

Report on Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague

Student:	Bc. Jakub Hostačný
Advisor:	RNDr. Matúš Baniar
Title of the thesis:	Non-linear classification as a tool for predicting tennis matches

OVERALL ASSESSMENT (provided in English, Czech, or Slovak):

Please provide your assessment of each of the following four categories, summary and suggested questions for the discussion. The minimum length of the report is 300 words.

Contribution

As summarized by the author, there are four main contributions of the thesis. First, utilized methods have not been frequently used in this type of analysis. Second, the list of used explanatory variables is much larger than in similar studies. Third, the covered period of 16 years is much larger than in other studies. And fourth, the thesis compares different types of betting strategies. Overall, the contribution of the thesis is significant and I can easily see a polished version of the thesis being published in a scientific journal.

Methods

Used methods are partly within the scope of the Data Science with R course taught at IES but most of them go well beyond. All methods are succinctly described and well presented not wasting much of the space.

Literature

The literature review is detailed and well structured. I personally prefer shorter literature reviews but in this case, I understand the point of having a larger one as the topic is not so well-known to the standard economics community.

Manuscript form

This is the point where I see probably the only weakness of the thesis. There are few typos and some equations and their descriptions are not quite as nice as they could have been. But mainly, the figures (5.1+) are quite ugly. That is only a minor issue, though.

Summary and suggested questions for the discussion during the defense

Shortly, this is an excellent thesis – straightforward motivation, well-structured, descriptive when needed, applicable results. As a lecturer of the Data Science course, I can say that this text can be easily presented as a template for a very well written data science thesis. Plus, as I have noted before, I believe this can be easily published in a nice scientific journal. I gladly suggest the thesis for the defense with the A grade. For the defense, I have only one question:

- Why only men's matches? Data availability or something else?

Report on Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague

Student:	Bc. Jakub Hostačný
Advisor:	RNDr. Matúš Baniar
Title of the thesis:	Non-linear classification as a tool for predicting tennis matches

SUMMARY OF POINTS AWARDED (for details, see below):

CATEGORY	POINTS
<i>Contribution</i> (max. 30 points)	30
<i>Methods</i> (max. 30 points)	30
<i>Literature</i> (max. 20 points)	20
<i>Manuscript Form</i> (max. 20 points)	18
TOTAL POINTS (max. 100 points)	98
GRADE (A – B – C – D – E – F)	A

NAME OF THE REFEREE:

doc. PhDr. Ladislav Krištofek, Ph.D.

DATE OF EVALUATION:

24.1.2018

Referee Signature