User Manual

Order Imprefection Prediction

Roya Sadeghimehr

July 13, 2023

Introduction:

This user manual serves as a comprehensive guide for utilizing the Celonis environment and our designed tool to estimate the likelihood of order imperfection. It provides step-by-step instructions, ensuring users can effectively navigate through the Celonis platform and perform the necessary tasks.

The manual begins with instructions on accessing the Celonis environment and selecting the Machine Learning option. The manual continues by explaining how to extract data from Celonis pools using a specific notebook and how to perform exploratory data analysis using another designated notebook.

The manual then guides users through the modeling process and provides instructions on how to utilize the notebook effectively.

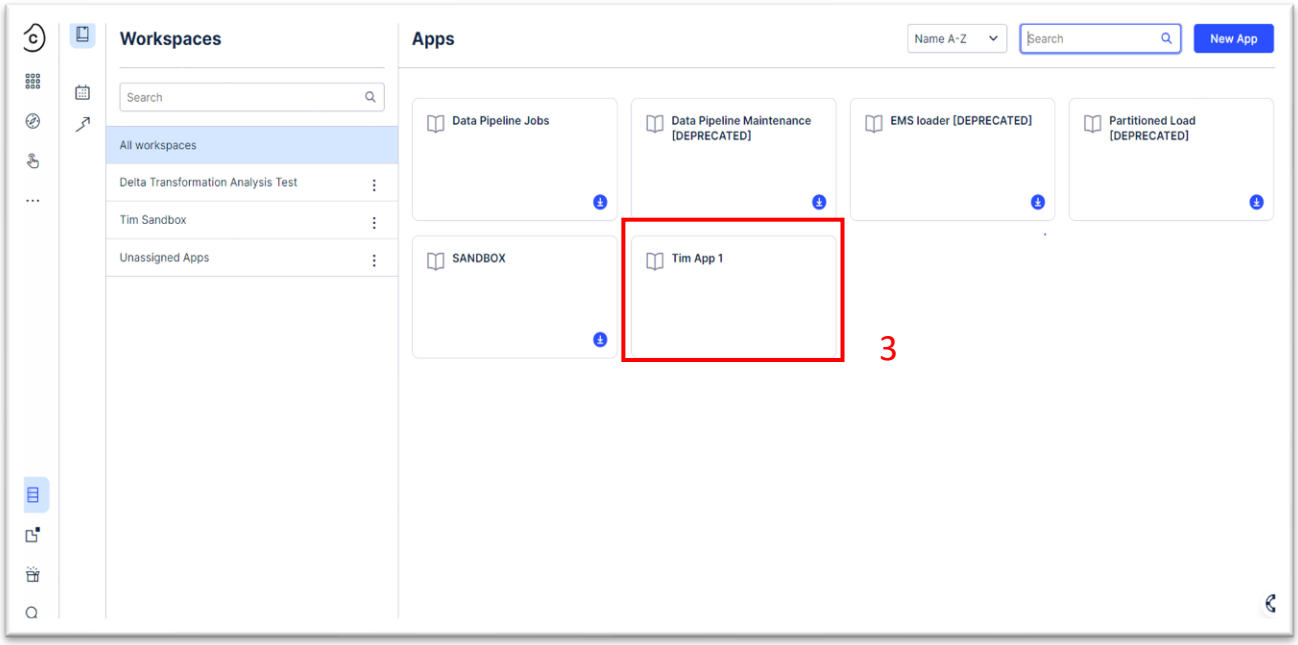
Finally, the manual instructs users on running the tool itself. It also emphasizes the importance of having the necessary dependencies available for smooth operation within the Celonis environment.

Throughout the manual, it is recommended to follow the provided sequence of steps to ensure a cohesive and successful user experience. By following these instructions, users will be able to utilize the Celonis environment and our designed tool to estimate the likelihood of order imperfection effectively and efficiently. The detailed description of the functionality of each notebook is provided in the “README.md” file.

User guidelines:

1. Accessing Celonis Environment
    - Open the Celonis environment and navigate to the left pane
    - Click on "Data" to proceed
2. Selecting Machine Learning
    - Locate the navigation bar and choose "Machine Learning" from the options

3. Choosing the Tim App
    - In the apps menu, select "Tim App 1" for further actions



4. Extracting Data from Celonis Pools
    - To extract data from Celonis pools, utilize the notebook titled "DataExtraction_LabellingComplete.ipynb"
5. Performing Exploratory Data Analysis
    - Conduct exploratory data analysis using the notebook named "EDA_DataCleaning_Complete.ipynb"

Note: Some datasets used in the notebook named "EDA_DataCleaning_Complete.ipynb" has been archived and are no longer accessible. Use the corrected dataset found in the data folder, please.

6. Non-Activity-Based and Activity-Based Modeling
    - Utilize the notebook labeled "Modelling.ipynb" for both non-activity-based and activity-based modeling

Note: The ROC Curve Analysis is an additional analysis that was done to check the model's correctness in predicting imperfection.

7. Running the Tool
    - For running the tool with both non-activity-based and activity-based approaches, access the "main.ipynb" file.
        - Enter the name of the test order file placed inside the data folder
        - Enter the name of the activity file placed inside the data folder

- Ensure the following dependencies are available: lib_withoutactivities, LR_model.joblib, SHAP_LR, lib_withactivities and SHAP_xgb_1

Note: It is recommended to follow the provided sequence of steps to ensure smooth operation within the Celonis environment.