



AN NLP PROJECT

FILTERING REALITY

FAKE NEWS DETECTION THROUGH
TEXT CLASSIFICATION

NEKKY L.

KATHERINE T.

ABOUT THIS PROJECT...

This project focuses on detecting fake news headlines using **Natural Language Processing (NLP)** and **Machine Learning** techniques. The dataset contains news headlines labeled as **0 (fake)** or **1 (real)**. The objective is to build **a classifier** that can automatically distinguish between the two.

PROJECT GOALS

BUILD A FAKE NEWS
CLASSIFIER

Train a machine learning model that can **distinguish** between real and fake news headlines.

PREPROCESS AND
TRANSFORM TEXT DATA

Apply NLP techniques such as **cleaning**, **tokenization**, and **TF-IDF vectorization** to convert raw text into features usable by classifiers.

COMPARE MODELS AND
EVALUATE PERFORMANCE

Experiment with different algorithms (e.g., **Logistic Regression**, **Naive Bayes**, **SVM**) and evaluate them using accuracy and classification metrics.

PREDICT ON UNSEEN
DATA

Use the trained model to assign **correct labels** (0 = fake, 1 = real) to the testing_data.csv file, ensuring the output matches the original format.

ASSESS GENERALIZATION
ABILITY

Estimate how well the model will **perform** on new, unseen data and reflect on **limitations** and potential improvements.

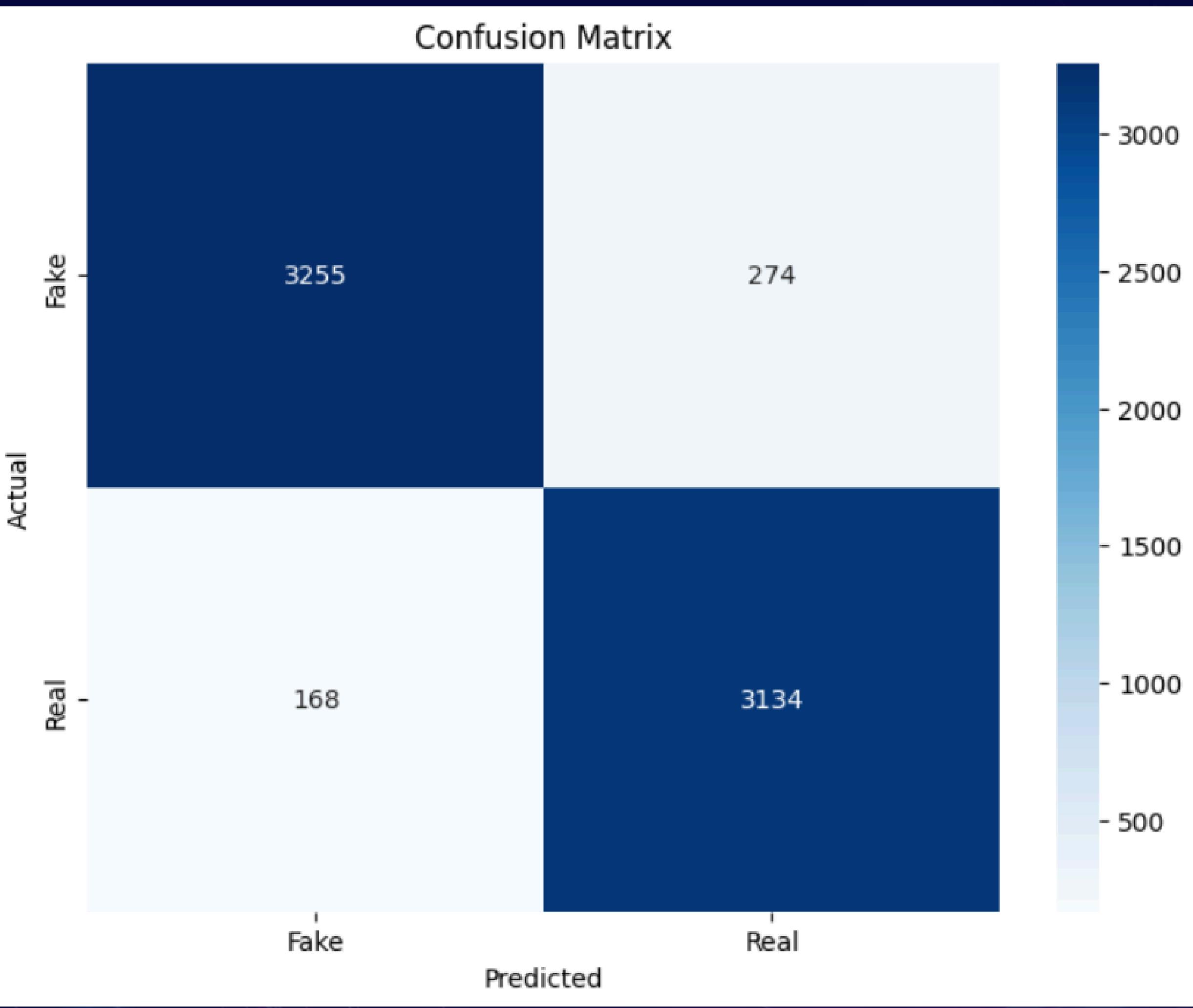
DATA PREPARATION

- Splitting raw lines into **labels** and **headlines**
- Removed punctuation & numbers
- Stemming & Lemmatization

FEATURE EXTRACTION

Converting text into numerical features using TF-IDF vectorization.

Confusion Matrix



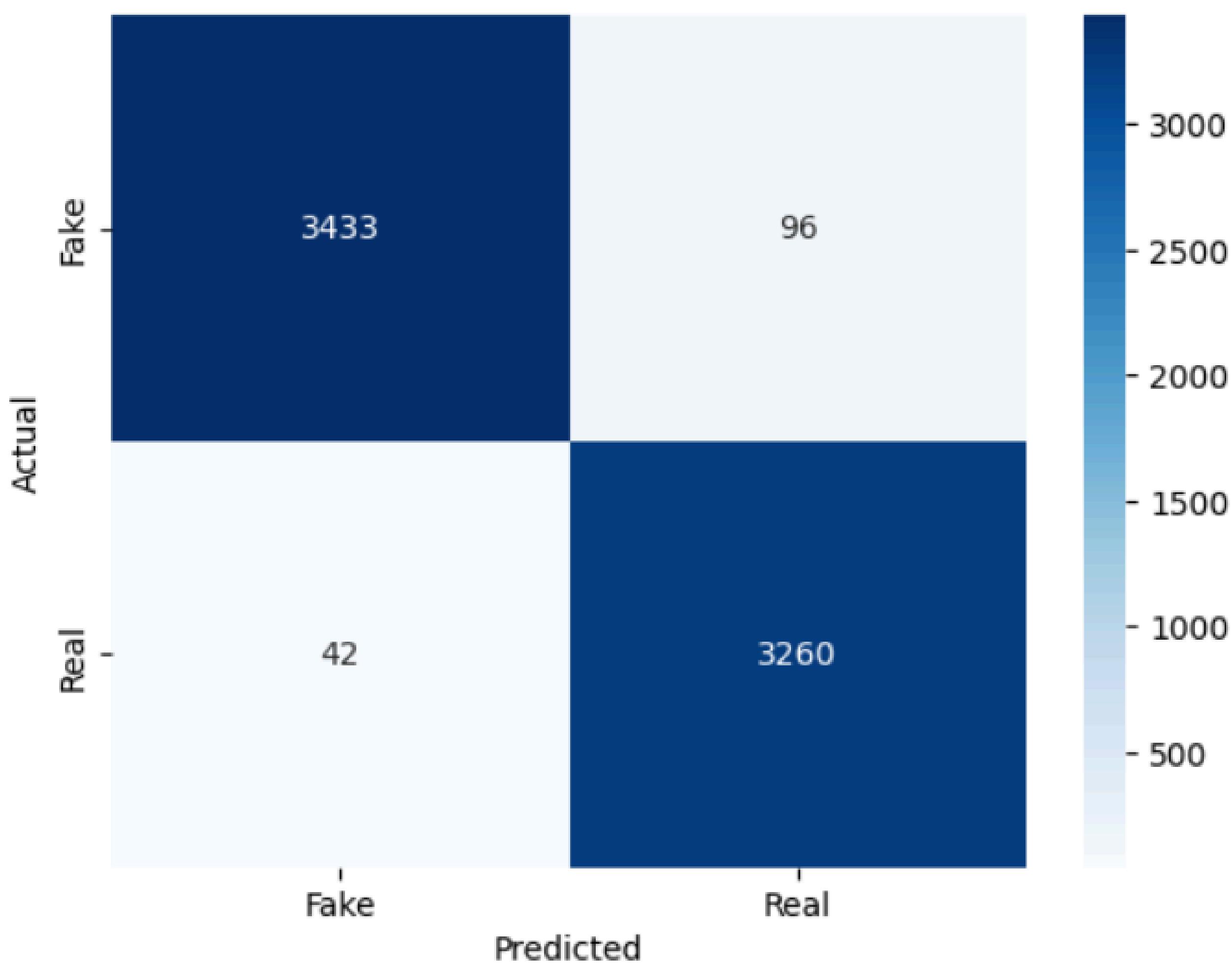
MODEL TRAINING

We experimented with different classifiers (Logistic Regression, Naive Bayes, SVM...) to get the highest accuracy we could.

- Logistic Regression Accuracy: 93.61%
- Naive Bayes Accuracy: 92.63%
- Linear SVM Accuracy: 93.58%

We then fine-tuned a pre-trained transformer on the dataset, using **HuggingFace models**, and tweaking it a bit so it fits our necessities. After training our model, with BERT and an optimizer such as AdamW, we got an **accuracy of 99%**!

Confusion Matrix



PREDICTION

WHAT WE DID

We fed the trained model new headlines, which it analyzed to predict whether they are fake or real. The text was preprocessed just like the training data, so the model could interpret it correctly.

WHY IT MATTERS

These predictions transform raw headlines into actionable insights, helping identify misinformation quickly. By keeping the original dataset format, the results can be easily integrated into reports or further analysis, making the process both practical and reliable.

IN THE FUTURE...

Expand the model to multilingual headlines for global reach.

Integrate real-time news streams for instant fake news detection.

Combine with fact-checking APIs to improve accuracy and trust.

Explore transfer learning from larger NLP models to boost performance.

Include social media signals to detect trending misinformation.

Develop user-friendly dashboards for journalists and readers.

Incorporate explainable AI so predictions are transparent and understandable.

THANK YOU!

ANY QUESTIONS?