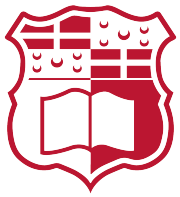

Anomaly Detection Using Machine Learning Techniques for Beam Injections from the SPS to the LHC at CERN

MARC FERRIGGI

Supervised by Dr. Gianluca Valentino

Department of Computer Science

Faculty of ICT

University of Malta

May, 2019

A FYP submitted in partial fulfilment of the requirements for the degree of B.Sc. (Hons.) Computing Science AND Statistics and Operations Research.

Statement of Originality

I, the undersigned, declare that this is my own work unless where otherwise acknowledged and referenced.

Candidate Marc Ferriggi

Signed _____

Date February 21, 2019

Acknowledgements

Abstract

Contents

1 Introduction **1**

2 Background and Literature Review **2**

2.1 Introduction 2

2.2 Understanding the Problem Domain 2

2.3 Feature Selection and Reduction Techniques 3

2.4 Unsupervised Anomaly Detection Techniques 3

References **4**

List of Figures

List of Tables

List of Abbreviations

CERN European Organization for Nuclear Research
IQC Injection Quality Check 2
LHC Large Hadron Collider
LS Logging Service 2
PCA Principal Component Analysis
SPS Super Proton Synchrotron

1

Introduction

2

Background and Literature Review

2.1 Introduction

2.2 Understanding the Problem Domain

The LHC is filled to the required maximum intensity by injecting electron bunch trains from the SPS through a transfer line using kicker magnets [CITE HERE]. This is a challenging task given the high energy of the beam, the very small apertures and the delivery precision's tight tolerances, thus multiple sensors are installed around the CERN particle accelerator complex [CITE HERE] which gather readings and data that can be used to check the quality of the injected beam. For this particular study, the sensors around the injection from the SPS to the LHC will be of particular interest. This data is stored using CERN's Logging Service (LS) [CITE HERE]. While many studies have been made using this logged data and lots of statistical tests have been done with regards to injection quality checks for the LHC (such as [CITE HERE] and [CITE HERE]), no literature was uncovered where researchers used unsupervised machine learning methods to analyse this data.

The Injection Quality Check (IQC) software currently installed has a set of hard-coded rules for detecting anomalies in the SPS-LHC injection [CITE HERE], however there are documented cases in the past where situations occurred which were outside the originally foreseen rules and were therefore not caught as anomalies.

2.3 Feature Selection and Reduction Techniques

PCA uses statistical and mathematical techniques to reduce the dimension of large data sets, thus allowing a large data set to be interpreted in less variables called principal components (Richardson [2009]). This non-parametric method can be used as a means of revealing the simplified structures underlying complex datasets with minimal effort. The fact that this technique is non-parametric gives it the advantage that each result is unique and only dependent on the provided data set since no parameter tweaking is required (Shlens [2014]) however, this is also a weakness of this technique as there is no way of exploiting prior expert knowledge on the data set.

2.4 Unsupervised Anomaly Detection Techniques

References

Mark Richardson. Principal Component Analysis, May 2009. Class Lecture.

Jonathon Shlens. A Tutorial on Principal Component Analysis, April 2014.