

Summary Report for Preliminary Data Validation and Baseline Model Evaluation

Project Title: AI-Based Fake News Detection System Using Machine Learning and Natural Language Processing in the Ugandan Context

1. Introduction

This summary presents the preliminary findings of the AI-Based Fake News Detection research project, focusing on **data availability, suitability, and initial model performance**. The primary objective at this stage was to address concerns regarding whether sufficient and reliable data could be obtained to support the development of an effective fake news detection system tailored to the Ugandan context.

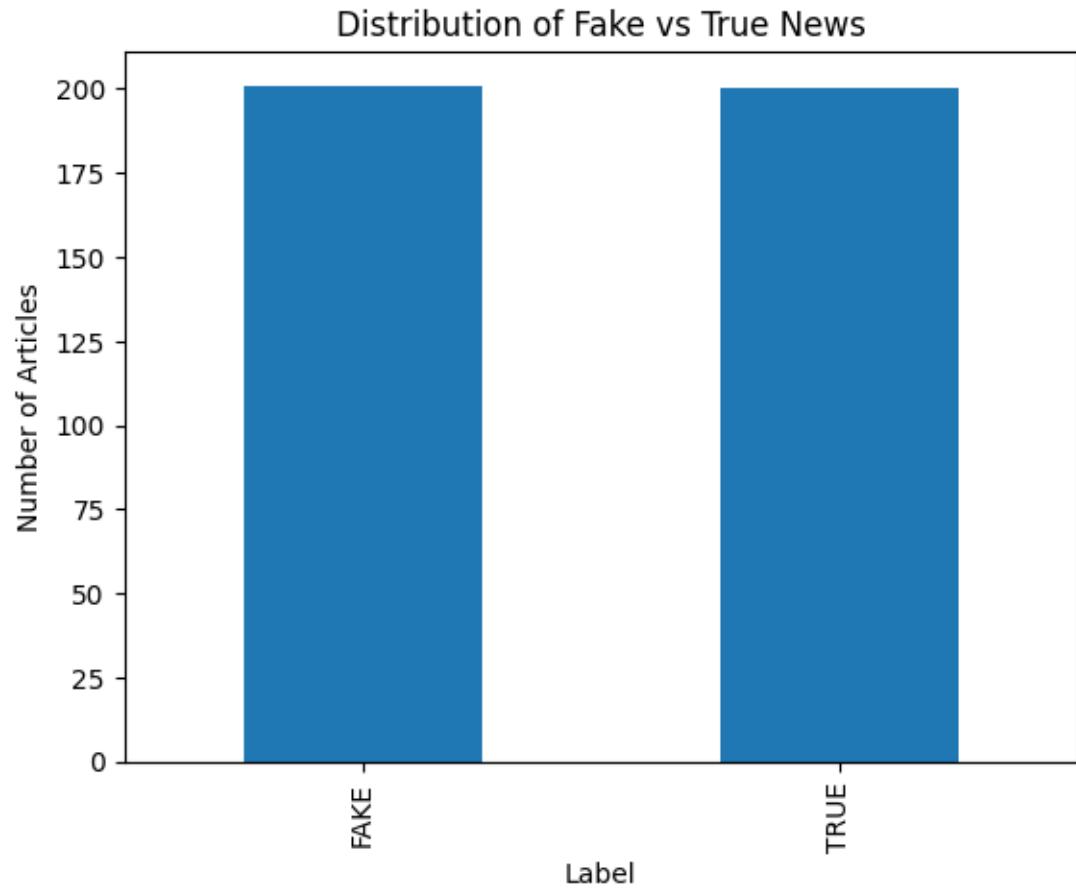
2. Dataset Collection and Preparation

A balanced textual dataset was manually collected from **credible fact-checking organizations** and **established news outlets**, ensuring data reliability and label correctness.

I collected a balanced dataset of 400 Ugandan news items. Exploratory analysis shows meaningful linguistic structure, balanced class distribution, realistic language use, and credible labeling sources. The data is sufficient and suitable for training machine learning models for fake news detection.

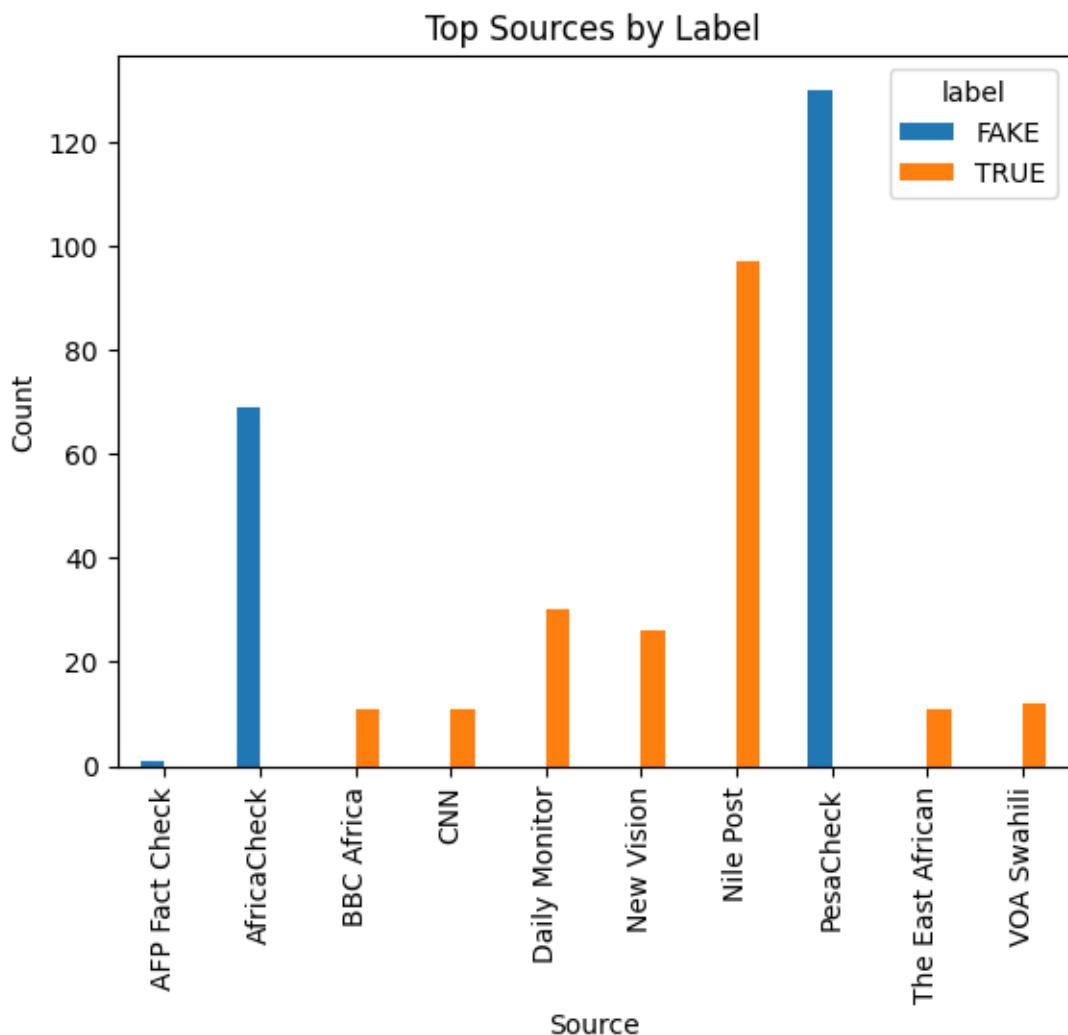
The dataset is perfectly balanced with an equal number of fake and true news articles (200 each). This balance prevents model bias toward a dominant class and supports fair evaluation of classification performance.

- **Total samples:** 400
- **Fake news:** 200 samples
 - Sources: AfricaCheck, PesaCheck



- **True news:** 200 samples
 - Sources: Daily Monitor, New Vision, Reuters, BBC Africa, VOA, AFP, CNN
- **Content type:** Text-based news headlines and short article summaries
- **Languages:** English and mixed English–Swahili content

- **Labels:** Binary classification (FAKE / TRUE)



The dataset was consolidated and locked as a master file (uganda_fake_news_v1.csv) to ensure reproducibility and data integrity throughout the research.

3. Exploratory Data Analysis (EDA)

Initial exploratory analysis confirmed:

- **Balanced class distribution**, eliminating class imbalance bias.
- Consistent text length and structure suitable for NLP processing.
- No critical missing values in labels or text fields after cleaning.

This analysis provided confidence that the dataset was appropriate for machine learning experiments.

4. Data Preprocessing and Feature Extraction

Text preprocessing steps included:

- Lowercasing
- Removal of punctuation and special characters
- Stopword removal
- Tokenization

Feature extraction was performed using **Term Frequency–Inverse Document Frequency (TF-IDF)**, which converts textual data into numerical vectors while preserving the importance of discriminative terms. TF-IDF was selected due to its effectiveness in text classification tasks and its suitability as a baseline approach.

5. Baseline Model Training

A **Logistic Regression** classifier was trained as the baseline machine learning model due to its interpretability, efficiency, and widespread use in text classification research.

The dataset was split into training and testing sets using an **80/20 stratified split** to preserve class balance.

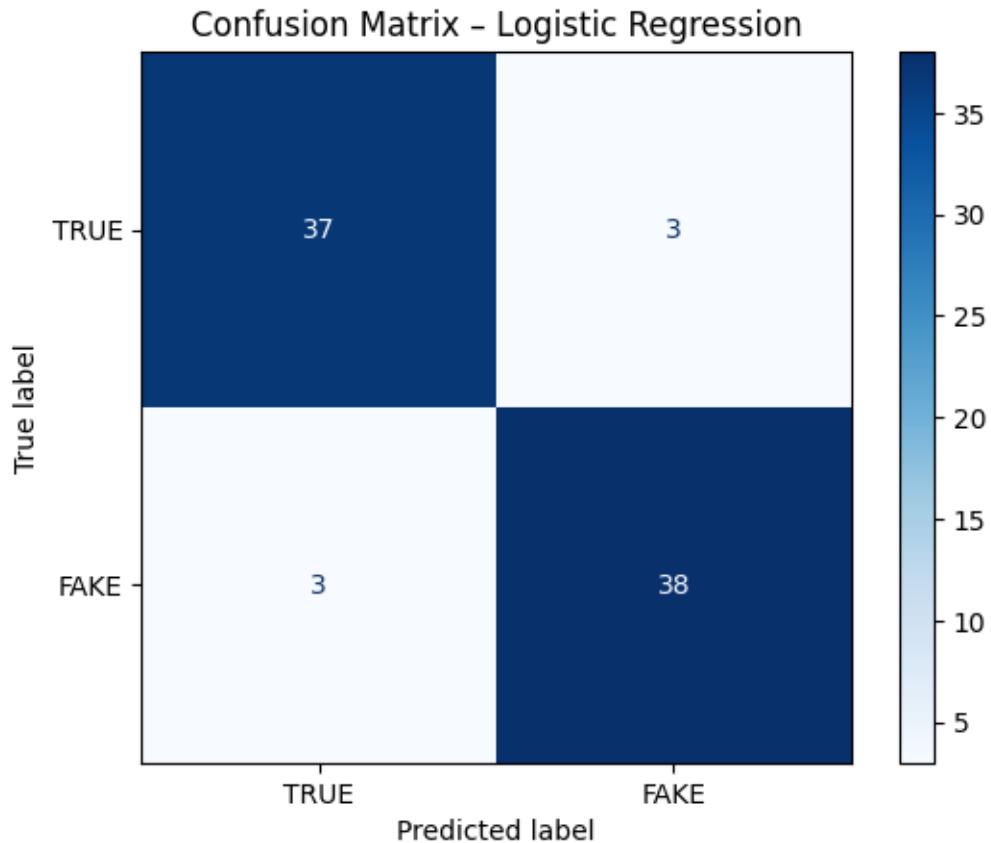
6. Model Evaluation Results

The trained model achieved the following performance metrics on the test dataset:

- **Accuracy:** 92.6%
- **Precision:** 92.7%
- **Recall:** 92.7%
- **F1-score:** 92.7%

The classification report demonstrated balanced performance across both fake and true news classes. The confusion matrix revealed a low number of

misclassifications, indicating strong generalization and minimal bias.



7. Interpretation of Results

The high and balanced evaluation metrics indicate that:

- The collected dataset is **sufficient in size and quality** for fake news detection tasks.
- The model successfully learned meaningful linguistic patterns distinguishing fake from true news.
- The Ugandan context data is suitable for training machine learning models despite linguistic diversity and informal expressions.

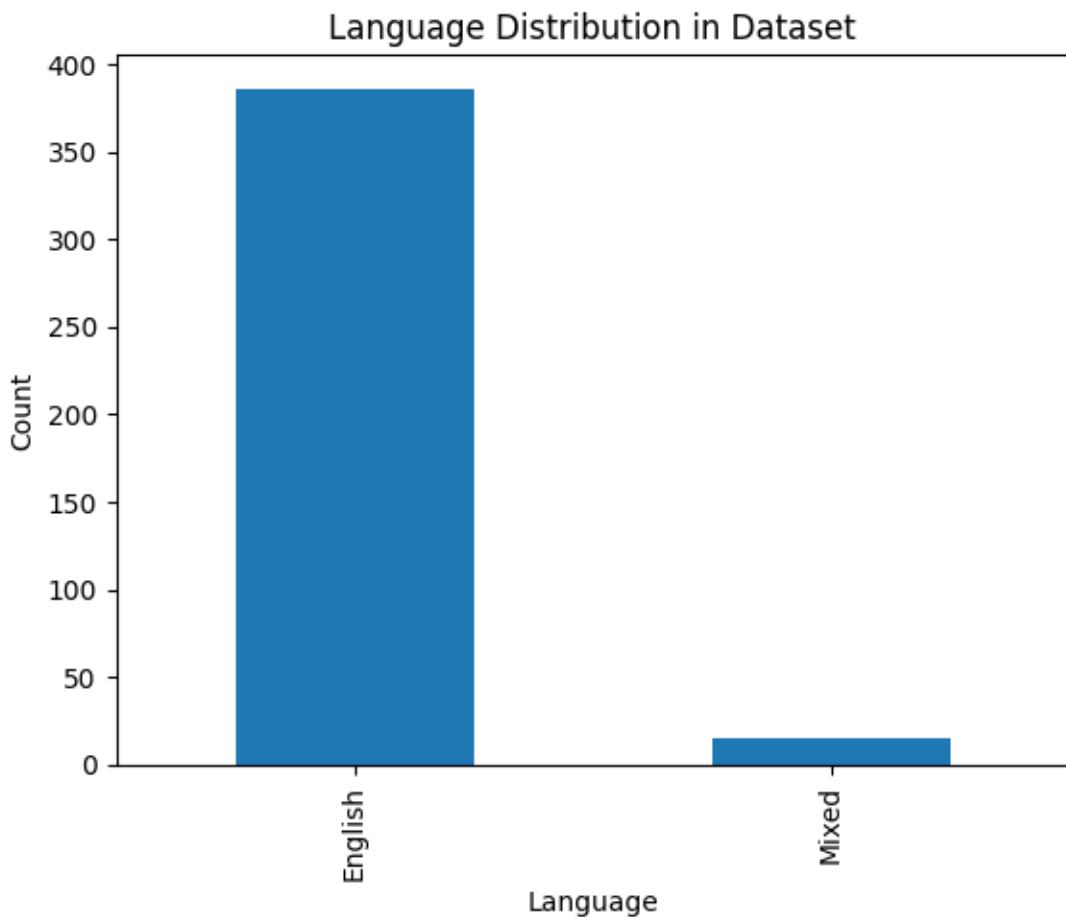
These findings directly address the initial concern regarding data adequacy and confirm the feasibility of continuing with the project.

8. Limitations

- The dataset size, while sufficient for baseline validation, remains moderate.
- The current study focuses on **text-only content**, excluding images, audio, and video.
- Local languages such as Luganda are underrepresented at this stage.

These limitations will be addressed in later phases of the research where feasible.

The dataset includes English and mixed-language content (English with Swahili/local language influence), reflecting real Ugandan online communication patterns. This supports the study's localization objective.



9. Conclusion

This preliminary phase successfully validated data availability and model feasibility. The results demonstrate that a localized AI-based fake news detection system for Uganda is both practical and achievable.

My next step would be:

- Expanding the dataset and language coverage
- Experimenting with advanced models (e.g., SVMs, deep learning, transformers)
- Integrating the trained model into a functional prototype system
- Conducting further evaluation and ethical analysis

