THESIS REVIEWER'S REPORT

I. IDENTIFICATION DATA

Thesis title:	**Classification of Intrapartum Fetal Heart Rate Signals**
Author's name:	**Mohamed Safwat**
Type of thesis :	bachelor
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Dept. of Cybernetics
Thesis reviewer:	Ing. Eduard Bakštein, Ph.D.
Reviewer's department:	Dept. of Cybernetics

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment	**ordinarily challenging**

How demanding was the assigned project?

The assignment describes a standard, well-defined machine-learning-oriented project on an extensive dataset.

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 student fulfilled the assignment appropriately, the primary aims have been achieved.

Methodology	**correct**

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

The student used several standard machine learning algorithms, including the logistic regression, SVM and KNN. Although scarcely described, I found the approach used for feature evaluation and cross-validation above average.

Technical level	**B - very good.**

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?

The student studied a machine-learning problem, applied appropriate methods for data preprocessing and evaluation and achieved reasonable results. The soundness of the depth of technical knowledge achieved by the student is at many points hard to assess, as the level of technical detail is highly variable throughout the thesis and some sections and concepts are described in a rather informal and vague way (e.g. the random forest and SVM classifiers, different approaches to crossvalidation). However, I very much appreciate the reconsideration of the preprocessing criteria, presented by the "hypotheses" in the results section, providing a nice discussion to the results.

Formal and language level, scope of thesis	**E - sufficient.**

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 is written in reasonable English with a rather high number of typos and grammar errors. I found the structure of the text quite misleading at several points, e.g. the very beginning of the thesis, starting with general ML introduction, or the Feature importance section introducing the random forest abruptly and without context or any other notice later in the text, or several concepts being referred to before being defined.

As for the formal aspects, I found other considerable shortcomings: The outline does not correspond to the actual content of the thesis (e.g. the "Introduction" section comes, strangely, after "Machine learning", which does not contain "Data" subsection, suggested by the outline). Also, the numbering is not nested (subheadings are numbered, while headings not), which, combined with the confusing outline, makes orientation in the text difficult.

The description of the dataset on p10 is unclear: it is not defined what is the meaning of a positive/negative sample (is it good/bad labor outcome? How is that measured/assessed?). If there are 2125 positive samples in the dataset and 86 are from the "first stage" and 65 from the "second stage", what is the rest and why can it not be used for the task? I could not

understand the description of the FIGO-based features, due to some grammar errors. Some abbreviations (e.g. the AUC) are used before they are defined.

Strangely, the author name on the title page does not match the name in the assignment (maybe a third name?)

Selection of sources, citation correctness	**C - good.**

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?

The thesis contains relatively few citations – additional citations would be helpful at many points in the thesis. However, the citations sources are well selected and clearly identify the referenced sources.

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.

As the introduction section stresses the interpretability of ML algorithms being crucial in medicine, I missed any kind of such evaluation in the results and discussion sections

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

Summarize your opinion on the thesis and explain your final grading. Pose questions that should be answered during the presentation and defense of the student's work.

I found the problem in the assessment to be appropriately addressed by the student, presenting also some nice additional thoughts for model evaluation. However, I found the level of the written text part quite unsatisfactory – hence the grade.

Q1: Did you calculate the feature values yourself from the raw recordings according to the methods defined in Georgoulas et. Al 2017 , or did the database already contain the feature values?

Q2: As stated in the database description section, the database used contains recordings based on two different physical principles: i) direct scalp ECG electrodes and ii) sonography-based recordings. Are the signals in these two groups really so similar that no specific preprocessing or different feature-sets are necessary? Have you verified whether using different treatment for the two data subgroups would bring any benefit in the classification or feature performance?

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

Date: **20.8.2020**   Signature: