2025-10-20 Meeting Agenda

🗓️ Date

20 Oct 2025

👥 Participants

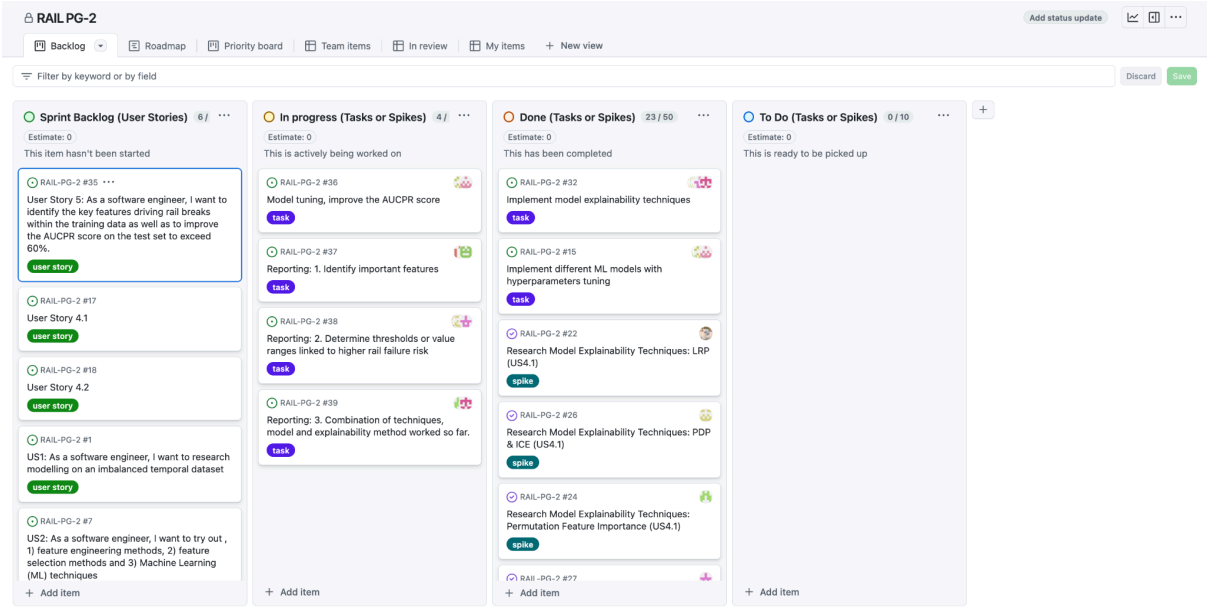
- RAIL PG-2 project team
    - Tao Xu a1937511
    - Sheng Wang a1903948
    - Jinchao Yuan a1936476
    - Zilun Ma a1915860
    - Di Zhu a1919727
    - Xin Wei a1912958
    - Yifan Gu a1909803
    - Tianhua Zhang a1915934
    - Zihan Luo a1916700
- Murtaza (Proxy Client)

🥅 Goals

- Github backlog overview
- Report progress
- Submission results
- Next steps
- QA

🗣️ Discussion topics

1. GitHub backlog overview



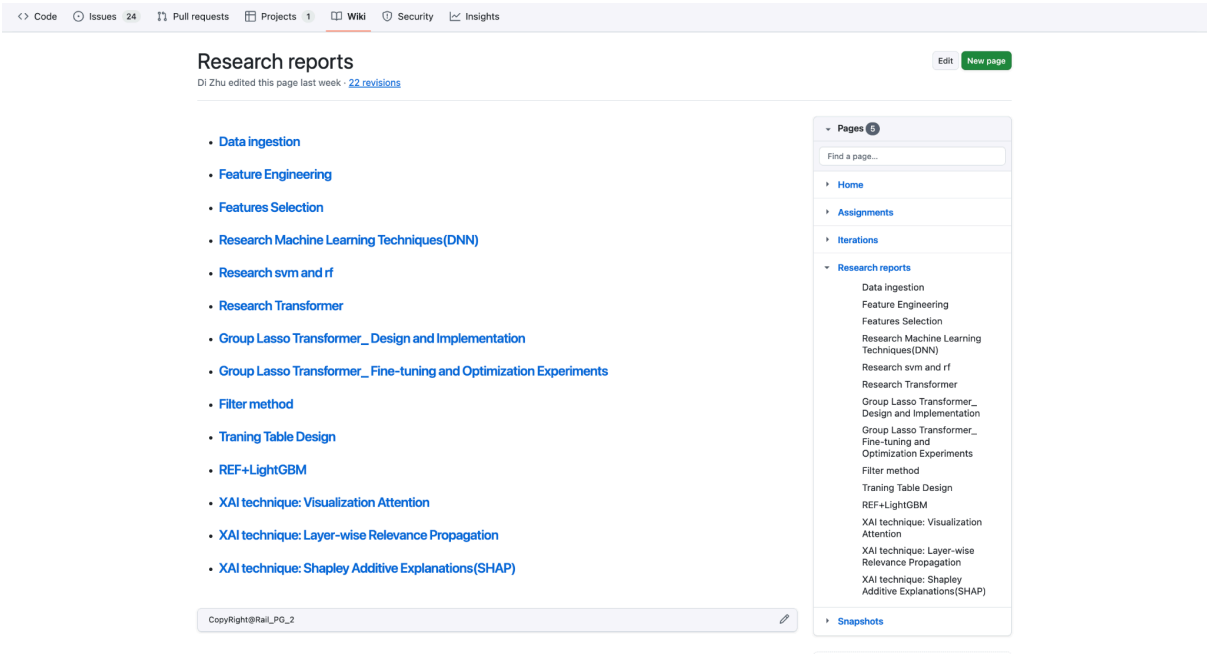
2. Report progress

2.1 XAI technique Research:

Each team member is assigned to research one XAI technique and share the result with the team. The research outcomes are added to the wiki page on GitHub.

Related tasks:

1. Research Model Explainability Techniques: EBM
2. Research Model Explainability Techniques: PDP & ICE
3. Research Model Explainability Techniques: Visualizing Attention
4. Research Model Explainability Techniques: Permutation Feature Importance
5. Research Model Explainability Techniques: LIME
6. Research Model Explainability Techniques: LRP
7. Research Model Explainability Techniques: Integrated Gradients
8. Research Model Explainability Techniques: Counterfactual Explanations (DICE)
9. Research Model Explainability Techniques: SHAP



2.2 XAI technique Implement:

XAI technique candidates:

- LIME
- SHAP
- DICE

ML model candidates:

- Group Lasso Transformer
- Transformer
- DNN
- SVM
- LightGBM

After the internal meeting of the team, **LIME, SHAP, and DICE** are chosen to be applied to the machine learning models.

Each team member chose a combination of model and XAI technique and implemented them. Shared the visual chart screenshot in github task comments.#32

2.3 Maintain the production line

2.4 Model tunning

3. Submission results

The best combination for AUC_PR is:

- Model: SVM
- XAI: SHAP
- Rebalanced technique: Class weighting
- Tuning Hyperparameters: Manual C selection
- Accuracy: 43.37%, ACC_PR: 61.49%, F1_score: 51.88%

The best combination for F1 is:

- Model: Group Lasso Transformer
- Rebalance technique: Stratified batch resampling (75% positive samples per batch) combined with weighted loss function (pos_weight = neg_count/pos_count)
- XAI: SHAP

  The best score of using SHAP explanation is to select top 10 features:

  Accuracy: 64.98%, ACC_PR: 46.00%, F1_score: 52.42%

- The best score for F1 is using gating technique to select top8 features:

  Accuracy: 64.32%, ACC_PR: 46.57%, F1_score: 57.68%

4. Next steps

Goal: In Sprint 5, our goal is to continue optimizing the model, maintain the **AUCPR** above 60, organize the experimental data, identify key features and risk ranges, and finally sort out the best model combination at present.

- Model tuning, improve the AUCPR score
- Identify important features
- Determine thresholds or value ranges linked to higher rail failure risk
- Sort out combination of techniques, model and explainability method worked so far

5. QA

Q: Details related to the final report and final presentation