

DETECÇÃO DE *FAKE NEWS* UMA ABORDAGEM DE APRENDIZADO DE MÁQUINA

Luciana da Silva Andrade¹

Alex M. G. Almeida²

¹Aluna : Luciana da Silva Andrade

²Professor : Alex marino Gonçalves de Almeida

Área de conhecimento (Tabela CNPq): 1.03.03.04-9 Sistemas de Informação.

Palavras-Chave: *Fake News*, Aprendizado de Máquina, Mineração de Textos.

RESUMO: *Fake News* têm um impacto imenso na nossa sociedade moderna. Detectar notícias falsas é um trabalho desafiador. Este trabalho propõe o uso de algoritmos de aprendizado de máquina para detectar *Fake News*. Cinco algoritmos populares foram usados nos experimentos: *Naïve Bayes*, *Multi Layer Perceptron*, *Random Forests*, *Decision Tree* e *Support Vector Machine*. O pré-processamento e as normalizações são etapas importantes para reduzir dimensionalidade dos dados antes da classificação. Os resultados mostram SVM e Random Forest classificadores mais apropriados, enquanto MinMax e Escala de Variância métodos de normalização produziram melhores resultados.

Fake News Detection a Machine Learning Approach

ABSTRACT: Fake news has an immense impact on our moderns society. Detecting fake news is an important step. This work proposes the use of machine learning techniques to detect fake news. Five popular algorithms are used in experiments: Naive Bayes, Multi-Layer Perceptron, Random Forests, Decision Tree and Support Vector Machine. Preprocessing and normalizations are essential steps for cleaning data before classification. The results show SVM and Random Forest as better discriminators while MinMax and Variance Scale as better normalization methods.

KEYWORDS: Fake News, Machine Learning, Text Mining.

INTRODUÇÃO

Nos dias de hoje, as informações *on-line* crescem de forma exponencial e sem precedentes, e no cotidiano das pessoas a procura por notícias e informações se dá cada vez mais pelas mídias digitais (NEWMAN et al., 2018). Pode-se observar este comportamento cotidiano com o crescimento das redes sociais *on-line* tais como Facebook, Twitter, Redit dentre outras, nas quais o usuário além da interação social acaba encontrando uma forma de veiculação de informação e notícias. O problema das notícias falsas ganhou imensa notoriedade nas eleições presidenciais americanas de 2016, já que muitos alegam que o resultado foi imensamente influenciado pela difusão maciça de notícias falsas (DEWEY, 2016; READ, 2016). As eleições presidenciais de 2018 e 2019 no Brasil (GRAGNANI, 2018) e Índia (TEAM, 2019) respectivamente, foram recheadas de campanhas de notícias falsas, com intenso uso de ferramentas de chats digitais e principalmente o Whatsapp.

A motivação desta pesquisa enquadra-se na elaboração de experimentos orientados a uma abordagem orientada às características de complexidade, sintáticas e psicológicas que serão extraídas dos *datasets* ISOT FAKENEWS¹ e LIAR FAKENEWS; e posterior a extração de características serão empregadas técnicas de normalização. Seguidos pela subsequente criação de modelos de aprendizado de máquina - *Multi Layer Perceptron* (MLP), *Random Forest* (RF), *Naïve Bayes* (NB), *Support Vector Machine* (SVM), *Decision Tree* (DT) - com intuito de aferir acurácia. Ao término deste trabalho pretende-se responder às seguintes perguntas: (i) *Existe prevalência de classificação quanto a um dataset?* ii) *As técnicas de normalização contribuem para classificação?* (iii) *Os métodos de normalização influenciam a classificação, se sim existe prevalência de algum método?;* (iv) *Existe prevalência de um algoritmo de aprendizado?.*

MATERIAIS E MÉTODOS

Tarefas

Nesta seção foram descritas as tarefas percorridas para execução deste trabalho. Na Figura 1 é possível observar cada uma das etapas.

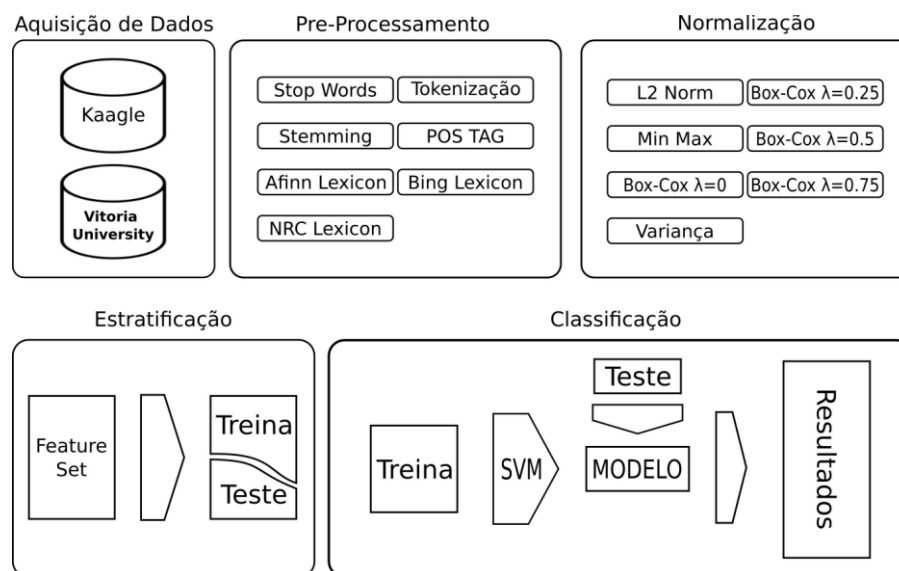


Figura 1: Diagrama de tarefas

Datasets

O *dataset* ISOT FAKENEWS contém dois tipos de artigos, falsos e reais. Este conjunto de dados foi coletado de fontes mundiais; as notícias verdadeiras foram obtidas rastreando artigos da agência de notícias Reuters. Quanto às notícias falsas, foram coletadas de diferentes fontes não confiáveis que foram sinalizadas pelas agências de checagem Politifact² e Wikipedia.

O LIAR FAKENEWS é um *dataset* disponível publicamente para detecção de notícias falsas. Ao longo de uma década foram rotuladas manualmente 12.800 notícias em vários contextos do Politifact, que fornece relatório de análise detalhado e links para documentos de origem para cada caso.

Aparato Ferramental

Dentre as ferramentas utilizadas, todas são de utilização da linguagem R³ e podem ser visualizadas na Tabela 1.

Tabela 1: Relação de ferramentas

Tipos	Ferramentas	Pacotes
Algoritmo	SVM	LiblineaR
Algoritmo	Naïve Bayes	naivebayes
Algoritmo	Random Forest	ordinalForest
Algoritmo	Decision Tree (J48)	Rweka
Algoritmo	MLP	RSNNS
Extração de características	POS TAGER	cleanNLP
Extração de características	AFINN Lexicon	textdata
Extração de características	NRC Lexicon	textdata
Extração de características	Bing Lexicon	textdata

Conjunto de Características

Para cada sentença dos documentos (*datasets*) foram extraídas considerando-se as naturezas de ordem psicológicas, estilísticas, e complexidade, conforme pode ser observado na Figura 2.

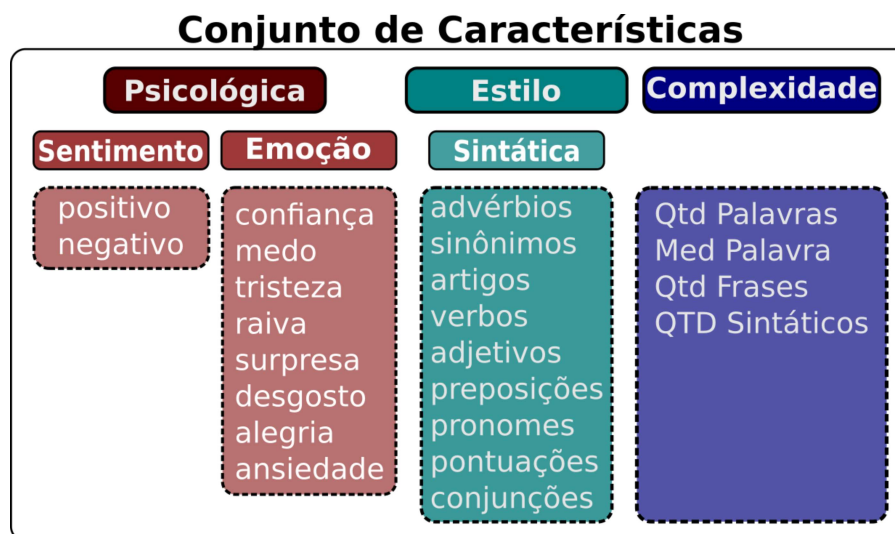
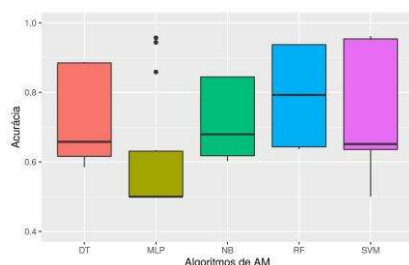


Figura 2: Conjunto de Características empregadas

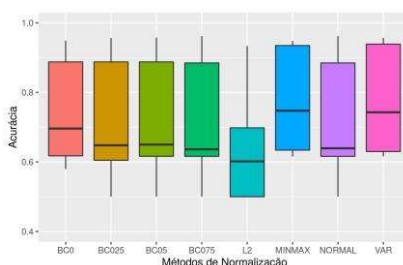
RESULTADOS E DISCUSSÃO

No intuito de arguir entorno do desempenho geral dos experimentos em relação aos *datasets* ISOT e LIAR é necessário voltar a atenção para Figura 3c. Podemos observar que as distribuições dos resultados em acurácia para os experimentos empregados no *dataset* ISOT supera consideravelmente os resultados obtidos empregados no LIAR. É notória a diferença dos resultados da acurácia mediana do ISOT=0.87134 e LIAR=0.61433. É possível ainda observar que todos resultados obtidos com o LIAR são equivalentes à região de *outlier* inferior do ISOT, ou seja, na região do primeiro quartil. Desta forma é possível afirmar que o conjunto de experimentos melhor se adequou ao ISOT.

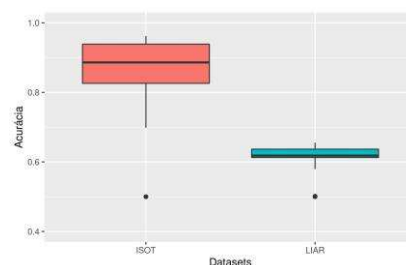
Para responder à pergunta (ii) “As técnicas de normalização contribuem para classificação?” é preciso observar a Figura 3b. A primeira consideração recai sobre o desempenho inferior observado no desempenho dos experimentos realizados com o método de normalização L2, notoriamente inferior aos demais. Uma segunda consideração nota-se a respeito da similaridade dos resultados obtidos com os experimentos realizados com o conjunto de características *in natura* - ausente de normalização – em



(a) Algoritmos ML



(b) Métodos Normalização



(c) Datasets

comparação aos experimentos com os demais métodos de normalização exceto o L2. Nota-se ainda que a mediana dos métodos BOX-COX $\lambda = 0.25$, $\lambda = 0.5$ e $\lambda = 0.75$ são muito similares ao resultado mediano da mediana dos experimentos *in natura*. Ainda é possível observar superioridade das medianas dos métodos Escala em Variança e MinMax em relação ao *in natura*. Por fim ainda é possível notar que os métodos e Escala em variação produziram menos resultados enquadrados como *outliers*. Diante destas observações podemos aferir que o método L2 contribuiu negativamente para classificação e que os

demais métodos pouco contribuíram na classificação, o que responde à pergunta (iii) “*Os métodos de normalização influenciam a classificação, se sim existe prevalência de algum método?*”.

Finalmente para responder à pergunta (iv) “*Existe prevalência de um algoritmo de aprendizado?*” volta-se a atenção para a Figura 3a. Pode-se notar que o melhor resultado é obtido com MLP, porém na região de *outlier*. Pode-se ainda notar que os melhores resultados obtidos com os classificadores SVM e RF se equipararam em acurácia ao melhor resultado obtido com *outlier* MLP.

CONCLUSÕES

No decorrer deste trabalho foi possível elaborar um conjunto de experimentos com objetivo de aplicar técnicas de classificação de *FAKE NEWS* empregando algoritmos de aprendizado de máquina. A extração de características de naturezas psicológicas, estilísticas e de complexidade demonstraram-se mais adequadas ao *dataset* ISOT. Quanto ao emprego do uso de métodos de normalização, estes não se mostraram determinantes na melhoria da classificação, mas ainda assim os métodos MinMax e Escala de Variança apresentaram um indicativo mínimo de melhoria. Por fim os classificadores SVM e RF produziram os melhores resultados em acurácia. Para eventual continuidade do trabalho a investigação de novas características e análise mais profunda da importância de cada uma das características na classificação é um bom caminho a percorrer.

REFERÊNCIAS

DEWEY, C. Facebook Fake-news Writer: ‘I Think Donald Trump Is in the White House Because of Me’. The Washington Post. 2016.

GRAGNANI, J. Um Brasil dividido e movido a notícias falsas: uma semana dentro de 272 grupos políticos no WhatsApp. 2018. <<https://www.bbc.com/portuguese/brasil-45666742>>. [Online; accessed 22-Março-2020].

NEWMAN, N. et al. AI, & nielsen, rk (2017). Reuters Institute digital news report 2017, 2018.

READ, M. Donald trump won because of facebook. New York Magazine, v. 9, 2016.

TEAM, R. C. India election 2019: The debunked fake news that keeps coming back. 2019. <<https://www.bbc.com/news/world-asia-india-47878178>>. [Online; accessed 22-Março-2020].

¹<https://www.uvic.ca/engineering/ece/isot/datasets/fake-news/index.php>

²<https://www.politifact.com/>

³www.r-project.org