**ACKNOWLEDGEMENT**

Working on the project was very inserting and really enhanced my knowledge.

I would like to express my sincere gratitude to my beloved **Principal Prof. Aruna P. Kamath** without whose permission, I would not be able to do my project and for taking keen interest for the students of BCA in providing useful guidelines and giving all the necessary facilities.

I extend thanks to my Guide **“ “ Assistant Professor and Project Guide** under Computer Department, SDM College of Business Management, for her valuable guidance and constant encouragement, which helped me in successfully completing my project.

Finally, extend thanks to my parents and friends who were directly or indirectly involved in the completion of my project.

Place: BANTWAL …………………….. Date: MAY 03, 2024 **PREETHAM D S**

RegNo.: 163CS21042

**EXECUTIVE SUMMARY**

An overview of my internship experience is given in this executive summary, which also highlights the important abilities, information, and successes I acquired while working for the company.

Internship gave me the chance to work in a stimulating and demanding atmosphere while contributing to a variety of projects and engaging with experts in the field. Through practical experience, I was able to learn practical skills that complemented my academic knowledge and a firm awareness of industry practises.

Technical expertise was one of the main things I needed to improve during my internship. I got to work with a variety of software tools and technologies, like React as front-end, Express as back-end and MongoDB as database. By honing my technical abilities and using them in practical situations, this experience helped me to improve my problem-solving abilities.

Throughout the internship, I actively engaged in various projects and successfully contributed to their completion. This included ‘**AMAZON REVIEW’** and ‘**PRODUCT REVIEW DETECTION**, which allowed me to apply my knowledge, demonstrate initiative, and deliver results. These achievements not only contributed to the organization's goals but also enhanced my confidence and professional development.

This project report describes about two projects, on which I have worked during internship. The first project is regarding ‘**AMAZON REVIEW,** which helped to learn how to build a simple sentiment analysis application. The second project is about ‘**PRODUCT REVIEWV DETECTION’,** in that I have developed a website to see weather the review is fake or real

**INTERNSHIP CERTIFICATE ISSUED BY THE ORGANIZATION:**

****

| **SI. No** | **Contents** | **Page No** |
| --- | --- | --- |
| 1. | COMPANY PROFILE  1.1 OVERVIEW OF THE COMPANY  1.2 VISION AND MISSION OF THE ORGANIZATION  1.3 ORGANIZATION STRUCTURE  1.4 ROLES AND RESPONSIBILITIES OF PERSONNEL IN THE ORGANIZATION  1.5 PRODUCTS AND MARKET PERFORMANCE | 1-6 |

| 2. | ASSESSMENT OF ON JOB TRAINING - 1 2.1 INTRODUCTION  2.2 AIM  2.3 OBJECTIVE  2.4 PURPOSE  2.5 SCOPE  2.6 DISADVANTAGES & ADVANTAGES         2.6.1 ADVANTAGES         2.6.2 DISADVANTAGES  2.7 USE CASE DIAGRAM  2.8 ACTIVITY DIAGRAM  2.9 MODULE      2.9.1 MACHINE LEARNING               2.9.1.1 ALGORITHMS USED               2.9.1.2 LIBRARIES USED: 2.10 REQUIREMENT SPECIFICATION 2.10.1 HARDWARE REQUIREMENTS  2.10.2 SOFTWARE REQUIREMENTS  2.10.3 LANGUAGES USED  2.11CONCLUSION | 4-10 |
| --- | --- | --- |
| 3. | DISCUSSION ON JOB TRAINING CASE-2 | 11-16 |
| 4. | RESUME OF STUDENT | 17-21 |
| 5 | PHOTO GALLERY | 22-25 |

**LIST OF FIGURES:**

|  |  |
| --- | --- |
| **FIGURE NO.** | **PAGE NO.** |
| Fig 1.1 Organization Structure  Fig 2:Activity diagram  Fig 3:Module diagram  Fig 2.2:activity diagram  Fig 2.3:Module diagram  Fig 4.1:Use case  Fig 4.1:Use case | 10  17  18  23  24  40  41 |

**LIST OF TABLES:**

|  |  |  |
| --- | --- | --- |
| **TABLE NO.** | **TABLE NAME** | **PAGE NO.** |
| 1.2 | Product And Market performance | 12 |

**ABBREVIATIONS/NOTATIONS/NOMENCLATURE:**

|  |  |  |
| --- | --- | --- |
| **SL NO.** | **ABBREVIATIONS/ NOTATIONS** | **DESCRIPTION** |
| 1. | VS Code | Visual studio code |
| 2. | JS | Java script |
| 3. | HTML | Hypertext Markup language |
| 4. | PD | Pandas |
| 6. | LE | Label Encoder |
| 6. | CV | TfidVectorizer |
| 7. | CSS | Cascading Style sheet |

**CHAPTER – 1**

## **COMPANY PROFILE**

**COMPANY PROFILE**

#### **Overview of the company**

Codelab System is a rapidly growing company in the field of computer application Implementation, solutions and services. Codelab System is a service provider of Web-based Development & Web based Software Development Solutions, Mobile Application Development, Graphic Design and Windows Applications. Codelab Systems is headquartered in Mangalore, with the Business development in UAE,Sudi Arabia and Qatar In a short span of 8+ years, our products as well as services & Solutions have been widely accepted by the global market. Today, Codelab Systems has the experience to undertake any IT development or deployment works on a single point responsibility basis. Our efficient and experienced team is greatest resource Intellect’s infrastructure Houses A-team of young and competitive professionals having experience n Web Designing and Software Development who are dedicated to providing high-end Solution to our clients. We develop software and web based applications with Latest Technologies. For web development projects, we also provide hosting And, domain Facility for customers, so they don’t need to bother about that. Our , products and services are user friendly with easy controls and are of Superior specifications. We are always proactive to fulfill client’s needs and requirements to the best possible extent Of their satisfaction. We manage interactive sessions with clients throughout the Project development

### **VISION AND MISSION OF THE ORGANIZATION**

**VISION:**

To help people and businesses throughout the world realize their full potential.Codelab inspiring vision statement seeks to support people. You can see its intention isn’t about business; it’s about people and giving those people the services to be their best selves. With this aim, Codelab has numerous initiatives. It’s a big supporter of inclusivity, diversity, environmental issues, and corporate responsibility. We are on a journey to be the trusted performance leader that unleashes the potential of data.

**MISSION:**

“To enable people and businesses throughout the world to realize their full  
 potential and to organize the world’s information and make it universally

accessible and useful.”

Codelab Systems provides customized package to suit the needs of every client and take into consideration the needs and requirements of each clients and plan different ideas to improve client’s business strategies.Every customer satisfaction is our business and we pay special attention to each clien. To provide best services.The main goal of our company is to provide best and innovative products that will help to drive potential

customers to their businesses

### **ORGANIZATION STRUCTURE**

An organization is a group of people who work together, like a neighborhood association, a charity, a union, or a corporation. You can use the word organization to refer to group or business, or to the act of forming or establishing something. Organizational structure (OS) is the systematic arrangement of human resources in an organization so as to achieve common business objectives. It outlines the roles and responsibilities of every member of the organization so that work and information flow seamlessly, ensuring the smooth functioning of

an organization.  
**TypesofOrganizationalStructure**

•Hierarchical  
 •Flat  
 •Flatarchy  
 •Functional  
 •Divisional  
 •Matrix  
In a flatarchy, there are little to no levels of management. A company using this structure could have only one manager in between its executive and all other employees. It is called a flatarchy because it is a hybrid of a hierarchy and a flat organization. This type of organizational structure is used more by smaller companies since they have fewer employees, though it can be used in companies of all sizes. While some companies grow out of this organizational structure, others continue to use it. Codelab systems have a Matrix organization structure, where teams report to multiple leaders. The matrix design keeps open communication between teams and can help companies create more innovative products and services. Using this structure prevents teams from needing to realign every time a new project begins.

### **1.4ROLES AND RESPONSIBILITIES OF PERSONNEL IN THE ORGANIZATION**

We have an expertise team that offer unique solutions. All the members of our team  
are professional, experienced and have in depth knowledge of the technology. Codelab  
Systems provides customized package to suit the needs of every client and take into  
consideration the needs and requirements of each clients and plan different ideas to  
improve client’s business strategies.

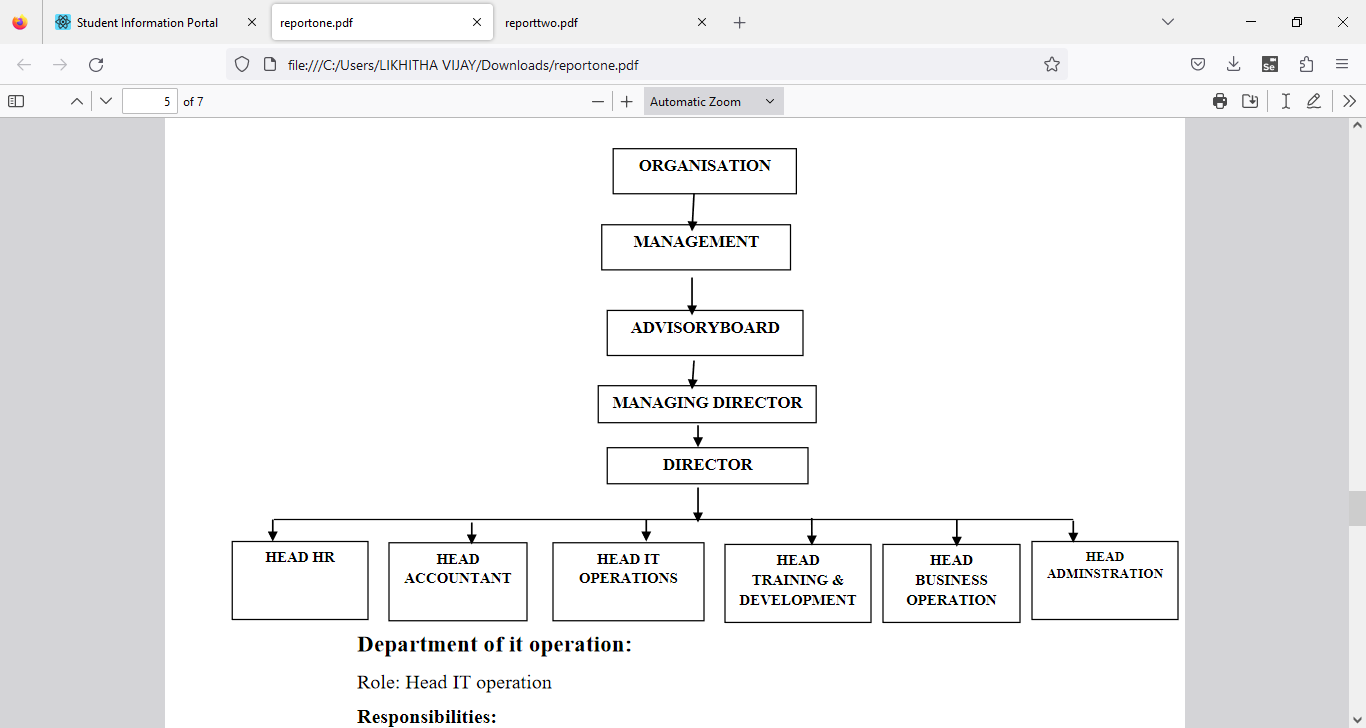


Figure 1.1 Organization Structure

**Department of it operation:  
Role:** Head IT operation  
**Responsibilities:**  
• Development of clients line project  
• Assigning tasks to the subordinate developers  
• Managing client meetings and maintaining a good relationship on stakeholde

**Department of Training and Development:**  
**Role:** Head of Training and Development  
**Responsibilities:**  
• Training to interns and newly joined employees & Work with new projects and

domain

**Department of HR:  
Role:** Head of HR  
**Responsibilities:**

• Maintaining Employees data & payroll calculations  
• Employee leave management & Interns internship program management.

**Department of Account:**  
**Role**: Head of Account.  
**Responsibilities:**  
• Keep track of daily Account & Maintaining Balance sheet and Income tax  
procedure

**Department of administration:**  
**Role:** Head Administration  
**Responsibilities:**  
• All the administration work such as file management, print, maintains data of  
computers and items. Arrangement of training program schedule.

**Department of Business operation:**  
**Role:** Head Administration  
**Responsibilities:**  
• Conducting market research, Contact and approach clients for live projects.  
• Communication with new clients and maintaining and managing social

### **PRODUCTS AND MARKET PERFORMANCE**

• MSS LODGE(INDIA) : MSS LODGE is a budget property located in the beautiful  
 city of Ujire.  
• SESCO : SESCO is one of the first enterprises in the electrical equipment sector,  
• QACADEMIA : Q-Academia providing wide range of career oriented IT Courses.

**SERVICES :**We believe in quality services

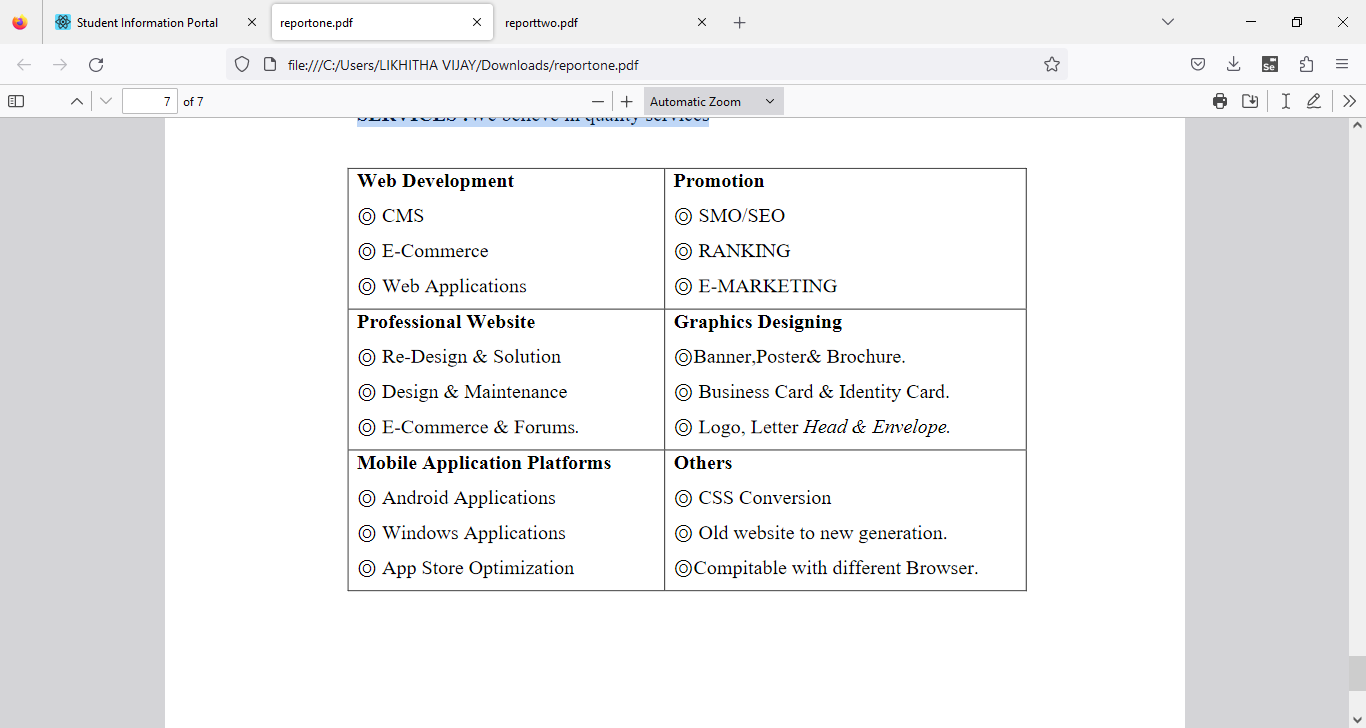


Figure 1.2 Product And Market performance

**CHAPTER- 2**

## **ASSESSMENT OF ON JOB TRAINING – 1**

## **CASE-1 ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**1.1 INTRODUCTION:**

As the commercial site of the world is almost fully undergone in online platform people is trading products through different e-commerce website. And for that reason reviewing products before buying is also a common scenario. Also now a day, customers are more inclined towards the reviews to buy a product. So analyzing the data from those customer reviews to make the data more dynamic is an essential field nowadays. In this age of increasing machine learning based algorithms reading thousands of reviews to understand a product is rather time consuming where we can polarize a review on particular category to understand its popularity among the buyers all over the world.

The objective of this paper is to categorize the positive and negative feedbacks of the customers over different products and build a supervised learning model to polarize large amounts of reviews. A study on amazon last year revealed over 88% of online shoppers trust reviews as much as personal recommendations. Any online item with large amounts of positive reviews provides a powerful comment of the legitimacy of the item. Conversely, books, or any other online item, without reviews puts potential prospects in a state of distrust. Quite simply, more reviews look more convincing. People value the consent and experience of others and the review on a material is the only way to understand others impression on the product. Opinions, collected from users’ experiences regarding specific products or topics, straightforwardly influence future customer purchase decisions. Similarly, negative reviews often cause sales loss. For those understanding the feedback of customers and polarizing accordingly over a large amount of data is the goal. There are some similar works done over amazon dataset. In did opinion mining over small set of datasets of Amazon product reviews to understand the polarized attitudes towards the products.

In our model, we used both manual and active learning approach to label our datasets. In the active learning process different classifiers are used to provide accuracy until reaching satisfactory level. After getting satisfactory result we took those labeled datasets and processed it. From the processed dataset we extracted features that are then classified by different classifiers. We used combination of two kinds of approaches to extract features: the bag of words approach and tf-idf & Chi square approach for getting higher accuracy.

1. **AIM:**

The world we see nowadays is becoming more digitalized. In this digitalized world e-commerce is taking the ascendancy by making products available within the reach of customers where the customer doesn’t have to go out of their house. As now a day’s people are relying on online products so the importance of a review is going higher. For selecting a product, a customer needs to go through thousands of reviews to understand a product. But in this prospering day of machine learning, going through thousands of reviews would be much easier if a model is used to polarize those reviews and learn from it. We used supervised learning method on a large scale amazon dataset to polarize it and get satisfactory accuracy.

1. **OBJECTIVE:**

* The objective of this research is to develop a supervised learning model that efficiently categorizes customer feedback into positive and negative sentiments for various products. In an era dominated by online commerce, where consumer decisions are heavily influenced by reviews, the need to streamline the analysis of vast amounts of feedback is paramount. By automating the sentiment polarization process, this study aims to provide consumers with valuable insights for informed purchasing decisions and assist businesses in understanding the reception of their products in the market.
* To achieve this objective, a combination of manual and active learning approaches will be utilized for labeling datasets. Active learning involves iteratively employing different classifiers to improve accuracy until a satisfactory level is attained. The labeled datasets will then undergo processing to extract relevant features. Feature extraction will utilize two approaches: the bag of words approach and tf-idf & Chi-square approach. These features will serve as input for various classifiers, ensuring accurate sentiment classification.
* Ultimately, the goal is to develop a robust model capable of efficiently categorizing customer feedback, thereby empowering both consumers and businesses in their decision-making processes within the dynamic landscape of online commerce.

1. **PURPOSE:**

* The purpose of this research is to address the pivotal role of online product reviews in influencing consumer behavior and purchasing decisions in the digital marketplace. With the increasing reliance on e-commerce platforms for trading goods, the significance of customer feedback has surged, as shoppers often turn to reviews to guide their buying choices. However, the sheer volume of reviews poses a challenge for effective analysis and interpretation.
* This study seeks to develop a supervised learning model that can automatically categorize customer feedback into positive and negative sentiments for diverse products. By leveraging machine learning algorithms, the aim is to streamline the process of sentiment analysis, making it more efficient and scalable. The ultimate objective is to empower both consumers and businesses with actionable insights derived from large-scale review data.
* Through a combination of manual and active learning approaches, the research endeavors to label datasets accurately, ensuring the training of robust classifiers. Active learning methodologies will be employed iteratively to enhance classification accuracy, culminating in a model capable of effectively polarizing sentiments across a wide array of products. Feature extraction techniques, including the bag of words approach and tf-idf & Chi-square approach, will be utilized to capture the nuanced characteristics of customer feedback and improve classification performance.
* By understanding and categorizing the sentiments expressed in customer reviews, this study aims to provide valuable intelligence to consumers, enabling them to make informed purchasing decisions. Likewise, businesses stand to benefit from insights into the reception of their products in the market, facilitating strategic decision-making and product development efforts. Ultimately, the research endeavors to contribute to the optimization of online shopping experiences and the enhancement of consumer satisfaction in the digital age.

1. **SCOPE:**

The scope of this research encompasses the development and implementation of a supervised learning model aimed at categorizing customer feedback into positive and negative sentiments across various products within the realm of e-commerce. Given the widespread reliance on online platforms for trading goods and the growing importance of customer reviews in influencing purchasing decisions, the study focuses on analyzing large volumes of review data to derive actionable insights.

1. **ADVANTAGES:**

* The script preprocesses the raw text data by removing stop words and cleaning the reviews, which helps in improving the quality of the input data for analysis.
* Through bar charts and pie charts, the script provides a visual representation of the distribution of positive and negative sentiments in the dataset, allowing for quick insights into sentiment proportions.
* The word cloud visualizations generated for both positive and negative sentiment categories offer a concise representation of the most frequent words used in reviews, aiding in understanding the key themes and sentiments expressed by customers.
* By using TF-IDF vectorization, the script transforms text data into numerical vectors, capturing the importance of words in reviews relative to the entire corpus. This approach helps in representing text data effectively for machine learning algorithms.
* The logistic regression model is a simple yet effective algorithm for binary classification tasks like sentiment analysis. It offers interpretability and can handle large feature spaces efficiently.
* The script evaluates the performance of the sentiment analysis model using accuracy score, providing a quantitative measure of how well the model performs on unseen data.
* The script utilizes the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance in the dataset, ensuring better generalization of the model by generating synthetic samples for the minority class.

**1.8 ACTIVITY DAIGRAM**

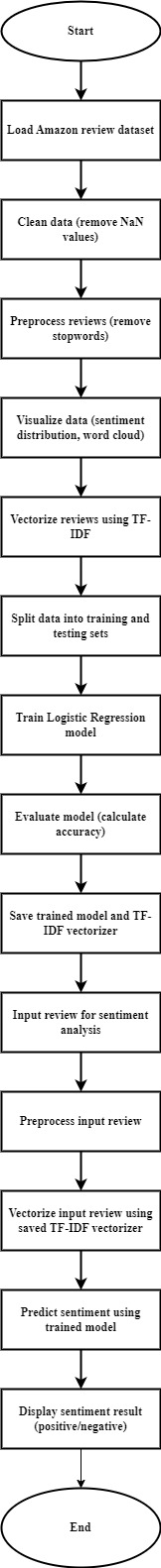


Fig 2:Activity diagram

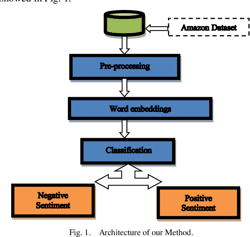
**1.9 MODULE**

Fig 3:Module diagram

### **1.10 REQUIREMENT SPECIFICATION:**

**1.10.1 HARDWARE REQUIREMENTS**

* RAM: 8 GB or higher
* Storage: 256 GB SSD or higher
* Network: Ethernet/Wi-Fi for internet connectivity
* Display: 15-inch monitor or larger
* Processor: Intel Core i5 or equivalent

**1.10.2 SOFTWARE REQUIREMENTS**

* Operating System: Windows 10 or Ubuntu 20.04 LTS
* Web Browser: Google Chrome or Mozilla Firefox
* Integrated Development Environment (IDE): Visual Studio Code, Python 3.8 or higher installed

**1.10.3 LANGUAGES USED**

* Front-end: HTML, CSS, JavaScript
* Back-end:Python, Flask

### **1.11 CONCLUSION:**

In conclusion, as online commerce continues to dominate global trade, the significance of customer reviews in influencing purchasing decisions has never been greater. Analyzing vast amounts of feedback to categorize and polarize sentiments towards products is essential in this landscape. With over 88% of online shoppers trusting reviews as much as personal recommendations, the volume and sentiment of reviews directly impact consumer trust and purchasing behavior. This study aimed to develop a supervised learning model to categorize positive and negative feedback, leveraging both manual and active learning approaches to label datasets. By employing various classifiers and feature extraction techniques, such as the bag of words and tf-idf & Chi-square methods, the model achieved satisfactory accuracy levels. Understanding and effectively polarizing customer feedback is crucial for businesses in maintaining consumer trust and competitiveness in the online marketplace.

**CHAPTER- 2**

## **ASSESSMENT OF ON JOB TRAINING - 2**

## 

## **CASE-2 ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**2.1 INTRODUCTION:**

Reviews are every user’s individual opinion regarding productsthey bought from the website along with the services provided to them by the company.The reviews they provide prove to be of extreme help to the future consumers or users,they help the users to make properdecision before purchasing a product or service.As users are provided with great offers and deals online they tend to use the e-commerce website to purchase the products,but sometimes there might be cases where the users are not provided a expected product or service,at this point the reviews of product comes in use. Reviews are important for both the company’s management and the user.There are cases where the company’s management team attempted to get positive reviews about the product to increase sales of the product,whereas on the other hand there were also attempts to degrade a genuine It is noticed that huge number of duplicate and reviews written by same users on same products and also different products.

**2.2 AIM:**

In today’s world, an e-commerce website’s healthygrowth is seen only when there is a huge presence of its user. The presence of users is seen in virtual world by the reviews they post after they use or buy the product. The reviews given to a single product determines whether the product holds credibility. In recent times it is seen that the companies tend to provide reviews to its own product in order to increase its sales, and on other hand there are some users who tend to degrade the product by giving fake negative reviews.Hence,a system which can detect such types of fake , be it positive or negative is needed.These reviews are also important in terms of the future progress of the ecommerce website a as this reviews Helps to overcome the negative aspects in the e-commerce website.

### **2.3 OBJECTIVE:**

### Implement measures to detect and minimize fake reviews, including duplicates and those from biased sources, to ensure that users can trust the authenticity of the feedback provided.

* Develop strategies to encourage genuine and helpful reviews from users, ensuring that future consumers can make informed decisions based on reliable information.
* Enhance transparency in the review system by clearly distinguishing between verified and unverified purchases, providing users with more accurate information about the credibility of each review.
* Implement mechanisms to identify and mitigate biased reviews, whether they are overly positive attempts to boost sales or malicious attempts to undermine competitors.
* Establish a feedback loop where users can provide constructive feedback directly to the company, enabling swift resolution of issues raised in reviews and fostering continuous improvement in product quality and service delivery.
* Provide guidance to users on writing helpful and constructive reviews, emphasizing the importance of honesty, clarity, and relevance in their feedback.
* Utilize data analytics tools to monitor review trends and patterns, identifying areas of improvement in products or services as well as emerging issues that need attention.
* Enhance customer support channels to address concerns raised in reviews promptly, demonstrating the company's commitment to customer satisfaction and resolving issues effectively.

### **2.4 PURPOSE:**

The purpose of the prepared statement is to highlight the significance of user reviews in the e-commerce industry. It emphasizes how reviews serve as valuable resources for both consumers and companies. Furthermore, it addresses the potential challenges posed by fake or duplicate reviews and the importance of maintaining authenticity and transparency in online product evaluations. Ultimately, the statement aims to underscore the pivotal role of reviews in facilitating informed decision-making for consumers and fostering trust and credibility in e-commerce platforms.

### **2.5 SCOPE:**

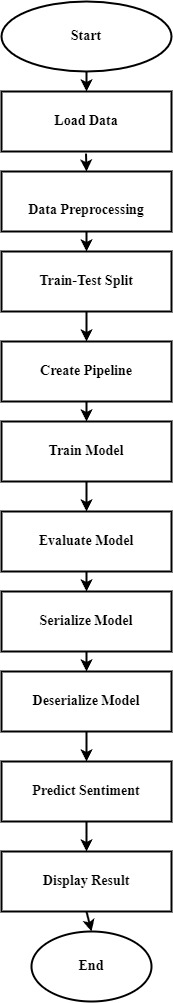
* Implementing improvements to the review system to enhance its usefulness for both users and the company.
* Implementing measures to ensure the authenticity and quality of reviews, reducing duplicate and fraudulent reviews.
* Developing resources to educate users on the importance of genuine reviews and how to provide constructive feedback.
* Ensuring transparency in the company's review collection process, including disclosing any incentives provided for reviews.
* Integrating user feedback mechanisms to address issues raised in reviews and improve products/services accordingly.
* Ensuring compliance with relevant laws and regulations governing online reviews, such as the Federal Trade Commission's guidelines on endorsements and testimonials.
* Utilizing data analytics to gain insights from reviews, such as identifying trends, common complaints, and areas for improvement.
* Integrating review feedback into customer support processes to address issues raised in reviews in a timely manner.
* Establishing processes for ongoing review system evaluation and improvement based on user feedback and industry best practices.
* Developing a communication strategy to inform users about changes to the review system and how their feedback is being utilized.
* Implementing systems to monitor review activity for fraudulent behaviour and providing regular reports on review metrics and trends.
* Exploring partnerships with reputable third-party review platforms to enhance the credibility of the company's review system.
* Providing training for employees involved in managing and responding to reviews to ensure consistency and professionalism in interactions with users.
* Ensuring that the review system is accessible to all users, including those with disabilities, through design and functionality considerations.
* Offering support for reviews in multiple languages to cater to a diverse user base.
* Designing the review system to accommodate growth in user activity and product/service offerings over time.

### **2.6 ADVANTAGES:**

* Your model relies solely on the text content of the reviews. This means it might miss out on other features like review length, writing style, sentiment of specific words, or metadata such as the time of the review, the reviewer's profile information, etc. Incorporating such features could enhance the model's performance.
* If your dataset has a significant class imbalance (i.e., many more original reviews than computer-generated ones, or vice versa), the model might become biased towards the majority class. Techniques like oversampling, under sampling, or using algorithms that handle class imbalance well (like ensemble methods) could mitigate this issue.
* Logistic regression is a good starting point, but it might not be the best-performing algorithm for your specific task. Experimenting with other algorithms such as Random Forest, Support Vector Machines, or even deep learning models like LSTM could lead to better results. Also, hyperparameter tuning via techniques like grid search could further optimize your model's performance.
* While accuracy is a common metric, it might not be the most suitable if the classes are imbalanced. You should consider using metrics like precision, recall, F1-score, or area under the ROC curve (AUC-ROC), especially in imbalanced classification tasks.
* Your model's performance might degrade when applied to real-world data that differs significantly from your training data. Regular updates and retraining on new data can help maintain its accuracy over time.
* Logistic regression provides coefficients for each feature, which offer interpretability. However, if you choose to move to more complex models, interpretability might decrease. Consider the trade-off between model performance and interpretability based on your requirements.

### 

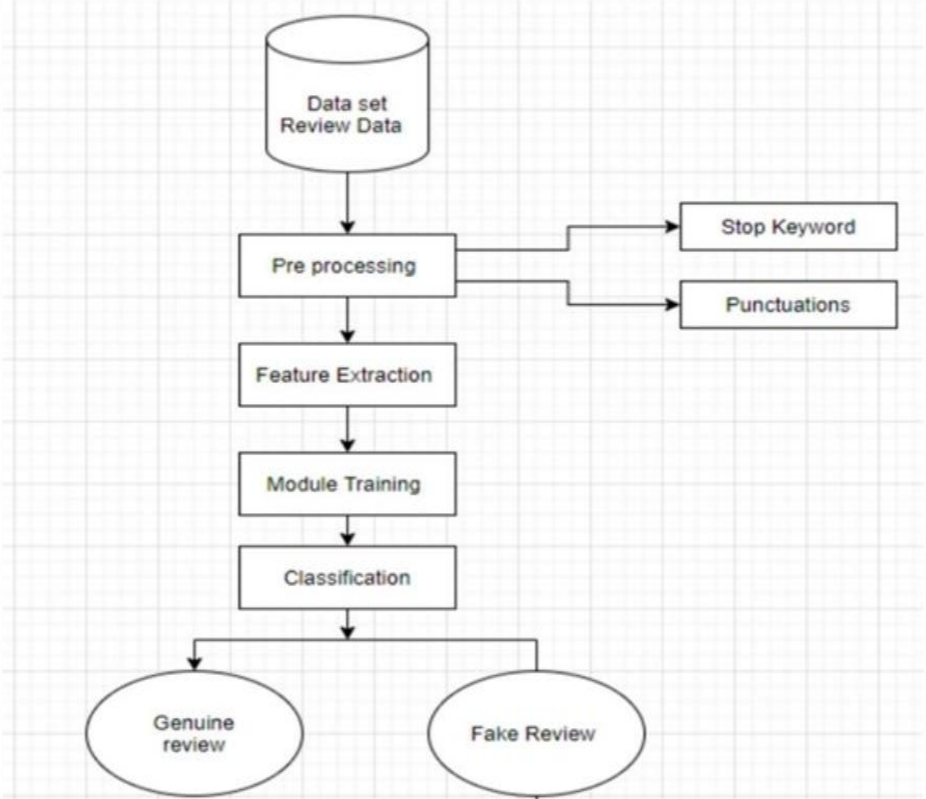
### **2.8 ACTIVITY DIAGRAM**



### 

Fig 2.2:activity diagram

### **2.9 MODULE:**

  
Fig 2.3:Module diagram

### 

### **2.10 REQUIREMENT SPECIFICATION:**

**2.10.1 HARDWARE REQUIREMENTS**

* RAM: 8 GB or higher
* Storage: 256 GB SSD or higher
* Network: Ethernet/Wi-Fi for internet connectivity
* Display: 15-inch monitor or larger
* Processor: Intel Core i5 or equivalent

**2.10.2 SOFTWARE REQUIREMENTS**

* Operating System: Windows 10 or Ubuntu 20.04 LTS
* Web Browser: Google Chrome or Mozilla Firefox
* Integrated Development Environment (IDE): Visual Studio Code, Python 3.8 or higher installed

**2.10.3 LANGUAGES USED**

* Front-end: HTML, CSS, JavaScript
* Back-end:Python, Flask

### 

### **1.11 CONCLUSION:**

# In conclusion, user reviews play a crucial role in informing future consumers about products and services offered by e-commerce websites. These opinions aid users in making informed decisions, particularly when considering the reliability and quality of a product or service. Reviews serve as a valuable feedback mechanism for both consumers and companies, offering insights into customer satisfaction and areas for improvement. However, the prevalence of duplicate or manipulated reviews poses challenges in maintaining the integrity of the feedback system. It is imperative for companies to prioritize transparency and authenticity in managing reviews to foster trust and credibility among consumers.

## 

## **CASE 2:IMPLEMENTATION**

**IMPORT LIBRARY:**

import numpy as np

import pandas as pd

# import seaborn as sns

# import matplotlib.pyplot as plt

# %matplotlib inline

import warnings, string

warnings.filterwarnings('ignore')

from sklearn.model\_selection import train\_test\_split, GridSearchCV

from sklearn.metrics import classification\_report, confusion\_matrix, accuracy\_score

from nltk.corpus import stopwords

from sklearn.feature\_extraction.text import CountVectorizer, TfidfTransformer

from sklearn.pipeline import Pipeline

from sklearn.linear\_model import LogisticRegression

**READ DATASET:**

df = pd.read\_csv('fake reviews dataset.csv')

df.head()

**OUTPUT:**



**DROPPING COLUMNS:**

df.dropna(inplace=True)

df.info()

**OUTPUT:**

**REMOVE PUNCTUATION:**

def text\_process(review):

nopunc = [char for char in review if char not in string.punctuation]

nopunc = ''.join(nopunc)

return [word for word in nopunc.split() if word.lower() not in stopwords.words('english')]

**TRAIN-TEST-SPLIT:**

review\_train, review\_test, label\_train, label\_test = train\_test\_split(df['text\_'],df['label'],test\_size=0.35)

**CREATING MODEL:**

LR\_model = Pipeline([

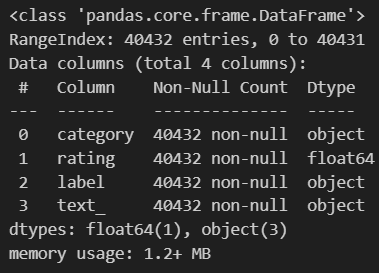
('bow',CountVectorizer(analyzer=text\_process)),

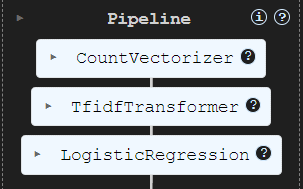
('tfidf',TfidfTransformer()),

('classifier',LogisticRegression())

])

LR\_model.fit(review\_train,label\_train)

**OUTPUT:**

****

**PREDICTING:**

lr\_pred = LR\_model.predict(review\_test)

lr\_pred

**OUTPUT:**

****

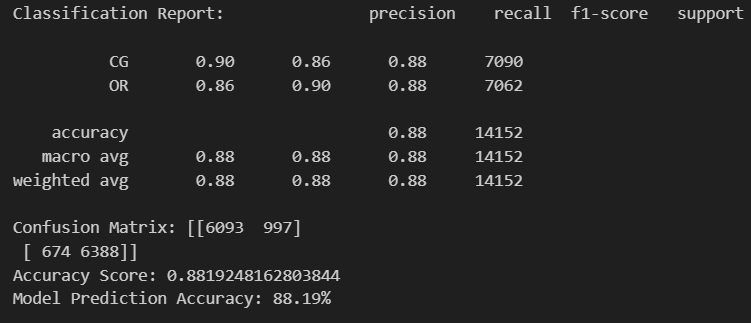
**ACCURACY:**

print('Classification Report:',classification\_report(label\_test,lr\_pred))

print('Confusion Matrix:',confusion\_matrix(label\_test,lr\_pred))

print('Accuracy Score:',accuracy\_score(label\_test,lr\_pred))

print('Model Prediction Accuracy:',str(np.round(accuracy\_score(label\_test,lr\_pred)\*100,2)) + '%')



import pickle

pickle.dump(LR\_model,open('LogisticRegression\_Model.pkl','wb'))

model = pickle.load(open('LogisticRegression\_Model.pkl','rb'))

def clean\_message(message):

stp\_words = stopwords.words('english')

clean\_message = " ".join(word for word in message[0].split() if word not in stp\_words)

return [clean\_message]

def predict\_sentiment(message\_text):

cleaned\_message = clean\_message(message\_text)

prediction = model.predict(cleaned\_message)

if prediction[0] == 'OR':

return 'Original Review'

else:

return 'Computer Generated Review'

#Now you can use the predict\_sentiment function to classify reviews

input\_message = input('Enter the Review :')

new = [input\_message]

result = predict\_sentiment(new)

print(f'The review is : {result}')

**OUTPUT:**

****

**CHAPTER-4**

**USE CASE-1 AND USE CASE-2**

### **USE CASE-1 DIAGRAM:**

Fig 4.1:Use case

### **USE CASE-2 DIAGRAM:**

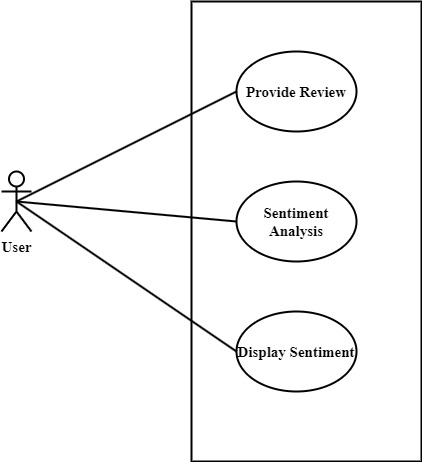
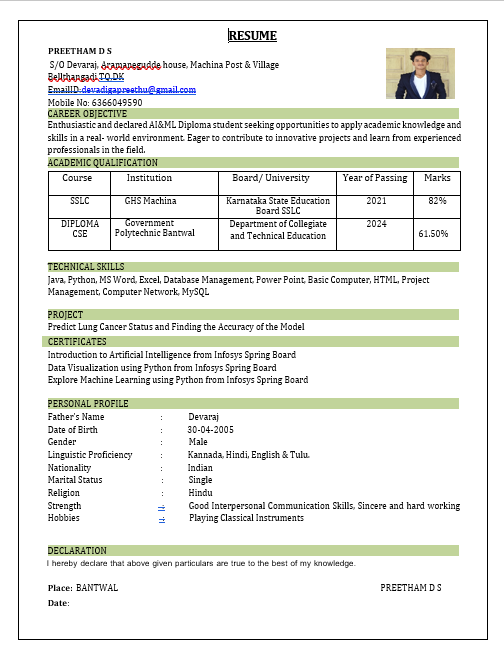


Fig 4.2:Use case

** 5. RESUME OF STUDENT**

## **6. PHOTO GALLERY**

### **SNAPSHOTS OF OJT – 1**

SNAPSHOTS OF OJT – 2

**7. CONCLUSION**

**8. FUTURE SCOPE AND FURTHER ENHANCEMENT  
OF THE PROJECT**

**9. REFERENCE**

[1] https://www.sciencedirect.com/science/article/abs/pii/S0747563213004007#preview-section-introduction

[2]

[3]

[4]

[5]

(paste research paper links)