International Journal of Scientific Research and Engineering Development-– Volume 7 Issue 4, July-Aug 2024

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

NEPALI FAKE NEWS DETECTION USING MACHINE LEARNING ALGORITHMS

Rakshya Sharma[1], Bidhya Sharma[2], Prashikshya Baral[3], Susmita Parajuli[4]
Department of Electronics, Communication and Information Engineering,
Tribhuvan University, Kathmandu, Nepal.
paudel.raxya123@gmail.com

--*************************----------------------------------

ABSTRACT:

The Internet's rapid development makes it possible for information to spread quickly through websites or social networks. Fake or unverified news spreads on social media and reaches thousands of users without anyone questioning its veracity. Misinformation or fabricated news that spreads on social media with the intention of harming a person, group, or agency is known as fake news. Various machine-learning techniques have been used to distinguish fake news from real. In this paper, we tried to achieve this using machine learning techniques and natural language processing. We use frequency-inverse document frequency as a feature extraction method and logistic regression, support vector machine and Naive Bayes as the classifier. We use a dataset with labels for fake and real news to train our model.

Keywords — **Nepali Fake News, Machine Learning, Natural Language Processing, Term frequency-inverse document frequency.**

--*************************----------------------------------

I. INTRODUCTION

The Internet has revolutionized journalism by serving as an essential publishing tool for journalists and providing a primary source for finding the latest news. In recent times, social networks have played a crucial role in spreading news rapidly, often surpassing the speed of traditional media outlets. However, while social networks offer numerous advantages, they also come with certain drawbacks. One significant drawback is the lack of fact-checking that often accompanies the spread of information on social media platforms. Unlike traditional media, where news organizations employ professional journalists who adhere to rigorous fact-checking processes, social media allows anyone to share information without any fact checking. Consequently, the validity and reliability of information circulating on social networks can be compromised. To mitigate this issue, individuals often rely on the collective effort of posting and sharing information from various sources to verify its accuracy, meaning that if a news article is shared many times, it is usually assumed to be true. However, even when a piece of news is shared multiple times with attention-grabbing headlines, it does not guarantee its accuracy. Inaccurate information can still spread rapidly across social media, resulting in the rapid spread of fake news and rumours.

The spread of false information on social media has become a significant social problem as it can profoundly impact a person's perspectives and understanding of a subject. The way humans form opinions are largely based on the information available to them, and if that information is inaccurate, it can lead to misinformed decisions and actions. This misinformation can be intentional, with some stories being designed as propaganda to mislead readers, create biased opinions, manipulate ideas, or generate economic incentives for the writer through clickbait tactics. An example of the spread of fake news is evident in the context of the recent COVID-19 pandemic. Misinformation has circulated regarding various aspects of the disease, including misconstrued understandings of its nature, false claims about remedies or cures, and baseless speculation about the course of the virus's spread. These instances of misinformation on social media highlight the potential harm that can arise when inaccurate information is disseminated widely and rapidly. While the Internet and social media platforms have revolutionized the dissemination of news, their unrestricted nature and lack of robust fact-checking mechanisms have contributed to the spread of false information. This issue has societal consequences, as misinformation can significantly influence people's views and understanding, leading to misguided decisions and actions.

II. OBJECTIVE

- To predict whether a news is true or fake using machine learning approaches.
- To make comparisons among Naïve Bayes classifier, SVM classifier and Logistic Regression classifier for fake news detection and apply the effective ones.

III. LITERATURE REVIEW

There has been a significant amount of research and study in classifying and detecting fake news. Lots of studies have focused on detecting and classifying bogus news on social media sites. Certain entities use social media with negative intentions by spreading fake news for monetary gain or creating bias or promoting absurdity even to defame certain political figures.

Vosoughi et al. investigated the features of news spread on social media; that is, the authors explored the distribution of news (rumours) on social media platforms such as Twitter and analysed how fake news differs from true news in terms of Twitter diffusion. The authors conducted a large-scale analysis of Twitter data, examining the dissemination of news stories and assessing their veracity. They collected around 126,000 stories shared by 3 million users and classified them as true or false using information from reputable fact-checking organizations. The study revealed that false information spreads significantly faster, farther, and more broadly than true information. Falsehoods were found to be 70% more likely to be retweeted than true news, and they reached users six times faster on average. The authors also examined the characteristics of individuals who spread false information and found that they tended to have a smaller number of followers, follow fewer people, and be less active on Twitter compared to those who shared true information. The paper examines the spread of fake news online by discussing the depth, size, maximum breadth, structural virality, mean breadth of true and false rumour cascades at different depths, number of unique Twitter users reached at any depth, and number of minutes it takes for true and false rumour cascades to reach depth and number of Twitter users.[1]

Similarly, H. Allcott and M. Gentzkow, through their journal, examined the role of social media platforms in the spread of fake news during the 2016 United States presidential election. The authors provide a comprehensive review of the existing literature on the subject, analysing the factors that contributed to the prevalence of fake news, its impact on the election, and the potential remedies to address this issue. They discuss the various mechanisms through which fake news spreads on social media, such as the algorithms used by platforms to personalize and recommend content to users, the influence of social networks and their echo chambers, and the potential role of clickbait headlines and confirmation bias.[2]

The 2018 Science article "The science of fake news" by D. M. J. Lazer, M. A. Baum, Y. Benkler, et al. offers a succinct summary of the research on the fake news phenomena. The writers examine fake news's many facets, such as its prevalence, consequences, and methods of distribution. Along with the psychological and cultural variables that contribute to the dissemination and acceptance of false information, they explore the function of social media platforms and the algorithms that define the information ecosystem. The article asks for continuous efforts to promote truth and accuracy in a media environment that is becoming more complicated and highlights the value of interdisciplinary study and collaboration to counteract the harmful impacts of false news.[3]

A thorough summary of the research on fake news detection using machine learning methods is given in the review article "Fake news detection using machine learning algorithms: Hussain, M., Zhang, and Ali's review, which appeared in Soft Computing in 2020. The decision trees, support vector machines, neural networks, and ensemble approaches used in the detection and classification of fake news are all covered by the writers. They examine the performance metrics used to judge these models' efficacy and talk about the features and datasets used for building and testing these models. The analysis draws attention to the difficulties and restrictions of current strategies, including data imbalance and the dynamic nature of fake news. Additionally, it suggests prospective study directions for the future, such as combining social and user behaviour in context to the detection process.[4]

Ahmed et al. extracted linguistic features such as n-grams from textual articles and trained multiple machine-learning models, including K-nearest neighbour (KNN), support vector machine (SVM), logistic regression (LR), linear support vector machine (LSVM), decision tree (DT), and stochastic gradient descent (SGD), with SVM and logistic regression achieving the highest accuracy that is 92%.[5]

Shu et al. went over a variety of veracity assessment approaches for detecting fake news on the internet. The study examines two primary groups of assessment methods: linguistic cue techniques and network analysis approaches. Combining the two comes about in a more robust hybrid approach to detecting bogus news on the internet. SVM or naive Bayes models are used to train classifiers using these language techniques.[6]

Mykhailo Granik et al. proposed a simple method for fake news detection, using naive Bayes classifier. The system was tested against data from Facebook's news posts. It gave an accuracy of 74% on the test data.[7]

Karishnu Poddar et al. the performance of various classifiers, including Logistic Regression, Naïve Bayes, Artificial Neural Network, Decision Tree, and Support Vector Machine, has been compared by the authors in [3]. The output of these classifiers was compared by the author after using a dataset from Kaggle. The author claims that the amount of the dataset has a significant impact on the model's performance. Of all the models employed, SVM with TF-IDF vectorizer appeared to have the highest accuracy, but NN had the lowest accuracy with any of the vectorizers. Additionally, the author claims that Logistic Registration is particularly helpful when the dataset is huge and that it works well in both TF-IDF and Countvectorizer.[8]

IV. SYSTEM DESIGN

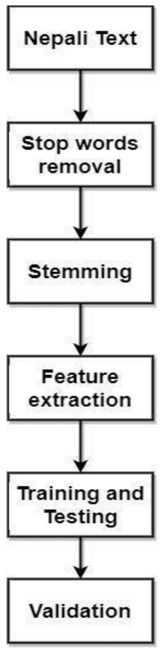

Fig. 1 Workflow Diagram

The fundamental process in the machine learning pipeline is gathering data for training the ML model. The accuracy of the predictions provided by ML systems is only as good as the training data. We collected data through various methods such as, web scraping from different news portals, manual collection and labelling of fake news from various websites and translation of fake news available on the web in English language. The collected data is labelled with a '1' or a '0', where 1 means true and 0 means fake.

Preprocessing text data is necessary in order to transform it into a format suitable for data modelling. Text data can be converted using a variety of widely used techniques, one of which is natural language processing techniques (NLTK). The steps included in Data processing (preprocessing) are: Data cleaning, Stop words removal, punctuation removal, and Stemming. By removing the unnecessary data from the information, this will reduce the actual information size. In the data cleaning step, we remove unwanted columns and find cells with null values and remove the row or replace the null values with empty string, depending on the type of data. In our project, we have two columns: news and label, so we drop rows with null values. Stop words are a set of commonly used words for constructing sentences. For example, in Nepali, "अझै", "अन्य", "जो", "तर", "प्रत्येक" are some stop words. As features for text classification, they are meaningless. We have used the Natural Language Toolkit – (NLTK) library to remove stop words from every item in the dataset so that more focus is given to other important words. Punctuations like commas are removed from text using regular expressions (regex) since they are not important. Stemming is a technique that reduces a word to its base word, called stem, aiding the process of text processing. For example, "घरमा" becomes "घर", "देशको" becomes "देश", "मलाई" becomes "म" after stemming.

Document data needs to be converted into numerical data because it cannot be computed. We have used Term Frequency-Inverse Document Frequency (TF-IDF) vectorizer for this. It is a statistical measure used in natural language processing and information retrieval to evaluate the importance of a word in a document. TF-IDF takes into account the frequency of a word in a document (Term Frequency) and the inverse frequency of the word in the entire corpus of documents (Inverse Document Frequency). This is done to give more weight to words that are frequent in a document, but rare in the entire corpus, as those words are more likely to be meaningful and informative. The TF-IDF score is used to rank documents based on their relevance to a given query in a search engine, or to extract key features from a corpus in text classification and clustering tasks.

First, we began preparing the data collection by eliminating characters and words that weren't needed or significant. Feature extraction using term frequency-inverse document frequency is the next stage. We looked into three different algorithms: Passive Aggressive Classifier, Naïve Bayes, and Support Vector Machines (SVM). We used the Python Natural Language Toolkit (NLTK) to implement these classifiers. The dataset was divided into preparation and testing sets. 20% of the dataset is used for testing purposes and the remaining 80% is used for preparation.

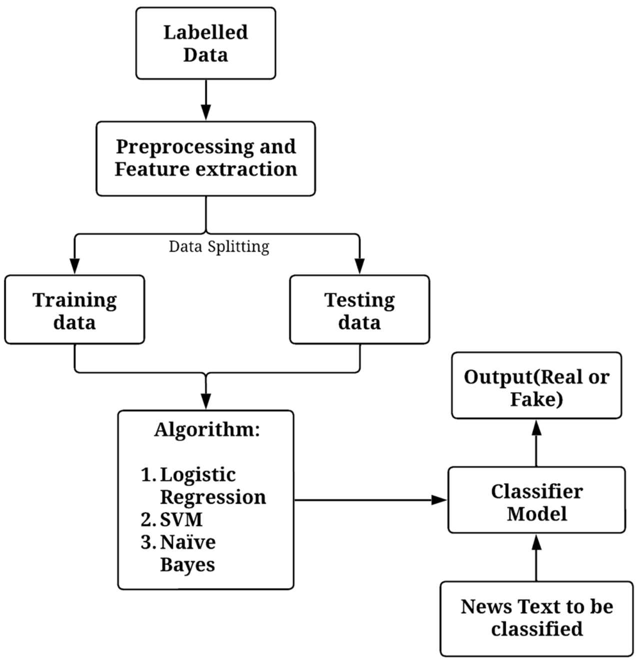

Fig. 2 System Block Diagram

In order to construct the model, we employed the following classification algorithms to categorize news articles into true and fake categories. These algorithms analyse the features and patterns within the data and assign a label to each instance, indicating whether it is likely to be fake or true.

A. *Naïve Bayes*

Naïve Bayes is a conditional probability model which can be used for labelling. The goal is to find a way to predict the class variable (B) using a vector of independent variables (A), i.e., finding the function f: A→B. To put it in probability words, the objective is to determine P(B|A), or the likelihood that B belongs to a particular class A. It is typically believed that B has two or more discrete values and is a categorical variable. This is a basic mathematical method to incorporate several factors' contributions when predicting the class of the subsequent data instance in the testing set. The assumption that no feature is dependent on any other characteristic is the limitation of Naïve Bayes.[11]

Input:

- A document d
- fixed set of classes C ={c1,c2}
- A training set of m hand-labelled document (d1, c1) ,..., (dm, cm)

Output:

- A learned classifier: γ: d→c

Bayes' Rule Applied to Documents (d) and Classes (c):

$$P(c|d)=\frac{P(d|c)P(c)}{P(d)}$$

B. *Support Vector Machine (SVM) classifier*

Support Vector Machines (SVM) are an arrangement of related supervised learning techniques operated for grouping and classification based on statistical approach. It establishes the optimal border for decision-making between vectors that are part of a specific group (category) and vectors that are not. This implies that the texts must be converted into vectors in order to fully utilize the potential of SVM text categorization. In our project we used Tfidfvectorizer to transform the text into numeric vectors. Vectors are (sometimes huge) lists of numbers which represent a set of coordinates in some space in SVM. SVM determines the optimal location of the "line" (also known as the best hyperplane) that splits the space into two subspaces: one for the vectors that fall within the specified category and another for the vectors that do not.[9]

The mathematically displayed expression for SVM can be represented as follows:
Given a training dataset D with N data points, each represented by a feature vector $x_i \in \mathbb{R}^n$ and a corresponding class label $y_i \in \{-1, 1\}$,

$$D = \{(x_i, y_i) \mid x_i \in \mathbb{R}^n, y_i \in \{-1, 1\}\}_{i=1}^{n}$$

The goal of SVM is to find a hyperplane that maximally separates the data points of different classes. The decision function of SVM can be expressed as:

$$f(x) = \text{sign}(w \cdot x + b)$$

where:
f(x) is the predicted class label for the input feature vector x, w is a weight vector that is perpendicular to the hyperplane, and b is the bias term. The weight vector w and bias term b are learned during the training process.

To find the optimal hyperplane, SVM aims to solve the following optimization problem:

$$\text{minimize: } \tfrac{1}{2}\|w\|^2 + C\sum \xi_i$$

$$\text{subject to: } y_i(w \cdot x_i + b) \geq 1 - \xi_i \text{ for all } i$$

where:
$\|w\|$ is the norm of the weight vector w.
C is a hyperparameter that controls the trade-off between maximizing the margin and minimizing the training errors.
ξ_i are slack variables that allow for misclassifications.

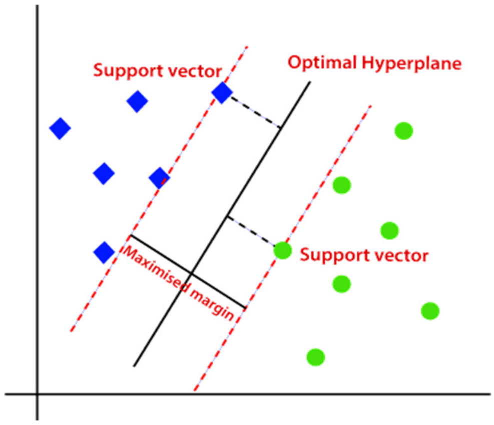

Fig. 3 SVM algorithm graph

C. *Logistic Regression Classifier*

Logistic regression is a statistical method used to develop machine learning models. It is used for binary classification problems. For instance, predicting whether an email is spam or not.[10]. The term logistic regression is derived from the "logistic function" (or sigmoid function) that it uses.

Sigmoid Function

Model

Output = 0 or 1

Hypothesis => Z = WX + B

H(x) = sigmoid (Z)

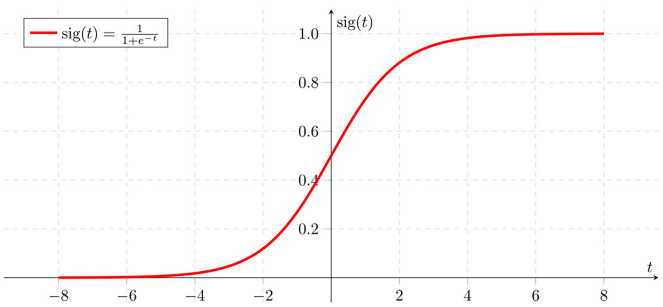

Fig. 4 Graph showing sigmoid function

The predicted value of Y will become 1 if 'Z' reaches infinity, and 0 if 'Z' reaches negative infinity.

V. PERFORMANCE EVALUATION

In order to evaluate the performance of the machine learning model, we have used different metrics on the datasets. The metrics consist of accuracy, precision, recall, and F1-score. These metrics are based on the confusion matrix, which is a performance measurement where the output can be two or more classes – in this case, the output will be fake and real news.

		Predicted Class	
		Positive	**Negative**
True Class	**Positive**	TP	FN
	Negative	FP	TN

Fig. 5 Confusion Matrix

The confusion matrix is a table with 4 different classifications of predicted and actual values: TP (true positive) determines the articles that have been correctly classified as real news, FN (false negative) determines the article is true and predicts it as fake news, FP (false positive) determines the article is fake and predicts it as real news and TN (true negative) determines the articles that have been correctly classified as fake news.

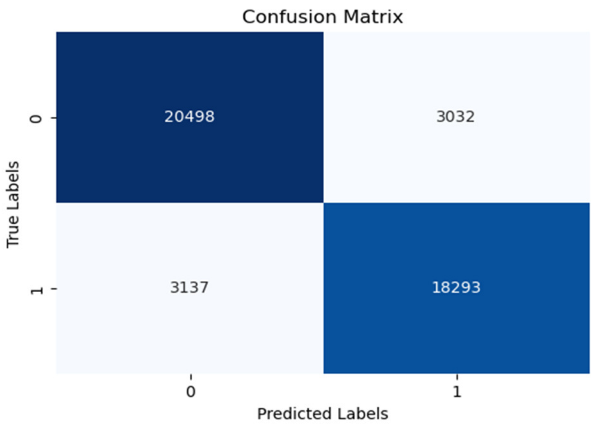

Fig. 6 Confusion Matrix of Naïve Bayes model

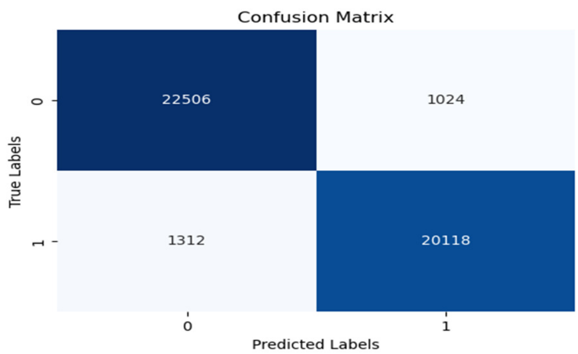

Fig. 7 Confusion Matrix of SVM model

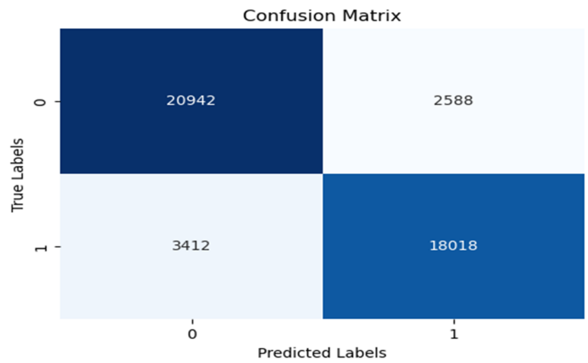

Fig. 8 Confusion Matrix of Logistic Regression model

VI. RESULT

After entering the website we created and navigating to the check news page, the input news which was to be checked was entered in the text box, and the model predicted that the news was fake or true according to the calculation from the model, which is displayed below the text box.

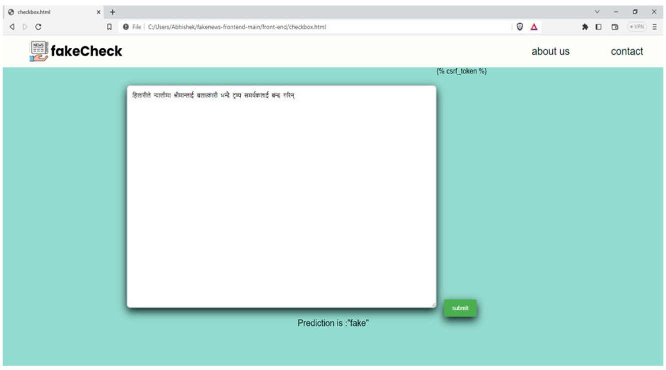

Fig. 9 News predicted as 'fake'

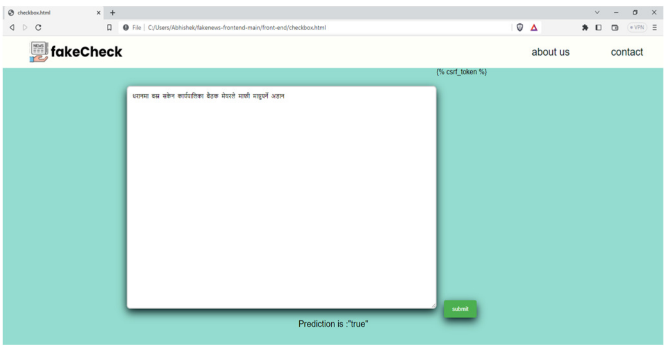

Fig. 10 News predicted as 'true'

Our project on fake news detection utilized Support Vector Machine (SVM), Logistic Regression, and Naïve Bayes models. Analysing the performance evaluation matrices, Support Vector Machine algorithm proved to be the most accurate and reliable in identifying fake news, followed closely by Logistic Regression and Naïve Bayes. This project contributes to combating misinformation by providing effective tools for detecting fake news articles, ensuring that users are better informed and can trust the information they consume online, thereby promoting a more accurate and truthful media landscape.

TABLE I
COMPARISON OF PERFORMANCE OF VARIOUS CLASSIFIERS

Classifier	**Accuracy**	**Precision**	**Recall**	**F1-score**
Naïve Bayes	0.87	0.86	0.85	0.85
SVM	0.89	0.95	0.94	0.95
Logistic Regression	0.86	0.87	0.84	0.86

ACKNOWLEDGEMENT

We would sincerely like to show our gratitude to our Head of Department of Electronics and Computer Engineering Er. Laxmi Prasad Bastola Sir for encouraging and allowing us to present the project on the topic "Nepali fake news detection using machine learning" at our department premises for the partial fulfillment of the requirements. Also, we are grateful towards our project supervisor Er. Smita Adhikari Madam for her active guidance, cooperation, encouragement and help in the project. We would also like to send our regards to faculty members, teachers and individuals who helped us with their support and suggestions for the making of this report and project.

REFERENCES

[1] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," Science, vol. 359, no. 6380, pp. 1146–1151, 2018.

[2] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," Journal of Economic Perspectives, vol. 31, no. 2, pp. 211–236, 2017.

[3] D. M. J. Lazer, M. A. Baum, Y. Benkler et al., "The science of fake news," Science, vol. 359, no. 6380, pp. 1094–1096, 2018.

[4] M. Hussain, W. Zhang and A. Ali, "Fake news detection using machine learning algorithms: A review." Soft Computing, vol.24, no.6, pp. 4143-4166, 2020.

[5] H. Ahmed, I. Traore, and S. Saad, "Detection of online fake news using n-gram analysis and machine learning techniques," in Proceedings of the International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, pp. 127–138, Springer, Vancouver, Canada, 2017.

[6] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media," ACM SIGKDD Explorations Newsletter, vol. 19, no. 1, pp. 22–36, 2017.

[7] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2017, pp. 900-903, doi: 10.1109/UKRCON.2017.8100379.

[8] K. Poddar, G. B. Amali D. and K. S. Umadevi, "Comparison of Various Machine Learning Models for Accurate Detection of Fake News," 2019 Innovations in Power and Advanced Computing Technologies (i-PACT), 2019, pp. 1-5, doi: 10.1109/i-PACT44901.2019.8960044.

[9] A. Kowalczyk."SVM- Understanding the math - the optimal hyperplane." SVM Tutorial. https://www.svm-tutorial.com/2015/06/svm-understanding-math-part-3/ (accessed May, 2023)

[10] S. Swaminathan."Logistic Regression- Detailed overview." TowardsDataScience. https://towardsdatascience.com/logistic-regression-detailed-overview-46c4da4303bc (accessed Oct, 2022)

[11] "Naïve Bayes classifiers." IBM. https://www.ibm.com/topics/naive-bayes (accessed Feb, 2023)

[12] A. Holovaty and J. K. Moss, *The Definitive Guide to Django: Web Development Done Right*, 2nd ed.,Apress Berkeley, CA, 2009.

[13] J. Moon."A Beginner's Guide to Splitting Machine Learning Datasets in Python." Medium. https://medium.com/@moonchangin/how-to-split-machine-learning-datasets-in-python-scikit-learn-d0a60cc30929#:~:text=Splitting%20the%20dataset%20To%20split%20the%20dataset%20into,which%20is%20used%20to%20evaluate%20the%20model%27s%20performance

