Automatic Summarisation of Product Reviews using Text Mining

Bibhudutta Moharana

M.Sc. in Big Data

UNIVERSITY *of* STIRLING
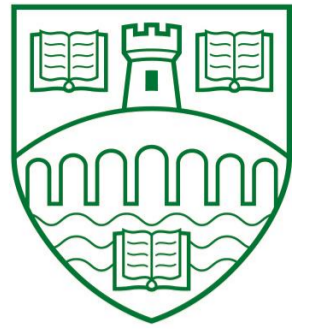
PROBLEM

Public perception of products has always been of great interest to companies. Tracking public opinion has usually been time consuming, requiring manual compilation of reviews and costly user service. It is impossible to read and comprehend ever growing number of reviews published on e-commerce websites.

CURRENT PRODUCT FILTERS
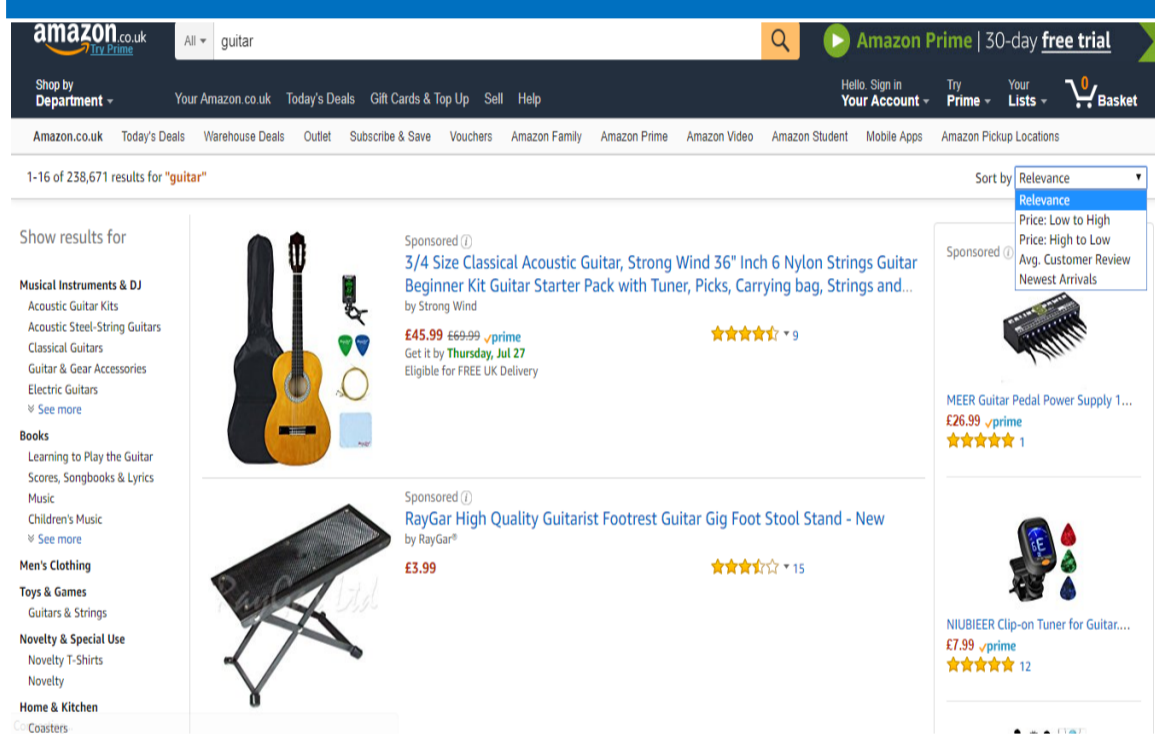
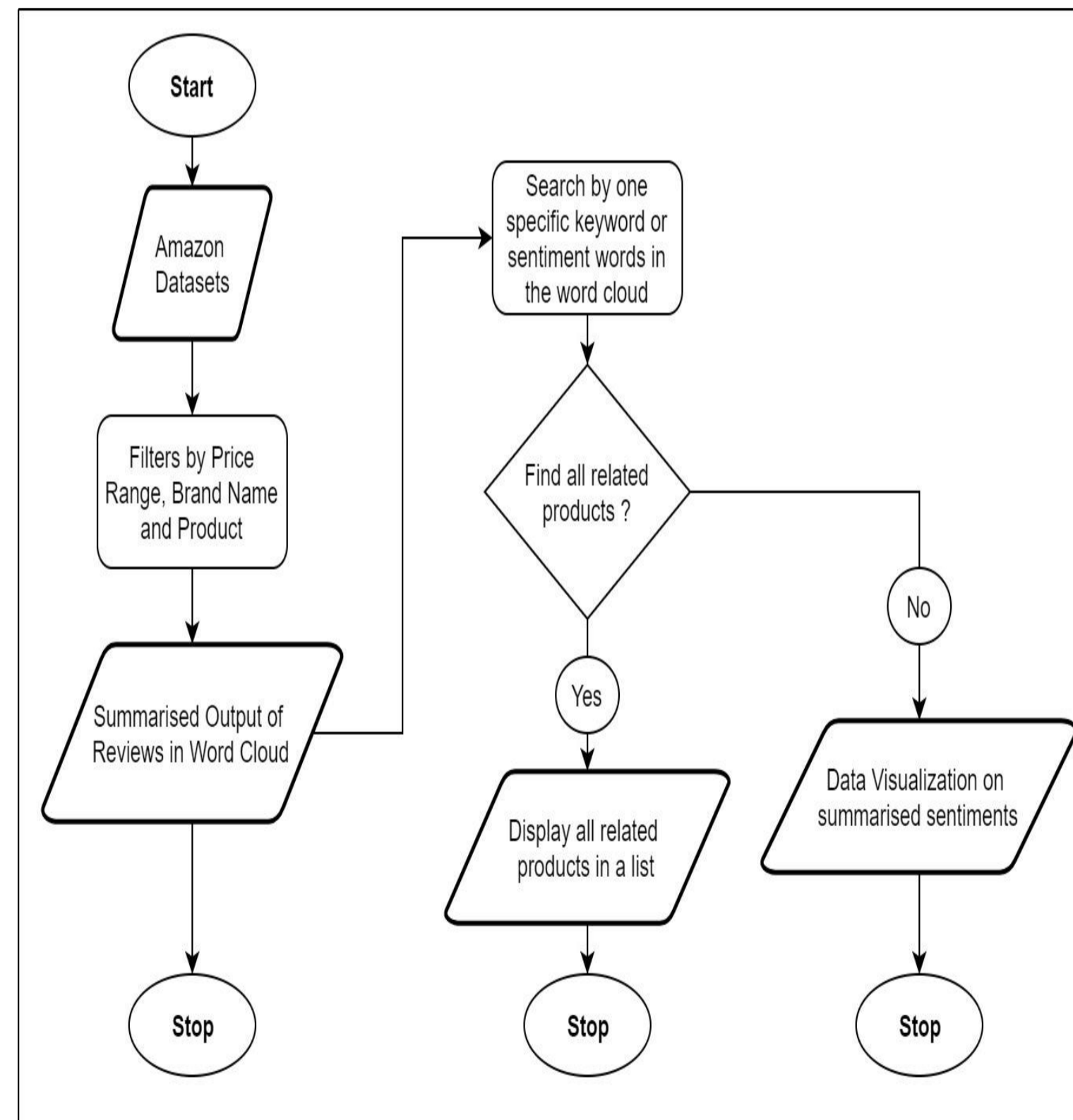### PROPOSED PRODUCT FILTERS
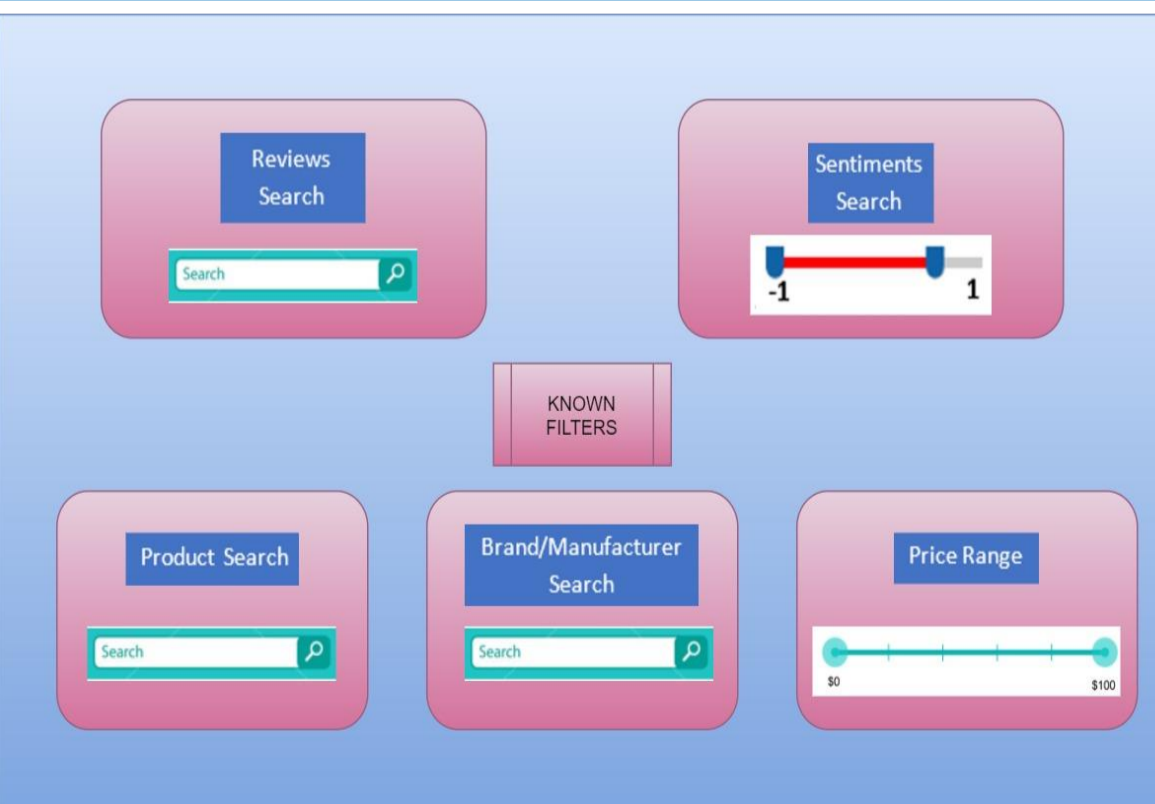
AIM

To automatically summarise large number of reviews by:

- 1. Extracting meaningful summary information.
- 2. Identifying the sentiments of a review.
- 3. The summarised reviews and identified sentiments are visualized as Word Cloud and Bar-graph respectively.

FLOWCHART OF PROPOSED METHODOLOGIES

WORKFLOW

1. A graphical user interface is created to filter products by Price, Brand and Product Type as well as reviews search and sentiments polarity. The filtered reviews are summarised as word cloud.

2. This stage involves linking the data from both the datasets i.e. Metadata about all the products and Reviews given by the customer with respect to each product. Python functions are implemented to filter the reviews according to Price, Brand and Product Type.

3. Data Visualisation and Sentiment analysis is carried out to allow deeper understanding of the information. The reviews are visualized as both 'Word Cloud' and bar graph showing the sentiments of the reviews.

WORD CLOUD OF SUMMARISED REVIEWS

SENTIMENT ANALYSIS OF REVIEWS

Technologies Used : amazon · python · mongoDB · {JSON} · Tk&π · matplotlib · Natural Language Analysis with Python NLTK