FRANCIS XAVIER®
ENGINEERING COLLEGE
AN AUTONOMOUS INSTITUTION
Vannarpettai, Tirunelveli - 627 003.
Department of Computer Science and Engineering

DECEPTIVE PRODUCT FEEDBACK IDENTIFICATION WITH ML

GUIDE NAME:

MRS. ELIZABETH PAULSHYAH

PRESENTED BY

NOEL ABRAHAM V (95072012063)

VIGNESH E (95072012321)

SRINIVASA DHARMAR (95072012318)

CONTENTS

- AREA OF THE PROJECT AND FIELD OF THE PROJECT
- PROBLEM STATEMENT
- LITERATURE SURVEY
- OBJECTIVES
- ABSTRACT
- INTRODUCTION
- EXISTING SYSTEM
- PROPOSED SYSTEM
- HARDWARE/SOFTWARE USED
- IMPLEMENTATION
- RESULT AND OUTPUT
- APPLICATIONS
- REFERENCES

AREA OF THE PROJECT AND FIELD OF THE PROJECT

AREA OF PROJECT : Machine Learning (ML)

FIELD OF PROJECT : E-Commerce

PROBLEM STATEMENT

- Authenticity of online product reviews crucial for consumer trust, yet deceptive feedback prevalent, jeopardizing platform credibility.

- Traditional manual detection methods inefficient for handling large review volumes, leading to time-consuming processes and inadequate results.

- Traditional ML relies on labeled data, struggles with semantic relationships. Deep learning excels in pattern recognition but demands resources and can be opaque in decision-making.

LITERATURE SURVEY

SI. NO	TITLE AND AUTHOR, YEAR	PUBLISHER & JOURNAL NAME	METHODOLOGY USED	DRAWBACKS
1.	Machine Learning Approaches for Fake Reviews Detection(2022)	IEEE	• Checked rating behaviour of the product. • Unnecessary bad or good reviews were eliminated	Relatively slower compared to LSA
2.	Exploring E-Commerce Product Experience Based on Fusion Sentiment Analysis Method (2022)	IEEE	• Uses NLP to analyse the opinion mined • Analysis of semantic sentiment	Lacks the tracking of redundant review

SI. NO	TITLE AND AUTHOR, YEAR	PUBLISHER & JOURNAL NAME	METHODOLOGY USED	DRAWBACKS
3.	IP spam detection using Machine Learning for Data Analytics	IEEE	• Tracks IP	Doesn't uses a standard ML model
4.	Opinion Mining Using Multi-Dimensional Analysis (2023)	IEEE	• Classify the opinion expression with ML and NLP • Recognition of emotion	Resource intensive NLP processes

OBJECTIVES

- Develop a Deceptive feedback identification system for online e-commerce platforms.
- Address the increasing impact of product reviews on consumer purchasing decisions.
- Create a technology-driven solution to automatically identify and filter out fake or misleading reviews.
- Ensure the authenticity of customer feedback, contributing to a fair online shopping environment.
- Contribute to the overall integrity of online marketplaces and e-commerce platforms.

ABSTRACT

"Deceptive Product Feedback Identification with ML" tackles the challenge of identifying fake reviews in online platforms. Leveraging machine learning, specifically Latent Semantic Analysis (LSA) to detect reviews, biased user promotions, IP address patterns, review floods, simultaneous similar reviews, and LSA for meaningful analysis. This technical solution ensures automated and robust deception detection, safeguarding the credibility of online product reviews.

INTRODUCTION

- Reviews on online websites play a vital role in sales of the product as before buying people try to get all the pros and cons of the product.
- The scope and need of online markets and e-commerce platforms are on the rise.
- The amount of feedbacks for products are present in detail for users to analyze the product they are buying.
- Users sometime bombard the review section with extreme comments which can work in favor or against the product.
- This project aims to take care of factitious review

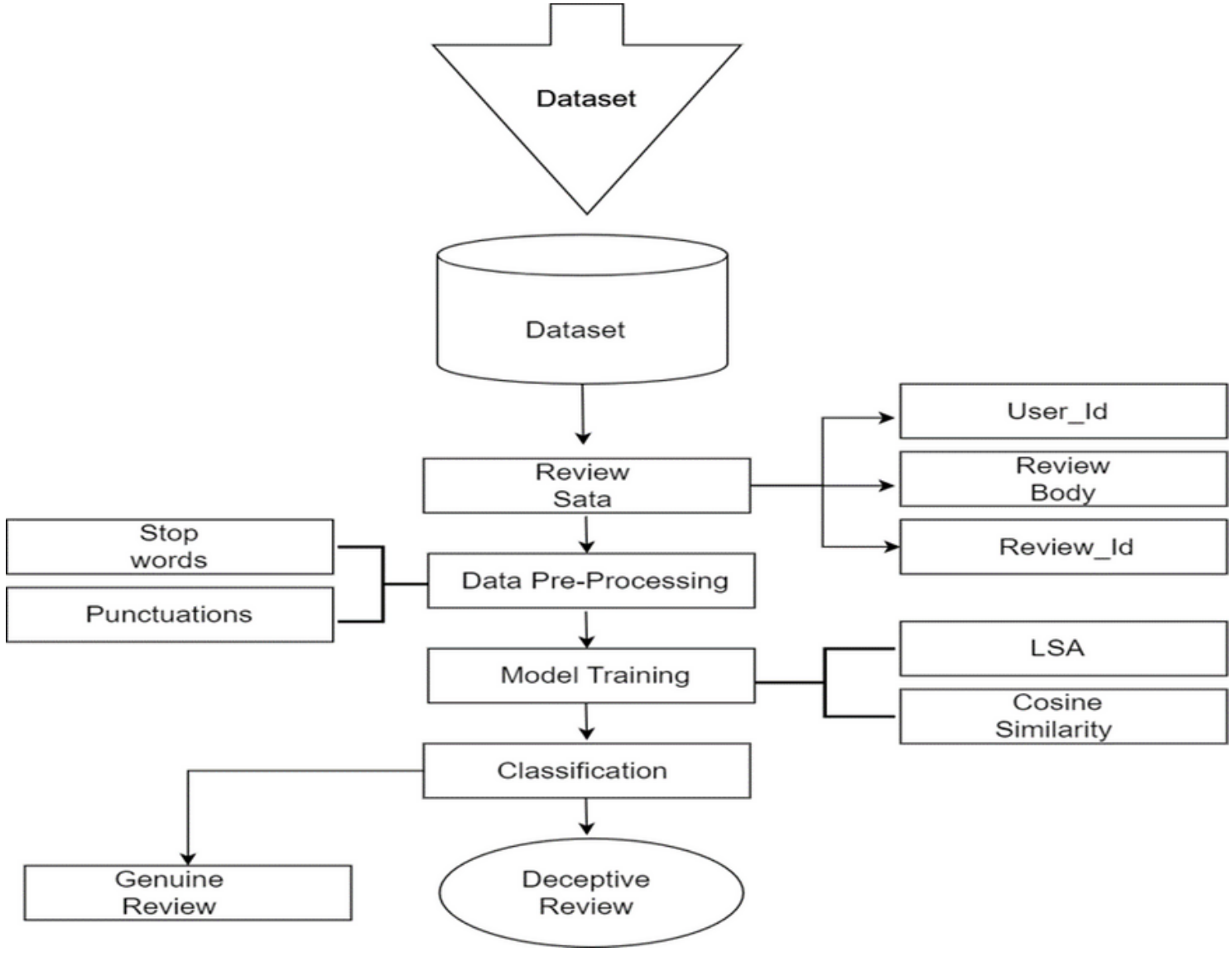
EXISTING SYSTEM

- User are not able to find out whether the review is genuine or fake. If the social media optimization team uses different IP address to send the same review, system fail to track the fake review.
- Brands can use their resources to wrongly increase the rating of their particular products. Same user can write multiple reviews from different accounts.
- The classification of user review methodologies is lesser and the model is slower and resource intensive

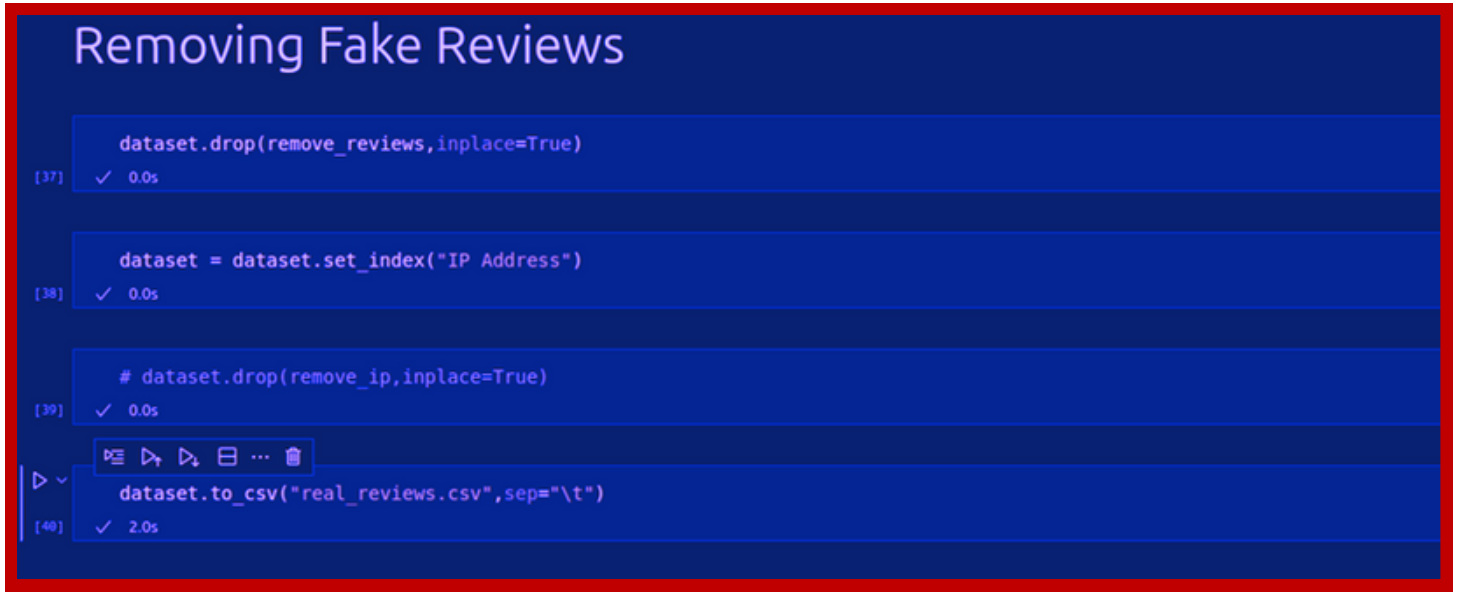
PROPOSED SYSTEM

- Reviews by customers on a product will act as our data on which methods will be applied.
- Utilizes LSA, Cosine Similarity and Sentiment analysis to process and categorize the reviews.
- For filtering the fake reviews from genuine one mining method will be carried out.
- Apply analysis algorithm to make insightful data analysis with methods natural language processing and Logistics Regression

PROPOSED SYSTEM-BLOCK DIAGRAM



EXPECTED OUTPUT



	marketpla	customer_	review_id	product_ic	product_p	product_ti	product_c	star_rating	helpful_vo	total_vote	vine	verified_p	review_he	review_body
2	US	20422322	R8MEA6IG	B00MC4CI	82850235	BlackVue [	Mobile_El	5	0	0	N	Y	Very Happ	As advertised. Everything works perfectly, I'm very h
3	US	40835037	R31LOQ8J	B00OQMF	82850235	GENSSI GS	Mobile_El	5	0	1	N	Y	five star	it's great
4	US	51469641	R2Y0MM9	B00QERR5	82850235	iXCC Multi	Mobile_El	5	0	0	N	Y	great cabl	These work great and fit my life proof case for the i
5	US	4332923	RRB9C05H	B00QUFTF	82850235	abcGoode	Mobile_El	4	0	0	N	Y	Work very	Work very well
6	US	44855305	R26I2RI1G	B0067XVN	5.63E+08	Generic Cá	Mobile_El	2	0	0	N	Y	Cameras h	Be careful with these products, I have bought severa
7	US	7846966	RY8DDL22	B00KA6CC	5.63E+08	Aweek® Ai	Mobile_El	3	0	1	N	Y	It appears	It appears to be good, but I'm still waiting for my oth
8	US	21299354	R2AT2426	B00MJCD6	5.63E+08	Sentey LS-	Mobile_El	3	0	1	N	Y	Didn't love	First one arrived as a brick. Wouldn't work, wouldn'
9	US	20422322	R8MEA6IG	B00MC4CI	5.63E+08	BlackVue [	Mobile_El	5	0	0	N	Y	Very Happ	As advertised. Everything works perfectly, I'm very h
10	US	20422322	R8MEA6IG	B00MC4CI	2.17E+08	BlackVue [	Mobile_El	5	0	0	N	Y	Very Happ	As advertised. Everything works perfectly, I'm very h
11	US	40835037	R31LOQ8J	B00OQMF	1.37E+08	GENSSI GS	Mobile_El	5	0	1	N	Y	five star	it's great
12	US	51469641	R2Y0MM9	B00QERR5	82850235	iXCC Multi	Mobile_El	5	0	0	N	Y	great cabl	These work great and fit my life proof case for the i
13	US	4332923	RRB9C05H	B00QUFTF	2.21E+08	abcGoode	Mobile_El	4	0	0	N	Y	Work very	Work very well but couldn't get used to not hearing
14	US	44855305	R26I2RI1G	B0067XVN	5.63E+08	Generic Cá	Mobile_El	2	0	0	N	Y	Cameras h	Be careful with these products, I have bought severa
15	US	7846966	RY8DDL22	B00KA6CC	7.14E+08	Aweek® Ai	Mobile_El	3	0	1	N	Y	It appears	It appears to be good, but I'm still waiting for my oth

HARDWARE/SOFTWARE USED

HARDWARE REQUIREMENT

- OS: Linux/Windows-10/MAC OS
- Processor : Atleast 4 cores CPU
- Main Memory : 8GB RAM
- Hard Disk : 25GB

SOFTWARE REQUIREMENT

- Python Language(>3.9)
- Jupyter Notebook
- Anaconda Environment
- Logistic Regression Library
- NLTK ,Sklearn , Pickle, Requests , Pandas

APPLICATIONS

- To be intergrated in the backend of an E-commerce Website which could help in categorizing the decptive reviews
- Identifying and addressing deceptive feedback promptly helps uphold brand reputation and integrity in the competitive e-commerce landscape.
- Monitors the reviews of each and every customers and analyzes it.
- Businesses can make more informed decisions regarding product development, marketing strategies, and customer service initiatives based on genuine feedback rather than misleading reviews.

REFERENCES

[1] MACHINE LEARNING APPROACHES FOR FAKE REVIEWS DETECTION: A SYSTEMATIC LITERATURE REVIEW 2022 MOHAMMED ENNAOURI AND AHMED ZELLOU

[2] EXPLORING E-COMMERCE PRODUCT EXPERIENCE BASED ON FUSION SENTIMENT ANALYSIS METHOD 2022 HUAQIAN HE , GUIJUN ZHOU , AND SHUANG ZHAO

[3] EVALUATING ANNOTATED DATASET OF CUSTOMER REVIEWS FOR ASPECT BASED SENTIMENT ANALYSIS 2021 DIMPLE CHEHAL PARUL GUPTA AND PAYAL GULATI

[4] OPINION MINING USING MULTI-DIMENSIONAL ANALYSIS 2023 SATARUPA BISWAS AND G. POORNALATHA

[5] TEXT MINING AND EMOTION CLASSIFICATION ON MONKEYPOX TWITTER DATASET: A DEEP LEARNING-NATURAL LANGUAGE PROCESSING (NLP) APPROACH RUTH OLUSEGUN 1 , TIMOTHY OLADUNNI AND HALIMA AUDU AND YAO HOUKPATI AND STAPHORD BENGESI


Thank
You