

# Fake Reviews Detection

NUCES-FAST

May 15, 2019

## Team Members:

- **Muhammad Moazzam Mushtaq** (P156060@nu.edu.pk)
- **Muhammad Raza** (P156097@nu.edu.pk)
- **Muhammad Ahmad** (P66111@nu.edu.pk)

**Supervisor: Dr. Muhammad Taimoor Khan**  
(taimoor.khan@nu.edu.pk)

# Table of contents

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

Future Work

References

Demo

Suggestions

- 1 Overview
- 2 Motivation
- 3 Project Work Flow Diagram
- 4 Data Pre-Processing
  - Data Sets
- 5 Feature Extraction
- 6 Classification
- 7 Work Break Down Structure
- 8 Future Work
- 9 References
- 10 Demo
- 11 Suggestions

# Overview

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

Future Work

References

Demo

Suggestions

- Fake reviews have become a dangerous prospect for online users.
- Fake reviews affected both consumers and companies.
- We have applied Natural Language Processing (NLP) techniques to extract features and supervised classification through Machine Learning (ML) algorithms to identify fake reviews.

# Motivation

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

Future Work

References

Demo

Suggestions

- This field is less explored.
- Reviews are extremely important for the reputation and revenue of an organization.
- It will be helpful in detecting the fake reviewers.

# Project Work Flow Diagram

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

Future Work

References

Demo

Suggestions



# Data Pre-Processing

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

Future Work

References

Demo

Suggestions

- Removing stop word.
- Tokenization.
- Lowering the capitalize words.

# Data Sets

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

Future Work

References

Demo

Suggestions

- **1. Consumer Reviews of Amazon Products<sup>1</sup>** (A list of over 34,000 reviews of Amazon products like the Kindle, Fire TV, etc.)
- **2. Consumer Reviews of Amazon Products<sup>2</sup>** (A list of over 21,000 reviews of Amazon electronic products)

---

<sup>1</sup><https://www.kaggle.com/datafiniti/consumer-reviews-of-amazon-products/version/>

<sup>2</sup>[https://github.com/aayush210789/Deception-Detection-on-Amazon-reviews-dataset/blob/master/amazon\\_reviews.txt](https://github.com/aayush210789/Deception-Detection-on-Amazon-reviews-dataset/blob/master/amazon_reviews.txt)



# Feature Extraction

## ■ TF-IDF.

(0, 7436)	0.17438300540878185
(0, 7364)	0.15716856486231465
(0, 7158)	0.22269705398615136
(0, 7099)	0.3070408458839408
(0, 6432)	0.258204578846634
(0, 6431)	0.24992759044753354
(0, 6368)	0.1727161316215619
(0, 5297)	0.20499407269013994
(0, 5257)	0.22018066449737095
(0, 5179)	0.21491349756085948
(0, 4524)	0.27060794251143616
(0, 4355)	0.221000100098099
(0, 4054)	0.23819493669953573
(0, 3930)	0.21419911283639081
(0, 2536)	0.14561667075869814
(0, 2272)	0.17175178726827556
(0, 1801)	0.16449766504421967
(0, 1648)	0.17860108731740834
(0, 1279)	0.24320049234936794

# Classification

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

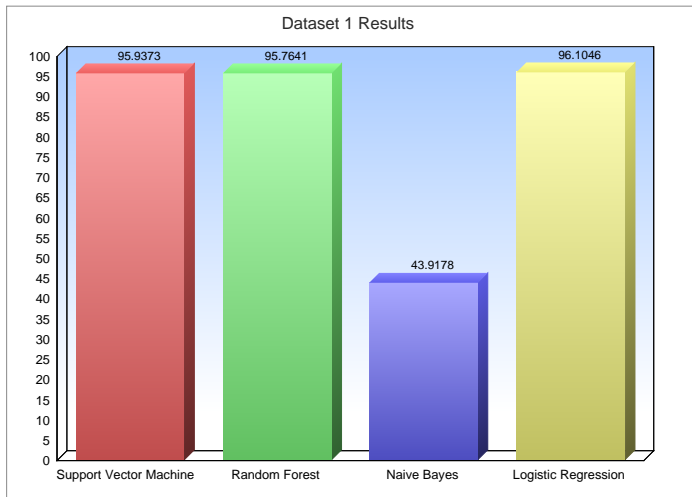
Future Work

References

Demo

Suggestions

## ■ 4 Fold Cross Validation.



# Classification

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

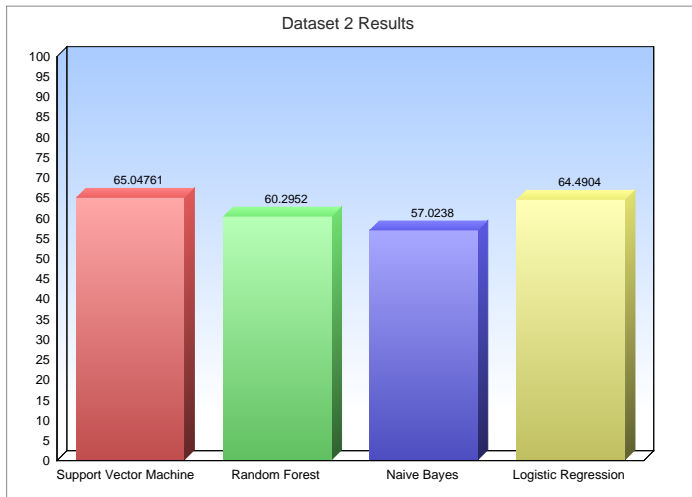
Future Work

References

Demo

Suggestions

## ■ 4 Fold Cross Validation.



# Work Break Down Structure FYP 1

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

Future Work

References

Demo

Suggestions

Proposal

Literature Review of Fake Reviews

Identification Of Features For Fake Review Classification

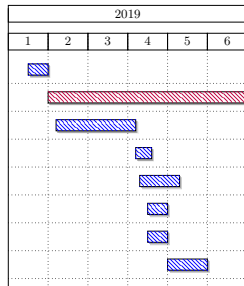
Planning

Chossing Techniques For Classification

Tunning Of Classifier Features

Testing and debugging

Documentation phase 1



# Future Work

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

**Future Work**

References

Demo

Suggestions

- We will try to improve the accuracy of our models.
- We will try to test other datasets to observe results.
- Some other models are also under consideration like Generative model.

# References

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

Future Work

References

Demo

Suggestions



Naveed Hussain, Hamid Turab Mirza, Ghulam Rasool, Ibrar Hussain, and Mohammad Kaleem.

Spam review detection techniques: A systematic literature review.

*Applied Sciences*, 9(5):987, 2019.



Huayi Li, Zhiyuan Chen, Bing Liu, Xiaokai Wei, and Jidong Shao.

Spotting fake reviews via collective positive-unlabeled learning.

In *2014 IEEE International Conference on Data Mining*, pages 899–904. IEEE, 2014.



Arjun Mukherjee, Bing Liu, and Natalie Glance.

Spotting fake reviewer groups in consumer reviews.

In *Proceedings of the 21st international conference on World Wide Web*, pages 191–200. ACM, 2012.

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

Future Work

References

**Demo**

Suggestions

# Demo

## Suggestions?

Fake Reviews  
Detection

Overview

Motivation

Project Work  
Flow Diagram

Data Pre-  
Processing  
Data Sets

Feature  
Extraction

Classification

Work Break  
Down  
Structure

Future Work

References

Demo

**Suggestions**