Teaching Forecasting and Prediction

Ceyhun Ozgur[1,*]**, Sanjeev Jha**[2] **and Madison Wallner**[3]

[1]*Emeritus Professor of Information and Decision Sciences, Valparaiso University, United States*
[2]*Associate Professor of Information and Decision Sciences, Valparaiso University, United States*
[3]*Class of 2024 Meteorology Undergraduate, Valparaiso University, United States*
Corresponding author: ceyhun.ozgur@valpo.edu

Abstract. This paper aimed to materialize how to teach forecasting and predictions to undergraduate students. There are two modes of thinking, including the determination of forecasters and forecast errors. To determine the best forecast, we have to use the methodology of minimizing sums of forecast errors, which is further discussed in this paper. An example of this is single exponential smoothing, which is rather difficult when it comes to determining the smoothing constant alpha. In management and administrative situations, the need for planning is great because the lead time for decision-making ranges from several years to a few days or even a few hours. The ability to predict many types of events seems, as natural today as, to be the accurate forecasting of weather conditions in a few decades.

Keywords: Teaching forecasting, exponential smoothing, forecast errors, prediction.

INTRODUCTION

The word "predict" comes from Latin descent, meaning to say or do in advance [14]. The word "forecast" comes from the English Germanic root, meaning to calculate some future event or condition [15]. Together, forecasting is the ability to predict the future. Teaching forecasting and prediction are done by utilizing the latest and most cost-efficient financial tools to optimize the financial cost of specific organizations and predict the financial situation of the specific organization to flourish [16]. The first mode is System one, which is automatic and fast and has constant low-level predictions. The second mode is System two, which is effortful and has complex computations. In this paper, we are going to discuss how we calculate the forecast error and how to teach forecasting using simple exponential smoothing and a linear regression procedure.

FORECAST ERROR

The analysis of the accuracy of forecasts over a set period of time series is of much importance, due to forecasting function that is evaluated as inappropriate techniques. Therefore, the link to economic evaluation is changed [3].

To determine the best forecast, we have to minimize the sum of the forecast errors. We first came up with the forecasted demand (Fi) and compared it with the actual demand (Ai). This is called the "Forecast Error," which is equivalent to e = Ai-Fi. Then, once we come up with the forecast error, we can then determine the sum of the forecast errors. This equation is given by (Ai-Fi). The firm or organization provides us with the actual demand, but it is up to us to establish the predicted demand. The method used by each business or organization to determine the anticipated demand varies. For instance, Fi=Ai-1 means that the predicted demand might be identical to the actual demand from the prior period.

SINGLE EXPONENTIAL SMOOTHING

Single exponential smoothing shows a class of forecasting methods that show a variety of the most successful forecasting methods that are reliant on exponential smoothing. There are multiple methods that can be considered exponential smoothing, dealing with the fact that forecasts are weighted combinations of commercial observations. This type of forecasting is based on determining the smoothing constant alpha [8]. This equation is given

by $Fi = \alpha Xi + (1-\alpha)Fi$. The hardest part of utilizing single exponential smoothing in forecasting and prediction is determining the smoothing constant alpha. The larger the value of α or closer to 1, the less smoothed it is, and the smaller the value of α or closer to 0, the more smooth it is. Since the aforementioned smoothing calculations are straightforward and only a minimal amount of historical information should be preserved from one prediction to the next, it is possible to employ higher-order polynomials as forecasting models [1].

Example 1: Let us take a situation with the following demand values 10, 12, 14, 15 and smoothing constant $\alpha = 0.1$. Let us determine the forecasted demand with these given values. Assume that the first forecast is the same as the actual demand, which is 10.

The forecast for 2019 is 10, so the forecast for 2020 is (0.1*10). So:

$$Forecast_{i+1} = (\alpha * Actual_i) + (1 - \alpha) * Forecast_i$$
$$Forecast_{20} = (\alpha * Actual_{19}) + (1 - \alpha) * Forecast_{19}$$
$$F_{20} = (0.1 * 10) + (1 - 0.1) * 10 = 10$$
$$A_{20} = 12$$

The Forecast for 2021 is:

$$Forecast_{i+1} = (\alpha * Actual_i) + (1 - \alpha) * Forecast_i$$
$$Forecast_{21} = (\alpha * Actual_{20}) + (1 - \alpha) * Forecast_{20}$$
$$F_{21} = (0.1 * 12) + (1 - 0.1) * 10 = 10.2$$

The Forecast for 2022 is:

$$Forecast_{i+1} = (\alpha * Actual_i) + (1 - \alpha) * Forecast_i$$
$$Forecast_{22} = (\alpha * Actual_{21}) + (1 - \alpha) * Forecast_{21}$$
$$F_{21} = (0.1 * 14) + (1 - 0.1) * 10.2 = 10.58$$

The Forecast for 2023 is:

$$Forecast_{i+1} = (\alpha * Actual_i) + (1 - \alpha) * Forecast_i$$
$$Forecast_{23} = (\alpha * Actual_{22}) + (1 - \alpha) * Forecast_{22}$$
$$F_{21} = (0.1 * 15) + (1 - 0.1) * 10.58 = 11.02$$

REGRESSION ANALYSIS

We can also determine the forecast error by determining the forecast with a different procedure. The most common procedure that is used is the linear regression procedure.

Regression analysis is one of the most widely used techniques for reviewing multi-factor data.

Its familiarity and ability to predict come from a conceptual process of prediction utilizing an equation expressing relationships between a variable of interest and a set of other predictability variables [17]. In a linear regression procedure, we have to come up with the equation for Fi. This can also be stated as $Y = a + bX$, where a is the y-intercept, b is the slope, X is the value of the forecast, and Y is the forecasted value of demand, which can also be stated as F.

Along with this, there is also robust regression. This is the same as linear regression but on a more "robust" scale. Some factors that are important when doing robust regression analysis are to protect against distortion by anomalous and good efficiency when the data come from the ideal Gaussian model, as well as from a range of other similar models [11]. Regardless of which regression we decide to utilize, they are all very similar to one another when it comes to forecasting and prediction.

Example 2: The waiting time (in minutes) for bank customers wanting to sign up for a loan was measured at four banks for all customers who arrived between 2:00 and 2:30 p.m. on a certain Wednesday. The results are shown below.

Bank Waiting Time (Minutes)

Bank 1 (5 Customers)	Bank 2 (4 Customers)	Bank 3 (7 Customers)	Bank 4 (6 Customers)
10	8	5	1
19	25	11	20
5	17	24	9
26	36	16	5
11		18	10
		29	12
		15	

Research Question

Are the mean waiting times for the bank customers the same for all four banks?

Answer: (Using MINITAB)

The mean waiting times for the bank customers are the same for all four banks because the F-value calculated is 1.76, which is less than 3.16 from the F-table. Therefore, there is no significant difference in waiting times between the four banks.

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	3	377.7	125.89	1.76	0.191
Error	18	1288.2	71.56		
Total	21	1665.8			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
8.45957	22.67%	9.78%	0.00%
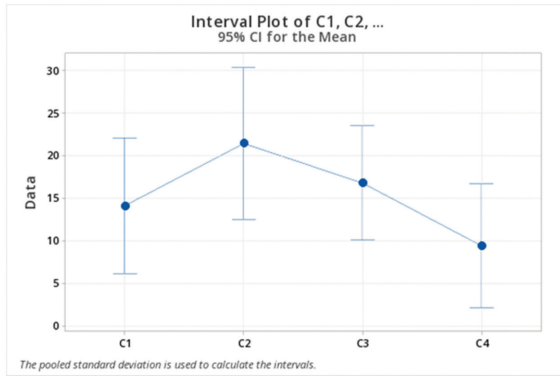
Figure 1. Interval plot of comparison of means.

Means (Figure 1)

Factor	N	Mean	StDev	95% CI
C1	5	14.20	8.29	(6.25, 22.15)
C2	4	21.50	11.90	(12.61, 30.39)
C3	7	16.86	7.95	(10.14, 23.57)
C4	6	9.50	6.47	(2.24, 16.76)

Pooled StDev = 8.45957.

Example 3: A real estate appraiser is developing a regression model to predict the selling price in a certain neighborhood. She is planning to use square footage as a variable, but she also wishes to incorporate the type of house (Colonial = 0, Tudor = 1, or Contemporary = 1) and whether the house has a brick front (brick = 1, no brick = 0). She obtains the following data.

Regression Analysis Coefficients

Area (sq. ft.)	Type	Brick	Price ($000)
2100	Contemporary	No	85
2400	Tudor	Yes	110
2396	Colonial	Yes	115
3000	Tudor	Yes	175
2850	Tudor	Yes	160
2500	Contemporary	No	135
2000	Colonial	No	79
2200	Colonial	No	89
2600	Contemporary	No	98
2400	Colonial	No	105
2900	Tudor	Yes	154

Research Question

(a) State the model that uses indicator variables.
(b) Estimate the model, using the given data.

(c) Interpret the values of the coefficients and constant terms.

Answer: (Using MINITAB)

(a) Variable Definitions:

Y = Price

$X1 = 1$, if contemporary

$X1 = 0$, otherwise

$X2 = 1$, if Tudor

$X2 = 0$, otherwise

$X3 = 1$, if brick

$X3 = 0$, if not brick

$X4$ = Area

(b) $Y = -85.7$, C1 $= 0.0803$, C2 $= -1.1$, C3 $= 3.3$, C4 $= 8.3$

(c) Constant term is -85.7, contemporary houses have an expected price increase of 0.0803, Tudor houses have an expected price increase of 1.1, the houses with a brick front have an expected price increase of 3.3, and each square foot increases the expected price of the house by 8.3.

Regression Equation

C2	C3	C4			
0	0	0	C5	=	$-85.7 + 0.0803$ C1
0	0	1	C5	=	$-77.4 + 0.0803$ C1
0	1	0	C5	=	$-82.4 + 0.0803$ C1
0	1	1	C5	=	$-74.1 + 0.0803$ C1
1	0	0	C5	=	$-86.8 + 0.0803$ C1
1	0	1	C5	=	$-78.5 + 0.0803$ C1
1	1	0	C5	=	$-83.5 + 0.0803$ C1
1	1	1	C5	=	$-75.2 + 0.0803$ C1

Coefficients

Term	Coefficient	SE Coefficient	T-Value	P-Value	VIF
Constant	-85.7	48.9	-1.75	0.130	
C1	0.0803	0.0219	3.66	0.011	2.48
C2					
1	-1.1	12.6	-0.08	0.935	1.66
C3					
1	3.3	18.2	0.18	0.862	4.09
C4					
1	8.3	17.2	0.48	0.648	3.88

Model Summary

S	R^2	R^2 (Adjusted)	R^2 (Predicted)
14.4032	88.35%	80.59%	

Analysis of Variance

Source	Degrees of Freedom	Adjusted Sum of Squares	Adjusted Means Squared	F-Value	P-Value
Regression	4	9441.8	2360.46	11.38	0.006
C1	1	2786.0	2786.04	13.43	0.011
C2	1	1.5	1.49	0.01	0.935
C3	1	6.8	6.81	0.03	0.862
C4	1	47.9	47.94	0.23	0.648
Error	6	1244.7	207.45		
Total	10	10686.5			

Fits and Diagnostics for Unusual Observations

Observations	C5	Fit	Residual	Standard Deviation of Residual
3	115.0	115.0	−0.0	
9	98.0	122.1	24.1	−2.21 R

Example 4: We are using air passengers' data to show an example of forecasting using R. The advantage of using R or Python is that the syntax can be reused with little changes for new data or new cases. Comment lines are prefixed with a # sign.

```
# Linear regression analysis air passenger dataset
ap = AirPassengers

# install forecast
install.packages("forecast")

# Load forecast package
library(forecast)

# structure of the dataset
str(ap)

# view the dataset
View(ap)

# Plotting raw time series data (Figure 2)
plot(ap, main = "Airpassengers Over Time", ylab = "Passengers")

# Creating training dataset
train = window(ap, end = c(1959,12))
```
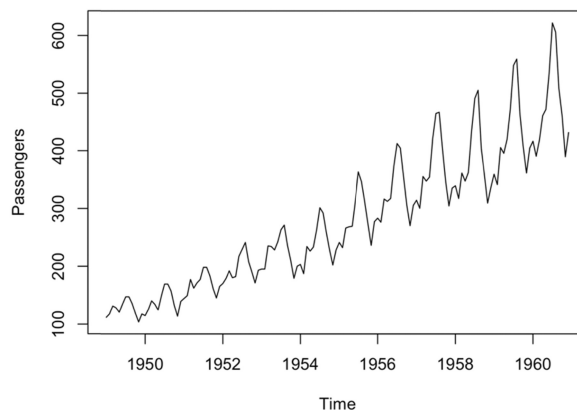

Airpassengers Over Time

Figure 2. Air passengers over time.

```
# start and end date training data
start(train)
end(train)
# Creating test dataset
test = window(ap, start = c(1960,1))
test
length(test)
# Model 1: Additive Trend and seasonality (Figure 3)
# Plotting the complete dataset
plot(ap, main = "Airpassengers Over Time", ylab = "Passengers")
model1 = tslm(train ~ trend + season)
summary(model1)
lines(model1$fitted.values, col = "red", lwd = 2)

# forecast
pred1 = forecast(model1, h = 12)
pred1
lines(pred1$mean, col = "blue", lwd = 2)

# accuracy
accuracy(pred1, test)
# Model 2: Multiplicative Trend and Seasonality (Figure 4)
# Plotting the complete dataset
plot(ap, main = "Airpassengers Over Time", ylab = "Passengers")
model2 = tslm(train ~ trend + season, lambda = 0)
summary(model2)
lines(model2$fitted.values, col = "red", lwd = 2)

# forecast
pred2 = forecast(model2, h = 12)
pred2
lines(pred2$mean, col = "blue", lwd = 2)

# accuracy
accuracy(pred2, test)
```
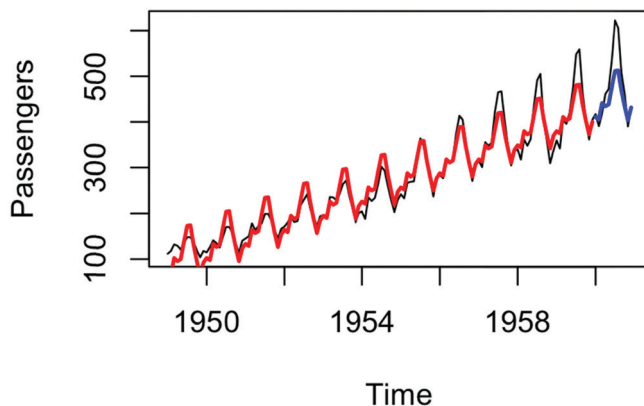
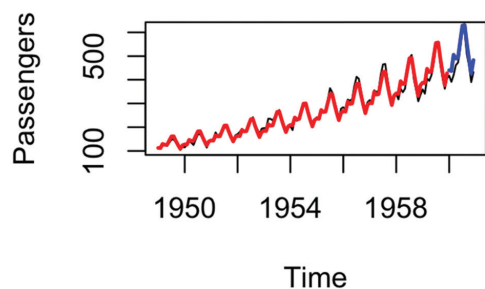**Figure 3.** Additive trend and seasonality.

Figure 4. Multiplicative trend and seasonality.

DISCRIMINANT ANALYSIS

Discriminant analysis is a complete measurement of the difference between a categorical variable and a set of related variables. Essentially, discriminant analysis involves pattern recognition, which is a key to forecasting and prediction [13]. However, there are two extremes when it comes to determinant analysis. The first being at the decision end of the scale, descriptive discriminant analysis intends to understand and predict the factors that determine group separation, and on the basis of a set of variables (entities) for which the group membership is already known. In contrast, the prescriptive discriminant analysis focuses on the classification of future variables (entities) of which the group membership is unknown. However, everyday discriminant analysis lies in between these two extremes, usually having to deal with no assignment or allocation when the groups are known to the researcher [20]. When dealing with discriminant analysis, it is important to note error rates, posterior probabilities, and basic notation when utilizing this forecasting method. There are various familiar techniques such as linear discriminant and quadratic rules that show the

relationship of estimates and mean and covariance parameters [18].

TEACHING FORECASTING AND PREDICTION OVER THE YEARS

The decision-making can be as few as hours, but as many as days or weeks. This is why the need for planning is great in management and administrative situations. Old methods such as time series and ratio methods are classified as "passive" methods with no control [21]. Through the time series, the use of regression and simulation is often "active", and they can also be "passive" at times. However, it should be noted that throughout the last few centuries, predicting has made significant advancements. The advancement of research has expanded our awareness of diverse environmental factors, which has boosted our capacity to forecast many different events. Today, the capacity to foresee a variety of occurrences appears as natural as how accurate weather forecasting will be in a few decades [23]. Despite these upgrades, two crucial points should be remembered. The first is that managers and others may not always immediately benefit from accurate forecasting, and the second is the distinction between external events that are out of one's control and internal ones that one can influence. Both kinds of events are necessary for a business to succeed, but predicting is only relevant to the former while decision-making is only relevant to the latter. The connection that joins them is planning.

EFFICIENT USE OF RESOURCES

The scheduling of production, transportation, money, employees, and other factors is necessary for the efficient utilization of resources. A key component of such scheduling is forecasts of the degree of demand for products, materials, labor, funding, or services. The lead time for purchasing new machinery and equipment, employing new employees, and sourcing raw materials might range from a few days to many years. The initial goal was to identify significant contributions, application areas, constraints, and research and application suggestions for the area's future study. The second objective was to explain control models in terms of production management [4].

Forecasting is crucial to examine the need for future resource requirements. Long-term and short-term forecasting differs among organizations, and they must determine what is used in long term. Such decisions depend on market opportunities, environmental factors, and the internal development of financial, human product, and technological resources. These determinations all require good forecasts and managers who can interpret the predictions and make appropriate decisions. Figure 5 shows the use of
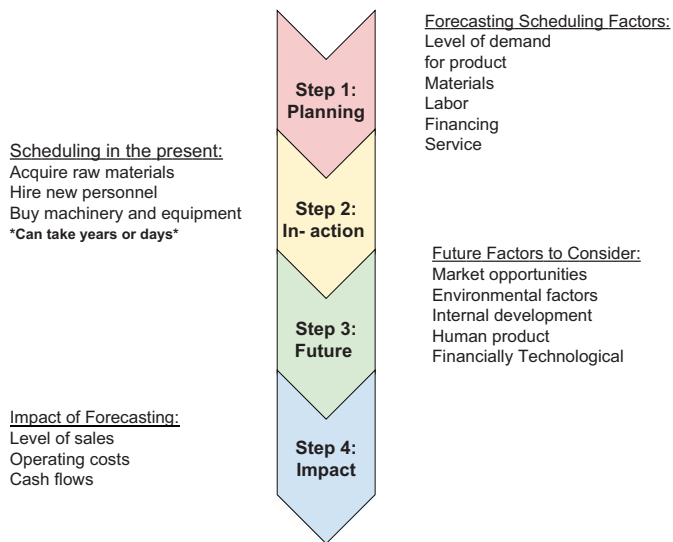
Figure 5. Use of resource flow chart.

the flow of resources from planning, to action, to future, and finally its impact.

Regardless of what problems are being studied, it is important to provide an efficient algorithm or an algorithm with worst-case and performance analyses. Similarly, computational experiments demonstrate that the heuristics developed are capable of generating near-optimal solutions for any problem desired [2]. With new models and new solution strategies, this can be extremely helpful for decision-making and for future research in areas of production scheduling and planning. This is important to support decision makers by improving their resource planning and scheduling system [6]. If forecasting is to be successful, it is essential to recognize the strong dependency that exists between the predictions of different divisions or departments. The following instances may be impacted by errors in sale predictions through a chain of events: operational costs, cash flows, inventory inventories, and price forecasts. Similar budgeting mistakes in estimating the amount of funding available to each division will have an impact on product development, equipment modernization, staff employment, and advertising expenses. In turn, this will have an impact on, if not really dictate, how much money will flow in and out of the business. Clearly, the many forecasting functions within an organization are highly interdependent.

DIFFERENCE BETWEEN ARTIFICIAL INTELLIGENCE AND MACHINE LANGUAGE

Artificial intelligence tries to determine the forecasted demand using the human brain while machine language can also determine the forecast error by using a computer program that functions like a machine. Artificial intelligence tries to develop learning models using the idea of the human brain, in particular, the ideas of neurons and connection, such as perceptions. This is also an area of computer science that emphasizes the creation of intelligent machine language that works and reacts like humans. With this, prediction is faster and cheaper, and it is a fundamental input to decision-making and a relatively newer business strategy. Machine learning is the study of algorithms and statistical models using a computer system to carry out a particular and specific intended activity. Essentially, this can be seen as a subset of artificial intelligence. However, machine learning algorithms build a mathematical model based on sample data in order to make predictions or decisions without being explicitly programmed to perform such a task [10]. This tries to develop learning models primarily based on new optimization methods such as gradient boosting. In particular, machine languages are essentially computer programs that function at the bit/byte level stressing machine components such as logic gates. The goal is to ensure that all computer programs are able to solve problems and achieve success [12]. Machine language or artificial intelligence is utilized to make predictions.

HUMAN ERROR AND SOCIETY IN FORECASTING AND PREDICTION

Humans are naturally worse when it comes to heuristics and biases, large datasets, and through complex interactions when forecasting and predicting. In fact, there is nothing new about tragic accidents being caused by human error. For example, human errors complicate forecasting because natural disasters often happen in the area of human disasters. Now, the nature and the scale of possible hazardous technology mean that human error can have adverse effects around the globe while working with large datasets [19]. However, human errors are better when it comes to small data sets such as children learning to recognize a face and the challenger space shuttle, as these are easy things to remember. In society, forecasting and prediction can be utilized to determine job loss, rise of inequality, involving few large companies, and few countries with an advantage. Regarding the primary cause of defects in forecasting, human error is the key to understanding, also to predicting, and avoiding them. There has been little research done to predict defects in forecasting on the basis of the cognitive errors that cause them [22].

DIVISION OF LABOR: HUMANS VS. MACHINES

Utilizing machines for prediction provides a better, faster, and cheaper prediction. Humans create their own predictions involving strategies such as anticipatory shipping for

companies and organizations. However, there are precautions for both types of divisions of labor.

The division of labor between humans and computer systems has changed for both technology and humans. There has been a shift from technologies of automation with the aim of disallowing human intervention at nearly all points in the system [5]. For machines, it is important to note that there may be biased datasets, small or incomplete datasets, and artificial intelligence for specific tasks. For humans, we have an enhanced value of human judgment. Computers are only capable of solving tasks within practices in which rules are specified by humans or patterns found in datasets. Humans have the capacity to change practices through creativity and innovation, as these processes are effective in nature. The paradox of this is that even though computers can make rapid change possible, human intervention is still a necessity [7].

CONCLUSION

In teaching forecasting and analytics along with prediction, there are many similarities. It is critical that academics understand the content that is currently being taught in these courses and develop the most optimal teaching model. These findings illustrate a current gap between teaching analytics and teaching forecasting with a significant variance among the topics taught across programs.

REFERENCES

[1] Brown, R. G., and Meyer, R. F. (1961). The fundamental theorem of exponential smoothing. *Operations Research*, 9(5), 673–685.

[2] Chen, Z. L., and Vairaktarakis, G. L. (2005). Integrated scheduling of production and distribution operations. *Management Science*, 51(4), 614–628.

[3] Davydenko, A., and Fildes, R. (2016). Forecast error measures: critical review and practical recommendations. *Business forecasting: Practical problems and solutions*, 34.

[4] Dolgui, A., Ivanov, D., Sethi, S. P., and Sokolov, B. (2019). Scheduling in production, supply chain and Industry 4.0 systems by optimal control: fundamentals, state-of-the-art and applications. *International Journal of Production Research*, 57(2), 411–432.

[5] Ekbia, H., and Nardi, B. (2014). Heteromation and its (dis) contents: The invisible division of labor between humans and machines. *First Monday*.

[6] Haase, K. (2012). *Lotsizing and scheduling for production planning* (Vol. 408). Springer Science and Business Media.

[7] Hammershøj, L. G. (2019). The new division of labor between human and machine and its educational implications. *Technology in Society*, 59, 101142.

[8] Hyndman RJ, Khandakar Y (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3), 1–22. doi:10.18637/jss. v027

[9] Hyndman, R., Koehler, A. B., Ord, J. K., and Snyder, R. D. (2008). *Forecasting with exponential smoothing: the state space approach*. Springer Science and Business Media.

[10] Khanzode, K. C. A., and Sarode, R. D. (2020). Advantages and Disadvantages of Artificial Intelligence and Machine Learning: A Literature Review. *International Journal of Library and Information Science (IJLIS)*, 9(1), 3.

[11] Li, G. (1985). Robust regression. *Exploring data tables, trends, and shapes*, 281, U340.

[12] McCarthy, John. "What is artificial intelligence." URL: http://www-formal.stanford.edu/jmc/whatisai.html (2004).

[13] McLachlan, G. J. (2005). *Discriminant analysis and statistical pattern recognition*. John Wiley and Sons.

[14] Merriam-Webster Dictionary. (2022). *forecast*. The Merriam-Webster.Com Dictionary. Retrieved 2022, from https://www.merriam-webster.com/dictionary/forecast

[15] Merriam-Webster Dictionary. (2022). *predict*. The Merriam-Webster.Com Dictionary. Retrieved 2022, from https://www.merriam-webster.com/dictionary/predict

[16] Mishra, S. (2018). Financial management and forecasting using business intelligence and big data analytic tools. *International Journal of Financial Engineering*, 5(02), 1850011.

[17] Montgomery, D. C., Peck, E. A., and Vining, G. G. (2021). *Introduction to linear regression analysis*. John Wiley and Sons.

[18] Rayens, W. S. (1993). Discriminant analysis and statistical pattern recognition.

[19] Reason, J. (1990). *Human error*. Cambridge university press.

[20] Silva, A., and Stam, A. (1995). Discriminant analysis.

[21] Sinuany-Stern, Z. (2021). Forecasting methods in higher education: An overview. *Handbook of Operations Research and Management Science in Higher Education*, 131–157.

[22] Strigini, L., and Huang, F. (2021). HEDP: A Method for Early Forecasting Software Defects based on Human Error Mechanisms.

[23] Suhardi, S., Widyastuti, T., Bisri, B., and Prabowo, W. (2019). Forecasting Analysis Of New Students Acceptance Using Time Series Forecasting Method. *Jurnal Akrab Juara*, 4(5), 10–23.