A BENCHMARK COMPARISON OF GAUSSIAN PROCESS REGRESSION, SUPPORT VECTOR MACHINES, AND ANFIS FOR MAN-HOUR PREDICTION IN POWER TRANSFORMERS MANUFACTURING

by

Kamil Işık, B.S.

Thesis

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

in

INDUSTRIAL ENGINEERING

in the

GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

of

GALATASARAY UNIVERSITY

Supervisor: Assoc. Prof. Dr. S. Emre ALPTEKİN

September 2022

ACKNOWLEDGEMENTS

This research has been financially supported by Galatasaray University Research Fund with the project number FBA – 2022 – 1091.

I sincerely thank to Assoc. Prof. Dr. Emre Alptekin for his guidance and supervision during this work, and also to my wife Kübranur Işık for her help and patience.

September 2022
Kamil IŞIK

TABLE OF CONTENTS

LIST OF ABBREVIATION .... **v**

LIST OF FIGURES TABLES .... **vi**

LIST OF TABLES .... **vii**

ABSTRACT .... **viii**

ÖZET .... **ix**

1. INTRODUCTION **1**

2. RELATED WORKS **3**

3. CASE STUDY **4**

4. MODELS **6**

4.1 Support Vector Machines 7

4.2 Gaussian Process Regression 8

4.3 Adaptive Neuro-Fuzzy Inference System 9

4.4 Evaluation Metrics 10

5. COMPARATIVE ANALYSES **11**

6. CONCLUSION **14**

7. REFERENCES **14**

LIST OF ABBREVIATION

SVM	**:** Support Vector Machine
SVR	**:** Support Vector Regression
ANN	**:** Artificial Neural Network
GPR	**:** Gaussian Process Regression
BP	**:** Backpropagation
ANFIS	**:** Adaptive Neuro-Fuzzy Inference System
MVA	**:** Mega-Volt Ampere
Kv	**:** Kilo Volt
CART	**:** Classification and Regression Trees
MLR	**:** Multiple Linear Regression
COCOMO	**:** Constructive Cost Model
ML	**:** Machine Learning
kNN	**:** k-Nearest Neighbor
MCDM	**:** Multi-criteria decision-making
DNN	**:** Deep Neural Network
GWO	**:** Grey-Wolf Optimizer
ABC	**:** Artificial Bee Colony
MCS	**:** Modified Cuckoo Search
DL	**:** Deep Learning
RMSE	**:** Root Mean Squared Error
MSE	**:** Mean Squared Error
MAE	**:** Mean Absolute Error
CRISP-DM	**:** Cross-industry standard process for data mining
$\boldsymbol{R^2}$	**:** Coefficient of Determinations

LIST OF FIGURES

Figure 1: MH Input in General Process 2

Figure 2: General Production Process Flow.. 3

Figure 3: Flow Network Diagram 4

Figure 4: ANFIS Architecture 11

Figure 5: The prediction result of testing set in SVM ($R^2 = 0.85$) 14

Figure 6: The prediction result of testing set in GPR ($R^2 = 0.87$) 15

Figure 7: The prediction result of testing set in ANFIS ($R^2 = 0.79$) 16

Figure 8: Evaluation metrics of testing set 16

Figure 9: The prediction result of another testing set in GPR with prediction interval 17

LIST OF TABLES

Table 1: Pearson correlation coefficients of different influence factors. 5

ABSTRACT

Production times affect the product valuations because it is directly relevant to the production cost. Therefore, accurate estimation of the production times is one of the key problems to be able to make correct product valuations and pricing. The man-hour unit is widely used for this purpose and it is taken into consideration while making product valuations before tendering phase of projects. It is especially important in tailor-made production which generally has a labor-intensive manufacturing environment because each product is a different project. When the man-hours are predicted by experts instead of systematic tools derived from local data, they often result in incorrect predictions which affect the ex-factory cost of the product. If the incorrect prediction has a negative deviation from the actual man-hour, it may result in a reduction of the profit margin which is a serious problem for all kinds of businesses or if the man-hour prediction is over-calculated to prevent this problem then another problem emerges and the cost-effectiveness may be lost for customers in tendering. Hence, the prediction of the man-hour without excess under-estimations and over-estimations based on systematic methods in labor-intensive manufacturing environments is important to be able to keep the overall profitability of factories. There have been several kinds of research to overcome the problems resulting from incorrect expert estimations in different industries such as shipbuilding and the aircraft industry. Also, it has been researched under the title of “effort estimation” in the software development field. However, there has not been any study directed to the Power Transformer manufacturing which has a tailor-made production mentality. In this study, man-hour predictions in Power Transformers manufacturing have been studied based on data-driven methods and machine learning applications. The results showed that the proposed forecasting system can be a good alternative to the existing ones.

Keywords: Man-hour prediction, predictive models, power transformers, project management

ÖZET

Üretim süreleri, üretim maliyeti ile doğrudan ilişkili olduğu için ürün değerlemelerini etkiler. Bu nedenle üretim sürelerinin doğru tahmin edilmesi, doğru ürün değerlemeleri ve fiyatlandırması yapabilmek için en önemli sorunlardan biridir. Adam-saat birimi bu amaçla yaygın olarak kullanılmakta ve projelerin ihale öncesi ürün değerlemeleri yapılırken dikkate alınmaktadır. Genelde emek yoğun bir üretim ortamına sahip olan, proje bazlı üretimlerde her ürün farklı bir proje olduğu için özellikle önemlidir. Adam-saatler yerel verilerden türetilen sistematik araçlar yerine uzmanlar tarafından tahmin edildiğinde, genellikle ürünün fabrika çıkış maliyetini etkileyen yanlış tahminlerle sonuçlanır. Yanlış tahminin gerçek adam-saatten negatif bir sapmaya sahip olması, her türlü işletme için ciddi bir sorun olan kar marjının düşmesine sebep olur. Bu sorunu önlemek için adam-saat tahmininin gereğinden fazla hesaplanması da başka bir sorun ortaya çıkararak müşteriler için ihale sırasında maliyet etkinliği azaltabilir. Bu nedenle, emek yoğun üretim ortamlarında sistematik yöntemlere dayalı aşırı düşük ve fazla tahminler olmadan adam-saat tahmini, fabrikaların genel karlılığını koruyabilmek için önemlidir. Gemi inşa ve havacılık endüstrisi gibi farklı sektörlerde yanlış uzman tahminlerinden kaynaklanan sorunların üstesinden gelmek için çeşitli araştırmalar yapılmıştır. Ayrıca yazılım geliştirme alanında "efor tahmini" başlığı altında aynı konu irdelenmiştir. Ancak proje bazlı üretim anlayışına sahip Güç Trafosu imalatına yönelik herhangi bir çalışma yapılmamıştır. Bu çalışmada, veriye dayalı yöntemler ve makine öğrenimi uygulamalarına dayalı olarak Güç Transformatörleri imalatında adam-saat tahminleri incelenmiştir. Sonuçlar, önerilen tahmin sisteminin mevcut olanlara iyi bir alternatif olabileceğini göstermektedir.

Anahtar Sözcükler: Adam-saat tahmini, Tahminleme modelleri, Güç trafoları, Proje yönetimi

1. INTRODUCTION

Effort estimation has been a challenge for various industries, especially for the ones the labor intensiveness is high, the production is customized and project-based like shipbuilding, aviation, and construction. Power Transformers (PT) manufacturing also has a similar environment and manpower is used intensively in production steps.

Power transformers with outputs from 2.5 MVA to over 1000 MVA are static electrical machines. They are capable of converting one input voltage level to another output voltage level using electromagnetic induction. Their manufacture inherently depends on human labor. Although attempts have been made to implement some automation, most tasks are still performed by humans. These machines have very solid constructions with only minor changes over the last few decades. And they have a high degree of customization. They require special designs for high power and special applications from a structural and electromagnetic point of view. For this reason, each product requires its own production process that takes into account all customer requirements. (Mendes et al., 2019) This type of production is called tailor-made production as a nomenclature throughout the industry.

The total manpower required to complete a PT is calculated as a unit of man-hour (MH). Man-hours can have different meanings in different circumstances. It refers to an integrated unit of human labor and time required to perform individual jobs required in the production process in this document. For example, if a job is to cut a silica steel plate with a width of 5 x 5 m and two people needed to work for 2 hours to complete the job, then the total man-hour would be 4 man-hours for this task. MH is an important cost-driving factor for production when the production environment is labor-intensive and the automation rate is low. Actually, it is one of the three main cost items together with material costs and energy costs. It is also needed for allocating resources and creating reasonable schedules. Therefore, it has a vital role in project management and planning.

The project managers generally want to know the MH input in the early stages of design and even before the design to be able to calculate the overall product cost for submitting offers and participating in the tendering held by clients because each unpredicted man-hour means extra cost. PT projects, or in other words their orders, are usually received through competitive bidding processes organized by a customer who generally is an electricity distribution authority in the public or private sector. After an announcement of new bidding, the PT manufacturers create a cost sheet. The cost sheet consists of the materials that will be consumed during the production and some technical characteristics of the PT. The MH is predicted after the cost sheet is prepared. The material amount and the electrical properties in the cost sheet are predictors for the total MH. Fig. 1 shows the general process and the place where MH is included.

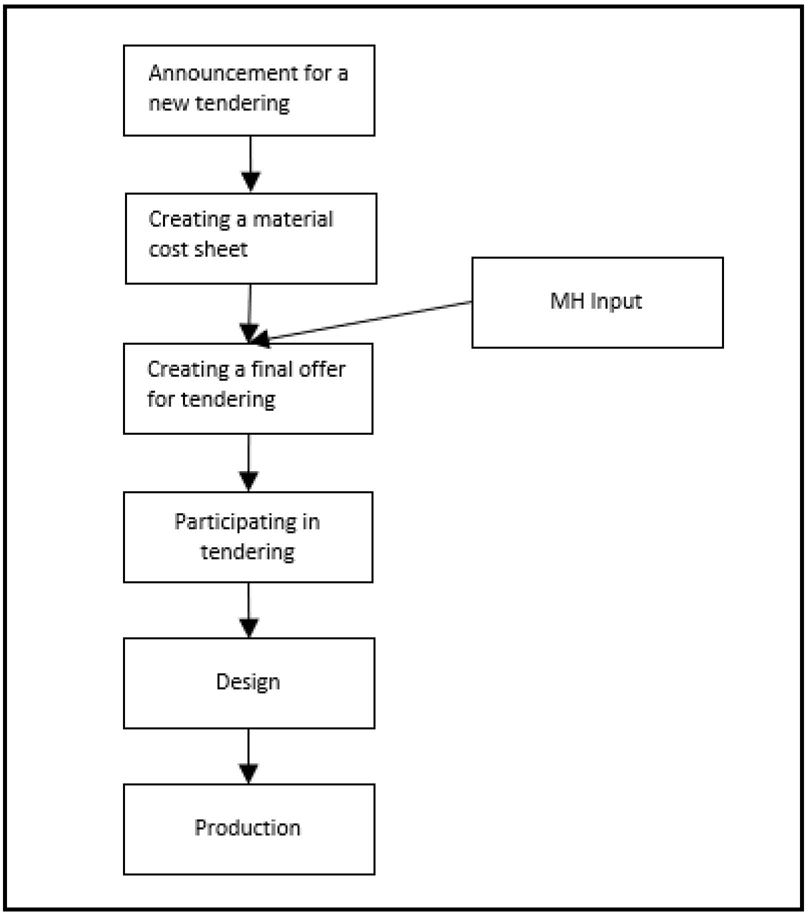

Fig. 1. MH Input in General Flow

After the tendering process, the contractor has the sole responsibility for the cost of the project. The selling price of a project should be greater than the cost for a profitable business and the selling price is determined in the bidding phase and it is indicated as a fixed price in the contract. And, this fixed price given to the client should include the cost and target profit. In this case, there is a possibility that each unpredicted MH will reduce the profit margin. Therefore, accurate estimation of MH is crucial. Under-estimations may result in profit loss.

The MH of each project is forecasted by comprehensively taking into account the specialties of the project, past experiences, and the environment. This prediction is made conventionally by experts comparing the possible orders' specifications with similar projects that have been produced before. The human expert uses several factors that can be thought of at each point to forecast the MH. However, most industrial processes show non-linear behavior, especially when the level of automation in production is relatively low. Therefore, even for identical projects, different man-hours can be forecasted depending on several different factors. (Lee et al., 1997) We can summarize the problems of pure expert estimation carries with the following points below:

1. First and foremost, the prediction is subjective. The expert forecast is finally realized by a human expert, which is why there is no guarantee for a stable forecast. There is also concern that factors that are difficult to implement could reduce the accuracy of forecasts (Hur et al., 2013)

2. There is a lack of a system to inherently link the factors required for forecasting and the forecasted values.

3. Expert estimation is not operationally efficient and effective. It takes a significant amount of time and effort to make forecasts.

4. It is not easy to exchange embedded know-how about forecasts. The know-how of the experts is not easy to quantify and for this reason, considerable previous experience is required to engage someone other than the applicable expert.

These problems can be overcome by systematizing the man-hour forecasting process. If the factors necessary for the forecast are quantified and a forecast model is created from them, a cost-effective, efficient and objective forecasting model can be implemented. Due to the nonlinear and fuzzy behaviors of project-based production environments, we especially focused on the models which can represent these behaviors. Therefore, we decided to try ANFIS and GPR as well as SVM in this study. ANFIS combines fuzzy logic with neural networks and GPR uses Gaussian Processes which are probabilistic. Both have an important chance to capture the nonlinear nature of project-based production environments.

The production of a PT can be considered to start when the copper conductors are loaded to winding machines. The silica steels are cut and stacked for core manufacturing almost simultaneously. Overall manufacturing steps like winding, core cutting & stacking, tank manufacturing, active parts assembly are shown in Fig.2.

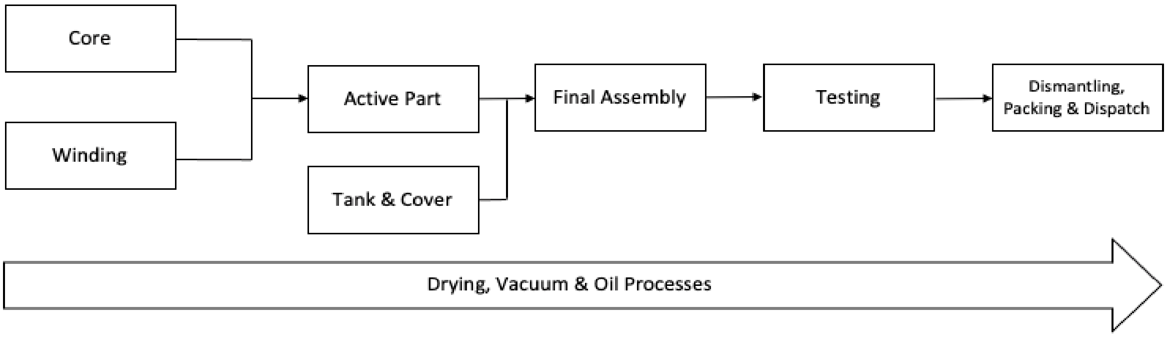

Fig. 2. General Production Process Flow

Each workstation in this process requires a certain effort to complete tasks and proceed with the next steps. The total effort of all stations required to produce the product equals to total man-hour of the project. The total man-hour means total direct labor force on the product, so just blue-collar employees' labor force is meant because while blue-collar employees' operating hours can be measured white-collar employees' working hours flexible and can not be strictly measured. This total man-hour directly affects the ex-factory cost of the product. So, it is important to predict this before submitting offers to customers because each man-hour approximately equals to 41$ at the studied factory including estimated overhead costs. The over-estimations or under-estimations directly affect the profit margins. Power transformers are strategic elements for every power supply network. Therefore, there is a competitive international market for these essential products. Optimizing the every cost input is very important for this reason. There numerous constraint during prediction of man-hour of a power transformer. Apart from the materials required for production and electirical properties of power transformers, there are also some other specifications that may affect the total effort to produce a power transformer such as phase number, connection, insulation and cooling types. These and some other technical details represent the technical diffuculty and noisiness of production of a power transformer. As such, the technical diffulcty of a power transformer described as above has not been considered by this study. It was handed over to models' prediction performance.

2 RELATED WORKS

Systematization of man-hour prediction has been attempted in many different industries. (Liu & Jiang, 2005) adopted the artificial neural network (ANN) as a predictor to develop a scheduling system for shipbuilding and confirmed the neural network model considerably outperforms multiple regression and simple regression. (Yu & Cai, 2005) predicted man-hours of aircraft assembly with BP neural networks and support vector machines (SVM) by concluding that SVM is more accurate with better fitness. (Hur et al., 2013) predicted man-hour with classification and regression trees (CART) and multiple linear regression (MLR) in the shipbuilding industry and compared them with expert predictions and concluded that the CART model had a lower forecast error than the MLR model. (Ishii et al., 2014) studied the role of man-hour in Engineering, Procurement and Construction projects. They developed simulations for strategies based on man-hour for order acceptance. As a result, they showed that the best strategy is maintaining the balance of MH for project execution and cost estimation. (Chou & Chang, 2001) introduced a variable indicating the level of technological advancement to allow a forecasting technique using historical production data to overcome the limitation that production technology cannot be integrated into shipyards over time.

The term “effort estimation” in the software engineering industry is widely used to forecast the cost of software projects. It applies to what we intend to do with man-hour prediction and it offers us a variety of studies due to plentiful studies conducted before. Effort estimation depends on local and historical data from previous projects in software development, and Constructive Cost Model (COCOMO) is one of the algorithmic and regression-based models most commonly used for software effort estimation. COCOMO divides projects into three categories depending on some characteristics such as software type, staff experience and size (Karaboğa & Kaya, 2019). It takes the project category into account when assigning initial values to the features. In a comprehensive literature review, (Wen et al., 2012) reviewed 84 primary studies on machine learning (ML) techniques for software effort estimation. Based on this study, it was found that eight types of ML techniques were applied when estimating software effort, and they concluded that machine learning models provide more accurate estimates compared to non-machine learning models. (Kirmani, 2017) also made a comprehensive literature review for software effort estimation in the early stages of software development. Additionally, (Asad & Gravino, 2019) performed a literature review based on the empirical studies published between 1991 and 2017. According to this systematic literature survey, the most used ML methods are ANN and SVM and these performed better than the other ML approaches.

The prediction of man-hour is also directly related to the cost estimation of the projects such as plant construction. Regarding the accuracy of the cost estimate, there are different types of research that were conducted before. (Oberlander & Trost, 2001) examined the determinants of the accuracy and evolution of the cost estimate and operated a system to predict the accuracy of the estimated costs. (Bertisan & Davis, 2008) analyzed the costs of 63 projects and statistically evaluates the accuracy of the estimated costs. (Nielsen & Brunoe, 2012) performed a statistical estimation method for bid purposes in an engineer-to-order environment where cost prediction is intensively resource-oriented. In particular, (Gerrard, 2000) point out that the cost estimation accuracy correlates positively with the volume of man-hour.

The application of machine learning techniques has gained considerable momentum to predict effort from 2006 onwards. The best-studied algorithmic approaches such as ANN, fuzzy logic, and nature-inspired algorithms have been continuously used in almost every aspect of effort estimation. (Karna et al., 2020) used k-Nearest Neighbor (KNN) clustering algorithm to effort modeling of agile software projects. (Azath et al., 2020) used modified Fuzzy C means clustering and gave various rules as input to the neural network which was incorporated with some optimization algorithms like modified cuckoo search (MCS), artificial bee colony (ABC), and hybrid ABC-MCS algorithms. (Sehra et al., 2018) developed a hybrid model to combine Multi-criteria decision-making (MCDM) and ML algorithm approach to predict the effort more correctly. The research community is recently considering deep learning (DL) as a potential solution for effort estimation. Deep Neural Network (DNN) enables the complex relationship between cost drivers and effort to be represented, making it a better option for effort estimating. (Muhammad et al., 2021) studied the application of meta-heuristic algorithms to obtain a meaningful and applicable parametric model for software effort estimation and applied the StrawBerry (SB) and Gray Wolf Optimizer (GWO) meta-heuristic algorithms to obtain a meaningful and applicable parametric model. The GWO algorithm was used for the starting weight optimization and SB for the learning rate optimization.

3 CASE STUDY

Historical data is an extensive reflection of the internal mechanism of change in a system. (Bing, 2014) Therefore, we need to use the available historical data to be able to build a man-hour prediction system. There are some kinds of data, related to power transformers produced before about their electrical and mechanical properties such as the power of transformer and weights of materials used. These kinds of data directly affect the size of the transformers. It is generally interpreted by experts, even if there is no empirical prediction method, that more power and more raw materials mean bigger transformers and need more production effort. This heuristic approach is not wrong but there doesn’t exist a quantitive assessment method for variables affecting the total man-hour of the project.

We have historic data of 1249 transformers in 395 different projects after pre-processing the data by eliminating or transforming the missing or unusual values and typos. The applied data mining process in this study uses Cross-industry standard process for data mining, known as CRISP-DM. It's a two-decade-old methodology that, according to many user polls and surveys, is still the de facto standard for developing insight discovery and data mining projects. (Martinez et al., 2021) This is an iterative process structured around six phases:

• Business understanding
• Data understanding
• Data preparation
• Modeling
• Evaluation
• Deployment

We will compare our method with the expert’s prediction. Hence, we will consider the factors that experts look for prediction. The relationship between these factors is not straightforward, making the prediction work a non-linear problem. These predictive model driving input parameters have important weight for the predictive accuracy of the model. So, the correlation analysis of related parameters is important to analyze their contribution and relevance. Furthermore, they can be influencing each other and some of them may have weak influence. The correlation analysis of historic data was made according to these assumptions and the results against total man-hours are shown in Table 1.

Table 1. Pearson correlation coefficients of various influencing factors.

Quantity	MVA	Kv	Copper	Insulation	Silica Steel	Oil	Radiators	Steel
-0.015	0.797	0.592	0.807	0.757	0.780	0.265	0.128	0.829

The quantity in the table means the number of transformers that will be produced in the scope of the same project. MVA and kV are relevant to the electrical properties of the transformer and they state the power and potential energy respectively. The other parameters represent the materials that will be consumed for production and their measure unit is kg. The Pearson coefficient values range from -1 to 1, and it measures the level of linear correlation among two variables. Generally, coefficient correlations bigger than 0.4 are considered as good positive correlations, and smaller than -0.4 as negative correlations. We can pay attention to the parameters having a significant correlation to reduce complexity and noisy interference. So, the driving factors would be MVA, Kv, and the weights of copper, insulation materials, silica steel, and steel. After the determination of driving parameters, the data should be normalized due to differences in magnitude and unit before regression models.

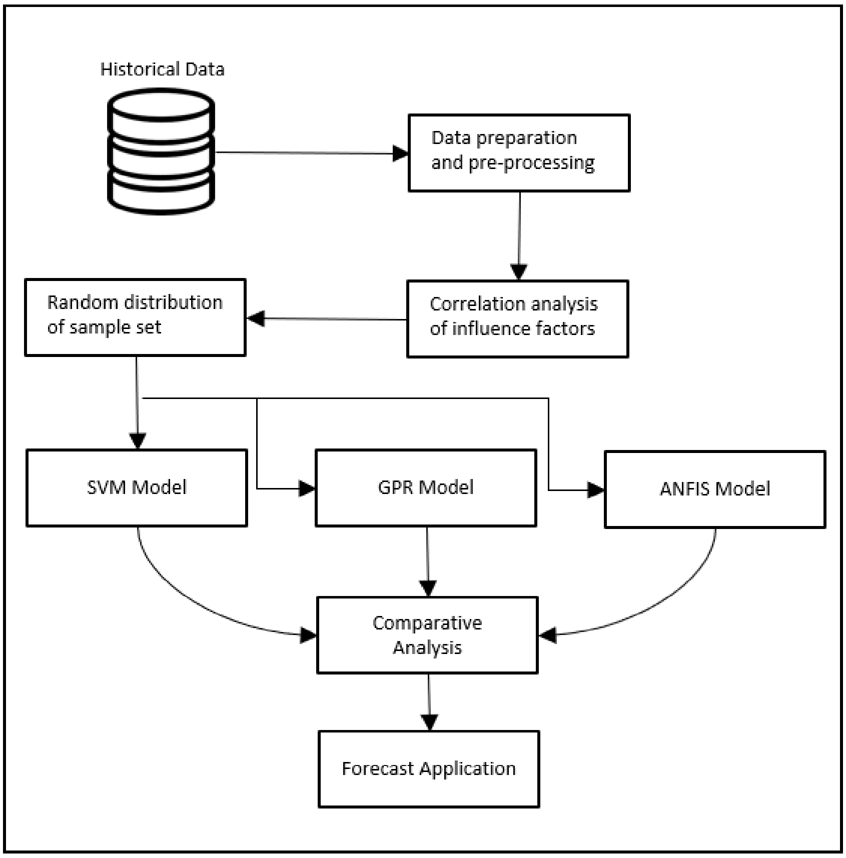

Fig. 3. Flow Network Diagram

4 MODELS

Machine learning algorithms provide the capability to understand and then predict the target outputs by themselves. The performance of an ML algorithm depends quite heavily on the selection of features and their training success. Three different ML algorithms have been used in this study and flow network diagram is given in Fig. 2.

4.1 SUPPORT VECTOR MACHINE (SVM)

SVMs are well known generally in classification problems. It is a parametric and supervised machine learning algorithm depending on statistical learning theory (Chun-Hua et al., 2010), (Vapnik & Cortes, 1995). SVM works by mapping data to a high-dimensional feature space, then the data can be categorized. A delimiter between the categories is found and the data is transformed with a kernel function by optimizing its regularization parameters so that the delimiter can be plotted as a hyperplane. Characteristics of new data can then be used for prediction. The SVM has evolved over time and has also been used in regression and outlier detection problems with nonlinear systems.

The ε-support vector regression (SVR) has been proposed by (Vapnik & Cortes, 1995) for regression tasks. While the objective is to minimize the sum of squared errors in most linear regression models like below.

$$\text{MIN} \sum_{i=1}^{n} (y_i - w_i x_i)^2 \tag{1}$$

where y_i is the output, w_i is the coefficient, and x_i is the feature; the objective function of the SVR is to minimize the coefficients of the regression when given a set of $\{x_i, y_i\}$ input-output sample pairs. Therefore, if it is not important how large errors are, as long as they are within an acceptable range, the SVR can be a better alternative. The objective function and constraint consisting of the error term of SVR are below.

$$MIN\, \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} |\xi_i| \tag{2}$$

$$\text{subject to: } |y_i - w_i x_i| \leq \varepsilon + |\xi_i| \tag{3}$$

where ε is the specified margin, ξ is slack variable for values falling outside of ε, the constant C is known as the regularization parameter and determines the tradeoff between the prediction line and error tolerance above ε. This is a simple illustration of the SVR. Readers can refer to (Vapnik & Cortes, 1995), (Smola, 1996) and (Smola & Schölkoph, 2004) for further details about solving the objective function in (2).

4.2 GAUSSIAN PROCESS REGRESSION (GPR)

Gaussian process regression (GPR) models are kernel based, non-parametric and probabilistic models. Following the definitions and derivations in (Williams & Rasmussen, 2006), a GPR model explains the response with explicit basis functions and latent variables from a Gaussian process (GP) which is a stochastic process and defined by random variables having a joint Gaussian distribution. The latent variable covariance function gives the smoothness of the response, and the basis functions reflect the inputs to the p-dimensional feature space. Therefore, GPR models are probabilistic and Bayesian. Contrary to many popular supervised ML algorithms that mandate the exact values for every parameter of the objective function, the Bayesian approach deduces a probability distribution over all possible values.

The above-mentioned general assumption is that target (y) would be defined by $f(x) + \xi$. Where $\xi \sim N(0, \sigma^2)$ meaning that the observational error is normal independent and identically distributed with a zero mean and variance of σ^2 and $f(x)$ is drawn from the Gaussian process (GP) and can be expressed as follows:

$$f(x) \sim GP(m(x), k(x, x')) \tag{4}$$

Where $m(x)$ is mean function which considered in many previous work as $m(x) = 0$ (Chu et al., 2005), (Brochu et al., 2010) and $k(x, x')$ is covariance function depending on a kernel function and the choice of which highly influences the quality of the prediction. The joint distribution between the training outputs $f = (f_1, f_2, \dots, f_n)$ and the test outputs $f_* = (f_{n+1}, f_{n+2}, \dots, f_{n+n_*})$ is given by:

$$\begin{bmatrix} f \\ f_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} M(X) \\ M(X_*) \end{bmatrix}, \begin{bmatrix} K(X, X) & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right) \tag{5}$$

Where $\mathcal{N}(\mu, \Sigma)$ is a multivariate Gaussian distribution, $X = (x_1, x_2, \ldots, x_n)$ corresponds to the training input and $X_* = (x_{n+1}, x_{n+2}, \ldots, x_{n+n_*})$ to the test inputs. $K(X, X_*)$ is the $n \; x \; n_*$ matrix of the covariances evaluated for all the (X, X_*) pairs. The predictive Gaussian distribution could be found by receiving the training data and the test data inputs:

$$f_* | X_*, X, f \sim \mathcal{N}(\widehat{M}(X_*), \widehat{K}(X_*, Y_*))$$
$$\widehat{M}(X_*) = M(X_*) + K(X_*, X)K(X, X)^{-1}(f - M(X)) \tag{6}$$
$$\widehat{K}(X_*, X_*) = K(X_*, X_*) - K(X_*, X)K(X, X)^{-1}K(X, X_*)$$

The GPR based regression approaches involves the idea of kernel functions to describe the covariance between any two function values. The kernel functions that will be used should be Positive semi-definite kernels (PSD). (Roman et al., 2021)

4.3 ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM (ANFIS)

Fuzzy inference system was introduced by (Jang, 1993). It is a sort of artificial neural network combined with fuzzy logic. It integrates both neural networks and fuzzy logic principles and therefore has the potential to combine the advantages of both in a single framework. ANFIS architecture is composed of five layers and it’s architecture is shown in Fig. 3.

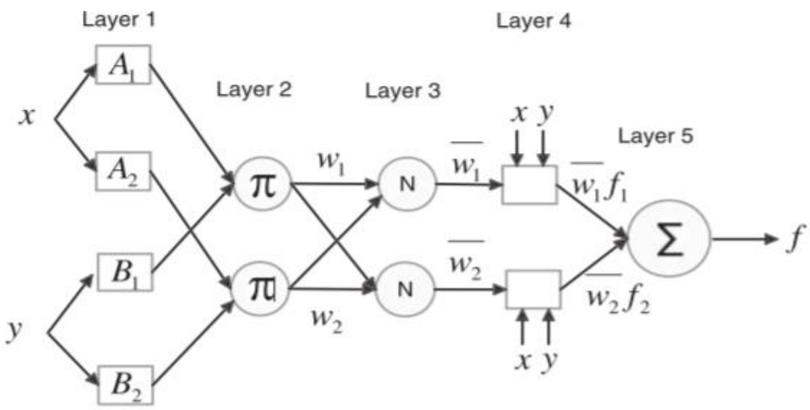

Fig. 4. ANFIS architecture. [24]

Layer1 – Fuzzification Layer

The main purpose of the first layer is fuzzification by using membership functions.

$$o_i^1 = \mu A_i(x), \qquad i = 1, 2 \tag{7}$$

$$o_i^1 = \mu B_{i-2}(y), \qquad i = 3, 4 \tag{8}$$

Layer2 – Fuzzy Rule Layer

Potential rules between the inputs are found by applying fuzzy intersection at the second layer. Therefore, it is responsible for generating the firing strength for the rules.

$$o_i^2 = w_i = \mu A_i(x)\mu B_i(y), \qquad i = 1, 2 \tag{9}$$

Layer3 – Normalization Layer

The weights obtained from the previous layer are normalized at the third layer.

$$o_i^3 = \overline{w_i} = \frac{w_i}{w_1 + w_2}, \qquad i = 1, 2 \tag{10}$$

Layer4 – Output Membership Layer

The fourth layer gets the normalized values and also the outcome parameter set as input. The number of outcome parameters of each rule is one more than the number of inputs.

$$o_i^4 = \overline{w_i} z_i = \overline{w_i}(a_i x + b_i y + c_i), \qquad i = 1, 2 \tag{11}$$

Layer5 – Defuzzification Layer

The final layer is for computing the overall output with summation of all incoming signal from layer 4.

$$o_i^5 = \sum \overline{w_i}\, z_i = \frac{\sum_i w_i z_i}{\sum_i w_i} \tag{12}$$

ANFIS can be trained with either a backpropagation algorithm or a hybrid method. While backpropagation uses gradient descent for optimization, a hybrid method generally consists of backpropagation for the parameters related to the least squares prediction and input membership functions for the parameters related to the consequence membership functions.

The gradient descent + least squares based hybrid algorithm is the first algorithm proposed for ANFIS and is used extensively today. It generally offers better performance than other derivatives-based approaches. (Karaboğa & Derviş, 2019)

4.4 EVALUATION METRICS

Accuracy is the primary criterion for assessing the models' performance success, and the commonly used evaluation metrics such as RMSE, MAE and R^2 are used both in comparison to each other and in evaluating the results of the predictive models.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - x_i| \tag{13}$$

The mean absolute error (MAE) is used to measure the closeness of the prediction to the actual outcome. The absolute error is the difference between the actual value and the forecasted value. MAE is the average of all absolute errors.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2} \tag{14}$$

Root Mean Squared Error (RMSE) is a quadratic evaluation rule that measures the mean magnitude of the error. It imputes a relatively high weight to big errors since the errors are squared before they averaged. RMSE and MAE are negatively-oriented metrics, which means small values are better.

$$R^2 = 1 - \frac{\sum(x_i - y_i)^2}{\sum(x_i - \bar{x}_i)^2} \tag{15}$$

Determination Coefficient (R^2) provides insight about how well a model can fit and forecast a set of measured data. Its value is between 0 and 1. The R^2 value approaching to 1 is a sign of better performance. (Chai & Draxler, 2014) showed that RMSE is more appropriate when the distribution of error is expected to be Gaussian. We expect also that the error will be Gaussian in this study due to expert predictions also generally Gaussian. However, still a combination of metrics is often required. Therefore, adding more measures is beneficial.

5 COMPARATIVE ANALYSES

We can begin the comparative analysis after completing the afore-mentioned CRISP-DM approach. We used cross-validation (v=5) for GPR and SVR to protect the models against overfitting. Cross-validation does this by partitioning the data set into folds and predicting accuracy on each fold. Additionally, we randomly seperated the valid data into training set (316 project) and testing set (79 project) when building the models. In this research, maximum iteration is 300 for SVM and GPR.

Optimized Kernel function type, regularization parameter C and the specified margin ε should be determined for SVM. The Gaussian, Linear, Quadratic and Cubic kernel functions have been used for hyperparameter optimization with Bayesian optimizer. The results show that the Quadratic kernel function which is:

$$K(x, y) = (x^T \times y + C)^2 \tag{16}$$

gives the most nearest predicitions. The ε value is 0.007 and C is 0.003. The forecasting results of SVM can be seen in Fig. 5. Linearly inseparable data is transformed into linearly separable one with the help of kernel. It is applied to each data instance to map the original nonlinear observations into a higher dimensional space where they become separable.

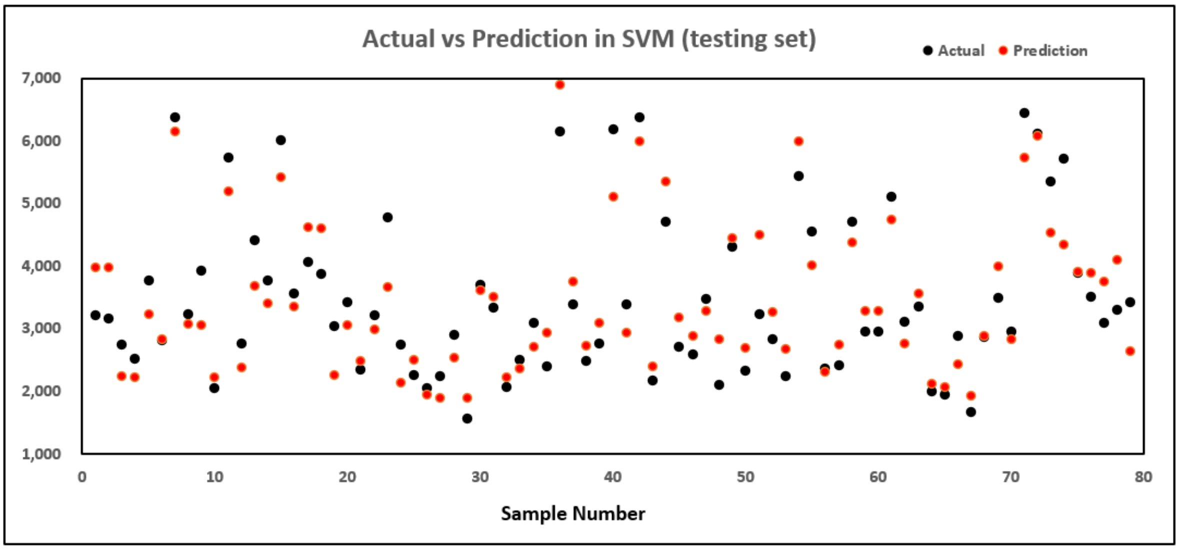

Fig. 5. The prediction result of testing set in SVM ($R^2 = 0.85$)

A hyperparameter optimization is also conducted for GPR's basis and kernel functions. The key advantage of GPR over other regression types is it's ability to represent $f(x)$ in a more flexible way. Therefore, data can have more influence on it's own form. It is seen that after a trial of different basis and kernel functions together with adequate search range for sigma value and kernel scale, the optimized basis function is Linear, and kernel function is Isotropic Exponential. The linear basis function is generalized form of linear regression. It regress the target on predictors' functions instead of pure predictors. Coefficients are predicted by least squares method in linear regression. Therefore, linear basis function is an expansion of the least squares which is:

$$Y = \phi\beta + \varepsilon \tag{17}$$

where $\phi_{ij} = \varphi_j(x_i)$ is design matrix. After least squares applying, coefficient vector β can be predicted with

$$\mathcal{B} = (\phi^T\phi)^{-1}\phi^T Y \tag{18}$$

Additionally, isotropic kernel function depends on the margin between the kernel arguments like below:

$$K(x, y) = f(\|x - y\|) \tag{19}$$

It means the direction of error is not very important for isotropic kernel. It is also suitable for our study because over and under estimations are nearly equally weighted. Lastly, kernel scale of GPR is 1 and $\sigma = 0.74$. Kernel scale is used to scale predictor matrix X and σ is the predicted noise of GPR's standard deviation. The forecasting results of GPR can be seen in Fig. 6. GPR is particularly proper for models which have uncertainty and it has a big chance for an accurate prediction under noisy environments. Even if we have quantitative data for MH prediction, the conditions in production environment are quite noisy because eventually the people are making the tasks and there are many unforeseen circumstantes during production. The GPR's probabilistic nature shows itself here as a good alternative for a stable prediction method in the long term.

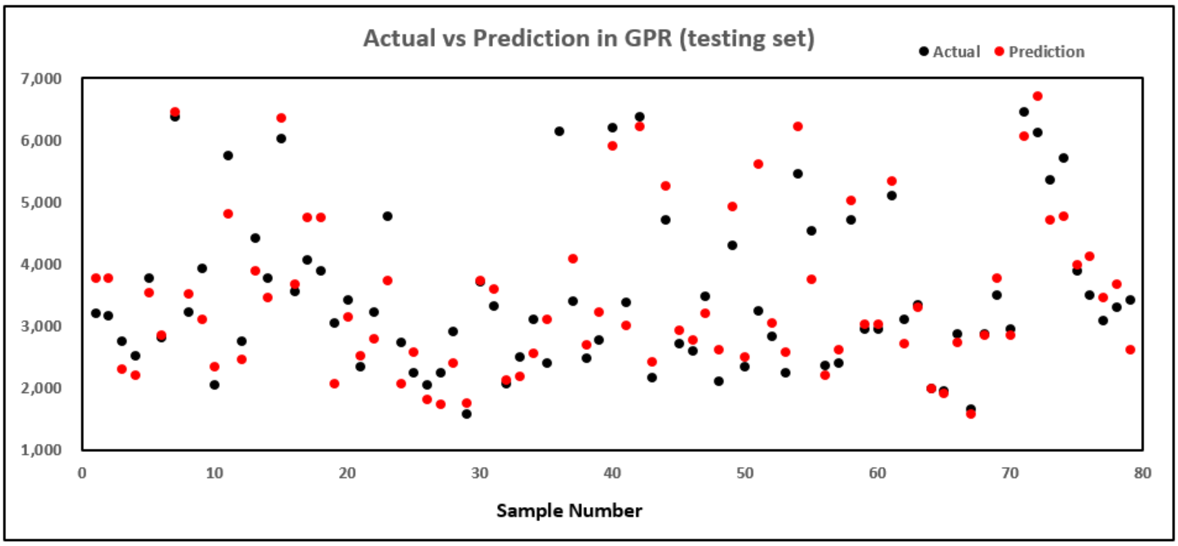

Fig. 6. The prediction result of testing set in GPR ($R^2 = 0.87$)

Similar to SVM and GPR, ANFIS parameters should also be determined. Overfitting is a known problem in ANFIS models. [28] Training process of ANFIS starts by deciding the each input variables, their membership function and number of epoch. Generally, datasets have their own maximum epochs numbers before they reach overfitting and the optimal epochs number can only be determined by running numerous experiments. To find the ideal ANFIS system, we trained the system on a variety of different parameters like recording pattern, epoch number, type and number of membership functions, number of inputs for better performance. We set the epoch numbers to 10 – 25 and 40 respectively in this research. After an analysis of the selecting different membership function types such as Gaussian ($gaussmf$) Triangular-shaped ($trimf$) and Generalized bell-shaped ($gbellmf$) and learning algorithms, the hybrid learning algorithm and Gaussian built-in membership function with “3*3*3*3*3*3” membership function numbers were used to train the ANFIS. Addionally, we used Linear output membership function with one output based on Takagi-Sugeno model. The forecasting result of ANFIS can be seen in Fig. 7. The performance of ANFIS seems lower than GPR at the first glance. It can be evaluated that even if IF-THEN rules of ANFIS have a great chance to describe the environment of a complex systems. The production environment of power transformers may not have so much complexity for ANFIS. A few misestimation of fuzzy regions determined by the fuzzy rules result big errors between actual and ANFIS prediction.

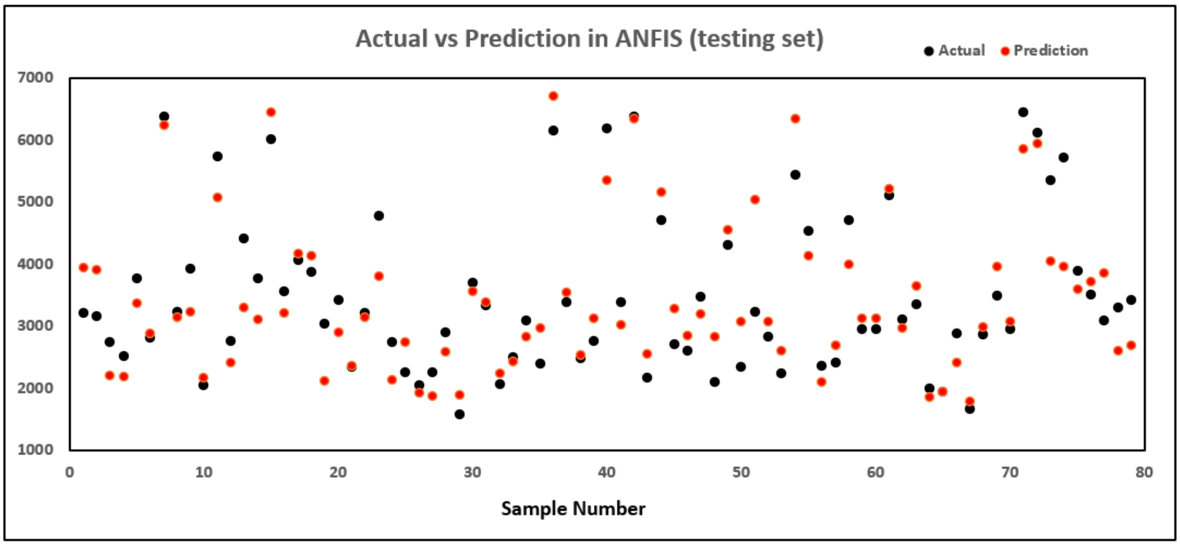

Fig. 7. The prediction result of testing set in ANFIS ($R^2 = 0.79$)

The results according to Determination Coefficients are indicative of the better performance of GPR as compared to SVM and ANFIS models. Fig. 8. shows the values of evaluation metrics in testing set.

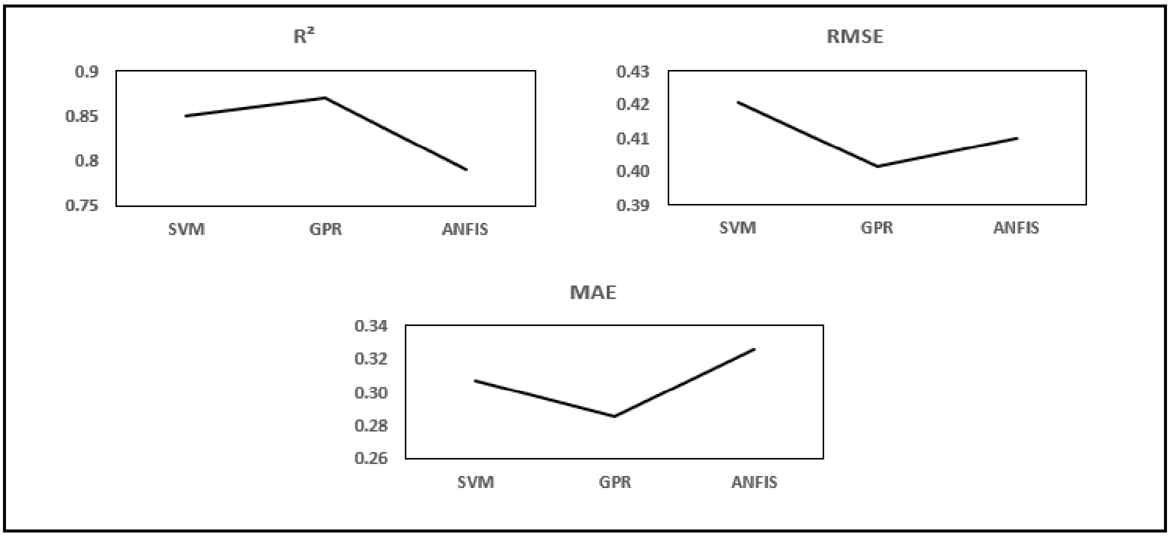

Fig. 8. Evaluation metrics of testing set

Even though the results are comparable, given that the Gaussian Processes have superior capabilities to represent highly nonlinear data relationships, its' outperforming the other two models isn't extraordinary. When given certain values of the independent variables for a new observartion, the range that eventually containing the value of the dependent variable is prediction interval. The independent variables in our study are the materials consumed during production and some electrical properties of the transformer. These are dependent on the specifications of the projects and the spesification of projects are dependent on the where the power transformer will be used. If power transformer will be used for example as a high voltage step-up power transformer just outside of a power plant. It may mean that the kVA of the transformer will be higher also and because of that the size of the transformer will be bigger due to much copper and silica steel will be consumed to yield that. Therefore, we can consider that the size of transformer is the main independent variable actually while trying to predict MH of a power transformer. We use raw material weight inputs as a dependent variable of size. The prediction interval of GPR with %95 confidence can be seen in Fig. 9. The prediction interval says in what range a future observation may be. According to Fig. 9, we can say that we are %95 confident that the actual value will be between upper and lower boundaries. It can be seen from figure that the real values are inside of the boundaries as well as predictions.

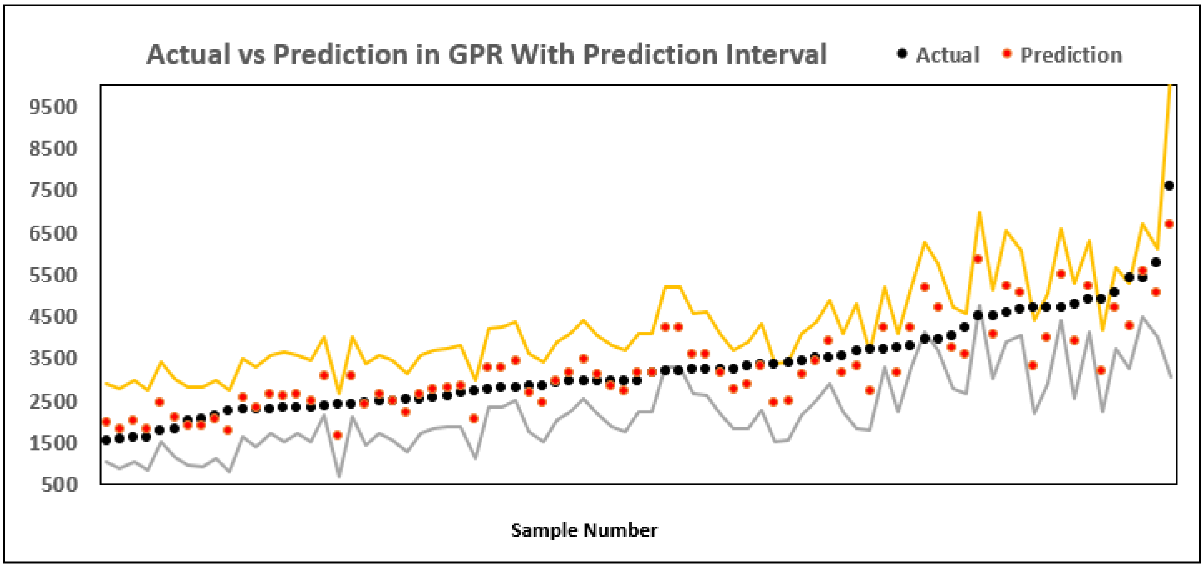

Fig. 9. The prediction result of another testing set in GPR with prediction interval

Finally, we can say that there are different types of transformers. This study has focused on three-phase, oil-insulated power transformers. They are used in electricity transmission network for long distances and they are not directly connected to the consumers. Therefore, load fluctuations and electricty losses are very less. The produced transformer type is surely important in these prediction cases because for example different connection types such as star and delta configurations can be important for required manpower.

6 CONCLUSION

Project-based productions need a high level customization and the total man-hour in the project-based production environments is a quite important cost element on total production cost. Therefore, its’ prediction before pricing and offering processes is crucial and prediction accuracy should be increased as much as possible for a profitable business. Human expert predictions are not totally satisfactory and have a lot of pains for a bunch of reasons such as slowness, low efficiency and big error. It also requires to assign a dedicated expert for prediction and this makes the prediction operation more difficult. Furthermore, this manual prediction operation causes frustration felt by planning expert. In this paper, ANFIS, GPR and SVM models were developed for a systemized man-hour predicton in Power Transformer production industry. We used data of 395 different projects. 316 of them were used for training and 79 projects were used as test dataset. Models were assessed with statistical indicators such as R^2, RMSE and MAE. GPR outperformed the others. The predictive model based on GPR can be effectively used in an acceptable error range, especially when compared to pure expert forecast. We didn’t take technical diffulcty ratings in this study and handed over it to models prediction power. Technical diffuculty may be changing according to transformer’s connection type, cooling type and insulation type or some other technical points.

Data analytics and machine learning are great leverages to improve the flexibility and efficiency. They are at the core of smart factories and Indusrty 4.0. We tried to improve a manual prediction system by systemitizing and making dependent it to data. As a result, the pains of manual prediction eliminated and the accuracy of production times have been improved as well. The main independent variable was actually the size of the transformer. The size of the transformer determines the required material amount and electrical properties or vice versa.

As future work, a variable indicating the extent of technological development can be introduced in order to overcome of being unable to include developments in production technology of power transformers. Additionally, if technical diffiulcty ratings of transformers can be quantified, they can be introduced to dataset and there can be a chance to improve models performances even better.

7 REFERENCES

A.M. Gerrard, Guide to Capital Cost Estimating, (4th ed.), Institute of Chemical Engineers, Warwickshire (2000)

Al-Hmouz A, Jun Shen, Al-Hmouz R, Jun Yan. Modeling and Simulation of an Adaptive Neuro-Fuzzy Inference System (ANFIS) for Mobile Learning. IEEE Transactions on Learning Technologies, Learning Technologies, IEEE Transactions on, IEEE Trans Learning Technol. 2012;5(3):226-237. doi:10.1109/TLT.2011.3

Asad, Ali. Gravino, Carmine: A systematic literature review of software effor prediction using machine learning methods. Journal of Software: Evolution and Process. July 2019.

Bing, D., 2014, “Reliability Analysis for Aviation Airline Network Based on Complex Network”, Journal of Aerospace Technology and Management, Vol. 6, No. 2, pp. 193-201. DOI: 10.5028/jam.v6i2.295

Brochu, E., Cora, V.., de Freitas, N.: A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. arXiv:1012.2599

C.K.I. Williams, C.E. Rasmussen. Gaussian processes for machine learning. MIT Press (2006).

Chou, C. C., & Chang, P. L. (2001). Modeling and analysis of labor cost estimation for shipbuilding: The case of China Shipbuilding Corporation, Journal of ship production, 17(2)

Chu, W., Ghahramani, Z.: Gaussian Processes for Ordinal Regression. Journal of Machine Learning Research 6(Jul), 1019–1041 (2005).

Chun-Hua Zhang, YingJie Tian, NaiYang Deng The new interpretation of support vector machines on statistical learning theory. Sci China, Ser A: Mathematics, 53(1)(2010), pp. 164

Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995

F. Martínez-Plumed et al., "CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories," in IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 8, pp. 3048-3061, 1 Aug. 2021, doi: 10.1109/TKDE.2019.2962680.

F. Tahir and M. Adil, "An Empirical Analysis of Cost Estimation Models on Undergraduate Projects Using COCOMO II," 2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE), 2018, pp. 1-5, doi: 10.1109/ICSCEE.2018.8538361.

G.D. Oberlender and S.M. Trost, Predicting accuracy of early cost estimates based on estimate quality, Journal of Construction Engineering Management (2001), pp. 173-182

Hrvoje Karna, Sven Gotovac and Linda Vicković (2020). Data Mining Approach to Effort Modeling on Agile Software Projects, Informatica 44 (2020) 231–239.

Hussain Azath, Marimuthu Mohanapriya, Somasundaram Rajalakshmi (2020). Software Effort Estimation Using Modified Fuzzy C Means Clustering and Hybrid ABC-MCS Optimization in Neural Network, J. Intell. Syst. 2020; 29(1): 251–263.

Ishii N, Takano Y, Muraki M. An order acceptance strategy under limited engineering man-hours for cost estimation in Engineering–Procurement–Construction projects. International Journal of Project Management. 2014;32(3):519-528. doi:10.1016/j.ijproman.2013.07.009

J. Bertisen and G.A. Davis, Bias and error in mine project capital cost estimation, The Engineering Economist, 53 (2) (2008), pp. 118-139.

J. -S.R. Jang. ANFIS: Adaptive Network-Based Fuzzy Inference System. IEEE Transactions on Systems, Man., and Cybernetics, Vol. 23, No. 3, pp 665-685, May 1993.

Jae Kyu Lee, Kyoung Jun Lee, Hung Kook Park, June Seok Hong, Jung Seung Lee (1997).

Developing Scheduling Systems for Daewoo Shipbuilding: DAS Project, European Journal of Operational Research 97, 380-395.

Karaboğa, Derviş. Kaya, Ebubekir. Adaptive network-based fuzzy inference system (ANFIS) training approaches a comprehensive survey. Artificial Intelligence Review. 52(4) 2263-2293

Kirmani Mudasir M (2017). Software Effort Estimation in Early Stages of Software Development: A Review, International Journal of Advanced Research In Computer Science.

Liu, B., & Jiang, Z. H. (2005). The man-hour estimation models & its comparison of interim products assembly for shipbuilding. International Journal of Operations Research, 2(1), 14–19.

Mendes H, Linhares C C, Coutinho C P, Tavares S M O, Teixeira M, Moura J P A, Novais C, Teixeira R, Marques Pinho A C, Meireles J F B, (2019). Smart Design and Manufacturing of Power Transformers Tanks, IEEE International Conference on Environment and Electrical Engineering

Minhoe Hur, Seung-kyung Lee, Bongseok Kim, Sungzoon Cho, Dongha Lee, Daehyung Lee (2013). A study on the man-hour prediction system for shipbuilding, Springer Science+Business Media New York.

Moosavi S.H. Samareh, Bardsiri V. Khatibi. Satin Bowerbird Optimizer: A new optimization algorithm to optimize ANFIS for software development effort estimation. Engineering Applications of Artificial Intelligence. Volume 60. April 2017. pp 1-15.

Roman, I. et al. (2021) 'Evolution of Gaussian Process kernels for machine translation post-editing effort estimation', Annals of Mathematics and Artificial Intelligence, 89(8–9), pp. 835–856. doi: 10.1007/s10472-021-09751-5.

S K Muhammad, J Farhana, G Sanaa, R Zobia, N Sheneela, A Wadood (2021). Metaheuristic Algorithms in Optimizing Deep Neural Network Model for Sofware Effort Estimation, IEEE Access, Vol 9, Pp 60309-60327 (2021)

Smola AJ, Schölkopf, A tutorial on support vector regression, Statistic and Computing, journal article vol. 14, no. 3, pp 199-222, August 01 2004.

Smola AJ. Regression estimation with support vector learning machines. Master's thesis. Munchen ,Germany: Technische Universitat München; 1996.

Sumeet Kaur Sehra, Yadwinder Singh Brar, Navdeep Kaur, Sukhjit Singh Sehra, (2018). Software effort estimation using FAHP and weighted kernel LSSVM machine, Springer-Verlag GmbH Germany, part of Springer Nature 2018.

T. Chai and R. R. Draxler, RMSE or MAE – Arguments against avoiding RMSE in the literature, Geoscientific Model Development (June 2014).

T.D. Brunoe and P. Nielsen, A case of cost estimation in an engineering-to-order company moving towards mass customisation, International Journal of Mass Customisation, 4 (3/4) (2012), pp. 239-254

Tingting Yu, Hongxia Cai (2014). The Prediction of the Man-Hour in Aircraft Assembly Based on Support Vector Machine Particle Swarm Optimization, J. Aerosp. Technol. Manag., São José dos Campos, Vol.7, No 1, pp.19-30, Jan.-Mar., 2015

Wen J, Li S, Lin Z, Hu Y, Huang C (2012). Systematic literature review of machine learning-based software development effort estimation models, Inf Softw Technol 54(1):41–59