Peer review summary

There are often unexpected errors or exceptions in textual medical data that cannot be written in advance. In order for the code to run smoothly, sometimes we had to handle it by using the try-except method. In one of the reviews, it was reasonably noted that cases that fit under the exception can be printed and reviewed in more detail. However, we are not planning to add implementation of this idea as it is proposed by the reviewer. The reason is as follows: if there are many such cases, the output may be too large and unreadable, thus is not user-friendly. Therefore, developing this idea, we are planning to create a log file in which the exceptions will be saved.

We have solved a similar problem of typos in doctor's reports with the help of regular expressions and rules that fight such misprints. Due to the fact that we had to list many different spellings of words, the code came out a bit messy. The reviewers objectively pointed this out to us and suggested refactoring the code. We will implement this suggestion in the near future. Also advice to look for a spelling correction library is useful. Such a library for sure may help correct general mistakes and misspellings in the clinicians reports, but are there libraries adapted for medical terms especially in Russian is an interesting question that requires more detailed search.

By the time of the second review, the part devoted to the generation of ECG features and its visualisation was not completely ready. Meanwhile, text data is sensitive and cannot be made publicly available. Therefore, the reviewers were not able to fully test our project. At the moment, all parts of the project are at the final stage and an open ECG dataset is attached for testing feature generation pipeline and comparing with our results.