

## Quantificando a Importância de Emojis e Emoticons para Identificação de Polaridade em Avaliações Online

Hildon L. de Paula<sup>1</sup>, Bruno Á. Souza<sup>1</sup>,  
Fabiola G. Nakamura<sup>1</sup>, Eduardo F. Nakamura<sup>1</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal do Amazonas (UFAM)  
Manaus – AM – Brasil

{help,bruno.abia,fabiola,nakamura}@icomp.ufam.edu.br

**Abstract.** *Virtual environments, such as online stores (e.g. Amazon, Google Play, Booking), promote a collaborative strategy for reviewing products and services. The users' opinions represent their degree of satisfaction regarding the reviewed item. The set of reviews of an item serves as a reputation index. Hence, the automatic identification of the user satisfaction, regarding an item, based on his/her textual review, is a tool of great economic and strategic potential for enterprises. In this context, the growing adoption of emojis and emoticons, boosted by the mobile devices and their Apps, the users increasingly adopt such a vocabulary to express their opinion and sentiments. In this work, we present a quantitative assessment of the richness of emojis/emoticons to predict the users' opinion in product reviews in collaborative systems. Our proposal uses the Bag of Words with Support Vector Machine to predict the users' opinion in a online review, taking into account the use of: (1) only words; (2) words and emojis/emoticons; and (3) only emojis/emoticons. For certain scenarios, considering the F1 metric, the use of words and emojis results in an efficacy of 0.70, compared to 0.65, when only words are used (traditional approach).*

**Resumo.** *Ambientes virtuais como lojas online de produtos e serviços (e.g. Amazon, Google Play, Booking) adotam uma estratégia colaborativa de avaliação e reputação onde os usuários classificam os produtos e serviços. A opinião do usuário representa o grau seu de satisfação em relação ao item avaliado. O conjunto de avaliações de um item é referencial de sua reputação/qualidade. Portanto, a identificação automática da satisfação do usuário em relação a um item, considerando sua avaliação textual, é uma ferramenta com potencial econômico singular. Neste contexto, com a popularização de emojis e emoticons, intensificada pelo uso de dispositivos móveis e seus aplicativos, os usuários adotam cada vez mais estes símbolos como parte do vocabulário utilizado para expressar opiniões e sentimentos. Neste trabalho, apresentamos uma avaliação quantitativa da representatividade de emojis/emoticons para a identificação de opinião e polaridade em ambientes online de avaliação colaborativa. A abordagem proposta quantifica o uso da técnica Bag of Words com SVM para determinar o grau de satisfação do usuário em relação a um item, considerando: (1) palavras e emojis/emoticons; (2) apenas palavras; (3) apenas emojis/emoticons. Particularmente, para cenários específicos o uso de emojis/emoticons para a análise de sentimentos chega a ter uma eficácia de 0,70, contra 0,65 quando utilizamos apenas as palavras, considerando a métrica F1.*

## 1. Introdução

Ambientes virtuais como lojas online (e.g. Amazon, Google Play, Booking) e redes sociais para **avaliação**<sup>1</sup> de locais, estabelecimentos e serviços (e.g. Trip Advisor, Foursquare, Yelp, Vivino), permitem que usuários avaliem e compartilhem suas experiências [Almeida et al. 2016b]. Esse compartilhamento de opiniões define de forma colaborativa a reputação tanto dos estabelecimentos quanto dos produtos/serviços disponibilizados. Essa realimentação dos clientes representa uma ferramenta importante para que as empresas possam identificar oportunidades de melhoria e crescimento. Portanto, identificar automaticamente a polaridade (ou sentimento) de uma avaliação (e.g. uma a cinco estrelas, ou simplesmente positivo/negativo) expressa por um usuário é um problema com grande potencial econômico e estratégico para empresas.

Uma abordagem natural é o uso de técnicas de Análise de Sentimentos. A Análise de Sentimento, também conhecida como Mineração de Opinião, é o campo de estudo que analisa as opiniões de pessoas, sentimentos, avaliações e emoções [Liu 2012]. Os trabalhos envolvendo técnicas de Análise de Sentimento utilizam duas abordagens principais para determinar a polaridade de textos: técnicas baseadas em aprendizagem de máquina [Montejo-Ráez et al. 2014; Ye et al. 2009] e técnicas baseadas em análise léxica [Martins et al. 2015; Wilson et al. 2009]. Enquanto a primeira abordagem utiliza modelos de aprendizagem supervisionada, que necessitam de exemplos rotulados (conjunto de treinamento) para classificar a polaridade de textos; a segunda abordagem usa léxicos (um conjunto de recursos linguísticos com semântica bem definida) para determinar tal polaridade.

Entretanto, o volume de informações geradas a partir destas avaliações é grande. Como consequência, conforme descrito por Jo e Oh [2011], é difícil definir a melhor combinação de aspectos para analisar essas avaliações, devido à diversidade de: (1) estilos de escrita; (2) tamanho dos textos; (3) caracteres e símbolos especiais (emojis e emoticons); e (4) idiomas utilizados. Particularmente, emojis e emoticons são cada vez mais utilizados para expressar emoções como contentamento e descontentamento. Portanto, esses recursos linguísticos inerentes a ambientes virtuais possuem relevância na identificação de polaridade em um texto [Terrana et al. 2014].

Este trabalho apresenta duas contribuições principais: (1) a quantificação do uso de *bag-of-words* com SVM (*support vector machine*) para quantificar a satisfação dos usuários em avaliações online e (2) a avaliação da riqueza linguística de emoticons e emojis nesses textos para expressar a satisfação desses usuários.

O restante do artigo está organizado como descrito a seguir. A seção 2 apresenta os trabalhos relacionados. A seção 3 detalha a abordagem proposta neste artigo para quantificar a satisfação dos usuários em textos que possuem emojis e emoticons. A construção da base de dados, os experimentos e resultados são detalhados na seção 4. A seção 5 apresenta as conclusões e direções futuras.

---

<sup>1</sup>Este trabalho utiliza o termo **avaliação** como tradução livre dos termos em inglês: *review* (e.g. Amazon, Google Play) e *tip* (e.g. Foursquare).

## 2. Trabalhos Relacionados

A análise de polaridade é um tema frequente em computação [Aisopos et al. 2012; Liu 2012; Pang e Lee 2008] devido à sua aplicabilidade e potencial econômico. As abordagens mais conhecidas utilizam aprendizagem de máquina [Martins et al. 2015; Pang et al. 2002] ou análise léxica [Serrano-Guerrero et al. 2015; Wilson et al. 2009].

No contexto de avaliações de dados online, Pang et al. [2002] são pioneiros no uso de aprendizagem de máquina com técnicas supervisionadas para classificar avaliações de filmes como positivas ou negativas. Os autores representaram cada revisão como uma *bag-of-words* e alcançaram uma acurácia superior a 82,9% com o *Support Vector Machine*, usando unigramas e bigramas. Turney [2002] utilizou o mesmo cenário de revisões, empregando o *PMI-Information Retrieval* (PMI-IR) para determinar a orientação semântica de uma revisão. Turney utilizou 410 revisões de diferentes domínios e obteve uma acurácia de 65,8%.

Moraes et al. [2013] aplicaram análise de polaridade em textos curtos, comparando métodos supervisionados e não-supervisionados (léxicos) para textos em inglês compartilhados no Foursquare<sup>2</sup> (cada *tip* coletada foi considerada como uma *bag-of-words*). Os experimentos foram realizados em duas bases de dados, obtendo como resultado um empate de 69,3% entre o *Naive Bayes* e o modelo léxico, levando em consideração o F1 médio para ambas as bases de dados.

Souza et al. [2016] avaliaram o uso de Análise de Sentimentos para identificar a polaridade (positivo, negativo e neutro) de *tweets* a respeito do processo de impedimento da presidente do Brasil. Durante os experimentos os autores compararam três métodos de aprendizagem de máquina (SVM, Multinomial Naive Bayes e Max Entropy) para inferir a polaridade dos *tweets*. Em seus resultado, o SVM apresentou uma medida F1 de 93,0% para identificar as classes positivo e negativo. Na classe neutra os resultados apresentam uma medida F1 de 84,6%.

Terrana et al. [2014] propuseram um método supervisionado para classificar um *tweet* como positivo, negativo ou neutro. Os autores consideram os emoticons presentes nos textos, evitando qualquer interferência humana durante o processamento de rotulação. Como contribuição, é apresentado um método com uma acurácia média de 72,0%, mostrando que é possível analisar os tweets em tempo real a partir dos emoticons.

Reis et al. [2016] investigaram oito métodos de análise de sentimentos, implementados e disponibilizados no sistema *iFeel* [Araújo et al. 2014], para inferir a polaridade em textos compartilhados no jornal Reuters Online<sup>3</sup>. Os autores concluíram que o método *Sentistrength* foi o que apresentou melhor eficácia, alcançando uma acurácia de 74,0%; além de ter de apresentar uma característica interessante, permitindo analisar a força da polaridade em textos em uma escala de -5 (muito negativo) a 5 (muito positivo).

---

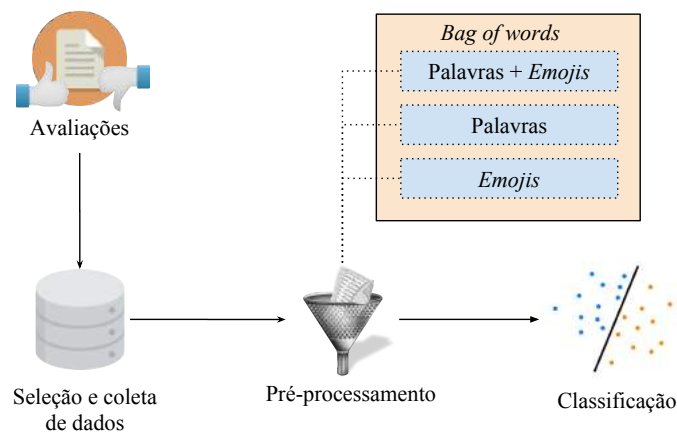
<sup>2</sup><https://pt.foursquare.com>.

<sup>3</sup><http://www.reuters.com>.

### 3. Abordagem Proposta

De forma sucinta, a abordagem proposta, ilustrada na figura 1, busca inferir a polaridade de avaliações a respeito de produtos e serviços publicados em ambientes virtuais, dando ênfase na utilização de emoticons e emojis e quantificando a riqueza linguística deste vocabulário para expressar polaridade e opinião. Para atingir esse objetivo a abordagem proposta inclui as seguintes etapas:

1. Seleção e coleta de dados;
2. Pré-processamento dos dados;
3. Classificação.



**Figura 1. Abordagem proposta.**

#### 3.1. Seleção e Coleta de dados

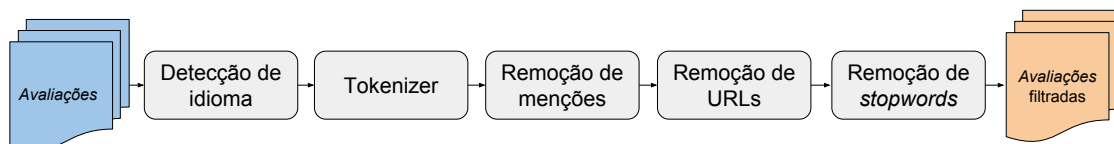
Ambientes virtuais de lojas online de produtos e serviços (e.g. Google Play e Amazon) são plataformas que permitem que usuários compartilhem suas opiniões e experiências através de avaliações públicas. A coleta destes dados públicos é particularmente interessante para validar técnicas de Análise de Sentimentos, pois estas avaliações possuem descrições textuais das opiniões e notas que quantificam com valores discretos (e.g. uma a cinco estrelas) a satisfação do usuário relatada no texto. Portanto, as próprias notas podem ser utilizadas como rótulos mais fidedignos das opiniões (sob a ótica do usuário). Como o objetivo deste trabalho é quantificar a riqueza de emojis e emoticons para determinar a polaridade de avaliações, foram consideradas apenas avaliações cujos textos possuíam pelo menos um emoji/emoticon. O detalhamento da base de dados e sua construção é detalhado na seção 4.

#### 3.2. Pré-Processamento

Após a coleta dos dados, os documentos (avaliações) são tratados para eliminar atributos não representativos das avaliações que foram coletadas. O pré-processamento proposto, ilustrado na figura 2, inclui as seguintes fases:

- A detecção de idioma descarta textos que não estejam em inglês;
- A função *Tokenizer* separa os documentos em *tokens* (segmentos de sentenças), permitindo processar os segmentos individualmente;

- A remoção de Menções elimina as citações a outros usuários e *hashtags* presentes nos documentos para que sejam avaliados apenas o texto e emojis/emoticons;
- As URLs foram removidas pelo mesmo motivo;
- A remoção de *stop words* elimina termos não representativos do documento (e.g. artigos, preposições e números).



**Figura 2. Etapas do pré-processamento.**

A seguir, as avaliações textuais são representadas usando a abordagem *bag-of-words* (**BoW**), onde cada documento é representado por um vetor de palavras que o compõe computando os valores de TF-IDF (*term frequency, inverted document frequency*) [Aisopos et al. 2012]. A abordagem geral de BoW com TF-IDF é a principal técnica utilizada em Análise de Sentimentos sobre textos e maiores detalhes podem ser conferidos no trabalho de Aisopos et al. [2012].

Para quantificar a riqueza linguística de emojis/emoticons são consideradas três análises:

1. **Palavras + emojis** - Uma representação de *bag-of-words* com ambos os termos;
2. **Palavras** - Uma representação de *bag-of-words* apenas com palavras;
3. **Emojis** - Uma representação de *bag-of-words* apenas com *Emojis*.

### 3.3. Classificação

Nessa etapa, utilizamos a representação dos dados criada pela BoW com TF-IDF e submetemos ao algoritmo de aprendizagem de máquina. Vale ressaltar, que uma característica importante de técnicas supervisionadas é que elas apresentam fundamentalmente o mesmo comportamento: dado um conjunto de dados de treinamento, composto por instâncias que são formadas por vários atributos previamente rotulados (polaridades) nas classes de interesse [Moraes et al. 2013], o modelo aprende características (base de treino) de cada classe e poderá ser usado para classificar outras amostras (base de teste).

O classificador adotado foi o SVM (*Support Vector Machine*) que, em uma avaliação prévia, apresentou maior eficácia.

## 4. Experimentos e Resultados

Nesta seção, é apresentada a metodologia experimental e são discutidos os resultados.

### 4.1. Metodologia

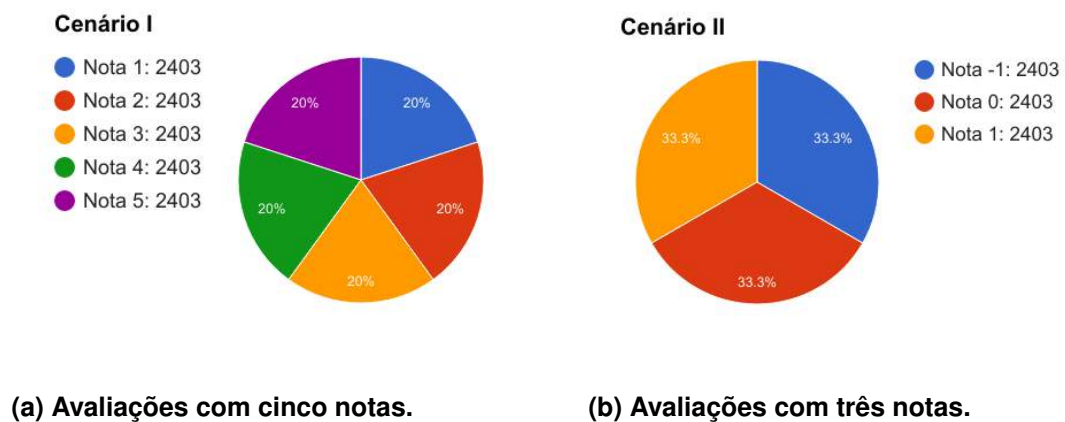
A base de dados foi construída com avaliações coletadas da Google Play (loja de aplicativos móveis para Android). As avaliações foram coletadas utilizando a Google API<sup>4</sup>, que permite captura de páginas de avaliações a partir do identificador do aplicativo e do

<sup>4</sup><https://github.com/facundoolano/google-play-scraper>.

idioma desejado. A Google API limita o acesso a 100 páginas de cada aplicativo por hora. Além das avaliações de aplicativos, a base inclui também avaliações publicadas na Amazon, disponibilizadas por Leskovec e Krevl [2014]; McAuley et al. [2015]. Todas as avaliações possuem uma nota entre um e cinco. Na base foram incluídas apenas avaliações com pelo menos um emoji/emoticon, dado o objetivo deste trabalho. O total de avaliações da Google Play é de 4.234 (35,3 %), e o número de avaliações da Amazon é de 7.781 (64,7 %).

A base de dados utilizada nos experimentos foi organizada de duas formas, representando dois cenários típicos para Análise de Sentimentos (figura 3):

- **Cenário I.** Esse cenário considera cinco possibilidades de nota para cada avaliação: nota um expressa a avaliação mais negativa possível; nota cinco expressa a avaliação mais positiva possível.
- **Cenário II.** Esse cenário considera três possibilidades de nota para cada avaliação: -1 (negativo), 0 (neutro) e 1 (positivo). A classe da nota -1 contém 50% de avaliações oriundas da nota 1 e 50% da nota 2 da base original, sorteados aleatoriamente. De forma análoga, a classe da nota 1 do cenário II contém 50% de avaliações oriundas da nota 4 e 50% da nota 5 da base original. A classe da nota 0 inclui as instâncias da nota 3 da base original.



**Figura 3. Distribuição de dados para os dois cenários considerados.**

Na fase de pré-processamento os emojis foram transformados em palavras (conforme ilustrado a figura 4), pois assim podem ser processados para a geração da *bag-of-words*. Essa transformação tem como objetivo agrupar emojis com o mesmo significado, simplificando o vocabulário.

Para cada tipo de pré-processamento, os experimentos foram realizados em dois cenários conforme descrito na seção anterior. No primeiro as avaliações tinham cinco rótulos, onde 1 significa nada satisfeito e 5 muito satisfeito. No segundo cenário, a classificação se dava apenas em negativo, positivo e neutro (-1,0, e 1) como a maioria dos trabalhos na área de detecção de polaridade [Almeida et al. 2016a]. Estes experimentos ajudaram a comparar o desempenho do método, e observar onde ele teve o melhor desempenho.





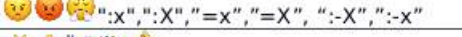



Emojis	Palavra
  ":)"", ":D", ":)", "=)", "=D", "=]", ":-)", ":-]"", ":3", '=3'	emojixsmile
 ":((", ":/", ":[", ":-(", ":-[", ":'(", ":'["", "=((", "=[("	emojixsad
 "<3"	emojixlove
 ":X", ":X", "=X", "=X", ":-X", ":-X"	emojixangry
	emojixpositive
	emojixnegative
 ...	emojixneutral

Figura 4. Substituição de *Emojis* por palavras.

Os experimentos foram conduzidos utilizando técnicas de representação de características tradicionais, tendo como entrada as avaliações textuais e suas respectivas pontuações. A seguir, o texto é pré-processado e é executada a vetorização e transformação em uma BoW utilizando TF-IDF (*term frequency–inverse document frequency*) para identificação dos termos significativos de cada texto, e por fim este vetor de características é fornecido como entrada ao SVM.

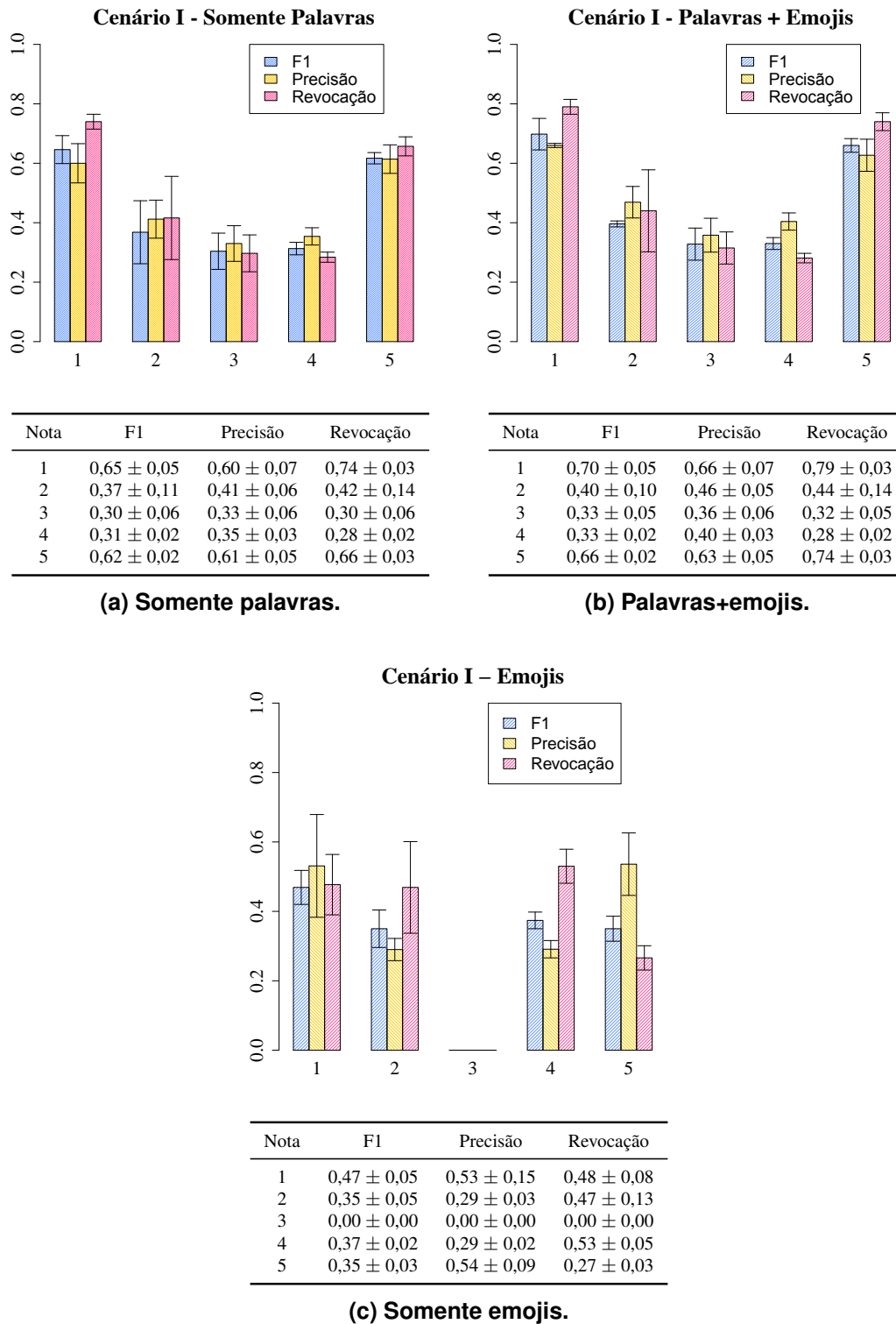
Visando quantificar a relevância de emojis e emoticons como recursos linguísticos (ainda que representem uma linguagem informal) para expressão de sentimentos e opiniões, cada cenário avalia a abordagem de Análise de Sentimentos considerando:

- **Somente Palavras.** Abordagem BoW tradicional que considera apenas palavras (emojis e emoticons que originalmente estavam presentes no texto são descartados).
- **Palavras+Emojis.** Abordagem que considera emojis e emoticons como palavras do vocabulário.
- **Somente Emojis.** Abordagem que considera somente os emojis e emoticons como palavras do vocabulário.

Os experimentos adotam uma validação cruzada de dez grupos (*ten fold cross validation*). Os valores médios apresentados foram complementados com os intervalos de confiança para um grau de 95% de confiança. Esses intervalos são apresentados nos gráficos através de barras de erros.

#### 4.2. Resultados e Avaliações

Os resultados e as avaliações estão organizados por cenário.



**Figura 5. Cenário I: F1, Precisão e Revocação.**

#### 4.2.1. Cenário I: Cinco Níveis/Notas

Considerando o uso de palavras somente, a **acurácia geral média** foi de **47,4% ± 4,05**, contra **50,7% ± 3,9** para palavras e emojis, representando um ganho médio de 6,9% (i.e.



3,3 pontos percentuais). Ao analisar a eficácia de classificação por classe, ilustrada na figura 5, fica evidente a maior facilidade para identificar avaliações com nota mínima e nota máxima. Em particular, ao incluir emojis/emoticons, o ganho em termos de F1 para a nota 1 é de 7,7% (0,65 contra 0,70). Outro fato importante é que ao utilizarmos apenas emojis, descartando outras palavras, o BoW não foi capaz de identificar a classe referente à nota 3 em nossa base, conforme ilustrado na figura 5(c). Além disso, o ganho ao combinar palavras e emojis é maior do que o uso isolado de emojis.

A matriz média de confusão para cada abordagem é apresentada na figura 6. Em ambos os casos, fica claro que a confusão/erro mais comum tem distância um. Por exemplo, a classe nota 3 é mais frequentemente confundida com nota 2 ou nota 4, mas raramente confundida com nota 1 ou nota 5, conforme evidencia a coluna nota 3 (prevista).

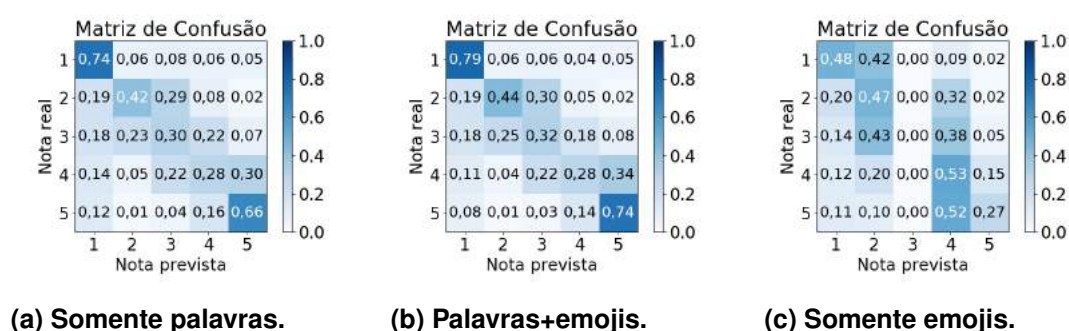


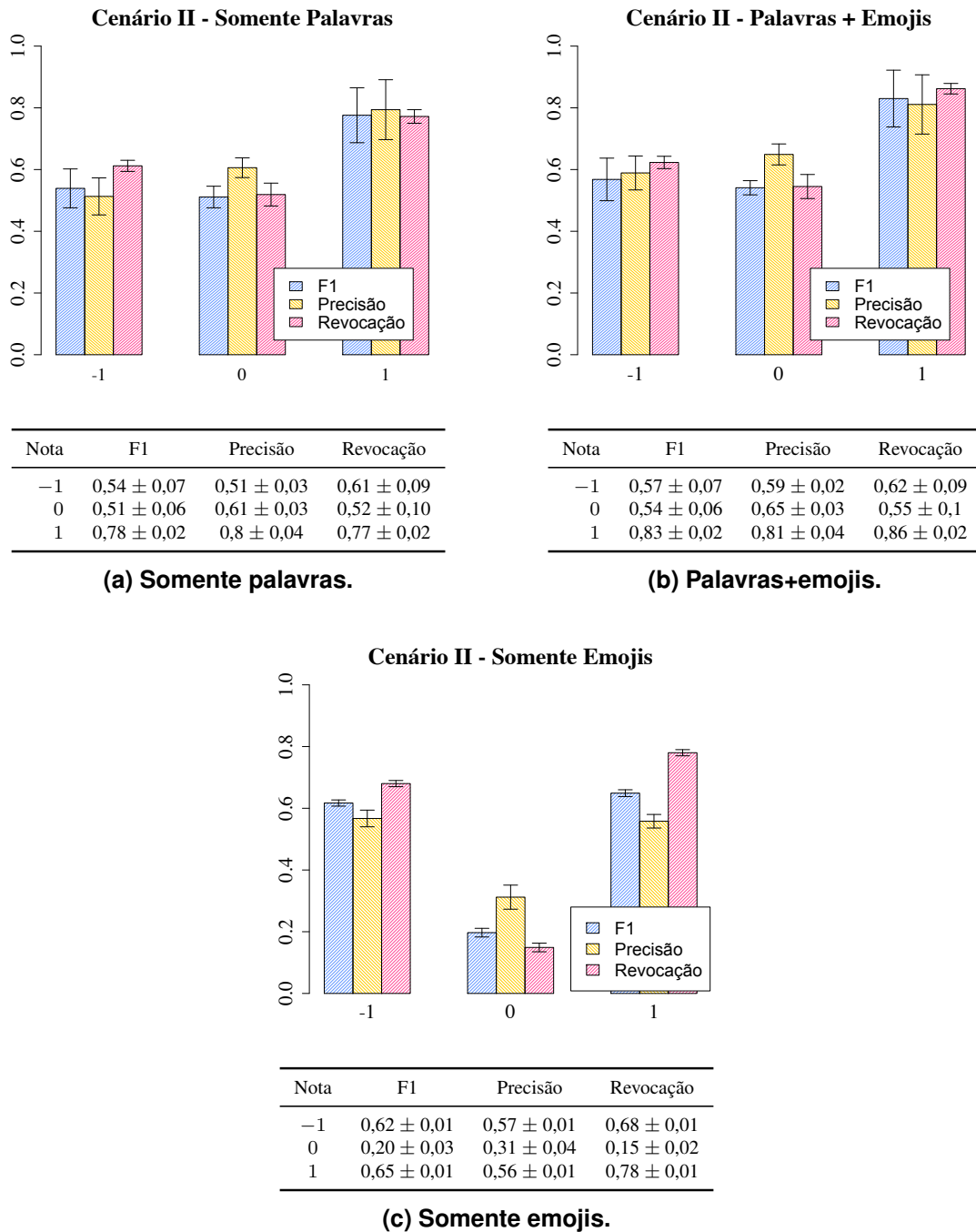
Figura 6. Cenário I: Matrizes de confusão.

Os resultados evidenciam a dificuldade para prever a nota de uma avaliação quando são considerados cinco valores possíveis. Essa dificuldade deve ser intensificada se a escala de notas for maior, e.g. de um a dez. Contudo, claramente os emojis/emoticons, utilizados para representar emoções, agregam informação relevante que é capturada pela abordagem BoW.

#### 4.2.2. Cenário II: Três Níveis

Considerando o uso de palavras somente, a **acurácia geral média** foi de **67,1% ± 1,7**, contra **62,9% ± 1,9** para palavras e emojis, representando um ganho médio de 6,7% (i.e. 4,2 pontos percentuais). A eficácia de classificação por classe é ilustrada na figura 7, fica evidente a maior facilidade para identificar avaliações com nota mínima e nota máxima. Contudo, o uso de emojis possui impacto menor, quando comparado ao problema de cinco classes do cenário I. Isso ocorre porque o problema de três classes é melhor resolvido pelo BoW, independentemente do uso de emojis/emoticons. Novamente, para três classes, o ganho ao combinar palavras e emojis também é maior do que o uso isolado de emojis.

A figura 8 apresenta as matrizes de confusão para as duas abordagens. As matrizes mostram que ao utilizar somente palavras (abordagem tradicional), o algoritmo confunde a classe 1 (positivo) como  $-1$  (negativo) mais do que a abordagem usando emojis: a célula (1,  $-1$ ) da matriz “somente palavras” (figura 8(a)) é mais densa que a da matriz “palavras+emojis”(figura 8(b)), representando uma redução de 53,3% de erro (0,15 contra 0,07).



**Figura 7. Cenário II: F1, Precisão e Revocação.**

## 5. Conclusão

Neste trabalho quantificamos a relevância de emojis/emoticons como vocabulário utilizado para expressão de opinião em avaliações online. Foram consideradas avaliações coletadas da Amazon e da Google Play. Dois cenários foram considerados assumindo escalas de cinco níveis (1 a 5) e três níveis (-1, 0 e 1). O modelo adotado para previsão de notas, com base nas avaliações, foi o BoW (*Bag of Words*). Os ganhos de eficácia na previsão chegaram a 7,7% (0,65 contra 0,70), considerando a métrica F1. O uso de emojis combinados às palavras trouxe para o modelo um ganho maior do que a eficácia do uso

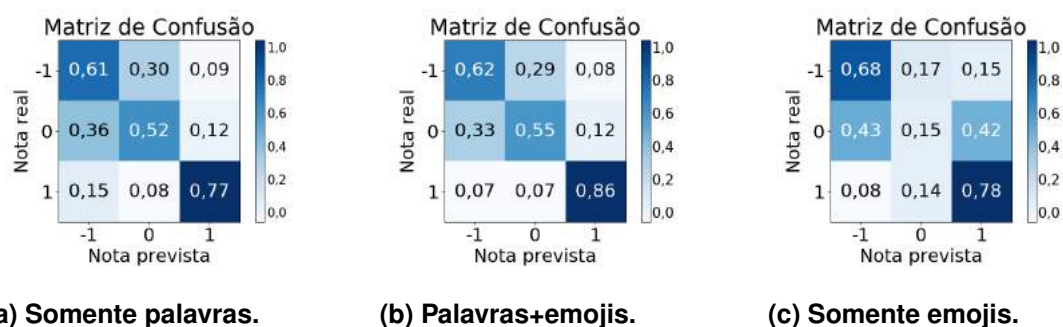


Figura 8. Cenário II: Matrizes de confusão.

isolado de emojis. Particularmente, no cenário de cinco classes, o uso isolado de emojis não foi capaz de identificar avaliações da classe neutra (nota 3). Outro fato interessante é que o uso de emojis/emoticons como parte do vocabulário do BoW ajuda a reduzir, principalmente, os erros mais extremos, tais como confundir nota -1 com 1, no cenário com três classes/notas. Portanto, emojis e emoticons são recursos linguísticos (linguagem online universal) com poder discriminante para previsão de opinião/polaridade/sentimento, principalmente, considerando opiniões extremas (muito negativas/positivas).

Como trabalhos futuros, pretendemos modelar o problema com uma estrutura hierárquica de classificação, resolvendo primeiro o problema de três classes (positivo, neutro, negativo) para então determinar sub-níveis das classes positivo e negativo. Pretendemos também aplicar técnicas de Teoria da Informação, representando as avaliações por distribuições discretas, para caracterizar as classes utilizando os conceitos de entropia e a divergência de Renyi.

## Referências

- Aisopos, F., Papadakis, G., Tserpes, K., e Varvarigou, T. (2012). Content vs. context for sentiment analysis: a comparative analysis over microblogs. In *Proceedings of the 23rd Conference on Hypertext and Social Media*, pages 187–196. ACM.
- Almeida, T. G., Souza, B. A., Menezes, A. A., Figueiredo, C., e Nakamura, E. F. (2016a). Sentiment analysis of portuguese comments from foursquare. In *Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web*, pages 355–358. ACM.
- Almeida, T. G., Souza, B. A., Menezes, A. A., Figueiredo, C. M., e Nakamura, E. F. (2016b). Sentiment analysis of portuguese comments from foursquare. In *Proceedings of the 22Nd Brazilian Symposium on Multimedia and the Web, Webmedia '16*, pages 355–358, New York, NY, USA. ACM.
- Araújo, M., Gonçalves, P., Cha, M., e Benevenuto, F. (2014). ifeel: a system that compares and combines sentiment analysis methods. In *Proceedings of the 23rd International Conference on World Wide Web*, pages 75–78. ACM.
- Jo, Y. e Oh, A. H. (2011). Aspect and sentiment unification model for online review analysis. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, WSDM '11*, pages 815–824, New York, NY, USA. ACM.
- Leskovec, J. e Krevl, A. (2014). SNAP Datasets: Stanford large network dataset collection. <http://snap.stanford.edu/data>.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167.

- Martins, R. F., Pereira, A., e Benevenuto, F. (2015). An approach to sentiment analysis of web applications in portuguese. In *Proceedings of the 21st ACM Brazilian Symposium on Multimedia and the Web*, pages 105–112. ACM.
- McAuley, J., Pandey, R., e Leskovec, J. (2015). Inferring networks of substitutable and complementary products. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794. ACM.
- Montejo-Ráez, A., Martínez-Cámara, E., Martín-Valdivia, M. T., e Ureña-López, L. A. (2014). Ranked wordnet graph for sentiment polarity classification in twitter. *Computer Speech & Language*, 28(1):93–107.
- Moraes, F., Vasconcelos, M., Prado, P., Almeida, J., e Gonçalves, M. (2013). Polarity analysis of micro reviews in foursquare. In *Proceedings of the 19th Brazilian symposium on Multimedia and the web (WebMedia)*, pages 113–120. ACM.
- Pang, B. e Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, 2(1-2):1–135.
- Pang, B., Lee, L., e Vaithyanathan, S. (2002). Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the 2002 ACL Conference on Empirical methods in natural language processing (ACL-02)*, pages 79–86. Association for Computational Linguistics (ACL).
- Reis, J., Miranda, M., Bastos, L., Prates, R., e Benevenuto, F. (2016). Uma análise do impacto do anonimato em comentários de notícias online.
- Serrano-Guerrero, J., Olivas, J. A., Romero, F. P., e Herrera-Viedma, E. (2015). Sentiment analysis: a review and comparative analysis of web services. *Information Sciences*, 311:18–38.
- Souza, B. A., Almeida, T. G., Menezes, A. A., Nakamura, F. G., Figueiredo, C. M., e Nakamura, E. F. (2016). For or against?: Polarity analysis in tweets about impeachment process of brazil president. In *Proceedings of the 22Nd Brazilian Symposium on Multimedia and the Web, Webmedia '16*, pages 335–338, New York, NY, USA. ACM.
- Terrana, D., Augello, A., e Pilato, G. (2014). Automatic unsupervised polarity detection on a twitter data stream. In *Semantic Computing (ICSC), 2014 IEEE International Conference on*, pages 128–134. IEEE.
- Turney, P. D. (2002). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pages 417–424. Association for Computational Linguistics.
- Wilson, T., Wiebe, J., e Hoffmann, P. (2009). Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. *Computational Linguistics*, 35(3):399–433.
- Ye, Q., Zhang, Z., e Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications*, 36(3):6527–6535.