

Detection of Fake News Using Machine Learning

Matej Koreň

Supervisor: Ing. David Hříbek

Motivation

The issue of fake news spreading is a growing problem in today's society. Machine learning provides a promising solution to this problem by allowing us to detect and flag fake news quickly and accurately. By analyzing vast amounts of data and learning to recognize patterns, these algorithms can help us identify false information and prevent its spread. By using machine learning to combat fake news, we can protect the integrity of information and ultimately create a safer and more informed society.



Implementation

With the use of standard machine learning and NLP libraries for Python (NLTK, pandas, ScikitLearn, Keras & others) a program which can take an online article and evaluates its credibility on pre-trained models was designed and created. Models, which were trained are Naive Bayes (NB), Support Vectors Machine (SVM), Decision Tree (DT) and Sequential Neural Network (SNN).

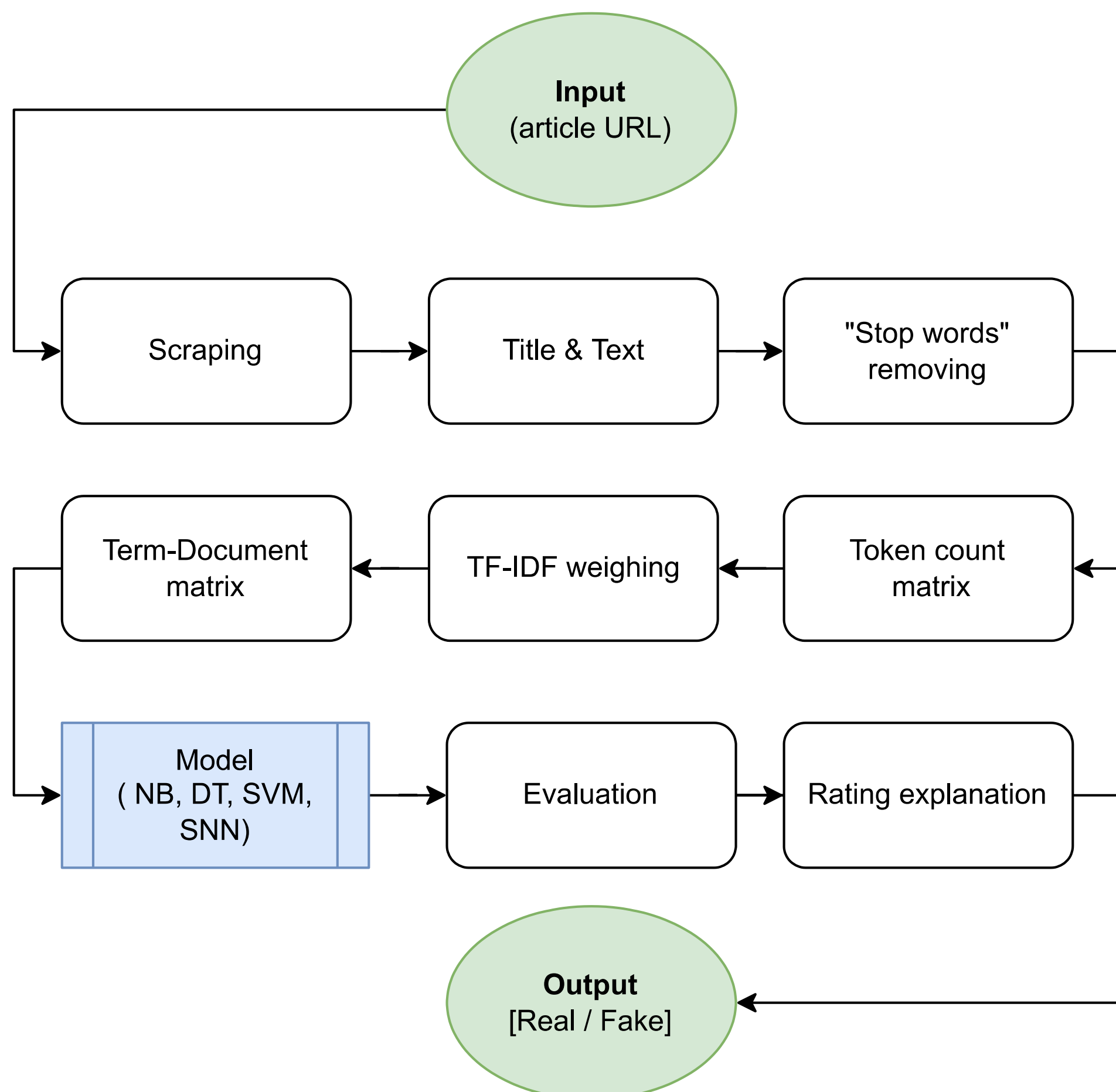


Fig. 1.: Fake News classifier data pipeline

Training & Testing

All of the mentioned models were trained and tested on multiple english datasets, usually reaching accuracy well over **85%**. They were also trained on the **Dezinfo SK** dataset, their evaluations and accuracies are in the table below.

Used model	Precision (Fake/True)	Recall (Fake/True)	F1- Score (Fake/True)	Accuracy
NB	0.88/0.79	0.82/0.85	0.85/0.81	83%
SVM	0.85/0.88	0.85/0.88	0.85/0.88	87%
DT	0.86/0.81	0.80/0.87	0.83/0.84	83%
SNN	0.94/0.85	0.89/0.92	0.91/0.88	90%

Tab. 1.: F1-scoring for each trained model

Article evaluation

With the use of Lime Text Explainer, we can see the weights of each word that the model considers important to overall validity of the article. For example, article from **nytimes.com** titled „Sudan Erupts in Chaos: Who Is Battling for Control and Why It Matters“ classified by SNN as „True“ was evaluated as seen below:

Sudan Erupts in Chaos: Who Is Battling for Control and Why It MattersGunshots erupted outside apartments and rockets screamed across city blocks. Smoke engulfed planes at the airport and shells crashed into a military tower. Two rival Sudanese generals have transformed a city of five million people into an arena for their personal war. The clashes have pitted a paramilitary group known as the Rapid Support Forces against the Sudanese Army, reflecting a longstanding rivalry between Sudan's two top generals who have been vying for dominance. The eruption of violence on Saturday in Sudan's capital and other parts of the country has dashed hopes that civilians could soon take leadership of a democratic government, the goal of mass protests four years ago. In 2019, Sudanese protesters and the military toppled the country's authoritarian leader, President Omar Hassan al-Bashir, offering hope to similar movements in Africa and the Arab world. After Mr. al-Bashir's rule ended, the military signed a power-sharing agreement, but then took over with a coup in 2021. One of Africa's largest countries, where the United States and its allies have tried to aid a transition to civilian control, is now reeling from a new crisis that many fear could become full-blown civil war.

Prediction probabilities

False	0.43
True	0.57

Fig. 2.: Article evaluation

Own Fake News Dataset

Dezinfo SK - Fake News Dataset was created as a part of a bachelor's thesis on this topic. Every article has its unique id, publishing date, title, text, URL of source, URL of proof and manually selected label.

Id	Date	Title	Text	Label
1	5.3.23	Ukrajinci pritvrdili ...	Ukrajinská armáda v posledn...	Fake
2	5.3.23	Bachmut je stále v ...	Situácia v meste Bachmut v ...	True

Tab. 2.: Simplified dataset snippet

To learn more about this dataset and its creation process, visit **Kaggle.com – Dezinfo SK – Fake News Dataset** or scan this QR code.

