

# Nanodegree Engenheiro de Machine Learning

## Proposta de projeto final

---

João Paulo Lopes de Souza

30 de maio de 2019

## Detecção automática de fake news na política brasileira

---

### Histórico do assunto

As pessoas assumem cada vez mais um papel ativo na disseminação da informação, consequência da popularização das redes sociais, fóruns e blogs na internet. Nessa quadra, são cada vez mais compartilhadas notícias e opiniões. Sendo que muitas delas são falsas e em determinados momentos podem ter alcance maior que as fontes reais<sup>1</sup>. Existem diferentes tipos de conteúdo enganoso, como explica Monteiro, 2018 [1] : “como notícias falsas, conhecidas popularmente como fake news, notícias satíricas, revisões falsas, boatos, etc.” As notícias falsas, que serão o foco neste trabalho, utilizam fatos forjados ou reais retirados do seu contexto original para desinformar o público. Isso normalmente acontece para beneficiar alguém.

Essa desinformação tem produzido enormes problemas em todo o mundo. Em especial para as democracias que tem suas eleições perturbadas para moldar os votos no sentido favorável à propagação de enganações.

Para tentar enfrentar esse problema da desinformação observamos grandes esforços em duas áreas em especial: 1. As agências de checagem de informação e 2. Sistemas automatizados para detecção de conteúdo falso. A primeira frente, as agências, realizam uma checagem manual para as notícias que surgem. Contudo, devido a sua natureza e a grande quantidade de notícias que surgem a cada dia, essa frente mostra-se necessária, porém ineficiente. A segunda frente, os sistemas automatizados, têm tentado realizar essa checagem de forma automática com o uso de machine learning. As iniciativas de checagem automática no Brasil são bastante tímidas e carecem de maior esforço de pesquisa.

---

<sup>1</sup> <<https://www.buzzfeed.com/br/alexandrearagao/noticias-falsas-lava-jato-facebook>>

## **Descrição do problema**

O presente trabalho pretende estudar como realizar detecção automática de notícias falsas, em especial, aquelas relacionadas à política que influenciam diretamente na democracia de diversos países e influenciam a vida de muitas pessoas. Nesse sentido, deve-se utilizar características verbais extraídos dos textos das notícias, realizar o processamento de linguagem natural - PLN e aplicar técnicas de machine learning para classificar as notícias em reais ou falsas.

## **Conjuntos de dados e entradas**

Utilizaremos o conjunto de dados reunidos por pesquisadores da USP [2] e chamado de *Cópus Fake.Br* [3]. Esse é o primeiro dataset detalhado sobre essa temática no Brasil.

O conjunto de dados *Cópus Fake.Br* possui 7200 notícias com seus respectivos metadados que foram coletadas no período de janeiro de 2016 a janeiro de 2018. Sendo 3600 notícias verdadeiras e 3600 notícias falsas com assuntos relacionados. Essas notícias possuem as seguintes categorias: política, TV e celebridades, sociedade e notícias cotidianas, ciência e tecnologia, economia e religião. O foco no presente trabalho serão as notícias da primeira categoria, política.

Existe um arquivo contendo o texto das notícias reais juntamente com seus títulos, outro arquivo com as notícias falsas e dois arquivos com os respectivos metadados. A seguir apresentamos as características presentes nos metadados: autor, link, categoria, data de publicação, número de tokens, número de palavras sem pontuação, número de tipos, número de links dentro das notícias, número de palavras em caixa alta, número de verbos, número de verbos no subjuntivo ou imperativo, número de substantivos, número de adjetivos, número de advérbios, número de verbos auxiliares, número de pronomes pessoais no singular nas primeiras e segunda pessoas, número de pronomes pessoais no plural da primeira pessoa, número de pronomes, pausas, número de caracteres, tamanho médio das sentenças, tamanho médio das palavras, percentagem de erros na escrita, emotividade e diversidade.

## **Descrição da solução**

Serão extraídas e analisadas das notícias características verbais e algumas métricas, em seguida será realizado processamento de linguagem natural - PLN e por fim serão aplicadas e avaliadas diferentes técnicas de machine learning para classificar as notícias em reais ou falsas.

## Modelo de referência (benchmark)

Será utilizado como modelo de referência o resultado obtido na pesquisa da USP [2]. O melhor e mais eficiente resultado foi obtido utilizando a técnica de bag of words e aplicando o algoritmo SVM.

## Métricas de avaliação

Serão utilizadas as mesmas métricas utilizadas no artigo [2] para avaliar o desempenho do modelo comparando-o com o benchmark: precision, recall, f-measure e accuracy.

## Design do projeto

- Entender melhor o dataset, realizar limpeza e pré-processamento nos dados. Tentar separar os títulos que estão incluídos no texto das notícias e tentar gerar novas métricas sobre eles. Filtrar os registros para utilizar apenas as notícias da categoria política;
- Separação do conjunto de dados em treino e testes;
- Realizar o treinamento e avaliação dos resultados. Pretende-se usar e avaliar outros algoritmos não abordados no trabalho [2] como o classificador gradient boost do scikit-learn e o XGBoost;

## Referências

1. Monteiro, R.A. (2018). Detecção Automática de Notícias Falsas. Trabalho de Conclusão de Curso. Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo. São Carlos-SP, November, 40p.<<https://drive.google.com/file/d/1dgazD5ZfFwY8aVPJUur4xE1kHwzSU8ET/view>>. Acesso em 30/05/2019.
2. Monteiro, R.A.; Santos, R.L.S.; Pardo, T.A.S.; Almeida, T.A.; Ruiz, E.E.S.; Vale, O.A. (2018). Contributions to the Study of Fake News in Portuguese: New Corpus and Automatic Detection Results. In the Proceedings of the 13th International Conference on the Computational Processing of Portuguese (PROPOR) (LNAI 11122), pp. 324-334. September, 24-26. Canela-RS/Brazil. <<http://conteudo.icmc.usp.br/pessoas/taspardo/PROPOR2018-MonteiroEtAl.pdf>> Acesso em 30/05/2019.
3. Fake.br-Corpus. <<https://github.com/roneysco/Fake.br-Corpus>>. Acesso em 30/05/2019.