Blood & Money

Detecting and Visualizing Blood Analyzer Anomalies

W205 Section 4 Final Project
James, Weixing, Annalaissa, Yang

Business Problem

- Our client is a blood analyzer device manufacturer
- They have thousands of customers throughout the nation
- By federal law, every blood analyzer device must be regularly tested to ensure a baseline in measurements
- Incorrect measurements can lead to false diagnosis
- How can we improve predicting performance problems with the machines?

Our approach

- We built a scalable data transformation pipeline using three aggregated client data sources:
  - Quality Control
  - Machine Raw
  - Sample Means
- Used Recurrent Neural Net with Tensorflow to build a predictive model
- And presented the analysis in two ways:
  - Tableau book to highlight statistical properties
  - Dynamic dashboard for visualizing outliers

Dataflow High-level

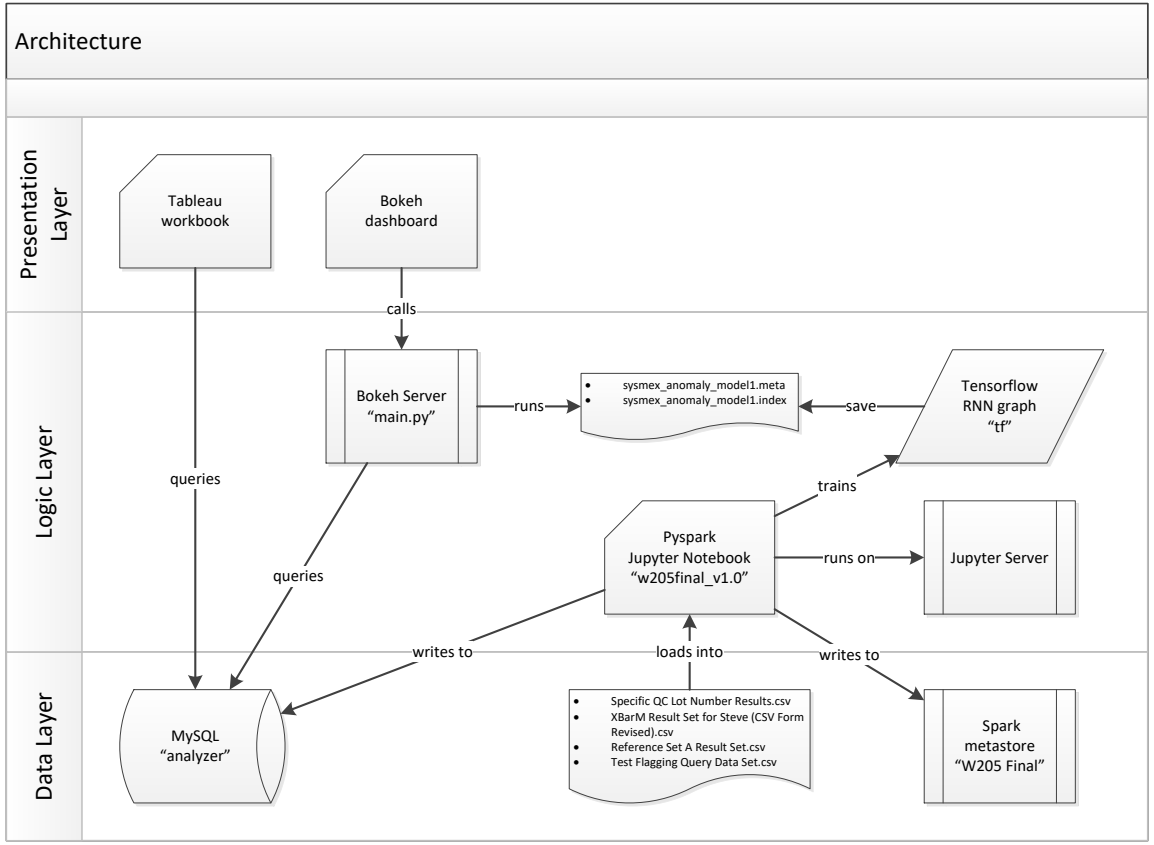
- Load raw csv

- Cleanse data
- Push to database

- Build predictive model

- Visualize data
- Show outliers

Components of the Project



Results

Bokeh dashboard 1 “main.py”

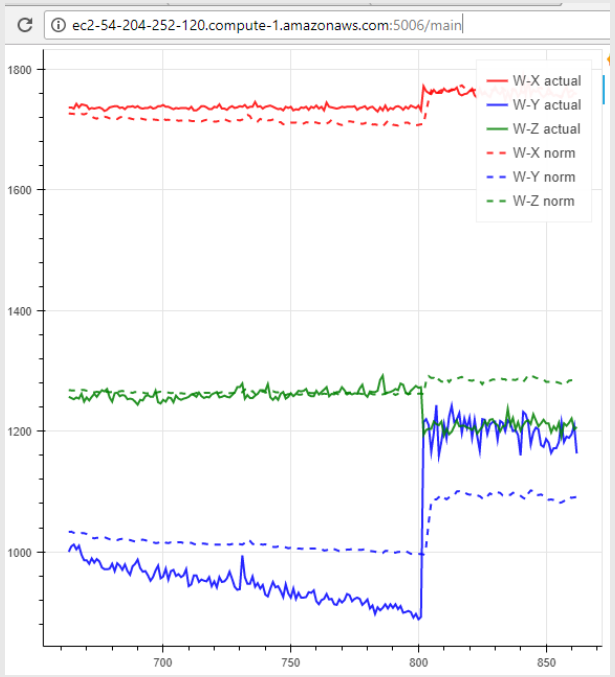
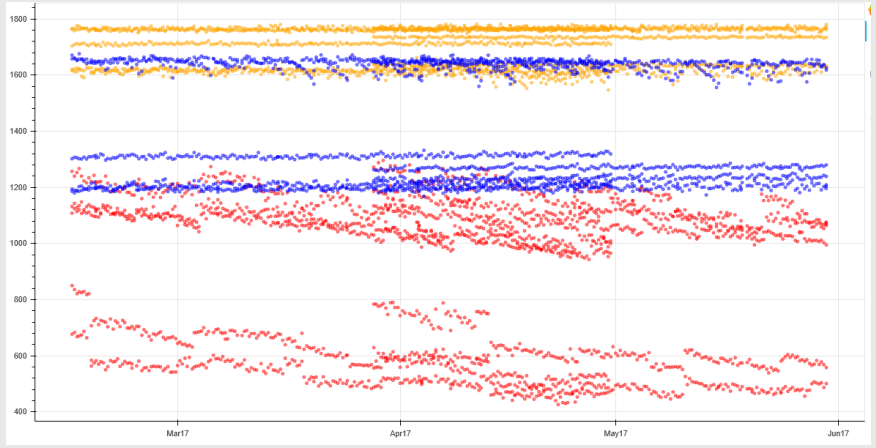


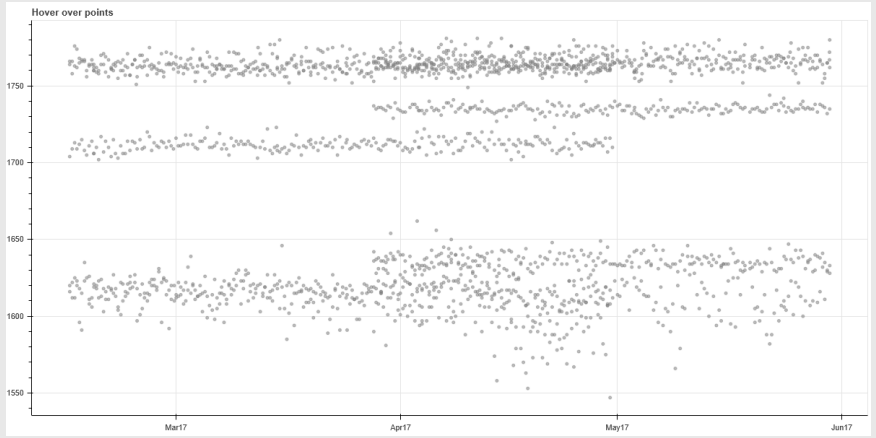
Tableau workbook “qc_data_dashboard.twb”

Results Continued

Bokeh dashboard 2 “qc_over_time.html”

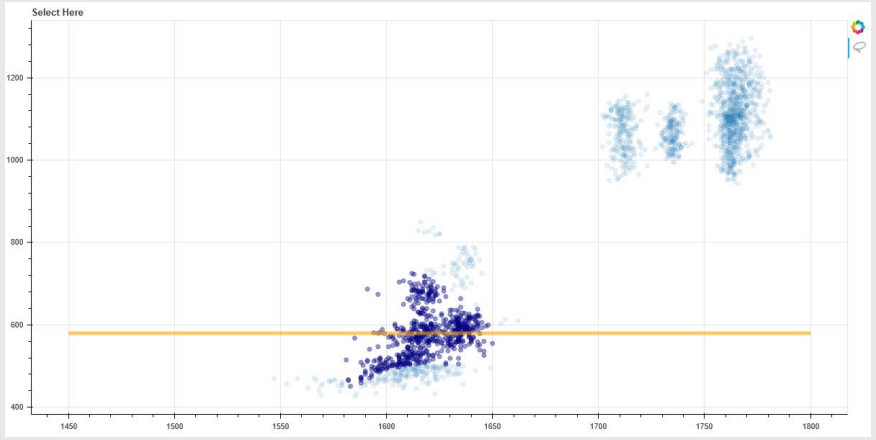


Bokeh dashboard 3 “wx_select.html”



Results Continued

Bokeh dashboard 4 “wx_wy.html”



Future Expansions

- Decreasing the overhead costs for diagnosing failing devices
- Scalable real time data ingestion, to complement the streaming visualization
- Leveraging a distributed architecture
- More efficient data transformations and modeling
- More robust architecture

Appendix: Dataflow Detailed