

Zadání maturitního projektu z informatických předmětů

Jméno a příjmení: *Matouš Volf*

Pro školní rok: *2024/2025*

Třída: *4. A*

Obor: *Informační technologie 18-20-M/01*

Téma práce: *Analýza textů internetových článků a příspěvků pomocí LLM*

Vedoucí práce: *doc. Ing. Jakub Šimko, PhD.*

Způsob zpracování, cíle práce, pokyny k obsahu a rozsahu práce:

Cílem projektu je analyzovat a klasifikovat textový obsah internetových novinových článků, příspěvků na sociálních sítích a jejich diskuzí. K tomuto účelu budou využity modely strojového učení a velké jazykové modely. Analýza bude provedena na existujících datových sadách.

Na textech budou pozorovány a vyhodnocovány znaky:

- téma,
- sentiment,
- vzájemný vztah,
- vzájemná shoda či neshoda.

Výstupem práce bude vědecký článek obsahující výsledky výzkumu a jejich interpretaci.

Stručný časový harmonogram (s daty a konkretizovanými úkoly):

září, říjen:

- nabytí teoretických znalostí potřebných k realizaci
- rešerše existujících prací zabývajících se stejným tématem
- průzkum dostupných nástrojů a modelů

listopad:

- čištění a preprocessing dat
- experimentace s využitím různých modelů a postupů

prosinec, leden:

- analýza samotná

únor, březen:

- interpretace a porovnání výsledků
- vyvození závěrů