Building and Monitoring a KubeFlow Machine Learning Pipeline Using AWS EKS, Prometheus, and Grafana

Ryan T. Peterson

Western Governors University

**Table of Contents**

[Proposal Overview 3](#_Toc441469365)

[Problem Summary 3](#_Toc441469366)

[IT Solution 3](#_Toc441469367)

[Implementation Plan 3](#_Toc441469368)

[Review of Other Work 3](#_Toc441469369)

[Relation of Artifacts to Project Development 4](#_Toc441469370)

[Project Rationale 4](#_Toc441469371)

[Current Project Environment 4](#_Toc441469372)

[Methodology 4](#_Toc441469373)

[Project Goals, Objectives, and Deliverables 5](#_Toc441469374)

[Goals, Objectives, and Deliverables Table 5](#_Toc441469375)

[Goals, Objectives, and Deliverables Descriptions 6](#_Toc441469376)

[Project Timeline with Milestones 6](#_Toc441469377)

[Outcome 7](#_Toc441469378)

[References 8](#_Toc441469379)

[Appendix A: 9](#_Toc441469380)

[Appendix B: 10](#_Toc441469381)

[Appendix C: 11](#_Toc441469382)

[Appendix D: 12](#_Toc441469383)

# Proposal Overview

AstroMined Labs is a stealth startup developing a plan to launch what many believe could be the most disruptive and challenging mission in human history. The company will be the first to conduct a detailed survey of the asteroid belt between Mars and Jupiter to find the most suitable candidates to mine for natural resources. Although the scale of this ambition is undoubtedly massive, AstroMined Labs plans to leverage machine learning and automation to maximize their chance of success while minimizing mission costs.

## Problem Summary

Current scientific belief is the asteroid belt contains between 1.1 and 1.9 million asteroids that are more significant than one kilometer (0.6 miles) in diameter and millions of more minor asteroids. (In Depth | Asteroids – NASA Solar System Exploration, n.d.) To minimize the time and money spent on the survey, AstroMined Labs will launch 500 microsatellites in an arcing pattern on each mission, with each microsatellite traveling towards a pre-determined sector of the belt. They will use their robust long-range sensors to gather data about the asteroids in their designated sector. Once they have covered half the distance between the Earth and the belt, the microsatellites will come to a consensus about the most promising sector, where they will converge to perform a detailed survey.

Since the microsatellites will be operating autonomously, they need to use as much of their energy and computing power for data collection and navigation. Therefore, Earth-based servers will conduct all computationally intensive machine learning model-building processes. The vast distance between Earth and the belt compounds the challenges presented by the sheer number of asteroids there, as communication delays between Earth and the belt can exceed 20 minutes. These factors combine to necessitate an extremely performant machine learning pipeline on the Earth-based servers.

## IT Solution

With petabytes of data streaming back to Earth daily, it will be essential to train machine learning models using a distributed solution to minimize the time between when new data is received, and updated models are sent back to the microsatellites. KubeFlow is a purpose-built project to enable large machine learning pipelines to operate in a distributed Kubernetes-based computing environment. As 100% uptime is essential to the success of this project, the ML Ops team will monitor every part of the KubeFlow infrastructure using Prometheus and Grafana to allow them to detect problems proactively before they worsen.

AstroMined Labs will use their on-premises server farm as the primary storage pool to minimize the cost of storing such a massive amount of data. Using EKS Anywhere combined with an Istio Service Mesh, they can achieve a seamless hybrid computing environment that maximizes performance while minimizing costs. The control plane used will be AWS Elastic Kubernetes Service (EKS) so that powerful GPU-enabled instances can do the computational heavy lifting. Prometheus and Grafana will also be deployed in a distributed manner for ease of management and scalability using Kubernetes.

## Implementation Plan

This section should contain the details of an implementation plan. Provide details on the different phases (if applicable). Explain how the project will be carried out and why it will be carried out in that manner. Discuss the plan for implementing the project.

# Review of Other Work

In this section (*suggested length of 3–4 pages*), review other works done by a third party that are relevant and directly relate to the project. Review at least four other works completed on the chosen topic. Summarize interviews, white papers, research studies, or other types of work by industry professionals. For example, if planning a server upgrade project, summarize articles that describe successful upgrade server solutions in other companies, comparisons of various server hardware and software products that are possible solutions, and any case studies that relate or are similar to the planned server upgrade.

Include the following in your review:

## Relation of Artifacts to Project Development

Provide a logical description of how each work reviewed relates to the proposed development of the project. Explain how each of the chosen works contextualize the problem or provide direction to the project.

# Project Rationale

Summaries should include the rationale for choosing this project, including what makes this problem interesting or significant. The Project Rationale section (*suggested length of 1–2 pages*) provides a rationale for the project. It should address the reasons for implementing the project, as described in the Proposal Overview. The rationale may include and expand on both business and technical reasons to support the implementation. Justify the selection of the project.

# Current Project Environment

This section (*suggested length of 2–3 pages*) describes and details the current project environment. It should also address specifically how the current state will set the direction for the definition and implementation of the proposed solution. Other details that support the description should also be included here as needed to support and succinctly define the project environment. Analyze the systems and describe the status of the project environment before the project began.

# Methodology

This section (*suggested length of 1–2 pages*) describes and details the specific methodology. The methodology is the process that the project will follow when it is implemented. Include specific details to adequately describe the steps that will take place to fully execute the project. Explain how a standard methodology (such as PDCA, ADDIE, SDLC, Prototyping or Agile) will be applied for the implementation of the project.

# Project Goals, Objectives, and Deliverables

In Project Goals, Objectives, and Deliverables (*suggested length of 3–5 pages*), provide a detailed explanation of the goals and objectives for the project, and explain what the project will provide. There should be a specific and clear identification of each goal and objective, as well as a clear and logical description with adequate detail of each of the goals and objectives. Address and explain the criteria that will be used to measure the objectives in order to demonstrate how each goal and objective will be measured. The project will include some sort of formal report. It will likely also include a technical IT product and/or a user’s manual or other documentation. Deliverables should provide a detailed logical explanation of what the project will provide to substantiate the work and completion of such. Include the following in the description:

## Goals, Objectives, and Deliverables Table

Every project has one or more goals. Each goal is supported by at least one project objective. Each objective is enabled by at least one project deliverable. Fill out this or a similarly organized table:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Goal | Supporting objectives | Deliverables enabling the project objectives |
| 1 | Summarize project goal 1 | 1.a. Describe project objective 1.a. | 1.a.i. Explain project deliverable 1.a.i |
| 1.a.ii. Explain project deliverable 1.a.ii |
| … |
| 1.b. Describe project objective 1.b. | 1.b.i. Explain project deliverable 1.b.i |
| 1.b.ii. Explain project deliverable 1.b.ii |
| … |
| … | … |
| … |
| … |
| 2 | Summarize project goal 2 | 2.a. Describe project objective 2.a. | 2.a.i. Explain project deliverable 2.a.i |
| 2.a.ii. Explain project deliverable 2.a.ii |
| … |
| 2.b. Describe project objective 2.b. | 2.b.i. Explain project deliverable 2.b.i |
| 2.b.ii. Explain project deliverable 2.b.ii |
| … |
| … | … |
| … |
| … |
| … | … | … | … |
| … |
| … |
| … | … |
| … |
| … |
| … | … |
| … |
| … |

## Goals, Objectives, and Deliverables Descriptions

Describe each of the project goals. Describe each objective. Explain how the objectives support the goals, and explain how the objectives will be achieved. Explain what types of deliverables the project will provide, and describe the key project deliverables expected by the end of the project. Provide a description of the project goals, objectives, and deliverables that aligns with the Goals, Objectives, and Deliverables Table.

# Project Timeline with Milestones

In this section (*suggested length of 1–2 pages*), provide a projected timeline with milestones for the project. These may be estimates that will most likely be adjusted, as many times the project will require adjustments during the development and implementation phases. Provide a projected timeline with milestones for the project, including the duration and start and end dates of each milestone. Include the following table:

|  |  |  |  |
| --- | --- | --- | --- |
| Milestone or deliverable | Duration  (hours or days) | Projected start date | Anticipated end date |
|  |  |  |  |
|  |  |  |  |

# Outcome

In this section, describe the anticipated project outcomes and explain how the success of the project will be measured once completed. Explain the expected project outcomes and describe the evaluation framework to be used once the project is completed to assess the project’s success and effectiveness.

# References

List all the outside sources that the narrative refers to in text. For information regarding in-text and reference list citations, please refer to the web link or visit the WGU Writing Center.

Smyth, A. M., Parker, A. L., & Pease, D. L. (2002). A study of enjoyment of peas. Journal of Abnormal Eating, 8(3), 120-125. Retrieved from

http://www.articlehomepage.com/full/url/

Bernstein, M. (2002). 10 tips on writing the living Web. A List Apart: For People Who Make Websites, 149. Retrieved from http://www.alistapart.com/articles/writeliving

Bell, T., & Phillips, T. (2008, May 6). A solar flare. Science @ NASA Podcast. Podcast retrieved from http://science.nasa.gov/podcast.htm

OLPC Peru/Arahuay. (n.d.). Retrieved April 29, 2011 from the OLPC Wiki: http://wiki.laptop. org/go/OLPC\_Peru/Arahuay

Plath, S. (2000). The unabridged journals. K. V. Kukil (Ed.). New York, NY: Anchor.

# Appendix A

# Title of Appendix

Put any supporting material in these appendices. Add additional or delete superfluous appendices as needed.

# Appendix B

# Title of Appendix

Put any supporting material in these appendices. Add additional or delete superfluous appendices as needed.

# Appendix C

# Title of Appendix

Put any supporting material in these appendices. Add additional or delete superfluous appendices as needed.

# Appendix D

# Title of Appendix

Put any supporting material in these appendices. Add additional or delete superfluous appendices as needed.