

# Motivácia k výberu témy bakalárskej práce

*Klasifikácia web stránok pomocou metód strojového učenia*

Vedúci práce: RNDr. Ľubomír Antoni, PhD.

Konzultant: RNDr. Stanislav Hrivňák, PhD.

Študent: Martin Bača

Vždy ma fascinovali počítače a veci, ktoré sa s nimi dajú dosiahnuť. Doteraz som bol limitovaný svojimi znalosťami, avšak posledné semestre strávené na Prírodovedeckej fakulte UPJŠ ma veľa naučili a myslím si, že som pripravený s pomocou môjho školiteľa a konzultanta začať vytvárať veci, ktoré pre mňa niekedy boli nedosiahnuteľné.

Hneď po prezentácii zástupcu firmy Global Logic som vedel, že téma, ktorú ponúkajú je pre mňa ako stvorená, pretože v sebe kombinuje zameranie môjho štúdia - dátovú analýzu, a prácu s webovými stránkami, ktorá je už dlhšie mojim koníčkom.

V minulosti som sa venoval podobným projektom, či už v rámci predmetu strojového učenia alebo vo voľnom čase, ale doteraz to vždy bolo v menšom rozsahu. Rozsah tejto práce je však jedna z mojich hlavných motivácií, keďže mám rád dlhodobú prácu na projektoch, ktoré ma bavia a naplňajú.

Cieľom práce je vytvoriť prehľad metód strojového učenia používaných na klasifikáciu webových stránok, scrapovaním dôležitých údajov zo stránok získať vhodnú množinu dát, na ktorej následne budem implementovať rôzne metódy strojového učenia, a na záver porovnať dosiahnuté výsledky s inými dostupnými štúdiami.

Napriek tomu, že podobné štúdie už boli vykonané, táto bude unikátna vzhľadom na to, že konkrétne kategórie klasifikácie vznikli na základe špecifických potrieb, čo so sebou prinesie nové výzvy a úskalia, ktoré bude treba prekonať.