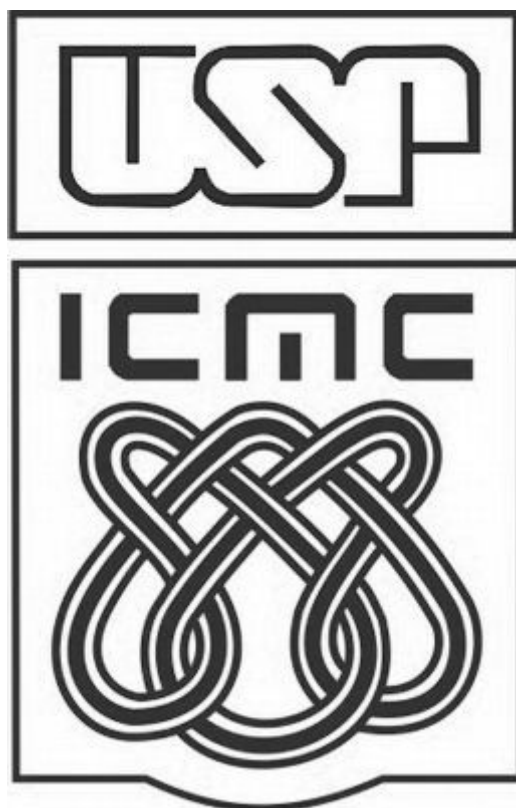


Universidade de São Paulo

Instituto de Ciências Matemáticas e de Computação

RELATÓRIO DO SEGUNDO TRABALHO DE INTELIGÊNCIA ARTIFICIAL PARTE 2

Vitor Albuquerque de Paula 8628220



1. Pré-processamento

1.1. Filtragem dos comentários

Primeiramente será feita uma filtragem dos comentários para que caracteres sem significado como vírgulas, por exemplo, sejam ignorados pela análise, melhorando sua eficiência. Além disso, considera-se que os números não possuem caráter discriminatório dentro do contexto analisado, ou seja, são irrelevantes na detecção de ofensas. Sendo assim, os números terão todos seus algarismos trocados por 9 a fim de que eles sejam considerados como palavras comuns (132 e 567 seriam ambos substituídos por 999, por exemplo) e, através do TF-IDF, se tornem irrelevantes.

1.2. Term frequency – Inverse Document Frequency

Após concluída a etapa de filtragem inicial será realizado um cálculo da frequência e contagem de termos em todos os comentários. O TF-IDF define a importância do termo dentro da coleção de comentários dando um peso aos termos num comentário. O peso é definido como o número de ocorrências do termo no mesmo comentário (D_i), modificada por uma escala de importância do termo (T_j), chamada de frequência inversa do documento que é maior quando um termo aparece num menor número de comentários diferentes. Por fim os termos que aparecerem em menor quantidade de comentários serão classificados como mais importantes e portanto não deverão ser ignorados, enquanto que os termos que aparecem com maior frequência em comentários diversos, como artigos, poderão ser ignorados já que não possuem relevância para classificar os comentários como ofensivos.

Vale notar que a escala utilizada deve ser normalizada a fim de evitar problemas causados pela repetição de uma mesma palavra num único comentário, o que aumentaria linearmente o seu peso no cálculo do TF-IDF.

1.3. Motivações

A filtragem será realizada para que os dados sem significado não sejam levados em consideração pelo algoritmo de classificação. Assim, apesar de simples, é uma parte importante para evitar resultados errôneos uma vez que o algoritmo TF-IDF se baseia em frequência de uma palavra.

Então a classificação do texto nos comentários com o TF-IDF é feita para organizar os dados numa estrutura com a qual é possível fazer a classificação entre as duas classes do problema (comentário ofensivo ou não ofensivo) utilizando o algoritmo de classificação escolhido, o Naive Bayes.

2. Técnica utilizada para análise de dados

2.1. Algoritmo de classificação

O algoritmo escolhido para a classificação dos dados é o Naive Bayes Gaussiano. O funcionamento do algoritmo se dá conforme os seguintes passos:

1. Converte-se o conjunto de dados numa tabela de frequência.
2. Cria-se uma tabela de probabilidade dos atributos encontrados em relação às classes no problema.
3. Por fim, é usada a equação Naive Bayesiana para calcular a probabilidade para cada classe. A classe com maior probabilidade acaba como resultado da previsão.

É notável que para a classificação de novos elementos o algoritmo considera tanto a probabilidade deles em todo conjunto quanto àquela dada o de cada classe cujos valores são utilizados na equação Naive Bayesiana.

2.2. Motivação para utilização desse algoritmo

A escolha do algoritmo foi feita entre vários frequentemente utilizados na área de classificação de texto, a partir dos quais procuramos um algoritmo simples embora eficiente. O Naive Bayes foi escolhido por apresentar essas características.

Apesar do seu funcionamento aparentemente simples, o algoritmo já funcionou muito bem em situações reais complexas e possui bases teóricas para tal. Além de se mostrar como um algoritmo simples que muitas vezes apresenta resultados tão bons quanto ou até melhores que algoritmos de classificação mais sofisticados, os quais são mais complexos e caros, como aqueles como regressão logística.

Entretanto, mesmo com suas qualidades, o algoritmo perde em performance quando comparado a outros algoritmos de classificação, como Boosted Trees e Random Forests.

3. Referências

- LUCCA, G; PEREIRA, A. I.; PRISCO, A.; Uma implementação do algoritmo Naïve Bayes para classificação de texto
<http://www.lbd.dcc.ufmg.br/colecoes/erbd/2013/0019.pdf>
- MOREIRA, D. Introdução a Naïve-Bayes. Disponível em:
<http://java.icmc.usp.br/moodle/mod/resource/view.php?id=2833>. Acesso em: 26 mai. 2016
- LOPES, M, C, S. Mineração de Dados Textuais Utilizando Técnicas de Clustering para o Idioma Português. Disponível em:
<http://www.inf.unioeste.br/~jorge/ARTIGOS%20INTERESSANTES/LING%20CDSTICA%20e%20MECANISMOS%20DA%20LINGUAGEM/Tese_coc_ufrj.pdf>. Acesso em: 21 mai. 2016
- LORIA, S. Tutorial: Finding Important Words in Text Using TF-IDF. Disponível em:
<<http://stevenloria.com/finding-important-words-in-a-document-using-tf-idf/>>. Acesso em: 21 mai. 2016