

ANÁLISE DE REGISTROS DE ACIDENTES DE TRABALHO DA ÁREA DA SAÚDE UTILIZANDO TÉCNICAS DE MINERAÇÃO DE DADOS

Letícia Daudt Silva (PIBIC/CNPq/FA/UEM), Gislaine Camila Lapasini Leal (Orientadora), Edwin Vladmir Cardoza Galdamez (Co-Orientador)
e-mail: ra103798@uem.br

Universidade Estadual de Maringá / Centro de Tecnologia/Maringá, PR.

Área: Engenharia de Produção
Subárea: Higiene e Segurança do Trabalho

Palavras-chave: acidentes de trabalho, mineração de dados, área da saúde.

Resumo:

Os acidentes de trabalho ocorrem em diversos setores em todo o mundo. No que se refere ao Brasil, é possível acessar tais registros de acidentes por meio do site de Comunicação de Acidentes de Trabalho, conhecido como CAT, o qual disponibiliza diversas bases de dados referentes à estas ocorrências em determinados períodos dos anos passados. Com isso, o objetivo principal é de analisar os registros da área da saúde, o método utilizado é a mineração de dados em um conjunto de dados do período de janeiro de 2018 até março de 2021, que após serem tratados e balanceados, resultaram em uma base de dados de 1712 registros. A partir disso, foram aplicadas doze técnicas de mineração de dados, seguido de uma comparação com cinco métricas, selecionando assim a técnica *Voting* como a mais adequada pois apresentou os melhores resultados.

Introdução

De acordo com o Ministério da Previdência Social em 2007, os acidentes de trabalho são aqueles decorrentes do exercício a serviço da empresa, sendo divididos em três categorias: típicos da atividade executada; de trajeto durante o percurso até o trabalho; ou doenças do trabalho ligadas ao ramo da atividade. Estes podem inclusive causar lesões corporais, distúrbios de forma permanente ou temporária, reduzindo a capacidade para o trabalho, gerando impacto emocional e econômico (BAKKE e DE ARAUJO, 2010).

No final de 2015, o Brasil se encontrou na 3ª posição dentro os países com mais registros de mortes por acidentes de trabalho, chegando a atingir 3 mil óbitos no ano, sem contar os que não foram comunicados.

Com base nos registros do Observatório Digital de Saúde, de 2017, a cada dez acidentes de trabalhos, um ocorre na área de saúde, principalmente no atendimento

hospitalar. Um estudo apontou que a taxa de acidentes de trabalho nesta área é 34% maior que em outros setores. O índice de ocorrências dentro da área da saúde que não são comunicadas no Brasil chega a 90% (ASCOM HUAC, 2019).

Inclusive, a NR-4 (Serviços Especializados em Engenharia de Segurança e em Medicina do Trabalho) classifica organizações que oferecem estes serviços de saúde como grau de risco 3, uma vez que expõe os pacientes e os profissionais a riscos variados (BAKKE e DE ARAÚJO, 2010).

Diante disso, o objetivo do trabalho é apresentar os resultados e conclusões obtidas através de técnicas de mineração de dados em um banco de dados de janeiro/18 a março/21 com foco na área da saúde, fornecido pelo CAT.

Materiais e Métodos

Para a estruturação de todo o raciocínio, primeiramente se pesquisou conceitos sobre segurança no trabalho, acidentes na área da saúde e a própria técnica de mineração de dados dentro da literatura. Em seguida, coletou-se dados históricos sobre estes acidentes em bancos de dados disponibilizados pelo CAT, do período de janeiro de 2018 até março de 2021, reunindo mais de um milhão de casos que foram tratados através do software Microsoft Excel para manter apenas ocorrências da área da saúde, resultando em 1712 registros, proporcionando análises primárias por meio de gráficos.

Com o objetivo de prever a ocorrência de óbitos e tendo a base de dados de acidentes específicos da área da saúde tratada e estruturada, realizou-se a preparação para a aplicação das técnicas de mineração de dados. Para isso, as variáveis que eram identificadas por texto foram transformadas em números binários; dentre as nove variáveis analisadas (gênero, localização, tipo de ocorrência, parte do corpo, natureza de lesão, agente causador, óbitos, CBO e idade), a de óbito fora considerada dependente e o restante independentes, assim realizou-se o balanceamento dos registros usando como referência o parâmetro de óbito e não óbito, resultando em 74 ocorrências.

Uma vez com estes pontos definidos, fora utilizado o Jupyter Notebook, um ambiente com modelos de programação na linguagem Python onde pode-se aplicar 12 técnicas de mineração para prever a ocorrência de óbitos. Os algoritmos utilizados foram: *Bagging*, *Extra Trees*, *Random Forest*, *Stacking*, *Voting*, *XGBoost*, *Decision Trees*, *K-Nearest Neighbors*, *Logistic Regression*, *Naive Bayes*, *Neural Networks* e *Support Vector Machine*.

Em cada um dos algoritmos, analisou-se cinco métricas, sendo elas: acurácia, precisão, *recall*, *F1 score* e curva ROC/AUC (*Receiver Operating Characteristic/Area Under Curve*), compilando em uma tabela para uma posterior comparação. Com isso, foi realizada uma análise dos resultados obtidos ao final da pesquisa.

Resultados e Discussão

Através da análise primária realizada por meio do primeiro tratamento da base de dados no Microsoft Excel, os colaboradores de faixas etárias entre 21 – 30 anos e de 31 – 40 anos, representando 31% e 38%, respectivamente, evidenciam ser a maior parcela de trabalhadores que sofreram acidentes de trabalho. Além disso, o gênero feminino marca presença em 82% dos casos.

Dentre os tipos de acidentes, os típicos apresentaram uma frequência de 73%, seguido dos de trajeto com 26% e por último os de doenças com 1%. Com relação os tipos de lesões, cinco se destacaram, sendo a maior delas o corte e a laceração (26,2%), atingindo diversas partes do corpo. Ademais, os locais destes acidentes se concentram em São Paulo e Maranhão.

Em uma análise secundária, com a aplicação das doze técnicas de mineração de dados, os resultados obtidos foram organizados na Tabela 1, cruzando a técnica utilizada com cinco métricas, sendo elas: acurácia, precisão, *recall*, F1 score e curva ROC/AUC, sendo destacado em negrito os melhores valores de cada um.

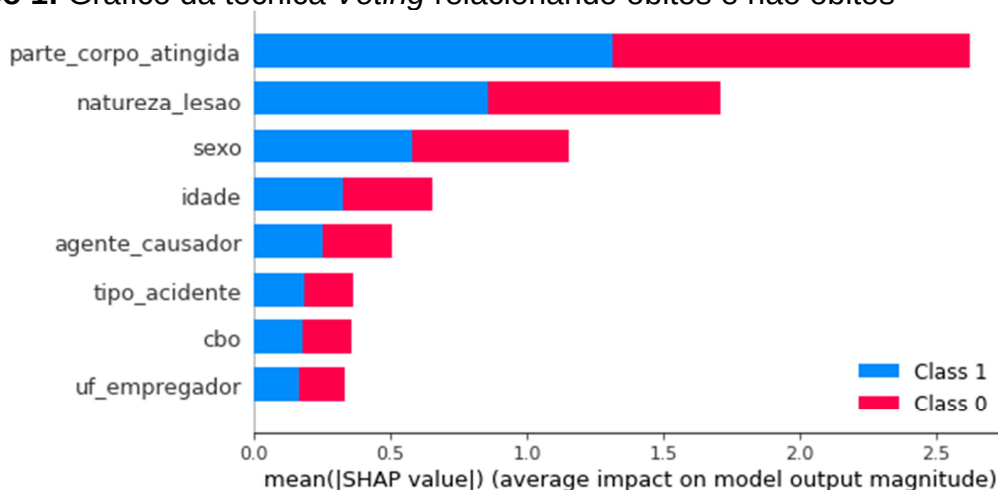
Tabela 1: Resultados das métricas e tempo computacional das técnicas

Técnica	Acurácia	Precisão	Recall	F1 score	ROC	Tempo (s)
Ensemble						
ET	0,85366	0,75000	0,93750	0,83333	0,90375	6,63440
RF	0,80488	0,78947	0,78947	0,78947	0,85646	9,58370
XGB	0,65000	0,77778	0,58333	0,66667	0,83333	3,95540
BA	0,75610	0,78261	0,78261	0,78610	0,81884	59,35950
VO	0,85000	0,83333	0,90909	0,86957	0,89899	4,71720
ST	0,78049	0,72414	0,95455	0,82353	0,85167	8,17490
Não ensemble						
SVM	0,65000	0,57143	0,88889	0,69565	0,73737	4,66960
LR	0,74520	0,75000	0,75000	0,75000	0,79048	15,57390
KNN	0,70732	0,59091	0,81250	0,68421	0,78000	7,25570
NB	0,70000	0,71429	0,83333	0,76923	0,85417	7,50070
DT	0,75610	0,78261	0,78261	0,78261	0,75242	17,99470
NN	0,73171	0,70833	0,80952	0,75556	0,83095	6,85790

Fonte: Autoria própria (2022)

Analisando as técnicas e cruzando com as métricas, a técnica *Voting* apresenta os melhores resultados de desempenho. Além disso, as categorias de atributos e suas respectivas relevâncias foram levadas em consideração, visto no Gráfico 1, o qual evidencia que a parte do corpo atingida e a natureza da lesão possuem maior influência nos óbitos. Da mesma forma que o tipo de acidente, CBO e localização do empregador não possuem intervenção considerável no cenário analisado.

Gráfico 1: Gráfico da técnica *Voting* relacionando óbitos e não óbitos



Fonte: Autoria própria (2022)

Conclusões

Diante do objetivo de identificar quais eram os principais fatores que influenciam nos acidentes de trabalho na área da saúde e com base na constatação de que colaboradores de 30 à 40 anos (38%), principalmente mulheres são vítimas frequentes destes acidentes majoritariamente de corte e laceração em diversas partes do corpo. A aplicação da mineração de dados entrou para proporcionar maiores insumos para tomadas de decisão mais assertivas, pois através da melhor técnica de mineração eleita, pode-se perceber que os maiores pontos de atenção são a parte do corpo atingida e a natureza da lesão, e por consequência são os que devem ser reavaliados constantemente a fim de diminuir cada vez mais os respectivos índices com planos de ações preventivos e garantir a segurança de todos.

Agradecimentos

Agradeço ao CNPq pelo apoio financeiro à pesquisa.

Referências

ASCOM HUAC, **Um a cada 10 acidentes de trabalho ocorre na área da saúde**, UFCG, abr/2010.

BAKKE, H.A., ARAÚJO, N.M.C. **Acidentes de trabalho com profissionais da saúde**, Produção, v. 20, n. 4, out./dez. 2010.