

Posudek vedoucího na bakalářskou práci Lenky Kovářové

Prostředí pro analýzu veteránského rankingu OB

Zařazení této práce je poněkud složitější. Pohybuje se na pomezí informatiky a statistiky. Jejím hlavním cílem bylo navrhnout statistické metody pro zpracování existujících dat a na základě tohoto zpracování vybrat nejvodnější výpočetní model rankingu. Softwarová část spočívala především v analýze a návrhu datového modelu tak, aby pro vlastní provádění potřebných statistických výpočtů už nebylo zapotřebí vytvářet žádné komplikované aplikační nadstavby.

Výsledný návrh datového modelu je popsán v kapitole 4. Další zpracování prokázalo, že tento model byl navržen vyhovujícím způsobem, jeho popis v textové části práce by mohl být poněkud podrobnější.

Datový model byl implementován a bylo nutné jej naplnit daty. Sběr dat nebyl úplně jednoduchý, mimo jiné i pro nutnost konverze z mnoha různých formátů vyskytujících se v několika odlišných variantách.

Po naplnění dat se těžiště práce přesunulo do statistické části. Studentka musela nejprve prozkoumat různé výpočetní modely, které se pro výpočty rankingu používají, a současně navrhla i některé vlastní přístupy. V této fázi došlo k doplnění databáze o atributy potřebné pro flexibilní práci s různými modifikacemi zkoumaných vzorečků.

Poslední částí teoretické přípravy bylo nalezení vhodných metod pro ověřování „správnosti“ jednotlivých vzorečků, resp. toho, jak daný vzoreček věrně mapuje existující realitu. Studentka přitom odvodila modifikaci Spearmanova korelačního koeficientu, která se pro daný účel ukázala jako nejvhodnější.

Následovala závěrečná fáze - zpracování statistických dat, selekce nejvhodnějších vzorečků a hledání jejich ideální parametrizace. Tuto část zpracovala studentka velmi pečlivě, možná zda ovšem práce poněkud postrádá „tah na branku“. Prospěla by jí rychlejší selekce vhodnějších a méně vhodných modelů vzorečků a doladování jemnější parametrizace už pouze na malé skupině vzorečků. To by umožnilo na počátku vycházet z ještě pestřejšího spektra možných modelů.

Pro prezentaci výsledků rankingu bylo vytvořeno jednoduché rozhraní přístupující pomocí WWW k datům v databázi. Toto rozhraní slouží v této fázi vlastně pouze pro účely statistického výzkumu, takže nemá potřebnou grafickou úroveň. Nicméně jeho funkčnost je postačující, prokazuje vhodnost použitého propojení s databází a v případě reálného provozu bude jednoduché doplnit je i v tomto směru.

Textová část práce má solidní úroveň, pouze by potřebovala místy poněkud formálnější tón.

Přes uvedené nedostatky by podle mého názoru tato práce měla být uznána za bakalářskou.

V Praze, dne 31.8.2009



RNDr. Libor Forst
SISAL MFF UK