Bachelor thesis opponent's review

CZECH TECHNICAL UNIVERSITY IN PRAGUE
Faculty of electrical engineering
Department of electrical power engineering
Technická 2, 166 27 Prague 6, Czech Republic

Master thesis: **Classification of sleep stages from polysomnography and actigraphy**

Author: **Hanseo Yoo**

Thesis supervisor: **Ing. Eduard Bakštein, Ph.D.**

Thesis opponent: **Ing. Vlastimil Koudelka, Ph.D.**

	Rating (1 – 5) (1 = best; 5 = worst):
1. Fulfillment of assignment requirements:	3
2. Systematic solutions of individual tasks:	2
3. Ability to apply knowledge and to use literature:	2
4. Thesis formal and language level:	3
5. Thesis readability and structuring:	3
6. Thesis professional level:	3
7. Conclusions and their formulation:	4
8. Final mark evaluation (A, B, C, D, E, F):	D
verbal:	**satisfactory**

Brief summary evaluation of the thesis (compulsory):

The project addresses an important and interesting topic, which has significant implications for advancing the field of automatic sleep scoring with minimal effort for both technicians and patients. While the overall structure of the text is logical, I have several comments regarding formal language, readability, and technical aspects. Some terms are defined after their first use (e.g., Theta z, Actigraphy mean), which can confuse readers. Additionally, certain terms (e.g., k-fold cross-validation, random tree approach to regularization) are not adequately explained. There are some typographical mistakes, such as the incorrect title "Error range" on page 35. The axes on figures lack units.

While ECG is correctly discussed as a significant predictor, Figure 14 suggests that the "time since sleep start" feature contributes most significantly to within-subject accuracy. This observation is not mentioned but represents an important and promising finding for potential applications.

None of the models performed sufficiently in the patient-based validation. Therefore, I disagree with the assertion in the discussion that the models demonstrated “a certain level of predictive performance.” The performance level was rather unsatisfactory.
In summary, while the project tackles a meaningful and interesting topic, the technical aspects of the thesis are relatively weak. However, the structure, conceptual framework, selected methods, and language are solid. I recommend that the thesis be accepted for defense.

Questions:

1. What is the key principle behind why a random forest prevents overfitting, despite utilizing more parameters than a single tree?
2. What is a possible scenario in which a model that performs satisfactorily well only on sample-based validation would be appropriate for application?

Date: 21.1.2025 Signature: