REAL-TIME PREDITIVE ANALYTICS ON OVERALL EQUIPMENT EFFECTIVENESS (OEE)

WONG RUI ZHEN

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

LAMPIRAN B: BORANG PENGESAHAN STATUS TESIS

Contoh halaman pengesahan

BORANG PENGESAHAN STATUS TESIS*

JUDUL: REAL - TIME PREDICTIVE ANALYTICS ON OVERALL EQUIPMENT EFFECTIVENESS (OEE)

SESI PENGAJIAN: 2016 / 2017

Saya WONG RUI ZHEN
(HURUF BESAR)

mengaku membenarkan tesis (PSM/Sarjana/Doktor Falsafah) ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. ** Sila tandakan (/)

______	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
______	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
✓	TIDAK TERHAD	

(TANDATANGAN PENULIS)

Alamat tetap: 1, Jalan Perdana 2/2, Taman Puchong Perdana, 47150 Puchong, Selangor.

Tarikh: 25. 8. 2017

(TANDATANGAN PENYELIA)

CHOO YUN HUOY

Nama Penyelia

Tarikh: 25 Aug 2017

CATATAN: * Tesis dimaksudkan sebagai Laporan Akhir Projek Sarjana Muda (PSM)
** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa.

REAL-TIME PREDITIVE ANALYTICS ON OVERALL EQUIPMENT EFFECTIVENESS (OEE)

WONG RUI ZHEN

This report is submitted in partial fulfilment of the requirements for the Bachelor of Computer Science (Artificial Intelligence)

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
UNIVERSITI TEKNIKAL MALAYSIA MELAKA
2017

DECLARATION

I hereby declare that this project report entitled

REAL-TIME PREDITIVE ANALYTICS ON OVERALL EQUIPMENT EFFECTIVENESS (OEE)

is written by me and is my own effort and that no part has been plagiarized without citations.

STUDENT : ______________________ Date : 25.8.2017
(WONG RUI ZHEN)

I hereby declare that I have read this project report and found this project report is sufficient in terms of the scope and quality for the award of Bachelor of Computer Science (Artificial Intelligence) with Honours.

SUPERVISOR : ______________________ Date : 25 Aug 2017
(ASSOC. PROF DR CHOO YUN HUOY)

DEDICATION

To my beloved parents, your love and support are my greatest inspiration upon accomplish the project.

To my dear friends, thanks to your motivation and support throughout this project.

To my dearest supervisor, Dr Choo Yun Huoy for being responsible, helpful and always spending your precious time to supervise me.

ACKNOWLEDGEMENTS

First of all, my greatest gratitude to my supervisor, Assoc. Prof Dr Choo Yun Huoy for the opportunity on the project topic and the assistance in the completion of this project.

Next, I wish to take this opportunity to express my sincere thanks to everyone for any moral support or physical support that given to me.

ABSTRACT

Overall equipment effectiveness (OEE) is a standard of performance measure on equipment productivity especially in manufacturing process. This metrics will helps the company to evaluate the performance of machines by identifying the underlying losses and operational effectiveness. Anomaly occurs when an instance deviates from a normal behaviour, in this case we meant when an OEE spike is being detected. A real time anomaly detection analytics is important for a business because some anomalies required immediate action, this will helps the management level for better and faster decision making especially on manufacturing scheduling. The methodology in this project involved 6 parts, preliminary studies, data preparation, attribute selection, model development, comparison analysis, and results validation. First, preliminary studies were done by reviewing literatures and self-questioning on the upcoming problems. Next, real industrial data is collected, compiled and undergo discretization to prepare the data for model training and testing process. After that, attribute selection was done by intuition, by expertise selection and wrapper methods such as Recursive Feature elimination (RFE) to reduce the dimensionality of dataset. In the model development process, Support Vector Regression (SVR), Linear Regression (LR) and Regression Tree (RT) classifiers were selected to develop a model for OEE value prediction and OEE spike prediction. The results shows that LR model by RFE selected attributes was performing best among others because it was able to achieved Root Mean Squared Error (RMSE) of 0.0013 and accuracy of 0.9892 when doing OEE value prediction and OEE spike classification respectively. In conclusion, the proposed model was able to predict OEE value and detect OEE spike (undesirable OEE drop), but the model can only tested with limited data (5 shifts) due to data constraints. Furthermore, this model is suggested to embed in smart manufacturing data analytics dashboard for operational monitoring purpose. Further work is suggested to collect actual real-time industrial data to test on the robustness of the proposed model. Last but not least, if the suitable dataset is successfully obtained, the next step is suggested propose a predictive analytic model where the model will be able to forecast future (one day ahead) OEE spike in real time, and provide actionable insights to notify user based on the causes of the spike occurrence.

ABSTRACK

Keberkesanan peralatan keseluruhan (OEE) adalah ukuran prestasi yang standard dalam produktiviti peralatan terutamanya proses pembuatan. Metrik ini akan membantu syarikat menilai prestasi mesin dengan mengenal pasti kerugian asas dan keberkesanan operasi. Anomali berlaku apabila suatu kejadian menyimpang dari tingkah laku biasa, dalam kes ini kita maksudkan apabila spek OEE dikesan. Analitik pengesanan anomali masa sebenar adalah penting untuk perniagaan kerana beberapa anomali memerlukan tindakan serta-merta, ini akan membantu tahap pengurusan untuk mendapat keputusan yang lebih baik dan cepat terutama pada pembuatan penjadualan. Metodologi dalam projek ini melibatkan 6 bahagian, kajian awal, penyediaan data, pemilihan atribut, pembangunan model, analisis perbandingan, dan pengesahan hasil. Pertama, kajian awal dilakukan dengan mengkaji literatur dan menyoal diri tentang masalah yang akan datang. Seterusnya, data perindustrian sebenar dikumpulkan, disusun dan menjalani pembicaraan untuk menyediakan data untuk latihan dan proses ujian model. Selepas itu, pemilihan atribut dilakukan dengan intuisi, melalui pemilihan kepakaran dan kaedah pembungkus seperti penghapusan Ciri Rekursif (RFE) untuk mengurangkan dimensi dataset. Dalam proses pembangunan model, Regresi Vektor Sokongan (SVR), Klasifikasi Regresi Linear (LR) dan Regresi Tree (RT) dipilih untuk membangunkan model untuk ramalan nilai OEE dan ramalan spek OEE. Hasilnya menunjukkan bahawa model LR oleh atribut terpilih RFE telah melakukan yang terbaik antara lain kerana ia dapat mencapai kesilapan akar min kesilapan (RMSE) 0.0013 dan ketepatan 0.9892 ketika melakukan ramalan nilai OEE dan klasifikasi spek OEE masing-masing. Kesimpulannya, model yang dicadangkan dapat meramalkan nilai OEE dan mengesan OEE spike (keturunan OEE yang tidak diingini), tetapi model hanya boleh diuji dengan data terhad (5 shift) disebabkan oleh kekangan data. Tambahan lagi, model ini dicadangkan untuk membenamkan dalam papan pemuka analisis data pintar untuk tujuan pemantauan operasi. Kerja lebih lanjut disarankan untuk mengumpul data perindustrian sebenar untuk menguji keteguhan model yang dicadangkan. Akhir sekali, sekiranya dataset yang sesuai berjaya diperoleh, langkah seterusnya dalam projek ini akan mencadangkan model analitik ramalan di mana model akan dapat meramalkan masa depan (satu hari ke depan) OEE spike dalam masa nyata, dan memberikan wawasan yang dapat dilihat Untuk memberitahu pengguna berdasarkan sebab-sebab kejadian spek.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AD	–	Anderson Darling
ANN	–	Artificial Neural Network
ANOVA	–	Analysis of Variance
BI	–	Business Intelligent
CDA	–	Confirmatory Data Analysis
DT	–	Decision Tree
EDA	–	Exploratory Data Analysis
FMEA	–	Failure Mode and Effect Analysis
IDE	–	Integrated Development Environment
IoT	–	Internet of Things
k-NN	–	k-Nearest Neighbour
KPI	–	Key Performance Indicator
MES	–	Manufacturing Execution System
MESA	–	Manufacturing Enterprise Solutions Association
ML	–	Machine Learning
MOM	–	Manufacturing Operations Management
OEE	–	Overall Equipment Effectiveness
RMSE	–	Root Mean Square Error
RT	–	Regression Tree
SMART	–	Specific, Measurable, Actionable, Realistic, Time-Based
SOM	–	Self-Organized Map
SVC	–	Support Vector Classification
SVM	–	Support Vector Machine
TPM	–	Total Productive Maintenance

CHAPTER I

INTRODUCTION

1.1 Introduction

In recent years, smart manufacturing has been discussed among the industrial sectors especially in production and manufacturing industry. This term was actually came from the fourth industrial revolution where the Germany government first revive the "Industrie 4.0" in year 2011 which promotes the modern communication technology, Internet of Things (IoT) via cyber physical systems and cloud computing. The process where the current physical system in the factory that communicate each other and form a network called Internet of Things, this network will monitor by the cyber physical system and cooperate with human in real time to do decentralized decision making. To be simplified, the computers and automations in the industry will all be connected together to form an internet, and remotely control by the decentralized computer operator that equipped with artificial intelligence and machine learning algorithms, making the computer decision as autonomous as possible, this will obviously reduce the efforts from the human operators.

The process that introduced by Industrial 4.0 is smart manufacturing, in order for a current manufacturing factory to achieve the latest industrial revolution, a lot of manpower is required to implement the system and a group of investors is needed to invest heavily to the project because of the new technologies that involved in the development, either side is indispensable. In realising the smart manufacturing process, big data and data analytics are playing important role in helping the decision making in the manufacturing process. This is because the analytic results were able to help in early detection of machine defects, production failures, forecasting future performances and etc. This will generates huge competitive values to the company such as helps in early prevention, operation scheduling and increase productivity of the manufacturing process. The intelligent from the raw data either from machines, logs or processes will help to illustrate a more informed and actionable decision, this will let the decision makers to have trust from the data and have higher confident in making a decision. Also, this decision making in a data-driven way will helps the process of manufacturing plants to operate better.

In today's manufacturing industrial, the management is bombarded with data and information that generated from the business line, this makes the management to heavily rely on some Business Intelligent (BI) software to track the performance of their organization, to improve visibility and efficiency of the business process, and to gain competitive intelligence and actionable information from the raw data in real time. It is important to understand the company requirement and objective to select the manufacturing metrics indicators to achieve the goal, the business objective is basically following the SMART concept – Specific, Measurable, Actionable, Realistic, and Time-Based. The Manufacturing Enterprise Solutions Association (MESA) organization has conducted research and 28 manufacturing metrics were identified to be the metrics that matter most in manufacturing industry, and the list of each metrics is shown in the table below.

Table 1.1 List of Manufacturing Metrics

Improving Customer Experience & Responsiveness	
1	On-Time Delivery to Commit
2	Manufacturing Cycle Time
3	Time to Make Changeovers
Improving Quality	
4	Yield
5	Customer Rejects/Return Material Authorizations/Returns
6	Supplier's Quality Incoming
Improving Efficiency	
7	Throughput
8	Capacity Utilization
9	Overall Equipment Effectiveness (OEE)
10	Schedule or Production Attainment
Reducing Inventory	
11	WIP Inventory/Turns
Ensuring Compliance	
12	Reportable Health and Safety Incidents
13	Reportable Environmental Incidents
14	Number of Non-Compliance Events / Year
Reducing Maintenance	
15	Percentage Planned vs. Emergency Maintenance Work Orders
16	Downtime in Proportion to Operating Time
Increasing Flexibility & Innovation	
17	Rate of New Product Introduction
18	Engineering Change Order Cycle Time
Reducing Costs & Increasing Profitability	
19	Total Manufacturing Cost per Unit Excluding Materials
20	Manufacturing Cost as a Percentage of Revenue
21	Net Operating Profit
22	Productivity in Revenue per Employee
23	Average Unit Contribution Margin

24	Return on Assets / Return on Net Assets
25	Energy Cost per Unit
26	Cash-to-Cash Cycle Time
27	EBITDA
28	Customer Fill Rate/On-Time delivery/Perfect Order Percentage

In our case, the requirement is to verify the Overall Equipment Effective (OEE) values, so we will use OEE as the manufacturing metrics in this project. Nakajima (1988) introduced Total Productive Maintenance (TPM), the metric in TPM is called Overall Equipment Effectiveness (OEE). OEE has three generic elements, Availability, Performance Efficiency, and Rate of Quality. The classic definition of Overall Equipment Effectiveness is shown below where A=Availability, PE=Performance Efficiency and RQ=Rate of Quality.

$$OEE = (A) * (PE) * (RQ)$$

OEE management shows an effective tool to improve manufacturing performance, a micro factory modelling process that defines workstation capabilities and tool theoretical outputs is required in OEE management process. Also, OEE management provides insight into the TPM losses, which identifies the improvement opportunities available. OEE is one of the Key Performance Indicator (KPI) that measure the overall throughput performance of production and test equipment in manufacturing, it used to monitor and understand the changes of equipment productivity by collecting the operating data. Not only that, OEE also important for other purposes such as to identify losses, identify available resources, improve the productivity of manufacturing equipment and provide support in making decision. The table below shows the six big losses which are the underlying losses from the OEE value. In this project, the OEE loss will be considered as there is spike occurs on the OEE metrics inside.

Table 1.2 Six Big Losses

OEE Metrics	Six Big Losses	Example Loss
Availability	Breakdown	Equipment failure Unexpected machine breakdown General Maintenance
	Setup and adjustment	Machine shortage Machine warm-up
Performance	Speed losses	Speed reduced during operation Operation efficiency Slow cycle
	Small stops and idling	Jams, Blocks, Obstructions Cleaning
Quality	Production defects	Product damage Scraps Repair
	Start-up rejects	Start-up product damage Improper Assembly

Since real time feedback is also a significant issue in Business Intelligent, because organizations need actionable insights faster than ever before to stays competitive, reduce risks, and capitalize on time-sensitive opportunities. Basically, real time analytics is analytics on live operational database and up to date to the current moment, so that the data can discover insights faster and act immediately upon problem occurs. Thus, we will considered real time data streaming in our case as well. Based on the above scenarios, in order to provide useful actionable insights to the company operation process, the obtained data have to be processed and run through some data analytic algorithms to generate useful and decisive information. Data analytics have 2 types of methodology, one is Confirmatory Data Analysis (CDA) and the other one is Exploratory Data Analysis (EDA). CDA do conventional analytical modelling by using traditional statistical metrics like confidences, significance and hypothesis to draw an estimated conclusion or to evaluate the hypothesis, examples such as Regression analysis, Analysis of Variance, and