# FAKE NEWS DETECTION
## & Financial Dashboarding

## PROJECT STATEMENT

To produce a text analytics algorithm for EBOS Reputation to determine the truthfulness of the articles found online, and to create an informative financial dashboard for EBOS' clients

## FINANCIAL DASHBOARDING

An informative dashboard was built so that the clients can utilize it to gain more useful information on a certain business, and to compare those businesses with industrial benchmarks, as part of their clients' due diligence process. The financial dashboard comprises three main views: **(i) Financial Information** - provides an overview of the financial performance of the companies & comparison with other companies, **(ii) Company Information** - description of companies & other qualitative information **(iii) Industry Comparison** - comparison of company ratios against industry benchmark ratios & countries

## TEXT ANALYTICS

As the objective is to generate accurate estimates of the probability that a given news article is fake, models were evaluated based on logarithmic loss and Brier score, which are metrics to compute errors in probability estimates.

**Logistic Regression:** This algorithm produces well-calibrated probability estimates and had the best performance on our dataset.

**Multinomial Naive Bayes:** An inherently probabilistic model, this algorithm performed moderately well.

**Deep Learning:** A convolutional neural network was constructed based on word embeddings, which are numerical representations capturing relationships between words.

The final model was an ensemble consisting of the Logistic Regression and Multinomial Naive Bayes models. This model was trained on the 2,000 words with the highest ANOVA F-Scores.

## DASHBOARD VISUALIZATIONS

PREPARED BY: ANJITHA RAMESAN, CHUA KIM YEN, KANG YI XI, NANDINI SANGEETHA NAIR, WONG JIA HUI VALERIE