**Project Proposal**

**Title:**

**Fake News Detection using TF-IDF and Machine Learning Models: A Comparative Study**

**1. Introduction**

Spreading of misinformation and fake news, particularly on social and news websites is a serious threat to society worldwide. Fake news impacts on societal behavior, political clothes, as well as the influences of opinion. Since digital information is becoming more and more easy to access and it is being generated in huge amounts, ensuring that there is a way of automatically detecting and filtering fake content has become very important.

Machine Learning (ML) offers good possibilities of classifying the content of news items in terms of being fake or not depending on the content. Properly prepared and preprocessed text data may be used to formulate and train classification models based on textual data, after converting its textual features to numerical features e.g. with the help of TF-IDF (Term FrequencyInverse Document Frequency). The models among them that have received a wide use are logistic regression and naive Bayes because they are simple, fast and efficient in assigning binary text classification tasks.

This study seeks to pre-process a publicly available fake news dataset; then create easy ML models to categorize news articles. Various models will be evaluated in comparison to understand which approach performs best by relying on text-based indicators such as TF-IDF.

**2. Problem Statement**

Though fake news detection is increasingly becoming a necessity, most developing nations and organizations cannot ensure that they have some good systems to detect false information. Verification of manual work is lengthy and nonsensical where the data of many volumes is concerned. Although there are deep learning and complex models, they are computationally demanding and need huge data.

The question that should be investigated is whether the simpler and computationally efficient ML models could remain effective when used in conjunction with suitable methods of text preprocessing and feature extraction, such as TF-IDF. This study will present an answer to that question by comparing Logistic Regression and Naive Bayes models over a publicly available dataset.

**3. Research Objectives**

The main objectives of this research are:

1. To preprocess and clean a publicly available fake news dataset for analysis.
2. To extract numerical features from text using the TF-IDF technique.
3. To develop simple classification models (Logistic Regression and Naive Bayes).
4. To evaluate and compare the performance of these models using common classification metrics.
5. To determine whether basic models with proper feature engineering can achieve acceptable accuracy.

**4. Research Questions**

* What preprocessing steps are essential to prepare text data for fake news detection?
* How effective is TF-IDF in transforming news articles into useful feature vectors?
* Which model—Logistic Regression or Naive Bayes—performs better for binary fake news classification using TF-IDF?
* What are the limitations of using simple models for this task?

**5. Literature Review**

Fake news detection is a growing field in Natural Language Processing (NLP). Studies by Shu et al. (2017) and Zhou & Zafarani (2018) show that text-based features are useful for detecting fake news in the absence of metadata like source credibility or user behavior.

TF-IDF remains one of the most popular techniques for feature extraction in text classification tasks. It weighs words according to their frequency in a document and their rarity across the corpus, emphasizing important and distinctive terms.

Naive Bayes classifiers have been widely used in spam and fake news detection due to their simplicity and performance with high-dimensional data. Logistic Regression is another strong linear classifier that performs well with TF-IDF features.

However, limited comparative studies have been conducted on fake news detection using simple models with TF-IDF features. This research will fill that gap by conducting a detailed comparative evaluation.

**6. Methodology**

**6.1 Dataset Selection**

* Use a publicly available dataset such as the Fake and Real News Dataset on Kaggle.
* The dataset typically contains article text, titles, authors, and labels (fake or real).

**6.2 Data Preprocessing**

1. **Data Cleaning**:
   * Remove missing values and duplicates.
   * Drop irrelevant columns (e.g., author if too sparse).
2. **Text Normalization**:
   * Lowercasing
   * Removing punctuation, stopwords, and special characters
   * Tokenization
   * Optional: stemming or lemmatization
3. **Splitting Data**:
   * Split into training and test sets (e.g., 80/20 split).

**6.3 Feature Extraction (TF-IDF)**

* Use TfidfVectorizer from scikit-learn to convert text into feature vectors.
* Adjust hyperparameters such as max\_features, ngram\_range, and min\_df for optimization.

**6.4 Model Development**

* **Logistic Regression**:
  + Binary classification model trained on TF-IDF features.
  + L2 regularization may be used to avoid overfitting.
* **Multinomial Naive Bayes**:
  + Suitable for word count or frequency-based features like TF-IDF.
  + Fast and effective with high-dimensional data.

**6.5 Evaluation Metrics**

* Accuracy
* Precision
* Recall
* F1 Score
* Confusion Matrix
* ROC-AUC Curve

Models will be compared based on these metrics on the test set.

**7. Expected Outcomes**

* A fully functional fake news detection pipeline using basic ML models and TF-IDF.
* Comparative analysis showing which model performs better for the task.
* Demonstration that even simple models can perform well when combined with the right text features.
* A possible basis for more advanced research involving ensemble or deep learning models.

**8. Tools and Technologies**

* **Language**: Python
* **Libraries**:
  + pandas, NumPy (data processing)
  + scikit-learn (ML models and TF-IDF)
  + matplotlib, seaborn (visualization)
* **Platform**: Jupyter Notebook / Google Colab
* **Dataset Source**: Kaggle or other public repositories

**9. Timeline**

| **Task** | **Duration** |
| --- | --- |
| Literature Review | Weeks 1–2 |
| Dataset Collection & Cleaning | Weeks 3–4 |
| Text Preprocessing & TF-IDF | Week 5 |
| Model Development | Weeks 6–7 |
| Model Evaluation | Week 8 |
| Analysis & Report Writing | Weeks 9–10 |

**10. Limitations**

* Simple models may fail in subtle or context-sensitive fake news scenarios.
* TF-IDF ignores word order and semantic meaning.
* Results may vary based on dataset size and balance.
* Advanced features like author credibility, social sharing, or sentiment are not considered.

**11. Future Work**

* Incorporate deep learning methods like LSTM or BERT for improved accuracy.
* Add metadata features such as publishing time, source, and engagement.
* Explore ensemble models like Random Forest and Gradient Boosting.
* Deploy a web-based demo or API for real-time fake news classification.

**12. References**

1. Shu, K., et al. (2017). Fake News Detection on Social Media: A Data Mining Perspective. *ACM SIGKDD Explorations*.
2. Zhou, X., & Zafarani, R. (2018). Fake News Detection: A Survey. *ACM Computing Surveys*.
3. Scikit-learn documentation: <https://scikit-learn.org/>
4. Fake News Dataset: https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset
5. Sebastiani, F. (2002). Machine Learning in Automated Text Categorization. *ACM Computing Surveys*.