

Echo Truth: A machine learning project focused on classifying

news as fake vs real

Kalif Byrd, Sarron Tadesse, Dev Raiyanni



number of rows: 2096
number of columns: 12

Problem Statement:

Wanted to tackle the fake news challenge; it is damaging to public trust

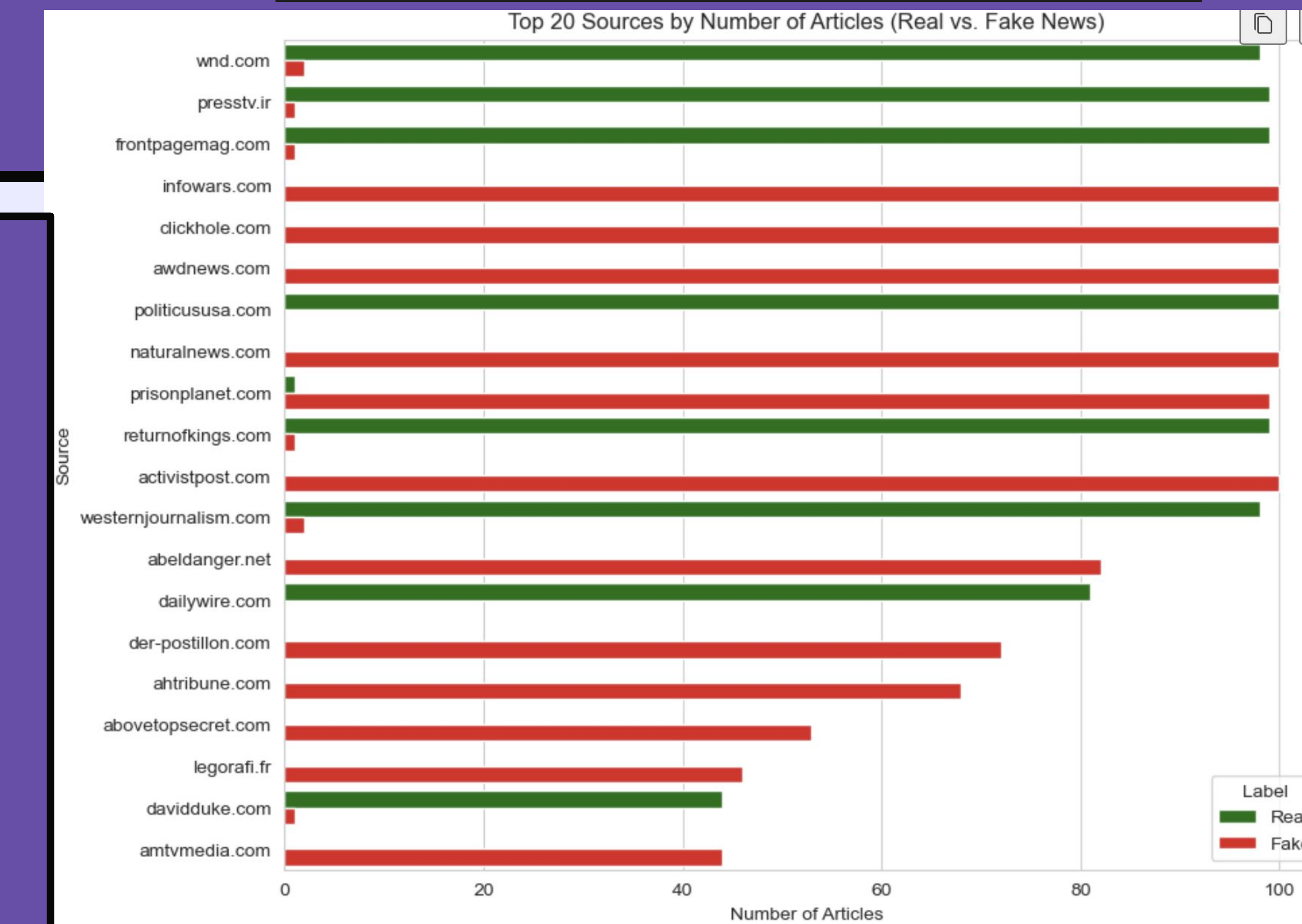
Solution Statement:

Create Echo Truth; a machine learning model that uses natural language processing techniques to classify news as being fake or real.

Technologies / Platforms Used

Dataset from:

- Kaggle
- Model Building:
- Python Libraries
- Model Deployment:
- Heroku



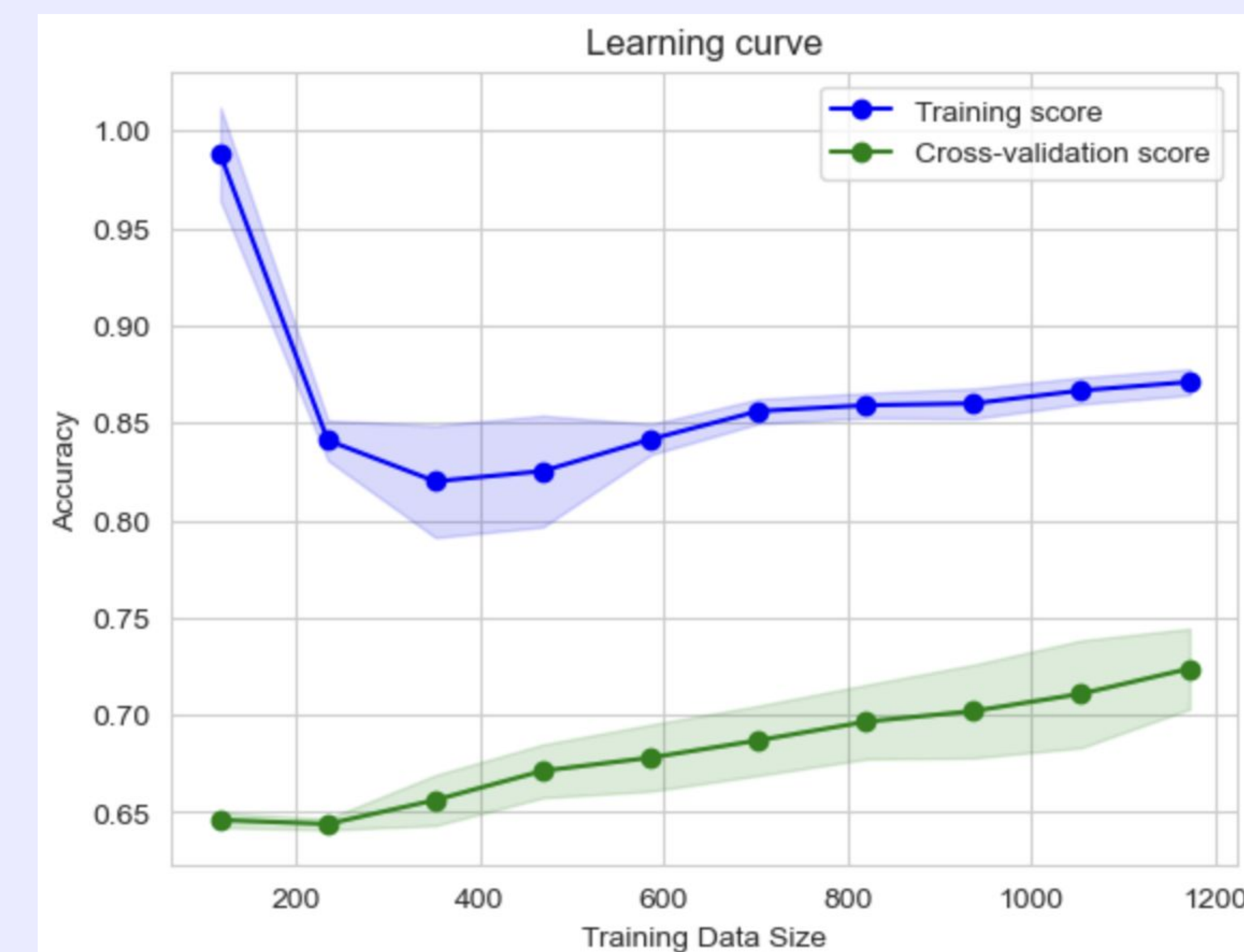
Process of the Project:



HEROKU

kaggle

- 1. Set Up Environment and Libraries:** Imported essential libraries (`pandas`, `numpy`, `matplotlib.pyplot`, `seaborn`, `scipy.stats`) and configured the Jupyter notebook for inline plotting to facilitate data manipulation and visualization.
- 2. Load and Explore Data:** Loaded the `news_articles.csv` into a Data Frame to assess the structure and initial characteristics of the data, exploring the first few entries and checking dimensions.
- 3. Clean and Preprocess Data:** Cleaned the dataset by handling missing values—filling text gaps with "No text available", categorical data with "Unknown", and numeric values with 0.0. Validated changes by rechecking for missing values.
- 4. Model Building and Analysis:** Engineered features suitable for modeling, selected and trained an appropriate machine learning model for classification, and evaluated its performance using standard metrics.
- 5. Visualization and Final Presentation:** Created visualizations using `matplotlib` and `seaborn` to illustrate the model's performance and the insights derived from the data analysis, facilitating an effective presentation of the results.



- We achieved a significant milestone by developing three distinct models, each demonstrating varying levels of accuracy.
- Greatest Model Accuracy was 85%