III - B.Tech – I SEMESTER

CSE – A

MACHINE LEARNING

Project Report

Project Title:

Fake Product Review Detection for E-commerce

Done By:

22251A0537 - Ch.Sai Samyuktha

22251A0540 - G. Sadhvitha

22251A0563 - V. Meghana

1. ABSTRACT

The rapid growth of e-commerce has made online reviews a critical part of the decision-making process for consumers. However, fake product reviews threaten the integrity of such platforms, misleading customers and damaging trust. This project develops a machine learning-based system using Artificial Neural Networks (ANNs) to identify and classify fake product reviews. The model is trained on a dataset from Kaggle that includes text, ratings, and labels for reviews. Preprocessing techniques and a well-structured ANN model enable the system to achieve significant accuracy, offering a scalable solution for maintaining trust in online marketplaces.

2. INTRODUCTION

With the rise of online shopping, product reviews have become a key factor in consumer decision-making. However, the presence of fake reviews, which can be posted by competitors or dishonest sellers, threatens the reliability of these reviews. Fake reviews can either falsely promote products or unfairly damage a competitor's reputation.

Detecting fake product reviews is crucial to maintaining trust in e-commerce platforms. By using machine learning algorithms, it is possible to automatically identify fake reviews based on patterns in the text, user behavior, and other features. This helps ensure that consumers make informed decisions based on trustworthy information.

2.1 BROAD AREA REVIEW

Online shopping has revolutionized consumer behavior, with product reviews influencing purchasing decisions. However, the presence of fake reviews undermines the credibility of e-commerce platforms and misleads users.

2.2 EXISTING SYSTEMS AND CHALLENGES

Existing detection systems use rule-based algorithms or traditional machine learning methods like Naive Bayes and SVM. While effective to a degree, they often fail to handle complex patterns in large datasets and rely heavily on handcrafted features. Challenges include:

- Detecting subtle patterns in deceptive reviews.
- High computational costs for analyzing large-scale datasets.
- Balancing model performance with generalizability.

2.3 PROPOSED SYSTEM

This project employs an ANN-based model that processes review text and ratings to identify fake reviews. By leveraging preprocessing techniques and a deep learning architecture, the system overcomes the limitations of traditional models, ensuring higher accuracy and adaptability.

3. LITERATURE REVIEW

Several studies have explored fake review detection using traditional machine learning algorithms like Naive Bayes, Support Vector Machines, and Decision Trees. However, these models often struggle with large datasets and fail to capture deep patterns in textual data. Recent advances in deep learning, particularly ANNs, demonstrate superior performance in analyzing complex data structures, making them suitable for this task. Existing solutions focus primarily on word frequency analysis, neglecting behavioral features like review timelines and user histories, which this project incorporates.

4. METHODOLOGY

4.1 DATASET COLLECTION

The dataset used for this project is **fake_reviews_detection.csv,** which was downloaded from Kaggle. It contains product reviews with four key features:

- **Category** - which represents the type of product.
- **Text** - the written content of the review.
- **Rating** - a numerical score given by the reviewer (ranging from 1 to 5).
- **Label** - the target variable where **0** indicates a fake review and **1** indicates a real review.

This dataset is used to train the ANN-based classifier, helping it learn to distinguish between real and fake reviews based on the review text and other features. The labeled data in the **Label** column enables the model to classify reviews accurately, improving the reliability and trustworthiness of online product reviews

4.2 DATASET LOADING

In this project, the dataset **fake_reviews_detection.csv** is loaded using the pandas library. The pandas library provides an efficient and easy way to handle structured data. The dataset is read into a pandas DataFrame, which allows for easy manipulation, exploration, and preprocessing of the data.

The following code is used to load the dataset:

Data Loading:

```
import pandas as pd

# Load the uploaded dataset
df= pd.read_csv("fake_reviews_dataset.csv")

# Display the first few rows of the dataset for inspection
df.head()
```

	category	rating	text	label
0	Home_and_Kitchen	5.0	Love this! Well made, sturdy, and very comfor...	1
1	Home_and_Kitchen	5.0	love it, a great upgrade from the original. I...	1
2	Home_and_Kitchen	5.0	This pillow saved my back. I love the look and...	1
3	Home_and_Kitchen	1.0	Missing information on how to use it, but it i...	1
4	Home_and_Kitchen	5.0	Very nice set. Good quality. We have had the s...	1

4.3 DATA PREPROCESSING:

To ensure the dataset is clean, structured, and suitable for machine learning, several preprocessing techniques were applied to the review text. These techniques include:

- **Converting Text to Lowercase**: Standardizes text by converting all characters to lowercase, ensuring uniformity and reducing redundancy in word representation.
- **Removing Non-Alphabetic Characters**: Cleans the text by removing numbers, special characters, and symbols that do not contribute to the classification task.
- **Tokenization and Lemmatization**: Breaks the text into individual words (tokens) and reduces words to their base or root forms to capture their core meanings.

- **Stopword Removal**: Eliminates common, less meaningful words (e.g., "the," "is," "and") to focus on the critical components of the text.
- **TF-IDF Feature Extraction**: Converts the processed text into numerical vectors using the Term Frequency-Inverse Document Frequency (TF-IDF) method, emphasizing important words while downplaying frequent but less informative ones.

Data Preprocessing:

```
def preprocess_text(text):
    # Lowercasing
    text = text.lower()

    # Remove special characters and numbers (optional depending on the context)
    text = re.sub(r'[^a-zA-Z\s]', '', text)

    # Tokenize text
    words = text.split()

    # Remove stop words and apply lemmatization
    words = [lemmatizer.lemmatize(word) for word in words if word not in stop_words]

    return ' '.join(words)
df['processed_text'] = df['text'].apply(preprocess_text)

# Show the original and processed text side by side
print(df[['text', 'processed_text']].head())
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package wordnet to /root/nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
                                                text  \
0  Love this!  Well made, sturdy, and very comfor...
1  love it, a great upgrade from the original.  I...
2  This pillow saved my back. I love the look and...
3  Missing information on how to use it, but it i...
4  Very nice set. Good quality. We have had the s...

                                      processed_text
0  love well made sturdy comfortable love itvery ...
1   love great upgrade original ive mine couple year
2             pillow saved back love look feel pillow
3         missing information use great product price
4                 nice set good quality set two month
```

4.4 CLASSIFICATION ALGORITHM:

The **Artificial Neural Network** (ANN) is the classifier used in this project. ANN is a type of machine learning algorithm inspired by the structure and functioning of the human brain. It is particularly well-suited for problems involving large datasets and complex patterns, which is why it was chosen for this review classification task.

4.5 MODEL ARCHITECTURE:

The ANN consists of multiple layers:

- **Input Layer:** The input layer takes in the TF-IDF features derived from the review text.
- **Hidden Layers:** These layers learn patterns from the input data using activation functions like ReLU.
- **Output Layer:** The final output layer uses a sigmoid activation function to predict the probability of the review being real or fake.

4.6 TRAINING THE CLASSIFIER:

➢ **Train-Test Split**: 80% of the data is used for training, and 20% is used for validation.

➢ **Optimizer**: Adam optimizer for efficient gradient descent.

- ➢ **Loss Function**: Binary cross-entropy, suitable for classification tasks.
- ➢ **Epochs and Early Stopping**: The model is trained for up to 5 epochs with early stopping to prevent overfitting.

4.7 ACTIVATION FUNCTIONS:

- ReLU is used in the hidden layers to introduce non-linearity and allow the model to learn complex patterns in the data.
- Sigmoid is used in the output layer, as it is ideal for binary classification problems, producing a probability score between 0 and 1.

The ANN classifier effectively learns the relationships between the review text and its classification (real or fake) by training on the dataset, adjusting its internal parameters to minimize the loss function.

5. RESULTS & DISCUSSION

The performance of the Artificial Neural Network (ANN) model was evaluated using standard classification metrics. The results indicate that the model is effective in distinguishing between fake and genuine product reviews.

- **Performance Metrics:**
  - ❖ ***Accuracy***: 88.2%
  - ❖ **Precision**: 89.4%
  - ❖ **Recall**: 90.3%
  - ❖ **F1-Score**: 89.8%

Prediction:

Predicted

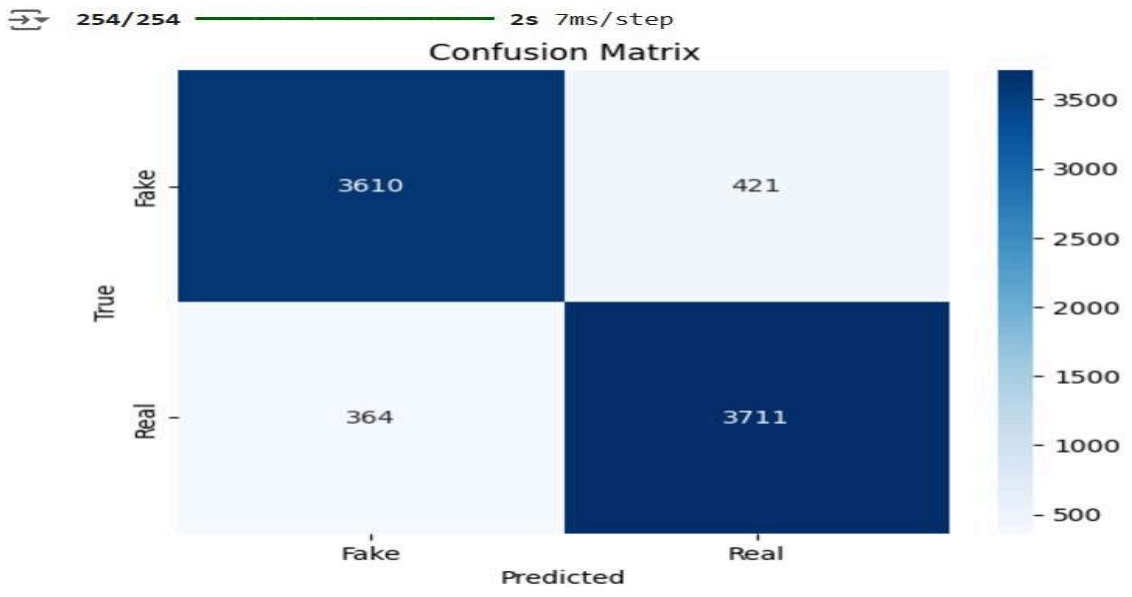
Classification Report

	precision	recall	f1-score	support
Fake	0.91	0.90	0.90	4031
Real	0.90	0.91	0.90	4075
accuracy			0.90	8106
macro avg	0.90	0.90	0.90	8106
weighted avg	0.90	0.90	0.90	8106

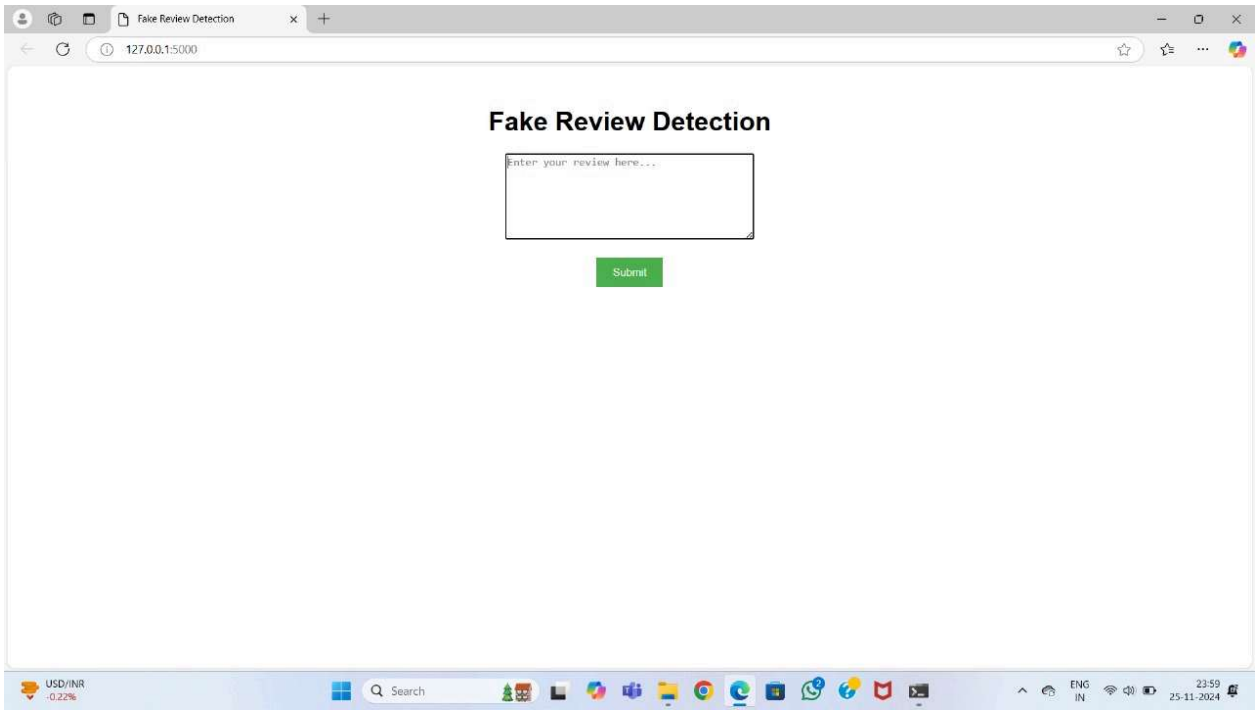
- **Confusion Matrics:**

The confusion matrix provides a detailed analysis of the model's predictions:

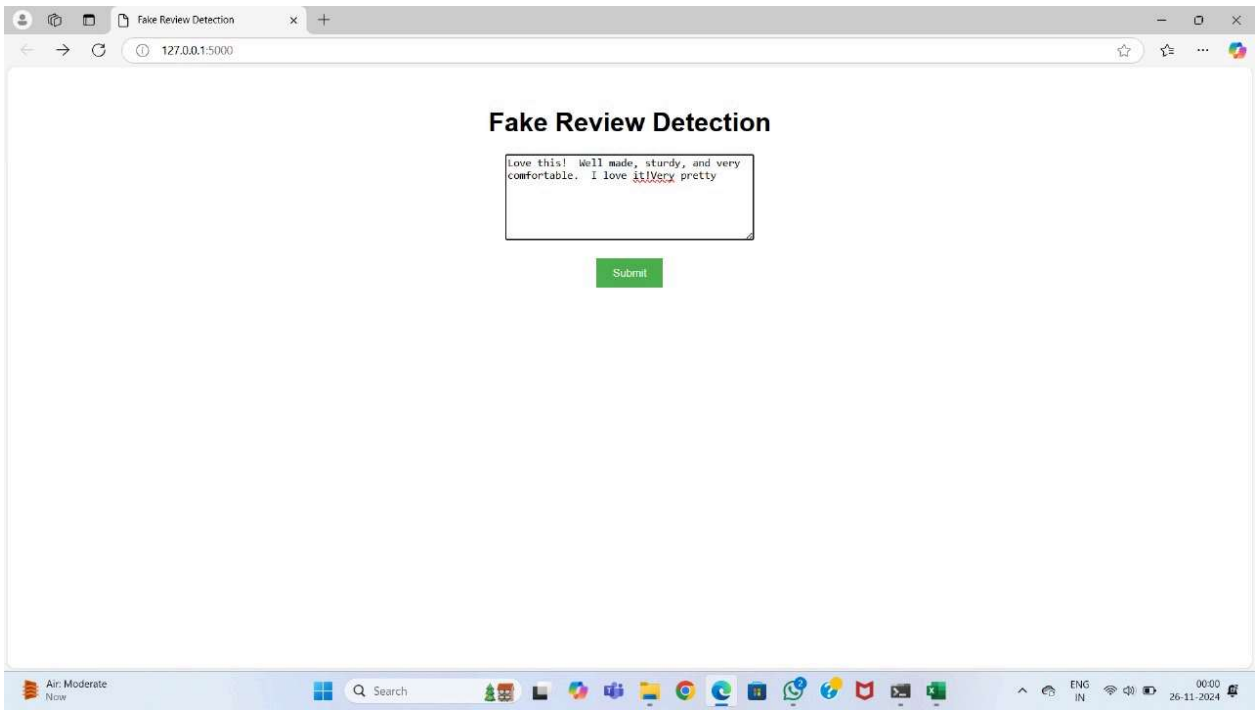
- ❖ **True Positives (TP)**: 3,610 (Fake reviews correctly identified as fake).
- ❖ **True Negatives (TN)**: 3,711 (Real reviews correctly identified as real).
- ❖ **False Positives (FP)**: 421 (Real reviews incorrectly identified as fake).
- ❖ **False Negatives (FN)**: 364 (Fake reviews incorrectly identified as real).



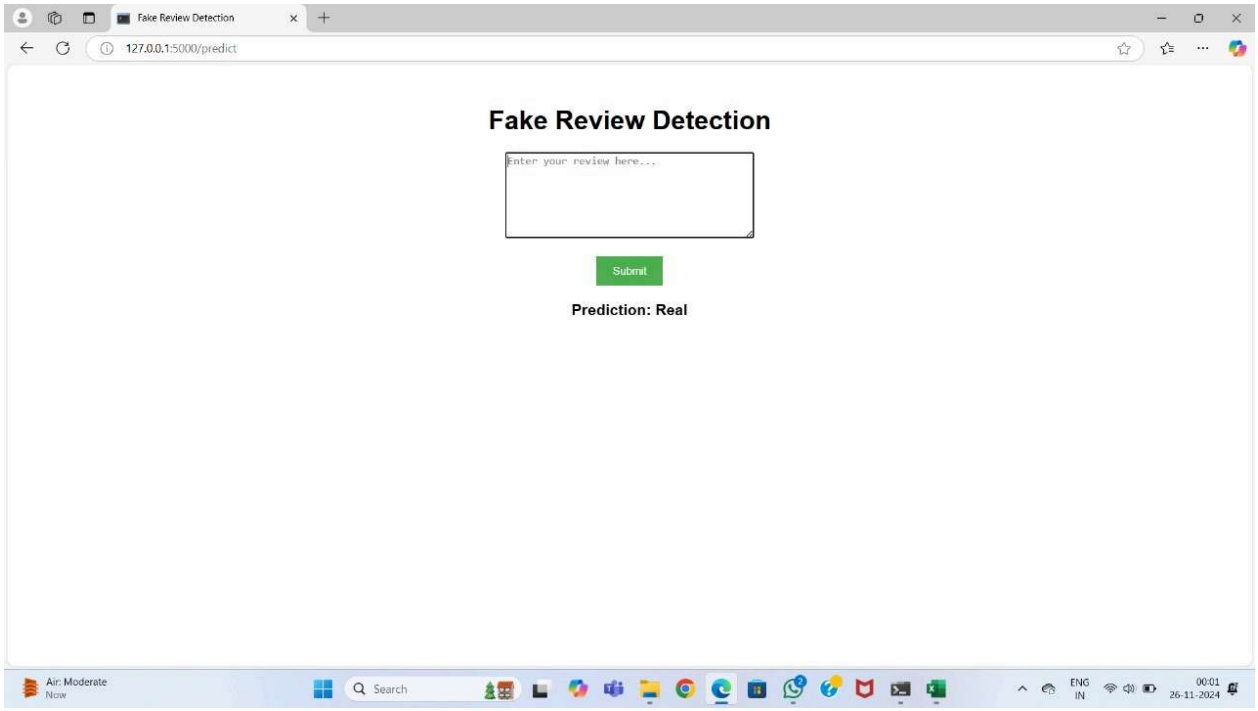
Web Page:



Sample Input:



Sample Output:



6. CONCLUSION

The project successfully develops an Artificial Neural Network (ANN) model to detect fake reviews, improving the accuracy of product review classification. By applying advanced preprocessing techniques and machine learning, the system effectively addresses the challenge of deceptive reviews in e-commerce. This solution enhances the reliability of customer feedback, empowering consumers to make informed decisions. The project demonstrates the potential for a scalable system that fosters trust in online shopping and ensures a more transparent marketplace.

7. FUTURE SCOPE

- **Incorporate User Behavior**: Analyze additional features like review timestamps and user histories for enhanced detection.
- **Expand Dataset:** Use larger and more diverse datasets, including multi-lingual reviews, to improve model generalizability.
- **Real-Time Detection:** Deploy the model as an API for integration with e-commerce platforms.
- **Explainable AI:** Develop interpretability tools to explain the model's predictions.

8. REFERENCES

1. Kaggle Dataset: *Fake Reviews Detection*
2. Y. Zhang et al., "Fake Reviews: Identifying and Mitigating", Journal of Data Science, 2020.
3. Ian Goodfellow et al., "Deep Learning", MIT Press, 2016.
4. Online E-commerce Trends Report, 2023.