

2025-10-09 Meeting Notes

Date

09 Oct 2025

Participants

- RAIL PG-2 project team
 - Tao Xu a1937511
 - Sheng Wang a1903948
 - Jinchao Yuan a1936476
 - Zilun Ma a1915860
 - Di Zhu a1919727
 - Xin Wei a1912958
 - Yifan Gu a1909803
 - Tianhua Zhang a1915934
 - Zihan Luo a1916700
- Murtaza (Proxy Client)

Discussion topics

1. Github backlog

- Introduced completed tasks in sprint 3, including the implementation of additional feature selection, feature engineering, and ML techniques for addressing imbalanced datasets.
- Divided tasks for sprint 4 based on the user story 4.1, including multiple research on XAI techniques.

2. Project progress

2.1 Feature engineering:

- Applied resampling techniques on three training tables, including stratified and random oversampling/undersampling.
- Refactored the feature engineering script, facilitating the addition of new features.
- Added lag features for all numeric columns to observe their impact on model performance
- Introduced interaction features (e.g., curvature \times speed, speed \times twist) to assess potential performance improvements.

2.2 Feature selection:

- Explored feature selection strategies with varying time windows, Group Lasso settings, and parameter tuning in Focal Loss for imbalance handling.
- Applied a new method, gating, achieving an F1 score of 57.68%.

2.3 ML model

- Combined models with datasets that were handled by different feature selection and feature engineering techniques.
- Tried out feature selection and feature engineering techniques delivered above.
- Tuned hyperparameters for three models, including SVM, DNN, and transformer.
- For the three models implemented in sprint 2, the highest F1 score for SVM is 52.51%, for DNN is 30.00%, and for Transformer is 53.11%.
- Implemented a new LightGBM model. Performed training on the overall dataset followed by merging, and got an F1 score of 44.27%.
- Implement a new Group Lasso transformer model, and achieved the overall highest F1 score 57.68%.

2.4 Maintain the production line

3. Sprint 3 results

Leadboard F1 score results:

- **Group Lasso Transformer**

be55df2e	RAIL-PG-2	Completed	7 days ago	68d781805425.csv	Competition 2 - Senna	Accuracy: 64.32%, AUC_PR: 46.57%, F1_Score: 57.68%
----------	-----------	-----------	------------	------------------	-----------------------	---

- **Transformer**

33b2dc2d	RAIL-PG-2	Completed	15 days ago	68cccf8b6106.csv	Competition 2 - Senna	Accuracy: 45.16%, AUC_PR: 44.17%, F1_Score: 53.11%
----------	-----------	-----------	-------------	------------------	-----------------------	---

- **DNN**

1634f6f4	RAIL-PG-2	Completed	19 days ago	68c77c9a4802.csv	Competition 2 - Senna	Accuracy: 58.74%, AUC_PR: 32.60%, F1_Score: 30.00%
----------	-----------	-----------	-------------	------------------	-----------------------	---

- **SVM**

81170c14	RAIL-PG-2	Completed	16 days ago	68cb7e137514.csv	Competition 2 - Senna	Accuracy: 44.21%, AUC_PR: 59.94%, F1_Score: 52.51%
----------	-----------	-----------	-------------	------------------	-----------------------	--

- **LightGBM**

b23f6603	RAIL-PG-2	Completed	8 days ago	68d5f0057351.csv	Competition 3 - The Defibrillator	Accuracy: 38.63%, AUC_PR: 54.19%, F1_Score: 51.62%
----------	-----------	-----------	------------	------------------	-----------------------------------	--

Based on F1 score, the best model in Sprint 3 is the Group Lasso Transformer

4. Blockers

- Multiple resample techniques are tried out in sprint 3, so we got up to 8 processed training datasets. The IF was unstable and took too much time for a single task. We didn't have enough time to run training and prediction tasks for each model on all of the datasets, so each model only tried parts of the resampled datasets.

5. Next plan

Goal: In sprint 4, we will explore and apply XAI techniques for our models to avoid overfitting and aim for an F1 score over 60% and a higher AUC_PR score.

- Identify and explore at least 5 model XAI techniques relevant to the project.
- Implement and try out at least 3 model explainability techniques.
- Continue hyperparameter tuning from sprint 3.
- Experiment on different feature engineering methods, feature selection methods and techniques to approach an imbalanced dataset

6. QA

Q: In the last sprint, we preprocessed the dataset by combining different resampling methods. However, this increased the workload for training and improving the models significantly. For the current sprint, could you provide us with any suggestions on how to effectively integrate XAI techniques with the model to prevent this kind of exponential growth in workload?

A: Focus on the model that has the best performance and try different XAI techniques on it.

7. Suggestions

- **Instruction On Future Improvement**

- Try to add more hyperparameters in grid search.
- Try to remove those particular rows with most of the null values and different features.