|  |
| --- |
| Urban Data Labs |
| **Real-Time Anomaly Detection for Building Sensors** |
| Master of Data Science Capstone |

|  |
| --- |
| Nathan Smith, Mitch Harris, Ryan Koenig  5-9-2021 |

# Executive Summary

Table of Contents

[Executive Summary 1](#_Toc73098790)

[1 – Introduction 3](#_Toc73098791)

[1.1 – Background 3](#_Toc73098792)

[1.2 – Project Scope and Objective 3](#_Toc73098793)

[2 – Study Data 4](#_Toc73098794)

[2.1 – SkySpark Database 4](#_Toc73098795)

[2.2 – Urban Data Lab InfluxDB 4](#_Toc73098796)

[2.3 – Campus Energy Centre Sensors 4](#_Toc73098797)

[2.4 – Data Anomalies 4](#_Toc73098798)

[3 – Anomaly Detection Framework 5](#_Toc73098799)

[3.1 – Options Considered 5](#_Toc73098800)

[3.2 – Framework Selected 5](#_Toc73098801)

[4 – Anomaly Detection Model 6](#_Toc73098802)

[4.1 – Model Approaches 6](#_Toc73098803)

[4.2 – Model Selected 6](#_Toc73098804)

[5 – Anomaly Detection Model Testing 7](#_Toc73098805)

[5.1 – Model Testing Approach 7](#_Toc73098806)

[5.1.1 – Manual Anomaly Labelling 7](#_Toc73098807)

[5.1.2 – Performance Criteria 7](#_Toc73098808)

[5.1 – Sensors Tested 7](#_Toc73098809)

[5.2 – Results 7](#_Toc73098810)

[5.3 – Model Variations 7](#_Toc73098811)

[5.4 – Recommended Model 7](#_Toc73098812)

[6 – Recommendations for Future Work 8](#_Toc73098813)

[6.1 – Anomaly Detection Framework 8](#_Toc73098814)

[6.2 – Anomaly Detection Model 8](#_Toc73098815)

[6.3 – Performance Monitoring 8](#_Toc73098816)

[7 – Conclusion 9](#_Toc73098817)

[8 - References 10](#_Toc73098818)

[Appendix A – Whatever 0](#_Toc73098819)

# 1 – Introduction

## 1.1 – Background

Describe UDL background and general database setup (SkySpark etc), some can likely be copied from proposal.

## 1.2 – Project Scope and Objective

Describe the scope of the project clearly discussing the project questions/goals

# 2 – Study Data

General section introduction of the following sections

## 2.1 – SkySpark Database

Overview of SkySpark

## 2.2 – Urban Data Lab InfluxDB

Overview of InfluxDB, discussing current state and setup

## 2.3 – Campus Energy Centre Sensors

Discussion of boiler sensors, grouping, some of the EDA stuff to go in here

## 2.4 – Data Anomalies

Discussion of what appears to show up as anomalies in the data, discussing how some of it’s different

# 3 – Anomaly Detection Framework

## 3.1 – Options Considered

Discuss the options that were looked at for how the real-time system could work with InfluxDB

## 3.2 – Framework Selected

Discuss what we selected and why and potentially note that what we used could potentially be changed to be faster

Need the diagram of the read/write/detect stuff as well as a diagram showing how we’re storing data in InfluxDB

Need to comment on how we’re storing model parameters

# 4 – Anomaly Detection Model

## 4.1 – Model Approaches

Lit review of various approaches goes in here with a discussion of what was looked at and considered (including commentary on LSTM)

## 4.2 – Model Selected

Discuss the approach we selected and the LSTM model in more detail with the cleaning pipeline (model diagram would show-up here), also potentially need a diagram of the LSTM we’re using

# 5 – Anomaly Detection Model Testing

## 5.1 – Model Testing Approach

General discussion of our approach

### 5.1.1 – Manual Anomaly Labelling

Discuss the manual labelling approach, what is achieves, what it doesn’t

### 5.1.2 – Performance Criteria

Discussion on the fact that we’re using an unsupervised approach and we’re generally using comparison with our labelling and visual assessment of datasets that weren’t labelled – need to note that the model may actually be better than our visual assessment

## 5.1 – Sensors Tested

This section describes the sensors we looked at – showing the datasets with some specific anomalies

Show the table of tests

## 5.2 – Results

Show the results here

## 5.3 – Model Variations

Anything we want to discuss in terms of other stuff we tried

## 5.4 – Recommended Model

Recommendation on the final model – default parameters etc

# 6 – Dashboard and Notifications System

## 6.1 – Dashboard

## 6.2 Notifications System

# 7 – Recommendations for Future Work

## 7.1 – Anomaly Detection Framework

What we did implement vs what would be better

## 7.2 – Anomaly Detection Model

What improvements could be looked at for the model

## 7.3 – Performance Monitoring

Recommendations on how the system should be monitored

# 8 – Conclusion

Final conclusion statements

# 9 - References

# Appendix A – Whatever