**Name:** DSBA Student

**Date:** 20 March, 2022

**Project Title:** Proof of Concept ML Pipeline in Google Vertex AI

**Mentor and Company:** Mentor Name, Company Name

**Period Covered**: 7 – 18 March 2022

**Project Objective:** Build, test, analyze viability, and improve performance of a proof-of-concept start-to-finish machine learning pipeline.

**Progress Summary:**

Work Completed.

* **Task 3.2.** (**COMPLETE)** CP-989: Make changes to `tx-prod` service to call the cloud function
  + **Summary:**
    - The `tx-prod` service is now using the cloud function to retrieve `is\_x` predictions in a completely serverless manner
    - Unit tests were written and passed for all modifications to the service
    - The full data reload test for one chosen vendor was comparable to or better than the previous `ml-x` microservice’s performance
    - The `ml-x` microservice can now be safely retired
  + **Lessons Learned:**
    - GCP cloud functions require identity tokens rather than access tokens, which can be retrieved from the metadata server and employed using the functions now defined in `tx-prod` daemon
  + **Risk & Mitigation:** Unit testing and authentication with Google for the cloud function took more time than anticipated, leading to a delay in the project of about 1 week. Additionally, based on feedback from a coworker on another concurrent Vertex pipeline effort, the time allotted for the pipeline construction has been increased by 1 week.
    - Mitigation: Expectations for the scope of Task 4, improving and expanding the model, has been adjusted.

Work Scheduled Next Sprint.

* **Task 3.3.** DMI-184: Assess ground truth data and measure the model’s current performance
  + Update query to pull most up-to-date training data from new table schemas
  + Compare new data pull to old master list of recognized products
  + Validate and update the master list as needed
  + Aggregate new unseen test data and reassess the model’s current performance as a baseline
* **Task 3.4.** DMI-183: Define, compile, and deploy the `is\_x` pipeline in Vertex AI
  + Build components to:
    - retrieve data,
    - engineer features,
    - train the model,
    - evaluate performance, and
    - deploy, conditional on meeting the performance threshold
  + Store reusable components in co-vertex-utils
  + Write and pass unit tests for all reusable components
* **Risks & Mitigations**
  + Risk: Some vendor data may not be sufficiently trustworthy to simply use all transactions with sales\_type=x as ground truth data
    - Mitigation: Quality control any vendors used for ground truth data, using either algorithmic means, manual inspection, or both
  + Risk: Google documentation on Vertex pipelines is sparse and somewhat disorganized, which introduces some uncertainty and hence some risk to the project timeline. Specifically, although the pipeline itself (Task 3) is the primary objective of the project, which will be completed, there may be less time available to devote to the subsequent improvement and expansion of the model (Task 4) due to troubleshooting on the pipeline construction and compilation.
    - Mitigation: Coordinate closely with coworkers working concurrently on other pipelines to maximized sharing of work and reusing of component structures.

**Work Package Updates:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **WP** | **Work Package Title** | **Scheduled Dates** | **Status** |
| **✓** | **~~1.~~** | **~~Onboarding~~** | **~~10-Jan – 11-Feb~~** | **Complete** |
| **✓** | **~~2.~~** | **~~CP-856: Test performance of `is\_x` model on Vertex Endpoint~~** | **~~17-Jan – 18-Feb~~** | **Complete** |
|  | **3.** | **CP-857: Build a proof-of-concept ML pipeline in Vertex that can effectively replace the current is\_x data pipeline** | **21-Feb – ~~25-Mar~~**  **8-Apr** | **In progress** |
| **✓** | ~~3.1.~~ | ~~CP-988: Create a Google Cloud Function to effectively replace the `ml-x` service~~ | ~~21-Feb – 4-Mar~~ | Complete |
| **✓** | ~~3.2.~~ | ~~CP-989: Make changes to tx-prod to call CF~~ | ~~7-Mar – 11-Mar~~ | ~~Complete~~ |
|  | 3.3. | DMI-184: Assess `is\_x` ground truth data and measure the model’s current performance | ~~14-Mar – 18-Mar~~  21-Mar – 25-Mar | Backlog |
|  | 3.4. | DMI-183: Define, compile, and deploy the `is\_x` pipeline in Vertex AI | ~~21-Mar – 25-Mar~~  28-Mar – 8-Apr | Backlog |
|  | **4.** | **Explore options for improving is\_x prediction performance and expanding to additional product categories.** | **11-Apr – 29-Apr** | **Backlog** |
|  | 4.1. | Explore model improvement | 11-Apr – 20-Apr | Backlog |
|  | 4.2. | Explore model expansion | 21-Apr – 29-Apr | Backlog |