

# ***SafetyRank: comparação de técnicas de aprendizado de máquina para classificação de alertas de segurança industriais***

**Wander Fernandes Júnior<sup>1</sup>, Karin Satie Komati<sup>1</sup>, Kelly Assis de Souza Gazolli<sup>1</sup>**

<sup>1</sup>Programa de Pós-Graduação em Computação Aplicada (PPComp)  
Campus Serra do Instituto Federal do Espírito Santo (Ifes)  
Serra - ES - Brasil

wanderfj@gmail.com; {kkomati, kasouza}@ifes.edu.br

**Abstract.** *In industrial areas the issuance of safety alerts after the occurrence of accidents is common. In this context, this work proposes a comparison of machine learning techniques for text classification, using a database prepared by the authors with safety alerts extracted from public documents obtained from the internet. Classical classifiers, KNN, SVM, Naive Bayes, Decision Tree, and Random Forest, were applied to the dataset. The best accuracy obtained by SVM with 79% followed by Random Forest with 75%. The results encourage further work, as a larger public database of accidents and safety alerts can increase the dissemination of knowledge and make it possible to reduce accidents at work.*

**Resumo.** *Nas áreas industriais emitem-se alertas de segurança em formato digital após a ocorrência de acidentes. Neste contexto, este trabalho propõe a criação de uma base de dados de alerta de segurança, bem como uma comparação de técnicas de aprendizado de máquina para classificação de textos. A base de dados foi elaborada pelos autores, através de coleta de documentos públicos obtidos da internet. Classificadores clássicos, KNN, SVM, Naive Bayes, Árvores de decisão e Floresta Aleatórias foram aplicados à base de dados, sendo a melhor acurácia obtida pelo SVM com 79% seguido da Floresta Aleatória com 75%. Os resultados estimulam a continuação do trabalho, pois uma maior base de dados pública de acidentes e alertas de segurança pode aumentar a divulgação do conhecimento e possibilitar uma diminuição dos acidentes de trabalho.*

## **1. Introdução**

De acordo com dados do Observatório Digital de Saúde e Segurança do Trabalho, de 2012 a 2018 o Brasil registrou 4,5 milhões acidentes e 16.455 mortes relacionadas ao trabalho. Neste mesmo período, os gastos de previdência associados aos acidentados atingiram R\$79 bilhões [SmartLab 2019].

Acidentes podem ocorrer com menor frequência se as lições aprendidas forem divulgadas, discutidas e registradas, para que as ações tomadas no passado estejam acessíveis às pessoas em situações similares [Kletz 1993]. Nas áreas industriais, é comum que sejam emitidos alertas de segurança em formato digital após a ocorrência de acidentes. Estes documentos de textos contém relatos da ocorrência e lições aprendidas que visam prevenir danos às pessoas e fortalecer a cultura de segurança por meio do compartilhamento de experiências. A Figura 1 ilustra alertas emitidos pela ANP

(Agência Nacional do Petróleo, Gás Natural e Biocombustíveis) [ANP 2018], CCPS (*Center for Chemical Process Safety*) [CCPS 2019], USCG (*United States Coast Guard*) [USCG 2019] e IADC (*International Association of Drilling Contractors*) [IADC 2019]. Outras instituições também emitem alertas de segurança, tais como BSEE (*Bureau of Safety and Environmental Enforcement*) [BSEE 2019] e NOPSEMA (*National Offshore Petroleum Safety and Environmental Management Authority*) [NOPSEMA 2019]. Todos os textos dos alertas estão na língua inglesa.



Figura 1. Exemplos de textos de alertas de segurança.

A Petrobras estabeleceu 10 regras de ouro da segurança as quais também relacionam-se às causas mais comuns de acidentes [Petrobras 2018]. A Figura 2 apresenta estas 10 regras de forma gráfica, que são descritas a seguir:

1. Permissão para Trabalho (em inglês *Permit to Work*): somente trabalhe com permissão de trabalho válida, liberada no campo e de seu total entendimento.
2. Isolamento de Energias (em inglês *Energy Isolation*): somente execute trabalhos em equipamentos ou instalações após certificar-se de que todas as fontes de energia tenham sido isoladas de forma segura.
3. Trabalho em Altura (em inglês *Work at Height*): somente execute trabalhos em altura com a utilização de cinto de segurança fixado em local seguro e previamente

- determinado.
4. Espaço Confinado (em inglês *Confined Space*): só entre em espaço confinado se autorizado, equipado e com treino específico.
  5. Atmosferas explosivas (em inglês *Loss of Primary Containment (LOPC)*): nunca entre em local com atmosfera explosiva. Obedeça sempre aos alarmes e à sinalização.
  6. Posicionamento Seguro (em inglês *Dropped Objects*): não acesse área isolada e nunca se posicione sob uma carga suspensa ou entre veículos, parados ou em movimento.
  7. Equipamentos de Proteção Individual (em inglês *Personnel Protective Equipment*): use sempre os EPI conforme recomendado.
  8. Atenção às Mudanças (em inglês *Management of Change*): fique atento aos riscos das mudanças e somente realize qualquer mudança que envolva pessoas, instalações, materiais ou procedimentos após análise e autorização.
  9. Segurança no Trânsito (em inglês *Road Safety*): respeite as leis de trânsito e pratique direção defensiva. Use o cinto de segurança, respeite os limites de velocidade, não use celular e se beber não dirija.
  10. Álcool e outras drogas (em inglês *Alcohol and Drugs*): nunca trabalhe sob efeito de álcool ou outras drogas.



**Figura 2. Dez regras de ouro da segurança, retirado de [Petrobras 2018].**

Outras regras de segurança também são usadas na indústria [IOGP 2019]. O “Içamento mecânico” (em inglês *Mechanical Lifting*) refere-se a minimizar os cuidados na utilização de equipamentos de içamento, tais como polias talhas. A “Integridade, corrosão e erosão” (em inglês *Integrity*) refere-se ao entendimento e monitoramento das condições de integridade da planta. E a regra “Controles de segurança de processo” (em inglês *Process Safety Controls*) trata das análises e autorizações necessárias para realizar procedimentos na planta de processo.

A classificação de alertas industriais baseada em suas respectivas regras de ouro pode facilitar a recuperação de informações, pois permite que documentos relevantes sejam mais facilmente identificados e aproveitados como referência para atividades similares no futuro.

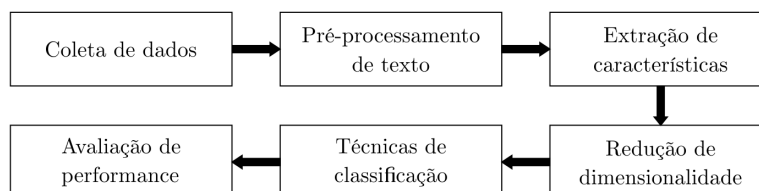
Uma base de dados de alertas públicos de segurança industrial foi elaborada pelos autores, os quais foram classificados em assuntos tendo como base as regras citadas acima. Esta base de dados de documentos foi categorizada em 8 classes: “Isolamento de

Energias”, “Trabalho em Altura”, “Espaço Confinado”, “Atenção às Mudanças”, “Posicionamento Seguro”, “Levantamento mecânico”, “Permissão para Trabalho” e “Integridade Mecânica”. O objetivo deste trabalho é a comparação de diferentes classificadores: Naive Bayes, SVM, KNN, Árvore de decisão e Floresta Aleatória para a categorização dos textos da base elaborada.

No que se refere à estrutura do artigo, na seção 2 são apresentados os principais conceitos e técnicas utilizados no trabalho. A metodologia empregada é descrita na seção 3, formando a base para a apresentação e análise dos resultados na seção 4. A seção 5 encerra o artigo com os comentários finais e conclusões.

## 2. Referencial Teórico

O processo de classificação de textos, ilustrado na Figura 3, pode ser dividido em seis etapas: coleta de dados, pré-processamento de texto, extração de características, redução de dimensionalidade, aplicação da técnica de classificação e avaliação de performance [Kadhim 2019].



**Figura 3. Etapas do processo de classificação de textos.**

A coleta de dados é a fase inicial, na qual é realizado o levantamento dos dados de texto a serem utilizados. Os conjuntos de dados em geral contém uma sequência de textos em documentos  $D = \{X_1, X_2, \dots, X_n\}$  onde  $X_i$  refere-se a um documento da base de dados. Cada documento é rotulado com uma classe de um determinado conjunto de  $k$  valores discretos. A etapa de pré-processamento de texto trata da preparação inicial do texto. É necessário um módulo de análise de conteúdo baseado em técnicas de processamento de linguagem natural, que permita transformar o texto bruto em representações mais significativas e que possam ser utilizadas de forma mais eficaz [Zhai and Massung 2016]. De acordo com [Allahyari et al. 2017], nesta etapa incluem-se a separação de palavras (tokenização), remoções de *stop words* e lematização.

Em seguida, na etapa de extração de características, os dados pré-processados são trabalhados para obter as características mais importantes para o processo de classificação. Conforme [Aggarwal and Zhai 2012], existem dois tipos principais de representação. O primeiro é denominado *bag of words*, na qual o documento é representado como um conjunto de palavras e suas respectivas frequências, de forma que a ordem das palavras no texto não é considerada. O segundo método é a representação na qual é possível capturar o contexto das palavras, denominado *word embedding*. De acordo com [Kowsari et al. 2019], as técnicas mais comuns são TF-IDF (*Term Frequency-Inverse Document Frequency*), Word2Vec (*Word to Vector*) e Glove (*Global Vectors for Word Representation*). Em geral, a etapa de extração de características geram vetores de alta dimensão, que impactam no tempo de processamento. Para a redução de dimensionalidade, as técnicas mais comuns são o PCA (*Principal Component Analysis*), o LDA (*Linear Discriminant Analysis*) e o NMF (*Non-negative matrix factorization*) [Kowsari et al. 2019].

Existem diversas técnicas de classificação, tais como Naïve Bayes, SVM (*Support Vector Machine*), KNN (*N-Nearest Neighbours*), árvores de decisão e diversos formatos de redes neurais [Lorena et al. 2000]. Na última etapa do processo de classificação de textos, é necessário avaliar a performance obtida. Esta avaliação é feita por meio de métricas obtidas a partir de uma matriz de confusão, tais como acurácia, revocação, precisão e/ou medida-F1. A acurácia indica a porcentagem de textos classificados adequadamente em toda a base de dados, ou seja, o quão frequente o classificador está correto.

### 3. Materiais e Métodos

A base de dados deste trabalho foi elaborada pelos autores a partir de alertas de segurança industrial extraídos de 80 (oitenta) documentos públicos em formato PDF obtidos na internet. Todos os documentos estão no idioma inglês e foram retirados dos sites da ANP, BSEE, CCPS, IADC, NOPSEMA e USCG. A base de dados está balanceada, com 10 (dez) documentos em cada uma das 8 (oito) classes.

Na etapa de pré-processamento de dados, foi usada a função **CountVectorizer** do scikit-learn. A função de vetorização consiste em transformar os textos em vetores de contagem de palavras (*tokens*), tendo sido desconsideradas todas palavras comuns que estejam presentes em mais de 95% dos textos e removidas (*stop words*) do idioma inglês.

Para extração de características foi utilizada a técnica TF-IDF, via função **TfidfTransformer** do scikit-learn. TF-IDF é uma medida estatística que tem o intuito de indicar a importância de uma palavra de um documento em relação a uma coleção de documentos. O valor de uma palavra, nessa representação, aumenta de acordo com o número de ocorrências do termo no documento e a raridade do termo na coleção [Jimenez et al. 2016].

Na aplicação dos classificadores os hiperparâmetros foram mantidos nos valores padrão da biblioteca Scikit-Learn [Pedregosa et al. ]. Não se usou técnica redução de dimensionalidade, dado que não houve problemas com o desempenho do sistema. Foi disponibilizado um repositório GitHub<sup>1</sup> com as implementações de código realizadas.

Baseado em distância, KNN efetua a classificação a partir de uma votação por maioria simples dos vizinhos mais próximos de cada ponto. Classificar um ponto, utilizando o KNN, pode ser resumido em três passos, segundo Neto et al. [Neto et al. 2017], que são (i) o cálculo da distância entre o exemplo que não é conhecido, com os demais exemplos do conjunto de treinamento, (ii) a identificação dos K vizinhos mais próximos e (iii) a utilização do rótulo da classe de vizinhos mais próximos para determinar o rótulo do exemplo desconhecido, usando um sistema de votação. Nos experimentos, foram usados os seguintes parâmetros: 5 vizinhos (K=5); todos os 5 vizinhos tem o mesmo poder de voto, isto é, não contém pesos diferentes; a distância usada foi a de Minkowski com potência 2.

Baseado em otimização, SVM constrói hiperplanos para separar diferentes classes no espaço. Esse algoritmo tem como vantagem ser eficaz em dados com muitas dimensões, mesmo que haja um número de instâncias menor do que o total de dimensões [Ceravolo et al. 2019]. Nesta função da biblioteca usada, o parâmetro 'C' controla o *trade-off* entre um limite de decisão suave e a classificação correta dos dados de trei-

---

<sup>1</sup><https://github.com/wanderfernandesjunior/safetyrank-text-classification>

namento, aumentar o valor de  $C$  pode levar ao *overfitting*. O valor usado foi de  $C=1$ . O SVM tradicional é um classificador binário, ou seja, classifica apenas duas classes. Para implementar um classificador multi-classes, podem ser empregadas a abordagem Um-Contra-Um (ovo, do inglês *One-vs-One*) e Um-Contra-Todos (ovr, do inglês *One Verse Rest*) [Pal 2008]. A opção usada foi a 'ovr', nesta abordagem são construídos classificadores binários que distinguem entre uma determinada classe entre as demais. Neste caso, se existem  $N$  classes, são construídos  $N$  classificadores binários, onde cada classificador é treinado para distinguir uma entre as outras  $N - 1$  classes. Para novas instâncias não rotuladas, é aplicado a abordagem do “o-vencedor-ganha”, onde o classificador com maior margem, que separa os hiperplanos é o vencedor.

Baseado em teoria probabilística, o Naive-Bayes, trabalha com a ideia de independência de atributos. Sendo considerado como ingênuo (*naive* em inglês), desconsidera a associação entre os atributos e os analisa como condicionalmente independentes. Naive-Bayes oferece bons resultados quando se tem disponível um conjunto de treinamento médio ou grande. Já o Processo Gaussiano é um método que utiliza distribuições de probabilidade para estimar a classe de um ponto através de inferência Bayesiana [Rasmussen 2003].

Baseada em busca, a árvore de decisão classifica os dados com base em regras inferidas a partir de seus atributos, tendo como vantagem facilitar o entendimento do modelo, bem como dos atributos mais relevantes. De modo geral, no método de árvore de decisão, o problema é representado como uma árvore, onde cada nó de entrada é analisado e dividido [Rodrigues et al. 2018].

A técnica é chamada de *ensemble* quando um conjunto de classificadores é treinado individualmente mas as decisões são tomadas de forma combinada. Métodos *ensemble* tendem a apresentar um menor *overfitting*. A Floresta Aleatória constrói diversas árvores em paralelo, que foram construídas com partes diferentes da base de treinamento, que posteriormente se combinam para obter uma única resposta como saída [Rodrigues et al. 2018].

A base de dados foi separada em 60% para treinamento e 40% para teste, a seleção foi aleatória. Foi utilizada a técnica de validação cruzada *k-fold* com  $k = 4$ , na qual é medido o erro de validação a partir do particionamento dos dados em conjuntos de teste e treino, visando diminuir a probabilidade dos dados serem divididos e testados de forma de forma inadequada.

#### 4. Resultados

A avaliação da performance foi feita usando a matriz de confusão, precisão, revocação, medida-F1 e acurácia. Todas as métricas se encontram no repositório. Para o artigo, resume-se a discussão sobre os dados de acurácia. Na Tabela 1 constam os resultados de acurácia obtidos na aplicação dos classificadores à base de dados de alertas de segurança.

O classificador que obteve a melhor acurácia nos testes foi o SVM (0,79) seguido da Floresta Aleatória (0,75). Interessante observar que o SVM é uma técnica linear e o KNN é não-linear, logo, é possível se conjecturar que os resultados dos vetores de características possuem um fronteira linear entre classes. Outra questão é a forma como o SVM resolve a questão de múltiplas classes, em que se constrói um hiperplano de separação de uma classe contra todas as outras.

**Tabela 1. Acurácia dos classificadores na base de dados de alertas de segurança**

Classificador	Acurácia	
	Treino	Teste
KNN	0,57	0,61
SVM	0,62	0,79
Naive Bayes	0,30	0,25
Árvore de Decisão	0,55	0,64
Floresta Aleatória	0,50	0,75

O fato do resultado de Naive Bayes ter apresentado acurácia baixa indica que as palavras, isto é, cada contagem de palavras, cada posição dos vetores de características não são independentes. Assim, as palavras aparecerem de forma repetida nos textos indica melhor qual é a classe. O resultado da técnica Floresta Aleatória ser melhor que a árvore de decisão só corrobora com a teoria.

Entretanto os resultados obtidos na base de alertas de segurança foram inferiores em relação aos resultados obtidos em outras bases de dados de textos, tais como na base 20-Newsgroups que chega a 0,86 de acurácia [Sun et al. 2009].

## 5. Considerações Finais

Este trabalho realizou uma comparação das técnicas de classificação de texto aplicadas à uma base contendo textos sobre acidentes elaborada pelos autores. Existem desafios associados à extração de informações e classificação de textos, pois as linguagens naturais normalmente são ambíguas e não obedecem a regras rígidas [Goodfellow et al. 2016].

Os resultados obtidos estimulam a continuação do trabalho, porém, é necessário aumentar o tamanho da base de dados e usar outras técnicas de classificação, bem como efetuar a calibração dos mesmos.

Acredita-se que é necessário que a base de dados seja pública, para maior divulgação do conhecimento sobre acidentes e alertas de segurança por parte das empresas e trabalhadores, o que possibilitaria a diminuição de acidentes de trabalho.

## Referências

- Aggarwal, C. C. and Zhai, C. (2012). *A Survey of Text Classification Algorithms*, pages 163–222. Springer US, Boston, MA.
- Allahyari, M., Pouriye, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., and Kochut, K. (2017). A brief survey of text mining: Classification, clustering and extraction techniques. *arXiv preprint arXiv:1707.02919*.
- ANP (2018). Alertas de segurança. Acessado em 18 dez. 2019.
- BSEE (2019). Bureau of safety and environmental enforcement. Acessado em 18 dez. 2019.
- CCPS (2019). Process safety beacon. Acessado em 18 dez. 2019.
- Ceravolo, I., Brasil, A. A., and Komati, K. (2019). Classifying readers with dyslexia from eye movements using machine learning and wavelets. In *ENIAC 2019*.

- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.
- IADC (2019). International association of drilling contractors - safety alerts. Acessado em 18 dez. 2019.
- IOGP (2019). International association of oil and gas producers. Acessado em 18 dez. 2019.
- Jimenez, S., Gonzalez, F. A., and Gelbukh, A. (2016). Mathematical properties of soft cardinality: Enhancing jaccard, dice and cosine similarity measures with element-wise distance. *Information Sciences*, 367:373–389.
- Kadhim, A. I. (2019). Survey on supervised machine learning techniques for automatic text classification. *Artificial Intelligence Review*, 52(1):273–292.
- Kletz, T. A. (1993). *Lessons from disaster: how organizations have no memory and accidents recur*. IChemE.
- Kowsari, K., Jafari Meimandi, K., Heidarysafa, M., Mendu, S., Barnes, L., and Brown, D. (2019). Text classification algorithms: A survey. *Information*, 10(4):150.
- Lorena, A. C., Gama, J., and Faceli, K. (2000). *Inteligência Artificial*. LTC, Rio de Janeiro.
- Neto, W. B. d. R., Jr., J. M. P. d. M., and Souza, R. V. L. (2017). Análise de dados obtidos através de um sistema de telemetria automotivo utilizando k-nn. *XIV Encontro Nacional de Inteligência Artificial e Computacional*, pages 960–971.
- NOPSEMA (2019). National offshore petroleum safety and environmental management authority. Acessado em 18 dez. 2019.
- Pal, M. (2008). Multiclass approaches for support vector machine based land cover classification. *arXiv preprint arXiv:0802.2411*.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12.
- Petrobras (2018). Sustentabilidade 2018. Acessado em 18 dez. 2019.
- Rasmussen, C. E. (2003). Gaussian processes in machine learning. In *Summer School on Machine Learning*, pages 63–71. Springer.
- Rodrigues, D. S. et al. (2018). A comparative analysis of loan requests classification algorithms in a peer-to-peer lending platform. In *Proceedings of the XIV Brazilian Symposium on Information Systems*, page 42. ACM.
- SmartLab (2019). Observatório de segurança e saúde no trabalho. Acessado em 18 dez. 2019.
- Sun, A., Lim, E.-P., and Liu, Y. (2009). On strategies for imbalanced text classification using svm: A comparative study. *Decision Support Systems*, 48(1):191–201.
- USCG (2019). United states coast guard - safety alerts. Acessado em 18 dez. 2019.
- Zhai, C. and Massung, S. (2016). *Text data management and analysis: a practical introduction to information retrieval and text mining*. Morgan & Claypool.