ADS-540 Group 3 Weekly Team Tracker

CMS Open Payments Risk Scoring & Anomaly Detection

Please copy and share this Google document with your project teammates.
Then, set general access sharing settings to "University of San Diego" and "Commenter" so your instructor can view or comment.

Project Group 3:

- Team member 1: Swapnil Patil
- Team member 2: Jamshed Nabizada
- Team member 3: [if applicable]
- Asana Board link: https://app.asana.com/1/952672460738672/project/1212851836514318/list/1212851844967962
- Project GitHub link: https://github.com/swapnilprakashpatil/aai540_3proj

In this Document:

1. **Teamwork Tracker**: For accountability, each team member should list briefly your weekly responsibilities whether you have completed the tasks or not. The point of this is to disperse work evenly among team members and ensure equitable work and continuous communication. Include the Team Tracker link in your Design Document draft in Module 2. Post the Team Tracker link to the *Team Project Update* in Canvas (Modules 3 thru 6.) This is a 2-3 minute task.

2. **Module Steps**: Recommended Module-by-Module project steps are found on the last pages in this document.

Weekly Teamwork Tracker

Module 2 Weekly Teamwork Tracker

Please copy and share a single Team Tracker document for your team. Then, individually update by Day 7 of Modules 2 thru 6. This should take 1-2 minutes for each person to update individually.
For Module 2, update your Design Document with your Team Tracker link

Teamwork Communication (Summary)

How many times have your team members met, exchanged project communications, or collaborated on work in the last week?
2 times.

Weekly Project Snapshot - Individual Contributions

As a team accountability practice and at-a-glance (for your instructor), please **summarize** your responsibilities and progress for the week in the table below.

Team Member 1 (Swapnil)	Team Member 2 (Jamshed)	Team Member 3 (Name)
List of contributions - Created Asana board - Created Github repo - Created tracker document - Working on EDA	List of contributions - Create s3 bucket for data - Create README for repo - Working on Feature engineering	List of contributions
Comments/ Roadblocks - None, our schedule is busy with work so we have to get creative on this project.	Comments/ Roadblocks - None, our schedule is busy with work so we have to get creative on this project.	Comments/ Roadblocks

Module 3 Weekly Teamwork Tracker

List briefly your weekly contribution to teamwork and any issues or roadblocks. Post the link to the *Team Project Update* in Canvas for this module. (This is a 2-3 minute task).

Teamwork Summary

How many times have your team members met, exchanged project communications, or collaborated on work in the last week?

Weekly Project Snapshot - Individual Contributions

Summarize your responsibilities and progress for the week in the table below.

Team Member 1 (Name)	Team Member 2 (Name)	Team Member 3 (Name)
List of contributions	List of contributions	List of contributions
Comments/ Roadblocks	Comments/ Roadblocks	Comments/ Roadblocks

Module 4 Weekly Teamwork Tracker

List briefly your weekly contribution to teamwork and any issues or roadblocks. Post the link individually to the *Team Project Update* in Canvas for this module. (This is a 2-3 minute task).

Teamwork Summary

How many times have your team members met, exchanged project communications, or collaborated on work in the last week?

Weekly Project Snapshot - Individual Contributions

Summarize your responsibilities and progress for the week in the table below.

Team Member 1 (Name)	Team Member 2 (Name)	Team Member 3 (Name)
List of contributions	List of contributions	List of contributions
Comments/ Roadblocks	Comments/ Roadblocks	Comments/ Roadblocks

Module 5 Weekly Teamwork Tracker

List briefly your weekly contribution to teamwork and any issues or roadblocks. Post the link to the *Team Project Update* in Canvas for this module. (This is a 2-3 minute task).

Teamwork Summary

How many times have your team members met, exchanged project communications, or collaborated on work in the last week?

Individual Contributions

Summarize your responsibilities and progress for the week in the table below.

Team Member 1 (Name)	Team Member 2 (Name)	Team Member 3 (Name)
List of contributions	List of contributions	List of contributions
Comments/ Roadblocks	Comments/ Roadblocks	Comments/ Roadblocks

Module 6 Weekly Teamwork Tracker

List briefly your weekly contribution to teamwork and any issues or roadblocks. Post the link to the *Team Project Update* in Canvas for this module. (This is a 2-3 minute task).

Teamwork Summary

How many times have your team members met, exchanged project communications, or collaborated on work in the last week?

Individual Contributions

Summarize your responsibilities and progress for the week in the table below.

Team Member 1 (Name)	Team Member 2 (Name)	Team Member 3 (Name)
List of contributions	List of contributions	List of contributions
Comments/ Roadblocks	Comments/ Roadblocks	Comments/ Roadblocks

Module-by-Module Steps

General Guidance

Below are **general** steps as guidance for expected teamwork for each module. Refer to the specific details in Canvas for the Team Project Update each week. .

Week 1

Individual Project Proposal/Group Sign-ups (Self-Enroll)

1. Propose a problem that is solvable with a Machine Learning System.
2. Write a RFC (Request for Comments) Document.
3. After completing your RFC document, post your RFC in the canvas discussion forum.
4. Comment on 2-3 projects that you find interesting and try to find a partner.
5. Confirm a team of 2-3 people to work on your final project with (Day 7).
6. Complete Group Sign-ups in Canvas (Day 7, ALL students, even those working alone)
7. Copy/share, and then update/post this tracker to the Team Project Update in Canvas (Day 7).

Week 2

Team Project Proposal, Data Engineering

1. With your team, agree on a problem that is solvable with a Machine Learning System.
2. Begin writing your Design Document. (See AAI-540 Design Document Template).
3. Set up an Asana board and Github to manage your project.
4. Collect a raw data set and store in an S3 Datalake.
5. Set up Athena tables to enable cataloging and querying of your data.
6. Complete indicated fields and update the design document draft to submit to the Team Assignment (see Canvas Assignment instructions).

Week 3

Training Data and Feature Engineering

1. Perform exploratory data analysis on your data in a SageMaker notebook.
2. Initialize a feature store.
3. Design the feature groups needed for your system.
4. Perform feature engineering on raw data and store features in feature groups.
5. Split your feature data into training (~40%), test (~10%) validation (~10%) datasets.
6. Reserve some data for “production data” (~40%).
7. Update and post the tracker to the Team Project Update.

Week 4

Model Development and Deployment

Check In #2 (See Check-In Outline for Deliverables)

1. Set up a benchmark model in SageMaker.
- The benchmark model should be simple, it can be a simple heuristic or a model with just a couple of features.
- The idea here is to have a baseline for further improvements, and to get a minimum viable. product out before you start improving your system.
2. Build, train and debug your ML model in a SageMaker model.
- This is your first real iteration on your model, it doesn't need to be perfect, you may want to revisit model development again once we implement CI/CD in module 6
3. Evaluate your model and compare against your simple benchmark model.
4. Deploy your model to SageMaker (Batch process or Real Time Endpoint).
5. Link your initial findings and codebase in your ML Design Document.
6. Update and post the tracker to the Team Project Update.

Week 5

Monitoring

1. Implement model monitors on your ML system.
2. Implement data monitors on your ML system.
3. Implement infrastructure monitors on your ML system.
4. Create a monitoring dashboard for your ML endpoint/job on Cloudwatch
5. Generate model and data reports on SageMaker.
6. Update and post the tracker to the Team Project Update.

Week 6

CI/CD

1. Implement CI/CD Pipeline to automate training, evaluation, and deployment.
    - CI/CD pipeline should have checkpoints to evaluate model performance, model code and system integration.
    - Your CI/CD pipeline should train with training data, and evaluate with testing data.
    - You should try to improve your initial model and run it through your CI/CD pipeline.
1. Update and post the tracker to the Team Project Update.

Week 7

Wrap Up Final Project

Group Deliverables Due:

1. Deliverable 1: ML Design Document
2. Deliverable 2: ML System Operation Validation
3. Deliverable 3: Codebase GitHub Repository

Individual Deliverable Due:

1. Peer Evaluation Form