

## Review of Bachelor's Thesis

**Student:** Uchytíl Albert  
**Title:** Recognition of Audio Events Using Deep Neural Networks (id 18850)  
**Reviewer:** Černocký Jan, doc. Dr. Ing., UPGM FIT VUT

- 1. Assignment complexity** **average assignment**  
The assignment is very generally defined and the difficulty corresponds to the choice of work the student and supervisor finally agreed on. In my opinion, the targeted classification of phonemes in TIMIT is moderately difficult.
- 2. Completeness of assignment requirements** **acceptable under reservation**  
The assignment was fulfilled, but what was finally done is a classification. For recognition, the implemented estimator of phoneme posteriors would need to be complemented with an appropriate model (most probably an HMM).
- 3. Length of technical report** **within minimum requirements**  
The report contains a chapter on the theoretical bases of classification and recognition (that could be shortened), but the description of the actual experiments is not sufficient, the report should contain detailed information about data, setups, parameters of feature extraction, metrics, etc.
- 4. Presentation level of technical report** **60 p. (D)**  
See the previous comment. The logical structure of the thesis is not very clear and titles of chapters are somewhat confusing.
- 5. Formal aspects of technical report** **60 p. (D)**  
The thesis is written in English - its level could be improved, but I appreciate that the student used English and tried to formulate his ideas himself, not by copying from sources. The math notation needs to be improved by consistently using appropriate notation for matrices and vectors. The figures need to be improved as well (axes, scaling of values and merging related figures into one).
- 6. Literature usage** **75 p. (C)**  
The list contains relevant literature and the student made best efforts to study bases of machine learning, however, from the structure of the thesis it is way too evident, that the quality of sources could be improved and that much MORE READING should be done. I recommend using good papers and PhD thesis rather than summaries randomly found on the internet...
- 7. Implementation results** **60 p. (D)**  
The output is a set of Python scripts for TIMIT phoneme classification. I appreciate that the student used a low-level toolkit and tried to understand the fundamentals of NN rather than adopting a "black-box" approach. The work would be usable pedagogically, but the report does not provide much information on what is where, how to use different functions, etc.
- 8. Utilizability of results**  
Itself, this BP is not much usable but it can provide a nice basis for further work in machine learning.
- 9. Questions for defence**
  1. Precise, how the learning rate was set during the training.
  2. Explain the different metrics in the text output on page 24 of the report.
- 10. Total assessment** **70 p. good (C)**  
The work aims at phoneme classification by DNNs. A set of scripts based on Theano was produced and a range of experiments was conducted on TIMIT. The report is however lagging behind the work done and would need substantial improvements both at content and formal levels.

In Brno 2. June 2016

.....  
signature