A FIELD PROJECT REPORT

on

"Sentiment-Based Insights into Amazon Musical Instrument Purchases"

Submitted

by

221FA04094 Ammulu

221FA04148 Manasa

221FA04256 Mokshagna

221FA04416 Ganesh

Under the guidance of

Maridu Bhargavi

Assistant Professor, Department of CSE

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

VIGNAN'S FOUNDATION FOR SCIENCE, TECHNOLOGY AND RESEARCH Deemed to be UNIVERSITY

Vadlamudi, Guntur.

ANDHRA PRADESH, INDIA, PIN-522213.

CERTIFICATE

This is to certify that the Field Project entitled **"Sentiment-Based Insights Into Amazon Musical Instrument Purchases"** that is being submitted by 221FA04094 (Ammulu), 221FA04148 (Manasa), 221FA04256 (Mokshagna), 221FA04416 (Ganesh) for partial fulfilment of Field Project is a bonafide work carried out under the supervision of Ms.M.Bhargavi, M.Tech., Assistant Professor, Department of CSE.

Guide name& Signature	Dr. S. V. Phani Kumar	Dr.K.V. Krishna Kishore
Assistant/Associate/Professor, CSE	HOD,CSE	Dean, SoCI

DECLARATION

We hereby declare that the Field Project entitled **"Sentiment-Based Insights Into Amazon Musical Instrument Purchases"** is being submitted by 221FA04094 (Ammulu), 221FA04148 (Manasa), 221FA04256 (Mokshagna), and 221FA04416 (Ganesh) in partial fulfilment of Field Project course work. This is our original work, and this project has not formed the basis for the award of any degree. We have worked under the supervision of Ms. M.Bhargavi, M.Tech., Assistant Professor, Department of CSE.

By

221FA04094(Ammulu),
221FA04148(Manasa),
221FA04256(Mokshagna),
221FA04416(Ganesh)

Date:

ABSTRACT

The project aims to analyse Amazon music instrument reviews to understand customer satisfaction in terms of key sentiments related to the products. Online shopping has made customer reviews increasingly relevant for both buyers and sellers. The project analyses unprocessed text data using Natural Language Processing (NLP), which includes steps like text cleaning, tokenization, stop-word removal, stemming, and lemmatization. Word embeddings or term frequency-inverse document frequency (tf-idf) are used to encode the large dataset. A method of sentiment analysis on user reviews is described, with the dataset going through preliminary NLP work. Sentiments are classified using various models, evaluated based on accuracy, cross-validation scores, and classification re- ports. Visualization graphics depict sentiment distribution, emphasizing the potential of NLP for better sentiment analysis performance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER-1

INTRODUCTION

1. INTRODUCTION

In recent years, sentiment analysis has gained great importance and has typ- ically been used in understanding customer behaviour, especially in e-commerce, where consumer reviews significantly impact selling strategies and business practices[1]. This paper aims to describe a way in which reviews from the users can be classified based on rating using various methods of technology to evaluate the emotions of the consumers. Homogeneous, multilabel, multi class data concerning musical instruments reviews is in this regard processed using NLP techniques such as stemming, stop word removal, and TF-IDF in order to change unstructured text data to a structured format that machine learning makes sense Some of the machine learning models that are executed include, Gradient Regression, Support Vector Classifier (SVC), K-Nearest Neighbours (KNN), Decision Tree, Random Forest, Gradient Boosting Machine (GBM), Naive Bayes, and LightGBM are a few examples of other. These models are produced using pre-established rating cut-off points, at which they are trained and evaluated to anticipate an outcome of the review being positive, neutral or negative. To assess the accuracy of the various models and thereby develop an effective model for the purposes of sentiment prediction, parameters such as accuracy scores, cross validation and classification reports are put into place. Also, the features distribution of ratings and other text features are presented in graphs to help capture the overall trend of user feedback better. This study's main goal is to evaluate the various machine learning models' efficacy in doing sentiment analysis and giving companies a trustworthy tool for examining client comments. It is evident that integrating different ma- chine learning classifiers with NLP techniques makes the analysis of user emotion very effective. This method makes the selection of the most optimal model for solving practical problems easier, as it is possible to perform comparisons of numerous models to evaluate how effective each scan classification algorithm is in completing sentiment analysis tasks[2]. This technique can help businesses categorize reviews more effectively and employ the results obtained to enhance customer service, create improved products, and finally increase the number of loyal customers. This will help them make better decisions in trying to enhance customer satisfaction and improve their product offerings. This is roughly the customer review for each of the 10,000 or so musical instruments present in the Amazon has to be predicted about their residual feelings either positive or negative. The raw data has been extracted from the Kaggle dataset for Amazon Musical Instruments [3]. The dataset consists of 9 columns and 10262 rows. The list of the columns is given in Table 1 and a short description is also attached for each column.

Attribute	Description
Reviewer ID	Unique identifier for reviewer
Asin	Unique identifier for product
Reviewer Name	Full name of the reviewer
Helpful	Helpfulness rating of the review
Review Text	Text submitted in the review by the user
overall	Rating of the product given by the user
summary	Summary or title of the review given by the user
Unix Review Time	Unix timestamp [4] for the Date and Time at which review was submitted
Review Time	Date and Time at which review was submitted

TABLE 1: The dataset used in the implementation contains a list of columns and their descript

CHAPTER-2

LITERATURE SURVEY

2. LITERATURE SURVEY

Literature review

Zhang and Yin focus on issues related to domain-specific sentiment analysis in e-commerce[1].Information Processing Management. The paper explores domain-specific problems in sentiment analysis related to musical instruments, offering direction for feature extraction in applied contexts involving complex items such as musical instruments Rashika Mishra and colleagues have performance of reaching 92.8using CNN based architecture with U-Net network on ISBI 2017 dataset. Redha Ali and others presented a way to use VGG19 and U-Net structure and achieved 93.6% on ISIC 2018 dataset. Use features extracted from grab cut segmentation of melanomas and SVM based study was conducted Wenliang et al., Xinyu W et al., and Xinyu X et al. [3] investigated Ama using conventional machine learning methods in conjunction with the deep learning algorithm RNN. Yadav et al. and Jain et al. [4] employed a non-standard feature-based method in 2020 to perform sentiment analysis on Flipkart and Amazon product reviews. In fact, it fared better than TF, TF-IDF, and Naive Bayes Meth . In 2017, Hu et al. [5] combined the use of the LSTM model with the application of keyword vocabulary to do sentiment analysis on brief texts.

Title	**Year**	**Dataset**	**Methods/Algorithms**	**Classifier**	**Accuracy**	**Limitations**
Overall and Feature Level Sentiment Analysis of Amazon Product Reviews Using Machine Learning Techniques and Web-Based Chrome Plugin	2022	products on Amazon, including 142.8 million reviews	These models comprise not only conventional algorithms such as Random Forest, Multinomial Naïve Bayes, Complement Naïve Bayes, Bernoulli Naïve Bayes, but also VADER Sentiment Analysis.	Random Forest Classifier with TF-IDF Vectorizer	NA	Difficulty in teaching computers to grasp sarcasm, potential improvements with BERT for higher accuracy
Research on Sentimen	2023	Online reviews of	BERT model for sentiment analysis, word frequency	BERT emotion model	Not specified in the	

t Analysis and Personalized Recommendation Based on Agricultural Product Reviews		Shenyou Songhua Egg from the shopping platform Jingdong	analysis, and word cloud drawing		provided context	
A Scalable Approach for Sentiment Analysis of Turkish Tweets and Linking Tweets to News	2016	6000 manually labeled Turkish tweets	Naive Bayes, Complementary Naive Bayes, Logistic Regression	Complementary Naive Bayes	79.38	Lower accuracy in mapping tweets to news items (40.3%), challenges with Turkish language complexity, and potential overfitting with bigram and trigram models.
conference paper,Eighth International Conference on Social Network Analysis, Management and Security (SNAMS)	2022	IMDB dataset which consists of 50 thousand movie reviews (25 thousand positive reviews and 25 thousand negative reviews)	Machine Learning Classifiers,Ensemble Method,Data Preprocessing Techniques,Vectorization Technique	Logistic Regression (LR) Naïve Bayes (NB) XGBoost (XGB) Random Forest (RF) Multilayer Perceptron (MLP)	89.9%	demonstrating superior performance compared to individual classifiers and existing methods.
Sentiment Analysis in Social Media	2019	Twitter Data,instagram Data,Facebook	Lexicon-Based Methods,Machine Learning-Based Methods,Preprocessing Methods,Ensemble	Probabilistic Models,Text Representat	85%	while accuracy provides a broad measure of effectiveness, it can be

and Its Application		Data,Reddit Data	Methods	ion Techniques, Ensemble Methods		misleading in cases of class imbalance or when the dataset contains complex, informal language.
Sentiment analysis using product review data	2015	the dataset used in the study by Fang and Zhan consists of over 5.1 million product reviews collected from Amazon.com. ,Over 5.1 million Amazon product reviews.	source machine learning soft-ware package in Python. The classification models selected for categorization are: Naïve Bayesian, Random Forest, and Support Vector Machine	Support Vector Machine,,Random Forest Classifier:Bayes' theorem, assuming independence between features	85%	Random Forest is the most effective classifier for sentiment polarity categorization, with future work needed to address review-level classification challenges and implicit sentiment detection.
Social Media Sentiment Analysis	2021	Twitter Data,instagram Data,Facebook Data,Reddit Data	NumPy,pandas,Matplotlib,seaborn,NLTK	Logistic Regression, Bernoulli Naïve Bayes,Random Forest Regression, ultinomial Naïve Bayes classifier	88.90%	This can be used with a graphic interface to deliver the same results with
journal paper published in the "International Journal	2024	Amazon Musical Instruments Reviews Dataset, which is	Convolutional Neural Network (CNN),Natural Language Processing (NLP) with SpaCy,Data Preprocessing,Data	TextCategorizer component from the SpaCy NLP library….This	94%	Class Imbalance,Overfitting,Limited Dataset Size,Simple Binary Classification

of Research in Circuits,		available on Kaggle. The dataset contains 10,262 rows and 9 columns	Visualization,Evaluation Metrics	component is integrated with a Convolutional Neural Network (CNN), which classifies the review text into positive or negative sentiment.		
Sentimental Analysis of Twitter Users from Turkish Content with Natural Language Processing	2022	public dataset from Beyaz (2021), **Custom Dataset (SentimentSet)**: Researchers manually created this dataset by collecting Turkish tweets	**Natural Language Processing (NLP)** techniques for sentiment analysis.	**Zemberek** and **NLTK Snowball**	87%	1.Turkish Language Nuances 2.Agglutinative Nature
A feature fusion and detection approach using deep learning for sentimental analysis and offensive	2024	dataset for three under-resourced Dravidian languages (Tamil, Kannada, and Malayalam) generate	1.ALBERT Tokenization 2.Feature Extraction 3.Feature Fusion 4.Sentimental Analysis and Offensive Text Identification:	To design an approach for sentiment analysis and offensive text detection from a code-mix language using HAN.	accuracy attained by the proposed model is 0.956	Offensive text detection include handling noisy data, language variations, and context ambiguity

text detection from code-mix Malayalam language		d from social media comments The dataset contains more than 60,000				
Sentimental analysis of Facebook reviews: Does hospitality matter in senior living?	2023	the official Facebook pages of 125 senior living communities in the U.S. These communities were not randomly selected; they represented leading companies in the U.S. senior living industry.	1. text mining and 2.sentiment analysis techniques to extract insights.	techniques, including: 1.Multivariate Linear Regression 2.Random Forest 3.Support Vector Machine Regression	Operate nearly 75% of professionally managed communities in the industry (Argentum, 2021).	Generalizability: The study's findings should not be generalized widely due to self-selection bias in Facebook reviews. Some individuals without internet access or awareness of Facebook reviews may be excluded Comparative Analysis
Journal of Information Processing and Management	2023	IMDb Movie Reviews dataset	Long Short-Term Memory (LSTM) networks	LSTM-based	87%	Limited to English language reviews, does not perform well with short texts.

CHAPTER 3

PROPOSED SYSTEM

3. PROPOSED SYSTEM

The methodology will include sentiment analysis and evaluation of a ma- chine learning model on text data regarding reviews of musical instruments including data preprocessing, feature extraction, training, and comparison.

3.1 Data Pre-processing:

We carried out a number of data pre-processing procedures, which are described below, in order to get the dataset ready for sentiment analysis: Handling Miss- ing data: The reviewerName and review have missing values. An empty string has been used in the text field for consistency and to avoid any errors during text processing. Preprocessing: Text Cleaning Normalizing the Textual Material Re- moving extra and redundant characters, conversion of text to lowercase, removal of stop words, tokenization, and stemming are the default text pre-processing function that we utilized. These steps in text pre-processing normalize the textual material and even reduce the number of vocabularies that might be subjected to proper analysis. The processed text was then fed into a new column called stemmed content. Sentiment Labelling: Four classes of sentiment were identified from the rating of the review (overall column): Positive (ratings 4 and 5), Neutral (rat- ing 3 and), Negative (ratings 1 and 2). This transformation was conducted using a custom method with the sentiment class of every review stored in a new column called Sentiment.

3.2 Data Augmentation:

This means that we put the stratified sampling methodology to good use in solving the class imbalance anomaly identified with our sentiment analysis data set. Class imbalance is a problem in multi-class classification wherein certain classes dominate the dataset, so if left unchecked, this would have significantly affected model performance. For this data set, we had the three classes of sentiment: neutral, negative, and positive. The given dataset is balanced after the data augmentation process. Since all the nine classes of the data are equal in size, the overall dataset is balanced as well. Therefore, it would improve the performance of the model if this balanced data is to be tested on it. This will perform better with data augmentation. Data augmentation is useful in overfitting cases, or those where the train data is fewer.

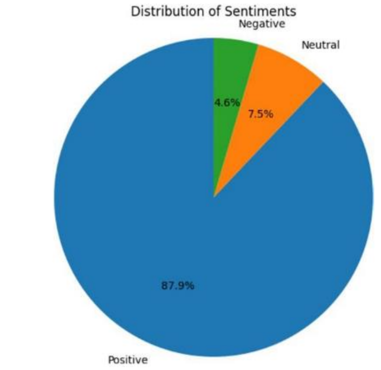

Fig. 1. Data distribution over 3 classifications

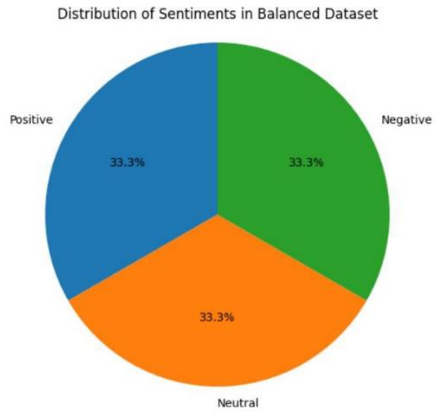

Fig. 2. After augmentation distribution of data over 3 classes

3.3 Feature Extraction:

One of the most important phases in improving context understanding is fea- ture extraction. The text needs to be transformed into usable features after proper cleaning and preparation for modelling. This method primarily aims at finding characteristics that positively affect the results of the categorization. In our study, we have applied two feature extraction machine learning methods, namely Bag-of-Words model and TF-IDF. (TF-IDF): We applied the technique TF-IDF to emphasize more relevant words and de-emphasize stop words. This technique gave a weight to each phrase based on how many times it appeared in any given document relative to how many times it occurred throughout the entire corpus. While doing so, the technique of TF-IDF highlights especially revealing words for the mood. Term Frequency: tf(t,d) =mount of words in d, count of t in d Document Frequency: df(t) = incidence of t in written records Inverse Document Frequency: df(t) = N(t) where df(t) = Document frequency of a term t N(t) = total number of documents that contain the term t Therefore, tfidf(t, d) = tf(t, d) * idf(t) The frequency of a phrase in a given document is determined by TF. It is only the total words in a text divided by the frequency with which a phrase occurs in that text. Often occurring terms in a manuscript are underlined prominently. IDF determines a term's rarity across a set of texts. Then, there would be penalties for terms that appear in every document. IDF and TF combined to form TF-IDF. To determine a term’s TF-IDF score in a document, calculate each term’s TF and IDF scores and multiply them together.

3.4 Model Selection:

We assessed many machine learning algorithms to cate- gorize sentiment into positive, negative, and neutral groups during the Model Selection stage. The following algorithms were taken into consideration: Logistic Regression: Very intuitive and familiar, applicable to multi class problems. SVMs: Function optimally even when the dimensions are very large. Random Forest: It is an ensemble technique that reduces overfitting and increases accuracy. Naïve Bayes: Good, feature- independent method for text categorization.

To test the performance of the model, we applied k-fold cross-validation testing over accuracy, precision, recall, and F1-score metrics. The implementation of the models was relatively easy using the scikit-learn toolkit, and the best model was further tuned for deployment purposes.

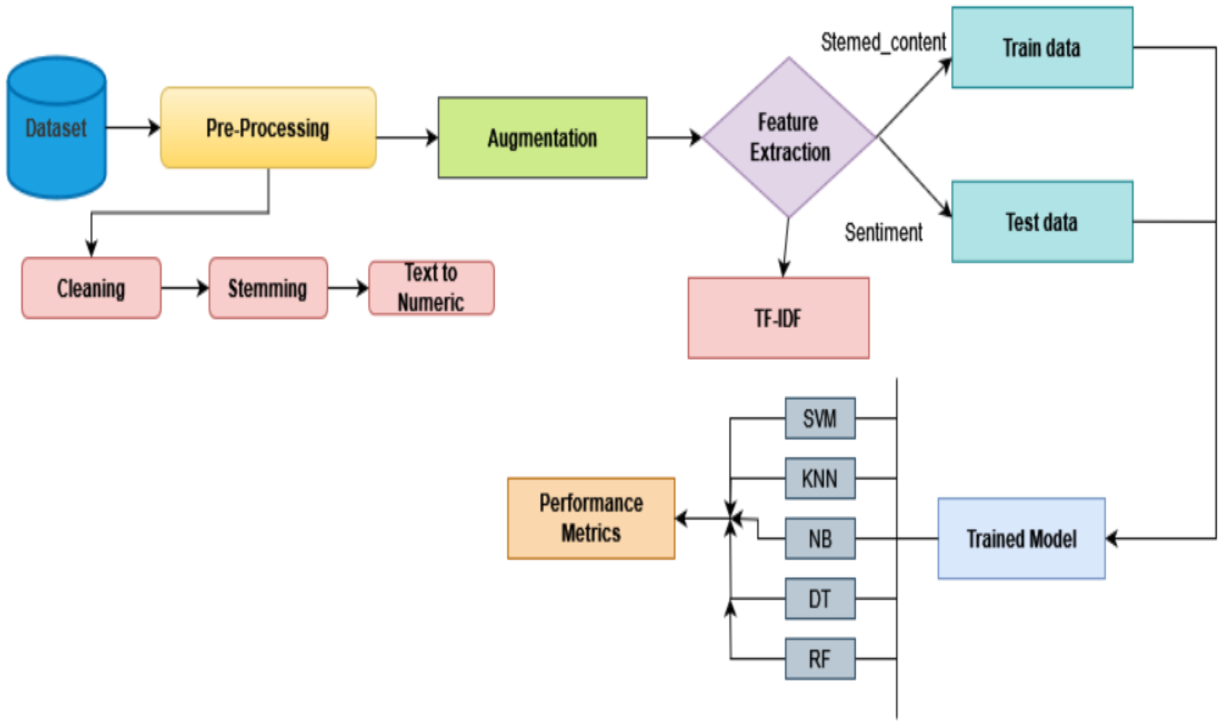

Fig. 3. Proposed model flowchart

3.4.1 Decision Tree Classifier:

Learning, they are mainly applied for classification problems, but this is not to say that regression problems cannot be solved with it. This is a tree structured classifier with internal nodes substituting for attributes of the dataset, whereas the nodes represent decision rules and leaf nodes represent each outcome.

F1 Score	Accuracy	Recall	Precision
0.8127	0.81	0.8122	0.8122

Fig. 4. Prediction Metrics for Decision tree

The accuracy produced is 81.9 when used with TF-IDF. Using hyperparameter tuning and cross-validation, a high accuracy of 96.00 was achieved, the highest accuracy in this investigation. Table 4 displays the prediction metrics for cross- validation.

3.4.2 Light GBM:

It is developed by Microsoft, an open- source, distributed, high-performance gradient boosting frame- work in focus on accuracy, scalability, and efficiency. Predicated on decision trees, it decreases memory utilization and enhances the efficacy of the model. In addition, LightGBM generates trees through methods involving histograms, which happens to be more efficient than all the other competing gradient boosting frameworks. These optimization methods give LightGBM an advantage, making it more efficient and more superior to other competing gradient boosting frameworks. Some other competitive optimizations include leaf-wise tree development and efficient data storage formats.

F1 Score	Accuracy	Recall	Precision
0.8404	0.8816	0.8816	0.8374

Fig. 5. Prediction Metrics for Light GBM

The accuracy produced is 88.16 when used with TF-IDF. Using hyperparameter tuning and cross-validation, a high accuracy of 95.33 was achieved, the highest accuracy in this investigation.

3.4.3 SVM(Support Vector Machine):

The support vector machine is a machine learning algorithm that uses either linear or nonlinear classification, strong regression, and outlier identification. Text classification, picture

F1 Score	Accuracy	Recall	Precision
0.8224	0.8787	0.8787	0.7729

classification, handwriting identification, face detection, spam detection, gene expression analysis, and anomaly detection are a few applications of SVM. Because SVMs can learn in a high-dimensional space and take nonlinear interactions into account, they are widely used. The accuracy produced is 87.8% when used with TF-IDF.

Fig. 6. Prediction Metrics for SVM

Using hyperparameter tuning and cross-validation, a high accuracy of 97.33% was achieved, the highest accuracy in this investigation. Table 6 displays the prediction metrics for cross-validation.

3.4.4 Gradient Boosting model:

In gradient boosting, a new model is learned at each step to minimize the loss function, which could be the previous model's cross-entropy or mean squared error. Gradient boosting, then, is a very effective boosting technique that turns a lot of weak learners into strong learners. Using the predictions from the current ensemble, the procedure calculates the gradient of the loss function at each iteration and trains a new weak model to maximize this gradient. Once the new model's predictions are included in the ensemble, the process is repeated until the stopping condition is met.

F1 Score	Accuracy	Recall	Precision
0.8348	0.8821	0.8821	0.8506

Fig. 7 Prediction Metrics for Gradient boosting

The accuracy produced is 88.2% when used with TF IDF. Using hyperparameter tuning and cross validation, a high accuracy of 96.00% was achieved, the highest accuracy in this investigation. Table 7 displays the prediction metrics for cross- validation.

3.4.5 Logistic Regression:

It applies algorithms, such as logistic regression, which is one of the most applied machine learning techniques. The approach uses an already existing set of independent factors when making a prediction of the category dependent variable. The accuracy produced is 87.8% when used with TF-IDF.

F1 Score	Accuracy	Recall	Precision
0.8271	0.8787	0.8787	0.8477

Fig. 8. Prediction Metrics for Logistic Regression

Using hyperparameter tuning and cross-validation, a highaccuracy of 97.33% was achieved, the highest accuracy in this investigation. fig 8 displays the prediction metrics for cross- validation

CHAPTER-4
IMPLEMENTATION

4. Implementation

The implementation phase covers the practical application of the proposed stock price prediction system, including setting up the environment, processing the data, and executing the models. The following sections detail the steps required for implementing stock price prediction using machine learning.

4.1 Environment Setup

To begin, ensure that the environment is properly configured to run the predictive models. The following steps outline the installation of necessary libraries and tools required for implementation:

1. **Programming Language**: The implementation is carried out using Python, a popular language for machine learning.
2. **Libraries**:
   - Pandas: For data manipulation and preprocessing.
   - NumPy: For numerical computations.
   - Scikit-learn: For implementing machine learning models.
   - Matplotlib: For data visualization.
3. **Installation**: Install the required libraries using pip:
   ```
   pip install pandas numpy scikit-learn matplotlib
   ```
4. **Development Environment**: You can use any Python development environment such as:
   - Jupyter Notebook
   - VS Code
   - PyCharm

4.2 Sample Code for Preprocessing and Model Operations

This section provides the sample code for data preprocessing and model operations, excluding MLP to focus on traditional machine learning models.

1. **Data Preprocessing**:

- **Load the Dataset**:

```
import pandas as pd
# Load the dataset
data = pd.read_csv('/content/drive/MyDrive/stock price2.csv')
```

- **Handle Missing Values**:

```
# Fill missing values with the mean of each column
data.fillna(data.mean(), inplace=True)
```

- **Handling Duplicates**:

```

# Find all duplicate rows except the first occurrence
duplicates = df[df.duplicated()]
print("Duplicate Rows:\n", duplicates)

duplicate_count = df.duplicated().sum()
print(f"Number of duplicate rows: {duplicate_count}")
```

- **Data Splitting**:

```
from sklearn.model_selection import train_test_split
```

```
# Features and target variable (assuming 'Close' is the target)
X = df_no_outliers.drop('Close', axis=1)  # Features
y = df_no_outliers['Close']  # Target

# Split the data into training and testing sets (80% training, 20% testing)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Print the shape of the training and testing sets
print("X_train shape:", X_train.shape)
print("y_train shape:", y_train.shape)
print("X_test shape:", X_test.shape)
```

1. **Model Building and Training**: The following is a sample of how to implement and train different machine learning models for predicting calories burned.

- **Linear Regression**:

```
from sklearn.linear_model import LinearRegression
# Initialize and train the model
lr_model = LinearRegression()
lr_model.fit(X_train, y_train)
```

- **Random Forest**:

```
from sklearn.ensemble import RandomForestRegressor
# Initialize and train the Random Forest model
rf_model = RandomForestRegressor(n_estimators=100, random_state=42)
rf_model.fit(X_train, y_train)
```

- **Gradient Boosting**:

```
from sklearn.ensemble import GradientBoostingRegressor
# Initialize and train the Gradient Boosting model
gb_model = GradientBoostingRegressor(n_estimators=100, learning_rate=0.1,
random_state=42)
gb_model.fit(X_train, y_train)
```

2. **Model Selection and Prediction**: After evaluating the models, choose the one with the best performance metrics and use it for predicting calories burned on new data.

- **Prediction Example**:

```
# Predict calories burned using the best model
best_model = rf_model  # Assuming Random Forest performed the best
new_data = [[85, 6, 30]]  # Example input: heart rate, MET, duration
scaled_data = scaler.transform(new_data)
predicted_calories = best_model.predict(scaled_data)
```

CHAPTER 5

METRICS

5. METRICS

The performance of our ensemble model is measured ac- cording to several essential metrics, which are spelled out in the section that follows:

4.1 Confusion Matrix:

This matrix is also responsible for documenting how accu- rate and incorrect the various predictions made by the model were when the evaluation was done on the test data. The Confusion Matrix is widely used to measure classification models, where it is used to predict a categorical label for each input instance. Some of the named quantities in it include True Positive(TP), True Negative(TN), False Positive(FP), and False Negative(FN).

4.2 Precision:

Precision is a measure of how well a model predicts favourable outcomes. The ratio of all the model's positive predictions to the actual positive forecasts is its definition.

$$\text{Precision} = A / A + B$$

4.3 Recall:

Recall measures a classification model's ability to identify each relevant instance in a dataset. The percentage of true positive (TP) cases relative to the total number of false negative (FN) and true positive (TP) cases is what matters.

$$\text{Recall} = A / A + C$$

4.4 F1-Score

The F1-score is used to evaluate a classification model's overall performance. The harmonic mean of precision and recall is what it is.

$$\text{F1 Score} = 2 \cdot \text{Precision} \cdot \text{Recall} / \text{Precision} + \text{Recall}$$

4.5 Accuracy

The performance of the model is evaluated based on its accuracy. It is computed as the ratio of all instances to all accurate occurrences.

Accuracy = A + D /A + D + B + C

VARIABLE DEFINITIONS

A = Correctly Predicted Positives (TP)

B = Incorrectly Predicted Positives (FP)

C = Missed Positives (FN)

D = Correctly Predicted Negatives (TN)

CHAPTER-6

RESULTS AND DISCUSSIONS

6. RESULTS AND DISCUSSIONS

This research study demonstrated that SVM, Na¨ıve Bayes, K-NN, LightGBM, Logistic Regression will get the best result on this dataset with accuracy ranges above 90+%. The essential aim of this research has been such that most reviews concerning the products should be correctly categorized and analyzed to let us determine which products customers love or detest. Therefore, we will now compare the classifiers that we used in our implementation with the ones already in use in this sector. Models which generalize well will be good and include logistic regression and SVM[. The model found to be overgeneralizing and the model reached the highest training accuracy.

TABLE 2

PERFORMANCE METRICS OF THE PROPOSED MODELS

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Decision tree	0.96	0.81	0.81	0.81
Light GBM	0.95	0.856	0.88	0.83
Gradient Boosting	0.96	0.85	0.88	0.83
Logistic regression	0.97	0.84	0.87	0.82
SVC	0.97	0.77	0.87	0.82
KNN	0.96	0.77	0.87	0.82

A comparison of the results from the second set in comparison with the first set from all three models shows an improvement. The largest change is found in Logistic Regression (from 0.88% to 0.97%).In the second set, KNN and Decision Tree are equally good at achieving an accuracy of 96%, with a slight edge to Decision Tree for the first set. The probable adjustments to model settings, data quality, feature engineering, and training protocols were done in light of the significant increase in accuracy achieved with the second set.

CHAPTER-7

CONCLUSION

7. CONCLUSION

The investigation concludes that feature engineering and preprocessing have a major influence over the performance of the machine learning model. Among all the used models, the best generalization skills were presented by the Gradient Boosting and LightGBM models, which are quite good for this kind of issue. In the future, more advanced NLP techniques can be further used to improve accuracy in sentiment predictions, ensemble learning to get better models, and hyperparameter tuning to optimize better.

REFERENCES

[1] Amazon Musical Instruments Reviews Dataset for musical instruments review and ratings in Kaggle (https://www.kaggle.com/eswarchandt/amazon-music-reviews) .

[2] Chattopadhyay, A., Basu, M. (2022). Unsupervised Learning Based Brand Sentiment Mining using Lexicon Approaches–A Study on Amazon Alexa. Indian Journal of Data Mining (IJDM), 2(1), 15-20.

[3] Sa´nchez-Franco, M. J., Arenas-Ma´rquez, F. J., Alonso-Dos-Santos, M. (2021). Using structural topic modelling to predict users' sentiment towards intelligent personal agents. An application for Amazon's echo and Google Home. Journal of Retailing and Consumer Services, 63, 102658.

[4] Khurana, P., Gupta, B., Sharma, R., Bedi, P. (2024). A sentiment-guided session-aware recommender system. The Journal of Supercomputing, 1- 40.

[5] Vishwakarma, S., Garg, D., Choudhury, T., Singh, T. P. (2023, December). Amazon Sales Sentiment Prediction and Price Forecasting Using Facebook Prophet. In International Conference on Cyber Intelligence and Information Retrieval (pp. 93-105). Singapore: Springer Nature Singapore.

[6] Marwat, M. I., Khan, J. A., Alshehri, D. M. D., Ali, M. A., Ali, H., Assam, M. (2022). Sentiment analysis of product reviews to identify deceptive rating information in social media: A sentideceptive approach. KSII Transactions on Internet and Information Systems (TIIS), 16(3), 830-860.

[7] Amorim, I. B. M. (2023). Using sentiment analysis to predict Amazon ratings: a comparative study using dictionaries approaches (Doctoral dissertation).

[8] Wankhade, Mayur, Annavarapu Chandra Sekhara Rao, and Chaitanya Kulkarni. "A survey on sentiment analysis methods, applications, and challenges." *Artificial Intelligence Review* 55, no. 7 (2022): 5731-5780.

[9] Taboada, Maite. "Sentiment analysis: An overview from linguistics." *Annual Review of Linguistics* 2, no. 1 (2016): 325-347.

[10] Hussein, Doaa Mohey El-Din Mohamed. "A survey on sentiment analysis

challenges." *Journal of King Saud University-Engineering Sciences* 30, no. 4 (2018): 330-338.

[11] Shah, Mrs Snehal, Miss Akshata Bhat, Miss Sumitra Singh, Miss Arya Chavan, and Master Aryan Singh. "Sentiment analysis." (2010).

[12] Devika, M. Dª, Cª Sunitha, and Amal Ganesh. "Sentiment analysis: a comparative study on different approaches." *Procedia Computer Science* 87 (2016): 44-49.

[13] Serrano-Guerrero, Jesus, Jose A. Olivas, Francisco P. Romero, and Enrique Herrera-Viedma. "Sentiment analysis: A review and comparative analysis of web services." *Information Sciences* 311 (2015): 18-38.

[14] Stine, Robert A. "Sentiment analysis." *Annual review of statistics and its application* 6, no. 1 (2019): 287-308.

[15] Madhoushi, Zohreh, Abdul Razak Hamdan, and Suhaila Zainudin. "Sentiment analysis techniques in recent works." In *2015 science and information conference (SAI)*, pp. 288-291. IEEE, 2015.

[16] Hartmann, Jochen, Mark Heitmann, Christian Siebert, and Christina Schamp. "More than a feeling: Accuracy and application of sentiment analysis." *International Journal of Research in Marketing* 40, no. 1 (2023): 75-87.

[17] Pawar, Anil Bhausaheb, M. A. Jawale, and D. N. Kyatanavar. "Fundamentals of sentiment analysis: concepts and methodology." *Sentiment analysis and ontology engineering: An environment of computational intelligence* (2016): 25-48.

[18] Rojas-Barahona, Lina Maria. "Deep learning for sentiment analysis." *Language and Linguistics Compass* 10, no. 12 (2016): 701-719.

[19] Das, Ringki, and Thoudam Doren Singh. "Multimodal sentiment analysis: a survey of methods, trends, and challenges." *ACM Computing Surveys* 55, no. 13s (2023): 1-38.

[20] Agarwal, Apoorv, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca J. Passonneau. "Sentiment analysis of twitter data." In *Proceedings of the workshop on language in social media (LSM 2011)*, pp. 30-38. 2011.

[21] Wang, Yequan, Aixin Sun, Jialong Han, Ying Liu, and Xiaoyan Zhu. "Sentiment analysis by capsules." In *Proceedings of the 2018 world wide web conference*, pp. 1165-1174. 2018.

[22] Ravi, Kumar, and Vadlamani Ravi. "A survey on opinion mining and sentiment analysis: tasks, approaches and applications." *Knowledge-based systems* 89 (2015): 14-46.

[23] Tang, Duyu, Bing Qin, and Ting Liu. "Deep learning for sentiment analysis: successful approaches and future challenges." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 5, no. 6 (2015): 292-303.

[24] Zhu, Linan, Zhechao Zhu, Chenwei Zhang, Yifei Xu, and Xiangjie Kong. "Multimodal sentiment analysis based on fusion methods: A survey." *Information Fusion* 95 (2023): 306-325.

[25] Kolkur, Seema, Gayatri Dantal, and Reena Mahe. "Study of different levels for sentiment analysis." *International Journal of Current Engineering and Technology* 5, no. 2 (2015): 768-770.