
1 Feedback

You have a compelling introduction that tells the reader your motivation for the report, after which you transition well to the specifics of the report. You have described and explained the euroSCORE II model very well, you also described the data profile and measurement variables well not making it too long while still conveying the important aspects. It seems like you have given some real thought of different ways of answering your research questions.

It seems like the project is very vast. There are so many attributes, a large and messy data set and a lot of things to predict and classify. It seems from the report that you are uncertain about what you have the time (and knowledge, as it is a complex subject with a lot to read up on) to do. For example, you write that you want to predict mortal and non-mortal complications post-op in a time span of 30-60 days, but the research questions that you propose would not reach this goal, since they do not mention anything about a time span. This also ties into how the reasoning behind model selections are not so thorough, as it seems like the model choices have not been finalized.

Even though you have described the euroSCORE II model well it is repeated a few times too many and some theory like machine learning and logistic regression does not need to be explained as your peers (AI students) already know it. Furthermore, in your data section there is no mention of imperfections in the dataset, but in the methods section it is stated that there are numerous, this should also be explained in the data section. You also tend to say "it is known" quite a lot without having sources that back this up, you are probably right with what you are saying, but it is hard to believe you when it is not backed up by a source. To be a bit nit picky you shouldn't write yourselves in the acknowledgements section.

Aside from the introduction to the research questions, the research questions themselves have a very feasible scope and make a lot of sense time-wise and in terms of actually ending up with a model that might be better than the logistic regression / euroSCORE II baseline. The report is also very well written so far. It seems like a very interesting and exciting project, and we very much look forward to seeing the results. Especially if they are positive, as the real world impact is very exciting.

Questions for meeting

What more, if any, could be written in State of the Art than the euroSCORE II and the other article, and what considerations have you made about what to keep in State of the Art?

Does the data section fully encompass all aspects of the data?

What parts of the research section do you think are feasible given the time limit?

How will you determine which models are worth investigating and/or using?

How will you make sure to convey that your model should or shouldn't be used as a medical device? Have you considered the ethics and the regulations regarding this?