Cyberbullying Detection on Twitter using Sentiment Analysis

Project Report submitted in the partial fulfilment

of

Bachelor of Technology

In

Computer Science and Business Systems

By

Avantika Jalote	**A061**
Karim Khan	**A069**
Kevin Mathew	**A068**
Mansi Nandkar	**A058**

Under the supervision of

Dr. Sandip Bankar

(Assistant Professor, Computer Engineering, STME)

SVKM's NMIMS University

(Deemed to be university)

SCHOOL OF TECHNOLOGY MANAGEMENT & ENGINEERING (STME)

Kharghar, Navi Mumbai - 410210

(2023-2024)

<u>CERTIFICATE</u>

This is to certify that the project entitled **"Cyberbullying Detection on Twitter using Sentiment Analysis"**, has been done by **Ms. Avantika Jalote, Mr. Karim Sohail Khan Patan, Mr. Kevin Mathew and Ms. Mansi Nandkar** under the guidance and supervision of **Dr. Sandip Bankar** & has been submitted in partial fulfilment of the degree of Bachelors in Technology in Computer Science and Business Systems of STME, SVKM's NMIMS (Deemed-to-be University), Navi Mumbai, India.

Project mentor (Internal Guide)

Examiner

Date:

Place: <u>Navi Mumbai</u>

(HoD)

ACKNOWLEDGEMENT

We express our heartfelt gratitude to our esteemed project mentor, **Dr. Sandip Bankar**, whose invaluable guidance, insightful feedback, and unwavering support have been instrumental throughout the duration of this project. His expertise, encouragement, and dedication have truly enriched our learning experience and empowered us to overcome challenges with confidence. We are deeply thankful for his mentorship and for sharing his expertise generously, which has undoubtedly contributed to the success of our final year project.

We extend our sincere appreciation to **Prof. Aditya Kasar**, the Head of the Department, for his constant encouragement, support, and belief in our abilities. His leadership and vision have provided us with a conducive environment to explore and innovate, fostering our academic and personal growth. We are grateful for his valuable insights, encouragement, and for always being approachable and supportive throughout our journey.

Furthermore, we would like to extend our gratitude to all the faculty members, staff, and the entire college administration for their support and provision of necessary infrastructure throughout the course of our project. Their collective efforts and dedication have created an enriching academic environment that has nurtured our intellectual curiosity and facilitated our learning and research endeavors. We are deeply thankful for their assistance, encouragement, and for fostering an atmosphere conducive to holistic development.

NAME	ROLL NO	SAP ID
Avantika Jalote	A061	70362019230
Karim Sohail Khan Patan	A069	70362019346
Kevin Mathew	A068	70362019345
Mansi Nandkar	A058	70362019221

ABSTRACT

Twitter is grappling with a significant and pervasive issue: cyberbullying. This form of online harassment poses a serious threat to users' well-being, necessitating a review that delves into its devastating impact on victims. The report also underscores the challenges posed by the ever-evolving tactics of online harassers, highlighting the urgent need for effective measures to address this problem.

One proposed solution is the implementation of a robust cyberbullying detection system, leveraging sentiment analysis. Such a system aims to proactively identify and address instances of cyberbullying, thereby empowering both Twitter users and the platform itself to foster a safer online environment. It advocates for a comprehensive comparative analysis of feature selection methods (Bag of Words, Term Frequency-Inverse Document Frequency (TF-IDF), and Count Vectorizer) commonly used in text classification tasks along with machine learning models (Naive Bayes, Random Forest and Support Vector Machine).

The best accuracy was 92.98% given by TF-IDF with SVM model which will be proposed further for the real-time prediction of tweets. This holistic assessment aims to offer insights into their relative effectiveness in cyberbullying detection, informing the selection of the most suitable method based on specific contextual requirements. Such a comparative study methodology will contribute valuable knowledge to the field and facilitate the development of more accurate and reliable cyberbullying detection systems.

TABLE OF CONTENTS

List of Figures......ii
List of Tables...... iii
Abbreviations iv
Chapter 1 Introduction......1
1.1 Background of the project topic......1
1.2 Motivation and scope of the report......2
1.3 Problem statement......2
1.4 Objectives 3
1.5 Salient contribution......3
1.6 Organization of report......4
Chapter 2 Literature Survey......5
Chapter 3 Methodology and Implementation 10
3.1 Block Diagram......10
3.2 Data Collection 11
3.3 Data Preprocessing......12
3.4 Balancing the data......13
3.5 Feature Extraction......13
3.6 Machine Learning Models 14
3.7 Deployment......16
Chapter 4 Result and Analysis......18
Chapter 5 Advantages, Limitations and Applications......41
5.1 Advantages......41
5.2 Limitations......41
5.3 Applications 41
Chapter 6 Conclusion and Future Scope 42
6.1 Conclusion 42
6.2 Future Scope 42
References......44
Appendix A: List of papers presented and published......46

List of Figures

Fig No	Name of the figure	Page No
3.1	Proposed Architecture	10
3.2	Data Preprocessing	12
3.3	Deployment	16
4.1	Classification report for CV with RF	20
4.2	Confusion matrix for CV with RF	20
4.3	ROC curve for CV with RF	21
4.4	Classification report for Count Vectorizer and SVM	22
4.5	Confusion matrix for Count Vectorizer and SVM	22
4.6	ROC curve for Count Vectorizer and SVM	23
4.7	Classification report for Count Vectorizer and Naïve Bayes	24
4.8	Confusion matrix for Count Vectorizer and Naïve Bayes	24
4.9	ROC curve for Count Vectorizer and Naïve Bayes	25
4.10	Classification report for TF-IDF and Random Fores	26
4.11	Confusion matrix for TF-IDF and Random Forest	26
4.12	ROC curve for TF-IDF and Random Forest	27
4.13	Classification report for TF-IDF and SVM	28
4.14	Confusion matrix for TF-IDF and SVM	28
4.15	ROC curve for TF-IDF and SVM	29
4.16	Classification report for TF-IDF and Naïve Bayes	30
4.17	Confusion matrix for TF-IDF and Naïve Bayes	30
4.18	ROC curve for TF-IDF and Naïve Bayes	31
4.19	Classification report for BoW using n-grams and Random Forest	32
4.20	Confusion matrix for BoW using n-grams and Random Forest	32
4.21	ROC curve for BoW using n-grams and Random Forest	33
4.22	Classification report for BoW using n-grams and SVM	34
4.23	Confusion matrix for BoW using n-grams and SVM	34
4.24	ROC curve for BoW using n-grams and SVM	35
4.25	Classification report for BoW using n-grams and Naïve Bayes	36
4.26	Confusion matrix for BoW using n-grams and Naïve Bayes	36
4.27	ROC curve for BoW using n-grams and Naïve Bayes	37
4.28	Accuracy Scores for all combinations	38
4.29	Precision Scores for all combinations	38
4.30	Recall Scores for all combinations	39
4.31	F1 Scores for all combinations	39

List of Tables

Table No	Name of the table	Page No
2.1	Literature Survey	
4.1	Comparison table	

Abbreviations

SVM	Support Vector Machine
RF	Random Forest
NB	Naïve Bayes
TF-IDF	Term Frequency-Inverse Document Frequency
BoW	Bag of Words
NLP	Natural Language Processing
SMOTE	Synthetic Minority Over-sampling Technique
AUC	Area Under Curve

Chapter 1

Introduction

1.1 Background of the project topic

Cyberbullying, a pervasive issue in today's digital age, poses significant challenges, particularly on social media platforms such as Twitter. The anonymity and ease of communication provided by online platforms have amplified the prevalence of cyberbullying, enabling individuals to engage in harmful behaviour without the fear of immediate consequences. This has led to a proliferation of instances involving harassment, intimidation, and abuse, targeting individuals based on various factors such as race, gender, sexual orientation, or differing opinions. The consequences of cyberbullying can be severe, leading to emotional distress, psychological harm, and in extreme cases, even self-harm or suicide [1]. Thus, addressing cyberbullying on Twitter has become a critical societal concern.

The impact of cyberbullying on individuals, including minors and vulnerable groups, cannot be overstated. Victims of cyberbullying often experience profound emotional and mental anguish, leading to decreased self-esteem, anxiety, depression, and feelings of isolation. Moreover, cyberbullying can have long-lasting effects on an individual's mental well-being, hindering their ability to engage in online social interactions freely. The pervasive nature of cyberbullying makes it imperative for stakeholders, including social media platforms like Twitter, policymakers, educators, and technology experts, to collaborate in finding effective solutions to combat this issue.

Despite the efforts made by platforms like Twitter to address cyberbullying through reporting mechanisms, community guidelines, and content moderation, the problem persists. This underscores the need for innovative and proactive approaches that leverage technology and data analysis to identify and mitigate instances of cyberbullying effectively [2]. By harnessing the power of machine learning, natural language processing, and sentiment analysis, it becomes possible to develop automated systems capable of detecting, monitoring, and intervening in cyberbullying incidents in real-time. Such solutions not only enable swift response to cyberbullying but also empower users to navigate online spaces safely and confidently.

1.2 Motivation and scope of the report

The motivation behind this project stems from the pressing need to address the pervasive issue of cyberbullying on Twitter and other social media platforms. As online communication continues to evolve, so do the methods and tactics employed by perpetrators of cyberbullying. The anonymity and accessibility of social media platforms have exacerbated the problem, making it increasingly challenging to detect and combat instances of cyberbullying effectively. By developing a robust system for detecting and classifying cyberbullying using sentiment analysis and machine learning techniques, this project aims to provide a proactive and automated approach to addressing this critical societal issue.

The scope of this report is centred around the implementation and evaluation of various sentiment analysis algorithms, machine learning models, and data preprocessing techniques tailored specifically for detecting cyberbullying on Twitter. This includes the exploration of different feature extraction methods, such as TF-IDF, Count Vectorizer and BoW using n-grams, to capture the nuanced language patterns indicative of cyberbullying behaviour. Additionally, the project will make use of machine learning models, including classifiers such as SVM, Naive Bayes, and RF, to effectively classify tweets into cyberbullying and non-cyberbullying categories.

Furthermore, the project's scope extends to the integration of real-time data collection and processing capabilities, allowing for timely interventions and responses to cyberbullying incidents as they occur. This involves the deployment of scalable and efficient data pipelines to ingest, preprocess, and analyse large volumes of Twitter data in near real-time. Additionally, the project will explore the implementation of automated alerting and reporting mechanisms to notify relevant stakeholders, including platform moderators and law enforcement agencies, of potential cyberbullying incidents requiring immediate attention.

1.3 Problem statement

Despite existing reporting mechanisms and policies, the sheer volume of tweets and the nuanced nature of cyberbullying make manual detection and intervention impractical.

Several machine learning models and feature extraction techniques have been tried out but no comprehensive comparative study has been presented. The aim is to compare three feature extraction techniques: TF-IDF, BoW, Count Vectorizer and three machine learning models: Naïve Bayes, Random Forest and SVM and present the combination which performs the best.

1.4 Objectives

- To detect cyberbullying by implementing sentiment analysis.
- Explore various feature extraction methods and evaluate the impact of different techniques on model performance.
- Investigate a range of machine learning models and assess the performance of each model in terms of accuracy, precision, recall, F1-score, and computational efficiency.
- Conduct a comprehensive comparative analysis of the performance of models against individual machine learning models and feature extraction methods.

1.5 Salient contribution

This project seeks to develop a system that can automatically detect and classify cyberbullying instances based on the sentiment and its content. This system can contribute significantly to the present online systems.

1. Timely Intervention: The system enables early detection of cyberbullying incidents on Twitter, allowing for prompt intervention to prevent escalation and provide support to victims.
2. Enhanced Accuracy: Leveraging machine learning models like RF and SVM, the system achieves high accuracy in classifying tweets as cyberbullying or non-cyberbullying, reducing the risk of misclassification and ensuring reliable detection.
3. Scalability and Efficiency: The automated nature of the system, coupled with its ability to handle large volumes of tweets, ensures scalability and efficiency in monitoring and detecting cyberbullying behaviour across the platform.

4. Real-time Alerts and Notifications: The system generates real-time alerts and notifications when cyberbullying is detected, enabling swift action by users, moderators, or platform administrators to address the issue promptly.
5. Data-driven Insights and Continuous Improvement: By analysing patterns and trends in cyberbullying data, the system provides valuable insights for understanding the dynamics of online harassment on Twitter. Additionally, through continuous learning and adaptation, the system improves its detection capabilities over time to stay ahead of evolving cyberbullying tactics.

1.6 Organization of report

Chapter 1 gives an insight into the background, motivation and scope of the project. It also contains the problem statement and the salient contributions to the field. A review of the research done until now is included in chapter 2. Chapter 3 deals with the methodology and implementation of different models. Result and Analysis are discussed in Chapter 4. The advantages, limitations and applications of the project are stated in chapter 5. This is followed by the conclusion and future scope, references and appendices.

Chapter 2

Literature Survey

The literature surrounding cyberbullying detection on social media platforms, particularly Twitter, is extensive and multifaceted, reflecting the growing concern over the prevalence and impact of online harassment and abuse. Researchers have employed various methodologies, including sentiment analysis, machine learning techniques, and deep learning models, to develop effective cyberbullying detection systems. In this literature review, we explore several seminal papers that contribute to the advancement of cyberbullying detection research. These papers offer insights into different approaches, algorithms, and technologies utilized to identify and mitigate cyberbullying behaviours on Twitter. Examining the findings and methodologies of each paper, we gain a comprehensive understanding of the current state of cyberbullying detection research and the challenges and opportunities in this evolving field.

The analysis of different types of cyberbullying, as discussed by B. R. Prathap et al. [1], involved gauging shifts in public sentiment towards criminal incidents and identifying emotions associated with various categories of unlawful acts. This approach provides a nuanced understanding of cyberbullying behaviours and their impact on social dynamics. The paper's failure to provide specific implementation and data analysis details raises concerns about reproducibility and hampers comprehensive evaluation of the methodology and results, indicating the need for improved transparency.

S.A. Mathur et al. [2] leveraged various machine learning classifiers such as Random Forest, gradient boosting, AdaBoost, among others, to analyse the scraped data using Selenium. By applying these classifiers, researchers aimed to extract meaningful insights from Twitter data, potentially uncovering patterns related to cyberbullying and other social phenomena. Random Forest's remarkable accuracy of 94% after tuning underscored its classification efficiency, whereas AdaBoost and Gradient Boosting exhibited lower accuracies, emphasizing the significance of selecting the appropriate ensemble method for specific classification tasks.

J. O. Atoum [3] explored cyberbullying detection through sentiment analysis using Naïve Bayes and Support Vector Machine (SVM) models. The study experimented with different N-gram language models to assess the effectiveness of these models, contributing valuable insights to the field of cyberbullying detection research. SVM classifiers outperformed Naïve Bayes in various evaluation metrics across n-gram models, with improved performance measures compared to previous studies, affirming their effectiveness in enhancing classification accuracy and reliability.

Using a combination of Term Frequency Inverse Document Frequency (TF-IDF), Count Vectorization, Bag of Words, Support Vector Machine (SVM), and Natural Language Processing (NLP), P. Dedeepya et al. [4] investigated the detection of cyberbullying using machine learning approaches. By employing multiple techniques, they aimed to enhance the accuracy and robustness of cyberbullying detection systems. SVM faced challenges in outlier misclassification, with a lack of an implementation solution provided, underscoring the importance of addressing outlier - related issues to enhance SVM-based classification model robustness.

M. M. Islam et al. [5] focused on using machine learning approaches such as Bag of Words and TF-IDF. They explored diverse machine learning methods to effectively categorize and identify cyberbullying content, contributing to the ongoing efforts in developing advanced cyberbullying detection tools. The TF-IDF method consistently outperformed the BoW, while SVM stood out as the top-performing machine learning algorithm, highlighting its effectiveness in text classification tasks and potential for enhancing model accuracy and robustness.

W. A. Prabowo et al. [6] highlighted the effectiveness of SVM model in detecting cyberbullying within software applications. By utilizing TF-IDF alongside SVM, the study demonstrates the potential of this hybrid approach in accurately categorizing cyberbullying content, offering valuable insights for future cyberbullying detection systems. The model's impressive accuracy (93%), precision (95%), and recall (97%), particularly when applied to Indonesian user data, signifies its relevance and adaptability in specific regional contexts, crucial for tailored NLP solutions.

A. Mody et al. [7] presented a hybrid approach to cyberbullying detection using sentiment analysis and machine learning techniques. By leveraging POS tagging, SentiWordNet, Naive Bayes, and Linear Support Vector Classifier (SVC), the worked to enhance the accuracy and reliability of cyberbullying detection systems through a multi-faceted analysis of tweet content. The study reported an overall classification accuracy of 70.3% for identifying cyberbullying tweets which is poor when compared with other models.

The Bag-of-Phonetic-Codes model proposed in the paper offered a novel approach to rectifying misspelled and censored words by utilizing the pronunciation of words as features [8]. Drawing from the Bag-of-Words concept, this model generated phonetic codes using the Soundex Algorithm, capturing text characteristics that can aid in improving the accuracy of text processing tasks. The imbalanced dataset and limited distinction between neutral tweets challenged the accuracy of traditional machine learning models like Support Vector Machine (55%) and Naïve Bayes (57%), urging the need for improved models and feature engineering in cyberbullying detection.

J. Yadav et al. [9] introduced a deep learning model for cyberbullying detection using a pre-trained BERT model with a single linear neural network layer as a classifier. The comparative analysis revealed the superior accuracy of deep learning models like BERT-based approaches in cyberbullying detection, emphasizing the value of advanced deep learning techniques in addressing complex NLP tasks.

The literature survey is compiled in Table 2.1 to present a bird eye's view of the progress made in research in this field.

Table 2.1 Literature survey

S.No.	Description	Year	Research Analysis
1	Analysis of different types of cyberbullying [1]	2018	The strategy seeks to gauge shifts in public sentiment towards criminal incidents and pinpoint the emotions tied to various categories of unlawful acts.
2	Analysis of Tweets with the help of Selenium for web scrapping [2]	2023	Various methodologies such as the Random Forest classifier, gradient

			boosting classifier, Ada boost, and others were tested.
3	Cyberbullying Detection Through Sentiment Analysis by Naïve Bayes and Support Vector Machine Models [3]	2020	They experimented with Naïve Bayes and Support Vector Machine evaluations across various N-gram language models.
4	Detecting Cyber Bullying using Support Vector Machine and Natural Language Processing Techniques [4]	2023	Term Frequency Inverse Document Frequency, Count Vectorization, Bag of Words, Support Vector Machine, and Natural Language Processing were used
5	Detecting Cyberbullying using Machine Learning Approaches [5]	2020	Two prominent techniques were employed for the examination: Bag of Words and Term Frequency Inverse Document Frequency. Diverse machine learning methods were utilized.
6	Detecting Cyberbullying using Machine Learning Approaches using Term Frequency Inverse Document Frequency and Support Vector Machine [6]	2020	The Support Vector Machine emerged as the most effective tool for categorizing cyberbullying within software applications.
7	Cyberbullying detection using a Hybrid Approach in Sentiment Analysis using POS tagging and Support Vector Machine [7]	2018	The text in each tweet undergoes analysis through two distinct methods: one relies on established knowledge using SentiWordNet, while the other employs machine learning with Naive Bayes and Linear Support Vector Classifier (SVC).

8	A Bag-of-Phonetic-Codes model, using the pronunciation of words as features to rectify misspelled and censored words [8]	2018	The system draws from the Bag-of-Words concept to gather text characteristics. It generates phonetic codes through the application of the Soundex Algorithm.
9	Cyberbullying Detection by Pre-Trained BERT Model - A deep learning model [9]	2020	It utilizes a pre-trained BERT model with a single linear neural network layer as a classifier.

Chapter 3

Methodology and Implementation

3.1 Block Diagram

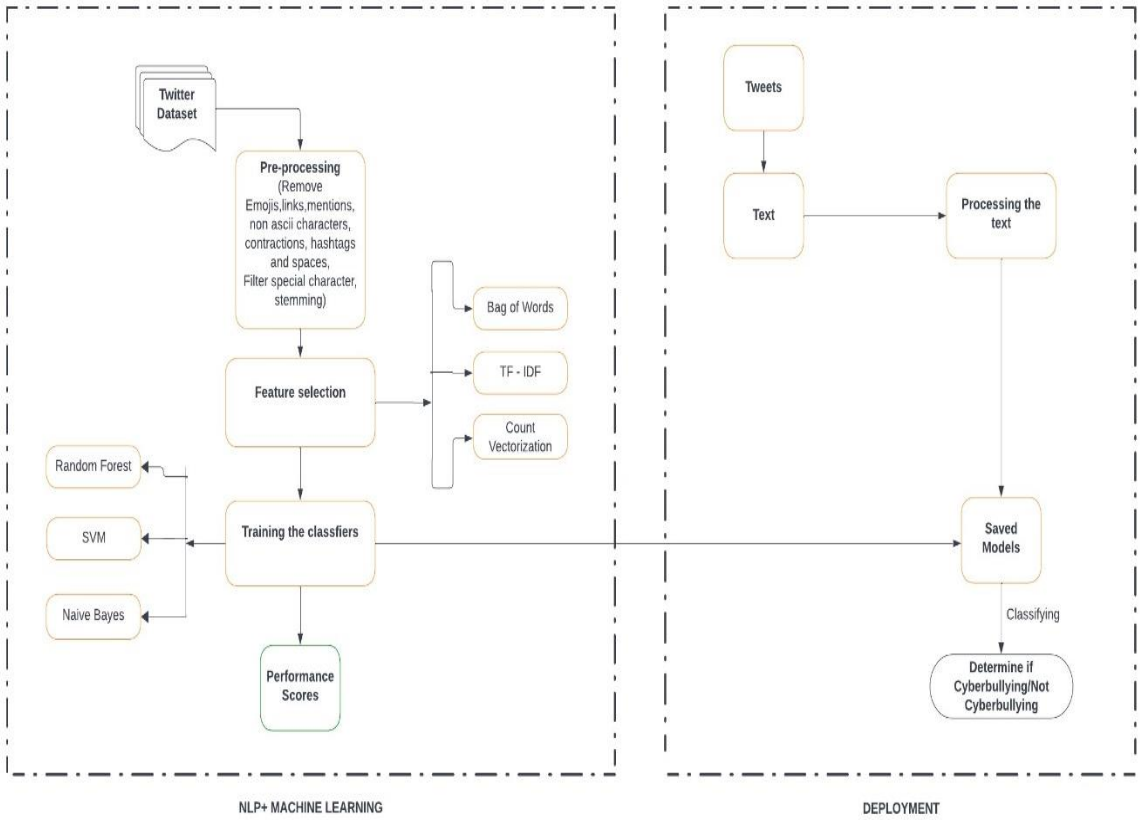

Fig. 3.1 Proposed Framework

The process of implementation includes:

1. **Data Collection:** Data is collected from various sources which include Kaggle and IEEE papers.
2. **Pre-processing:** The tweets are then pre-processed to clean the text data. This involved removing emojis, links, mentions, non-ASCII characters, contractions, hashtags and extra spaces. We also filtered special characters and carried out stemming.
3. **Balancing the dataset:** Synthetic Minority Over-sampling Technique (SMOTE) is applied to generate synthetic samples for the minority class by interpolating between existing minority class samples, thereby balancing the class distribution.

4. **Feature Selection:** Next, features are selected from the text using BoW with n-grams, TF-IDF and Count Vectorizer.
5. **Machine Learning Model Training:** Different machine learning models are trained on the labelled data. These models you include: Random Forest, SVM, Naive Bayes
6. **Performance Scores:** After training the models, their performance is evaluated using performance scores like precision, recall, and F1-score to measure the effectiveness of a machine learning model.
7. **Classification:** Once a model is determined to have the best performance, it is used to classify new tweets as cyberbullying or not cyberbullying.
8. **Deployment:** Finally, the system is deployed to be used in a real-world setting.

3.2 Data Collection

In data collection for cyberbullying research, there is a significant diversity in the data sources used in different research papers. Some studies draw on publicly available data from platforms such as Kaggle, while others prefer real-time data collection. In addition, many researchers are taking the painstaking route of manually preparing their data, while some use a combination of these approaches. These datasets, rich in textual content, often contain emojis, peculiar characters, and unconventional text, which require careful preprocessing to ensure their suitability for analysis.

The data was taken from the paper published by Jason Wang, Kaiqun Fu, Chang-Tien Lu [10] in 2020. This dataset is a compilation of six datasets. After balancing the data, the dataset had 48000 tweets with tags as: Cyberbullying, not cyberbullying, age, ethnicity, gender, religion and other.

The classification was changed to binary classification of Cyberbullying and not cyberbullying. 5500 rows from another Kaggle dataset were further appended to add more instances of not cyberbullying class [11].

3.3 Data Preprocessing

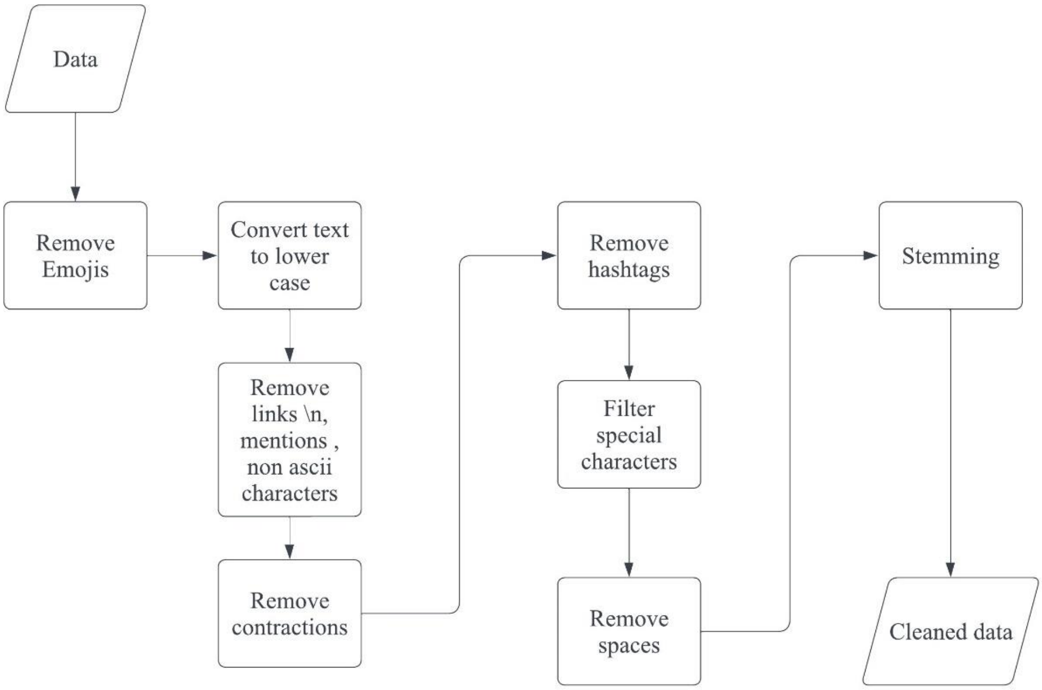

Fig. 3.2 Data Preprocessing

The steps for preprocessing are as follows:

1. **Remove Emojis:** The first step involved removing emojis from the data set.
2. **Convert text to lowercase:** This step converts all uppercase letters in the text data to lowercase.
3. **Remove Hashtags:** Hashtags are irrelevant here and are removed.
4. **Remove links, mentions, non-ASCII characters:** Removed links embedded within the text data, as well as mentions (usernames preceded by an "@" symbol) and non-ASCII characters.
5. **Filter special characters:** Removed special characters from the data set like punctuation marks, mathematical symbols, and other non-alphanumeric characters.
6. **Remove contractions:** Contractions, like "don't" or "can't" are expanded into their full forms.

7. **Remove spaces:** Extra spaces from the text data are removed.
8. **Stemming:** Words are reduced to their base form using stemming.

3.4 Balancing the data

There is a significant difference between the tweets available for the two classes: cyberbullying and not cyberbullying. SMOTE was used to balance the classes.

3.5 Feature Extraction

After SMOTE, we moved on to Feature extraction where we used three techniques: Bag of Words (n-grams), TF-IDF, and Count Vectorizer. When combined with diverse machine learning algorithms and data preprocessing methods, these yielded varying accuracies in sentiment analysis and text classification tasks.

1. **The Bag of Words (BoW)**

   Bag of Words (BoW) is a text representation technique where a document is represented as a bag (multiset) of its words, disregarding grammar and word order. N-grams are contiguous sequences of N items (words in this context) from a given text. So, BoW with N-grams considers not only single words but also sequences of words up to length N in a document.

   In this project, BoW with N-grams is utilized as a feature extraction technique for text data. By converting text into a matrix of token counts (occurrences of words or N-grams), it creates a numerical representation of text data suitable for machine learning algorithms. This allows classification models to process and analyse text data effectively by capturing the presence and frequency of words and N-grams in each document.

2. **Term Frequency-Inverse Document Frequency (TF-IDF)**

   TF-IDF is a numerical statistic that reflects the importance of a word in a document relative to a collection of documents. It is commonly used in natural language processing and information retrieval tasks.

   In this project, TF-IDF is utilized to convert raw text data into numerical feature vectors. Each word in the text is assigned a weight that represents its significance in the document relative to the entire corpus. This process helps in capturing the unique characteristics of each document and is particularly useful in sentiment analysis, where the presence of certain words can indicate the sentiment expressed in the text. The TF-IDF transformed features are then used as input to machine learning models for cyberbullying detection on Twitter.

3. **Count Vectorizer**

   Count Vectorizer is a feature extraction technique used in natural language processing. It converts a collection of text documents into a matrix of token counts. Each row in the matrix represents a document, and each column represents a unique word in the corpus. The value in each cell denotes the frequency of the corresponding word in the document.

   Here, Count Vectorizer is employed to convert the text data into numerical features. Each tweet is transformed into a vector of word counts, capturing the frequency of each word in the tweet. These word count vectors are then used as input features for machine learning algorithms to train models for cyberbullying detection.

3.6 Machine Learning Models

The models that we chose to implement in combination with the three feature extraction techniques are as follows:

1. **Random Forest**

   Random Forest builds multiple decision trees during training. Each tree in the forest operates independently and makes a prediction, and the final prediction is determined by a majority vote or averaging.

   In this project, Random Forest is employed as a machine learning model for cyberbullying detection. It is utilized to classify text data (tweets) into different categories of cyberbullying based on their content. By training on pre-labelled data, Random Forest learns patterns in the text features and is then able to predict the sentiment (Cyberbullying or Not Cyberbullying) of unseen tweets with high accuracy.

2. **Support Vector Machine (SVM)**

   Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification and regression tasks. Its objective is to find the optimal hyperplane that best separates different classes in the feature space while maximizing the margin between the classes.

We utilized SVM as a classifier to distinguish between different types of cyberbullying based on tweet content. It learns to classify tweets into predefined categories such as Cyberbullying or not cyberbullying by analysing the features extracted from the text data. SVM aims to find the decision boundary that best separates the tweets into Cyberbullying or not, thus enabling effective detection and classification of cyberbullying instances on Twitter.

3. **Naïve Bayes**

Naive Bayes is a simple probabilistic classifier based on Bayes' theorem with the assumption of independence between features. Despite its simplicity, it often performs well in text classification tasks.

In this project, Naive Bayes is used as one of the machine learning models for sentiment analysis of Twitter data. It leverages the frequency of words in the cleaned text data to classify tweets whether they are cyberbullying or not. By assuming independence between the occurrence of words in tweets, Naive Bayes efficiently learns and predicts the sentiment of tweets based on their textual content.

3.7 Deployment

After a comparative analysis to find which model combination yields the best result, the relevant feature extraction vectorizer and Machine Learning model is downloaded as joblib files. The deployment diagram is as follows:

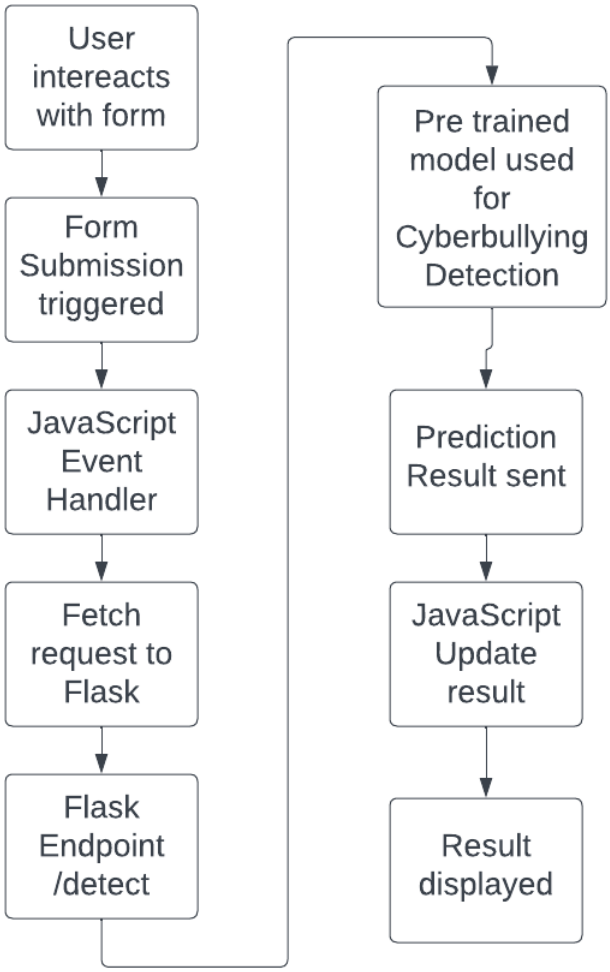

Fig. 3.3 Deployment

1. **User interacts with form:** The process begins with a user filling giving an input.
2. **Form submission triggered:** Once the user submits the form, the process is triggered.
3. **JavaScript Event Handler:** A JavaScript event handler is then employed.

4. **Prediction Result sent:** The JavaScript code then sends the text entered in the form for a prediction. This involves sending the text to a pre-trained model, which is designed to identify cyberbullying text.
5. **Fetch request to Flask:** This step involves sending a request to a Flask endpoint, likely on a server. The web application is designed to make the prediction on the server-side.
6. **Flask Endpoint /detect:** The request is sent to a specific Flask endpoint. The endpoint is designed to receive the text data and use the pre-trained model to make a cyberbullying detection.
7. **JavaScript Update Result:** The response from the Flask endpoint containing the prediction result (i.e., whether the text is cyberbullying) is then sent back to the JavaScript code. The JavaScript code updates the result accordingly.
8. **Result Displayed:** Finally, the JavaScript code displays the result on the user's web page.

Chapter 4

Result and Analysis

The performance measures used for comparing models:

a) **Confusion Matrix**

The utilization of a confusion matrix constitutes a fundamental aspect in the assessment of the efficacy of a classification model. It furnishes a comprehensive overview of the model's forecasts in comparison to the genuine ground truth across diverse categories. The components of a confusion matrix encompass:

- True positives (TP): occur when the model accurately predicts a positive data point.
- True negatives (TN): occur when the model accurately predicts a negative data point.
- False positives (FP): occur when the model predicts a positive data point incorrectly.
- False negatives (FN): occur when the model mispredicts a negative data point.

b) **Accuracy**

The measure of accuracy pertains to the ratio of correctly classified instances relative to the total instances under evaluation. Within the project's scope, accuracy serves as a reflection of the overall efficacy of the cyberbullying detection model in accurately discerning both cyberbullying and non-cyberbullying tweets. A heightened accuracy score indicates the model's adeptness in correctly categorizing many tweets, thereby offering dependable insights into the prevalence of cyberbullying on the Twitter platform.

$$\text{Accuracy} = \frac{TN+TP}{TN+FP+TP+FN} \dots\dots\dots \quad (1)$$

c) **Precision**

Precision denotes the ratio of true positive instances among all instances predicted as positive by the model, showcasing the model's capacity to avert misclassification of non-cyberbullying tweets as cyberbullying. Within the project, precision signifies the model's dependability in precisely identifying tweets containing cyberbullying content. A substantial precision score indicates the model's proficiency in displaying a minimal rate of false positives, thus ensuring that highlighted tweets genuinely signify cyberbullying behaviour.

$$\text{Precision} = \frac{TP}{FP+TP} \ldots\ldots\ldots \quad (2)$$

d) Recall

Recall, also recognized as sensitivity, gauges the ratio of true positive instances that were accurately identified by the model out of all genuine positive instances. In the project's context, recall illustrates the model's capability to capture and correctly categorize all instances of cyberbullying within the dataset. A notable recall score highlights the model's efficacy in identifying the majority of cyberbullying instances, thereby reducing the likelihood of false negatives and guaranteeing a comprehensive coverage of cyberbullying content on Twitter.

$$\text{Recall} = \frac{TP}{TP+FN} \ldots\ldots\ldots \quad (3)$$

e) F1 score

The F1 score represents the harmonic mean of precision and recall, furnishing a well-rounded evaluation of the model's performance. It considers both false positives and false negatives, rendering it a valuable metric for assessing the overall efficacy of the cyberbullying detection model. In the project, a substantial F1 score indicates the model's achievement of a balance between precision and recall, showcasing robust performance in accurately detecting cyberbullying content while minimizing misclassifications.

$$\text{F1 score} = 2* \frac{Precision*Recall}{Precision+Recall} \ldots\ldots\ldots \quad (4)$$

In depth analysis of all 9 combinations formed.

1. Count Vectorizer and Random Forest

Count vectorizer with random forest was one of the better performing models with 90.14% accuracy score, 90.26% precision score, 90.14% recall score, 90.13% F1 score.

```
Random Forest Classification Report:
              precision    recall  f1-score   support

           0       0.88      0.93      0.90      7270
           1       0.93      0.87      0.90      7370

    accuracy                           0.90     14640
   macro avg       0.90      0.90      0.90     14640
weighted avg       0.90      0.90      0.90     14640

Accuracy Score for Random Forest Classifier: 0.9014
Precision Score for Random Forest Classifier: 0.9026788761683529
Recall Score for Random Forest Classifier: 0.9014344262295082
F1 Score for Random Forest Classifier: 0.9013788949612742
```

Fig 4.1 Classification report for CV with RF

The confusion matrix has the following statistics: There are 6749 true negatives, 6448 true positives, 521 false positives, and 922 false negatives.

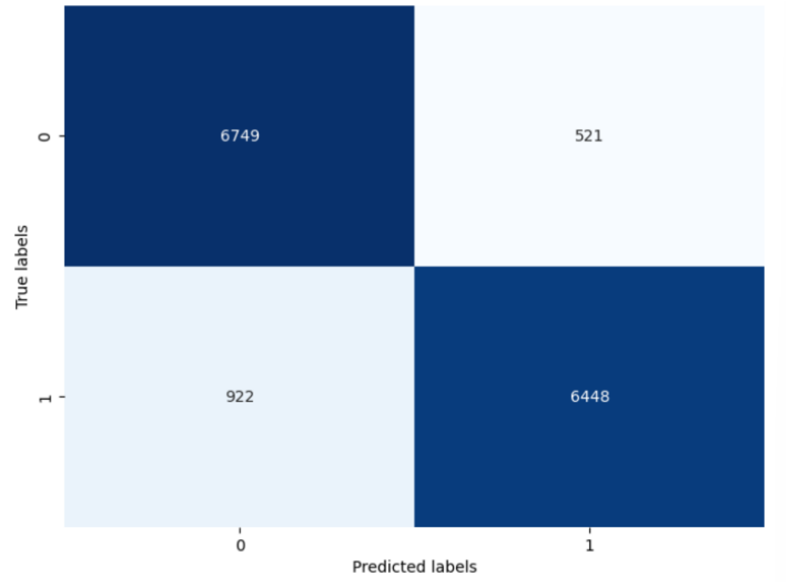

Fig 4.2 Confusion matrix for CV with RF

A steeper ROC curve indicates better performance, and the Area Under the Curve (AUC) value, denoted as 0.97 here, quantifies the model's overall performance; a higher AUC suggests better discrimination between the classes. The high AUC value of 0.97 indicates that the classifier has excellent predictive performance.

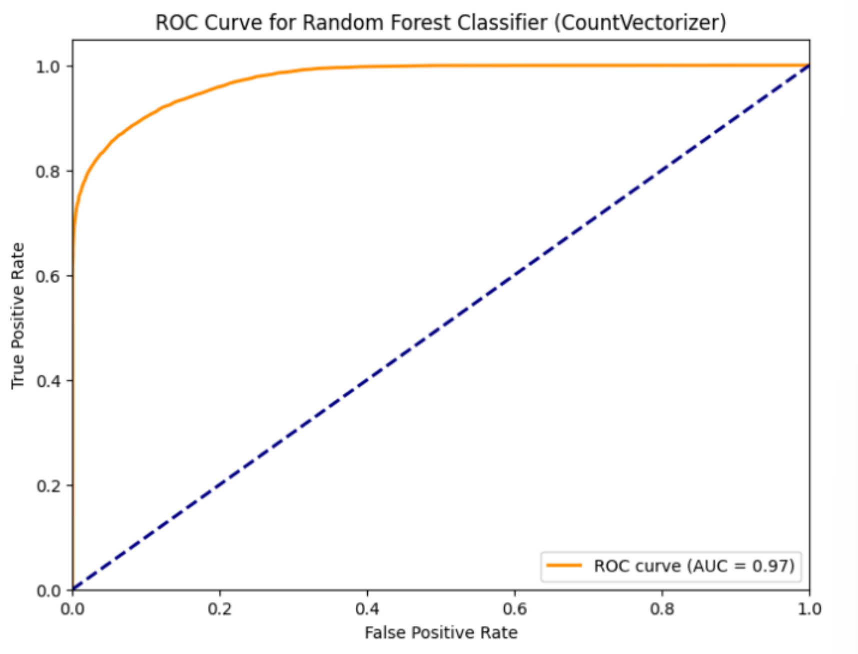

Fig 4.3 ROC curve for CV with RF

2. Count Vectorizer and SVM

Count vectorizer with SVM was the third best performing models with 91.62% accuracy score, 91.69% precision score, 91.62% recall score, 91.62% F1 score.

```
SVM Classification Report:
              precision    recall  f1-score   support

           0       0.90      0.94      0.92      7270
           1       0.93      0.90      0.92      7370

    accuracy                           0.92     14640
   macro avg       0.92      0.92      0.92     14640
weighted avg       0.92      0.92      0.92     14640

Accuracy Score for SVM Classifier: 0.9162

Precision Score for SVM Classifier: 0.9168697610467742
Recall Score for SVM Classifier: 0.916188524590164
F1 Score for SVM Classifier: 0.9161667864884624
```

Fig 4.4 Classification report for Count Vectorizer and SVM

The confusion matrix has the following statistics: 6802 true negatives, 6611 true positives, 468 false positives, and 759 false negatives.

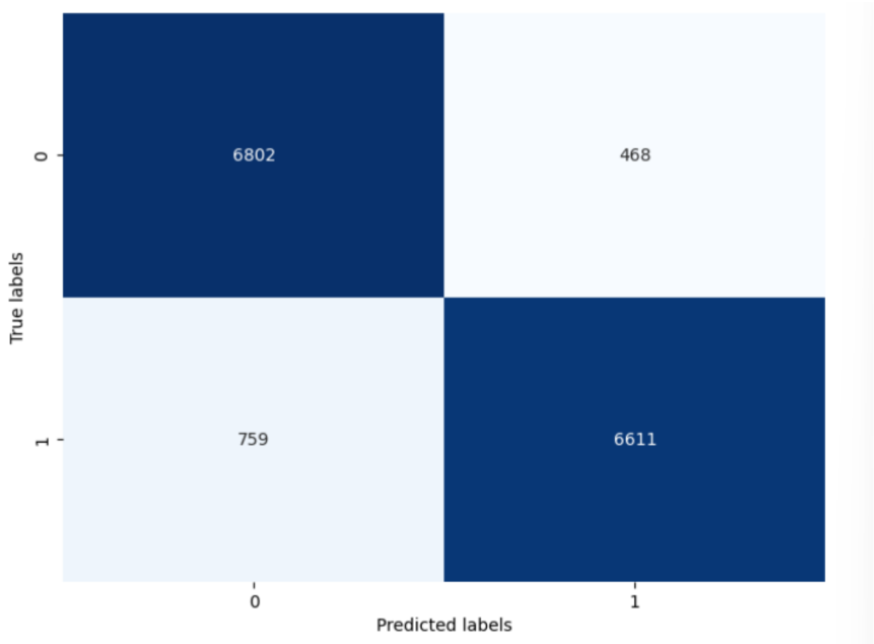

Fig 4.5 Confusion matrix for Count Vectorizer and SVM

The Area Under the Curve (AUC) value, denoted as 0.98, quantifies the model's overall performance with a higher AUC suggests better discrimination between the classes.

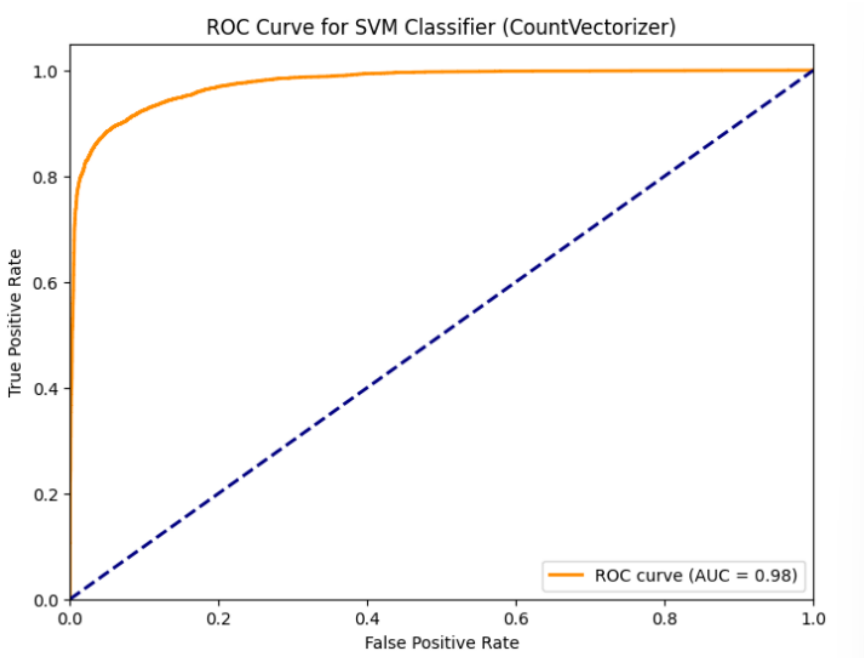

Fig 4.6 ROC curve for Count Vectorizer and SVM

3. Count Vectorizer and Naïve Bayes

Count vectorizer with Naïve Bayes was one of the poor performing models with 83.10% accuracy score, 83.67% precision score, 83.10% recall score, 83.02% F1 score.

```
Naive Bayes Classification Report:
              precision    recall  f1-score   support

           0       0.88      0.76      0.82      7270
           1       0.79      0.90      0.84      7370

    accuracy                           0.83     14640
   macro avg       0.84      0.83      0.83     14640
weighted avg       0.84      0.83      0.83     14640

Accuracy Score for Naive Bayes Classifier: 0.8310
Precision Score for Naive Bayes Classifier: 0.8366847036742341
Recall Score for Naive Bayes Classifier: 0.8310109289617487
F1 Score for Naive Bayes Classifier: 0.8302011482463786
```

Fig 4.7 Classification report for Count Vectorizer and Naïve Bayes

The confusion matrix has the following statistics: 5522 true negatives, 6614 true positives, 1718 false positives, and 756 false negatives.

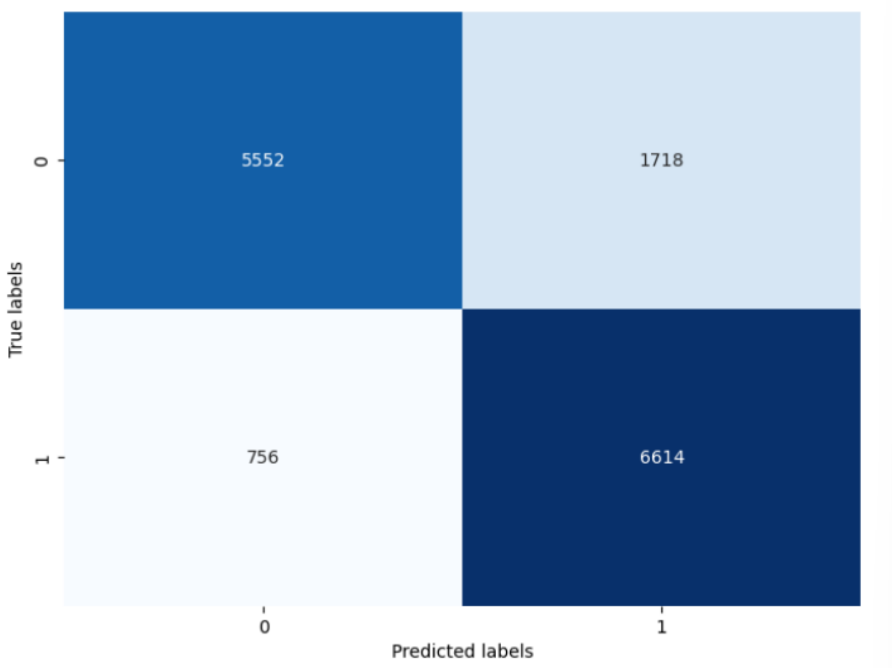

Fig 4.8 Confusion matrix for Count Vectorizer and Naïve Bayes

The ROC curve for the Naive Bayes classifier with Count Vectorizer representation demonstrates strong performance, with an area under the curve (AUC) of 0.92.

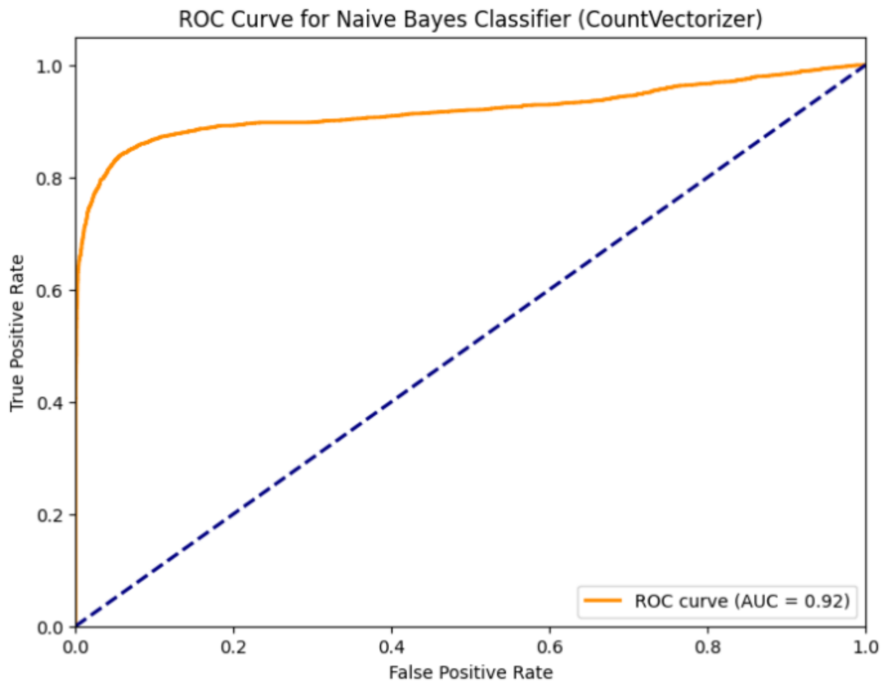

Fig 4.9 ROC curve for Count Vectorizer and Naïve Bayes

4. TF-IDF and Random Forest

TF-IDF with random forest was one of the better performing models with 92.31% accuracy score, 92.69% precision score, 92.30% recall score, 92.29% F1 score.

```
Random Forest Classification Report:
              precision    recall  f1-score   support

           0       0.89      0.97      0.93      7270
           1       0.97      0.88      0.92      7370

    accuracy                           0.92     14640
   macro avg       0.93      0.92      0.92     14640
weighted avg       0.93      0.92      0.92     14640

Accuracy Score for Random Forest Classifier: 0.9231
Precision Score for Random Forest Classifier: 0.926903328847879
Recall Score for Random Forest Classifier: 0.9230874316939891
F1 Score for Random Forest Classifier: 0.9229420257805973
```

Fig 4.10 Classification report for TF-IDF and Random Forest

The confusion matrix has the following statistics: 7051 true negatives, 6463 true positives, 219 false positives, and 907 false negatives.

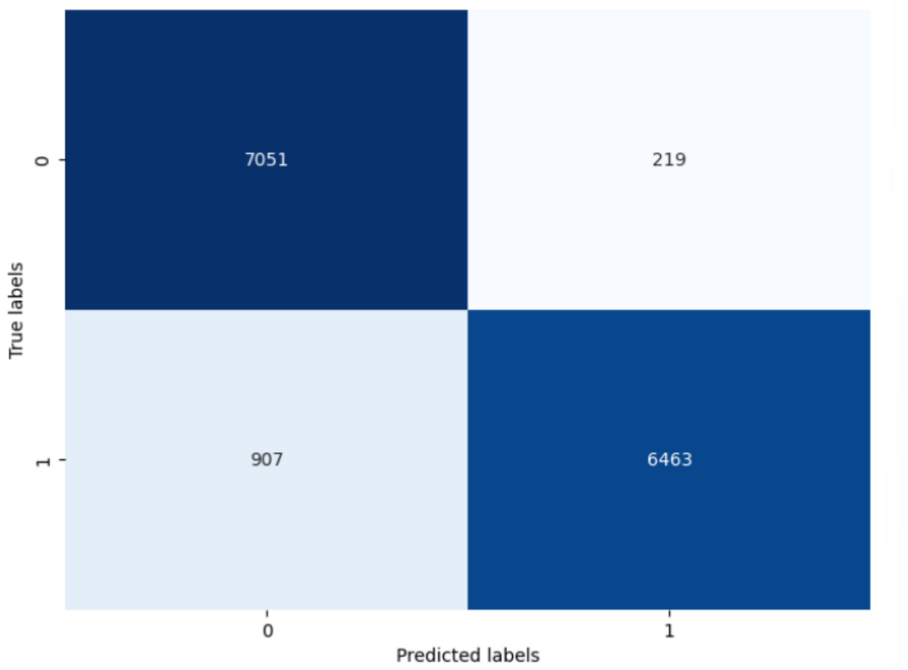

Fig 4.11 Confusion matrix for TF-IDF and Random Forest

The ROC curve for the Random Forest classifier with TF-IDF representation exhibits excellent performance, with an area under the curve (AUC) of 0.98.

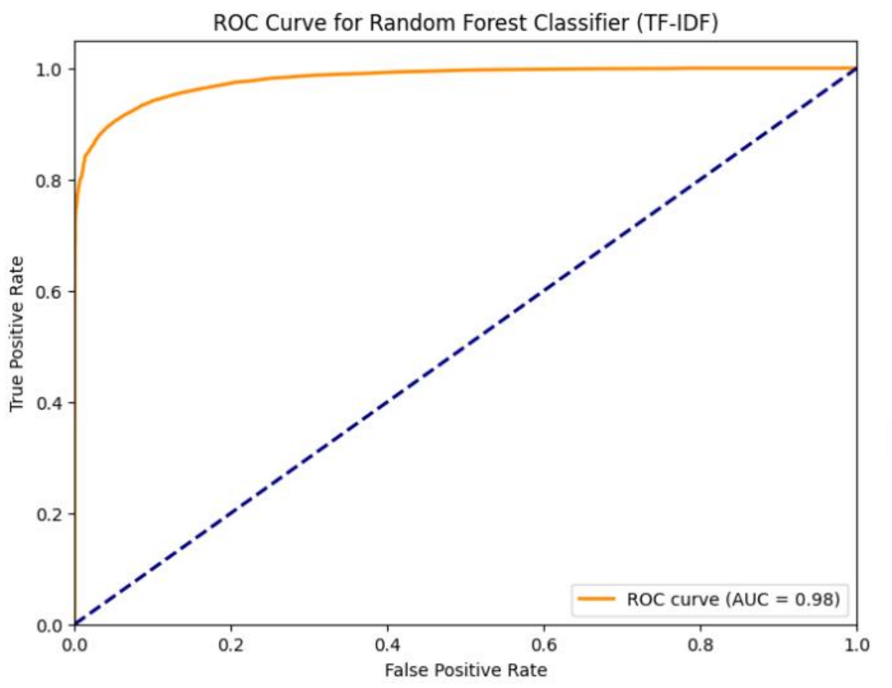

Fig 4.12 ROC curve for TF-IDF and Random Forest

5. TF-IDF and SVM

TF-IDF with SVM was the best performing models with 92.98% accuracy score, 92.99% precision score, 92.98% recall score, 92.98% F1 score.

```
SVM Classification Report:
              precision    recall  f1-score   support

           0       0.92      0.94      0.93      7270
           1       0.94      0.92      0.93      7370

    accuracy                           0.93     14640
   macro avg       0.93      0.93      0.93     14640
weighted avg       0.93      0.93      0.93     14640

Accuracy Score for SVM Classifier: 0.9298
Precision Score for SVM Classifier: 0.9299190986950254
Recall Score for SVM Classifier: 0.9298497267759562
F1 Score for SVM Classifier: 0.9298500472037285
```

Fig 4.13 Classification report for TF-IDF and SVM

The confusion matrix has the following statistics: there are 6801 true negatives, 6812 true positives, 469 false positives, and 558 false negatives.

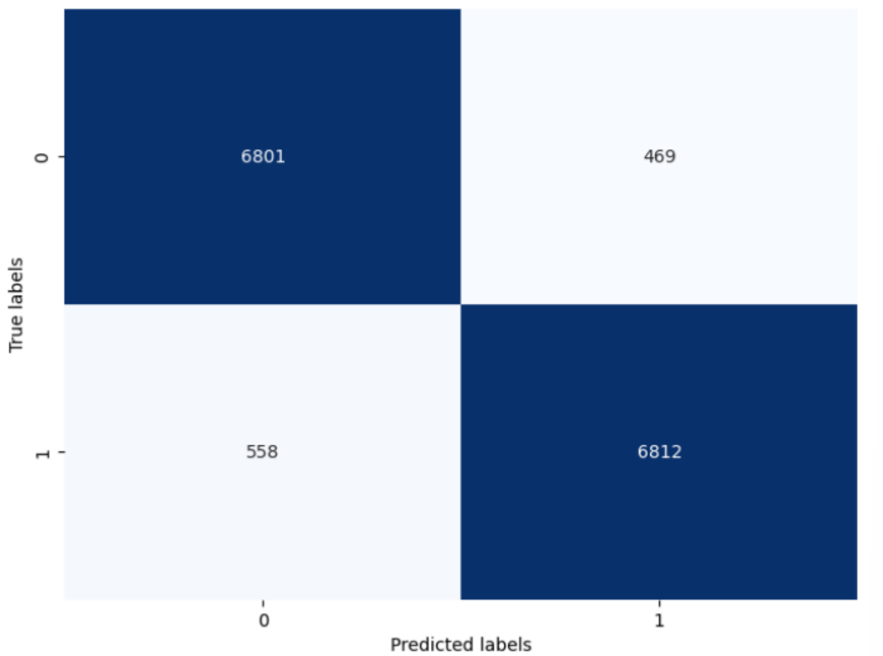

Fig 4.14 Confusion matrix for TF-IDF and SVM

The ROC curve for the SVM classifier with TF-IDF representation illustrates excellent performance, with an AUC of 0.98. This indicates a high ability of the classifier to distinguish between true positive and false positive rates, reflecting strong predictive capabilities across various thresholds.

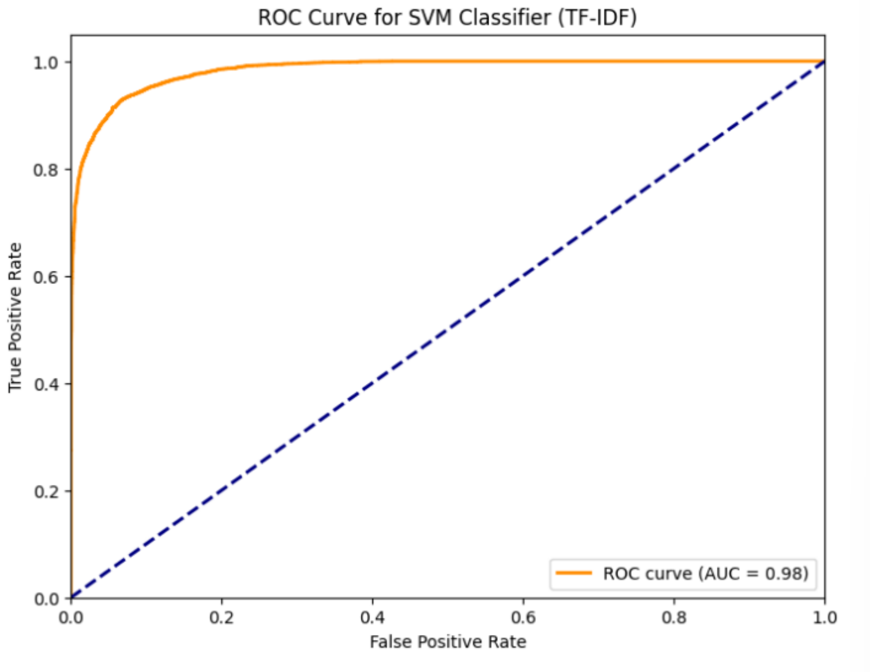

Fig 4.15 ROC curve for TF-IDF and SVM

6. TF-IDF and Naïve Bayes

TF-IDF with Naïve bayes had 84.85% accuracy score, 84.27% precision score, 84.85% recall score, 84.78% F1 score.

```
Naive Bayes Classification Report:
              precision    recall  f1-score   support

           0       0.89      0.79      0.84      7270
           1       0.81      0.90      0.86      7370

    accuracy                           0.85     14640
   macro avg       0.85      0.85      0.85     14640
weighted avg       0.85      0.85      0.85     14640

Accuracy Score for Naive Bayes Classifier: 0.8485
Precision Score for Naive Bayes Classifier: 0.8526867924370832
Recall Score for Naive Bayes Classifier: 0.8484972677595628
F1 Score for Naive Bayes Classifier: 0.8479766615454507
```

Fig 4.16 Classification report for TF-IDF and Naïve Bayes

The confusion matrix has the following statistics: there are 5757 true negatives, 6665 true positives, 1513 false positives, and 705 false negatives.

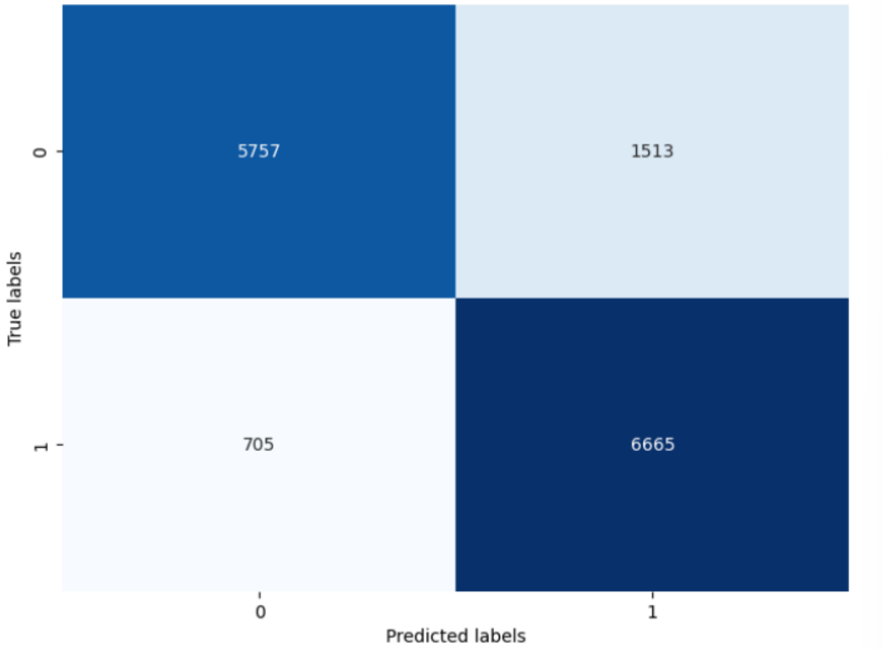

Fig 4.17 Confusion matrix for TF-IDF and Naïve Bayes

The ROC curve for the Naive Bayes classifier with TF-IDF representation indicates strong performance for both Class 0 and Class 1, with an area under the curve (AUC) of 0.94 for each class. This suggests that the classifier has a high ability to distinguish between the positive and negative classes.

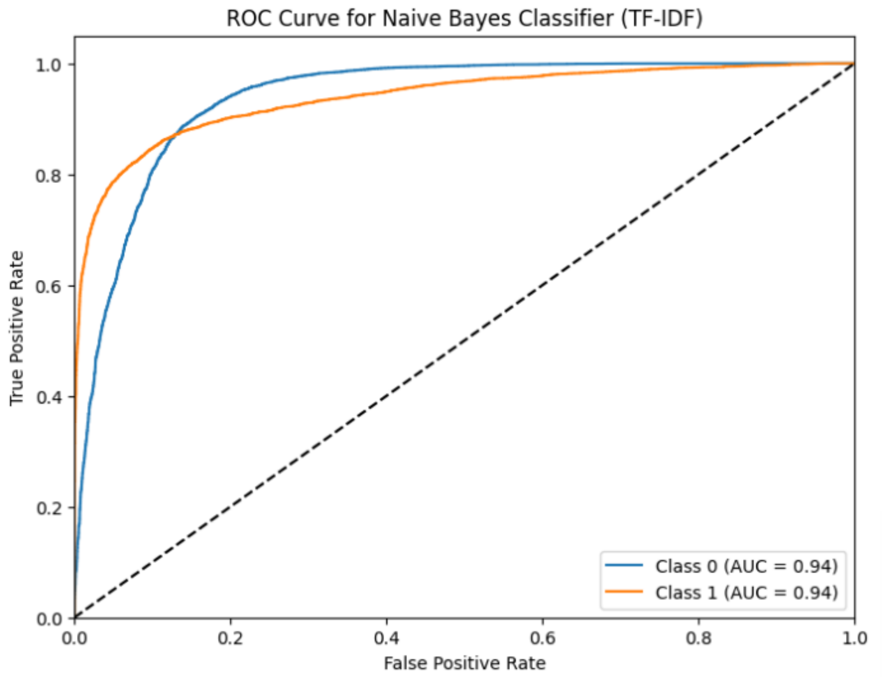

Fig 4.18 ROC curve for TF-IDF and Naïve Bayes

7. BoW using n-grams and Random Forest

BoW using n-grams and random forest had an 89.25% accuracy score, 89.69% precision score, 89.25% recall score, 89.22% F1 score.

```
Random Forest Classification Report:
              precision    recall  f1-score   support

           0       0.85      0.94      0.90      7270
           1       0.94      0.84      0.89      7370

    accuracy                           0.89     14640
   macro avg       0.90      0.89      0.89     14640
weighted avg       0.90      0.89      0.89     14640

Accuracy Score for Random Forest Classifier: 0.8924863387978142
Precision Score for Random Forest Classifier: 0.8969497118233968
Recall Score for Random Forest Classifier: 0.8924863387978142
F1 Score for Random Forest Classifier: 0.8922270059182499
```

Fig 4.19 Classification report for BoW using n-grams and Random Forest

The confusion matrix has the following statistics: there are 6868 true negatives, 6198 true positives, 402 false positives, and 1172 false negatives.

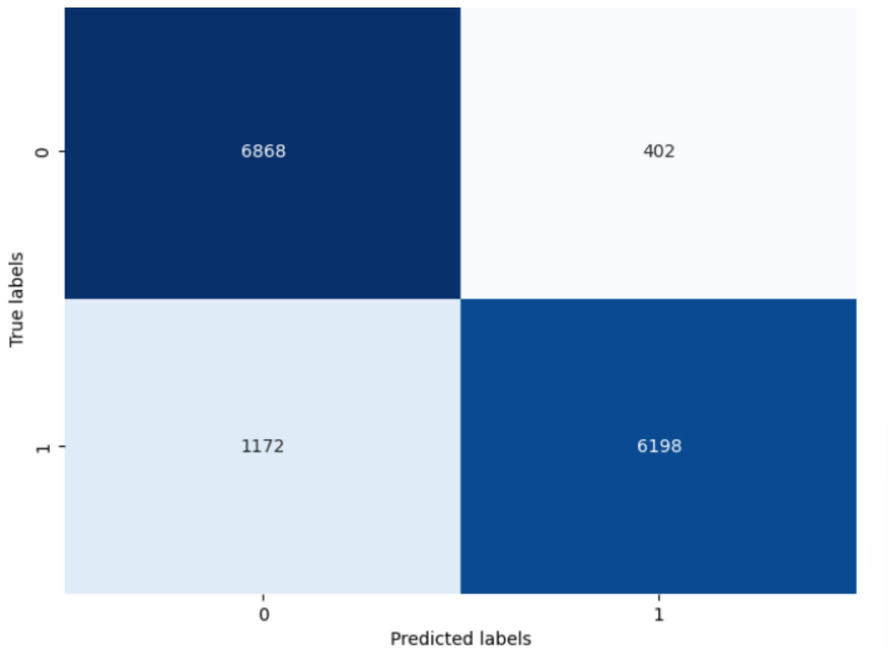

Fig 4.20 Confusion matrix for BoW using n-grams and Random Forest

With an AUC of 0.97, this classifier demonstrates excellent predictive performance.

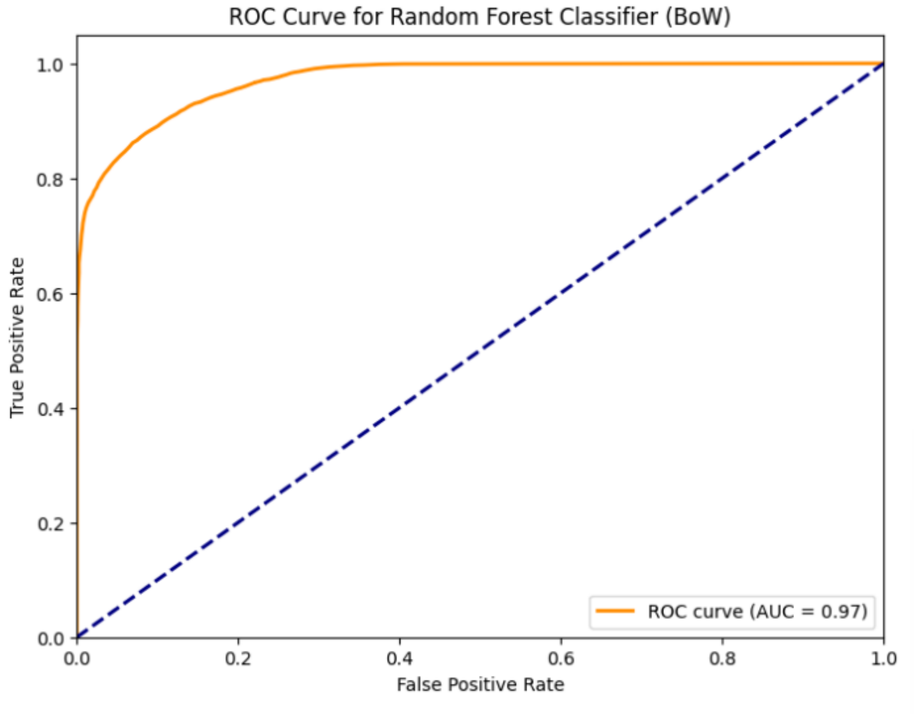

Fig 4.21 ROC curve for BoW using n-grams and Random Forest

8. BoW using n-grams and SVM

BoW using n-grams and SVM had a 90.93% accuracy score, 90.94% precision score, 90.93% recall score, 90.93% F1 score.

```
SVM Classification Report:
              precision    recall  f1-score   support

           0       0.91      0.91      0.91      7270
           1       0.91      0.91      0.91      7370

    accuracy                           0.91     14640
   macro avg       0.91      0.91      0.91     14640
weighted avg       0.91      0.91      0.91     14640

Accuracy Score for SVM Classifier: 0.9093579234972677
Precision Score for SVM Classifier: 0.9093728423373346
Recall Score for SVM Classifier: 0.9093579234972677
F1 Score for SVM Classifier: 0.9093589296170186
```

Fig 4.22 Classification report for BoW using n-grams and SVM

The confusion matrix has the following statistics: there are 6626 true negatives, 6687 true positives, 644 false positives, and 683 false negatives.

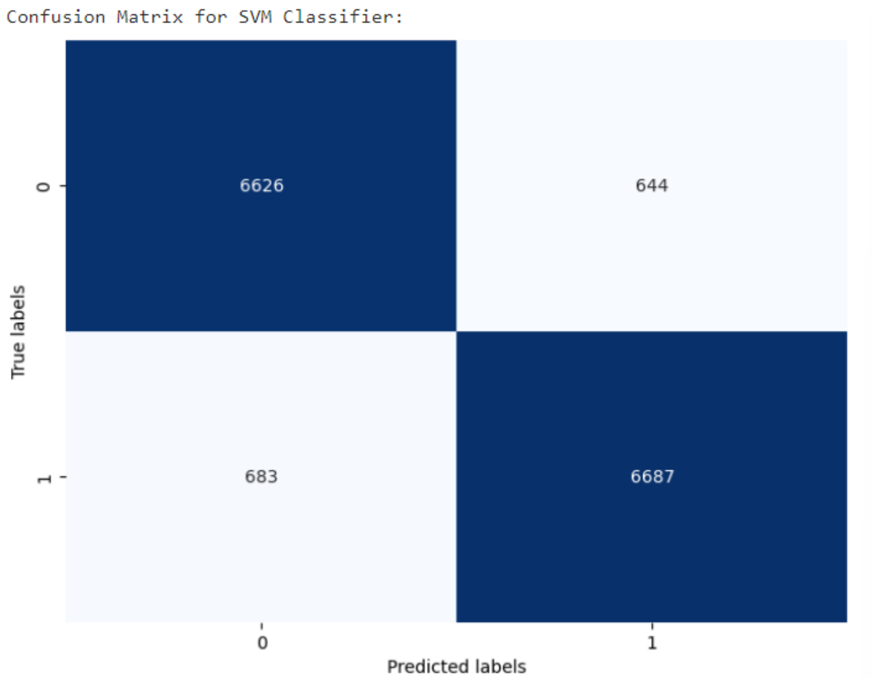

Fig 4.23 Confusion matrix for BoW using n-grams and SVM

With an AUC of 0.98, BoW using n-grams and SVM exhibits excellent predictive performance.

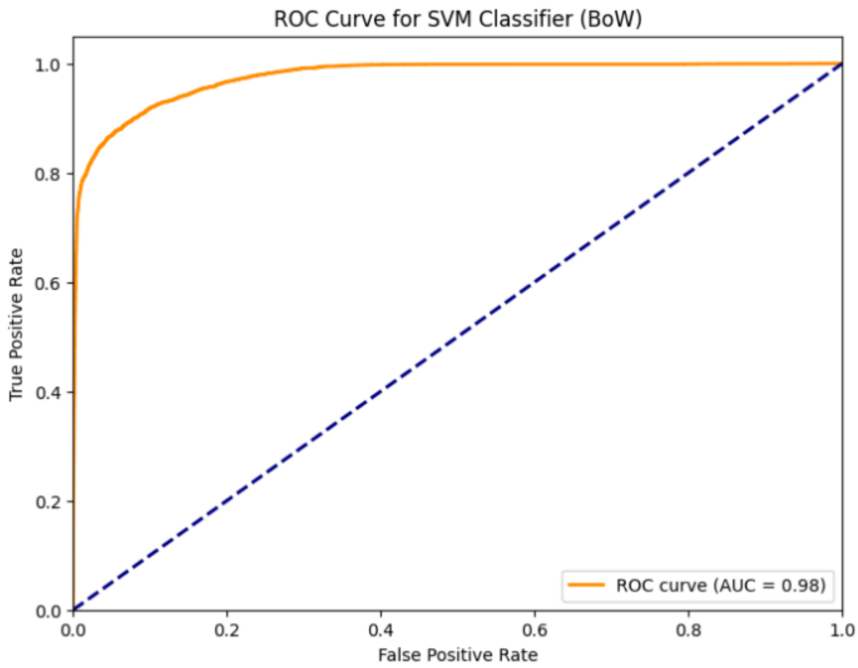

Fig 4.24 ROC curve for BoW using n-grams and SVM

9. BoW using n-grams and Naïve Bayes

This combination was one of the worst performing models with 76.52% accuracy score, 79.22% precision score, 76.52% recall score, 75.93% F1 score.

```
Naive Bayes Classification Report:
              precision    recall  f1-score   support

           0       0.88      0.61      0.72      7270
           1       0.70      0.92      0.80      7370

    accuracy                           0.77     14640
   macro avg       0.79      0.76      0.76     14640
weighted avg       0.79      0.77      0.76     14640

Accuracy Score for Naive Bayes Classifier: 0.7652322404371584
Precision Score for Naive Bayes Classifier: 0.7922467573177625
Recall Score for Naive Bayes Classifier: 0.7652322404371584
F1 Score for Naive Bayes Classifier: 0.7593350023445067
```

Fig 4.25 Classification report for BoW using n-grams and Naïve Bayes

The confusion matrix has the following statistics: there are 4431 true negatives, 6772 true positives, 2839 false positives, and 598 false negatives

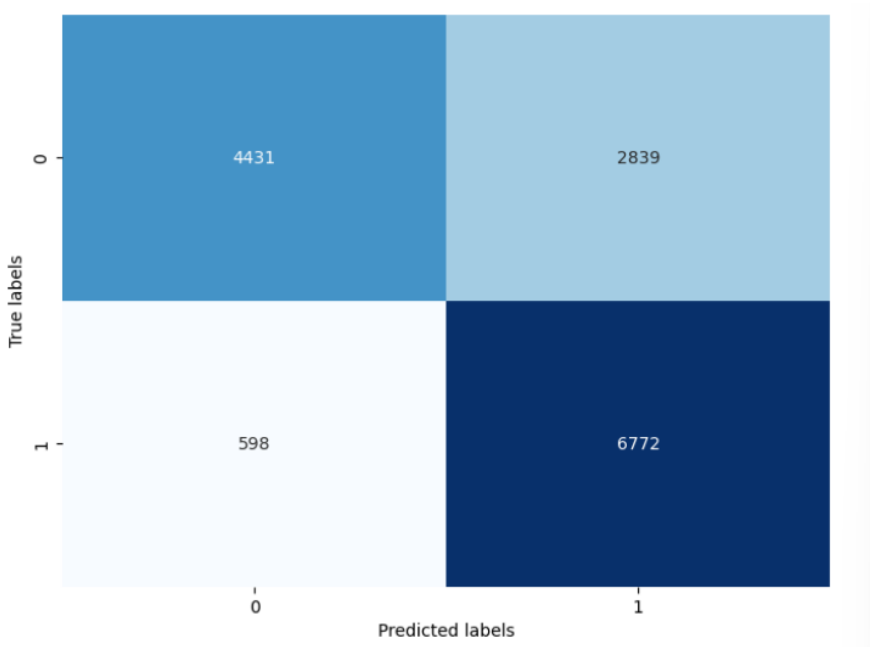

Fig 4.26 Confusion matrix for BoW using n-grams and Naïve Bayes

With an AUC of 0.91, this classifier demonstrates good predictive performance.

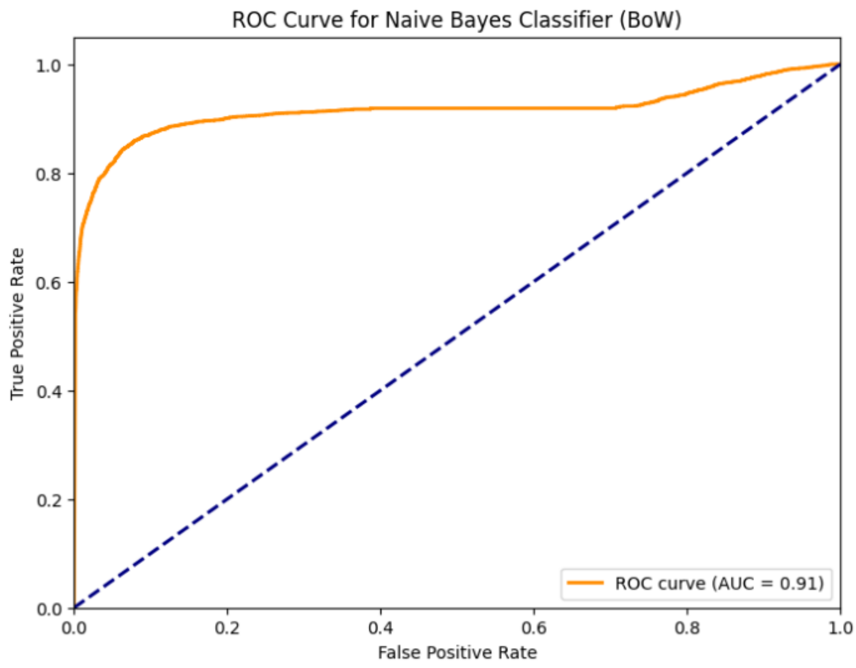

Fig 4.27 ROC curve for BoW using n-grams and Naïve Bayes

After analysing various combinations, their performance measures are combined and presented in table 4.1.

Table 4.1 Comparison table

	Count Vectorizer			TF-IDF			Bag of Words (n-gram)		
	SVM	RF	NB	SVM	RF	NB	SVM	RF	NB
Accuracy	91.62	90.14	83.10	92.98	92.31	84.85	90.94	89.25	76.52
F1 Score	91.62	90.13	83.02	92.99	92.29	84.80	90.94	89.22	75.93
Precision	91.69	90.26	83.67	92.99	92.69	85.27	90.94	89.69	79.22
Recall	91.62	90.14	83.10	92.98	92.31	84.85	90.94	89.25	76.52

The bar graph charts provide the comparative study analysis of each performance measure (Accuracy, Precision, Recall and F1 Score) in all combinations to evaluate the best model.

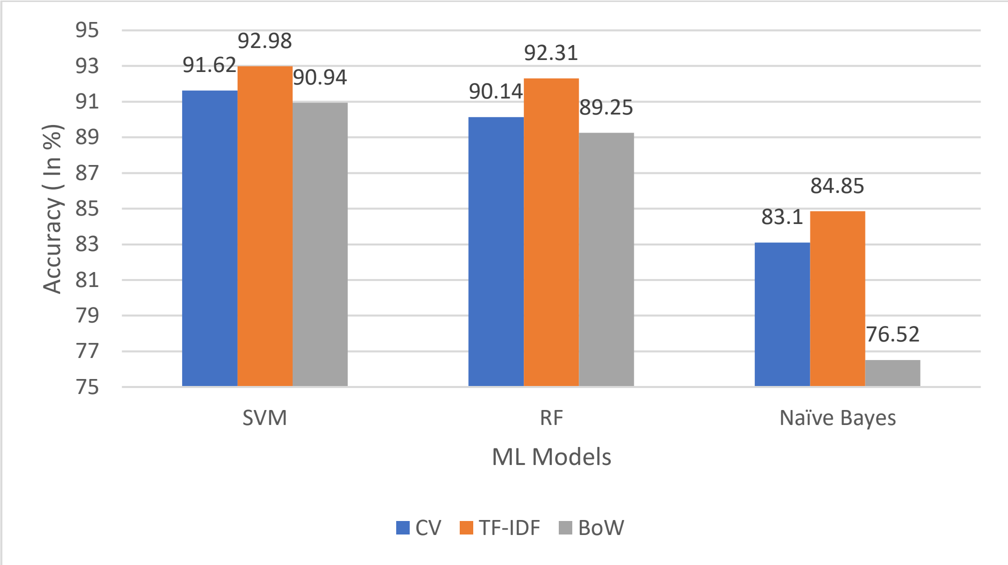

Fig 4.28 Accuracy Scores for all combinations

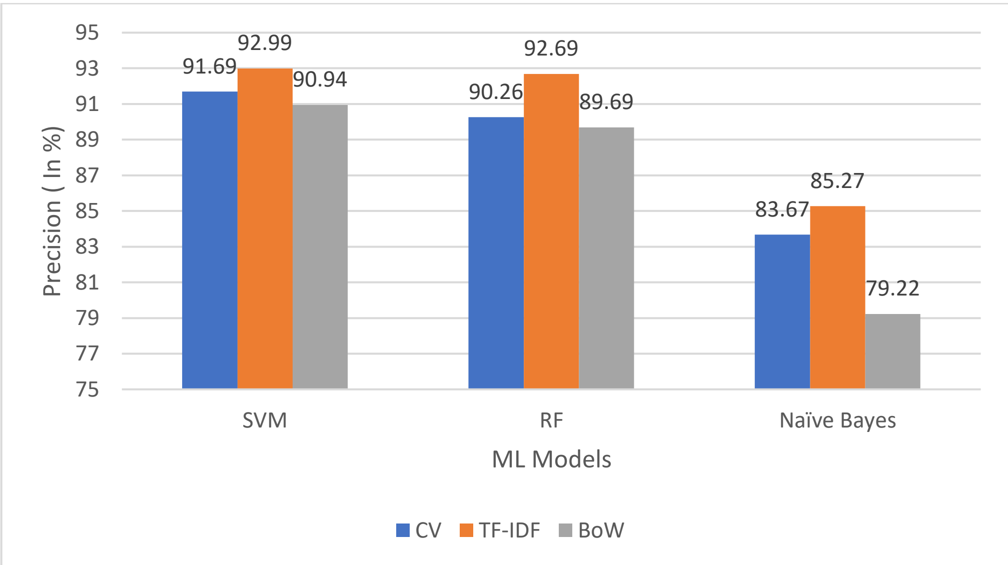

Fig 4.29 Precision Scores for all combinations

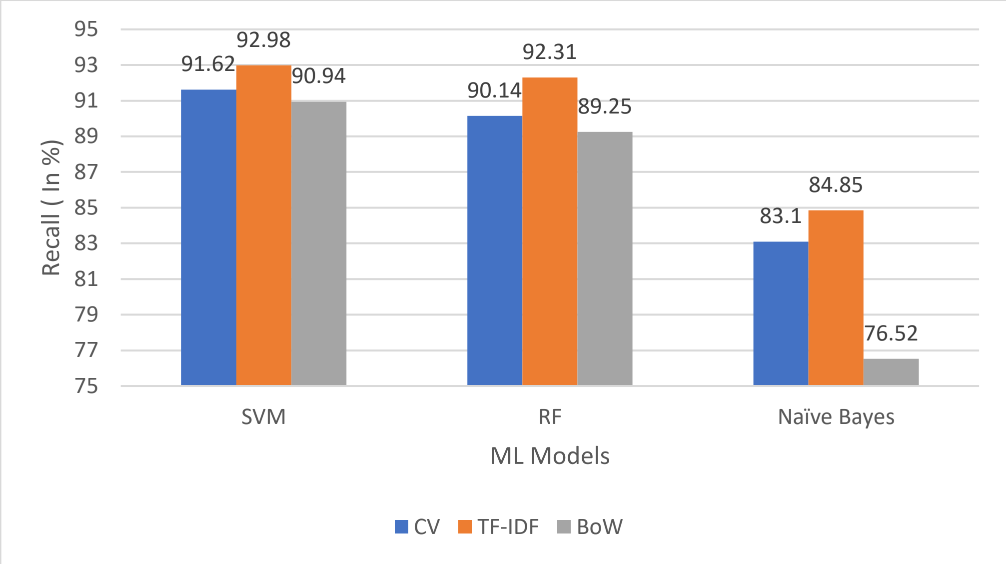

Fig 4.30 Recall Scores for all combinations

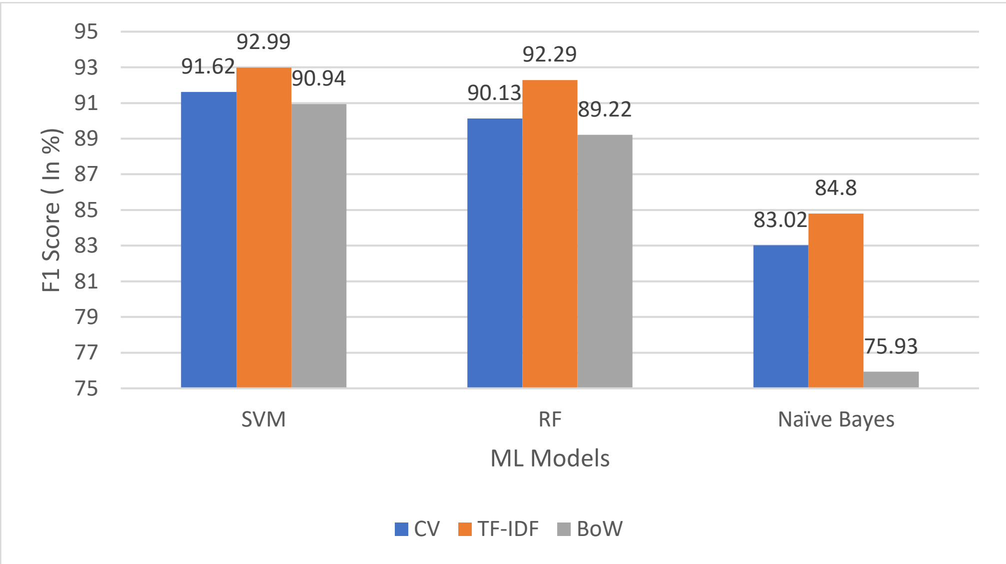

Fig 4.31 F1 Scores for all combinations

Following a comprehensive evaluation of multiple models for cyberbullying detection, SVM model utilizing TF-IDF for feature extraction emerged as the most effective choice with a 92.98% accuracy. Cross-validation was conducted to verify the accuracy of the model, with both accuracy and cross-validation scores showing close alignment. Consequently, the accuracy of the SVM model was duly validated. This model was then deployed for practical use in detecting cyberbullying. It accepts input in the form of strings and employs its trained knowledge to classify whether the input statement exhibits characteristics of cyberbullying or not.

Chapter 5

Advantages, Limitations and Applications

5.1 Advantages

1. The project aims to develop a robust cyberbullying detection system using sentiment analysis, which can proactively identify and address instances of cyberbullying on Twitter, contributing to a safer online environment.
2. The system enables early detection of cyberbullying incidents on Twitter, allowing for prompt intervention to prevent escalation and provide support to victims.

5.2 Limitations

1. The project relies on the effectiveness of the selected feature selection methods (Bag of Words, TF-IDF, Count Vectorizer) and machine learning models (Naive Bayes, Random Forest, Support Vector Machine), which may have limitations in accurately detecting cyberbullying instances.
2. The project utilizes a dataset compiled from various sources, which may introduce biases or inconsistencies in the data.

5.3 Applications

1. The developed cyberbullying detection system can be implemented on Twitter to proactively identify and address instances of cyberbullying, fostering a safer online environment
2. The system enables early detection of cyberbullying incidents on Twitter, allowing for prompt intervention to prevent escalation and provide support to victims
3. The project's web application, which utilizes a Flask endpoint and a pre-trained model, can be used to make real-time predictions on the server-side, identifying cyberbullying text and classifying tweets into different categories of cyberbullying based on their content

Chapter 6

Conclusion and Future Scope

6.1 Conclusion

The project signifies a notable progression in the continual endeavours to combat cyberbullying on Twitter. Through the utilization of sentiment analysis and machine learning methodologies, a solution that exhibits the ability to accurately identify cyberbullying with an exceptional precision rate of 92.98% utilizing Support Vector Machine (SVM) with Term Frequency-Inverse Document Frequency (TF-IDF) was found.

Moreover, the project goes beyond simple detection by presenting a practical and user-friendly solution that can be implemented at a local level, facilitating instant recognition of cyberbullying material. This equips both platform moderators and users with the ability to tackle instances of cyberbullying promptly and proactively, thereby nurturing a more secure and inclusive digital sphere on Twitter.

Ultimately, the objective is to foster a climate of empathy, comprehension, and reciprocal regard, ensuring that Twitter endures as a platform where individuals can openly express themselves devoid of apprehension of mistreatment or harassment. Through ongoing exploration, cooperation, and advocacy, we persist in our commitment to advancing the mission of preventing cyberbullying and creating a more secure online environment for all users.

6.2 Future Scope

There are various opportunities for prospective advancements. Extending the scope of the project to include diverse social media platforms and virtual communities beyond Twitter is a plausible course of action. In addition, the integration of the system directly into current online platforms shows potential for automatically identifying and categorizing instances of cyberbullying, thereby promoting a more secure online milieu. Engaging in partnerships with social media corporations and institutions to incorporate this system within their safety protocols would mark a significant progression. Ultimately, to ensure sustained effectiveness of the system, continuous monitoring and enhancements are imperative to adapt to the evolving strategies employed by online offenders.

References

[1] B. R. Prathap and K. Ramesha, "Twitter Sentiment for Analyzing Different Types of Crimes," 2018 International Conference on Communication, Computing, and Internet of Things (IC3IoT), Chennai, India, 2018, pp. 483-488, doi: 10.1109/IC3IoT.2018.8668140.

[2] S. A. Mathur, S. Isarka, B. Dharmasivam, and J. C. D., "Analysis of Tweets for Cyberbullying Detection," 2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2023, pp. 269-274, doi: 10.1109/ICSCCC58608.2023.10176416.

[3] J. O. Atoum, "Cyberbullying Detection Through Sentiment Analysis," 2020 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2020, pp. 292-297, doi: 10.1109/CSCI51800.2020.00056

[4] P. Dedeepya, P. Sowmya, T. D. Saketh, P. Sruthi, P. Abhijit and S. P. Praveen, "Detecting Cyber Bullying on Twitter using Support Vector Machine," 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2023, pp. 817-822, doi: 10.1109/ICAIS56108.2023.10073658.

[5] M. M. Islam, M. A. Uddin, L. Islam, A. Akter, S. Sharmin and U. K. Acharjee, "Cyberbullying Detection on Social Networks Using Machine Learning Approaches," 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), Gold Coast, Australia, 2020, pp. 1-6, doi: 10.1109/CSDE50874.2020.9411601.

[6] W. A. Prabowo and F. Azizah, “Sentiment Analysis for Detecting Cyberbullying Using Term Frequency Inverse Document Frequency and Support Vector Machine”, J. RESTI (Rekayasa Sist. Teknol.Inf.), vol. 4, no. 6, pp. 1142 -, Dec. 2020.

[7] A. Mody, S. Shah, R. Pimple, and N. Shekokar, "Identification of Potential Cyber Bullying Tweets using Hybrid Approach in Sentiment Analysis," 2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), Msyuru, India, 2018, pp. 878-881, doi: 10.1109/ICEECCOT43722.2018.9001476.

[8] A. Shekhar and M. Venkatesan, "A Bag-of-Phonetic-Codes Model for Cyber-Bullying Detection in Twitter," 2018 International Conference on Current Trends

towards Converging Technologies (ICCTCT), Coimbatore, India, 2018, pp. 1-7, doi: 10.1109/ICCTCT.2018.8550938.

[9] J. Yadav, D. Kumar, and D. Chauhan, "Cyberbullying Detection using Pre-Trained BERT Model," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 1096-1100, doi: 10.1109/ICESC48915.2020.9155700.

[10] J. Wang, K. Fu, C.T. Lu, “SOSNet: A Graph Convolutional Network Approach to Fine-Grained Cyberbullying Detection,” Proceedings of the 2020 IEEE International Conference on Big Data (IEEE BigData 2020), December 10-13, 2020

[11] https://www.kaggle.com/datasets/soorajtomar/cyberbullying-tweets

Appendix A: List of papers presented and published

[1] A. Jalote, K. S. Khan Patan, K. N. Mathew, M. Nandkar and P. More, "A Review on Cyberbullying Detection on Twitter Dataset," 2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), Bengaluru, India, 2024, pp. 1101-1107, doi: 10.1109/IDCIoT59759.2024.10467658.