A MINI-PROJECT REPORT

ON

**”Summarization & Sentiment Analysis of German Amazon Reviews”**

BY

# **Aayush Singh**

# **Deepanshu Sonparote**

# **Deepali Zutshi**

Under the guidance of

**Internal Guide**

**Prof. Suresh Mestry**



Juhu-Versova Link Road Versova, Andheri(W), Mumbai-53

Department of Computer Engineering

University of Mumbai

April- 2020



Juhu-Versova Link Road Versova, Andheri(W), Mumbai-53

**CERTIFICATE**

# Department of Computer Engineering

This is to certify that

1. Aayush Singh
2. Deepanshu Sonparote
3. Deepali Zutshi

Have satisfactory completed this project entitled

**“Summarization & Sentiment Analysis of German Amazon Reviews”**

Towards the partial fulfilment of the

**FINAL YEAR BACHELOR OF ENGINEERING**

**IN**

**(COMPUTER ENGINEERING)**

**as laid by University of Mumbai**.

|  |  |
| --- | --- |
| **Guide**  **Prof. Suresh Mestry** | **H.O.D**  **Dr. Satish Y. Ket** |
| **Principal**  **Dr. Sanjay Bokade** | |

**Project Report Approval for B. E.**

This project report entitled” **Summarization & Sentiment Analysis of German Amazon Reviews*”*** by ***Aayush Singh, Deepanshu Sonparote and Deepali Zutshi*** is approved for the degree of ***Bachelor of Computer Engineering***.

**Examiners:**

1. ­­­­­­­\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. ­­­­­­­­­­­­\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date:

Place:

**Declaration**

We wish to state that the work embodied in this project titled ”Summarization & Sentiment Analysis of German Amazon Reviews” forms our own contribution to the work carried out under the guidance of ”Guide Name” at the Rajiv Gandhi Institute of Technology.

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

**(Students Signatures)**

Aayush Singh (B-858) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Deepanshu Sonparote (B-859) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Deepali Zutshi (B-870) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Abstract

The end of the previous decade led to a booming expansion of Amazon as a leading retail company. With the internet coverage reaching out to remote locations, the visits on such online retailers have increased exponentially. As a general tendency, people look at suggestions and feedback from the crowd to decide on whether to purchase commodities online. The paper utilizes German as its core language of analysis. This paper describes a technique to provide users with concise and accurate reviews of the product. It also provides a general outlook, which the prior customers have, of the product. The technique utilizes the classification of original and summarized reviews to training a model for review sentiment analysis using various Machine Learning models. The model, being one of a kind, utilizes data gathered by web scraping. The Encoder-Decoder LSTM model operates on this to create a summary of the reviews. The model performs abstractive summarization to retain the sense of statements. The same operations are performed on the original reviews. These two summaries are then fed to a sentiment analyzer to generate polarity of the reviews.

**Contents**

|  |  |  |
| --- | --- | --- |
| **List of Figures** | | **i** |
| **List of Tables** | | **ii** |
| **1** | **Introduction**   * 1. Introduction Description . . . . . . . . . . . . . . . . . . .   2. Organization of Report . . . . . . . . . . . . . . . . . . . . | **1**  1  1 |
| **2** | **Literature Review**  2.1 Survey Existing system . . . . . . . . . . . . . . . . . . . .  2.2 Problem Statement and Objective . . . . . . . . . . . . . . .  2.2.1 Objectives . . . . . . . . . . . . . . . . . . . . . . . . | **2**  2  2  2 |
| **3** | **Proposed System**  3.1 Data Collection . . . . . . . . . . . . . . . . . . . . . . . .  3.2 Data Pre-processing . . . . . . . . . . . . . . . . . . . . .  3.2.1 Cleaning . . . . . . . . . . . . . . . . . . . . . . . . .  3.2.2 Applying NLP Techniques for Pre-processing . . . . . . .  3.3 Text Summarization . . . . . . . . . . . . . . . . . . . . . .  3.3.1.Word Embeddings . . . . . . . . . . . . . . . . . . . .  3.3.1.1. Vectorization . . . . . . . . . . . . . . . . . . .  3.3.2. Encoder-Decoder LSTMseq2seq Model . . . . . . . . . .  3.3.2.1. Training Phase . . . . . . . . . . . . . . . . . .  3.3.2.2. Inference Phase . . . . . . . . . . . . . . . . .  3.4 Sentiment Analysis . . . . . . . . . . . . . . . . . . . . . .  3.4.1. Count Vectorizer . . . . . . . . . . . . . . . . . . . .  3.4.2. Predicting Models . . . . . . . . . . . . . . . . . . . . | **3**  3  3  3  4  4  4  5  5  6  7  7  7  8 |
| **4** | **Results** | 9 |
| **5** | **Conclusion and Future Work** | 10 |
|  | **References** | 11 |