Esoon Ko
IT Högskolan

Project

Group work: 2-3 individuals in group

Two possible tasks for you to choose from:

TASK 1

Scenario:

You are a new Data Team at a small firm Something Weather AB (SWAB).
The firm has built an ML model as a POC with a dataset they purchased.
They want your team to build the company's data platform so that their ML model can use the data from the platform for future training. They also want your team to deploy the model so that their end user can use the model.
Your team is given a deadline of 19 September to both launch the products and hold a presentation on your products, workflow process and why you did things the way you did.

For grade G:
- Build a data platform that takes in data from the given public API and processes data for consumption by the ML model for future training or analysis. The platform must have schedules for the pipelines as well as possibility to monitor and logging.
- Deploy the given ML model to the cloud for use by end user.

Other Requirements:
- Use agile methods during the project. Document the proof of the methods you used.
- Containerize services through docker or other containerization technology such as Kubernetes.

Group work: 2-3 individuals in group

Two possible tasks for you to choose from:

TASK 1

Scenario:

You are a new Data Team at a small firm Something Weather AB (SWAB).
The firm has built an ML model as a POC with a dataset they purchased.
They want your team to build the company's data platform so that their ML model can use the data from the platform for future training. They also want your team to deploy the model so that their end user can use the model.
Your team is given a deadline of 19 September to both launch the products and hold a presentation on your products, workflow process and why you did things the way you did.

For grade VG:
- Build a data platform that takes in data from the given public API and processes data for consumption by the ML model for future training or analysis. The platform must have schedules for the pipelines as well as possibility to monitor and logging.
- Deploy the given ML model to the cloud for use by end user.
- Find an additional relevant data API and create an additional data pipeline. Use the data from this new data source in combination with the existing data in order to gather new insight through a dashboard. Optionally train a new ML model that uses both data sources. Examples include: Bike traffic data.

Other Requirements:
- Use agile methods during the project. Document the proof of the methods you used.
- Containerize services through docker or other containerization technology such as Kubernetes.
- Apply CI/CD techniques and automated testing + linting into your project.

Group work: 2-3 individuals in group

Two possible tasks for you to choose from:

TASK 1

Materials:
https://drive.google.com/drive/folders/1ArHYeS7L3rRXM_Sgmst1sDkmSVDmDPXV?usp=sharing

Code:
Github Repo - https://github.com/esoonko/Data-engineering-AI23
lectures/code/project/ml-model

Group work: 2-3 individuals in group

Two possible tasks for you to choose from:

TASK 2

Scenario:

You are a new Data Team at a small firm {INSERT YOUR FICTIONAL COMPANY AB}.
The firm wants to build ML functionality as part of their business model.
They want your team to create a ML POC with any data API of your choice.
They want your team to also build the company's data platform so that their ML model can use the data from the platform for future training. They also want your team to deploy the model so that their end user can use the model.
Your team is given a deadline of 19 September to both launch the products and hold a presentation on your products, workflow process and why you did things the way you did.

For grade VG:
- Use any data API of your choice to gain insight and train a new ML model to solve a problem you identified.
- Build a data platform that takes in data from the given public API and processes data for consumption by the ML model for future training or analysis. Create analysis dashboard from this data. The platform must have schedules for the pipelines as well as possibility to monitor and logging.
- Deploy the given ML model to the cloud for use by end user.

Other Requirements:
- Use agile methods during the project. Document the proof of the methods you used.
- Containerize services through docker or other containerization technology such as Kubernetes.
- Apply CI/CD techniques and automated testing + linting into your project.

Group work: 2-3 individuals in group

Pair up with people that have:

- Similar grade ambition
- Location