Final Project Evaluation Guidelines

The final project aims to consolidate all the learnings from this course and apply them to a real problem end-to-end. The choice of problem is open. Any relevant problem that can be MLOps-ified can be considered as the final project. The evaluation of the final project is based on the coverage of learnings deployed into the process and the viva-voce performance of the students. The breakup of points is as below:

1. Having a data (pre)processing pipeline implemented via Apache Spark. The preprocessing can include data acquisition automation, data cleansing, data munging, data transformation and vectorization, but not limited to. [10 points]
2. If the pipeline orchestration is also incorporated via Apache Airflow [8 points]
3. Creation of a structured project documentation in Github with notes of how to use the project source to reproduce the results. [7 points]
4. Push all the source code, models, metrics, reports, figures into the structured project in Github + LFS with release/development version numbers [10 points]
5. Setting up an experimentation tracking page in MLFlow, where the parameters, arguments, metrics, data files and other necessary components are tracked. [10 points]
6. Creation of REST API endpoints for the model building part of your project [10 points]
7. Add Prometheus instrumentation to the REST API functions to capture metrics [7 points]
8. Create a Grafana dashboard to display and visualize the analytics outcomes of the captures metrics [8 points]
9. Dockerize the instrumented REST API for easy deployments with resource limitations and port mapping [10 points]
10. **Viva-voce:**
    a) Ability to explain the components [5 points]
    b) Ability to explain the results [5 points]
    c) Ability to list limitations and potential ways to address the limitations [5 points]
    d) Ability to defend the design/architecture of the project [5 points]

Best Wishes.