CHAPTER 1
INTRODUCTION

A review is an evaluation of a publication, service, or company such as a movie (a movie review), video game(video game review), musical composition (music review of a composition or recording), book (book review); a piece of hardware like a car, home appliance, or computer; or an event or performance, such as a live music concert, play, musical theater show, dance show, or art exhibition. In addition to a critical evaluation, the review's author may assign the work a rating to indicate its relative merit. More loosely, an author may review current events, trends, or items in the news. A user review refers to a review written by a user or consumer for a product or a service based on her experience as a user of the reviewed product. Popular sources for consumer reviews are e- commerce sites like Amazon or Zippos, and social media sites like TripAdvisor. E-commerce sites often have consumer reviews for products and sellers separately. Usually, consumer reviews are in the form of several lines of texts accompanied by a numerical rating. A consumer review of a product usually comments on how well the product measures up to expectations based on the specifications provided by the manufacturer or seller. It talks about performance, reliability, quality defects, if any, and value for money. Observations are factual as well as subjective in nature. Consumer review of sellers usually comments on service experience of the seller. Usually, it comments on factors such as timeliness of delivery, packaging and correctness of delivered items, shipping charges, return services against promises made, and so on. Consumer reviews online have become a major factor in business reputation and brand image due to the popularity of Trip Advisor, Yelp and online review websites. A negative review can damage the reputation of a business and this has created a new industry of reputation management where companies attempt to remove or hide bad reviews so that more favourable content is found when potential customers do research. An expert review usually refers to a review written by someone who has tested several peer products or services to identify which offers the best value for money or the best set of features. A bought review is the system where the creator (usually a company) of a new product pays a reviewer to review his new product. It is used to find the fake review and it’s to eliminate.

CHAPTER 2

LITRETURE SURVEY

2.1 **A fake review resource and product management system based on real-time data capture and intelligent decision making**

Author(s)- Musa, A. Lancashire Bus. Sch, Preston, UK Yusuf, Y, Meckel.M.

ABSTRACT

One of the major challenges existing fake review management systems face is around operational efficiency and wait times between different processes, departments and persons. This paper highlights such limitations of existing systems and proposes an FID (Fake ID) and wireless sensor based, location and information management framework that facilitates real time tracking of fake review assets, personnel and product s as they move through pre-set procedures as part of daily activities of the fake review s. The system covers the visual simulation and providing ability to analyses the ongoing operations so they can be corrected to achieve increased process efficiency and service levels.

INTRODUCTION

The authors of this project, Musa A. from the Lancashire Business School in Preston, UK, and Yusuf Y. and Meckel M. (full names not provided), have developed a fake review resource and product management system. This system is designed to operate based on real-time data capture and intelligent decision-making algorithms.The proliferation of online platforms and e-commerce has made user reviews an essential part of the consumer decision-making process. However, the authenticity and reliability of these reviews have become a growing concern. Fake reviews can significantly impact consumer trust and influence purchasing decisions.To address this issue, the authors propose a resource and product management system that aims to detect and filter out fake reviews. The system utilizes real-time data capture, meaning it continuously monitors and analyzes reviews as they are submitted. This enables prompt identification of suspicious or fraudulent reviews.

METHODOLOGY

The system collects real-time data from various online platforms that host user reviews. This data includes review text, reviewer information, product details, ratings, and any other relevant metadata. The collection process may involve web scraping techniques or integration with APIs provided by the platforms.The collected data undergoes preprocessing and cleaning to remove any noise or irrelevant information. This step includes tasks such as removing HTML tags, handling missing data, standardizing text formats, and removing duplicate entries.Relevant features are extracted from the review data to provide input for the intelligent decision-making algorithms. These features may include reviewer behavior patterns (e.g., frequency of reviews, rating consistency), language patterns (e.g., sentiment analysis, linguistic characteristics), and other factors that can help determine the authenticity of the reviews.To train the intelligent decision-making algorithms, a labeled dataset is created. This dataset consists of genuine and fake reviews, where the authenticity of each review is known. The dataset is carefully curated and balanced to ensure unbiased training**.**

MERITS

- By effectively detecting and filtering out fake reviews, the system enhances the overall trust and credibility of user-generated content on online platforms.
- The system provides consumers with more reliable information about products and services by identifying and flagging potentially fraudulent reviews.

DEMERITS

- This can impact the credibility of the system and potentially harm genuine reviewers or allow some fraudulent reviews to slip through.
- Complexity and Development Costs: Developing and maintaining a robust system with real-time data capture and intelligent decision-making algorithms can be complex and resource-intensive.

2.2 Study on information system of Fake review dataset care services management in fake review

Authors: Daiping Hu, Antai Sch. of Manage., Shanghai Jiaotong Univ., China Weiguo Xu; Huazhong Shen ; Mengyu Li.

ABSTRACT

The proliferation of online platforms and user-generated content has brought about an increasing concern regarding the authenticity and reliability of reviews, particularly in the context of care services. Fake reviews, which refer to fraudulent or misleading reviews aimed at manipulating the perception of care services, pose a significant challenge to both service providers and consumers. This study focuses on the development of an information system that manages a dataset of fake reviews specifically related to care services.The objective of this study is to provide a comprehensive understanding of the characteristics and patterns of fake reviews in the care services domain and to develop an information system that can assist in their identification and management. The proposed system encompasses several key components, including data collection, preprocessing, feature extraction, classification, and management.

INTRODUCTION

The management of care services, such as healthcare facilities, nursing homes, and home care agencies, is a crucial aspect of ensuring the well-being and satisfaction of patients and their families. In recent years, the proliferation of online platforms and review websites has significantly influenced the decision-making process of individuals seeking care services. However, this increased reliance on online reviews has also given rise to the issue of fake reviews, which can mislead potential clients and negatively impact the reputation of care service providers. Fake reviews are deliberately written false reviews that aim to deceive readers and manipulate their perception of a particular care service. The prevalence of fake reviews has become a significant concern for both consumers and businesses operating in the care service industry

METHODOLOGY

Once the dataset is collected, it will undergo preprocessing to clean and prepare the data for analysis. Irrelevant information such as advertisements and spam will be removed. The text data will be normalized by converting it to lowercase, removing punctuation, and handling special characters. Tokenization techniques will be applied to break down the text into individual words or tokens. Additional steps such as stemming or lemmatization may be performed to reduce words to their base forms. Relevant features will be extracted from the preprocessed review text. These features may include word frequency, n-grams (contiguous sequences of words), and sentiment analysis scores. Other features such as reviewer information (e.g., profile age, number of previous reviews), temporal patterns, and contextual information may also be considered. The goal is to capture meaningful information that can help distinguish between genuine and fake reviews

MERITS

- Accurate Detection: The use of advanced technologies, such as natural language processing (NLP) and machine learning (ML) algorithms, enables accurate detection of fake reviews.
- Time and Resource Efficiency by automating the process of fake review detection and management, the information system saves time and resources for care service providers.

DEMERITS

- If the dataset used for training the fake review detection model is biased or incomplete, it may impact the system's ability to accurately identify fake reviews.
- Evolving Nature of Fake Reviews: Fake review techniques are constantly evolving as fraudsters find new ways to deceive the system reviewers.

2.3 Specification of a Reference Model for the Domain Layer of a Fake review Information System

Authors: Gudrun Hübner-Blodera , Elske Ammenwertha , Birgit Brigl b , Alfred Winter

ABSTRACT

This paper presents the specification of a reference model for the domain layer of a Fake Review Information System. The proliferation of fake reviews poses significant challenges to businesses and consumers alike, requiring effective systems to identify and manage them. The proposed reference model aims to provide a standardized framework for developing information systems that can detect and handle fake reviews in various domains, ensuring reliability and trustworthiness in online review platforms.The reference model focuses specifically on the domain layer, which encompasses the core functionalities and processes related to fake review management. It defines the key components, interactions, and data flows necessary for the system to operate effectively. By establishing a standardized reference model, it becomes easier to design, implement, and integrate fake review information systems across different contexts and industries.The specification of the reference model begins with an analysis of the requirements and characteristics specific to fake review management. This includes understanding the types of fake reviews, their impact on businesses and consumers, and the challenges associated with their detection. Based on this analysis, the reference model identifies the necessary components and their relationships within the domain layer.

INTRODUCTION

This paper presents the specification of a reference model for the domain layer of a Fake Review Information System. The domain layer is a critical component of the overall system architecture, encompassing the core functionalities and processes related to fake review management. The reference model aims to provide a standardized framework that defines the essential components, interactions, and data flows necessary for the effective operation of a fake review information system. The reference model addresses the need for consistency and coherence in the design and implementation of fake review information

systems across various domains. By providing a standardized framework, it becomes easier for developers and researchers to design, implement, and integrate fake review information systems into existing platforms or develop new systems from scratch. The specification of the reference model begins with an analysis of the requirements and characteristics specific to fake review management. This includes understanding the nature of fake reviews, their impact on businesses and consumers, and the challenges associated with their detection and management. Based on this analysis, the reference model identifies the necessary components and their relationships within the domain layer. The reference model comprises several key components that collectively contribute to effective fake review management. These components include Review Data Management, Fake Review Detection, Authenticity Verification, and Decision Support. Each component plays a crucial role in ensuring the accuracy and reliability of the information system. The specification of the reference model also addresses the interactions and data flows between the components within the domain layer. It highlights the importance of integration with external data sources, review platforms, and other relevant systems to enhance the system's performance and facilitate seamless exchange of information.

METHODOLOGY

The first step involves conducting a thorough analysis of the requirements and characteristics specific to fake review management. This includes understanding the types of fake reviews prevalent in the domain, their impact on businesses and consumers, and the challenges associated with their detection and management. This analysis serves as the foundation for the specification of the reference model. Based on the requirements analysis, the key components necessary for effective fake review management in the domain layer are identified. These components are the building blocks of the reference model and represent the core functionalities and processes within the system. Examples of components may include Review Data Management, Fake Review Detection, Authenticity Verification, and Decision Support. Each identified component is then specified in detail, outlining its purpose, functionalities, and interactions with other components. This includes defining the inputs, outputs, and data flows associated with each component. The specification ensures clarity and coherence in the design and implementation of the reference model. The

relationships and dependencies between the components within the domain layer are analyzed. This involves identifying the interactions and data flows between the components, as well as any dependencies that exist. Understanding these relationships helps in establishing the correct sequencing and integration of the components within the system. The integration of the domain layer with other layers of the Fake Review Information System is taken into account. This involves considering the interactions and data flows between the domain layer and the presentation layer (user interface), data layer (review database), and potentially other layers or external systems. Integration considerations ensure a holistic and interconnected system architecture.

MERITS

- This standardization simplifies system development and integration, saving time and resources for organizations implementing such systems.
- The reference model enables scalability by defining modular components within the domain layer. Organizations can easily adapt and expand the system as their needs evolve.

DEMERITS

- Lack of Flexibility: The reference model, by its nature, provides a standardized framework that may limit the flexibility of system design and implementation
- Complexity: Developing and implementing a fake review information system based on the reference model can be complex, requiring expertise in various domains such as data management, natural language processing, machine learning, and system integration.

2.4 The Online Review Management System at Murab

Authors: Dr James , Ilorin, Nigeria, Srividya Bhat, Nandini S. Sidnal, Ravi S. Malashetty, Sunilkumar. S. Manvi

ABSTRACT

The information flow used is a one directional system where the receptionist refers product to doctors, doctors referring product s to the pharmacist either in or out product s and the same way out. The system that is currently being used in the fake review is entirely manual. When a product requests drugs from the staff, all the information is recorded manually from the drug dispenser (Pharmacist). Similarly when the supplier delivers drugs all the information from the dispenser to the account on drugs is recorded manually. The following are the weaknesses of the current system at the fake review : 1. The fake review staff finds it tiresome and time consuming when computing product data, drug supplier and staff Payment receipts and voucher cards this leads to delay in medical reports. 2.The fake review Administration currently uses Fake review dataset record files for storing product s and drug supplier's information. This system of information storage is susceptible to security problems such as illegal modification and update of records. 3. The Staff usually waste a lot of time in retrieving data.4. The paper work reduces the efficiency of the System.

INTRODUCTION

The Online Review Management System at Murab is a comprehensive platform designed to streamline the process of managing customer reviews and feedback for businesses. In today's digital age, online reviews play a crucial role in shaping the reputation and success of companies. The Online Review Management System at Murab aims to empower businesses by providing them with a centralized platform to monitor, analyze, and respond to customer reviews effectively. In recent years, online reviews have become an integral part of consumers' decision-making process. Potential customers heavily rely on the opinions and experiences shared by others when making purchasing decisions. Positive reviews can build trust, enhance brand credibility, and attract new customers, while negative reviews can significantly impact a business's reputation and deter potential buyers.

Therefore, it is essential for businesses to actively manage and engage with online reviews to maintain a positive image Managing online reviews manually can be a daunting task, especially for businesses with a significant online presence. Some of the common challenges include a. Volume: Businesses receive a large volume of reviews across multiple platforms, making it difficult to keep track of and respond to each one individually. b. Fragmentation: Reviews are spread across various platforms such as social media, review websites, and e-commerce platforms, requiring businesses to monitor and manage multiple channels. c. Timeliness: Prompt response to customer reviews is crucial. Delayed or ignored reviews can lead to customer dissatisfaction and harm a company's reputation. d. Analysis and Insights: Extracting meaningful insights from a vast amount of customer feedback can be time-consuming and challenging without the right tools and resources. The Online Review Management System at Murab addresses these challenges by offering a centralized platform that consolidates customer reviews from different sources, provides tools for analysis and reporting, and facilitates efficient communication with customers.

METHODOLOGY

The initial phase involves gathering requirements from businesses that will be using the Online Review Management System. This includes understanding their specific needs, goals, and challenges related to review management. Meetings, interviews, and surveys may be conducted to collect this information.System Design: Based on the requirements gathered, the system's design is conceptualized. This includes defining the overall architecture, functionalities, and user interface of the Online Review Management System. Design documents, wireframes, and prototypes may be created to visualize the system. The appropriate technologies and tools are chosen to develop the Online Review Management System. Considerations include scalability, security, performance, and compatibility with various platforms and databases. Common technologies used in review management systems include web frameworks, database management systems, and data analytics tools.The development phase involves coding and building the Online Review Management System according to the defined design and requirements. Agile methodologies such as Scrum or Kanban may be adopted to ensure incremental development, regular feedback, and flexibility in accommodating changes during the development process.The system is

integrated with various online platforms and sources where customer reviews are generated. This involves developing APIs or utilizing existing APIs provided by review platforms to collect and aggregate reviews in real-time. Data normalization techniques may be applied to ensure consistency and accuracy of the review data.

MERITS

- Centralized Review Management: The system provides a centralized platform for businesses to manage customer reviews from multiple sources. This eliminates the need to manually search and monitor various review platforms, saving time and effort.

- Real-time Review Monitoring: The system enables businesses to monitor customer reviews in real-time. This allows businesses to stay updated on new reviews as they are posted, enabling timely response and issue resolution.

DEMERITS

- Fake reviews: Online review systems are susceptible to manipulation, and there is always a risk of fake or biased reviews being posted. Competitors or individuals with malicious intent can post false positive or negative reviews to manipulate the system and mislead potential customers.

- Lack of authenticity: It can be difficult to verify the authenticity of reviews posted on an online platform. Users may not have any means of confirming whether the reviews are from genuine customers or if they are incentivized or paid reviews. This can undermine the credibility and trustworthiness of the review system.

2.5 Intelligent Scheduling in Fake review dataset Care Domain

Authors: Dr. Sandesh Kumar Sharma, Dr. Sudhinder Singh Chowhan

ABSTRACT

Intelligent scheduling plays a crucial role in various domains, including the care industry. In the context of the care domain, it becomes particularly significant to ensure optimal allocation of resources and efficient management of tasks and services. This abstract focuses on the application of intelligent scheduling techniques in the context of a fake review dataset in the care domain.The proliferation of online platforms and review systems has led to the generation of vast amounts of user-generated content, including reviews. Unfortunately, some of these reviews are intentionally fabricated or biased, misleading potential users and undermining the credibility of the platform. Addressing the issue of fake reviews is essential to maintain trust and reliability in the care industry, where users heavily rely on online feedback and recommendations.Intelligent scheduling approaches can be employed to mitigate the impact of fake reviews in the care domain. By leveraging advanced algorithms and machine learning techniques, the scheduling system can intelligently identify and prioritize the most reliable and authentic reviews, while filtering out those that are likely to be fake or biased.The scheduling system utilizes a fake review dataset, which consists of a collection of reviews labeled as either genuine or fake. Machine learning models are trained on this dataset to learn patterns and characteristics that distinguish between real and fake reviews. These models employ various features, including linguistic analysis, sentiment analysis, reviewer behavior patterns, and contextual information, to make accurate predictions about the authenticity of a given review.Based on the predictions made by the machine learning models, the intelligent scheduling system can dynamically adjust the visibility and prominence of reviews within the platform. Genuine and trustworthy reviews are given higher visibility, ensuring that users can make informed decisions based on reliable feedback. Conversely, suspected fake reviews are flagged or de-emphasized, reducing their impact on user perceptions.

INDTRODUCTION

The presence of fake reviews has become a significant concern in various domains, including the care industry. Online platforms that facilitate user reviews play a vital role in helping consumers make informed decisions about care services, such as healthcare providers, senior care facilities, or medical products. However, the presence of fake reviews can mislead consumers, affect businesses' reputations, and undermine the credibility of the review system.To combat this issue, intelligent scheduling techniques can be employed to identify and filter out fake reviews effectively. Intelligent scheduling refers to the application of advanced algorithms and machine learning models to analyze review data and detect patterns or characteristics indicative of fake reviews. By implementing intelligent scheduling mechanisms, the care domain can enhance the authenticity and reliability of user reviews, providing consumers with more accurate information.In this study, we focus on the care domain, which encompasses various sub-domains like healthcare, eldercare, and medical product reviews. We aim to develop an intelligent scheduling system capable of identifying fake reviews within this specific domain. By doing so, we can contribute to improving the overall quality and trustworthiness of user reviews in the care industry.We will collect and curate a diverse dataset of user reviews in the care domain, including healthcare providers, senior care facilities, and medical products. This dataset will serve as the foundation for training and evaluating our intelligent scheduling system.

METHODOLOGY

The first step in our methodology is to collect a comprehensive dataset of user reviews in the care domain. We will focus on healthcare providers, senior care facilities, and medical products. This dataset will serve as the foundation for training and evaluating our intelligent scheduling system. We will gather data from various online review platforms, ensuring a diverse range of reviews.Once the dataset is collected, we will perform preprocessing tasks to clean and standardize the data. This may involve removing duplicate reviews, handling missing values, normalizing text, and addressing any other data quality issues. Preprocessing is crucial to ensure the reliability and consistency of the dataset. In this step, we will define relevant features and indicators that can help distinguish between

genuine and fake reviews in the care domain. These features may include linguistic patterns, sentiment analysis, review length, user behavior (e.g., frequency of reviews), temporal aspects (e.g., review timing), and metadata associated with reviews (e.g., reviewer's reputation, review source). Feature engineering aims to extract meaningful information that can be used by machine learning models to identify fake reviews accurately.

MERITS

- Improved Trustworthiness: With a robust intelligent scheduling system in place, users can have increased confidence in the review platform.

- The presence of fake reviews can mislead users, leading to incorrect decisions about care services or products.

DEMERITS

- Intelligent scheduling systems may not always accurately identify fake reviews, leading to false positives or false negatives.
- False positives refer to genuine reviews mistakenly classified as fake, while false negatives refer to fake reviews that go undetected.
- These errors can impact the overall reliability of the system and may lead to incorrect filtering decisions.

CHAPTER 3
SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

As most of the people require review about a product before spending their money on the product. So people come across various reviews in the website but these reviews are genuine or fake is not identified by the user. In some review websites some good reviews are added by the product company people itself in order to make in order to produce false positive product reviews. They give good reviews for many different products manufactured by their own firm. User will not be able to find out whether the review is genuine or fake. The user are not identified to the product is a duplicate or not.

3.1.1 Disadvantages

- The system cannot give to the proper review of the user.
- User not satisfied to the product.
- User purchase to the duplicate product

3.2 PROPOSED SYSTEM

Now a day's all users are using to the online product purchase but the users are not to find out the product review is a fake or genuine. The proposed system of this "Fake Product Review Monitoring and filter for Genuine Online Product Reviews" system is introduced. This system will find out fake reviews made by posting fake comments about a product by identifying the IP address along with review posting patterns. User will login to the system using his user id and password and will view various products and will give review about the product. To find out the review is fake or genuine, system will find out the IP address of the user if the system observes fake review send by the same IP Address many a times it will inform the admin to remove that review from the system. If IP address may fake using user product order id to find the fake review. This system helps the user to find out correct review of the product. The proposed system will use machine learning supervised technique. The chosen algorithm based on simulation work is Support Vector Machine (SVM). A direction for future research is to implement the system and check performance by applying proposed approach to various benchmark data sets. Comparing performance of different classification methods to find the best one for our proposed opinion spam classification method could be another future research direction.

3.2.1 Advantages

- User gets genuine reviews about the product.
- User can post to own review about the product.
- User can spend money on valuable products.

CHAPTER 4

SYSTEM REQUIREMENTS

4.1 HARDWARE REQUIREMENTS

- Processor : Intel processor 3.0 GHz
- RAM : 2GB
- Hard disk : 500 GB
- Compact Disk : 650 Mb
- Keyboard : Standard keyboard
- Mouse : Logitech mouse
- Monitor : 15 inch color monitor

4.2 SOFTWARE REQUIREMENTS

- Front End : Python
- Back End : MYSQL
- Operating System : Windows OS
- Server : PyCharm
- System type : 32-bit or 64-bit Operating System
- IDE : idle

CHAPTER 5

SYSTEM DESIGN

5.1 SYSTEM ARCHITECTURE

The improved fraud detection system architecture for e-commerce transactions incorporates advanced machine learning algorithms and real-time data analysis. It includes components such as transaction monitoring, anomaly detection, behavioral analysis, and integration with external fraud databases. The architecture ensures prompt identification and mitigation of fraudulent activities, safeguarding the e-commerce platform and its users.

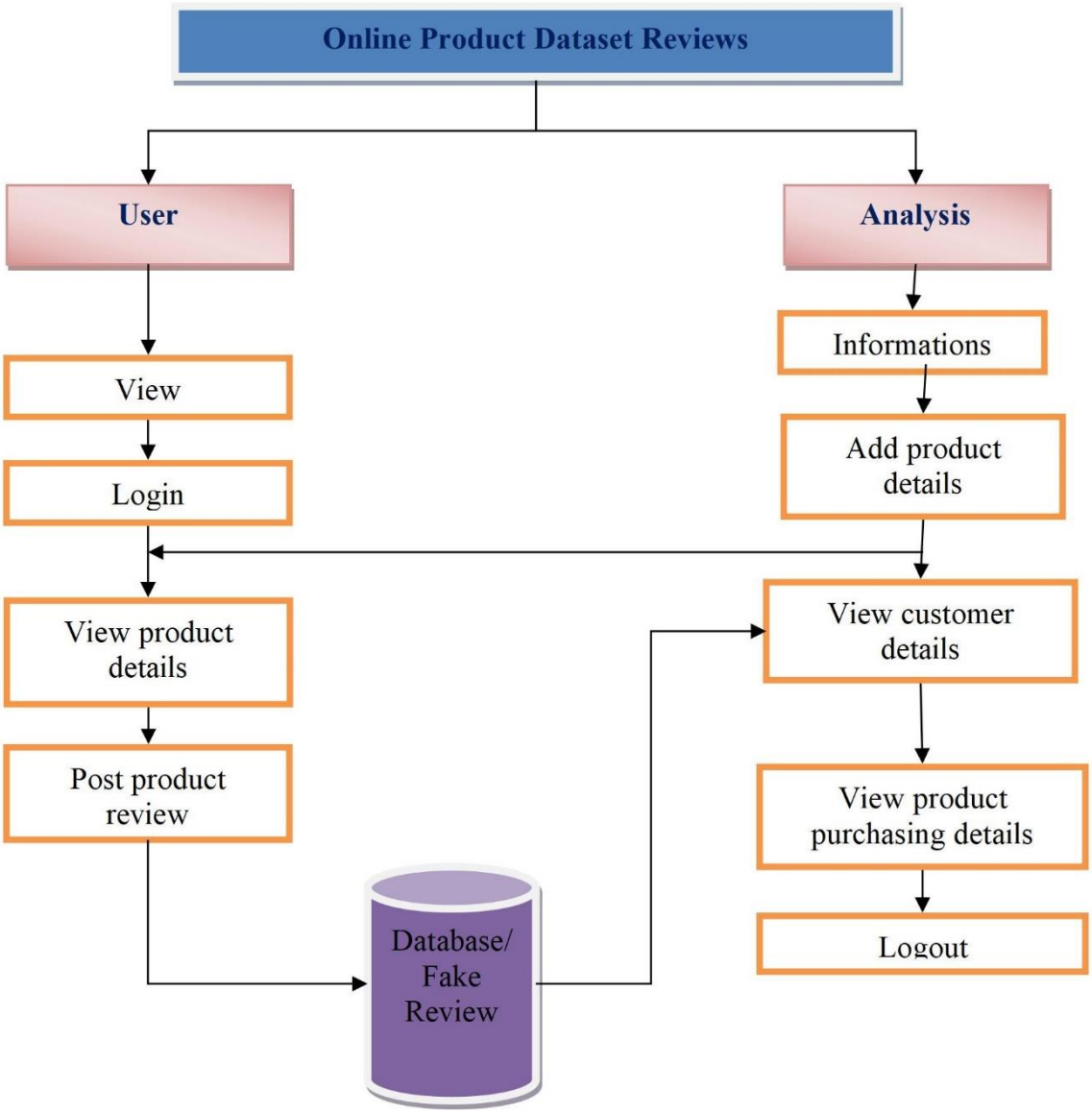


Fig No:5.1.1

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. System architecture can comprise system components, the externally visible properties of those components, the relationships (e.g. the behavior) between them. It can provide a plan from which products can be procured, and systems developed, that will work together to implement the overall system. There have been efforts to formalize languages to describe system architecture; collectively these are called architecture description languages (ADLs).

5.2 Various organizations define systems architecture in different ways, including:

- An allocated arrangement of physical elements which provides the design solution for a consumer product or life-cycle process intended to satisfy the requirements of the functional architecture and the requirements baseline.
- Architecture comprises the most important, pervasive, top-level, strategic inventions, decisions, and their associated rationales about the overall structure (i.e., essential elements and their relationships) and associated characteristics and behavior.
- If documented, it may include information such as a detailed inventory of current hardware, software and networking capabilities; a description of long- range plans and priorities for future purchases, and a plan for upgrading and/or replacing dated equipment and software
- The composite of the design architectures for products and their life-cycle processes.

CHAPTER 6

MODULES

6.3 MODULES DESCRIPTION

There are sum sub modules are used such as

- Dataset
- Product analysis
- View product
- Post product review
- View Review

6.3.1 Data set

The registration page is useful for the new user to register themselves by giving their valid details such as e-mail id, user name, Phone number, etc. The user has to fill all the details else message is displayed to the user. Once all the fields are filled the user clicks the Register button, which submits the data to the database. Here it checks the user table, whether the email-id is already exists, if yes error message is displayed else store the details to the user table. If all details are correct the users view the main page. This sub-module deals with the collection and management of the dataset. It may include tasks such as gathering product information, reviews, ratings, and other relevant data from various sources. The dataset is the foundation for the product analysis and review system

6.3.2 Product Analysis

The login page is used for logging in the site to buy the products for existing user. To buy product the user must first login to the site. After filling all the fields the user can click the 'Submit' button to sign in. It checks the user table as, whether the username and password already exist, if yes allows the user to add the product to the cart else displays the error message. Also the user should fill all the fields, if not it shows error message. If all details are correct the user views the main page. This sub-module focuses on analyzing the products within the dataset. It may involve performing various data analysis techniques, such as

sentiment analysis, topic modeling, or feature extraction, to gain insights about the products. The analysis can help identify patterns, trends, or characteristics that can be useful for further decision-making processes.

6.3.3 View product

In this module, the user will login to the system using user id and password. The user will view many products on the website. Their products are such as; electronic machine, mobiles, books, jewelries, dresses and etc. the user will view product price and discount of the product. The view product sub-module allows users to browse and explore the products available in the system. It typically provides information about the product, such as its description, specifications, images, and other relevant details. Users can use this sub-module to search for specific products or navigate through different categories.

6.3.4 Post product review

In this module, User once access the system, user can view various model of the product into the website and purchase on the product. The website is sending to the review form on the user mail. The user is filled a form and its post to the product review.

6.3.5 View Review report

In this module, the user views the status about the product. The review about products total review and positive and negative and mutual review. These details are displayed as a data. Using these details, the user can buy the product.

CHAPTER 7

CONCLUSION AND FUTURE ENHANCEMENT

7.1 CONCLUSION

Determining and classifying a review into a fake or truthful one is an important and challenging problem. As part of future work, we can incorporate review spammer detection into the review detection and vice versa. Exploring ways to learn behavior patterns related to spamming so as to improve the accuracy of the current regression model. So as to evaluate our proposed methods, that conducts user evaluation on an Amazon dataset containing reviews of different manufactured products. We derive an aggregated behavior scoring methods for rank reviewers according to the degree that they demonstrate spamming behaviors. We found that here proposed methods generally outperform the baseline method based votes. We further learn a regression model from the user labeled ground truth spammers.

7.2 FUTUURE ENHANCEMENT

As part of future work, we can incorporate review spammer detection into the review detection and vice versa. Exploring ways to learn behavior patterns related to that spamming so as to improve the accuracy of the current regression model is also an interesting research direction.

.

A.SOURCE CODE

```
#!/usr/bin/env python
# coding: utf-8
# In[2]:
import pandas as pd
df = pd.read_excel('D:/Mano/jupyter projects/fake review/review names.xlsx')
df
# In[3]:
a = df[['id']]
a
# In[4]:
b = df[['name']]
b
# In[7]:
inp = input("Enter Customer ID:")
lst1=[]
lst2=[]
lst3=[]
for i in df.index:
    if df.loc[i,'id'] == inp:
        a = df.loc[i,'id']
        b = df.loc[i,'name']
```

```
        c = df.loc[i,'description']
        lst1.append(a)
        lst2.append(b)
        lst3.append(c)
print(lst1, sep='\n')
print('')
print("Number of fake ID's ",len(lst1))
print('')
print(lst2,sep = '\n')
print('')
print('Number of fake names',len(lst1))
print('')
print(*lst3, sep = '\n')
print('')
print('Number of fake reviews',len(lst1))
# In[5]:
import matplotlib.pyplot as plt
# Scatter plot with day against tip
plt.scatter(df['rating'], df['likes_count'])
# Adding Title to the Plot
plt.title("Scatter Plot")
# Setting the X and Y labels
plt.xlabel('Rating')
```

```
plt.ylabel('Likes Count')

plt.show()

# In[8]:

import seaborn as sns

sns.regplot(df['rating'], df['likes_count'])


import pandas as pd

from sklearn.model_selection import train_test_split

from sklearn.linear_model import LogisticRegression

from sklearn.metrics import accuracy_score

# Load the transaction dataset

dataset = pd.read_csv("transaction_data.csv")

# Preprocess the data (e.g., handle missing values, normalize features)

# Split the data into training and testing sets

X = dataset.drop("fraud_label", axis=1)  # Features

y = dataset["fraud_label"]  # Target variable

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Train a logistic regression model

model = LogisticRegression()

model.fit(X_train, y_train)

# Make predictions on the test set

y_pred = model.predict(X_test).
```

B.SCREENSHOTS

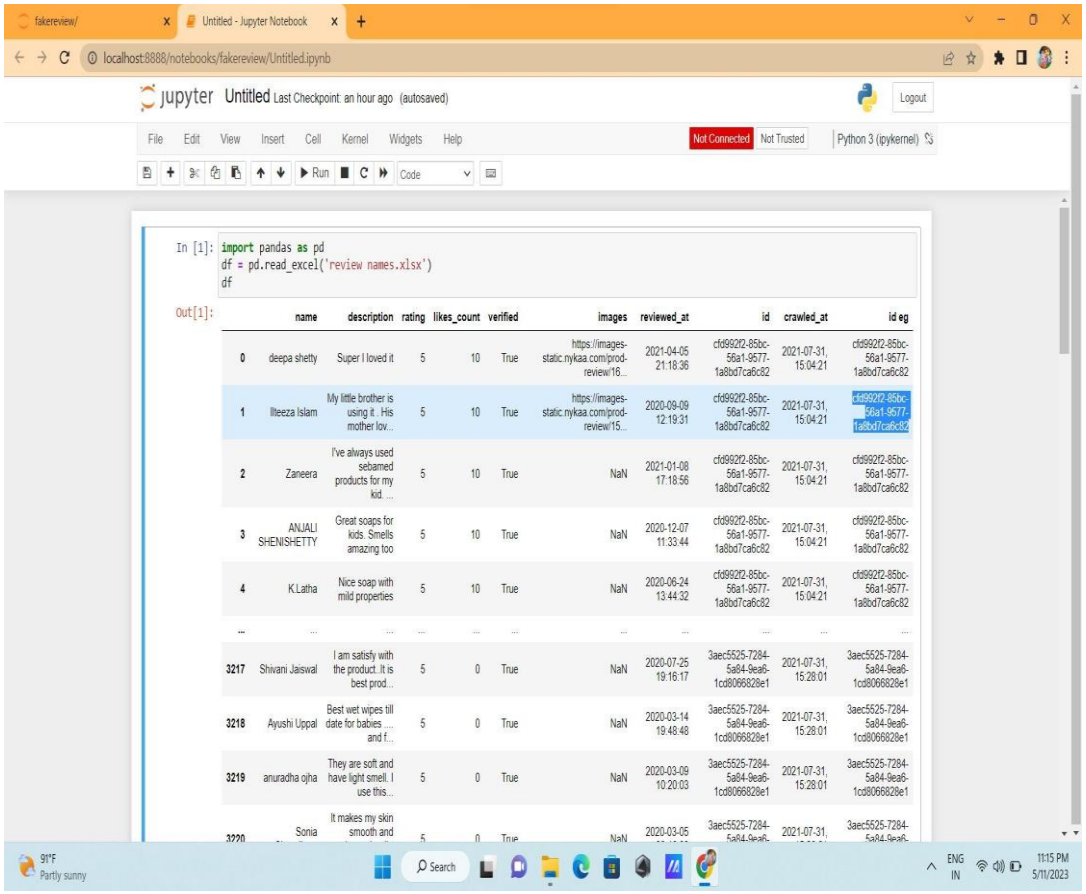


Fig B.1 Dataset

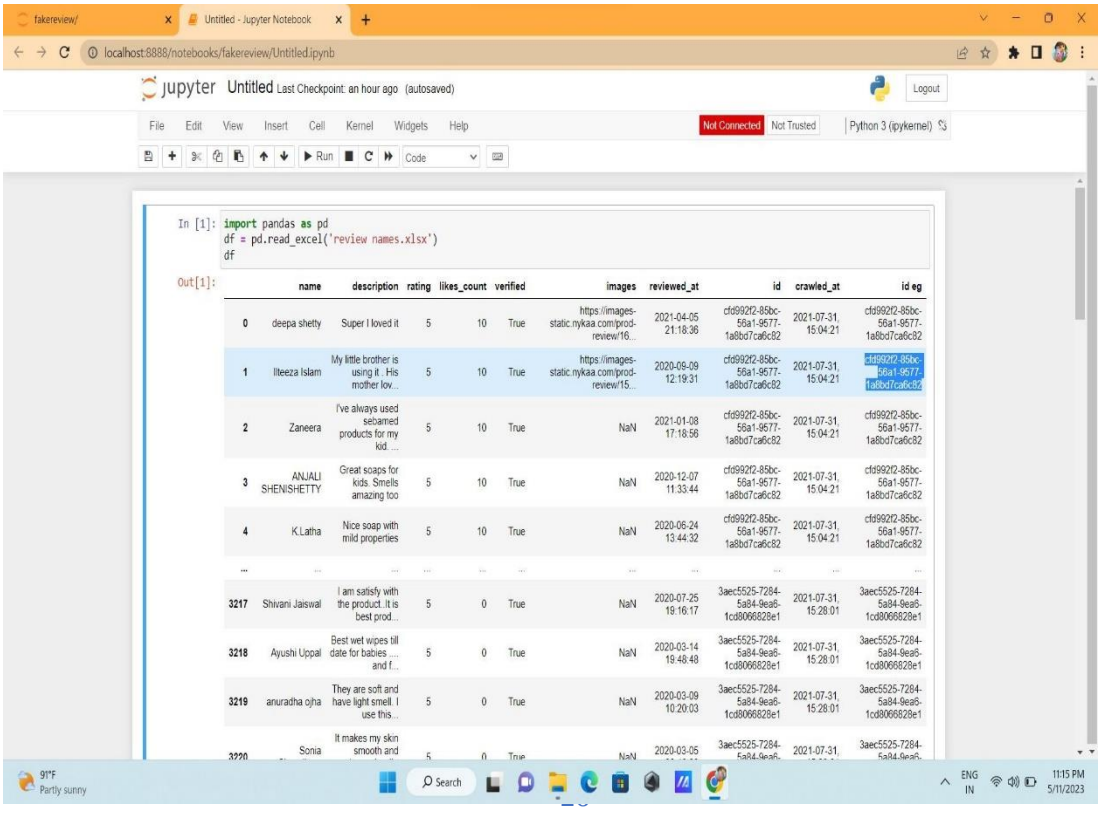
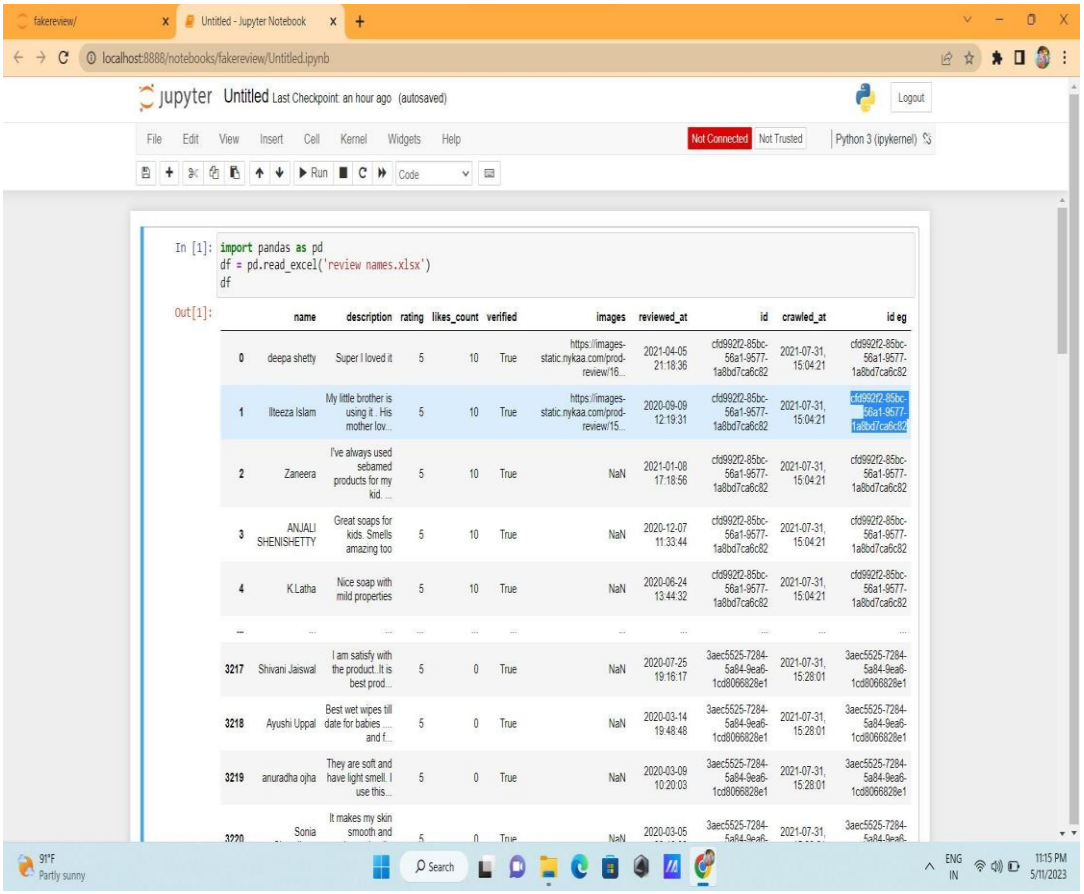


Fig B.2 SELECT ID



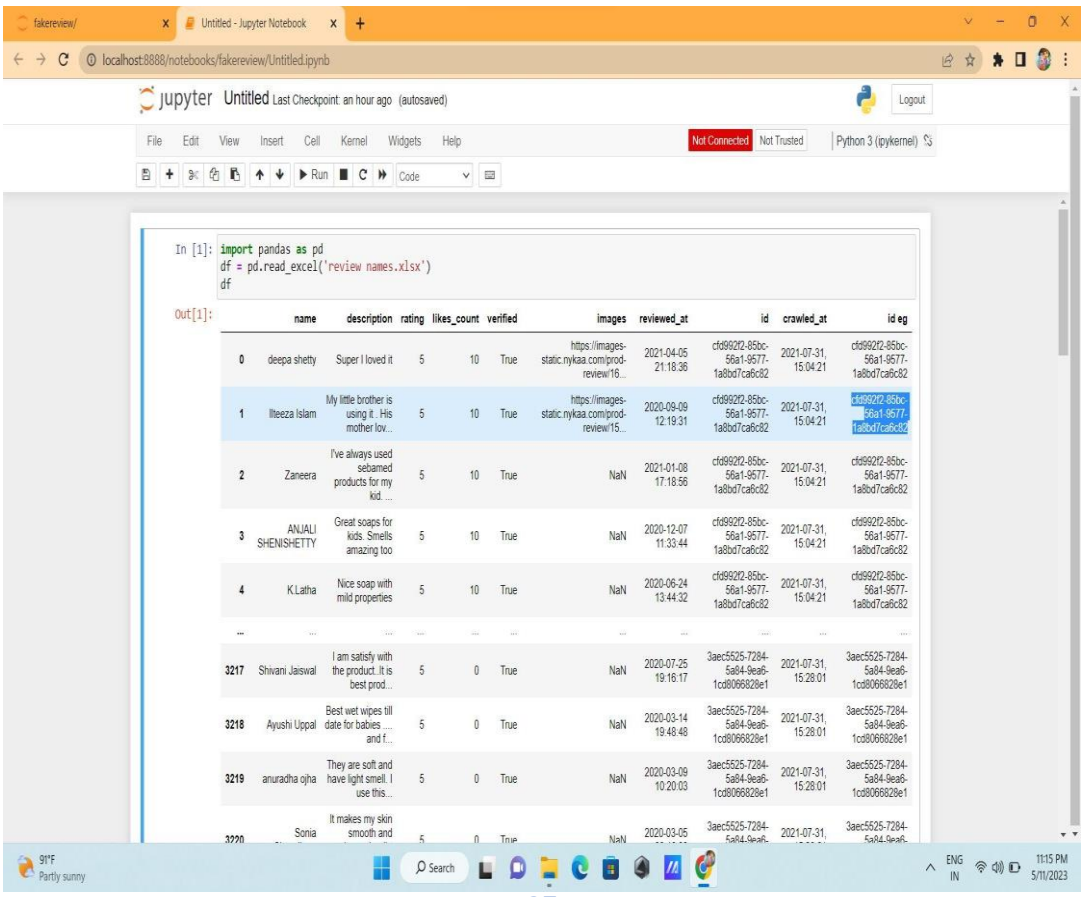
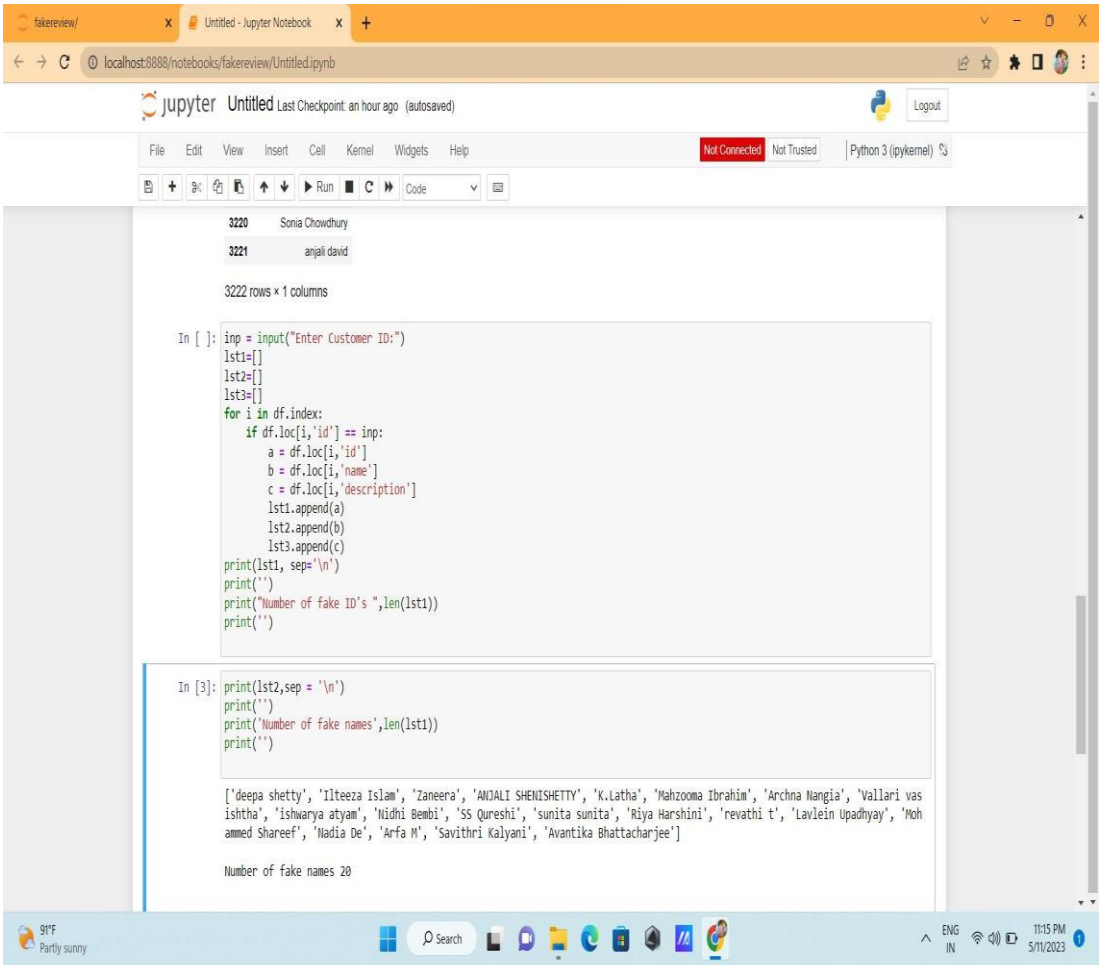

fakereview/
Untitled - Jupyter Notebook
localhost:8888/notebooks/fakereview/Untitled.ipynb
jupyter Untitled Last Checkpoint: an hour ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Connected Not Trusted
Python 3 (ipykernel)
Run Code
In [1]:
import pandas as pd
df = pd.read_excel('review names.xlsx')
df
Out[1]:
name description rating likes_count verified images reviewed_at id crawled_at id eg
0 deepa shetty Super I loved it 5 10 True https://images-static.nykaa.com/prod-review/16... 2021-04-05 21:18:36 cfd992f2-85bc-56a1-9577-1a8bd7ca6c82 2021-07-31, 15:04:21 cfd992f2-85bc-56a1-9577-1a8bd7ca6c82
1 Ilteeza Islam My little brother is using it . His mother lov... 5 10 True https://images-static.nykaa.com/prod-review/15... 2020-09-09 12:19:31 cfd992f2-85bc-56a1-9577-1a8bd7ca6c82 2021-07-31, 15:04:21 cfd992f2-85bc-56a1-9577-1a8bd7ca6c82
2 Zaneera I've always used sebamed products for my kid. ... 5 10 True NaN 2021-01-08 17:18:56 cfd992f2-85bc-56a1-9577-1a8bd7ca6c82 2021-07-31, 15:04:21 cfd992f2-85bc-56a1-9577-1a8bd7ca6c82
3 ANJALI SHENISHETTY Great soaps for kids. Smells amazing too 5 10 True NaN 2020-12-07 11:33:44 cfd992f2-85bc-56a1-9577-1a8bd7ca6c82 2021-07-31, 15:04:21 cfd992f2-85bc-56a1-9577-1a8bd7ca6c82
4 K.Latha Nice soap with mild properties 5 10 True NaN 2020-06-24 13:44:32 cfd992f2-85bc-56a1-9577-1a8bd7ca6c82 2021-07-31, 15:04:21 cfd992f2-85bc-56a1-9577-1a8bd7ca6c82
3217 Shivani Jaiswal I am satisfy with the product. It is best prod... 5 0 True NaN 2020-07-25 19:16:17 3aec5525-7284-5a84-9ea6-1cd8066828e1 2021-07-31, 15:28:01 3aec5525-7284-5a84-9ea6-1cd8066828e1
3218 Ayushi Uppal Best wet wipes till date for babies .... and f... 5 0 True NaN 2020-03-14 19:48:48 3aec5525-7284-5a84-9ea6-1cd8066828e1 2021-07-31, 15:28:01 3aec5525-7284-5a84-9ea6-1cd8066828e1
3219 anuradha ojha They are soft and have light smell. I use this... 5 0 True NaN 2020-03-09 10:20:03 3aec5525-7284-5a84-9ea6-1cd8066828e1 2021-07-31, 15:28:01 3aec5525-7284-5a84-9ea6-1cd8066828e1
3220 Sonia It makes my skin smooth and 5 0 True NaN 2020-03-05 3aec5525-7284-5a84-9ea6- 2021-07-31, 3aec5525-7284-5a84-9ea6-
91°F Partly sunny
Search
ENG IN
11:15 PM 5/11/2023

Fig B.2 SELECT ID



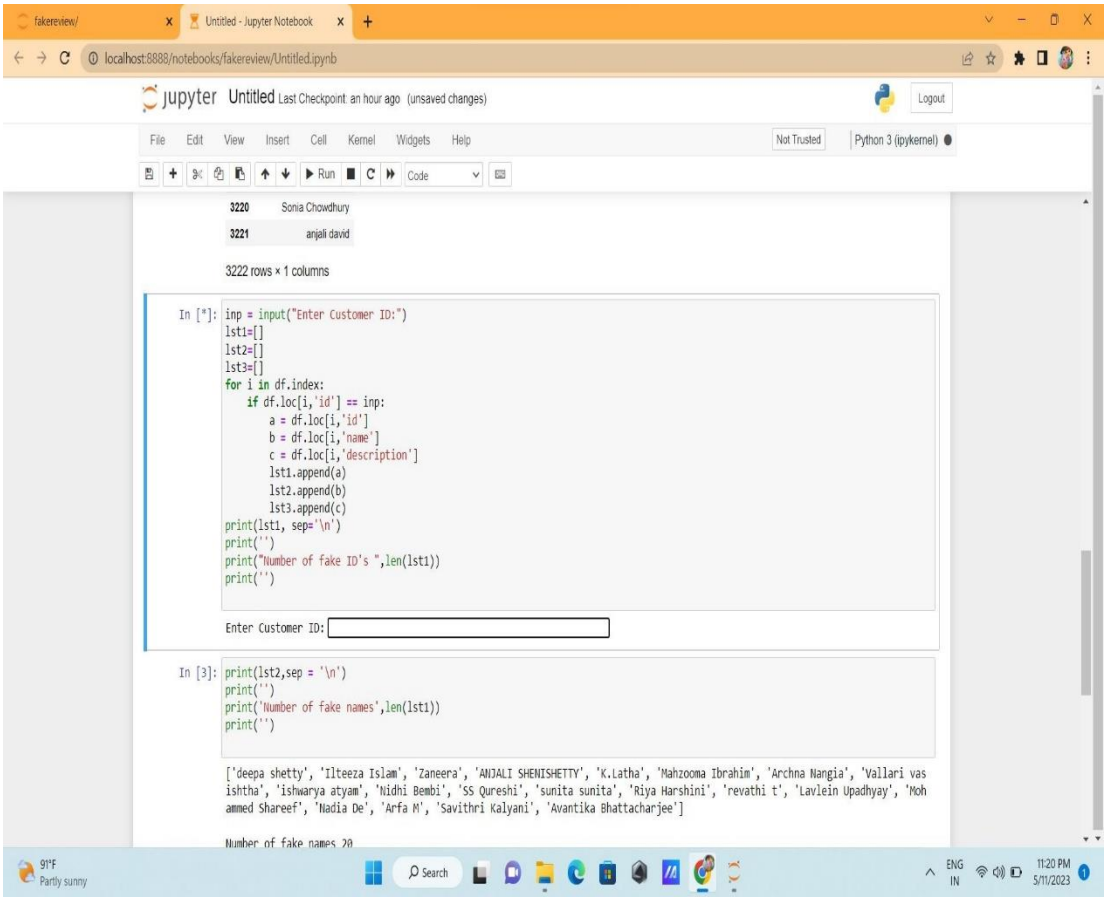
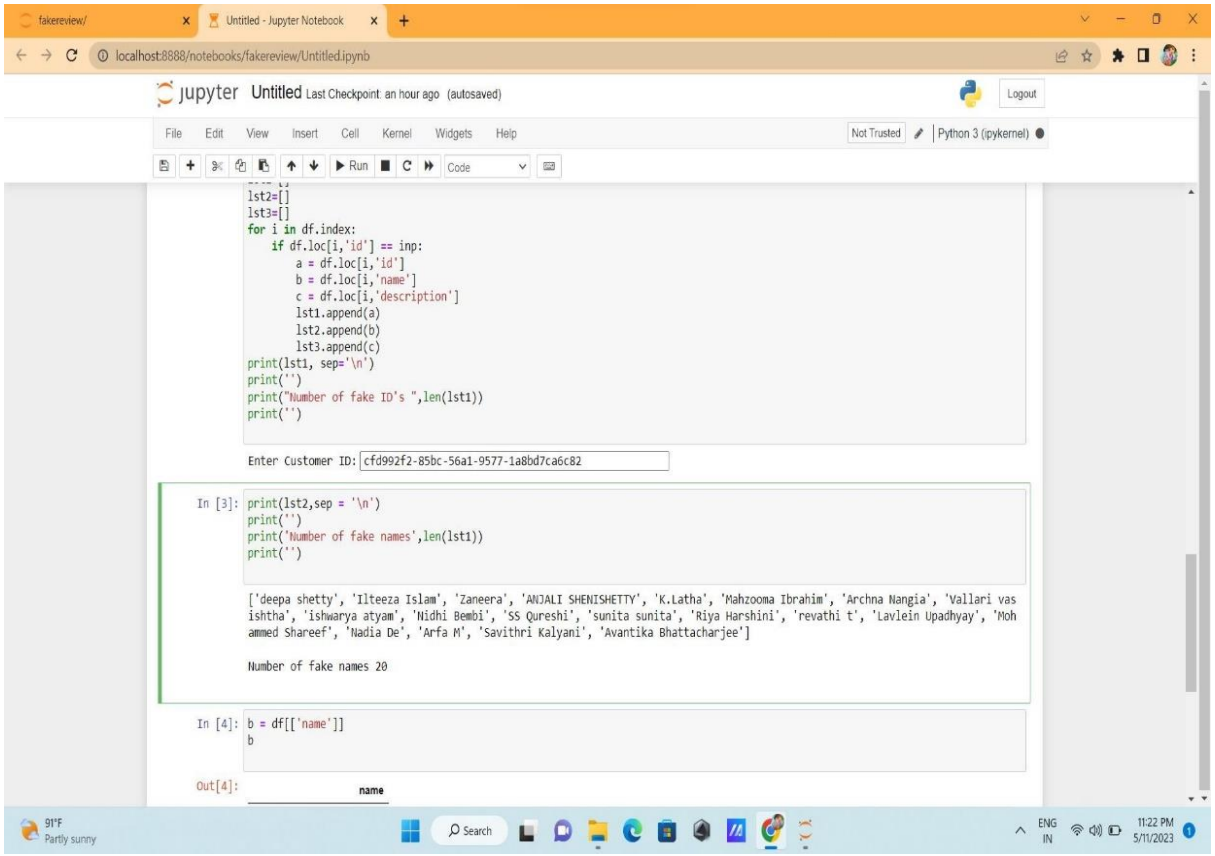


Fig B.2 NUMBER OF FAKE REVIEWS



REFERENCES

1. Yuejun Li, Xiao Feng, Shuwu Zhang,“Detecting Fake Reviews Utilizing Semantic and Emotion Model” IEEE2016
2. Kolli Shivagangadhar, Sagar H, SohanSathyan, Vanipriya C.H ,“Fraud Detection in Online Reviews using Machine Learning Techniques.” IJCER-2015.
3. Sushant Kokate, Bharat Tidke, “Fake Review and Brand Spam Detection using J48 Classifier.” IJCSIT, 2015.
4. Amir Karam, Bin Zhou, “Online Review Spam Detection by New Linguistic Features.” UNIVERSITY OF MARYLAND BALTIMORE COUNTY“15
5. Kolhe N.M., Joshi M.M.Jadhav A.B., Abhang P.D., “Fake Reviewer Groups“ Detection System.” IOSR-JCE 2014.
6. Arjun Mukherjee, Vivek Venkataraman, Bing Liu, NatalieGlance,“Fake Review Detection:Classification and Analysis of Real and Pseudo-Reviews.”\-CS“13