PHOTOVOLTAIC (PV) SYSTEM OUTPUT POWER FORECASTING USING SUPPORT VECTOR MACHINES (SVM) TECHNIQUE

WONG WAI LEONG

BACHELOR OF ELECTRICAL ENGINEERING WITH HONORS
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2019

PHOTOVOLTAIC (PV) SYSTEM OUTPUT POWER FORECASTING USING SUPPORT VECTOR MACHINES (SVM) TECHNIQUE”

WONG WAI LEONG

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

Faculty of Electrical Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2019

DECLARATION

I declare that this thesis entitled "PHOTOVOLTAIC (PV) SYSTEM OUTPUT POWER FORECASTING USING SUPPORT VECTOR MACHINES (SVM) TECHNIQUE" 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 candidature of any other degree.

Signature : --

Name : --

Date : --

APPROVAL

I hereby declare that I have checked this report entitled “PHOTOVOLTAIC (PV) SYSTEM OUTPUT POWER FORECASTING USING SUPPORT VECTOR MACHINES (SVM) TECHNIQUE” and in my opinion, this thesis it complies the partial fulfillment for awarding the award of the degree of Bachelor of Electrical Engineering with Honours

Signature	:	....
Supervisor Name	:	....
Date	:	....

DEDICATIONS

To my beloved mother and father

ACKNOWLEDGEMENTS

My Final Year Project is in smooth progress with the help and contribution from many of others. Therefore, I would like to show my appreciation and extend my sincere thanks to all of them. First of all, I would like to thank the Almighty God for granting me health and knowledge so that I could have finished Final Year Project.

Next, I would like to express my very great appreciation to my supervisor, En. Kyairul Azmi Bin Baharin for the support and guidance which can help me to finish my project on time. I very appreciate that he taught me a lot of knowledge about my project and lead me to the path of success.

Furthermore, I would like to offer my special thanks to all the lecturers for willing to support and guide me by teaching me a lot of knowledge and techniques to overcome the challenge and complete my Final Year Project.

Last but not least, I wish to thank my family and friends that support me and lend me a hand when I seeking for help. I appreciate that there are family and many friends who always got my back throughout the whole progress of Final Year Project.

ABSTRACT

The use of Solar Photovoltaic (PV) system for power generation have expanded rapidly for the past few years. However, the growth of the solar PV system is also causing problems for the management of the power distribution as the operator have to always maintain the stability of the power grid between power generation and power distribution. Therefore, solar power output forecasting have become an important task to focus on to overcome the problems of using solar PV system for power generation. A solar power output prediction model is developed in this project to predict the day ahead hourly power output by using the Support Vector Machines (SVM) method. The prediction model is developed based on the data and module technology of the Solar Lab of Faculty of Electrical Engineering (FKE) in University Teknikal Malaysia Melaka (UTeM). The prediction model is designed by training the prediction model using local data with regression learner application in MATLAB software version R2017b. The results indicate that using SVM model to forecast solar power output is valid and the accuracy of the prediction is satisfied. The predictor variables used to trained the predictive model is analyzed. Irradiance and Module Temperature are the most dominant variables that will give a large impact to the accuracy of the trained predictive model to perform day ahead solar PV power output forecating.

ABSTRAK

Penggunaan sistem Solar Fotovolta (PV) untuk penjanaan kuasa telah berkembang dengan pesat sejak beberapa tahun yang lalu. Walau bagaimanapun, pertumbuhan sistem PV solar juga menyebabkan masalah pengurusan pengedaran kuasa kerana pengendali harus sentiasa mengekalkan kestabilan grid kuasa antara penjanaan kuasa dan pengagihan kuasa. Oleh itu, tumpuan kepada peramalan hasil tenaga solar telah menjadi objektif yang penting untuk mengatasi masalah menggunakan sistem PV solar untuk penjanaan kuasa. Model ramalan hasil tenaga solar telah dibina dalam projek ini untuk meramal penghasilan kuasa setiap jam sehari ke depan dengan hanya menggunakan kaedah *Support Vector Machines* (SVM). Model ramalan dibina berdasarkan data dan modul teknologi yang didapati di Makmal Solar Fakulti Kejuruteraan Elektrik (FKE) di Universiti Teknikal Malaysia Melaka (UTeM). Model ramalan direka bentuk dengan melatihnya menggunakan data tempatan dengan aplikasi '*Regression Learner*' dalam MATLAB versi R2017b. Kesimpulannya, keputusan yang didapati adalah dengan menggunakan model SVM untuk ramalan penghasilan kuasa solar adalah sah dan ketepatan ramalan adalah memuaskan. Pembolehubah peramalan yang digunakan untuk melatih model ramalan dianalisa. Keamatan cahaya matahari dan suhu modul adalah pembolehubah yang paling dominan yang akan memberi impak yang terbesar kepada ketepatan model ramalan yang telah dilatih untuk melaksanakan ramalan hasil tenaga solar bagi sehari ke depan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

PV	-	Photovoltaics
GHI	-	Global Horizontal Irradiance
SVM	-	Support Vector Machines
SVR	-	Support Vector Regression
NNs	-	Neural Networks
RFs	-	Random Forests
HMM	-	Hidden Markov Model
GBR	-	Gradient Boosted Regression
PPF	-	Past-Predicts-Future
KNN	-	K Nearest Neighbor
AR	-	Autoregressive
MLR	-	Multi Linear Regression
ANN	-	Artificial Neural Network
GWC	-	Generalized Weather Classes
WSPR	-	Weather Statuses Pattern Recognition
LS	-	Least Square
RBFNN	-	Radial Basis Function Neural Network
ARIMA	-	Autoregressive Integrated Moving Average
SARIMA	-	Seasonal Autoregressive Integrated Moving Average
GRNN	-	Generalized Regression Neural Network
NWP	-	Numerical Weather Predictiion
ESDLS	-	Evolutionary Seasonal Decomposition Least Square
ELM	-	Extreme Machine Learning
WSVM	-	Weighted Support Machines
LSSVM	-	Least-Square Support Vector Machines
FFA	-	Firefly Algorithm
GP	-	Genetic Programming
LVQ	-	Learning Vector Quantization
WD	-	Wavelet Decomposition
WT	-	Wavelet Transform
GA	-	Genetic Algorithm
PSO	-	Particle Swarm Optimization
SOM	-	Self-organizing Map
MSE	-	Mean Square Error
RMSE	-	Root Mean Square Error
MAE	-	Mean Absolute Error
CSI	-	Clear Sky Index
RH	-	Relative Humidity

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Motivation

Nowadays, large scale power plant that generate electricity by renewable energy has been utilized all around the world. There are many types of renewable energy used to generate electricity in power plant and one of the most common used renewable energy for power generation for the power grid are solar energy as the solar photovoltaic (PV) system had growth rapidly in worldwide. Based on the review of the cumulative installed capacity for solar photovoltaic system, by the end of 2017 the cumulative installed capacity had reached 398 Gigawatts (GW) in whole world [1].

On the other hand, there are also solar PV systems used to generate electricity in Malaysia. Based on Sustainable Energy Development Authority Malaysia (SEDA), by the end of 2017 the cumulative installed capacity of solar PV systems had achieved around 380 Megawatts (MW) [2]. This shows that the solar PV systems also growth rapidly in Malaysia as Malaysia is a country that located at the equator causing Malaysia geographically exposed to the sunlight for quite a long period of time in a year.

However, using solar energy as the source of electricity creates quite some problems causing the management of the electricity become difficult. Therefore, solar power forecasting is used to improve the energy management system for grid planning, scheduling, maintenance and the balance between power generation and power consumption [3]. Solar forecasting is to predict or estimate the electricity power output generated in the future and this can be performed in several method and the method used in this project is machine learning method.

The machine learning method used in this project is Support Vector Machines (SVM) method. SVM is a supervised machine learning method that can used to solve classification or regression problems. Therefore, a prediction model that can predict the solar power output which coincided with the solar panel module technology of Malaysia is to build in this project through simulation by using MATLAB R2017b. The prediction model is trained with the data of predictor variables that obtained from local solar PV system and this prediction model should predict an accurate day ahead hourly power output based on give input data.

1.2 Problem Statement

The power output production of a photovoltaic (PV) system is unstable as the power output is mainly affected by the global solar irradiance and other factors such as weather condition and PV module characteristics. Therefore, the ability to forecast the power output precisely has become the main problem to overcome so that fluctuations can be anticipated before it occurs and necessary mitigation measures can be executed.

The main problem in this project is that currently there is no adequate investigation on the practicality of forecasting power output in tropical climates such as in Malaysia. PV performance depends strongly on local weather and environment. Predictive models used in other countries cannot be applied directly due to this limitation.

How to prioritize which parameter that must be used for prediction is one of the main problem too. This is important in designing a predictive model that corresponds with Malaysia PV systems for solar power forecasting.

There are a multitude of available machine learning methods from to choose from. One of them, SVM has been proven to perform well in solar power output forecasting in many previous researches. However, the performance of the prediction depends on the type of data and system information available. Hence, there is a need to evaluate the validity of using SVM to build a prediction model using local data.

1.3 Objective

There are a few objectives have to be achieved during the analysis of the validation of Solar PV Forecasting by SVM predictive model. Below shows the objectives of this project.

i) To analyze the potential of Support Vector Machine (SVM) in PV forecasting.

ii) To predict the power output of the PV system in the solar lab at FKE UTeM.

iii) To determine the validity and necessity of the predictor variables to train the predictive model.

1.4 Scope and Limitation

The analysis of the solar PV forecasting with machine learning is a big field to work into. Therefore in this project, there are some scope and limitation that are made to narrow down the scope of work as shown below.

a) To perform solar power forecasting by using Support Vector Machines (SVM) algorithms.

b) Perform simulation by using the regression learner in MATLAB software.

c) Using the data collected from the PV system of FKE in 2016 to train a predictive model with SVM.

d) The input data is restricted to two parts. The first part is the hourly power output from PV and module temperature. This data is obtained from the 6 kWp monocrystalline system in FKE. The second part comprise meteorological parameters such as irradiance, ambient temperature, relative humidity, wind speed and amount of rain.

CHAPTER 2

LITERATURE REVIEW

2.1 Solar Forecasting

As the solar photovoltaic (PV) technology had growth rapidly these few years, the photovoltaic has become the main target in the electricity market and this cause the solar photovoltaic power plant become more popular.

2.2 Machine Learning

Machine learning is an artificial intelligence that classified in the subfield of computer science that can learn and solve problems which are impossible to be represented by explicit algorithms or equations [4]. Machine learning algorithms can analyze and interpret the input and output data and determined the relation between the input data to perform classification or prediction based on given data. Machine learning can be classified into two categories which are supervised learning and unsupervised learning.

2.2.1 Unsupervised Learning

Unsupervised learning is a learning model that group and interpret the dataset consist only input data to determine the hidden pattern in the data. The main unsupervised learning technique is clustering, which group the data based on the common pattern or characteristics. Clustering also divided into two categories which are hard clustering and soft clustering. Hard clustering is that the every of the data point is either completely belong to a cluster or vice versa while soft clustering is to assign data point into more than one cluster with a probability [4]. Table 2.1 show the common hard clustering and soft clustering algorithms.

Table 2.1 Algorithms of Hard Clustering and Soft Clustering

Hard Clustering	**Soft Clustering**
k-Means	Fuzzy c-Means
k-Medoids	Gaussian Mixture Model
Hierarchical Clustering	
Self-Organizing Map	

2.2.2 Supervised Learning

Supervised learning is a model that required the known information of both input data and desired output data to learn and find a general relation between the inputs and outputs [5]. In supervised learning, the training data is obtained from a set data of training examples. The supervised learning model will analyze the training data and developed a predictive model. Supervised learning can classified into two big categories which are classification and regression. Classification techniques is used to predict discrete responses while regression techniques is used to predict continuous responses. Table 2.2 Show the algorithms that categorized under classification and regression.

Table 2.2 Algorithms for Classification and Regression

Classification	**Regression**
Support Vector Machine	Linear Regression
Discriminant Analysis	Generalized Linear Model
Naive Bayes	Support Vector Regression
Nearest Neighbor	Gaussian Process Regression model
	Ensemble Methods
	Decision Trees
	Neural Networks

2.2.3 Support Vector Machines (SVM)

Support Vector Machine (SVM) is a supervised learning model under classification categories that can analyzed the input data and determined the linear decision boundary (hyperplane) to classify all data of one class from another class [6], [7]. After the data sets is separated linearly, the best hyperplane is achieved by finding the largest margin for the nearest data point between two classes and the classification model is achieved. The SVM classification algorithms can be modified and extended with more function in machine learning. The use of SVM for regression is first introduced by Vladimir N. Vapnik in 1996 and this method is known as Support Vector Regression (SVR)[5]. Figure 2.1 shows how the SVM separated the non-linear input data into two category by hyperplane.

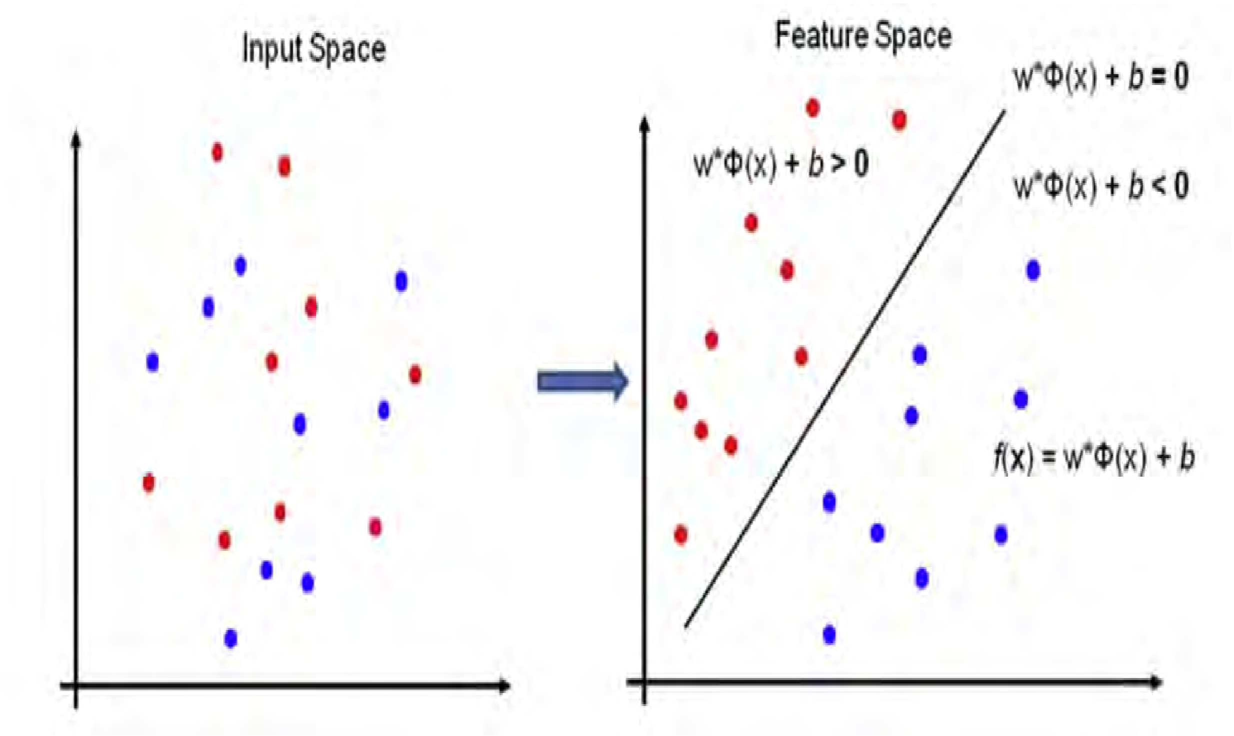

Figure 2.1 Classification of Non-Linear Input by Support Vector Machine [6]

2.2.4 Support Vector Regression (SVR)

The Support Vector Regression (SVR) is the predictive model that solve regression problem based on SVM algorithms [5]. SVR show an excellent performance in non-linear regression and time series prediction tasks [5], [7], [8]. The SVR developed a model that depend and learn on a subset of training data with a large number of predictor variables that allow the SVR model to determine the non-linear decision boundaries and predict a continuous response [8]. The purpose of using SVR algorithms is same as SVM that is minimize the error and maximize the hyperplane margin in term of prediction. Therefore, SVR algorithms is chosen as the method to use in this project as SVR is a suitable method to apply in solar power forecasting to predict the power output of the power plant.

2.3 Solar Forecasting Predictive Model

There are a lot of existed machine learning based predictive model for the power output or solar irradiance forecasting such as Neural Networks (NNs), Random Forests (RFs), Hidden Markov Model (HMM),Gradient Boosted Regression (GBR), Support Vector Machine (SVR) and etc. On the other hand, there are also some other predictive model based on Extreme Machine Learning (ELM) and Quantile Regression Forests as machine learning tools. From the previous research, there are comparison done by the researcher to determine which method is the better predictive model for solar forecasting.

2.3.1 Comparision between SVR and other Predictive Model

As mentioned above, Support Vector Regression (SVR) is one of the most common solar forecasting predictive model for power output or solar irradiance prediction. However, from the previous research we know that there are some other ready exist predictive model too. Therefore, it is important for us to determine which method give the best result in solar