

Definition of Done MLOps

December 6, 2021

1 Introduction

One of the deliverables within this project is a MLOps pipeline used for the periodic extracting and processing of climate data, and the training, validating and uploading of a machine learning model. This document describes what steps the MLOps pipeline should contain and when it is considered done.

2 What should the MLOps pipeline do?

The MLOps pipeline should be able to periodically collect new data from a variety of sources. This data should then be processed. After processing a machine learning model must be trained. To ensure the trustworthiness of the model an accuracy test must be passed before it's deployed to our application.

3 When do we consider the MLOps pipeline to be "done"?

The MLOps pipeline is considered done when the following pipeline steps are implemented:

1. Obtain the dataset
2. Clean the dataset
3. Process the dataset
4. Train model
5. Validate that the model accuracy is above a certain threshold.
6. Upload trained model to prototype