COMP 4983: Project Description (Additional Information)

Now that you had a chance to perform data exploration and preprocessing on the training set, you shall proceed with the implementation, in Python, of the machine learning workflow which includes multiple iterations of data preprocessing, model training, model testing/tuning and model deployment. The following provides additional information on the project deliverables, including the due dates and the submission folders on the BCIT Learning Hub.

I Trained Model Assessment Checkpoints

To assess the performance your trained model(s), each team will be limited to a total of forty (40) submissions on the test set over the course of your development, with no more than ten (10) submissions for each checkpoint and no carryover across checkpoints. Submissions exceeding the submission limit will not be processed.

Performance assessments feedback will be provided on Fridays (for submissions up to 1530 Tuesdays) and on Saturdays (for submissions up to 1530 Fridays). Team performance assessment ranking will be posted under (News | Leaderboard) on the BCIT Learning Hub.

Deliverables:

For each of the four (4) checkpoints,

- Each team will use the features provided in *testset.csv* to predict the claim amount and submit the predicted claim amounts using the filename *checkpointnumber_groupnumber_submissionnumber.csv* for performance evaluation using the file format provided in *submission.csv*. In each of your submissions, you must include the Mean Absolute Error (MAE) on the training set in the upload comments.
- You will submit the source code and executable including detailed instructions on how to build and run the executable. Note that you will receive a mark of zero if your code does not compile or execute on the (freshly-imaged) instructor terminal in the option lab (SE12-306). Ensure your source code is adequately commented and submit using the filename *sourcecode_groupnumber_checkpointnumber.zip*.

The GROUP submission folders and deadlines for each checkpoint on the BCIT Learning Hub (Project Insurance Submission | …) are:

[Tue, Nov 1, 2022 @1530] Trained Model Assessment Checkpoint #1
[Tue, Nov 8, 2022 @1530] Trained Model Assessment Checkpoint #2
[Tue, Nov 15, 2022 @1530] Trained Model Assessment Checkpoint #3
[Fri, Nov 18, 2022 @1530] Trained Model Assessment Checkpoint #4

II Final Deployed Model Assessment

The final deployed model will be assessed using the competition set. The team with the best performing final deployed model on the competition set will be crowned the winner of this project. Team competition ranking for the groups will be announced after the final presentations. The top ranking team will earn a mark of 5.0%. Each team rank below will earn a mark of 1.0% less, i.e., 4.0% for 2nd, 3.0% for 3rd, etc. If a team fails to submit the final deployed model, a mark of 0.0% will be awarded.

Note: You will submit the source code and executable including detailed instructions on how to build and run the executable. Your program, a trained model, must take as input *competitionset.csv* (same format as *testset.csv*) and output *predictedclaimamount.csv* (same format as *submission.csv*). You will receive a mark of zero if your code does not compile or execute on the (freshly-imaged) instructor terminal in the option lab (SE12-306).

Deliverable:

Ensure your source code is adequately commented and submit using the filename *sourcecode_groupnumber_finalmodel.zip* to the BCIT Learning Hub. In your submission, you must include the Mean Absolute Error (MAE) on the training set in the upload comments.

The GROUP submission folder and deadline for the final deployed model on the BCIT Learning Hub (Project Insurance Submission | …) is:

[Tue, Nov 22, 2022 @1530] Final Deployed Model Assessment

III Project Presentation

Each group is allocated 30-mins for the project presentation, which includes AV setup (~1 mins), presentation time (~24 mins) and Q&A/discussion (~5 mins). All students must participate in the project presentation. In the presentation slidedeck, you must also include

- a citation slide listing all external resources
- a team member contribution slide that clearly indicates the project components that each team member has contributed to

The project presentation will be graded based on content, clarity, organization and delivery quality.

Deliverables:

You will submit the slidedeck using the filename *slidedeck_groupnumber.pptx* and all other supplemental files used in the material preparation, including figures, source code and executable with detailed instructions in *slidedecksupplementalfiles_groupnumber.zip*.

The GROUP submission folder and deadline for the slidedeck on the BCIT Learning Hub (Project Insurance Submission | …) is:

[Tue, Nov 29, 2022 @0930] Presentation

IV Peer Evaluation

You have the opportunity to confidentially evaluate the relative contributions of your team members toward the entire project. The guideline, along with the evaluation form can be downloaded from the BCIT Learning Hub (Project Material | Peer Evaluation).

You will submit the completed peer evaluation form (one submission per student only) using the filename *lastname_firstname.xlsx*. If you do not submit your evaluation form on time or fail to adhere to the evaluation method, your overall project mark will be reduced by 15%. No exceptions.

Deliverable:

The INDIVIDUAL submission folder and deadline for the completed peer evaluation form on the BCIT Learning Hub (Project Insurance Submission | …) is:

[Tue, Nov 29, 2022 @1730] Peer Evaluation

V Grading Breakdown

All work submitted is subject to the standards of conduct as specified in BCIT Policy 5104. Late submissions will not be accepted.

The grading breakdown for the project is as follows:

Project (20%):

Checkpoint #1 Submission	0.5%
Checkpoint #2 Submission	0.5%
Checkpoint #3 Submission	0.5%
Checkpoint #4 Submission	0.5%
Final Deployed Model Submission	0.5%
Competition Ranking	5.0%
Model Development	5.0%
Presentation	7.5%

Peer Evaluation (10%)