**Fake News Detection with Dynamic Model Updates Based on Classifier Comparison**

**INTRODUCTION**

The proliferation of digital platforms has led to an unprecedented rise in the dissemination of fake news, resulting in widespread misinformation and public mistrust. Detecting fake news has become critical in maintaining the integrity of information consumed by the public. Existing fake news detection systems primarily focus on static model training, where models are trained on a fixed dataset and remain unchanged over time. However, the dynamic and evolving nature of misinformation demands systems that can adapt continuously.

Currently, fake news detection models suffer from performance degradation as new styles, formats, and types of misinformation emerge. Static models become outdated, leading to reduced accuracy and reliability. Furthermore, there is a noticeable gap in the industry and research: the absence of dynamic model updating mechanisms. To date, minimal efforts have been directed towards dynamically updating fake news detection models to maintain relevance with real-time data.

Addressing this gap, the present study proposes a system for dynamic model updating, wherein the model is periodically retrained with newly collected real and fake news articles. By doing so, the model remains aligned with current trends, improving predictive accuracy and robustness over time. The proposed methodology systematically integrates data collection, preprocessing, model training, evaluation, and dynamic updating to address this pressing research gap.

**NEED FOR STUDY / MOTIVATION / BACKGROUND**

With the increasing sophistication of misinformation campaigns, fake news has evolved beyond traditional detection methods. Static detection models, once deployed, are not updated with new data unless manually retrained, leading to obsolescence and decreased detection capability.

Key motivations for this study include:

* **Evolution of Fake News**: Language patterns, news topics, and misinformation strategies evolve rapidly, rendering static models ineffective.
* **Lack of Dynamic Updating**: Current industry solutions and research implementations largely ignore the need for continuous model updating.
* **Real-World Applicability**: A dynamic updating system ensures the model remains contextually accurate and robust, enhancing its real-world applicability.
* **Resource Optimization**: Automated dynamic updating reduces the manual overhead associated with retraining and redeployment of models.

Thus, there is a critical need for developing a dynamic fake news detection framework that can adapt to real-time changes in news patterns, providing a more reliable solution than existing static models.

**RESEARCH OBJECTIVES**

The objectives of this research are outlined as follows:

1. **To identify and address the gap** in existing fake news detection systems concerning dynamic model updating.
2. **To design a dynamic fake news detection system** capable of periodic model retraining with newly collected data.
3. **To implement a comprehensive methodology** comprising data collection via web scraping, text preprocessing, feature extraction using TF-IDF, and machine learning-based classification.
4. **To evaluate multiple models** (Support Vector Machine, Logistic Regression, Naïve Bayes) and select the best-performing model based on key metrics such as accuracy, precision, recall, and F1-score.
5. **To integrate dynamic model updating** by retraining the selected model on a weekly basis with updated datasets to ensure sustained detection performance.

**Literature Review:**

Table –

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Ref | Title | Publication Year | Methodology | Positive points of publication | Gaps in publication work |
| 1. | Detecting and Mitigating the Dissemination of Fake News: Challenges and Future Research Opportunities – Wajiha Shahid, Bahman Jamshidi, Saqib Hakak, Haruna Isah, Wazir Zada Khan, Muhammad Khurram Khan, Kim-Kwang Raymond Choo (Journal) | 2024 | Here’s an even shorter version, focused only on methodology:  The paper uses GloVe embeddings with a CNN model (FNDNet) for fake news detection, a CNN-LSTM model for stance detection, and applies feature extraction with PCA and Chi-Square for dimensionality reduction. | The authors addressed most common and critical challenges encountered by other researchers to make an efficient classifier  They also addressed solutions to those challenges. | The paper lacks real-time implementation like the authors could have implemented any one proposed model for clear understanding. |
| 2. | Enhancing Fake News Detection by Multi-feature classification – Ahmed Hashim Jawad Almarashy,  Mohammad-Reza Feizi-Derakhshi, Pedram Salehpour (Journal) | 2023 | The paper combines TF-IDF (global features), CNN (spatial features), and BiLSTM (temporal features) using early fusion, and classifies the result with a Fast Learning Network (FLN) for fake news detection. | The authors proposed a model that integrates global, spatial and temporal features of text for enhancing the accuracy  The hybrid approach accurately classifies fake news, even when dealing with datasets having large number of training samples | The model is tested only on English Language dataset. |
| 3. | Real-Time Fake News Detection Using Big DataAnalytics and Deep Neural NetworkMuhammad Babar, Member, IEEE, Awais Ahmad, Member, IEEE,Muhammad Usman Tariq, Member, IEEE, and Sarah Kaleem | 2023 | The paper uses a hybrid N-gram and LSTM model for fake news detection, deployed on a big data platform for parallel and distributed processing to achieve real-time, accurate classification. | The authors combined DNNs and blockchain for a highly accurate and secure fake news detection system, ensuring source credibility and traceability  They built a real-time detection system using big data analytics, enabling fast,  large-scale news analysis with minimal delays. | The model does not focus on language nuances such as sarcasm, satire, and ambiguous phrasing, leading to misclassification. |
| 4. | Fake News Detection Using Enhanced BERT Shadi A. Aljawarneh | 2022 | Here’s the short summary of the methodology:  The paper fine-tunes the BERT model on a fake and real news dataset to enhance its detection capabilities, achieving 99.96% accuracy and outperforming other models. | The optimized BERT model improves language understanding, reducing errors from misleading or satirical headlines.  Enhanced BERT outperforms traditional models, boosting accuracy and feature extraction for better fake news detection.  3 | It does not address the high computational cost of BERT,making real-time large-scale detection difficult. |
| 5. | Fake News in Virtual Community, Virtual Society, and Metaverse: A Survey  Jinxia Wang , Stanislav Makowski, Alan Cieślik, Haibin Lv , Senior Member, IEEE,  and Zhihan Lv , Senior Member, IEEE | 2023 | The paper provides a survey on fake news in virtual communities and the metaverse, analyzing its manifestation in single-modal and multimodal forms, and reviewing detection methods. It also discusses future directions for intelligent detection and information security in these environments. | The authors explore fake news challenges in emerging digital spaces, highlighting new threats in virtual societies and the metaverse.  The paper provides a comprehensive comparison of existing fake news detection techniques, identifying key gaps and future research directions. | The study provides a broad survey of fake news in virtual environments but lacks practical implementation strategies for real-time detection and mitigation.  Solution: Developing AI-driven real-time monitoring systems tailored for virtual communities and the metaverse can help detect and counter fake news more effectively. |

**Gap Areas:**

The current research lacks the feature if dynamic model updating, so we are proposing a system for dynamic model updating, that is updating the model on regular intervals to keep it up to date with current news articles to produce better prediction outcome.

**Methodology:**

The project will employ a systematic methodology that includes collection of data, preprocessing, training models, evaluation, and taking the best model into consideration for dynamic model updating. We are addressing the noted research gaps of dynamic model updating in the methodology. The proposed system architecture is shown in Fig. 1.

1. Data Collection

We are collecting real news data from legitimate news websites through web scrapping and labelling them as 0. The fake news is scrapped through some legitimate websites that have a category for identified fake news, the rest of the fake news are created by using any GenAI model, prompting it to create a paraphrased fake version of some real news articles. The fake news articles data is labelled as 1. The details that we are scrapping from news websites are: article title, description and the date on which the article was scraped.

|  |  |  |
| --- | --- | --- |
| Sr no. | Feature name | Description |
| 1 | title | Title if the news article |
| 2 | text | Description if the news article |
| 3 | label | Label of the news article; fake: 1 and real: 0 |
| 4 | date | Date on which the article was web scrapped |

Table 1. Dataset Features

1. Text Preprocessing

To prepare textual data for analysis, a preprocessing function is developed, which involves removing stop-words, tokenizing text, converting to lowercase, eliminating nonalphabetic characters, and applying stemming or lemmatization. The dataset that we have created goes through pre-processing function.

1. Feature extraction

Feature extraction from textual data can be accomplished by using word embeddings or TF-IDF vectorization (Term Frequency-Inverse Document Frequency). By converting text into numerical attributes, we can determine the semantic content and word importance of each document. In our implementation, we utilized the TfidfVectorizer from the scikit-learn library, configuring it with specific parameters to enhance feature extraction:​

* N-gram Range: Set to (1, 2) to capture both unigrams and bigrams, allowing the model to consider individual words and pairs of consecutive words, thereby incorporating some contextual information.​
* Maximum Document Frequency (max\_df): Set to 0.9 to exclude terms that appear in more than 90% of the documents, as such terms are likely to be non-informative and may not contribute to distinguishing between documents.​
* Minimum Document Frequency (min\_df): Set to 2 to ignore terms that appear in fewer than two documents, reducing the impact of rare terms that may not be relevant for the analysis.​

The vectorizer was fitted to the corpus and the textual data was transformed into a TF-IDF weighted term-document matrix, which served as the input for subsequent analytical processes.

1. Model Training

The pre-processed features are divided into training and testing groups using a suitable splitting technique, such as a random split or cross-validation. We have utilized random split. When selecting a machine learning model, the goals of the research and the characteristics of the dataset are taken into account. We have trained the data on three ML models, Support Vector Machine (SVM), Logistic Regression (LR) and Naïve bayes (NB).

Support Vector Machine (SVM) is an effective classification method that searches for the best hyperplane in a multidimensional space to separate different classes. It is robust against outliers and efficiently handles both linear and non-linear decision boundaries through the use of kernel functions. By maximizing the margin between classes, SVM improves generalization on unseen data. Its adaptability to different data types and the ability to control overfitting through regularization parameters make SVM a popular choice for various classification tasks.

Logistic Regression is a widely used statistical method for binary and multiclass classification problems. It models the probability of a data instance belonging to a particular class using a logistic (sigmoid) function. The model estimates the relationship between input features and the log-odds of the outcome, making it both interpretable and computationally efficient. Logistic Regression performs best when the classes are linearly separable and includes regularization techniques to prevent overfitting, allowing it to generalize well to new data.

Naive Bayes is a simple yet powerful probabilistic classification algorithm based on Bayes’ Theorem. It assumes that the features are conditionally independent given the class label, which simplifies the computation significantly. Despite its 'naive' assumption, Naive Bayes often performs competitively, particularly in text classification and spam detection tasks. It is highly efficient, requires a small amount of training data, and works well even when the independence assumption is somewhat violated in practice.

After training these models, we calculate the evaluation metrics such as accuracy, precision, recall and f1- score. The model with best accuracy is selected as the final model for predicting fake news.

For dynamic model updating, we are scrapping the real news from websites on daily basis and appending it to our dataset, and training the model weekly on this newly obtained dataset to improve the accuracy of the model based on latest news.

**Assumptions used:**

We are assuming that the news taken from legitimate news website is real or true.

**Result and Analysis:**

The study examines the LR, SVM and NB methodologies' performance using our dataset. The models' performance is evaluated using key performance indicators like F1-score, recall, precision, and accuracy. The method with best accuracy is used for dynamic model updating.

It is observed that SVM performs better compared to LR an NB so it is selected as the final model. This SVM model is then trained periodically to increase accuracy for predicting latest fake news. Table 2, 3, 4 and 5 show the evaluation metrics for SVM such as accuracy, precision, recall and f1-score, for two weeks of data. The model was trained on new collected data once a week. The graphs show the same visualization of evaluation metrics.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Train Test Split**  **80-20** | Accuracy | Precision | Recall | F1-score |
| Week 1 | 0.7826 | 0.7647 | 0.9286 | 0.8387 |
| Week 2 | 0.6700 | 0.7500 | 0.5660 | 0.6452 |

Table 2. Evaluation metrics for SVM with 80-20 train-test split

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Train Test Split**  **70-30** | Accuracy | Precision | Recall | F1-score |
| Week 1 | 0.7143 | 0.6786 | 0.9500 | 0.7917 |
| Week 2 | 0.6711 | 0.7183 | 0.6375 | 0.6755 |

Table 3. Evaluation metrics for SVM with 70-30 train-test split

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Train Test Split**  **60-40** | Accuracy | Precision | Recall | F1-score |
| Week 1 | 0.6739 | 0.6486 | 0.9231 | 0.7619 |
| Week 2 | 0.6131 | 0.6304 | 0.5743 | 0.6010 |

Table 4. Evaluation metrics for SVM with 60-40 train-test split

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Train Test Split**  **50-50** | Accuracy | Precision | Recall | F1-score |
| Week 1 | 0.7193 | 0.6818 | 0.9375 | 0.7895 |
| Week 2 | 0.6371 | 0.6230 | 0.6333 | 0.6281 |

Table 5. Evaluation metrics for SVM with 50-50 train-test split

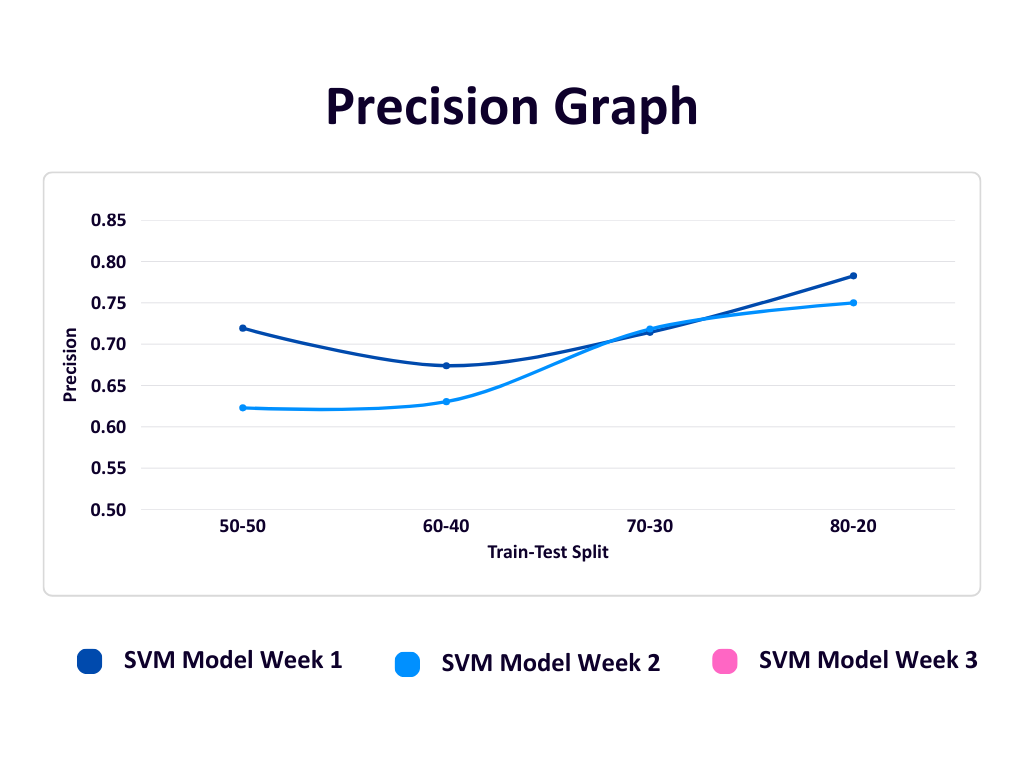


Figure 2

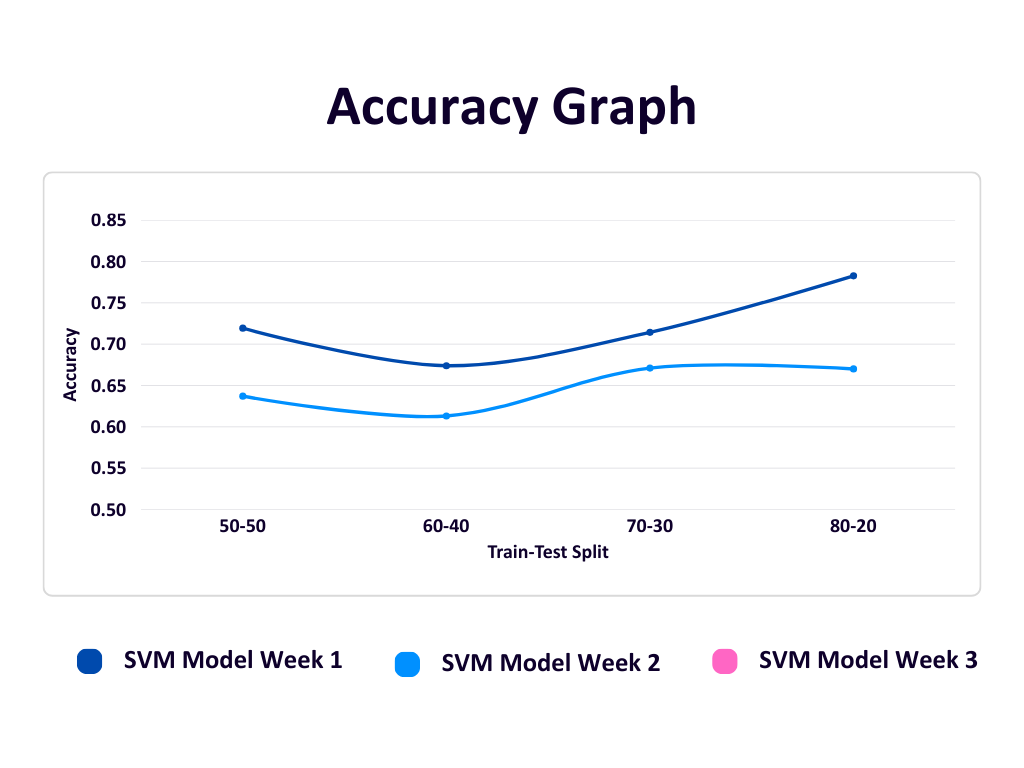


Figure 3

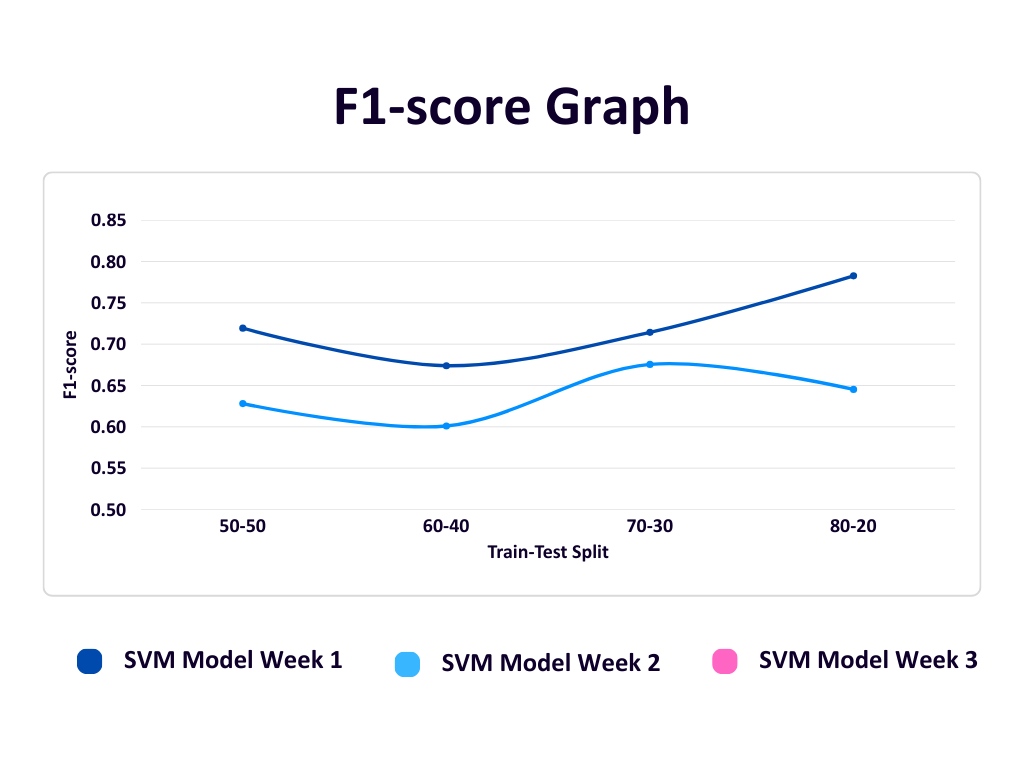


Figure 4

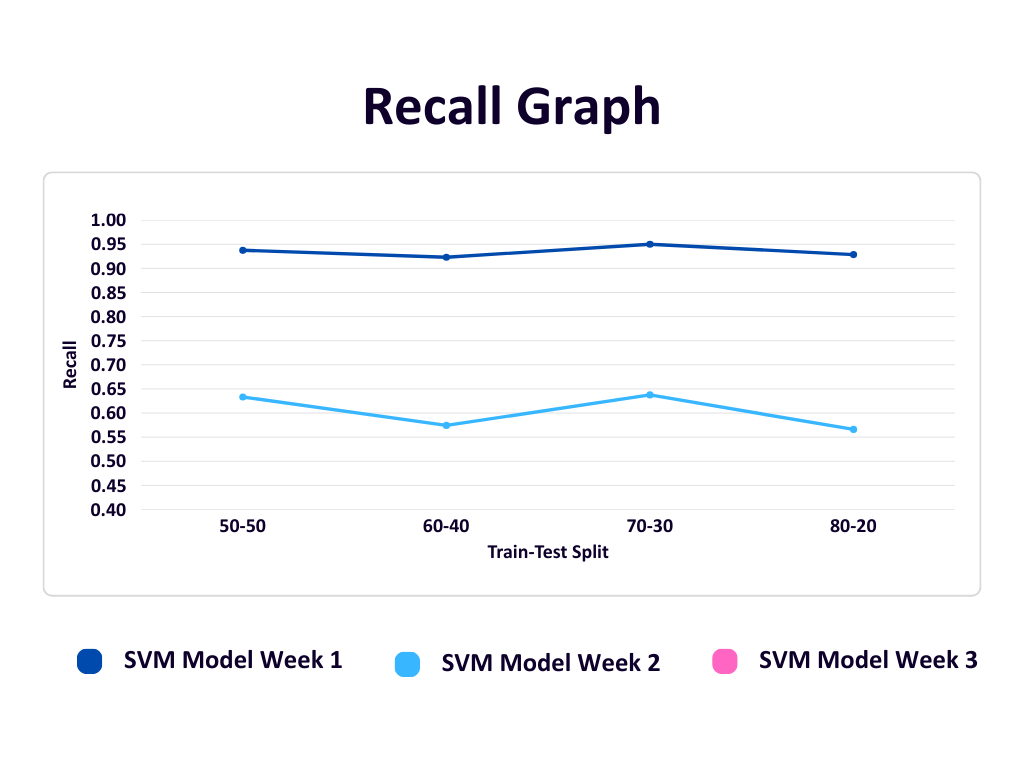


Figure 5

**Conclusion**

In this paper, we developed a fake news detection system using machine learning techniques such as Logistic Regression, Support Vector Machine (SVM), and Naive Bayes. The system evaluates news article text to determine its authenticity. To ensure ongoing relevance and accuracy, we implemented dynamic model updating by selecting the best-performing algorithm based on accuracy evaluations. The model is retrained using newly collected data obtained through web scraping, allowing it to adapt to emerging trends and linguistic patterns in fake news. This dynamic approach significantly enhances the model’s reliability and robustness over time, making it a practical solution for combating the evolving challenge of misinformation.

**Future Scope**

The current system provides a strong foundation for further enhancements aimed at improving the adaptability and effectiveness of fake news detection. One potential direction is the automation of model selection, wherein the optimal algorithm can be dynamically chosen based on comprehensive performance metrics such as precision, recall, and F1-score, rather than accuracy alone. Furthermore, expanding the scope of web scraping to incorporate diverse data sources, multilingual content, and multimedia formats, such as images and videos, can substantially enrich the training dataset and enhance the model's generalization capabilities. The integration of advanced deep learning architectures, including Long Short-Term Memory (LSTM) networks and transformer-based models like BERT, may also contribute to improved contextual understanding and classification accuracy.