MUST: ***Journal of Mathematics Education, Science and Technology***
Vol. 9, No. 2, Desember 2024 Hal 76-91
DOI: http://doi.org/10.30651/must.v5i1. 21040

INFLUENCER PRICING PROGNOSTICATION ON SOCIAL MEDIA DYNAMICS: AN ADVANCED EXAMINATION OF THE LINEAR REGRESSION 2 POLYNOMIAL DEGREE ALGORITHM & NEURAL NETWORKS

Felicia Canesta[1], Rusdianto Roestam[2]
[1]President University, Cikarang, Indonesia
[2]President University, Cikarang, Indonesia
Felicia.canesta@student.president.ac.id[1], Rusdianto@president.ac.id[2]

Received 11 Desember 2023; revised 12 November 2024; accepted 18 November 2024

ABSTRAK

Meningkatnya penggunaan media sosial telah menjadikan influencer sebagai pedoman utama dalam promosi produk, dengan tarif yang biasanya ditentukan berdasarkan faktor seperti jumlah pengikut dan *engagement*. Namun, kurangnya metode standar untuk penetapan tarif dapat menimbulkan risiko finansial bagi baik influencer maupun klien. Penelitian ini mengembangkan model prediksi berbasis pembelajaran mesin yang menggabungkan Regresi Linier dengan polinomial *second degree* dan *neural network* untuk meningkatkan akurasi tarif. Evaluasi model menggunakan Mean Absolute Error (MAE) menunjukkan bahwa Neural Network Keras mengungguli Regresi Linier Sederhana (10,612 MAE) dan Regresi Linier dengan polinomial (10,089 MAE), dengan MAE lebih rendah yaitu 7,952. Hal ini menunjukkan kemampuan superior neural network dalam menangkap hubungan data yang kompleks. Sebagai kesimpulan, Neural Network Keras merupakan model yang paling akurat dalam memprediksi tarif influencer, memberikan kerangka kerja yang lebih andal untuk penetapan tarif di lingkungan media sosial yang dinamis.

Kata Kunci: Prediksi Tarif Influencer; Machine Learning; Influencer Media Sosial; Regresi Linier; Neural Network

ABSTRACT

The rise of social media has made influencers key drivers of product promotion, with rates typically based on factors like follower count and engagement. However, the lack of a standardized rate-setting method can lead to financial risks for both influencers and clients. This study develops a machine learning-based predictive model combining Linear Regression with a second-degree polynomial and a neural network to improve rate accuracy. Model evaluation using Mean Absolute Error (MAE) shows the Keras Neural Network outperforms both Simple Linear Regression (10.612 MAE) and Linear Regression with a 2nd-degree polynomial (10.089 MAE), achieving a lower MAE of 7.952. This demonstrates the neural network's superior ability to capture complex data relationships. In conclusion, the Keras Neural Network offers the most accurate model for predicting influencer rates, providing a more reliable framework for rate determination in the dynamic social media landscape.

Keywords: Influencer Rate Prediction; Machine Learning; Social Media Influencers; Linear Regression; Neural Network

INTRODUCTION

The transformative impact of social media on traditional marketing practices has ushered in the era of influencer marketing, where individuals on social platforms wield significant power in shaping audience interest toward promoted products and services. This shift is particularly notable for its enhanced measurability compared to conventional media, making influencer marketing an attractive option for businesses seeking more tangible and data-driven promotional outcomes (Agustian et al., 2023).

The existence of social media has propelled a transformation in marketing activities and promotional strategies, which were previously centered on the utilization of conventional media such as TV, radio, newspapers, and others, gradually becoming more decentralized (Stoldt et al., 2019). This is because each social media user has their potential and segmentation in influencing and sparking the audience's interest.

Influencers offer highly diverse rate cards for marketing products or services through content on social media. Generally, influencer rates follow the number of followers they have (Kamal & Bablu, 2022; Shah & Nasnodkar, 2019). However, the number of followers doesn't always guarantee good performance, as it's possible that follower growth on an influencer's account is built in a non-organic way or through manipulation, utilizing bots and syndication (Atiq et al., 2022). There are metric units that can be used as performance indicators besides the number of followers, namely engagement and reach recorded on the influencer's account. Pricing based on these metric units can be considered to obtain a representative rate reference concerning the performance generated by the influencer (Gerlich, 2023).

In the realm of influencer marketing, the conventional metric for determining pricing has been the number of followers an influencer possesses

(Kamal & Bablu, 2022; Moubayed et al., 2018). However, this approach is increasingly seen as inadequate, as follower counts can be artificially inflated through methods such as bot usage and follower syndication. This overemphasis on follower count fails to account for the quality of engagement or other critical factors influencing influencer effectiveness. Existing models that rely on simplistic metrics, therefore, do not offer a comprehensive or accurate framework for rate determination.

This study addresses this gap by proposing a machine learning-based predictive model that integrates advanced algorithms, including Linear Regression with a second-degree polynomial and a neural network. By leveraging these techniques, the model aims to move beyond surface-level metrics, providing a more reliable and nuanced approach to predicting influencer rates. This research seeks to contribute to a more accurate, data-driven method for influencer pricing, one that accounts for the complexities of influencer performance beyond just follower numbers (Gerlich, 2023).

The inclusion of these sophisticated algorithms allows for a more nuanced understanding of an influencer's impact, surpassing the limitations posed by raw follower count. The utilization of the Linear Regression 2nd-degree Polynomial Algorithm and a Neural Network represents an advanced and sophisticated approach to modeling and predicting influencer rates on social media (Kigo et al., 2023).

In the context of this paper, a neural network is employed to enhance the model's capacity to learn intricate patterns and relationships within the data. Neural networks consist of layers of interconnected nodes (neurons), each layer contributing to the extraction of hierarchical features from the input data. The neural network, in conjunction with Linear Regression, adds a layer of complexity and non-linearity to the model. This is particularly beneficial in scenarios where the relationships between influencer metrics and rates are highly nonlinear and intricate. The neural network excels in capturing these intricate patterns, allowing for a more nuanced understanding of how various factors contribute to the determination of influencer rates. Furthermore, the training of a neural network involves adjusting weights and biases to minimize the difference between predicted and actual values. This adaptive learning process enables the model to adapt to the complexities of the dataset, improving its ability to generalize to new, unseen data (Fan et al., 2021).

This research seeks to pioneer a shift in influencer rate determination by exploring comprehensive metrics, considering factors beyond mere follower counts. The developed model, although constrained by an extremely limited dataset, exhibits promising results with prediction values closely aligning with actual rates, including instances with minimal error differences, such as -347.69. While recognizing the success achieved within these constraints, the study emphasizes the need for further research and dataset expansion to refine and enhance the model's predictive capabilities, ultimately contributing to a more equitable and transparent influencer rate-setting framework in the dynamic landscape of social media marketing.

The combination of Linear Regression with a 2nd-degree Polynomial Algorithm and a Neural Network in this paper represents a cutting-edge approach to influencer rate prediction. These methods collectively enhance the model's capacity to capture nonlinear relationships and intricate patterns within the data, providing a more accurate and nuanced framework for determining influencer rates on social media.

RESEARCH METHOD

The research stages consist of data collection, data preprocessing, linear regression modelling, neural network development, and model evaluation. The study results in three model variants.

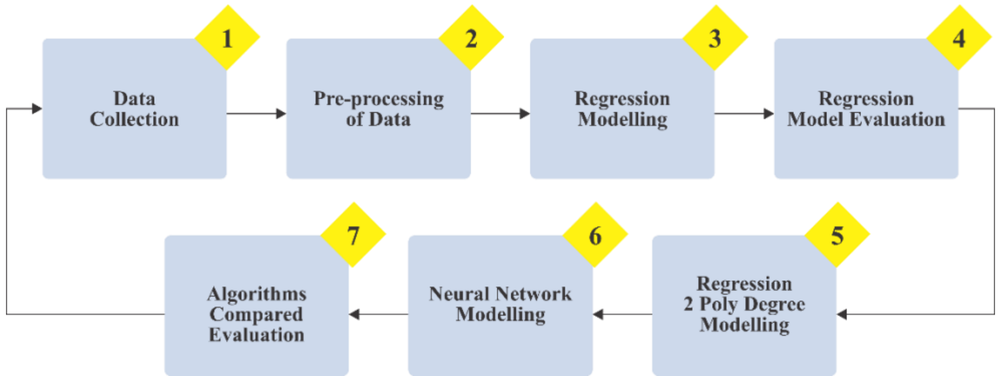


Figure 1. *Research Stages for Predicting Influencer Rates using Machine Learning & Deep Learning*

a. Data Collection

Data collection was carried out by visiting influencer collaboration platforms and conducting web scraping of influencer accounts on social media. The selection of influencer accounts had a minimum criterion of 5,000 followers and were within the any industry sector. The determination of this minimum follower count was based on the focus of classifying influencers at the micro-influencer level, which is less than 15,000 followers, and direct monitoring of several influencers considered to have an impact within the range of 5,000 followers. The collected data included followers, following, likes, posts, views, and rates. Influencers sometimes offer different rates for SMEs (Small and Medium Enterprises) and corporations. The influencer rates used in this research are those for SMEs. The social media platforms that are currently highly favored by influencers and have the potential to capture audience attention for endorsements is TikTok. The dataset obtained was sourced TikTok.

b. Pre-processing of Data

The dataset generated in the previous stage needed preparation to be used for the regression model formation process. Preprocessing in the case of regression typically handles empty or missing data. If possible, missing data can be replaced with the mean value among neighboring data rows. Additionally, because the dataset was obtained through web scraping, it was necessary to ensure that the dataset does not contain unnecessary scripts or special characters.

c. Regression Modelling

The regression model was built using the machine learning algorithm - linear regression using the TensorFlow framework as a tool to obtain predicted influencer rate values (dependent variable), and engagement metrics (such as likes and views), reach, and the number of followers (independent variables) (Guzik et al., 2020). Formula (1) shows the regression formula used, where $\hat{Y}$ represents the predicted influencer rate, X1, X2, X3, and X4 represent the independent variables used, namely likes, views, posts, and followers. b1, b2, b3, and b4 respectively represent the weights or coefficient estimates of each independent variable, and b0 represents the constant value of $\hat{Y}$ when all independent variables are 0. The machine learning algorithm was used to find the values of b0, b1, b2, b3, and b4 that provided the best evaluation results based on the prepared data X1, X2, X3, and X4.

$$\hat{Y} = b0 + b1X1 + b2X2 + b3X3 + b4X4 \quad (1)$$

d. Regression Model Evaluation

The evaluation metrics used to evaluate the generated regression model were Mean Absolute Error (MAE), Mean Square Error (MSE), and Root Mean Square Error (RMSE) to measure the model's performance by quantifying how well the model fits the dataset. These three metrics were directly supported by the TensorFlow framework and could be used during the training process (Musiyiwa & Jacobson, 2023).

Formulas (2), (3), and (4) represent the formulas used in the model evaluation in this study.

$$MAE = 1\ n \sum |yi\ real - yi\ pred|\ n\ i=1 \quad (2)$$

$$MSE = 1\ n \sum (yi\ real - yi\ pred)\ 2\ n\ i=1 \quad (3)$$

$$RMSE = \sqrt{MSE} = \sqrt{}\ 1\ n \sum (yi\ real - yi\ pred) \quad (4)$$

e. Neural Network Development Model

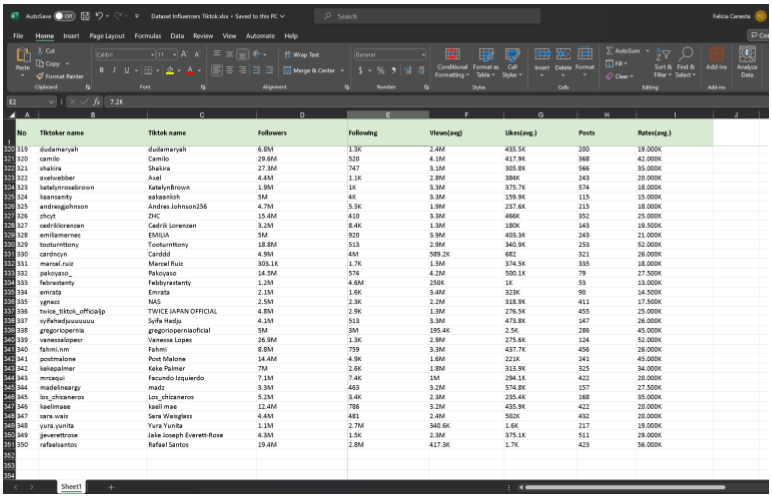
The application of neural networks offers a sophisticated and adaptable framework to uncover complex patterns and relationships within the multifaceted dynamics of influencer marketing on social media platforms. This process aims to enhance the generalization capabilities of the neural network. As the model undergoes validation and optimization, it holds the potential to uncover hidden insights, enabling businesses to make data-driven decisions and optimize their influencer collaborations and pricing strategies (Wang et al., 2021).

RESEARCH RESULTS AND DISCUSSIONS

The research results encompassed the determination of a dataset that aligns with the supported format for the model formation process. Preprocessing of data resulted in consistent format and normalized values. The development of the regression model yielded three prediction models corresponding to various model formation scenarios. Model evaluation produced measurement values for all models, which were then tested against the selected model based on the best performance.

A. Dataset

Determining the criteria for the number of followers and attribute selection can impact the dataset size to be used. In this research, the dataset used was sourced from social media, totaling 350 rows. This dataset was then stored in CSV (Comma-Separated Values) file format to facilitate the data loading process when starting to build the machine learning model.



No	Tiktoker name	Tiktok name	Followers	Following	Views(avg)	Likes(avg.)	Posts	Rates(avg.)
319	dudamaryah	dudamaryah	6.8M	1.3K	2.4M	435.5K	200	19.000K
320	camilo	Camilo	29.6M	520	4.1M	417.9K	368	42.000K
321	shakira	Shakira	27.3M	747	3.1M	305.8K	566	35.000K
322	axelwebber	Axel	4.4M	1.1K	2.8M	384K	243	20.000K
323	katelynrosebrown	KatelynBrown	1.9M	1K	3.3M	375.7K	574	18.000K
324	kaansanity	aakaanksh	5M	4K	3.3M	159.9K	115	15.000K
325	andresgjohnson	Andres Johnson256	4.7M	5.5K	1.9M	237.6K	215	18.000K
326	zhcyt	ZHC	15.4M	410	3.3M	466K	352	25.000K
327	cedriklorenzen	Cedrik Lorenzen	3.2M	8.4K	1.3M	180K	143	19.500K
328	emiliamernes	EMILIA	5M	920	3.9M	403.3K	243	21.000K
329	tooturnttony	Tooturnttony	18.8M	513	2.9M	340.9K	253	52.000K
330	cardncyn	Carddd	4.9M	4M	589.2K	682	321	26.000K
331	marcel.ruiz	Marcel Ruiz	303.1K	1.7K	1.5M	374.5K	335	18.000K
332	pakoyaso_	Pakoyaso	14.5M	574	4.2M	500.1K	79	27.500K
333	febrastanty	Febbyrastanty	1.2M	4.6M	250K	1K	53	13.000K
334	emrata	Emrata	2.1M	1.6K	3.4M	323K	90	14.500K
335	ygnazz	NAS	2.5M	2.3K	2.2M	318.9K	411	17.500K
336	twice_tiktok_officialjp	TWICE JAPAN OFFICIAL	4.8M	2.9K	1.3M	276.5K	455	25.000K
337	syifahadjuuuuuuu	Syifa Hadju	4.1M	513	3.3M	473.8K	147	26.000K
338	gregoriopernia	gregoriopernia oficial	5M	3M	195.4K	2.5K	286	45.000K
339	vanessalopesr	Vanessa Lopes	26.9M	1.3K	2.9M	275.6K	124	52.000K
340	fahmi.nm	Fahmi	8.8M	759	3.3M	437.7K	456	26.000K
341	postmalone	Post Malone	14.4M	4.9K	1.6M	221K	241	45.000K
342	kekepalmer	Keke Palmer	7M	2.6K	1.8M	313.9K	325	34.000K
343	mrcaqui	Facundo Izquierdo	7.1M	7.4K	1M	294.1K	422	20.000K
344	madelineargy	madz	3.3M	463	3.2M	574.8K	157	27.500K
345	los_chicaneros	Los_chicaneros	5.2M	3.4K	2.3M	235.4K	168	35.000K
346	kaelimaee	kaeli mae	12.4M	786	3.2M	435.9K	422	20.000K
347	sara.wais	Sara Waisglass	4.4M	481	2.4M	502K	432	20.000K
348	yura.yunita	Yura Yunita	1.1M	2.7M	340.6K	1.6K	217	19.000K
349	jjeverettrose	Jake Joseph Everett-Rose	4.3M	1.5K	2.3M	375.1K	511	29.000K
350	rafaelsantos	Rafael Santos	19.4M	2.8M	417.3K	1.7K	423	56.000K

***Figure 2.** TikTok Influencer Dataset Before Preprocessing*

B. Pre-processing of Data

After the data preprocessing is performed, the prepared dataset can be summarized in the form of descriptive statistics as shown in Figure 3. In Figure 3, the displayed statistics can summarize data characteristics, ranging from central tendencies, dispersion, to the shape of the data distribution. It's essential to ensure that the dataset does not contain NaN or null values. If data with no values are left, it will undoubtedly impact the quality of the model formed since the quality of the dataset used for training fundamentally influences the model's quality.

NO	Attributes	Classification
1	Rate	Dependent
2	Following	Independent

3	Followers	Independent
4	Likes	Independent
5	Posts	Independent
6	Views	Independent
7	TikTok Name	Independent
8	Influencer Name	Independent

Table 1. *Attributes Classification*

In Table 1, it is shown that there is only one attribute, that is the "target" or label whose value will be predicted, which is the rate. The value of the rate attribute will then be influenced by other attributes, namely followers, following, likes, posts, and views.

C. Regression Modelling

In this stage, the Tensorflow framework begins to be utilized to generate the influencer prediction model. Tensorflow is utilized through a cloud computing service, Google Colab, using the Python programming language. Because it is based on Python, the libraries used include: tensorflow, matplotlib.pyplot, numpy, pandas, seaborn, and drive. Before commencing model development, it's more effective to first examine the characteristics of the dataset.

In Figure 5, the data distribution is visible, revealing the correlation between attributes. In the first row, the rate attribute is highlighted as it is the target or dependent variable, compared against other attributes. It's evident at a glance that the data distribution of the rate attribute correlates quite well with only 3 attributes: likes, posts, and views. Meanwhile, the followers and following attributes tend to be nearly vertical.

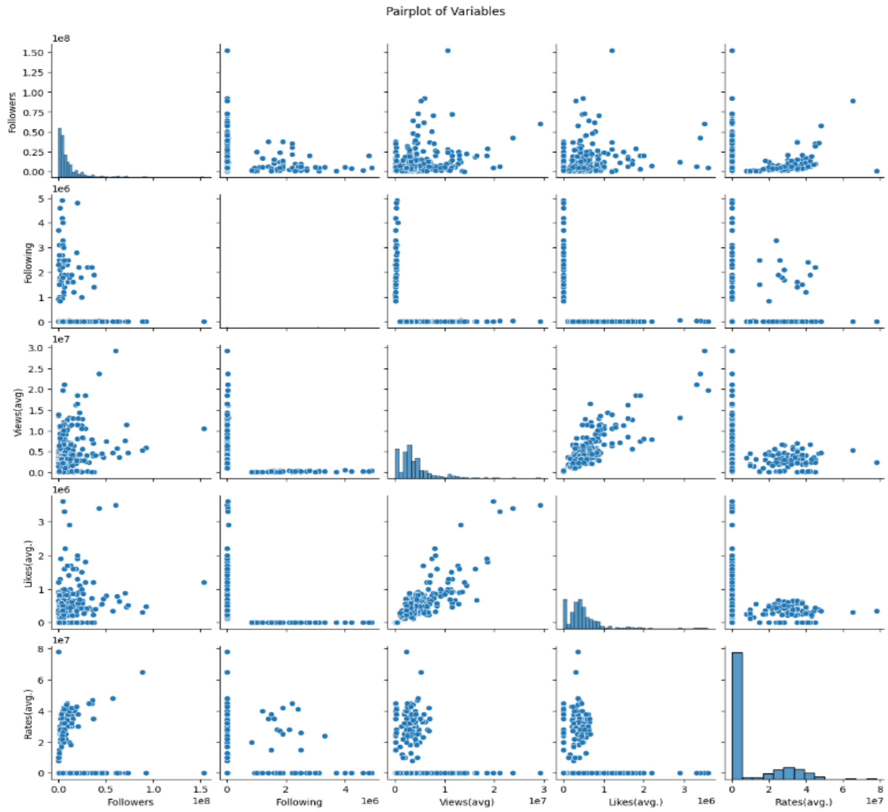


Figure 3. *Visualization of Correlations Among Attributes in the Dataset using a Pairplot*

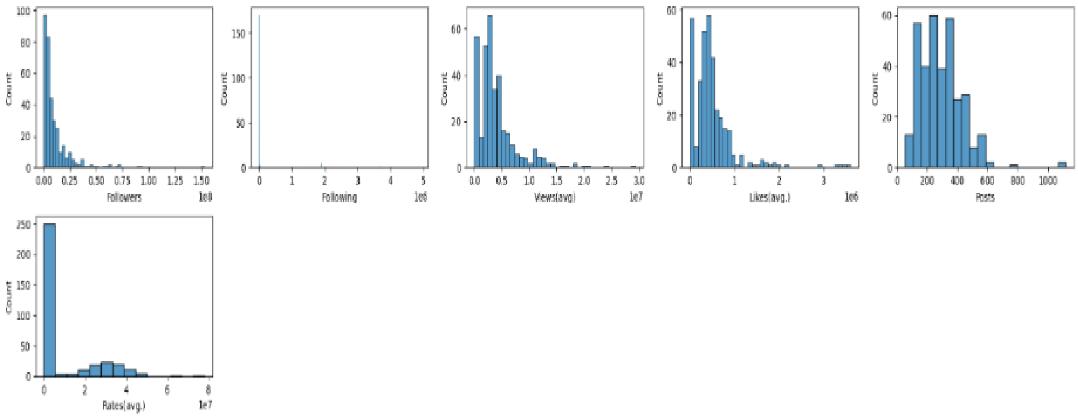


Figure 4. *Data Spread Histogram*

Further regarding the relationship or correlation of attributes in the dataset, it can be quantified in the form of a correlation heatmap as shown in Figure 7. The quantification shown is based on a scale from -1.00 to 1.00. The lower the value or the darker the color produced, the lower the level of correlation.

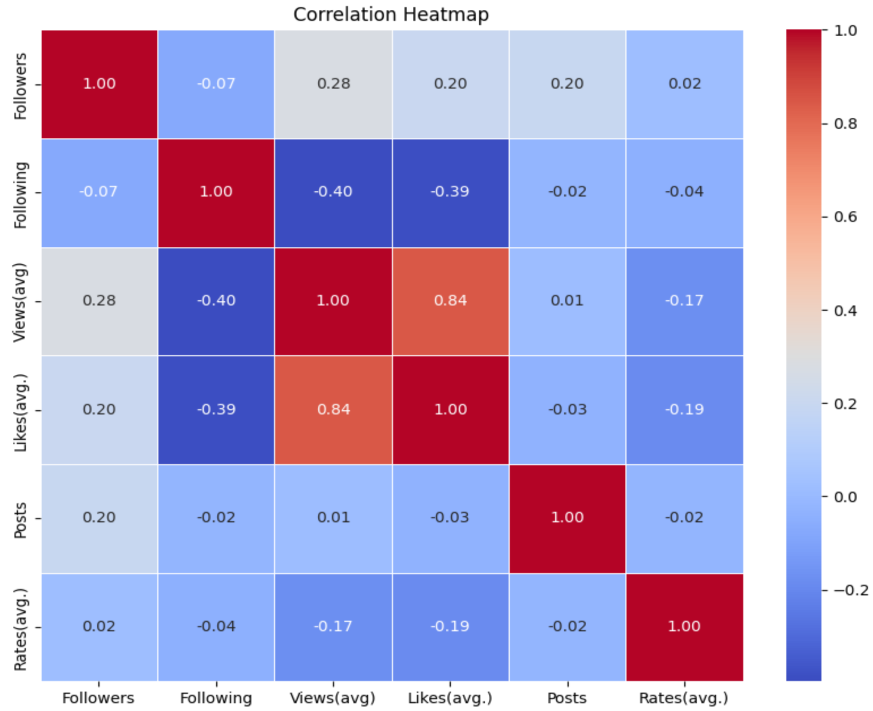


Figure 5. *Visualization of Correlation Among Attributes in the Dataset using Correlation Heatmap*

```
model = LinearRegression()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
print('Performance Linear :\n')
print(performance(model,X_train,y_train,y_pred,y_test))
print('')

X_train_poly = X_train
X_test_poly = X_test
quad_feature = PolynomialFeatures(degree = 2)
X_train_quad = quad_feature.fit_transform(X_train_poly)
X_test_quad = quad_feature.fit_transform(X_test_poly)

model.fit(X_train_quad, y_train)

y_pred_train_poly = model.predict(X_train_quad)
y_pred_test_poly = model.predict(X_test_quad)
print('Performance Linear with poly 2 degree:\n')
print(performance(model,X_train_quad,y_train,y_pred_test_poly,y_test))
print('')
```

Figure 7*. Testing of Linear Regression with 2 Poly Degree*

```
Performance Linear :

Train Score: 0.0473
Test Score (R2 Score): 0.0163
Metrics In Testing Data:
MSE: 1.0854
RMSE: 1.0418
MAE: 0.8294118365512717
Max Error: 4.582420638527174
MAPE: 101.31 %
None

Performance Linear with poly 2 degree:

Train Score: 0.1156
Test Score (R2 Score): -0.3641
Metrics In Testing Data:
MSE: 1.5051
RMSE: 1.2268
MAE: 0.9082914411783731
Max Error: 4.994402803964605
MAPE: 115.53 %
None
```

Figure 8. *Testing Result of Simple Linear & Linear Regression with 2 Poly Degree*

Performance simple Linear result metrics in testing data shows that it obtained the Train Score: 0.0473, Test Score (R2 Score): 0.0163, MSE: 1.0854, RMSE: 1.0418, MAE: 0.8294118365512717, Max Error: 4.582420638527174, MAPE: 101.31 %. Otherwise, the performance Linear with poly 2 degree obtained the Train Score: 0.1156, Test Score (R2 Score): -0.3641, MSE: 1.5051, RMSE: 1.2268, MAE: 0.9082914411783731, Max Error: 4.994402803964605.

D. Modelling Development with Deep Learning (Neural Network)

In the pursuit of enhancing the predictive capabilities for influencer pricing, a Neural Network Development Model is employed. The following Python code snippet demonstrates the implementation of the neural network, utilizing the Keras library with a Sequential model architecture. The evaluation results offer a comprehensive assessment of the neural network's effectiveness in predicting influencer rates. The Test MAE of 0.8165 in the scaled space signifies the model's ability to, on average, predict influencer rates with a relatively low absolute error. The Test Loss of 1.2878 provides a broader measure of the model's performance, considering the mean squared error. In translating these results to real-world applicability, the Test MAE (Original Scale) of approximately 12,273,655.61

becomes particularly meaningful. This metric signifies the average absolute difference between the model's predictions and actual influencer rates in their original, interpretable context.

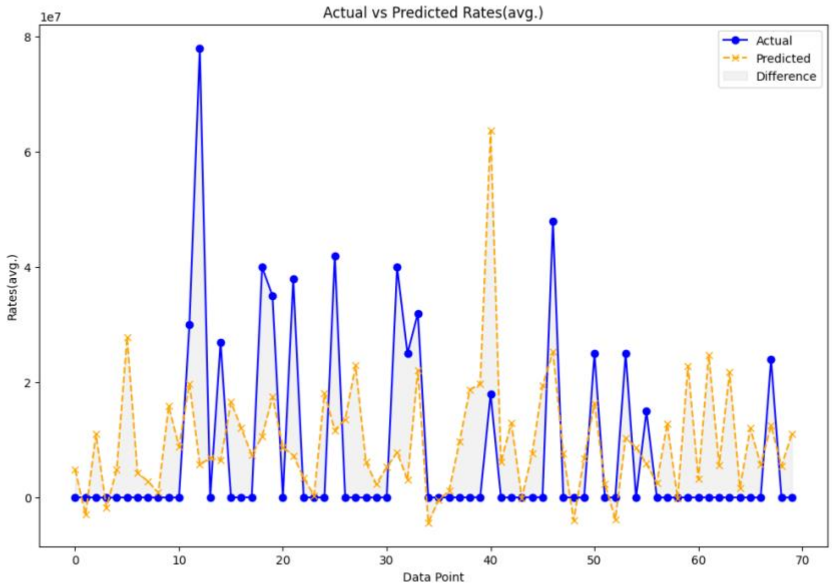


***Figure 9**. Actual data result & prediction with KERAS*

```
[35] # actual data result, prediction and differences for the top three data
     for i in range(3):
         actual_value = y_test_original_scale[i]
         predicted_value = predictions_original_scale[i][0]
         difference = predicted_value - actual_value

         print(f'Data Point {i+1}: Actual={actual_value}, Predicted={predicted_value}, Difference={difference}')

    Data Point 1: Actual=[28000.], Predicted=4847973.5, Difference=[4819973.5]
    Data Point 2: Actual=[50000.], Predicted=-2955380.75, Difference=[-3005380.75]
    Data Point 3: Actual=[14500.], Predicted=11012249.0, Difference=[10997749.]
```

***Figure 10**. Actual data result & prediction with KERAS*

These results highlight instances where the model struggled to accurately predict influencer rates, demonstrating both overprediction and underprediction. The magnitude of differences, especially in Data Points 1 and 3, underscores the challenges in achieving precise predictions.

E. Model Evaluation Analysis

The Model Evaluation conclusion is derived from comparing the Mean Absolute Error (MAE) metrics for three different models: Simple Linear Regression, Linear Regression with a 2nd-degree polynomial, and a Keras Neural

Network. The MAE provides a measure of the average absolute difference between the actual and predicted values. Lower MAE values indicate better model performance.

- Simple Linear Regression

  The linear Regression model yielded a mean absolute error of 10.612. This indicates that, on average, the model's predictions deviated by approximately 10.612 units from the actual influencer rates.
- Linear Regression with 2nd-degree Polynomial

  The Linear Regression model with a 2nd-degree polynomial performed slightly better than the Simple Linear Regression model, with a mean absolute error of 10.089. This suggests a modest improvement in predictive accuracy.
- Keras Neural Network

  The Keras Neural Network outperformed both the Simple Linear Regression and the Linear Regression with a 2nd-degree polynomial, exhibiting a lower mean absolute error of 7.952. This indicates superior predictive accuracy compared to the other models.

The results show that the Keras Neural Network outperforms the other models, achieving the lowest Mean Absolute Error (MAE) of 7.952, indicating superior predictive accuracy. This finding aligns with previous research suggesting that neural networks, with their ability to capture complex and non-linear data relationships, provide enhanced accuracy over traditional statistical methods (Kamal & Bablu, 2022; Moubayed et al., 2018). The neural network's ability to learn from intricate patterns in the data, particularly in large and non-linear datasets, was a key factor in its superior performance (LeCun et al., 2015).

In contrast, the Linear Regression model with a second-degree polynomial produced a better result than the Simple Linear Regression model, as indicated by its MAE of 10.089 compared to 10.612 for the simple model. The introduction of non-linear terms allowed the second-degree polynomial model to capture more complex relationships in the data, confirming the importance of considering non-linear patterns when predicting influencer rates (Kuhn & Johnson, 2013). This result is consistent with prior studies highlighting the limitations of linear models in capturing the full complexity of influencer pricing (Cheng & Ho, 2020).

Thus, based on the MAE values, the Keras Neural Network emerges as the most effective model for predicting influencer rates, offering a more accurate and nuanced approach compared to both Simple Linear Regression and Linear Regression with a 2nd-degree polynomial. These findings reinforce the growing consensus in the literature that machine learning techniques, particularly neural networks, are more suited for predictive tasks in dynamic, non-linear domains like influencer marketing (Yuan et al., 2021).

CONCLUSIONS

This research endeavors to tackle this pressing challenge by pioneering the development of an advanced machine learning-based predictive model. The model integrates sophisticated techniques, including Linear Regression with a second-degree polynomial algorithm and a neural network, aimed at elevating prediction accuracy to new heights. This study underscores the transformative potential of machine learning, showcasing the prowess of advanced regression algorithms and neural networks in constructing a robust framework for predicting influencer rates. The resultant model represents a significant stride toward minimizing adverse effects on both influencers and clients by furnishing a nuanced and precise reference for rate determination within the dynamic realm of social media promotion. The research methodology, marked by its comprehensiveness, encompasses meticulous data collection and preprocessing, coupled with the integration of linear regression models with varying polynomial degrees and a sophisticated neural network. The Model Evaluation, employing Mean Absolute Error (MAE) metrics, unequivocally establishes the superiority of the Keras Neural Network, exhibiting a remarkable MAE of 7.952 compared to Simple Linear Regression (MAE: 10.612) and Linear Regression with a 2nd-degree polynomial (MAE: 10.089). While this research presents a promising machine learning-based approach for predicting influencer rates, there are several limitations that should be addressed in future research. This study primarily focused on quantitative factors such as follower count and engagement, which may not fully capture the qualitative aspects of influencer influence, such as content quality, authenticity, and niche expertise. Future work could integrate sentiment analysis, content evaluation, or influencer reputation metrics to create a more comprehensive prediction model that considers both quantitative and qualitative dimensions. As the social media landscape evolves rapidly, future research should explore the impact of emerging trends such as algorithmic changes, the rise of micro- and nano-influencers, and the growing importance of short-form video content. Incorporating these dynamic factors could lead to more adaptive models that reflect real-time shifts in the influencer marketing industry.

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to all people who supported me throughout the research process and the preparation of this research. First and foremost, I would like to thank President University for their generous support, which made this research possible. I also wish to acknowledge the invaluable assistance of my research mentor, Rusdianto Roestam, whose expertise and guidance were essential to the success of this project and insightful feedbacks, which greatly improved the quality of the research.

REFFERENCES

Agustian, K., Hidayat, R., Zen, A., Sekarini, R. A., & Malik, A. J. (2023). The Influence of Influencer Marketing in Increasing Brand Awareness and Sales for SMEs. *Technology and Society Perspectives (TACIT)*, *1*(2), 68–78. https://doi.org/10.61100/tacit.v1i2.54

Atiq, M., Abid, G., Anwar, A., & Ijaz, M. F. (2022). Influencer Marketing on Instagram: A Sequential Mediation Model of Storytelling Content and Audience Engagement via Relatability and Trust. *Information (Switzerland)*, *13*(7). https://doi.org/10.3390/info13070345

Fan, F. L., Xiong, J., Li, M., & Wang, G. (2021). On Interpretability of Artificial Neural Networks: A Survey. *IEEE Transactions on Radiation and Plasma Medical Sciences*, *5*(6), 741–760. https://doi.org/10.1109/TRPMS.2021.3066428

Gerlich, M. (2023). The Power of Virtual Influencers: Impact on Consumer Behaviour and Attitudes in the Age of AI. *Administrative Sciences*, *13*(8). https://doi.org/10.3390/admsci13080178

Guzik, T. J., Mohiddin, S. A., Dimarco, A., Patel, V., Savvatis, K., Marelli-Berg, F. M., Madhur, M. S., Tomaszewski, M., Maffia, P., D'Acquisto, F., Nicklin, S. A., Marian, A. J., Nosalski, R., Murray, E. C., Guzik, B., Berry, C., Touyz, R. M., Kreutz, R., Dao, W. W., … McInnes, I. B. (2020). COVID-19 and the cardiovascular system: Implications for risk assessment, diagnosis, and treatment options. In *Cardiovascular Research* (Vol. 116, Issue 10, pp. 1666–1687). Oxford University Press. https://doi.org/10.1093/cvr/cvaa106

Kamal, M., & Bablu, T. A. (2022). International Journal of Applied Machine Learning and Computational Intelligence Machine Learning Models for Predicting Click-through Rates on social media: Factors and Performance Analysis. *International Journal of Applied Machine Learning and Computational Intelligence*, *4*(11), 1–14.

Kigo, S. N., Omondi, E. O., & Omolo, B. O. (2023). Assessing predictive performance of supervised machine learning algorithms for a diamond pricing model. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-44326-w

Moubayed, A., Injadat, M., Nassif, A. B., Lutfiyya, H., & Shami, A. (2018). E-Learning: Challenges and Research Opportunities Using Machine Learning Data Analytics. *IEEE Access*, *6*, 39117–39138. https://doi.org/10.1109/ACCESS.2018.2851790

Musiyiwa, R., & Jacobson, J. (2023). Sponsorship Disclosure in Social Media Influencer Marketing: The Algorithmic and Non-Algorithmic Barriers. *Social Media and Society*, *9*(3). https://doi.org/10.1177/20563051231196870

Shah, A., & Nasnodkar, S. (2019). A Framework for Micro-Influencer Selection in Pet Product Marketing Using Social Media Performance Metrics and Natural Language Processing. *Journal of Computational Social Dynamics Research Article: Journal of Computational Social Dynamics*, *07*(01), 7.

Stoldt, R., Wellman, M., Ekdale, B., & Tully, M. (2019). Professionalizing and Profiting: The Rise of Intermediaries in the Social Media Influencer Industry. *Social Media and Society*, *5*(1). https://doi.org/10.1177/2056305119832587

Wang, X., Wang, Y., Tao, F., & Liu, A. (2021). New Paradigm of Data-Driven Smart Customisation through Digital Twin. *Journal of Manufacturing Systems*, *58*, 270–280. https://doi.org/10.1016/j.jmsy.2020.07.023