ČVUT

ČESKÉ VYSOKÉ
UČENÍ TECHNICKÉ
V PRAZE

POSUDEK VEDOUCÍHO
ZÁVĚREČNÉ PRÁCE

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Predikce výsledků ve hře Counter-Strike s pomocí strojového učení
Jméno autora:	Ondřej Švec
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Vedoucí práce:	Gustav Šír, Ph.D.
Pracoviště vedoucího práce:	Katedra počítačů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání   **průměrně náročné**

Hodnocení náročnosti zadání závěrečné práce.

Jde o velmi standardni format zadani problemu z oblasti strojoveho uceni.

Splnění zadání   **splněno**

Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.

Student splnil vse co bylo zadano.

Aktivita a samostatnost při zpracování práce   **A - výborně**

Posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven. Posuďte schopnost studenta samostatné tvůrčí práce.

Student byl od zacatku velmi samostatny, prestoze pro nej byly mnohe koncepty z teto domeny zcela nove. Pravidelne reportoval a konzultoval, a sam aktivne prichazel s resenimi nastalych problemu.

Odborná úroveň   **B - velmi dobře**

Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.

Prace neprinasi z hlediska technicke odbornosti nic obzvlast noveho, v kontextu dane e-sport domeny jde však o docela unikatni pocin, kdy student sesbiral doposud nejvetsi a nejdetailnejsi datovou reprezentaci, a jeho analyza a vysledky tez predci predchozi odborne prace v teto domene (CS:GO).

Formální a jazyková úroveň, rozsah práce   **D - uspokojivě**

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Toto je nejslabsi cast - student se kvuli necekanym technickym problemum zasekl v experimentalni casti a podcenil casovou narocnost sepsani samotne zaverecne prace. Struktura textu tak pusobi trochu uspechanym a ne zcela dotazenym dojmem, napr. popis experimentu, ucicich reprezentaci a vysledku mohl byt mnohem detailnejsi. Nicmene, student psal kvalitni anglictinou a vsechny zasadni kroky jsou v textu uvedene.

Výběr zdrojů, korektnost citací   **B - velmi dobře**

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

V praci je pomerne vice odkazu na ruzne online materialy, namisto odbornych clanku, nez je zvykem, nicmene je to pochopitelne vzhledem k resenemu problemu, ktery je velmi cerstvy a teprve nabira na popularite v odbornych kruzich.

Další komentáře a hodnocení
Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.
Viz celkove hodnoceni.

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE
Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení.

Prace mela za cil pouzit relativne bezne techniky strojoveho uceni v relativne netradicni e-sport domene kompetitivni online hry CS:GO. Student nebyl s technikami strojoveho uceni dopredu moc obeznamen, proto obzvlast ocenuji jeho vytrvalost při samostudiu a pokrocilost nekterych vyslednych neuralnich modelu. Vysledky jsou slusne, za hranice stavajicich praci, prestoze nektere jednodussi modely se ukazaly nad ocekavani dobre.
Slabou strankou je textovy vystup, ktery student casove podcenil. Silnou strankou je pak predevsim vytvoreny dataset, ktery muze poslouzit vznikajici odborne komunite kolem teto nove domeny.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm B - velmi dobře.

Datum: 12.1.2022   Podpis:

I. IDENTIFICATION DATA

Thesis title:	**Predicting Counter-Strike Game Outcomes with Machine Learning**
Author's name:	**Ondřej Švec**
Type of thesis :	bachelor
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Cybernetics
Thesis reviewer:	Vojtěch Jindra, MSc
Reviewer's department:	Department of Computer Science

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment	challenging

How demanding was the assigned project?

The main goal of the assignment was to perform a predictive analysis of Counter Strike: Global Offensive matches. Achieving this includes a rather broad set of tasks such as data scraping, data analysis, experimenting with different models and optimization of their hyper-parameters. Hence, I consider the project to be rather time-consuming.

Fulfilment of assignment	fulfilled

How well does the thesis fulfil the assigned task? Have the primary goals been achieved? Which assigned tasks have been incompletely covered, and which parts of the thesis are overextended? Justify your answer.

The assignment has been fulfilled, although, as the author mentions in section 7.1 Future improvements, it would be interesting to see how the models would perform with different rating algorithms than Elo, especially given that publications have shown that for example TrueSkill consistently manages to outperform Elo.

Methodology	correct

Comment on the correctness of the approach and/or the solution methods.

Generally speaking, the models were chosen reasonably, and their application was indeed correct. However, some parts of the chosen methodology could be done more thoroughly:

- Section 2.4 Player ranking – Elo talks about extending the Elo rating algorithm to rating teams instead of individual players. This is achieved by perceiving a team as an individual, and the team's rating is calculated by averaging the ratings of its players. For updating players' ratings, the change in the team's rating is simply propagated to each of its players. This feels like an ad-hoc solution and its viability is never shown. For example, the author could have considered propagating the ratings with weights calculated based on given player's rating relative to the team's rating, so if given team lost, players with the highest ratings in the team would suffer higher loss of their rating and vice versa. The consequence of such approach would be that players' ratings of given team would converge towards the team's rating, which intuitively makes sense.
- On another note, in section 2.4 Player ranking – Elo, the equation for calculating one's rating shows a "magic" constant of 400 and its presence is never explained by the author. When inventing the Elo algorithm, the constant was selected by Arpad Elo because with this specific number, the formula best approximated the distribution of the chess data. Since the project is not dealing with chess data, it would be interesting to see what the optimal value for this constant would be.
- Section 2.5.3 Loss function delves into how crucial role a loss function plays in ML models. Then, the author proceeds to introduce binary cross-entropy loss with reasoning that it is very popular for binary classification tasks. While I agree, I would argue that several loss functions should have been explored, especially given the rather weak argument for binary cross-entropy.

Technical level	A - excellent.

Is the thesis technically sound? How well did the student employ expertise in the field of his/her field of study? Does the student explain clearly what he/she has done?

My previous comments on methodology were mostly about what could have been on top of what was done, rather than what could have been done differently. Hence, I find it important to say that what was done was, in my eyes, employed very well and I did not find any potential improvements.

Formal and language level, scope of thesis	A - excellent.

Are formalisms and notations used properly? Is the thesis organized in a logical way? Is the thesis sufficiently extensive? Is the thesis well-presented? Is the language clear and understandable? Is the English satisfactory?

The thesis was very well structured, and I found it very easy to follow. The formal language was well above average with an occasional typo, but that happens even to the best of us.

Selection of sources, citation correctness	A - excellent.

Does the thesis make adequate reference to earlier work on the topic? Was the selection of sources adequate? Is the student's original work clearly distinguished from earlier work in the field? Do the bibliographic citations meet the standards?

Citations are done correctly, and the choice of the sources seems adequate. The comprehensive list of references shows how much work has been done to study the topic rigorously before making a contribution.

Additional commentary and evaluation (optional)

Comment on the overall quality of the thesis, its novelty and its impact on the field, its strengths and weaknesses, the utility of the solution that is presented, the theoretical/formal level, the student's skillfulness, etc.

I believe that the work is overall very solid. The author took a lot of time to comprehend the current state of the art, obtained a dataset large enough for his work, applied several ML models and correctly evaluated the results.

I would appreciate if the thesis was more result oriented. For example, it is my feeling that the hyper-parameter optimization was not done very thoroughly, yet it is one of the factors that distinguish a bad model from a good model. Moreover, the selected baseline was a simple random decision weighted by the distribution of the data, hence beating the baseline was not a very competitive task.

III. OVERALL EVALUATION, QUESTIONS FOR THE PRESENTATION AND DEFENSE OF THE THESIS, SUGGESTED GRADE

The thesis is very comprehensive, generally well structured and the author's ability to express his thoughts concisely makes it easy to follow. The author showed that he put a lot of effort into researching and understanding the field. The process of obtaining the necessary data was thought-through as well as the analysis that followed. Applying different models on the dataset could have used a little bit more work; namely hyper-parameter optimization and testing out different loss functions when applicable.

For the reasons mentioned above, I suggest marking this thesis with B (very good), however I find it important to mention that the work is certainly closer to an A than to a C.

Questions for the defense:
- As mentioned in methodology, the extension of the Elo rating algorithm for teams was, in my eyes, rather straightforward. Can you reason why did you choose to do it this way? Have you considered other approaches? Do you think the performance of your algorithms would change had you chosen a different way to handle Elo for teams?

- Does the constant of 400 in the formula for calculating Elo rating fit your data? What steps would you take to find the optimal value for this constant?
- What other loss functions could be used for a binary classification task? Did you consider any other?
- In section 4.1.1 Matches, you state that it is believed, yet not proven, that more matches result in more realistic results (i.e., the underdog has a lower chance of winning). Cannot we simply refer to the law of large numbers to show this? If no, could you explain why, and if yes, can you give an example of how would you apply the law of large numbers to show this?
- In section 4.1.1 Matches, you mention that BO2 are usually used for league tournaments. Out of curiosity more than anything else, do you know why is that? What is it that makes BO2 superior for leagues, but inferior for other formats?
- In section 5.1.5 Data processing, you explain the normalization process of your training and validation datasets. What parameters do you use for normalizing the test dataset?
- In section 6.3 Test set evaluation, you state that for Random Forests, the accuracy for the training dataset is 99.8% and only 63.0% for the test dataset. That is a rather huge difference. Did you further explore why is that?

The grade that I award for the thesis is **B - very good.**

Date: **19.1.2022**   Signature: