Daraz Sentiment Analysis

Project Report

Session: 2018 − 2022

Submitted by:

Noor-ul-hassan   2018-CS-604

Mohsin Ikram   2018-CS-617

Talha Nouman   2018-CS-626

M.Saad Fareed   2018-CS-632

Supervised by:

Dr.Irfan Yousuf

Department of Computer Science, New Campus

University of Engineering and Technology

Lahore, Pakistan

Contents

List of Figures

List of Tables

Chapter 1

Project Overview

1.1 Introduction

Daraz is the leading online marketplace in South Asia, empowering tens of thousands of sellers to connect with millions of customers. Daraz provides immediate and easy access to 22 million products in 100+ categories to deliver more than 2 million packages every month to all corners of the country. Daraz is a mall, a marketplace, and a community for its customers. It is also a university for entrepreneurs, and every month it educates more than 2,000 new sellers on ecommerce operations. To overcome the logistics challenge in its markets, Daraz has built its own logistics company specifically designed for e-commerce operations - Daraz Express (known as DEX) - to raise the standards for the industry. In 2018, Daraz was acquired by Alibaba Group, and is proud to carry its part of the mission to 'make it easy to do business anywhere in the era of the digital economy'. As part of the Alibaba ecosystem, Daraz is leveraging Alibaba's global leadership and experience in technology, online commerce, mobile payments, and logistics to drive growth in its markets. Most recently, Daraz is interested in understanding its existing consumer base. Therefore, for our Experiential Learning Program (ELP) Daraz has entrusted us with the project, "Analysis of User Sentiment" in which the research agenda is to analyze brand awareness and understand reasons for brand sentiment pertaining to current users and customers both. The methodology would encapsulate both qualitative and quantitative research through field and online avenues. The most general use of sentiment analysis is in ecommerce activities. Websites allows their users to submit their experience about shopping and product qualities. They provide summary for the product and different features of the product by assigning ratings or scores. Customers can easily view opinions and recommendation information on whole product as well as specific

product features. Graphical summary of the overall product and its features is presented to users.

1.2 Our Project

The project is all about sentiment analysis of Product reviews of Daraz.pk. In our Project we have the target category is Electronics Accessories.

Following are the steps that we follow to complete our project, these are listed below:

1. Data Scraping of Electronics Accessories from Daraz.

2. Pre-processing on Scrap data so that we have better results.

3. Visualization of Data for better understanding.

4. Sentiment analysis

1.2.1 Data Scraping

In first step we do Scrapping of the Electronics Accessories data from Daraz using python selenium and beautifulsoup. following are the list of data that we scrap from our Data Scraper.You can check DarazDataComplete.csv file for more details.

- Product Number

- Brand Name

- Brand Title

- URL

- Average Rating

- Review

1.2.2 Pre-processing

Data preprocessing can refer to manipulation or dropping of data before it is used in order to ensure or enhance performance, and is an important step in the data mining process.In our project we do preprocessing to remove all the products that don't have any reviews.
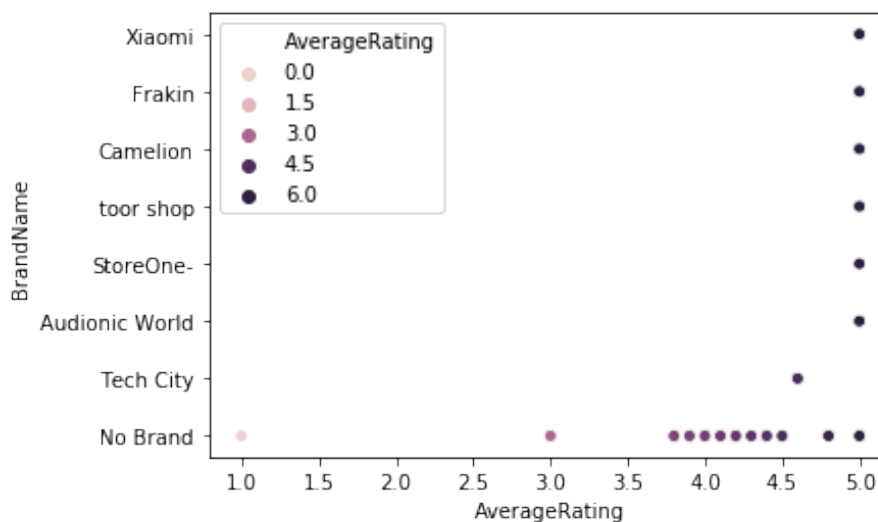
1.2.3 Visualization

Data visualization gives us a clear idea of what the information means by giving it visual context through maps or graphs.

FIGURE 1.1: Review Rating analysis
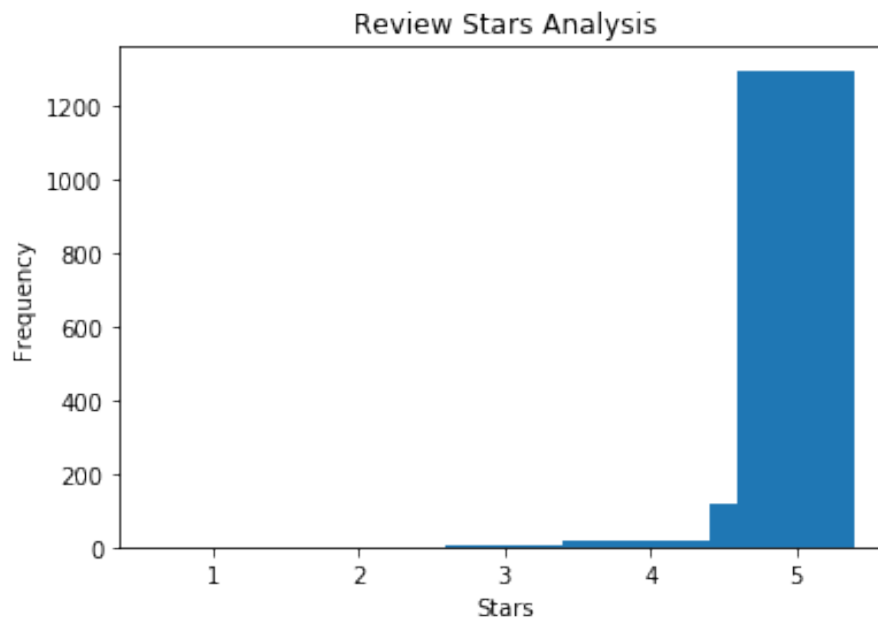
This graph show Rating frequency analysis. In this graph we have products group by stars rating.

FIGURE 1.2: Brand Rating analysis

This graph show brand rating analysis. In this graph we have products group by Average rating of different brands.

1.2.4 Sentiment Analysis

Sentiment analysis is contextual mining of text which identifies and extracts subjective information in source material.

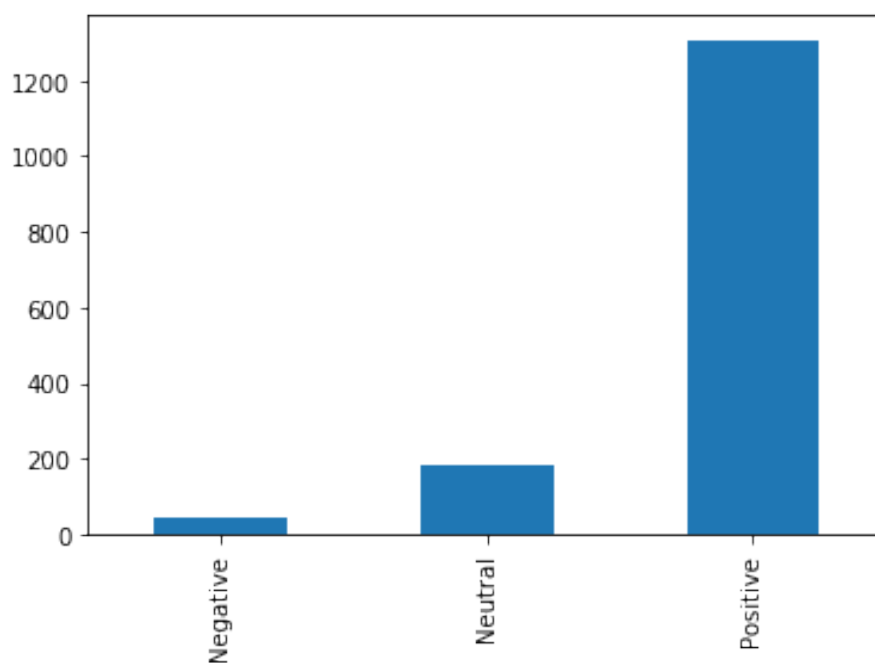FIGURE 1.3: Sentiment analysis Review
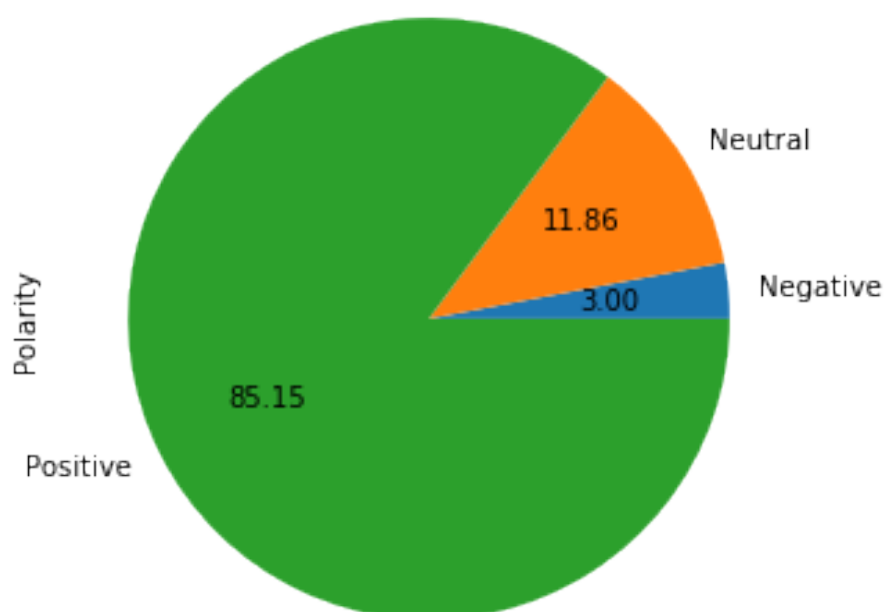

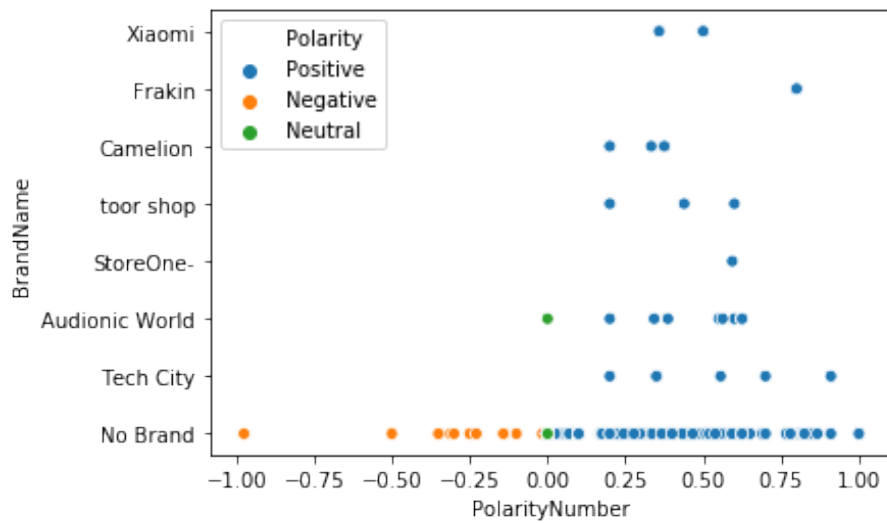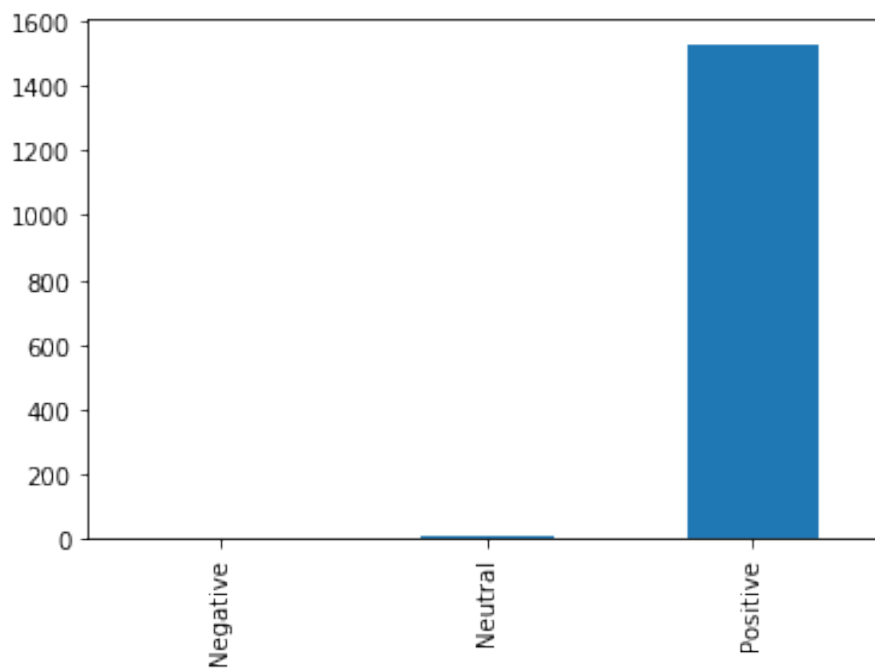FIGURE 1.4: polarity Sentiment analysis

FIGURE 1.5: Review Rating analysis

1.2.5 Rating based Analysis

FIGURE 1.6: Rating Based analysis

This graph show rating based analysis. In this graph The products whose rating is less than 3 are group into negative, products with greater than 3 rating are goes into the group of positive sentiment group and all other goes into neutral sentiment products.

References