

# Posudek oponenta bakalářské práce

Autor/Autorka

Jan Ondříček

Název práce

Vybrané prediktivní modely pro výsledky zápasů NBA

Studijní obor

Matematika a finanční studia

Oponent práce

Ing. Patrice Marek, Ph.D.

## Splnění cílů práce:

nadstandardně  velmi dobře  splněny  s výhradami  nebyly splněny

## Odborný přínos práce:

nové výsledky  netradiční postupy  zpracování výsledků z různých zdrojů  shrnutí výsledků z různých zdrojů  bez přínosu

## Matematická (odborná) úroveň:

vynikající  velmi dobrá  průměrná  podprůměrná  nevyhovující

## Věcné chyby:

téměř žádné  rozsahu přiměřený počet  méně podstatné, větší množství  podstatnější, větší množství  závažné

## Grafická, jazyková a formální úroveň:

vynikající  velmi dobrá  průměrná  podprůměrná  nevyhovující

## Slovní hodnocení

Autor se ve své práci zabývá rankingovými modely pro předpovídání vítěze v americké nejvyšší basketbalové lize NBA. Pro zpracování si zajistil reálná data od sezony 2002/03 do sezony 2016/17 a provedl kontrolu a základní statistické zpracování. V teoretické části se věnuje popisu třech rankingových modelů – konkrétně jde o Keenerovu metodu, Elo rating a PageRank –, které následně aplikuje na reálná data a optimalizuje některé jejich parametry. Uvedené modely a programy pro srovnání byly vytvořeny v prostředí programu MATLAB. Na závěr jsou modely porovnány jak mezi sebou, tak v některých částech i proti sázkové kanceláři.

Práce je napsána přehledně, neobsahuje téměř žádné věcné chyby a obdržené výsledky lze považovat za nové. Grafickou, jazykovou a formální úroveň rovněž považuji za vynikající.

## Konkrétní chyby v práci a případné poznámky:

- Obecně platí, že popisky v obrázcích by bylo vhodné zvětšit (např. v obrázku 5.3 na str. 31 jsou indexy již nečitelné).
- Odkaz na tabulku 2.1a má být uveden jako 2.1 (str. 4).
- Ve druhém odstavci na str. 9 nemá být čárka za spojením „rating každého týmu“.
- V kapitole 5.1.1 by bylo vhodné připomenout, na které způsoby výpočtu je aplikován výraz (3.5) a na které nikoliv.
- Na konci str. 11 a 35 nemá být osamocený řádek.
- Obrázek 3.3 na str. 15 by měl být zpracován lépe (z grafického pohledu).
- Ve vzorci (3.20) na str. 17 je použito značení obvyklé jednotkovou maticí, ne pro matici samých jedniček.
- V poznámce pod čarou na str. 24 má být „konečné“, místo „končné“.

- V kapitole 5 (str. 28) je stanoveno, že modely budou porovnávány na základě průměru. Tato volba by měla být blíže rozebrána, tj. zda se jedná o vhodné kritérium. V této souvislosti je zajímavý obrázek 5.3 (str. 31), kde je díky použití průměru vybrána jako nejlepší metoda „score“, ale při pohledu na minimální hodnoty by nevhodnější byla metoda „0/1“. Zde by tedy bylo vhodné prozkoumat alespoň mediány metod.
- S výše uvedeným souvisí i to, že by bylo zajímavé srovnat úspěšnost modelů např. v první a druhé polovině sezony, tedy zda model potřebuje na naučení nějakou sadu zápasů. Dle výstupů je navíc podobné srovnání obsaženo v MATLABovském kódu.
- V části 5.1.3 je uvedeno, že na základě testování byla zvolena hodnota parametru  $d = 0.75$ , ale nejsou zde uvedeny žádné výsledky, které toto tvrzení podporují.
- V první větě posledního odstavce na str. 40 má být slovo „existenci“, místo „existence“.
- V poslední odrážce na str. 45 nemá být čárka před slovem „při“.
- V poslední větě na str. 47 chybí slovo „vektor“.

Práci navrhuji hodnotit stupněm **výborně**.

#### Otázky k obhajobě

1. Jak bylo řešeno stěhování (prodej) týmů do jiného města v případě prvního odhadu sezony?
2. V kapitole 5.2.2 uvádíte srovnání úspěšnosti predikce v závislosti na vzdálenosti ratingů. Bylo by zde nějakým způsobem možno zakomponovat i podobné orientační srovnání se sázkovou kanceláří? Stejně tak i v kapitole 5.2.3.



Patrice Marek, 4. 6. 2018