Trustpilot reviews for business sentiment

- Made by Kasper Overgaard Dahl & Markus Agnsgaard
Link to Github: https://github.com/TheAgns/ExamProjectDataScience

Project statement:

The challenge we are going to address is to compare the sentiment of customer evaluations for organizations/businesses using sentiment analysis and topic modeling on customer reviews from Trustpilot and then use logistic regression with our data to try to predict an unknown sentence that the model hasn't seen before.

Customer evaluations are a crucial source of information for organizations to understand the experiences of their clients and to pinpoint areas in need of development. Topic modeling may assist to find the main themes or subjects that consumers are discussing, while sentiment analysis can help to automatically categorize reviews as positive, negative, or neutral. Comparing the tone of customer evaluations for various firms may reveal insightful information about customer preferences and assist companies in gauging their performance against that of their rivals.

Project solution:

We have scraped reviews from Trustpilot and then made a program that employs NLP methods such as Logistic Regression and Naïve Bayes Classifier to do sentiment analysis. Furthermore, we used BERTopic for topic modeling to categorize each review into a topic. The program would also contrast the tone of customer evaluations for the companies and display the findings using a variety of data visualization tools, including Neo4j, word clouds, bar charts, confusion matrices, clusters in scatter plots and more.

These findings can help a business to adjust to the market and get an insight into the company's competitors; what they do good and bad.