

I. IDENTIFIKAČNÍ ÚDAJE

| | |
|----------------------------|---|
| Název práce: | Predikce neúspěšného dokončení povinných předmětů a vyhodnocení finanční náročnosti výuky |
| Jméno autora: | Adam Johanides |
| Typ práce: | diplomová |
| Fakulta/ústav: | Fakulta elektrotechnická (FEL) |
| Katedra/ústav: | Katedra počítačů |
| Oponent práce: | Ing. Denis Baručič |
| Pracoviště oponenta práce: | Katedra kybernetiky |

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

| Zadání | náročnější |
|---|------------|
| Zadání se skládá ze tří částí: (1) Seznámení se s rozsáhlou metodikou Kometa, jež má být použita pro výpočet finanční náročnosti výuky. (2) Využití strojové učení pro predikci úspěšnosti studia a počtu studentů v povinných předmětech. (3) Navržení a implementace informačního systému, jež dokáže získávat potřebná data, zpracovávat je a zobrazovat. Každá z těchto částí je dle mého názoru výzvou, a proto hodnotím zadání jako <i>náročnější</i> . | |

| Splnění zadání | splněno s menšími výhradami |
|---|-----------------------------|
| Student analyzoval metodiku Kometa a pomocí lineárního programování formuloval problém nalezení minimální finanční náročnosti. Pro predikci úspěšnosti studia byl použit lineární klasifikátor naučený z dat ze studijního systému KOS. Zadaný informační systém byl navržen a implementován. Klasifikátor bohužel existuje pouze ve formě Jupyter notebooků, a tudíž není přímo integrovaný do systému, což znemožňuje kompletní automatizaci. | |

| Zvolený postup řešení | vynikající |
|--|------------|
| Práce začíná rešerší existujících přístupů. Následuje analýza a formulace problémů a shrnutí cílů. V další části student představuje relevantní data a statisticky je zpracovává. Později jsou představena řešení jednotlivých částí, jejichž výsledky jsou nakonec shrnuty a vyhodnoceny. Tento postup se mi jeví jako systematický a vhodný. | |

| Odborná úroveň | A - výborně |
|--|-------------|
| Práce využívá postupy softwarového inženýrství a základní metody optimalizace nebo strojového učení (logistická regrese). Student pracuje s poměrně rozsáhlou sadou dat (ve smyslu dimensionalit i počtu vzorků), kterou statisticky zpracovává. Napočítané statistiky vizualizuje a správně interpretuje. Odevzdaný zdrojový kód je přehledný a dodržuje doporučené praktiky. | |

| Formální a jazyková úroveň, rozsah práce | D - uspokojivě |
|---|----------------|
| Práce je napsána v anglickém jazyce s drobnými gramatickými chybami, které však nebrání porozumění textu. Na několika místech text zasahuje do okrajů stran, ale jinak je z hlediska typografie kvalitní. Objevují se malé nedokonalosti, např. Figure 4.5 následuje po Figure 4.6, odkazování v sekci 3.2 na „dříve popsaná data v sekci 4.2“ či chybné značení definic (např. definice 3.1.2 není definice). Z hlediska obsahu jsem text vnímal jen velmi obtížně a hodnotím jej jako nejslabší část práce. Sekce 3 představuje značení používaných atributů, které je velmi nepřehledné: pro značení jsou použity zkratky, které ale referují k českým názvům, přičemž v textu jsou názvy anglicky; obvykle mají tyto zkratky index, který někdy konkretizuje atribut (např. P_{par} pro počet paralelek), někdy určuje příslušnost atributu k entitě (např. ZH_k pro započtené hodiny katedry k) a jindy se tyto indexy kombinují. Subjektivně jsem měl problém sledovat souvislosti v textu, např. sekce 2.1 se poměrně intenzivně věnuje problému batohu a jak jej využít pro určení finanční náročnosti, což zmiňuje ještě sekce 3.1.2, ale v navržené ILP formulaci v sekci 5.1.2 už souvislost není zřejmá. Obecně je rozsah práce dostatečný, ale informačnímu systému se text věnuje jen stručně, a tak je čtenář odkázán převážně na příložený zdrojový kód. | |

Výběr zdrojů, korektnost citací

A - výborně

Student čerpá z odborných učebnic či článků a několika internetových zdrojů. Domnívám se, že výběr zdrojů odpovídá charakteru práce. Citace jsou řádně vyznačeny.

Další komentáře a hodnocení

Výsledky predikce nejsou příliš přesvědčivé, ale naznačují, že vztah mezi historickými daty o studentech a budoucími výsledky studia existuje, a tak by mohlo být zajímavé v tomto směru navázat budoucí práci.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Student prokázal schopnost analyzovat komplexní problém a navrhnout vhodné řešení, ale zvolená notace a struktura textu degraduje kvalitu předložené práce.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **C - dobře**.

Otázky k obhajobě:

1. Co by integrace klasifikátoru přímo do informačního systému znamenala? Šlo by použít i jiný klasifikátor než logistickou regresi?
2. Domníváte se, že navýšení počtu dat pro učení klasifikátoru by významně zvýšilo jeho výkon? Jaké další faktory by mohly přispět?

Datum: 15.1.2021

Podpis: