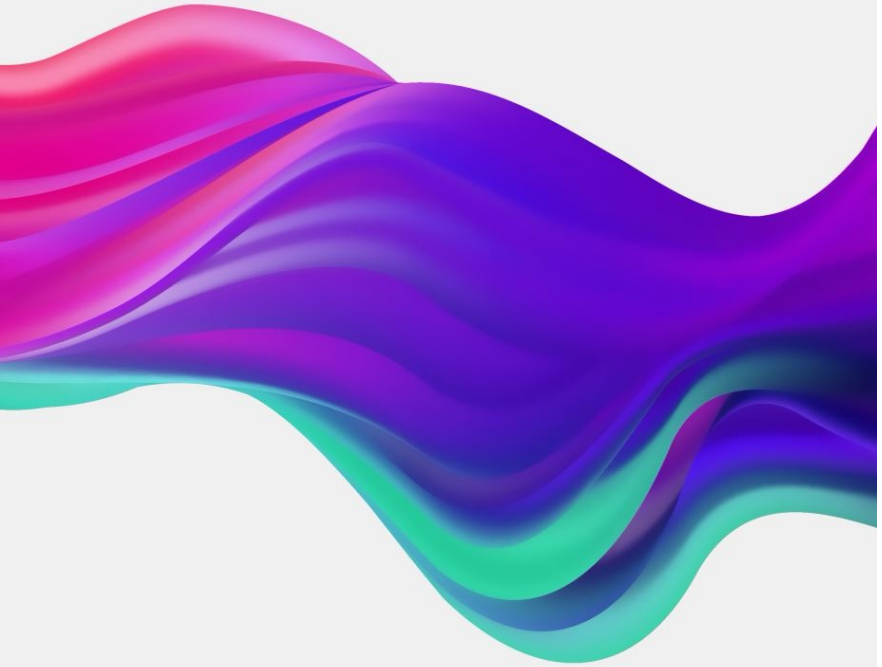


Title: Sentiment Analysis on Large Scale Amazon Product Reviews

Author: Tanjim Ul Haque, Nudrat Nawal Saber, Faisal Muhammad Shah

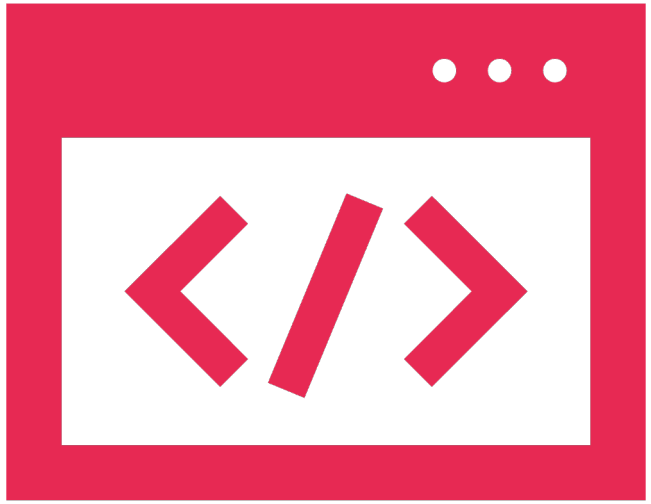
Year: 2018

Conference/Journal Name: IEEE International Conference on Innovative Research and Development, May 11-12, 2018, Bangkok, Thailand.

Student Name: Muthu Selvam (801276057)

Introduction

In the era of burgeoning e-commerce, consumer decisions are increasingly influenced by online reviews. This project, titled "**Sentiment Analysis on Amazon Product Reviews**," aims to delve into the world of sentiment analysis to decipher the sentiments embedded in vast Amazon product review datasets.



Objectives & Goals

Primary Objective:

- Develop a supervised learning model for sentiment analysis on Amazon product reviews.

Specific Goals:

- Achieve accuracy exceeding 90% in sentiment classification.
- Explore and implement feature extraction methods such as Bag of Words, TF-IDF, and Chi Square.
- Compare and evaluate various classifiers, including Linear Support Vector Machine, Multinomial Naïve Bayes, Stochastic Gradient Descent, Random Forest, Logistic Regression, and Decision Tree.

Problem Statement

The existing challenge revolves around the need for a reliable and time-efficient model to analyse sentiments in a large-scale Amazon product review dataset.

Dataset:

- Utilize a labelled Amazon product dataset provided by researchers.
- Focus on three categories: Electronics, Cell Phone & Accessories, and Musical Instruments.
- Dataset contains approximately 48,500 reviews, categorized as 21,600 for mobile phones, 24,352 for electronics, and 2,548 for musical instruments.

Dataset

Dataset: Amazon product reviews (categories: Electronics, Cell Phone & Accessories, Musical Instruments)

https://www.kaggle.com/datasets/abdallahwagih/amazon-reviews

Methodology

Active Learning Approach:

- Implement a pool-based active learning process for iterative dataset labeling.
- Utilize user feedback to enhance model accuracy.

Feature Extraction Methods:

- Employ Bag of Words approach focusing on nouns and adjectives.
- Apply TF-IDF and Chi Square methods for effective feature extraction.

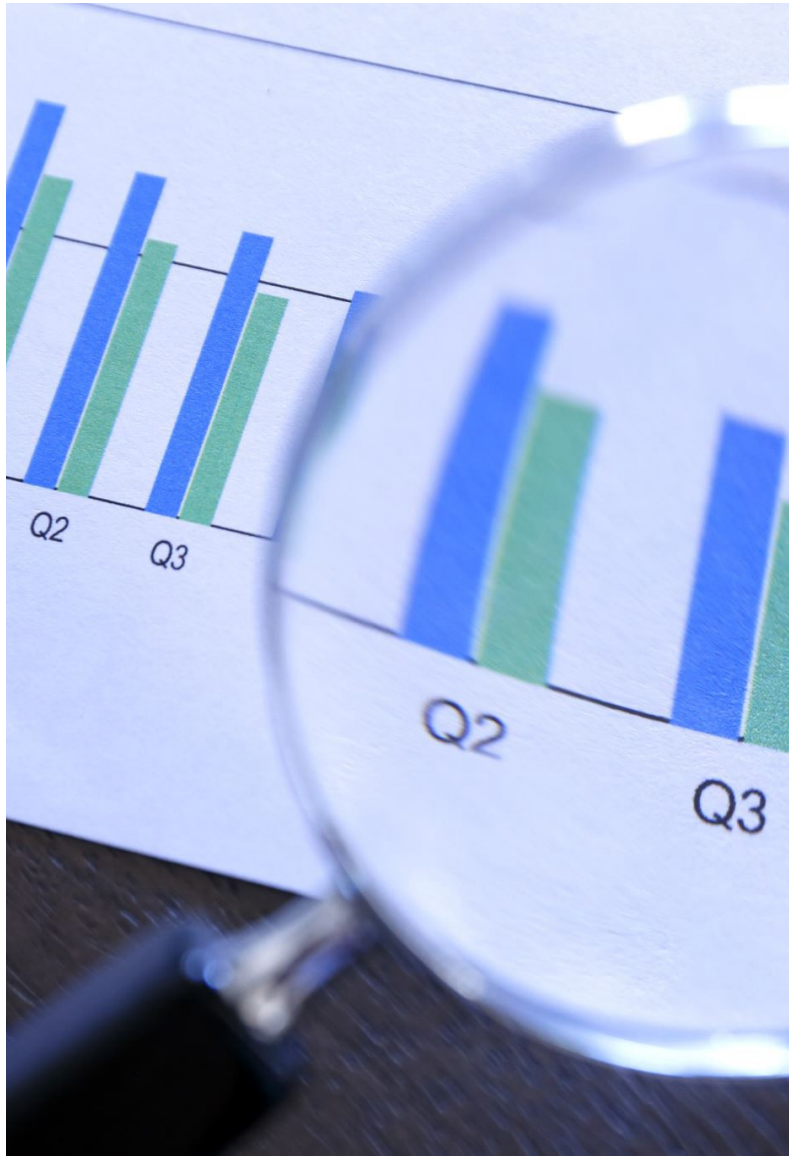
Evaluation

Classifier Comparison:

- Evaluate the performance of multiple machine learning algorithms through 10-fold cross-validation.
- Analyse accuracy, precision, recall, and F1 score.

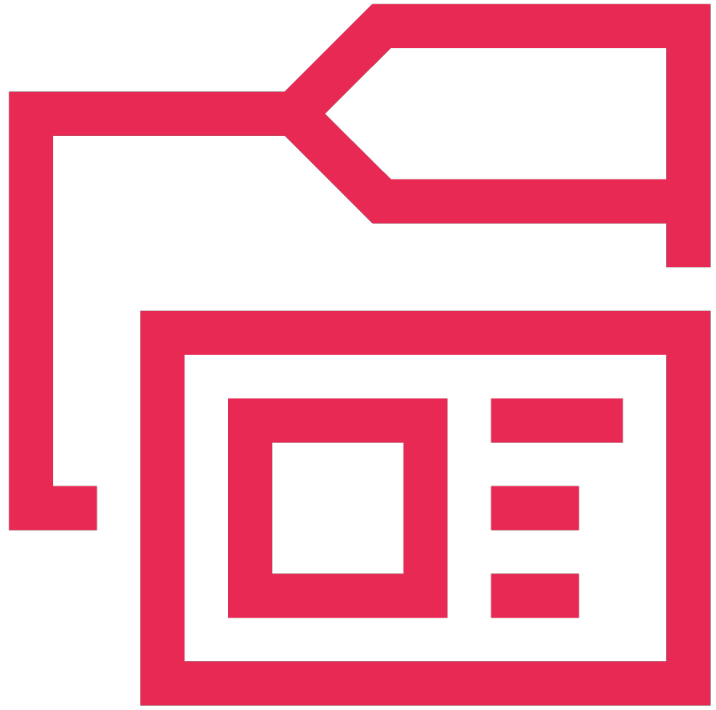
Comparative Analysis:

- Compare the proposed model's accuracy with related works in sentiment analysis on Amazon product reviews.



Related Works

- Overview of existing research in sentiment analysis on product reviews.
- Mentioned paper: "Sentiment Analysis on Large Scale Amazon Product Reviews" (2018).
- Survey on the area: Brief summary of other related works in sentiment analysis, focusing on three primary papers.



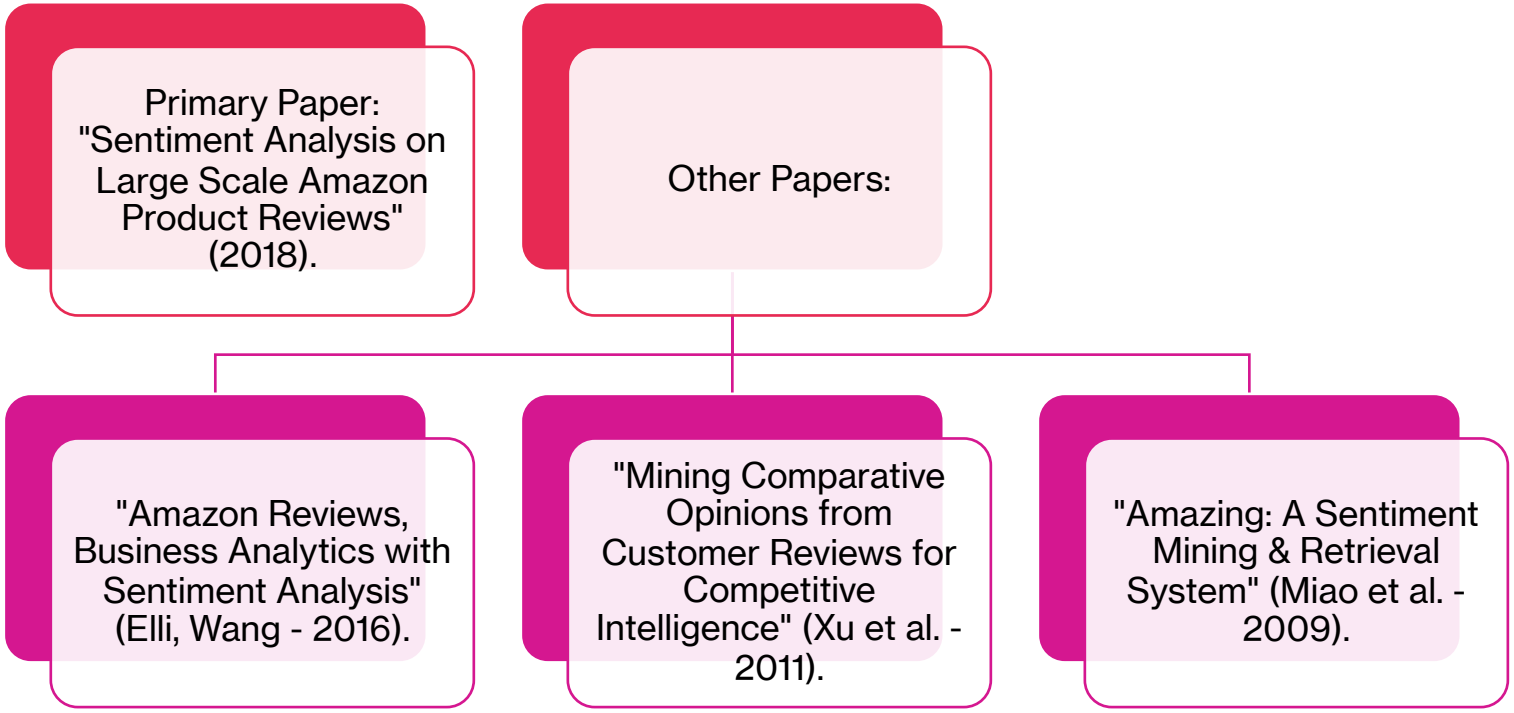
Citation and References Slide (MLA Style)

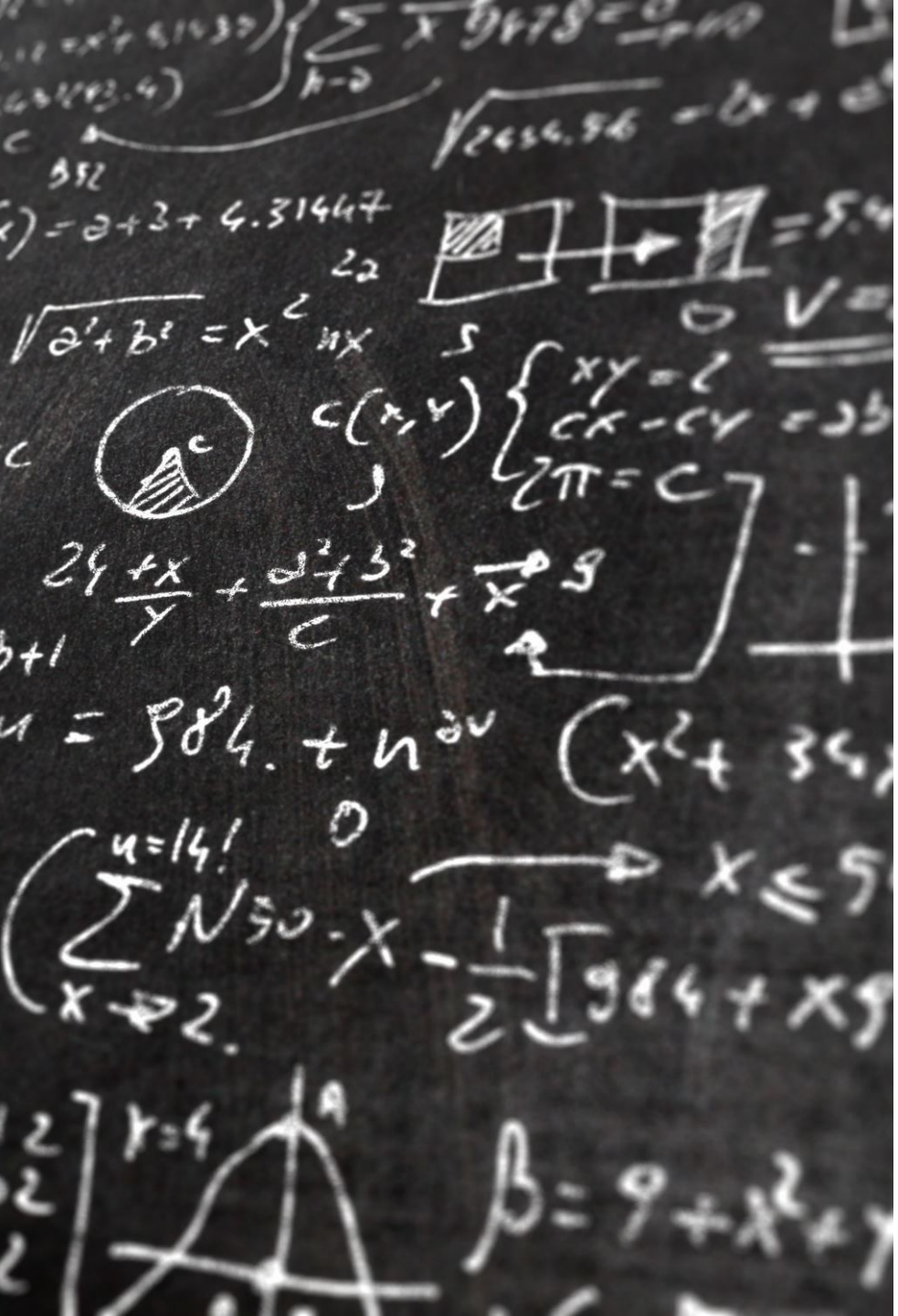
Haque, Tanjim Ul, et al. "Sentiment Analysis on Large Scale Amazon Product Reviews." *2018 IEEE International Conference on Innovative Research and Development, 11-12 May 2018, Bangkok, Thailand.* 2018, pp. 1-7. DOI: 10.1109/IRCD.2018.8477249.

References

- Haque, Tanjim Ul, Nudrat Nawal Saber, and Faisal Muhammad Shah. "Sentiment Analysis on Large Scale Amazon Product Reviews." 2018 IEEE International Conference on Innovative Research and Development, 11-12 May 2018, Bangkok, Thailand. 2018, pp. 1-7. DOI: 10.1109/IRCD.2018.8477249.
- Xu, Yun, Xinhui Wu, and Qinxia Wang. "Sentiment Analysis of Yelp's Ratings Based on Text Reviews." 2015.
- Rain, Callen. "Sentiment Analysis in Amazon Reviews Using Probabilistic Machine Learning." Swarthmore College, 2013.
- Shaikh, Tahura, and Deepa Deshpande. "Feature Selection Methods in Sentiment Analysis and Sentiment Classification of Amazon Product Reviews." 2016.
- Chen, Weikang, et al. "Amazon Review Classification and Sentiment Analysis." International Journal of Computer Science and Information Technologies 6.6 (2015): 5107-5110.
- Miao, Q., Li, Q., & Dai, R. "AMAZING: A sentiment mining and retrieval system." Expert Systems with Applications 36.3 (2009): 7192-7198.
- Xu, Kaiquan, et al. "Mining comparative opinions from customer reviews for Competitive Intelligence." Decision support systems 50.4 (2011): 743-754.

Other Papers on Similar Problem





Method Reproduction

Paper to Reproduce: "Sentiment Analysis on Large Scale Amazon Product Reviews" (2018)

Methodology:

- Active learning approach for iterative dataset labeling.
- Feature extraction methods: Bag of Words, TF-IDF, and Chi Square.
- Multiple classifiers: Linear Support Vector Machine, Multinomial Naïve Bayes, Stochastic Gradient Descent, Random Forest, Logistic Regression, and Decision Tree.

Delivery & Timeline

- March 8: Literature Review and Understanding Existing Methods
- March 22: Data Collection and Preprocessing
- April 5: Active Learning for Dataset Labeling
- April 5 : Feature Extraction and Model Training
- April 19: Evaluation and Results Analysis and Presentation
- April 29: Finalizing Project Report

Expectations

- Gain hands-on experience in sentiment analysis.
- Understand the practical implementation of active learning.
- Enhance skills in feature extraction and model evaluation.
- Contribute to the existing body of knowledge in sentiment analysis on product reviews.
- A sentiment analysis model with accuracy exceeding 90%.
- Insights into the effectiveness of different feature extraction methods and classifiers in the context of Amazon product reviews.

Conclusion

The proposed project aims to address the crucial task of sentiment analysis in Amazon product reviews, offering valuable insights for decision-making in the e-commerce landscape.