A HYBRID FORECASTING MODELS USING LEAST SQUARE SUPPORT VECTOR MACHINE (LSSVM) AND GENETIC ALGORITHM (GA)

THIAN SWEE LIANG

BACHELOR OF ELECTRICAL ENGINEERING WITH HONOURS
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2021

A HYBRID FORECASTING MODELS USING LEAST SQUARE SUPPORT VECTOR MACHINE (LSSVM) AND GENETIC ALGORITHM (GA)

THIAN SWEE LIANG

A report submitted
in partial fulfilment of the requirements for the degree of
Bachelor of Electrical Engineering with Honours

Faculty of Electrical Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2021

DECLARATION

I declare that this thesis entitled "A HYBRID FORECASTING MODELS USING LEAST SQUARE SUPPORT VECTOR MACHINE (LSSVM) AND GENETIC ALGORITHM (GA) is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in the candidature of any other degree.

Signature : .....

Name : THIAN SWEE LIANG

Date : 28/6/2021

APPROVAL

I hereby declare that I have checked this report entitled "A Hybrid Forecasting Models Using Least Square Support Vector Machine (Lssvm) And Genetic Algorithm (GA)" and in my opinion, this thesis it complies with the partial fulfilment for awarding the award of the degree of Bachelor of Electrical Engineering with Honours

Signature : .....

Supervisor Name : Dr. Intan Azmira Binti Wan Abdul Razak

Date : 5/7/2021

DEDICATIONS

To my beloved mother and father

ACKNOWLEDGEMENTS

In preparing the report, I received a lot of help from my friends and supervisor. They have contributed to help me for more understanding about my project. I would like to express my specials thanks to Dr Intan Azmira Binti Wan Abdul Razak act as my project supervisor who give me a lot of technical and knowledge support. I would like to convey my heartfelt thanks to my academic advisor Dr. Azrita Binti Alias for the guidance and advice. Without their support, the project would not to presented here.

Next, I would like to express my greatest thanks to my parents. Without their financial support, understanding, patience, and love, this report would not to exist here.

Lastly, I would like to thank my classmates and my friend for their opinions and their views during the process of completing the project.

ABSTRACT

The project is to develop electricity price forecasting models. Forecast price of electricity become an important issue in the electricity power generation field. The prediction price of electricity price can help the electricity provider decide their electricity price strategy. It can also help the electricity consumer know their electricity price more accurately and estimate their daily electricity usage a day ahead of 24-hour cost. The challenge face on forecasting prices is more on the accuracy and the efficiency of the forecasting model. Therefore, some researchers have developed a forecasting model by using the algorithm to optimize the forecast output. Hence, a combination of the Least Square Support Vector Machine (LSSVM) and Genetic Algorithm (GA) forecast model was developed in this project. The parameter used in the project was the hourly Ontario electricity price (HOEP) and Ontario demand in years 2010 provided by The Independent Electricity System Operator (IESO) website.This type of hybrid forecasting model will be tested on the Ontario power market, which is report as the most volatile market worldwide, to compare with the existing model and provide better accuracy of forecast price and performance. The hybrid forecasting model has completed developed come out with the forcasting results and performance on the hybrid forecasting model itself.

ABSTRAK

Projek ini adalah membina sesuatu model ramalan untuk meramalkan harga penggunaan elektik. Ramalan harga elektrik merupakan satu isu yang amat penting dalam bidang penjanaan kuasa. Oleh itu, mempunyai ramalan harga elektrik akan bermanfaat kepada industri penjanaan kuasa dapatkan info penggunaan kuasa elektrik masa depan, ia juga akan membantu pengguna elektrik untuk mengetahui harga penggunanan elektrik yang lebih tepat dan lanjut bagi masa yang depan seperti mengetahui ramalan harga elektrik untuk hari keesokan. Walau bagaimanapun, cabaran atas model ramalan yang telah dicipta adalah ketepatan dan kecekapan model tersebut. Oleh itu, sesetengah penyelidik menggunakan pelbagai algritma untuk optimumkan hasil yang dapat dalam model ramalan yang diciptakan. Dengan ini, model ramalan yang mempunyai gabungan antara algoritma 'Least Square Support Vector Machine (LSSVM)' dan algoritma genetik akan dihasikkan dalam projek ini. Data yang digunakan dalam projek ini adalah harga elektrik setiap jam Ontario pada tahun 2010 yang dikongsikan oleh The Independent Electricity System Operato (IESO) di halaman web rasmi mereka. Tujuan kombinasi dua algoritma ini akan mengoptimumkan ketepatan dan kecekapan hasilnya, dan ianya akan diguna kaji di tempat Ontario dengan menggunakan data kuasa elektrik di Ontario. Sebagai pengetahuan Ontario mempunyai gelaran antara harga elektrik yang paling tidak stabil di dunia, dengan mengambil kesempatan untuk mengkaji ketepatan keluaran dari projek yang sedia ada. Model ramalan gabungan antara algoritma 'Least Square Support Vector Machine (LSSVM)' dan 'algoritma genetik' berjaya dibina dengan mengeluarkan keluaran anggaran yang diteteapkan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

IESO	-	The Independent Electricity System Operator
LSSVM	-	Least Square Support Vector Machine
SVM	-	Support Vector Machine
MAE	-	Means Absolute Error
MAPE	-	Means Absolute Percentage Error
GA	-	Genetic Algorithm
FA	-	Fuzzy ARTMAP
FF	-	Firefly Algorithm
WT		Wavelet Transform

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background

Electricity is a flow of charge, and it was one of the essential daily needs for humans nowadays. Based on the data and statistics done by International Energy Agency, world electricity consumption for the year 2018 was archived over 24000 TW/h, the consumption of electricity still increasing every year.

Since the early of years 1990, Europe, Canada, North America, and Australia started introducing a competitive electricity market due to deregulation in the countries. By transforming competitive electricity market launching, the electricity was traded based on two common categories, which are sold under market rules spot and derivative contracts. This type of trade was made a most significant change on the traditional monopolistic and government-controlled power sectors.

Price forecasting becomes an important strategy, especially for the power generation company. Future electricity price forecasting can help the power generation company optimize the output and gain a good profit. The same as the electrical power consumption can also benefit from an electricity price forecast to control their energy consumption. At the same time, it will become more manageable for the consumer to choose which electricity supplier they need to choose.

Some researchers have successfully developed a forecasting model with a Support Vector Machine or hybrid forecasting model to predict the electricity price. However, forecasting electricity price was more challenging than load or demand because some factors can directly affect the cost of electricity, such as coal, fuel price, etc. In this project, the Least Square Support Vector Machine with optimization of the Genetic algorithm will be used. The project will be developed to improve the accuracy and the performance compared with the previous model.

1.2 Problem Statement

Nowadays, several forecasting models with different algorithms or methods exist in the market, such as straight-line method, moving line method, simple linear regression method, multiple regression method, etc.

The way to developing a good forecasting model will easier the user during the electricity price forecasting. The forecasting model should need a stable or friendly user platform for development to ensure that the forecasting model can be complete when running forecasting work with good forecasting results.

The multiple regression analysis methods with the support of LSSVM was one of the forecasting models available in the market. The result comes out with the forecasting model become a challenge to get an accurate forecasting result. The forecasting model also can improve its accuracy by supporting other optimization methods such as the Genetic algorithm.

1.3 Objective

The aim of this project to:

i. Develop hybrid Forecasting model (LSSVM – GA) with MATLAB.
ii. Optimize LSSVM parameters (gamma and sigma) by Genetic Algorithm.

1.4 Project Scope

i. The implementation of LSSVM technique as main forecast engine.

ii. The input used for the forecast is the historical price and demand of electricity per hours.

iii. The output forecast for day-ahead 24-hour prices.

iv. The objective function of the model is Mean Absolute Percentage Error (MAPE) and MAE; where minimum error should be achieved to obtain good predictive accuracy.

v. The optimization process is held by the Genetic Algorithm (GA) to optimize the parameters of LSSVM and reduce the number of input for prediction.

vi. The model developed is tested in the Ontario electricity market, which is one of the most volatile energy markets in the world

1.5 Thesis Outline

The thesis is organized into five chapters, which is:

Chapter 1 explained the background of the project and followed with the problem statement. In the background of the project, the overview of the project was discussed. After that, the problem statement was to state the problem more clearly and followed with the objective and scope of the project.

Chapter 2 will describe the literature review LSSVM and Genetic algorithm, which contained thesis, book or journal information related to the project needed to develop.

Chapter 3 discussed methodology for the LSSVM, Training data selection, testing data selection, Mean Absolute Error (MAE) and Means Absolute Percentage Error (MAPE).

Chapter 4 discuss the data after the forecasting model made the prediction and comparisons between the LSSVM, LSSVM-GA and other existing forecasting models.

Chapter 5 will discuss the conclusion and the recommendation based on the project.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter will describe about the literature review on Support Vector Machine, Least Square Support Vector Machine (LSSVM), Genetic Algorithm. The information on this chapter come from various source such as thesis, books, journal, article and other academic websites. All the information were cited clearly and benefit on the road of complete the project.

2.2 Support Vector Machine

Support Vector Machine is one of the machines learning techniques and it was introduced by Vapnik [1]. The Support Vector Machine (SVM) commonly used for classifying or separating of two data. As an example using in the forecasting model to separating the demand and the price of electricity.

A simple example for easier understanding, when it had two of the data as shown as the figure 2.2.1 below, the data in the graph was mixing on the graph. In this case, separation step will be done by the SVM. SVM will separate the different data with the line or can called hyperplane as show in the **Figure 2-1**.

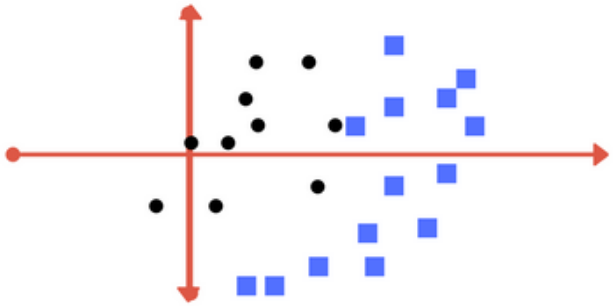

Figure 2-1: Data Before Classification

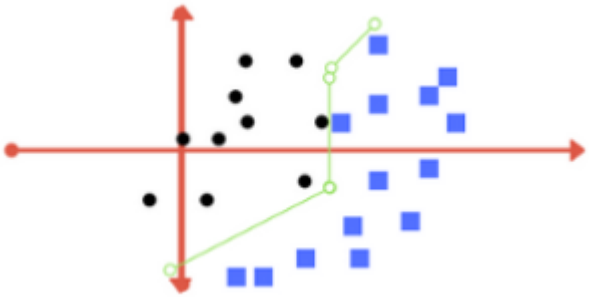

Figure 2-2: Data After Classification

When the operation of separating data done by SVM, there was few parameters need to be considered in our forecasting project especially the Gamma and Sigma. Gamma can be understanding as the how far or close of the separation line needed to consider on the separating line calculation. It can be known as the how long of our forecasting data range needed to be considered in the forecasting model to produce our output needed. Sigma also an important parameter guided forecasting model to archived an accurate result. Sigma can be explained as the **Figure 2-3** shown the different of the high sigma(left) and low sigma(right) of the data separating.

Figure 2-3:Different Of The High Sigma(Left) And Low Sigma(Right) Of The Data Separating

Lastly, the more important role of the SVM in the project was to produce a good margin during the forecasting process. A good margin of SVM needed to considered the equability of two data, closer gamma parameter and low sigma parameter as well.

2.3 Least Square Support Vector Machine (LSSVM)

LSSVM is the least-squares of SVM. It was first proposed by J.A.K. Suykens and J. Vandewalle as a classifier in 1999. [2] [3] Unlike the inequality constraints introduced in the standard SVM, LSSVM proposed equality constraints in the formulation. LSSVM solves a system of linear equations instead of quadratic programming (QP) problem that improves the computational speed. The solution is transformed from solving a quadratic program to a set of linear equations known as the linear Karush–Kuhn–Tucker (KKT) systems. [2] [4]

LSSVM become an efficient method to become a choice to apply on the forecasting model. Refer to the previous researcher, the LSSVM forecasting model is better than the SVM forecasting model during the training process. [5]

For improving the performance, a suitable LSSVM method in forecasting data can be enhanced as combined with the other type of algorithm such as the Genetic Algorithm.

2.4 Genetic Algorithm (GA)

Genetic algorithm has reflected the process as a natural selection where the fittnest individual are selected for evolution process via reproduction. The genetic algorithm has 5 main phase on in process which stand for initial population, fitness function, selection, crossover and mutation. [6] [7]

The every single phase has their own characteristic, let start with the initial population phase. Initial population phase was a beginning step of transformation in this phase, it will have a set of individual (problem) then follow with a set of genes will joined into a string to form a chromosome. In other word means that a problem(individual) will be solved by a solution (Chromosome).

The next phase was represent as fitness function. Fitness function was a pahase to determine how to fit an individual. It means that these phases has the ability which able the an individual to compete with the other individual. For an example, genes A had more taller genetic compare genes B.

Next, Selection phase was represent the select step on the fitness individual to fulfil the enquiry. Same as the meaning on a situation had a lot of the individual, researcher will select the individual which had more closer the fitness score for reproduction.

In the transformation of the genetic algorithm, crossover also had happened on it. Crossover means the random chosen at genes. Lasty was the mutation. Mutation had low random probability happened because it will exchange the genes.

As the phases of genetic algorithm explain before, the concept on GA will act as 2nd step of optimization of data to increase the accuracy of the forecast mode.

CHAPTER 3

METHODOLOGY

3.1 Introduction

The design of the hybrid forecasting model by MATLAB will be explained in detail. The project flowchart also will present in this chapter. The Support Vector Machine (SVM), Least Square Support Vector Machine (LSSVM), and Genetic Algorithm (GA) forecast process will be explain more detail in this chapter. The mean absolute error (MAE) and Mean Absolute Percentage Error (MAPE) calculation method will show in this chapter to observes performance of designing forecast model.