**British Airways sentiment analysis and Predictive Modeling of Customer Bookings**

**Business objectives:**

Customers are more empowered than ever because they have access to a wealth of information at their fingertips. This is one of the reasons the buying cycle is very different to what it used to be. Today, if you’re hoping that a customer purchases your flights or holidays as they come into the airport, you’ve already lost! Being reactive in this situation is not ideal; airlines must be proactive in order to acquire customers before they embark on their holiday.

**Problem definition:**

Scrape data to gain company insight and build a machine learning model capable of predicting customers buying behavior

**Implementation process:**

* **Web scraping:** the first thing to do is to scrape reviews data from the web.
* **Data analysis:** Once we have your dataset, you need to prepare it. The data will be very messy and contain purely text.
* **Explore and prepare the dataset:** understand the different columns of the dataset, some statistics of the dataset and prepare the dataset for predictive modeling.
* **Train a machine learning model:** we will train a machine learning model to be capable of predicting the outcome, which is a customer making a booking.
* **Model evaluation:** after training the model, we will evaluate it performance by conducting cross-validation and outputting appropriate evaluation metrics.

Reference:

**Problem statement:** British Airways

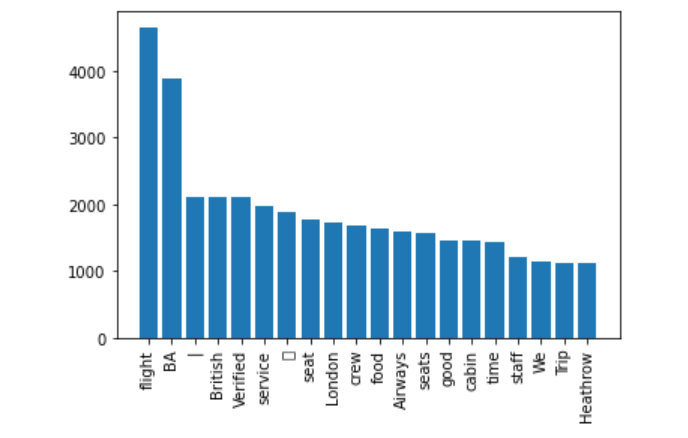
**Data**: [https://www.airlinequality.com/airline-reviews/british-airways]](https://www.airlinequality.com/airline-reviews/british-airways%5D)

**GitHub: https://github.com/armandjucelin/British-Airways-sentiment-analysis-and-Predictive-Modeling-of-Customer-Bookings**

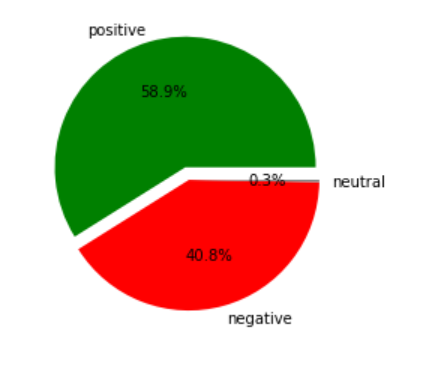
Tech Stack Used:

* Python,
* Machine Learning

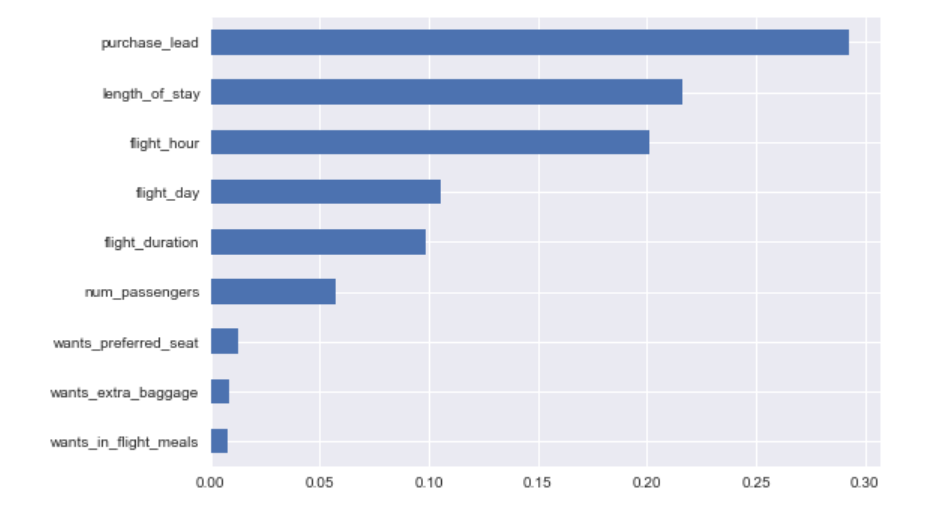
**Results:**

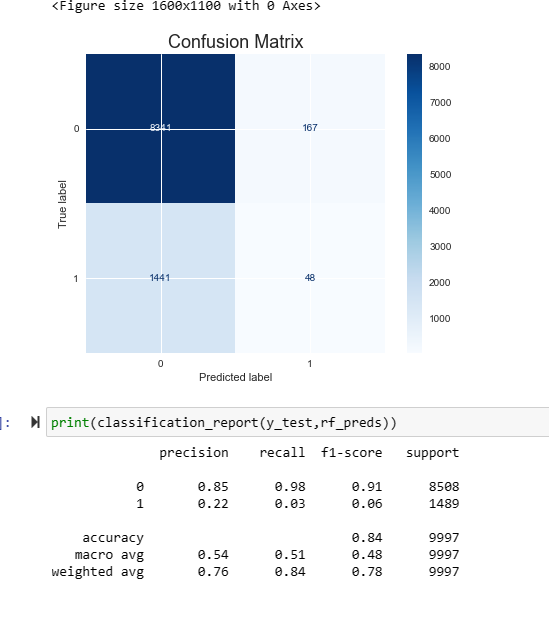
**Key topics:** The frequency chart below shows that the key topics within the reviews include “service”, “seat”, “crew” and “staff”, showing that people are actively talking about their experience and staff. 

**Sentiment:** Out of 1000 reviews, 546 were positive, 437 were negative and 17 were neutral. This means the majority of reviews were polarized as either positive or negative.



**Evaluation**

* The most important variable in the model was purchase lead, that is the time between purchase and departure.
* Information about the flight, e.g. flight time and duration were also significant, however booking origin of the customer was not important.
* The accuracy of the model was approximately 0.76 (Precision) and 0.84 (Recall), showing that this model requires more improvement. I suggest adding more customer-centric features into the model. 



**Application:**

* Acquired customer before they embark on their holiday for British airway using data and predictive model.
* Apply analytical skills to make tangible impact by providing recommendations, tools, and models that drive key business decisions.

**Strategy for future actions:**

* **Update and refine:** continuously monitor and update your web scraping process to ensure your data remains up to date. analyses the result regularly and refine your strategy base on the new insight.
* **Deploy and monitor the model:** Deploy and trained model into our production environment and continuously monitor it performance. regularly update the model as new data becomes available to ensure its predictions stay relevant and reliable.