Anomaly detection and prediction in operational log data

January 2021

1. Description of duties:

1.1. Mohammed Ibrahim (WAFH4I)

- Data exploration on Jupyter Notebook: Initial exploration of shared data to understand the fields and get some summary statistics.
- Setting up Kafka Connect to read data into Kafka topic on local machine.
- Kafka producer module to load data from csv to Kafka topic using Java.
- Converting epoch field to get DateTime info (Kafka stream).
- Kafka consumer module to stream data from Kafka to Cassanra.
- Using Kafka connect Cassandra to connect between Kafka and Cassandra.
- Setting up Cassandra DB on local machine.
- Storing Log data on Cassandra.
- Code repository management on Github: creating repository, reviewing pull requests and access management.
- Writing report.
- Preparing presentation material.

1.2. Ammar Reslan (YNBLGQ)

- Data preprocessing on Colab Notebook: adding the time stamp and extracting new colums daylight, Dayof the Week, weekday.
- Creating lstm anomaly detector function.
- Explore, create visualization charts using matplotlib visualizations
- Bulid a LSTM model, train it, and save model weights to be used from the detector function.
- Contribute in Writing report and preparing presentation material.
- Participating in project team's meetings.

1.3. Sakina Hajiyeva (DKXCU0)

- pyKafka Consumer, visualization, SARIMA.

1.4. Vitalii Naumov (FY46IN)

- Docker, pyKafka Producer, MongoDB.