I launched demonstration.ipynb and reproduced the test example following the documentation provided by the team.

Compared to the previous review the whole process of launching the demo became even more convenient and is well documented. The provided functionality allows for easy usage of the user's own voice records and generated text. Also, the notebook includes the code for training the model from scratch using the LibriSpeech dataset.

The results produced by the model are of the same quality as during the previous review: the text is clearly distinguishable, although with a little bit of noise.

Although the work done by the team shows impressive results I doubt that the project is suitable for real-world problems. First and foremost is the quality of the generated audio: it contains noise and is certainly not on the level of quality of today's modern speech generation systems. Secondly, the configuration of the training pipeline is lacking, i.e. it is not possible to introduce new methods, datasets, in a declarative style. Finally, there is no project management framework (such as Kedro) that would’ve shaped both the project structure and the code to a particular style.

Still, the aforementioned cons are basically non-existent if the project were to be used as a starting point in research in the area of speech generation: the project is nicely documented, contains a pipeline for launching the training that can be easily extendable (e.g. new datasets, methods, etc). The only possible shortcoming of the project is the lack of metrics other than human perception for evaluating the quality of the generated speech.