



Review report of a final thesis

Reviewer: RNDr. Vladimíra Sečkárová, Ph.D.
Student: Roman Isaev
Thesis title: Statistical and sentiment-based prediction of UFC results
Branch / specialization: Knowledge Engineering
Created on: 10 June 2023

Evaluation criteria

1. Fulfillment of the assignment

- ▶ [1] assignment fulfilled
- [2] assignment fulfilled with minor objections
- [3] assignment fulfilled with major objections
- [4] assignment not fulfilled

The final thesis follows the objectives of the assignment; objectives are formulated correctly and fulfilled sufficiently.

2. Main written part

95 /100 (A)

The aim of this final thesis is to reliable prediction of future Ultimate Fight Championship (UFC) results. Chapter 1, Introduction, contains motivation and aims of this thesis. In Chapter 2, descriptions of two machine learning (ML) methods used for classification are given. Chapter 3 focuses on sentiment; related definitions and overview of several sentiment classification methods and currently used technologies are given. Chapter 4 gives overview of how different authors dealt with UFC forecasting by using ML. Chapter 5 lists possible sources for semantic-based approach and gives details of semantic data acquisition. Chapter 6 firstly gives overview of (summary) statistics describing various aspects of the fights. Then, fight data acquisition, data pre-processing and feature description is described. In Chapter 7, after details of several evaluation metrics and model selection, the proposed method for UFC forecasting (a combination of ML particular ML method and semantic-approach) is revealed. This novel method is applied to small real dataset and brief commentary on the results is given. Final Chapter contains conclusion and suggestion for future work.

Below are comments and suggestions:

C1: Breaking longer portions of text would improve readability (applicable to p. 21 for example).

C2: Both, British and American English are used in the thesis - choose one. Example: neighbour and neighbor.

C3: Missing figures on p. 29?

C4: Overview of kNN method missing - other 'baseline' models used in Section 7.2.1. were described in preceding chapters.

C5: Section 7.3.2: winner evaluation metrics - please include references showing their use in other fields/works. The first one is (without the percentage) a commonly used proportion of successes for example.

3. Non-written part, attachments

100 /100 (A)

Python files, provided together with written part of the thesis, contain no errors and all outputs are visible.

Comments:

C1: A single .ipynb file containing the whole code with sections following the order of the topics in the thesis would be appreciated.

4. Evaluation of results, publication outputs and awards

98 /100 (A)

A novel method for forecasting UFC results is proposed and application to real data shows promising results (Chapter 7). Greater discussion regarding its performance could have been included. The proposed algorithm could have large impact in a range of applied fields such as Data Science, Applied Statistics, etc.

The overall evaluation

96 /100 (A)

The presented final thesis, written in English, is well-structured and has number of typos as expected for the length of the thesis. I really enjoyed reading it. A greater discussion regarding the performance of proposed method could have been included. I recommend this thesis for the defense.

Práci doporučuje k obhajobě.

Questions for the defense

Q1: Section 7.2.2: Methods perform well when original label is 1 (positive). But both confusion matrices have high counts of false positives (label 1 when true is 0); in case of CatBoost less than 1/3 was identified as true negative (label 0 when true is 0). And thus, there seems to be bias towards label 1 present - can you comment on why this behaviour occurs?

Q2: Section 7.3.3: One would assume that the accuracies for metric 'Positive tweet counts' (fighter with higher total number of positive tweets is predicted as winner) and metric 'Negative tweets counts' (fighter with lower total number of negative tweets is predicted as winner) are similar. Surprisingly, the accuracy for 'positive' is much lower - why do you think this occurs?

Instructions

Fulfillment of the assignment

Assess whether the submitted FT defines the objectives sufficiently and in line with the assignment; whether the objectives are formulated correctly and fulfilled sufficiently. In the comment, specify the points of the assignment that have not been met, assess the severity, impact, and, if appropriate, also the cause of the deficiencies. If the assignment differs substantially from the standards for the FT or if the student has developed the FT beyond the assignment, describe the way it got reflected on the quality of the assignment's fulfilment and the way it affected your final evaluation.

Main written part

Evaluate whether the extent of the FT is adequate to its content and scope: are all the parts of the FT contentful and necessary? Next, consider whether the submitted FT is actually correct – are there factual errors or inaccuracies?

Evaluate the logical structure of the FT, the thematic flow between chapters and whether the text is comprehensible to the reader. Assess whether the formal notations in the FT are used correctly. Assess the typographic and language aspects of the FT, follow the Dean's Directive No. 52/2021, Art. 3.

Evaluate whether the relevant sources are properly used, quoted and cited. Verify that all quotes are properly distinguished from the results achieved in the FT, thus, that the citation ethics has not been violated and that the citations are complete and in accordance with citation practices and standards. Finally, evaluate whether the software and other copyrighted works have been used in accordance with their license terms.

Non-written part, attachments

Depending on the nature of the FT, comment on the non-written part of the thesis. For example: SW work – the overall quality of the program. Is the technology used (from the development to deployment) suitable and adequate? HW – functional sample. Evaluate the technology and tools used. Research and experimental work – repeatability of the experiment.

Evaluation of results, publication outputs and awards

Depending on the nature of the thesis, estimate whether the thesis results could be deployed in practice; alternatively, evaluate whether the results of the FT extend the already published/known results or whether they bring in completely new findings.

The overall evaluation

Summarize which of the aspects of the FT affected your grading process the most. The overall grade does not need to be an arithmetic mean (or other value) calculated from the evaluation in the previous criteria. Generally, a well-fulfilled assignment is assessed by grade A.