

# Report on Bachelor Thesis

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

<b>Student:</b>	<b>Tomáš Turlík</b>
<b>Advisor:</b>	<b>doc. PhDr. Ladislav Krištofuk, Ph.D.</b>
<b>Title of the thesis:</b>	<b>Neural networks and tree-based credit scoring models</b>

## **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**

The main result of the thesis is that it does not seem that the more complex methods can beat the standard logistic regression, at least for the two datasets. Even though I find it a little bit confusing in the Results section that the in-sample (training) results are not presented as well to see whether one or more of the methods overfit.

### **Methods**

The utilized methods are advanced ones, mostly of the master's level rather than the bachelor's level (as the thesis mostly follow what has been learnt during Data Science with R course). As these – neural networks and tree-based methods – are usually not known to a mainstream economics/finance reader, they described in detail, which makes perfect sense. The methodological section is well-written, the measures are described as well. No issues here.

### **Literature**

The most important literature is covered. As the data mining methods are pretty much starting the mainstream finance, it is always easy to find appropriate references. Parts of the methodology section could be a bit better referenced, and figures should have been referenced as well.

### **Manuscript form**

The text has a logical structure and it reads well in general. There grammatical mistakes here and there but these are not disruptive. What I do not like much is the subsections structuring which is not ideal, e.g. there is Section 3.1.1 but no Section 3.1.2 or further so that there is no need for this level of sectioning. In the similar vein, there are too many subsections – we are down to the third level and in the third level, there is a hidden fourth maybe even fifth level that is not numbered. This all disturbs the reader. In addition, the figures are not appropriately referenced (I believe they are not of the author).

### **Summary and suggested questions for the discussion during the defense**

Question for the defense:

- Bit of a philosophical one – how do you think the fact that institutions use scoring models affects the efficiency of such models for future use?

# Report on Bachelor Thesis

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

<b>Student:</b>	<b>Tomáš Turlík</b>
<b>Advisor:</b>	<b>doc. PhDr. Ladislav Krištofuk, Ph.D.</b>
<b>Title of the thesis:</b>	<b>Neural networks and tree-based credit scoring models</b>

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

<b>CATEGORY</b>	<b>POINTS</b>
<i>Contribution</i> (max. 30 points)	25
<i>Methods</i> (max. 30 points)	28
<i>Literature</i> (max. 20 points)	18
<i>Manuscript Form</i> (max. 20 points)	13
<b>TOTAL POINTS</b> (max. 100 points)	<b>84</b>
<b>GRADE</b> (A – B – C – D – E – F)	<b>B</b>

**NAME OF THE REFEREE:** doc. PhDr. Ladislav Krištofuk, Ph.D.

**DATE OF EVALUATION:** 28.8.2018

---

**Referee Signature**