

# Anomaly Detection In Financial Data

Daniel Grimm, Gayrat Rakhimov, Suchith Shetty

December 2020

## 1 Description of duties performed by project team members

### 1.1 Gayrat Rakhimov (IGMQ0T)

- Data preprocessing on Jupyter Notebook: modifying balance and exchange rates data
- Setting up Kafka Connect to read balance data into 'balance\_topic' stream
- Reading Twitter tweets feed from Twitter API
- Merging balance data and exchange data
- Filtering balance data according to balance differences
- Setting up remote ElasticSearch on a third-party host (Bonsai)
- Storing Twitter data on ElasticSearch
- Code repository management on Github: creating repository, reviewing pull requests and access management
- Participating in project team's meetings
- Writing report
- Preparing presentation material

### 1.2 Daniel Grimm (BU44BJ)

- Configuring and setting up Docker containers for Elasticsearch and Kibana
- Connecting pipeline: local installation + docker
- Parsing data from balance.csv
- Creating multiple anomaly detector functions

- Creating kibana visualizations
- Code repository management on Github: reviewing and creating pull requests
- Participating in project team's meetings

### 1.3 Suchith Shetty (HIUH69)

- Initial exploration of shared data to understand the fields and get some summary statistics
- Kafka producer module to load data from csv to Kafka topic using Python
- Setting up Elasticsearch and Kibana on a third-party host (Bonsai) and in local machine
- Kafka consumer module to stream data from Kafka to Elastic Search
- Explore, create visualization charts and compile them to create output dashboard in Kibana
- Contribute in stream transformations for anomaly detection and data engineering to make output streams compatible for easy visualization
- Contribute in report writing and preparing presentation material