', '?', '?'
SnowballStemmer	@mais um empreg nov , amém .	@rayssast n repar que eu te mand um beij ? ?
RSLPStemmer	@mais um empreg nov , amém .	@rayssast n repar que eu te mand um beij ? ?
Lematização Spacy	@mais Um empregar novo , amém .	@rayssaster n reparar que eu te mandar um beijar ? ?
Lematização Hunspell	@mais um empregar novo , amém .	@rayssaster n reparar que eu te mandar um beijo/beijar ? ?

Tabela 3.5: Exemplos de *stemming* e lematização, para a língua portuguesa.

<i>Tweet</i> inglês	@mais A new job, amen.	@rayssaster n noticed I sent you a kiss ??
<i>Tokens</i>	'@mais', 'a', 'new', 'job', ',', 'amen', ','	'@rayssaster', 'n', 'noticed', 'i', 'sent', 'you', 'a', 'kiss', '?', '?'
PorterStemmer	@mais a new job , amen .	@rayssast n notic i sent you a kiss ? ?
Stemming Lancaster	@mais a new job , am .	@rayssaster n not i sent you a kiss ? ?
Lematização Spacy	@mais a new job , amen .	@rayssaster n notice i send you a kiss ? ?

Tabela 3.6: Exemplos de *stemming* e lematização, para a língua inglesa. O *tweet* em inglês foi traduzido através do Google Tradutor.

consecutivas, três ou mais vezes, por apenas uma das letras. Por último, testámos *stemming* e lematização para ambas as línguas. Para a língua portuguesa, fazendo recurso da biblioteca NLTK, utilizámos o *stemmer* SnowballStemmer, que suporta mais de 10 línguas, e o RSLPStemmer, exclusivo para o português. No entanto, por esta ser uma língua com uma flexão rica, ambos reduzem demasiado as palavras retornando raízes que, na maioria das vezes agrupam demasiados conceitos, por exemplo os verbos poder e podar seriam considerados o mesmo. A lematização é, portanto, a melhor opção nesta língua. Implementámos e testámos dois lematizadores: o SpaCy e o Hunspell. No caso do Hunspell, contrariamente ao SpaCy, é necessário proceder previamente à tokenização das palavras e fornecer-lhe o POS das mesmas, para que haja contexto e nas palavras que possuem diversos lemas, como acontece com “beijo” no segundo *tweet* da Tabela 3.5, seja possível escolher o mais correto. Mediante as experiências realizadas, os resultados descritos por Lars Nieradzik no seu *blog*¹, onde compara os mesmos lematizadores, e tendo em conta que o Hunspell necessitaria de um bom dicionário e contexto prévio, escolhemos o SpaCy. Para a língua inglesa usámos o PorterStemmer e o LancasterStemmer. As experiências desenvolvidas demonstram que *stemming* não só é bastante mais rápido do que a lematização como apesar de, por vezes, reduzir bastante as palavras, funciona bem para o inglês, principalmente utilizando ambos os *stemmers* em conjunto. Porém, é essencial para a tarefa de classificação que o lema de cada palavra seja encontrado. A lematização é o processo mais adequado, uma vez que o *stemmig* não utiliza informação POS e que quando se aplica um lematizador palavra a palavra, embora se obtenha ambiguidade, a velocidade é semelhante. O SpaCy foi, mais uma vez, o eleito, uma vez que para além das funcionalidades internas que incorpora (tokenização das palavras e fornecimento do POS das mesmas), apresenta bons resultados em textos estruturados. A Tabela 3.6 mostra resultados das experiências para a língua inglesa.

¹<https://lars76.github.io/nlp/lemmatize-portuguese/>

Capítulo 4

Abordagem Baseada em Léxicos

Este Capítulo detalha os processos, experiências e decisões realizadas seguindo a abordagem baseada em léxico para a classificação de emoções, desde a construção do classificador e diversos léxicos utilizados, ao percurso dos *tweets* até serem classificados.

4.1 Arquitetura do Classificador

A Figura 4.1 reflete um esquema geral do classificador desenvolvido. O módulo concebido recebe um *tweet*, executa o pré-processamento para facilitar a sua classificação e adiciona o conjunto de emoções de todas as palavras do *tweet* e presentes no léxico, retornando as várias emoções presentes no *tweet* analisado. Esta contagem contabiliza as emoções existentes no *tweet*, descartando a frequência das mesmas.

Neste trabalho foram testados alguns léxicos, porém a totalidade dos dados foi classificada utilizando o léxico NCR original versão portuguesa e tanto os *tweets* como as suas emoções associadas guardados numa base de dados.

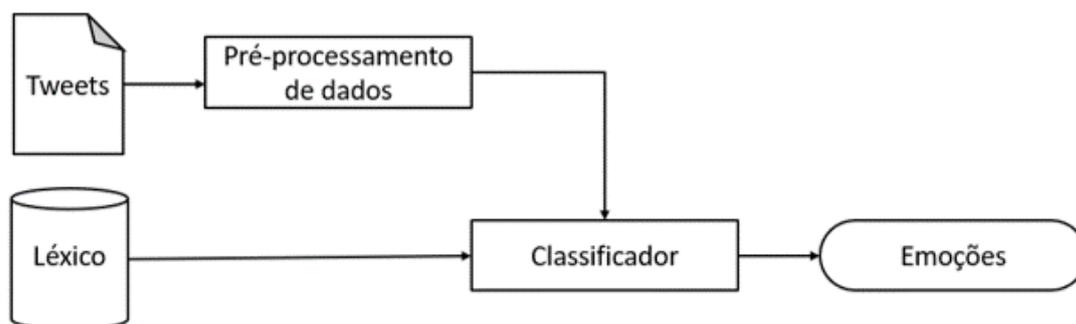


Figura 4.1: Esquema geral do Classificador.

Inglês	Português	Pos.	Neg.	Raiva	Antec.	Nojo	Medo	Alegria	Tristeza	Surp.	Conf.
aback	Surpresa	0	0	0	0	0	0	0	0	0	0
abacus	ábaco	0	0	0	0	0	0	0	0	0	1
abandon	abandono	0	1	0	0	0	1	0	1	0	0
abandoned	abandonado	0	1	1	0	0	1	0	1	0	0
abandonment	abandono	0	1	1	0	0	1	0	1	1	0
abate	diminuir	0	0	0	0	0	0	0	0	0	0
abatement	redução	0	0	0	0	0	0	0	0	0	0
abba	abba	1	0	0	0	0	0	0	0	0	0

Tabela 4.1: Exemplo do léxico NRC Word-Emotion com tradução em português.

4.2 Léxicos

O léxico escolhido foi o NRC Word-Emotion Association, estudado de forma a perceber os seus pontos fortes e fracos. A qualidade do léxico determina o potencial do classificador, pelo que foram desenvolvidos dois novos léxicos numa tentativa de obter melhores resultados.

Esta Secção descreve e informa sobre os diferentes léxicos utilizados neste trabalho.

4.2.1 Léxico NRC Original

O NRC Word-Emotion Association Lexicon, também conhecido como Emotion lexicon ou EmoLex, foi desenvolvido para a língua inglesa por Mohammad e Turney [22] através de anotação manual utilizando *crowdsourcing*.

O EmoLex considera dois sentimentos (positivo e negativo) e oito emoções básicas (raiva, nojo, medo, alegria, tristeza, surpresa, confiança e antecipação), atribuindo a cada palavra as emoções e sentimentos adequados, como ilustrado na Tabela 4.1. É constituído por 14182 palavras, das quais 2312 têm um sentimento positivo e 3324 têm um sentimento negativo. No que diz respeito às emoções, 1247 têm raiva associada, 839 antecipação, 1058 nojo, 1476 medo, 689 alegria, 1191 tristeza, 534 surpresa e 1231 confiança. Uma palavra pode ter 0 ou até 8 emoções atribuídas, sendo que 9719 não têm nenhuma e 2 têm todas as emoções.

Este léxico foi traduzido em Novembro de 2017, pelos seus autores, para mais de 40 idiomas utilizando o Google Tradutor, todas as versões encontram-se disponíveis *online*¹. Existe um grande número de palavras incorretamente traduzidas para o português, e mais de 250 palavras sem tradução [21]. A versão inglesa contém palavras como: *amour*, *alimantation* e *aloha*.

¹<https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

Inglês	Português	Pos.	Neg.	Raiva	Antec.	Nojo	Medo	Alegria	Tristeza	Surp.	Conf.
abandon	abandonar	0	1	0	0	0	1	0	1	0	0
abandon	abandono	0	1	0	0	0	1	0	1	0	0
abhorrent	abominável	0	1	1	0	1	1	0	0	0	0
abhorrent	repugnante	0	1	1	0	1	1	0	0	0	0
hurtful	doloroso	0	1	2	0	2	2	0	2	0	0
painful	doloroso	0	1	2	0	2	2	0	2	0	0
parasite	parasita	0	1	0	0	2	2	0	0	0	0

Tabela 4.2: Excerto do léxico melhorado, com os respectivos pesos.

4.2.2 LEX-I — Google Tradutor e DeepL

Os sistemas de tradução automática melhoraram significativamente desde 2017.

Com o objetivo de melhorar a tradução portuguesa do EmoLex e aumentar a fiabilidade da mesma, a versão fornecida em inglês foi traduzida utilizando o Google Tradutor e o DeepL, uma vez que são as plataformas de tradução mais utilizadas e reconhecidas.

As traduções que obtivemos de ambas as plataformas deram origem a um novo léxico (Lex-I) com um total de 18920 palavras. Apenas 1429 palavras das 14182 do EmoLex português não estão incluídas no Lex-I, porque segundo estes tradutores não correspondem a nenhuma tradução das palavras originais em inglês.

De forma a obtermos um valor de confiança associado, atribuímos às emoções presentes nas 9424 palavras com traduções iguais peso 2 e às emoções presentes nas 4748 palavras com traduções divergentes peso 1. A Tabela 4.2 apresenta um excerto do léxico melhorado, onde são visíveis palavras que originaram a mesma tradução e outras palavras que originaram traduções diferentes. Algumas das palavras com traduções diferentes divergem apenas no género (exemplo: agravada e agravado), no acordo ortográfico (abjecto e abjeto), no número (conselho e conselhos), na raiz da palavra (afixar e afixo) e na acentuação (acadêmico e académico).

As mais de 250 palavras que originalmente não haviam sido traduzidas, foram substituídas pela própria palavra ou por uma possível tradução fornecida por estes tradutores. É de notar que cerca de 500 palavras permaneceram iguais ao léxico NRC original, pois o mesmo contém palavras escritas que não estão em inglês (*amour* e *aloha*); existem palavras que se escrevem da mesma forma em inglês e português (total, animal e zebra); nomes próprios (Toby e Billy) ou são estrangeirismos (*byte*, *viking* e *versus*).

4.2.3 LEX-II — MUSE

A rede social Facebook encontra-se disponível em mais de 100 línguas e, no contexto da Inteligência Artificial, criou a biblioteca MUSE (*Multilingual Unsupervised and Supervised*

Inglês	Português	Pos.	Neg.	Raiva	Antec.	Nojo	Medo	Alegria	Tristeza	Surp.	Conf.
abandon	abandone	0	1	0	0	0	1	0	1	0	0
abandon	abandonem	0	1	0	0	0	1	0	1	0	0
abandon	abandonar	0	1	0	0	0	1	0	1	0	0
abandon	abandona	0	1	0	0	0	1	0	1	0	0
abandon	abandono	0	1	1	0	0	1	0	1	1	0
abandonment	abandono	0	1	1	0	0	1	0	1	1	0
painful	dolorosa	0	1	1	0	1	1	0	1	0	0
painful	dolorosos	0	1	1	0	1	1	0	1	0	0
painful	doloroso	0	1	1	0	1	1	0	1	0	0
parasite	parasita	0	1	0	0	1	1	0	0	0	0
parasite	parasitas	0	1	0	0	1	1	0	0	0	0

Tabela 4.3: Excerto do LEX-II, utilizando o MUSE como recurso de tradução.

Embeddings) [11], baseada em *fastText*, para compreender, representar e classificar os milhares de dados inseridos na plataforma, em diferentes línguas. Tal é aplicado para fins como: anúncios publicitários que vão de encontro as preferenciais necessidades do utilizador; quais temas/assuntos estão mais relacionados e qual o seu público alvo; número de caracteres ideal de uma publicação.

A MUSE suporta dois métodos de geração dos alinhamentos: aprendizagem supervisionada, utilizando os dicionários ou identificando caracteres da palavra idênticos; e não supervisionada, não usando dados paralelos. Oferece 110 dicionários bilingue de verdade fundamental bastante amplos, incluindo de inglês para 44 línguas diferentes e vice versa, e *word embeddings* para 30 línguas divergentes, representando cada palavra num único espaço vetorial.

Neste trabalho, é utilizado o dicionário inglês-português, composto por 92504 palavras em inglês e por 108696 em português, para traduzir o léxico NRC. Estes têm 11147 palavras inglesas em comum, as 3036 restantes do NRC não foram incluídas, por isso, neste dicionário (por exemplo: a palavra "*hurtful*" está incluída no NRC mas não na MUSE, apesar da sua tradução estar incluída nesta última, porque também é a tradução "*painful*" considerada pelo MUSE).

A MUSE apresenta uma tradução média, para as palavras comuns, de 1,76 e no máximo 6 palavras, que deram origem a um novo léxico em português (Lex-II). Observou-se que diferentes palavras inglesas tinham a mesma tradução, nestes casos o valor contabilizado das emoções foi a média, sendo posteriormente arredondado para 0 ou 1. O Lex-II é constituído por 16713 palavras distintas. Na Tabela 4.3 podemos observar algumas palavras com as respetivas emoções consideradas por este léxico e a ilustração dos dois casos mencionados anteriormente.

4.3 Resultados

O EmoLex está escrito em inglês e o receio de perda de informação incentivaram seis experiências diferentes, baseadas em léxicos, para a detecção de emoções em *tweets* portugueses utilizando:

1. A tradução original portuguesa do EmoLex;
2. O LEX-I;
3. O LEX-II;
4. A versão original inglesa do EmoLex e o Google Tradutor para traduzir os *tweets* para inglês;
5. A versão original inglesa do EmoLex e o DeepL para traduzir os *tweets* para inglês;
6. E uma combinação de todos os classificadores anteriores, sendo a classificação atribuída pela maioria.

Nas tabelas 4.4 e 4.5 cada classificador é comparado usando as anotações manuais descritas na Secção 3.3.

De forma a simplificar a observação do desempenho dos diferentes léxicos e a compreender qual produz melhores resultados decidiu-se combinar as anotações numa única referência. No entanto, como existe um baixo nível de acordo entre os anotadores, procurou-se explorar ao máximo a combinação das mesmas e, em vez de uma, criaram-se três referências com perspetivas diferentes. Para uma palavra ser considerada com determinada emoção, na primeira referência basta que uma das pessoas a assinale, na segunda ambas as pessoas têm de concordar, e na última consideram-se os valores do classificador corretos se, pelo menos, uma das pessoas concordar.

A taxa de acerto é geralmente elevada devido à natureza desequilibrada dos dados, mas a precisão, a cobertura, e a medida f1 fornecem informações mais úteis sobre a complexidade da tarefa.

Nas Tabelas 4.6, 4.7 e 4.8 observa-se a avaliação das experiências tendo em conta as três referências.

O LEX-I mostra uma melhoria relativamente ao EmoLex, devido à melhoria das ferramentas utilizadas.

Ao contrário do esperado o LEX-II não apresenta melhores resultados do que o LEX-I, muito provavelmente devido à restrição feita ao dicionário original MUSE e à não inclusão de uma quantidade significativa de palavras do léxico do EmoLex. Todavia, mostra melhor cobertura do que a experiência um, apesar de por vezes uma precisão inferior que leva a uma medida f1 semelhante.

Considerando a anotação manual 1 (AM-1) como referência

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Experiência 1: NRC original PT								
Taxa de acerto	0,836	0,763	0,838	0,809	0,672	0,725	0,813	0,733
Precisão	0,296	0,071	0,213	0,011	0,373	0,330	0,128	0,107
Cobertura	0,367	0,356	0,303	0,250	0,284	0,358	0,282	0,342
Medida f1	0,328	0,119	0,250	0,021	0,322	0,344	0,176	0,163
Experiência 2: Lex-I								
Taxa de acerto	0,822	0,694	0,819	0,776	0,668	0,712	0,802	0,621
Precisão	0,283	0,066	0,185	0,018	0,389	0,324	0,133	0,111
Cobertura	0,413	0,444	0,303	0,5	0,364	0,398	0,324	0,566
Medida f1	0,336	0,116	0,23	0,034	0,376	0,357	0,189	0,185
Experiência 3: Lex-II								
Taxa de acerto	0,787	0,711	0,824	0,75	0,664	0,709	0,809	0,682
Precisão	0,223	0,073	0,208	0,016	0,396	0,339	0,105	0,105
Cobertura	0,385	0,467	0,348	0,500	0,422	0,473	0,225	0,421
Medida f1	0,283	0,127	0,261	0,031	0,408	0,395	0,143	0,168
Experiência 4: NRC original EN - GT								
Taxa de acerto	0,817	0,718	0,845	0,806	0,679	0,706	0,813	0,703
Precisão	0,282	0,066	0,246	0,016	0,419	0,293	0,113	0,123
Cobertura	0,440	0,400	0,360	0,375	0,433	0,328	0,239	0,474
Medida f1	0,344	0,113	0,292	0,030	0,426	0,310	0,154	0,195
Experiência 5: NRC original EN - DPL								
Taxa de acerto	0,827	0,726	0,850	0,819	0,679	0,715	0,825	0,694
Precisão	0,281	0,065	0,239	0,022	0,414	0,296	0,139	0,111
Cobertura	0,376	0,378	0,315	0,500	0,400	0,303	0,282	0,434
Medida f1	0,322	0,110	0,272	0,042	0,407	0,300	0,186	0,177
Experiência 6: Fusão das cinco experiências								
Taxa de acerto	0,829	0,736	0,847	0,804	0,687	0,722	0,838	0,704
Precisão	0,282	0,067	0,233	0,015	0,426	0,329	0,153	0,113
Cobertura	0,367	0,378	0,315	0,375	0,396	0,368	0,282	0,421
Medida f1	0,319	0,114	0,268	0,030	0,411	0,347	0,198	0,178

Tabela 4.4: Métricas de avaliação das experiências, considerando a anotação manual 1 (AM-1) como referência.

Considerando a anotação manual 2 (AM-2) como referência

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Experiência 1: NRC original PT								
Taxa de acerto	0,840	0,768	0,779	0,812	0,656	0,710	0,785	0,754
Precisão	0,207	0,045	0,331	0,021	0,416	0,427	0,109	0,156
Cobertura	0,346	0,357	0,236	0,444	0,282	0,360	0,183	0,481
Medida f1	0,259	0,079	0,275	0,041	0,336	0,391	0,137	0,236
Experiência 2: Lex-I								
Taxa de acerto	0,824	0,695	0,766	0,777	0,648	0,701	0,776	0,636
Precisão	0,201	0,04	0,308	0,022	0,416	0,417	0,121	0,134
Cobertura	0,395	0,429	0,253	0,556	0,346	0,399	0,226	0,658
Medida f1	0,267	0,073	0,278	0,043	0,378	0,408	0,158	0,222
Experiência 3: Lex-II								
Taxa de acerto	0,791	0,704	0,771	0,747	0,628	0,69	0,785	0,701
Precisão	0,16	0,031	0,329	0,012	0,392	0,407	0,099	0,141
Cobertura	0,37	0,321	0,275	0,333	0,372	0,442	0,161	0,544
Medida f1	0,223	0,057	0,300	0,023	0,382	0,424	0,122	0,223
Experiência 4: NRC original EN - GT								
Taxa de acerto	0,823	0,721	0,788	0,805	0,659	0,681	0,787	0,706
Precisão	0,218	0,040	0,369	0,016	0,444	0,364	0,100	0,133
Cobertura	0,457	0,393	0,270	0,333	0,408	0,318	0,161	0,494
Medida f1	0,295	0,073	0,312	0,030	0,425	0,340	0,123	0,210
Experiência 5: NRC original EN - DPL								
Taxa de acerto	0,831	0,731	0,797	0,816	0,659	0,686	0,795	0,703
Precisão	0,199	0,042	0,393	0,017	0,440	0,364	0,111	0,132
Cobertura	0,358	0,393	0,258	0,333	0,379	0,291	0,172	0,494
Medida f1	0,256	0,076	0,312	0,032	0,407	0,323	0,135	0,208
Experiência 6: Fusão das cinco experiências								
Taxa de acerto	0,833	0,739	0,786	0,803	0,657	0,699	0,806	0,719
Precisão	0,197	0,040	0,350	0,015	0,434	0,404	0,115	0,144
Cobertura	0,346	0,357	0,236	0,333	0,359	0,353	0,161	0,519
Medida f1	0,251	0,071	0,282	0,030	0,393	0,377	0,134	0,226

Tabela 4.5: Métricas de avaliação das experiências, considerando a anotação manual 2 (AM-2) como referência.

Experiência 1: NRC original PT

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,811	0,757	0,754	0,809	0,609	0,679	0,765	0,727
Precisão	0,333	0,103	0,394	0,027	0,522	0,491	0,186	0,218
Cobertura	0,312	0,354	0,228	0,357	0,273	0,338	0,212	0,390
Medida f1	0,323	0,159	0,289	0,050	0,358	0,400	0,198	0,280

Experiência 2: Lex-I

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,801	0,690	0,745	0,776	0,611	0,672	0,762	0,635
Precisão	0,327	0,093	0,377	0,031	0,521	0,478	0,208	0,206
Cobertura	0,361	0,431	0,251	0,500	0,335	0,372	0,263	0,588
Medida f1	0,343	0,153	0,301	0,059	0,408	0,418	0,232	0,305

Experiência 3: Lex-II

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,770	0,703	0,748	0,748	0,605	0,677	0,761	0,682
Precisão	0,271	0,094	0,389	0,024	0,509	0,489	0,164	0,203
Cobertura	0,354	0,415	0,265	0,429	0,372	0,432	0,182	0,456
Medida f1	0,307	0,154	0,315	0,045	0,430	0,459	0,173	0,281

Experiência 4: NRC original EN - GT

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,808	0,712	0,767	0,804	0,626	0,654	0,763	0,695
Precisão	0,359	0,092	0,446	0,026	0,546	0,436	0,167	0,212
Cobertura	0,424	0,385	0,265	0,357	0,388	0,309	0,182	0,456
Medida f1	0,389	0,148	0,332	0,049	0,453	0,362	0,174	0,289

Experiência 5: NRC original EN - DPL

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,804	0,720	0,768	0,817	0,622	0,657	0,775	0,686
Precisão	0,322	0,091	0,444	0,033	0,541	0,437	0,194	0,199
Cobertura	0,326	0,369	0,237	0,429	0,360	0,284	0,204	0,434
Medida f1	0,324	0,146	0,310	0,062	0,432	0,344	0,199	0,273

Experiência 6: Fusão das cinco experiências

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,808	0,73	0,761	0,802	0,622	0,674	0,788	0,698
Precision	0,331	0,095	0,417	0,026	0,543	0,48	0,214	0,208
Cobertura	0,326	0,369	0,228	0,357	0,347	0,341	0,204	0,434
Medida f1	0,329	0,151	0,295	0,048	0,424	0,399	0,209	0,281

Tabela 4.6: Métricas de avaliação das experiências, utilizando a primeira referência.

Experiência 1: NRC original PT

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,865	0,774	0,863	0,812	0,719	0,756	0,833	0,760
Precisão	0,170	0,013	0,150	0,005	0,268	0,266	0,051	0,045
Cobertura	0,500	0,375	0,396	0,333	0,304	0,408	0,296	0,579
Medida f1	0,254	0,026	0,217	0,011	0,285	0,322	0,087	0,084

Experiência 2: Lex-I

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,845	0,699	0,840	0,777	0,705	0,741	0,816	0,622
Precisão	0,157	0,013	0,116	0,009	0,284	0,263	0,046	0,039
Cobertura	0,543	0,500	0,354	0,667	0,397	0,458	0,296	0,789
Medida f1	0,244	0,026	0,175	0,018	0,331	0,334	0,080	0,074

Experiência 3: Lex-II

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,808	0,712	0,847	0,749	0,687	0,722	0,833	0,701
Precisão	0,112	0,010	0,148	0,004	0,280	0,257	0,039	0,042
Cobertura	0,457	0,375	0,458	0,333	0,446	0,507	0,222	0,684
Medida f1	0,179	0,020	0,223	0,008	0,344	0,341	0,067	0,080

Experiência 4: NRC original EN - GT

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,832	0,727	0,866	0,807	0,712	0,733	0,837	0,714
Precisão	0,141	0,015	0,169	0,005	0,317	0,222	0,047	0,044
Cobertura	0,522	0,500	0,458	0,333	0,489	0,352	0,259	0,684
Medida f1	0,222	0,028	0,247	0,010	0,385	0,272	0,079	0,083

Experiência 5: NRC original EN - DPL

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,854	0,737	0,879	0,818	0,716	0,744	0,845	0,711
Precisão	0,158	0,015	0,188	0,006	0,312	0,223	0,056	0,044
Cobertura	0,500	0,500	0,458	0,333	0,451	0,324	0,296	0,684
Medida f1	0,240	0,030	0,267	0,011	0,369	0,264	0,094	0,083

Experiência 6: Fusão das cinco experiências

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,854	0,745	0,872	0,805	0,722	0,747	0,856	0,725
Precisão	0,148	0,012	0,167	0,005	0,316	0,253	0,053	0,049
Cobertura	0,457	0,375	0,417	0,333	0,44	0,401	0,259	0,737
Medida f1	0,223	0,023	0,238	0,01	0,368	0,311	0,089	0,092

Tabela 4.7: Métricas de avaliação das experiências, utilizando a segunda referência.

Experiência 1: NRC original PT

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,887	0,794	0,894	0,816	0,772	0,805	0,854	0,802
Precisão	0,333	0,103	0,394	0,027	0,522	0,491	0,186	0,218
Cobertura	0,662	0,821	0,633	0,714	0,460	0,560	0,604	0,869
Medida f1	0,443	0,183	0,485	0,052	0,489	0,523	0,284	0,349

Experiência 2: Lex-I

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,872	0,723	0,878	0,782	0,766	0,794	0,844	0,687
Precisão	0,327	0,093	0,377	0,031	0,521	0,478	0,208	0,206
Cobertura	0,712	0,875	0,640	0,875	0,547	0,605	0,655	0,952
Medida f1	0,448	0,168	0,474	0,060	0,534	0,534	0,316	0,338

Experiência 3: Lex-II

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,838	0,736	0,883	0,754	0,754	0,787	0,852	0,750
Precisão	0,271	0,094	0,389	0,024	0,509	0,489	0,164	0,203
Cobertura	0,671	0,844	0,690	0,750	0,594	0,662	0,543	0,912
Medida f1	0,386	0,170	0,498	0,047	0,548	0,563	0,253	0,332

Experiência 4: NRC original EN - GT

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,869	0,748	0,902	0,811	0,777	0,781	0,855	0,763
Precisão	0,359	0,092	0,446	0,026	0,546	0,436	0,167	0,212
Cobertura	0,735	0,862	0,690	0,714	0,622	0,516	0,556	0,912
Medida f1	0,482	0,166	0,542	0,050	0,582	0,472	0,256	0,343

Experiência 5: NRC original EN - DPL

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,878	0,757	0,909	0,823	0,777	0,788	0,865	0,757
Precisão	0,322	0,091	0,444	0,033	0,541	0,437	0,194	0,199
Cobertura	0,671	0,857	0,667	0,750	0,588	0,484	0,596	0,908
Medida f1	0,435	0,165	0,533	0,063	0,564	0,459	0,293	0,327

Experiência 6: Fusão das cinco experiências

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
Taxa de acerto	0,880	0,766	0,902	0,809	0,780	0,798	0,877	0,770
Precisão	0,331	0,095	0,417	0,026	0,543	0,48	0,214	0,208
Cobertura	0,653	0,828	0,641	0,714	0,574	0,56	0,583	0,922
Medida f1	0,439	0,170	0,505	0,05	0,558	0,517	0,313	0,339

Tabela 4.8: Métricas de avaliação das experiências, utilizando a terceira referência.

	Raiva	Antecip.	Nojo	Medo	Alegria	Tristeza	Surpresa
Percentagem média de acordo entre pares	89,8%	86,4%	91,1%	86,8%	87,3%	85,6%	89,3%
Média de Kappa de Cohen entre pares	0,621	0,655	0,617	0,599	0,671	0,601	0,591
Alpha de Krippendorff	0,620	0,655	0,617	0,598	0,672	0,601	0,591

Tabela 4.9: Coeficientes de fiabilidade para a classificação.

Ambas as experiências quatro e cinco alcançam um bom desempenho. Em comparação com o LEX-II, têm pior precisão, mas melhor cobertura.

A experiência seis foi desenvolvida para melhorar a classificação utilizando todos os classificadores ao mesmo tempo. Embora, por vezes, a precisão seja ligeiramente superior, não é a melhor opção porque o benefício é baixo para compensar o tempo extra que leva.

A referência que tem os melhores valores é a última porque é a mais flexível. Os resultados são menores nas outras referências porque o acordo entre as pessoas que escreveram os dados é baixo. O método NRC EN - GT original tem melhor precisão para a referência um e três e na referência dois é o método NRC EN - DPL original. Todos os valores de referência concordam que o Lex-I fornece melhores valores de cobertura. Uma vez que a medida f1 é uma média ponderada entre a precisão e a cobertura, as experiências com os melhores resultados são as mencionadas, mas depende especificamente de qual a métrica, ora de precisão ora de cobertura, mais elevada.

As experiências mostram que as emoções mais fáceis de prever são a alegria e a tristeza, porque não só são as mais expressas no *Twitter*, mas como as que geram o menor desacordo, tal como demonstrado pelo acordo entre os dois anotadores.

As mais difíceis são o medo e a antecipação, porque são menos comuns nos *tweets*, pelo que a probabilidade de acertar é menor do que nas emoções mais comuns. Além disso, estas emoções podem ser percebidas de forma diferente de pessoa para pessoa. Por vezes usamos erradamente a palavra "medo" para expressar emoções semelhantes de menor intensidade (por exemplo:"insegurança") ou dizemos que antecipamos algo, mas na realidade apenas o deduzimos tendo em conta algo que vimos ou sabemos.

Em geral, os classificadores mostram valores baixos/moderados, contudo, tal era esperado devido à pequena amostra em termos de variedade entre as emoções.

A Tabela 4.9 apresenta a concordância entre os resultados das experiências, exceto a sexta.

A métricas Alpha de Krippendorff e a média de pares Kappa de Cohen e mostram um baixo acordo nas emoções confiança, medo e surpresa, e um acordo moderado em todas as outras.

As diferentes experiências mostram valores bastante semelhantes, como é evidenciado pela concordância entre elas, isto porque ambas se baseiam no mesmo léxico.

Capítulo 5

Abordagem Baseada em Aprendizagem Automática

A Aprendizagem Automática é uma área de investigação, dentro da Inteligência Artificial, onde se extrai conhecimento a partir de um *dataset* [39]. Existem duas categorias:

- Aprendizagem supervisionada - o algoritmo recebe dados etiquetados e aprende com eles;
- Aprendizagem não-supervisionada - o algoritmo recebe apenas dados de entrada e deve estudá-los, perceber a sua estrutura e encontrar padrões.

Neste Capítulo é detalhado o trabalho elaborado com recurso a uma abordagem de aprendizagem automática para a classificação de emoções em *tweets*, incluindo a descrição dos dados utilizados, as experiências e decisões necessárias para a construção do classificador, o algoritmo utilizado e os resultados obtidos.

5.1 Classificação Automática do *Dataset*

O classificador baseado no léxico EmoLex, descrito no Capítulo 4, é utilizado para atribuir automaticamente as emoções estudadas ao *dataset* de 16 milhões de *tweets* descrito na Secção 5.2. Detetou pelo menos uma emoção em cerca de 57% dos *tweets*, o que corresponde a cerca de 9007723, mas um grande número de *tweets* foi classificado com mais do que uma emoção.

A Figura 5.1 revela o número de *tweets* com cada uma das emoções consideradas, onde a emoção predominante é tristeza (3762983 *tweets*), seguida pela antecipação (3309393 *tweets*), e surpresa é a emoção menos predominante (2183907 *tweets*).

A Figura 5.2 apresenta o número de *tweets* pelo número de emoções presentes nos mesmos e por emoção. Em geral, as pessoas não experienciam várias emoções ao mesmo

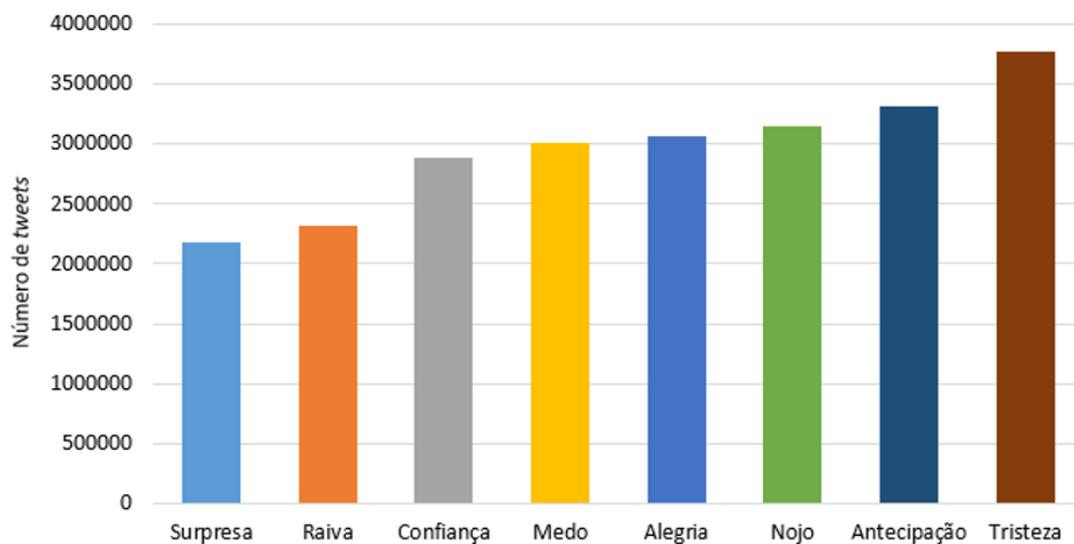


Figura 5.1: Número de *tweets* classificados por emoção.

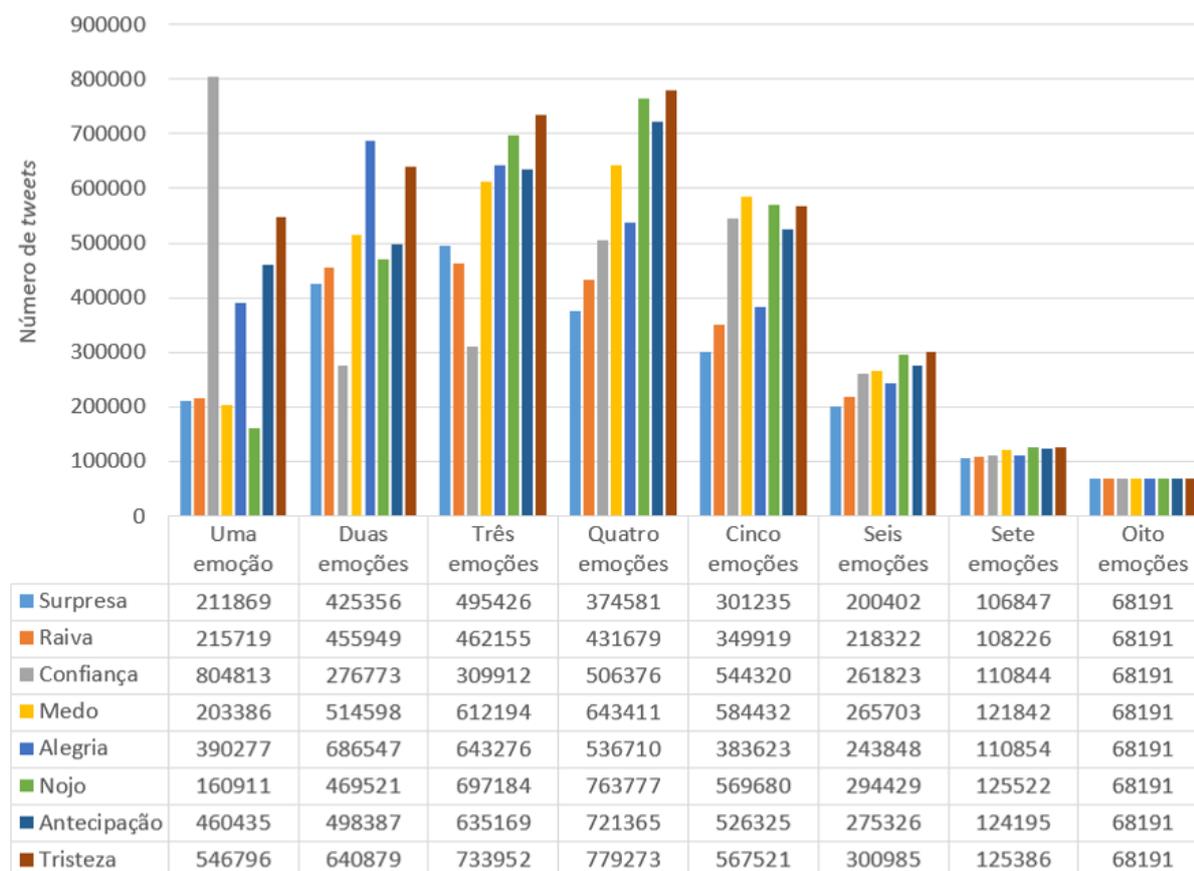


Figura 5.2: Número de *tweets* por emoção e por número de emoções.

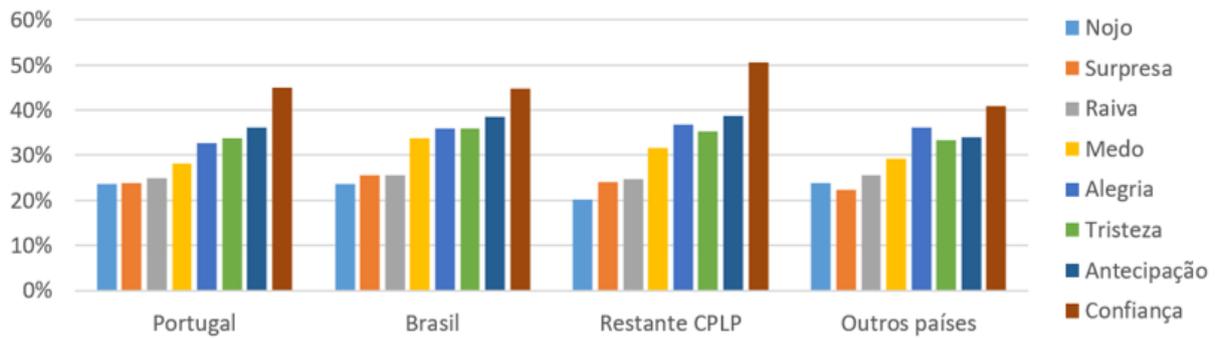


Figura 5.3: Porção de *tweets* classificados automaticamente com, pelo menos, uma emoção, por país.

tempo, pelo que, como era de esperar, há mais *tweets* com menos emoções. E, a probabilidade, nesta rede social em particular, de muitas publicações expressarem o que as pessoas sentem no momento é alta. No entanto, quando os indivíduos expressam as suas opiniões escritas, tendem a pensar mais sobre o assunto e, dependendo da complexidade do mesmo, a expressar mais do que uma emoção. A maioria dos *tweets*, cerca de 30%, têm quatro emoções e cerca de 19% têm uma emoção. Verificou-se que as combinações mais frequentes por número de emoções são: alegria e confiança; antecipação, alegria e confiança; raiva, repugnância, medo e tristeza; antecipação, medo, alegria, surpresa e confiança; antecipação, medo, alegria, tristeza, surpresa e confiança; raiva, antecipação, medo, alegria, alegria, tristeza, surpresa e confiança. O classificador conseguiu atribuir pelo menos uma emoção a 118046 *tweets* do Brasil, 6435 *tweets* de Portugal, 148 *tweets* de países PALOP, e outros 3459 *tweets* de outros países.

A Figura 5.3 mostra a ocorrência de cada emoção por país. Confiança é a emoção mais predominante em todos os países. Portugal e Brasil mostram a mesma tendência para todas as emoções. A raiva e o nojo são as emoções menos predominantes em todos os casos. Tristeza é a emoção mais dominante para Portugal e para o Brasil, mas é ultrapassada pela alegria nos PALOP e pela antecipação noutros países.

5.2 Classificador Baseado em Aprendizagem Supervisionada

O classificador desenvolvido engloba três fases de treino, representadas na Figura 5.4, e dois modelos de Regressão Logística que aprendem de forma supervisionada.

A abordagem foi escolhida com base na revisão de literatura e por ser aplicada e apropriada para problemas de classificação de categorias, entre outros; por possuímos um *dataset* com *tweets* etiquetados — recurso necessário; e pelo algoritmo corresponder às nossas necessidades: tem como propósito o cálculo da probabilidade de uma variável dependente binária (emoção), associada a variáveis independentes (características extraídas do *tweet* em análise), sem natureza específica.

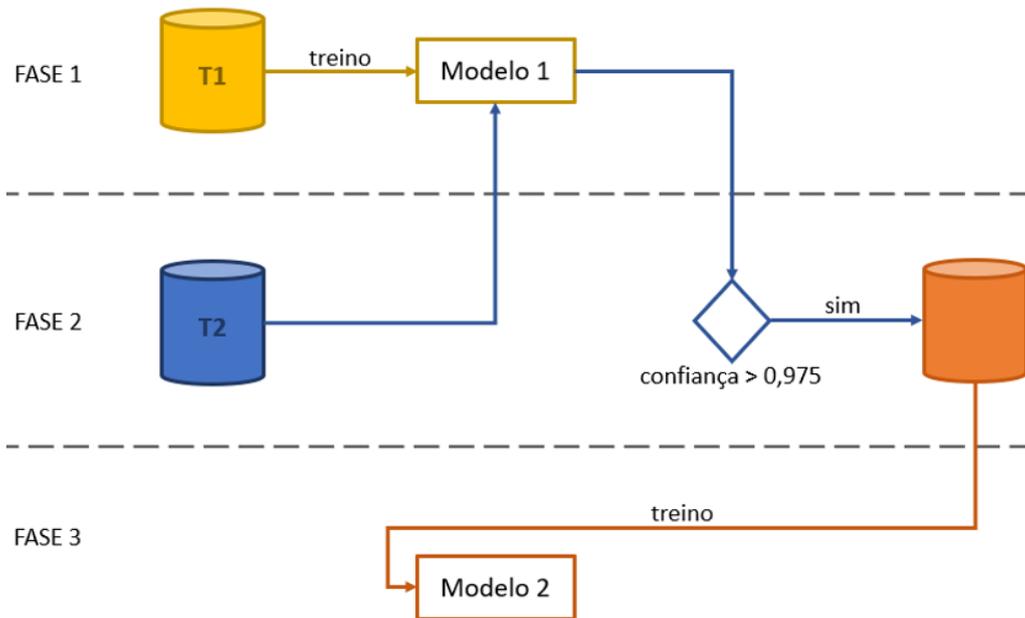


Figura 5.4: Esquema das fases de treino da abordagem baseada em aprendizagem supervisionada.

Como mencionado anteriormente, o *dataset* contém a totalidade dos 16 milhões de dados e foi etiquetado pela abordagem descrita no Capítulo 4. Este foi dividido em quatro conjuntos de dados, para esta abordagem: AM — Anotação Manual, com 1000 *tweets*; T1 com 5 milhões; T2 com, aproximadamente, 9 milhões; e AA — Avaliação Automática, com 2 milhões.

Na primeira fase, foi criado o Modelo 1 de Regressão Logística e dividiu-se o T1 em dados para treino (80%) e para teste (20%), de seguida transformou-se os *tweets* do conjunto de treino em *features* e utilizou-se para treinar o modelo criado. Os 20% restantes dos dados foram utilizados para testar o modelo. Testou-se este modelo com mais três conjuntos de dados: AM, com o propósito de validar os dados do classificador; AA, para testar com dados diferentes e para comparação com o próximo modelo; e T2, para a segunda fase. Todos os testes têm o propósito comum de perceber o sucesso da aprendizagem do classificador. Os *outputs* dos testes realizados passaram, individualmente, pelo processo de avaliação. Este processo recebe as emoções previstas e as anotadas, manualmente (REF) ou pelo trabalho feito e descrito no Capítulo anterior, referentes aos *tweets* analisados e apresenta os resultados das métricas de coeficientes de concordância consideradas.

Na segunda fase, usaram-se os resultados provenientes do teste com os dados T2 e extraiu-se as emoções classificadas com maior confiança (probabilidade superior a 0,975) e os correspondentes *tweets*, dando origem ao *dataset* T3.

Na última fase, criou-se um novo modelo de Regressão Logística e dividiu-se o T3 em dados para treino (80%) e para teste (20%), de seguida transformaram-se os *tweets* do conjunto de treino em *features* e utilizaram-se para treinar o modelo criado. Os 20% restantes

dos dados foram utilizados para testar o modelo. Este modelo foi testado com os *datasets* AM e AA, com o mesmo propósito da primeira fase e, também, passam pelo mesmo processo de Avaliação. Os resultados dos todos os testes realizados na primeira e última fase encontram-se descritos na Secção 5.5.

5.3 Regressão Logística

A Regressão Logística é o algoritmo escolhido para a aprendizagem supervisionada abordada neste Capítulo, por corresponder à classificação de Máxima Entropia para eventos independentes [7].

Um modelo de Máxima Entropia estima a probabilidade condicional, com base nas *features* correspondentes a x não necessariamente contínuas, de uma variável categórica e/ou aleatória. Neste caso a variável é binária e portanto as classes (C) consideradas são 0 (não tem emoção) e 1 (tem emoção), se o valor for igual ou superior a 0,5 a previsão é 1, caso contrario é 0. Assim sendo, o modelo obtém:

$$\arg \max_{y \in C} P(y, x) \quad (5.1)$$

Esta probabilidade é calculada através da combinação linear entre as *features* (x) e o seu peso (p), sendo n o número total dos mesmos:

$$p_0 + p_1x_1 + \dots + p_nx_n = \sum_{i=1}^n p_ix_i \quad (5.2)$$

Para garantir que o resultado deste somatório é positivo e uma probabilidade, ou seja, um valor entre 0 e 1, é aplicada a função exponencial e um fator normalizador:

$$p(y|x) = \frac{\exp(\sum_{i=1}^n p_ix_i)}{\sum_{y'} \exp(\sum_{i=1}^n p_ix_i)} \quad (5.3)$$

O modelo também associa as *features* às classes consideradas, através de funções $f_i(y, x)$. A equação final utilizada para a criação de um modelo de Máxima Entropia com o objetivo de obter uma classificação de 0 e 1, é:

$$p(y|x) = \frac{\exp(\sum_{i=1}^n p_if_i(y, x))}{\sum_{y'} \exp(\sum_{i=1}^n p_if_i(y', x))} \quad (5.4)$$

Total Dados	Features	Raiva	Antecip.	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
1M	15k	0,946	0,943	0,954	0,945	0,951	0,963	0,942	0,947
1M	25k	0,954	0,946	0,960	0,950	0,954	0,968	0,942	0,950
2M	25k	0,960	0,951	0,966	0,955	0,957	0,973	0,951	0,953
2M	40k	0,964	0,952	0,969	0,957	0,958	0,973	0,953	0,955
3M	25k	0,962	0,951	0,967	0,958	0,958	0,972	0,952	0,955
3M	40k	0,968	0,954	0,972	0,962	0,960	0,975	0,955	0,957
4M	40k	0,968	0,955	0,973	0,962	0,961	0,976	0,957	0,958
4M	45k	0,969	0,955	0,974	0,962	0,961	0,977	0,958	0,959
5M	40k	0,969	0,955	0,974	0,963	0,961	0,977	0,958	0,959
5M	45k	0,971	0,955	0,975	0,964	0,961	0,977	0,958	0,959

Tabela 5.1: Combinações de diferentes quantidades de *tweets* e número máximo de *features*, com os respetivos valores da medida f1.

5.4 Análise do Impacto do Tamanho do Conjunto de Treino e do Número de *Features*

Para se poder treinar um modelo é necessário escolher o número de dados a utilizar e o número máximo de *features* para se transformar os *tweets* no formato pretendido e aceite pela Regressão Logística. Para tal, foram executadas algumas experiências com diferentes quantidades de dados e *features*.

Na Tabela 5.1 observa-se o resultado de algumas dessas experiências, sendo que é possível obter uma percentagem superior, da métrica f1, aumentando o número de dados ou o número máximo de *features* que representam os mesmos.

No entanto e como previsto, a combinação ótima - tempo/previsão do resultado - é diferente de emoção para emoção, porque a complexidade, os dados específicos de treino e mesmo a visão da emoção, de pessoa para pessoa, são diferentes.

Optou-se por utilizar em todos os conjuntos de teste 40 mil para o número máximo de *features*, escolhidas internamente; conjuntos com, pelo menos, 2 milhões de dados; e 5 milhões de dados como base para o segundo modelo, visto este ser uma tentativa de melhor o desempenho do Modelo 1.

5.5 Resultados

Como mencionado na Secção 5.2 foram realizados dois testes de avaliação a cada um dos modelos: Avaliação Automática (AA) e REF. O *dataset* AA permite comparar os modelos entre si, enquanto REF combina as anotações manuais (AM-1 e AM-2, descritas na Secção

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
AA - Avaliação automática								
Precisão	0,963	0,947	0,970	0,955	0,957	0,974	0,955	0,955
Cobertura	0,969	0,960	0,973	0,968	0,964	0,978	0,961	0,958
Medida f1	0,966	0,954	0,972	0,962	0,960	0,976	0,958	0,957
REF								
Precisão	0,348	0,102	0,412	0,021	0,526	0,491	0,195	0,219
Cobertura	0,686	0,828	0,675	0,667	0,475	0,562	0,633	0,885
Medida f1	0,462	0,182	0,512	0,040	0,499	0,524	0,298	0,351

Tabela 5.2: Métricas de avaliação do Modelo 1.

	Raiva	Antecipação	Nojo	Medo	Alegria	Tristeza	Surpresa	Confiança
AA - Avaliação automática								
Precisão	1,000	0,932	0,980	0,952	0,947	0,980	0,945	0,931
Cobertura	0,996	0,886	0,861	0,876	0,902	0,893	0,886	0,890
Medida f1	0,998	0,909	0,916	0,912	0,924	0,934	0,914	0,910
REF								
Precisão	0,339	0,102	0,426	0,307	0,500	0,510	0,195	0,222
Cobertura	0,621	0,821	0,628	0,254	0,446	0,543	0,625	0,867
Medida f1	0,439	0,181	0,508	0,278	0,472	0,526	0,297	0,354

Tabela 5.3: Métricas de avaliação do Modelo 2.

3.3) numa única referência — sempre que os anotadores não concordarem o valor considerado é o retornado pelo classificador do Capítulo 4. Ambos detém o objetivo comum de avaliar a aprendizagem ocorrida durante o treino dos modelos e perceber o desempenho dos mesmos.

As Tabelas 5.2 e 5.3 mostram, para cada um dos modelos, os valores das métricas de avaliação dos testes realizados.

O primeiro modelo demonstra uma grande unanimidade, por emoção, nas três métricas utilizadas com valores superiores a 95% no teste AA e uma média de 96% de precisão e de 96,6% de cobertura. As emoções tristeza e nojo foram as previstas com maior sucesso, e antecipação, confiança e surpresa as com menor. O que é um facto curioso e aponta, através da correlação entre o número de *tweets* classificados por emoção no dataset utilizado para testar o modelo (Tabela 5.1) e a percentagem de acerto nas emoções, para a real facilidade ou dificuldade da extração de uma emoção específica, como é o caso dos ótimos resultados de raiva que é a segunda emoção menos presente no *dataset* de treino e os não tão bons de antecipação mesmo bastante presente no *dataset* de treino. Relativamente à

REF, a precisão apresenta na sua maioria valores baixos, especialmente para medo, com uma média de 28,9% de precisão e a cobertura indica valores médio altos com uma média de 67,6% de cobertura. Segundo esta referência o Modelo 1 foi melhor sucedido a prever as emoções tristeza e nojo, e pior a prever medo e antecipação. O teste AA realizado ao Modelo 2 apresenta bons resultados, uma média de 95,8% de precisão e de 89,9% de cobertura. A emoção classificada com mais êxito foi a raiva, e as com menos sucesso foram a antecipação, a confiança, o medo e a surpresa. Os resultados obtidos através do teste composto pelos *tweets* anotados manualmente indicam uma média de 32,5% de precisão e 60,1% de cobertura. A REF demonstra um maior sucesso na classificação das emoções tristeza e nojo, e um menor nas emoções medo e antecipação.

Comparando os dois modelos e relativamente ao *dataset* de teste AA, ambos apresentam melhores resultados a nível de precisão e encontram-se praticamente ao mesmo nível, sendo que o melhor depende da emoção a analisar. Contrariamente ao esperado, segundo a média da medida f1, o Modelo 1 (96,3%) é melhor que o Modelo 2 (92,7%). Estes valores podem ser explicados pela possível diminuição da ocorrência das emoções no *dataset* de treino, no entanto a principal razão advém possivelmente do facto do *dataset* de treino do Modelo 1 ter sido etiquetado pelo classificador desenvolvido no Capítulo 4 bem como os dados do *dataset* de treino AA. Existe uma exceção, a emoção raiva revela excelentes e melhores resultados com o Modelo 2, com uma cobertura quase total, uma precisão de 100% de acerto no que diz respeito aos *tweets* classificados.

Os resultados do teste relativo à REF revelam que, considerando a média da medida f1, o segundo modelo (38,2%) é melhor que o primeiro (35,9%). Era esperada uma maior diferença entre os resultados, possivelmente o facto de no caso dos anotadores não concordarem a emoção referência ficar de acordo com a etiquetagem do classificador, pode ter beneficiado o Modelo 1. Por outro lado, este mesmo facto revela uma melhoria real na aprendizagem do Modelo 2. É de destacar a emoção medo, porque é a única que apresenta valores tão desfasados. Mesmo com poucos casos da emoção medo entre os 1000 *tweets* anotados, o Modelo 2 conseguiu com bastante sucesso classificar *tweets* com medo e estes, segundo a REF, conterem realmente esta emoção. De forma a aumentar a precisão das emoções seria necessário equilibrar a quantidade das mesmas na REF, dando especial atenção às emoções mais difíceis de classificar/reconhecer.

Os resultados são medianos, o que era esperado devido à dificuldade e complexidade da tarefa, e porque até os seres humanos têm dificuldade em classificar emoções em texto, comprovado pelo baixo nível de acordo entre os anotadores (Tabela 3.2). Assim sendo, é perceptível que cada emoção deva ser estudada e analisada individualmente porque podem coexistir ao mesmo tempo e ser melhor percebidas por métodos de classificação diferentes.

Em geral, analisando os resultados e as possíveis influências e causas do *dataset* de treino e da referência utilizada, bem como o número de emoções presentes no *dataset* inicial, conclui-se que o Modelo 2 apresenta uma ótima aprendizagem em muitas das emoções.

Capítulo 6

Conclusões e Trabalho Futuro

Esta dissertação foi sustentada pelo interesse nas áreas (PLN, Inteligência Artificial e às emoções associadas ao ser humano, que acabam por ser o nosso estímulo para quase tudo) e pela curiosidade traduzida nas questões de investigação apresentadas no início da mesma. Apresenta, descreve e compara os dois trabalhos realizados com o objetivo de classificar emoções em *tweets* e responder às mesmas questões.

O modelo de emoções seguido em ambos os trabalhos é o de Plutchik. Este afirma existirem oito emoções básicas (raiva, antecipação, nojo, medo, alegria, tristeza, surpresa e confiança), as quais são estudadas e relatadas neste documento.

Devido à falta de *datasets* anotados com emoções foi necessário proceder-se ao pré-processamento dos *tweets* sobre a totalidade do *dataset*, através de algumas decisões que tornaram a sua informação mais clara e permitiram isolar as palavras e descobrir a raiz das mesmas para se chegar às suas diferentes formas gramaticais e assim aumentar a possibilidade da sua correspondência no léxico. Garantiu-se, assim, que os dados foram interpretados da melhor maneira e as suas informações e *features* extraídas corretamente.

No primeiro trabalho é desenvolvido um classificador automático que segue uma abordagem baseada em léxico e associa as suas emoções às palavras analisadas, tendo em conta o vocabulário e as informações que possui. Foram realizadas seis experiências com base em quatro léxicos (o EmoLex, versão inglesa e portuguesa, e dois léxicos refinados), avaliadas recorrendo a três referências. A referência 1 demonstra que a experiência 2 obteve o melhor desempenho em relação às restantes, com uma média de *media f1* de 28%. A referência 2 afirma que as experiências 4 e 5 detém o melhor desempenho com uma média de *media f1* de 17%. A referência 3 mostra que as experiências 2 e 4 apresentam o melhor desempenho com uma média de *media f1* 36%. Tendo em conta que a experiência 2 corresponde a um dos léxicos refinados e a 3 e 4 ao EmoLex, versão inglesa, aplicado sobre *tweets* portugueses traduzidos para inglês, admite-se que os dois métodos, atingem valores semelhantes, e um pode ser escolhido em detrimento do outro dependendo do problema. As descobertas sugerem, também, que algumas emoções são mais fáceis de detetar

do que outras. Alegria, tristeza, raiva e nojo foram as emoções previstas com mais sucesso utilizando esta abordagem. As mais difíceis são o medo e a antecipação,

O segundo trabalho centra-se na construção de um classificador automático que aprende de forma supervisionada. Para tal e tendo em conta a tarefa, o modelo seguido é o modelo de Máxima Entropia ou algoritmo de Regressão Logística. Foram realizadas duas experiências com recurso a dois modelos de Regressão Logística, avaliados por intermédio de dois *datasets* de treino. O primeiro foi treinado com parte dos dados do *dataset* etiquetados pelo classificador desenvolvido no primeiro trabalho. O segundo modelo foi treinado com dados etiquetados pelo primeiro modelo, numa tentativa de obter melhores resultados. O *dataset* Avaliação Automática (AA) apresenta uma média de medida f1 para o Modelo 1 de 96,3% e para o Modelo 2 de 92,7%. O *dataset* REF exibe uma média de medida f1 para o Modelo 1 de 35,9% e para o Modelo 2 de 38,2%. Conclui-se que o Modelo 1 apresenta um melhor desempenho no *dataset* de teste AA, no entanto os resultados são mais equilibrados no *dataset* de teste REF. Contrariamente ao esperado, os resultados obtidos sugerem que, na generalidade das emoções e das métricas, o primeiro modelo é melhor que o segundo. No entanto, os resultados com o *dataset* REF são semelhantes e mostram que o Modelo 2 é ligeiramente superior. As emoções previstas com mais sucesso usando esta abordagem foram a tristeza, o nojo e a raiva. E, as mais difíceis são a antecipação, o medo, a surpresa e a confiança.

Comparando as abordagens estudadas, através das experiências feitas e dos resultados apresentados, conclui-se que a abordagem que possibilita uma melhor previsão das emoções existentes em textos, especificamente informais e não estruturados, é a que tem por base a aprendizagem supervisionada.

As três referências utilizadas para avaliar o primeiro trabalho e o *dataset* REF usado no segundo, tiveram origem num *dataset* de 1000 *tweets* anotado manualmente.

Com o final destes trabalhos, os dados e os resultados estudados, construiu-se com sucesso dois classificadores de emoções e adquiriu-se o conhecimento necessário para responder às questões de investigação.

A maior alteração ao trabalho pensado inicialmente é a exploração de *emojis* e *emoticons* como uma característica extra na identificação e previsão das emoções. Após a exploração inicial dos *emojis* e *emoticons* e do seu uso em *tweets* foram perceptíveis as irregularidades existentes e causadas devido ao mau uso das mesmas. A maioria das pessoas que os utilizam não conhecem os seus significados utilizando-os incorretamente ou dando-lhes um significado elas próprias, que pode ou não ser de conhecimento geral na sua faixa etária ou social.

Como trabalho futuro aponta-se os seguintes pontos:

- Exploração de *emojis* e *emoticons*. Apesar da exploração inicial e conclusão, pode adicionar contexto, no sentido emocional, aos textos e são realmente utilizados em

grande escala nas redes sociais. Pelo que, possibilitem a obtenção de bons resultados;

- Anotação manual de um *dataset* maior e por um número ímpar de anotadores, pelo menos três. Permitindo obter resultados mais significativos e a dominância de uma das classes (emoção ou não emoção);
- Utilização técnicas de pré-processamento diferentes. A distância entre palavras, por exemplo, pode ser uma boa adição na previsão das emoções;
- Diferentes algoritmos na abordagem supervisionada. Como observado existem técnicas e abordagens que funcionam melhor para umas emoções do que para outras, e seria um ponto interessante para perceber essa diferença;
- Uma aplicação a um caso real. O afinamento do trabalho para um caso real oferece contexto e pode adicionar alguma informação preciosa, mesmo que não diretamente para a previsão em si e numa fase inicial, mas para o estudo das emoções e da exploração das suas *features* ótimas.

Bibliografia

- [1] Sattam Almatarneh e Pablo Gamallo. “A lexicon based method to search for extreme opinions”. Em: *PLoS ONE* 13.5 (2018), pp. 1–19. ISSN: 19326203. DOI: [10.1371/journal.pone.0197816](https://doi.org/10.1371/journal.pone.0197816).
- [2] Nurul Anisa, Sri Winarsih e Catur Supriyanto. “07873824.Pdf.Pdf”. Em: (2016), pp. 130–133.
- [3] Muhammad Zubair Asghar et al. “Performance Evaluation of Supervised Machine Learning Techniques for Efficient Detection of Emotions from Online Content”. Em: (ago. de 2019). DOI: [10.20944/preprints201908.0019.v1](https://doi.org/10.20944/preprints201908.0019.v1). URL: <http://dx.doi.org/10.20944/preprints201908.0019.v1>.
- [4] Stefano Baccianella, Andrea Esuli e Fabrizio Sebastiani. “SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining”. Em: *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. Valletta, Malta: European Language Resources Association (ELRA), mai. de 2010. URL: http://www.lrec-conf.org/proceedings/lrec2010/pdf/769_Paper.pdf.
- [5] Carla Bazzanella. “Emotions, Language, and Context”. Em: *Emotion in dialogic interaction*. 2004. Cap. Advances i, pp. 55–72.
- [6] Aaron Ben-Ze'ev. *The subtlety of emotions*. MIT Press, 2000.
- [7] Adam L. Berger, Stephen A. Della Pietra e Vincent J. Della Pietra. “A Maximum Entropy Approach to Natural Language Processing”. Em: *Computational Linguistics* 22.1 (1996), pp. 39–71. URL: <https://www.aclweb.org/anthology/J96-1002>.
- [8] Peter J. Bradley, Margaret M. ; Lang. *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Rel. téc. University of Florida, 1999.
- [9] E. Cambria, Daniel J. Olsher e Dheeraj Rajagopal. “SenticNet 3: A Common and Common-Sense Knowledge Base for Cognition-Driven Sentiment Analysis”. Em: *AAAI*. 2014.
- [10] Erik Cambria et al. “AffectiveSpace : Blending common sense and affective knowledge to perform emotive reasoning”. Em: *WOMSA* (2009), pp. 32–41.

- [11] Alexis Conneau et al. “Word Translation Without Parallel Data”. Em: *arXiv preprint arXiv:1710.04087* (2017).
- [12] Luis Duarte, Luís Macedo e Hugo Oliveira. “Exploring Emojis for Emotion Recognition in Portuguese Text”. Em: *Progress in Artificial Intelligence*. Ed. por Paulo Moura Oliveira, Paulo Novais e Luís Paulo Reis. Cham: Springer International, 2019, pp. 719–730. ISBN: 978-3-030-30244-3.
- [13] Vachagan Gratian e Marina Haid. “BrainT at IEST 2018 : Fine-tuning Multiclass Perceptron For Implicit Emotion Classification”. Em: (2018), pp. 243–247.
- [14] Eva Hudlicka. “Guidelines for Designing Computational Models of Emotions”. Em: *International Journal of Synthetic Emotions 2.1* (jan. de 2011), pp. 26–79. ISSN: 1947-9093. DOI: [10.4018/jse.2011010103](https://doi.org/10.4018/jse.2011010103). URL: <http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/jse.2011010103>.
- [15] Angela Jusupova, Fernando Batista e Ricardo Ribeiro. “Characterizing the Personality of Twitter Users based on their Timeline Information”. Em: *Atas da 16 Conferência da Associação Portuguesa de Sistemas de Informação*. 2016. URL: <https://ciencia.iscte-iul.pt/publications/characterizing-the-personality-of-twitter-users-based-on-their-timeline-information/30881>.
- [16] Sophia Lee. “A Linguistic Analysis of Implicit Emotions”. Em: *Chinese Lexical Semantics*. 2015, pp. 185–194. ISBN: 9783319271934.
- [17] “Linguistic inquiry and word count: LIWC2001”. Em: 2001. ISBN: 9781609607418. DOI: [10.4018/978-1-60960-741-8.ch012](https://doi.org/10.4018/978-1-60960-741-8.ch012).
- [18] Maria Inês Maia e José Paulo Leal. “An Emotional Word Analyzer for Portuguese”. Em: 17 (2017), pp. 1–14. ISSN: 21906807. DOI: [10.4230/OASICS.SLATE.2017.17](https://doi.org/10.4230/OASICS.SLATE.2017.17). URL: <http://www.tei-c.org/index.xml>.
- [19] Saif M. Mohammad. “From once upon a time to happily ever after: Tracking emotions in mail and books”. Em: *Decision Support Systems* 53.4 (2012), pp. 730–741. ISSN: 01679236. DOI: [10.1016/j.dss.2012.05.030](https://doi.org/10.1016/j.dss.2012.05.030).
- [20] Saif M. Mohammad e Svetlana Kiritchenko. “Using hashtags to capture fine emotion categories from tweets”. Em: *Computational Intelligence* 31.2 (2015), pp. 301–326. ISSN: 14678640. DOI: [10.1111/coin.12024](https://doi.org/10.1111/coin.12024).
- [21] Saif M. Mohammad e Peter D. Turney. *Crowdsourcing a Word-Emotion Association Lexicon*. 2013. arXiv: [1308.6297 \[cs.CL\]](https://arxiv.org/abs/1308.6297).
- [22] Saif M Mohammad e Peter D Turney. “Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon”. Em: *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, Los Angeles, California, June 2010* (2010), pp. 26–34. URL: papers3://publication/uuid/06948E35-4B20-4C3A-BD8C-CAEA7A6377F4.

- [23] Silvia M.W. Moraes et al. "Comparing approaches to subjectivity classification: A study on Portuguese tweets". Em: *Computing Processing of the Portuguese Language*. Vol. 9727. Springer International, 2016, pp. 86–94. ISBN: 9783319415512. DOI: [10.1007/978-3-319-41552-9_8](https://doi.org/10.1007/978-3-319-41552-9_8). URL: https://books.google.pt/books?id=T%7B%5C_%7Dd6DAAAQBAJ%7B%5C%7Dpg=PA86%7B%5C%7Dlpg=PA86%7B%5C%7Ddq=Comparing+Approaches+to+Subjectivity+Classification:+A+Study+on+Portuguese+Tweets%7B%5C%7Dsource=bl%7B%5C%7Dots=Jim9Gw08a8%7B%5C%7Dsig=ACfU3U3vaCyMtU29cooHM7nL7ccYBhMAaQ%7B%5C%7Dhl=pt-PT%7B%5C%7Dsa=X%7B%5C%7Dved=2ahUKEwi73-TJ3tLhAhWxx4UK.
- [24] Cataldo Musto, Giovanni Semeraro e Marco Polignano. "A comparison of lexicon-based approaches for sentiment analysis of microblog posts". Em: *CEUR Workshop Proceedings* 1314 (2014), pp. 59–68. ISSN: 16130073.
- [25] Kimberly A. Neuendorf. *The Content Analysis Guidebook*. 2002.
- [26] Petra Kralj Novak et al. "Sentiment of emojis". Em: *PLoS ONE* 10.12 (2015). ISSN: 19326203. DOI: [10.1371/journal.pone.0144296](https://doi.org/10.1371/journal.pone.0144296). URL: <http://www.perceptionanalytics.net>.
- [27] P. Ekman. "An Argument for Basic Emotions". Em: *Cognition and Emotion* 6.May 2013 (1992), pp. 169–200.
- [28] James W. Pennebaker et al. *Linguistic Inquiry and Word Count: LIWC2015*. Sage Publications, 2015. DOI: [10.4018/978-1-60960-741-8.ch012](https://doi.org/10.4018/978-1-60960-741-8.ch012).
- [29] James W. Pennebaker et al. *The Development and Psychometric Properties of LIWC2007*. 2007, pp. 1–22. URL: <http://www.liwc.net/LIWC2007LanguageManual.pdf>.
- [30] Robert Plutchik. "A psychoevolutionary theory of emotions". Em: *Social Science Information* 21.4-5 (1982), pp. 529–553. DOI: [10.1177/053901882021004003](https://doi.org/10.1177/053901882021004003). eprint: <https://doi.org/10.1177/053901882021004003>. URL: <https://doi.org/10.1177/053901882021004003>.
- [31] Aisulu Rakhmetullina, Dietrich Trautmann e Georg Groh. "Distant supervision for emotion classification task using emoji2emotion". Em: *CEUR Workshop Proceedings* 2130 (2018). ISSN: 16130073.
- [32] Jaime Redondo et al. "The Spanish adaptation of anew (Affective Norms for English Words)". Em: *Behavior Research Methods* 39.3 (2007), pp. 600–605. ISSN: 1554351X. DOI: [10.3758/BF03193031](https://doi.org/10.3758/BF03193031).
- [33] Phillip Shaver et al. "Emotion Knowledge: Further Exploration of a Prototype Approach". Em: *Journal of Personality and Social Psychology* 52.6 (1987), pp. 1061–1086. ISSN: 00223514. DOI: [10.1037/0022-3514.52.6.1061](https://doi.org/10.1037/0022-3514.52.6.1061).

- [34] Silva Fernando, Carvalho Ariadne e Roman Norton. "Identifying Emotions in Tweets related to the Brazilian Stock Market". Tese de doutoramento. Institute of Computing, 2016. URL: <https://pt.slideshare.net/FernandoJVdaSilva/identifying-emotions-in-tweets-related-to-the-brazilian-stock-market>.
- [35] Ana Paula Soares et al. "The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese". Em: *Behavior Research Methods* 44.1 (2012), pp. 256–269. ISSN: 1554351X. DOI: [10.3758/s13428-011-0131-7](https://doi.org/10.3758/s13428-011-0131-7).
- [36] Carlo Strapparava e Alessandro Valitutti. "WordNet Affect: an Affective Extension of WordNet". Em: *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*. Lisbon, Portugal: European Language Resources Association (ELRA), mai. de 2004. URL: <http://www.lrec-conf.org/proceedings/lrec2004/pdf/369.pdf>.
- [37] Min Wang e Xiaobing Zhou. "Yuan at SemEval-2018 Task 1: Tweets Emotion Intensity Prediction using Ensemble Recurrent Neural Network". Em: *Proceedings of The 12th International Workshop on Semantic Evaluation* (2018), pp. 205–209. DOI: [10.18653/v1/S18-1031](https://doi.org/10.18653/v1/S18-1031). URL: <http://aclweb.org/anthology/S18-1031>.
- [38] Wenbo Wang et al. "Harnessing twitter 'big data' for automatic emotion identification". Em: *Proceedings - 2012 ASE/IEEE International Conference on Privacy, Security, Risk and Trust and 2012 ASE/IEEE International Conference on Social Computing, SocialCom/PASSAT 2012* (2012), pp. 587–592. ISSN: 978-1-4673-5638-1. DOI: [10.1109/SocialCom-PASSAT.2012.119](https://doi.org/10.1109/SocialCom-PASSAT.2012.119).
- [39] Louis A. Wehenkel. *Automatic Learning Techniques in Power Systems*. 1998.
- [40] Janyce Wiebe, Theresa Wilson e Claire Cardie. "Annotating expressions of opinions and emotions in language". Em: *Language Resources and Evaluation* 39.2-3 (2005), pp. 165–210. ISSN: 1574020X. DOI: [10.1007/s10579-005-7880-9](https://doi.org/10.1007/s10579-005-7880-9).