

Opponent Review

for the master thesis

"Comparative Study of Fuzzy Logic, Artificial Neural Network, and Neuro-Fuzzy System in Medical Diagnosis – An Approach towards a Medical Expert System"

Reviewer: Prof. Dr. Phillipp Torkler
Student: Harvey Ngoe Kolle
Matriculation No: 22103050

Summary

The major motivation for the thesis is the permanent need, but limited access to trained physicians in various areas around the globe for precise medical diagnosis. To address the need, the thesis suggests to aid medical diagnosis by AI systems and compares fuzzy logic, artificial neural networks (ANNs), and neuro-fuzzy-systems - a combination of fuzzy logic and ANNs - for medical diagnosis. Fuzzy logic and fuzzy-systems are analyzed in this thesis due to the hypothesis that fuzzy-systems provide a better model interpretability compared to other machine learning approaches. Public available data sets for diabetes, heart disease and hepatitis are used for comparison of the different AI systems.

Review

Mr. Kolle's thesis is clearly structured and written in a precise language. Research questions and goals are clearly stated and motivated. The arguments for the need for explainable AI in certain applications like medical diagnosis are convincing.

The "Methods" section puts much emphasis on various membership functions of the used fuzzy system, but lacks for some important information for all used AI approaches. For instance, no details for the design of membership functions are given. It is not evident, how and why thresholds for the membership functions were derived. The design of the

used ANN (e.g. how many hidden layers, how many neurons per layer etc.) is missing which complicates a thorough comparison between used approaches.

The "Results" section reports accuracy, sensitivity and specificity for all three AI systems for each of the three diseases. Other common metrics are not reported. No confidence intervals or standard deviations are reported for any reported result. Reported results for the neuro-fuzzy system seem to be based on training data and thus, would overstate the reported performance. For instance, while performance values for the ANN are correctly reported for a test set of the data, the performances for the neuro-fuzzy system are reported on total data sets where 80% of that data set was previously used to train the neuro-fuzzy system. Commonly applied best practices like cross validation aren't used.

Results are discussed clearly, but suffer from remarks mentioned above. The hypothesis of an improved explainability could have been discussed in more detail. Obviously, a rule-based system consisting of only a couple of rules is interpretable. The neuro-fuzzy system derived for the heart disease data set consists of 221 rules. It remains an open question, how such a large system would improve interpretability.

Overall, I rate this as an average master thesis.

Deggendorf, 08.02.2023

Prof. Dr. Phillipp Torkler