**Resumo Apresentações 26/08/19**

O projeto “Análise de Sentimentos Rápida e Precisa com um Classificador Naive Bayes Aprimorado” utilizou o artigo “Fast and Accurate Sentiment Classification Using an Enhanced Naive Bayes Model” proposto em 2013 na 14th International Conference IDEAL 2013 (Intelligent Data Engineering and Automated Learning). Como insumo, o grupo utilizou dados do Kaggle com 50.000 comentários de avaliação de filmes etiquetados em avaliações positivas e negativas em uma mesma proporção (dados balanceados).

Durante a preparação dos dados, o grupo aplicou a tokenização das frases, remoção de sinais de pontuação, utilização de n-gramas, manipulação da negação, seleção de features (maiores frequências e PMI – Pointwise Mutual Information) e, por fim, a aplicação de um algoritmo de Naive Bayes tradicional e um Bernoulli Naive Bayes Classifier.

Um dos grandes desafios enfrentados pelo grupo foi justamente o poder computacional necessário para realizar todo o pipeline de preparação e treinamento dado que, por vezes, a memória RAM não foi suficiente para manter a cpu em boas condições. Com a feature selection, o tempo de processamento foi reduzido em grandes quantidades. Por fim, o grupo alcançou resultados com grande diferença para o modelo aplicando feature selection (métricas menores) para sem feature selection (métrica maiores).

No projeto “Survey on Hate Speech Detection Using Natural Language Processing”, o grupo apresentou um tema relacionado a pesquisa sobre detecção de fala de ódio usando processamento de linguagem natural. Como motivação, o grupo mencionou o crescimento das estruturas das redes sociais, juntamente com a acessibilidade massiva à internet e a polarização social e política.

Como métodos de preparação, o grupo utilizou bag-of-words, n-gramas de nível de caracteres, modelos estatísticos (LDA), Regressão Bayesiana e captação do contexto da frase. Na apresentação dos resultados, o grupo mostrou as principais palavras presentes em discursos de não-ódio (love, posit, smile, healthi, fun, life, etc) e também em discursos de ódio (Trump, polit, allahsoi, liber, libtard, sjw, etc).