# Fake news detection

Cieľ tohto projektu je rozdeliť fake news od naozajstných správ.

Funguje to tak že to spracuje veľké množstvo fake news a real news a podľa toho usúdi či je niečo fake news. Nikdy to nieje 100% percentné pretože to usudzuje podľa predošlích udalostí čo sa už stali.

## Dataset

Dataset je prevzatý z tejto webovej stránky [Link je tu!](https://www.kaggle.com/datasets/bhavikjikadara/fake-news-detection).

Nachádzajú sa tam 2 súbory csv s fake news a real news ktoré sa už staly.

Fake news csv má 23468 riadkov a 101 stĺpcov

Real news csv má 21418 riadkov a 38 stĺpcov

## Postup

Krok 1 - Načítanie dát

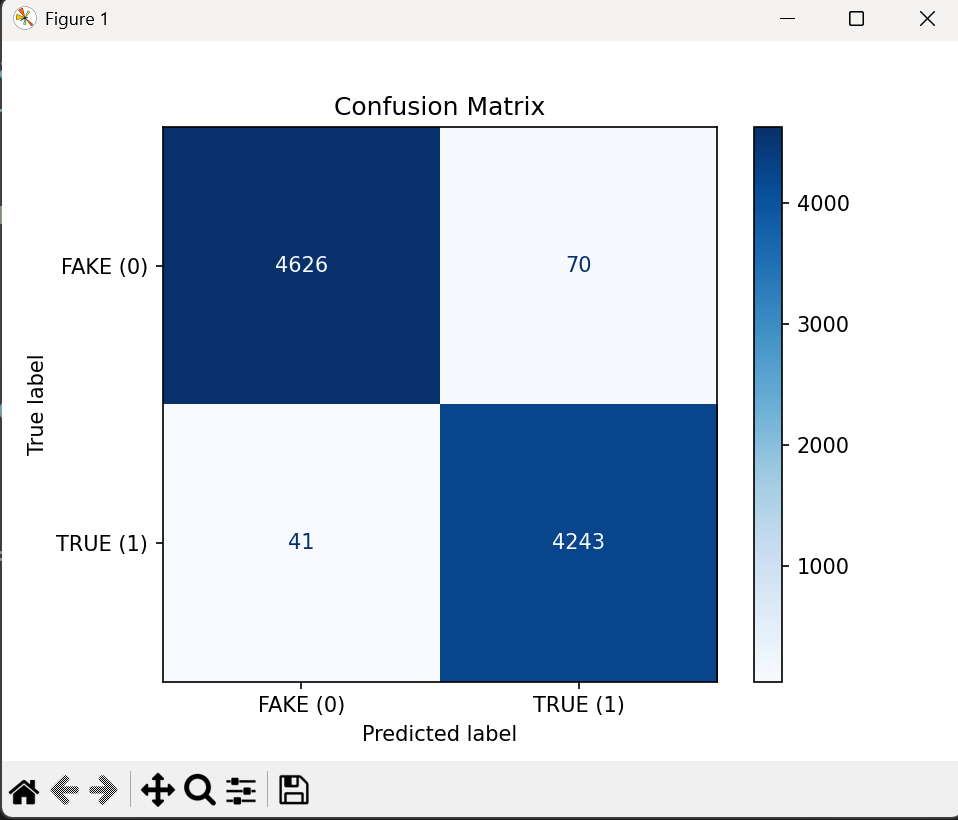
Krok 2 – Príprava textu

Krok 3 – Rozdelenie dát a vektorovanie

Krok 4 – Tréning 1 jednoduchého modelu

Krok 5 – Vyhodnotenie

## Výsledky



Train accuracy: 0.9914805946879002

Test accuracy: 0.9876391982182628

precision recall f1-score support

0 0.99 0.99 0.99 4696

1 0.98 0.99 0.99 4284

accuracy 0.99 8980

macro avg 0.99 0.99 0.99 8980

weighted avg 0.99 0.99 0.99 8980

# Záver

Model nikde nezlyháva, vyšlo to tak že model je na 98% presný aj keď by to malo byť na 99%.

Ako ďalši by som chcel vyskúšať Wake detection model čo znie dosť zaujímavo