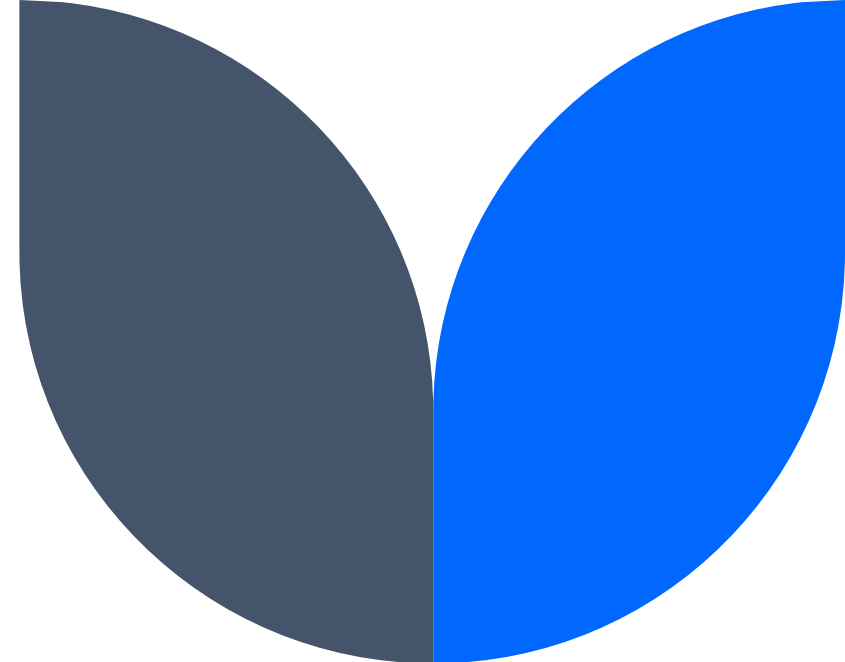
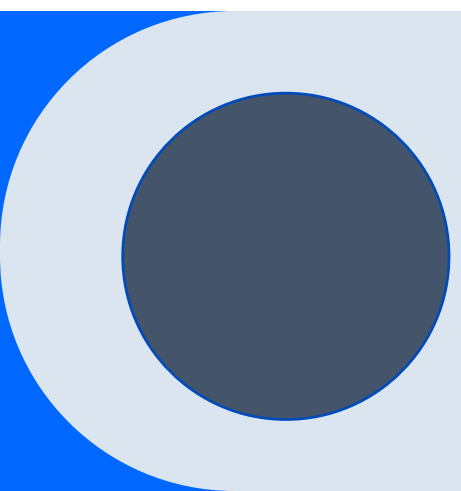




# HL'ADANIE ANOMÁLIÍ V TEXTOCH POMOCOU NEURÓNOVÝCH SIETÍ

Autor: Bc. Olha Koliesnikova

Vedúci práce: doc. RNDr. Gabriela Andrejková, CSc



# Motivácia

- Odhaliť plagiátorstvo
- Zabezpečenie kvality textových údajov - identifikovať chyby, nezrovnalosti alebo odláhlé hodnoty vo veľkých súboroch textových údajov

# Ciele

1. Spracovať prehľad známych výsledkov a použitých metód v oblasti hľadania anomálií v textoch.
2. Vytvoriť metódu (systém algoritmov) pre detekciu anomálií v textoch používajúci neurónovej siete.
3. Navrhnutú metódu aplikovať na textových dátach (vlastných alebo verejne dostupných) a výsledky vyhodnotiť porovnaním s inými metódami.

# 1. Prehľad existujúcich výsledkov a metód v oblasti hľadania anomálií v textoch

- Štatistické metódy
- Prístupy založené na pravidlách
- Strojové učenie
- Neurónové siete
- Článok Z.Szoplák, A.Almarimi, A.Salem, G.Andrejková *"Anomaly detection in texts using sentence embeddings"*

## 2. Vytvoriť metódu (systém algoritmov) pre detekciu anomálií v textoch používajúci neurónovej siete.

- Použitie reprezentácie textu
- Architektúra neurónovej siete
- Trénovanie s neanomálnymi dátami
- Identifikácia anomálií

### 3. Navrhnutú metódu aplikovať na textových dátach a výsledky vyhodnotiť porovnaním s inými metódami.

- Výber textových dát
- Trénovanie a testovanie
- Porovnanie s inými metódami
- Hodnotenie výkonu

**Ďakujem za  
pozornosť**