A Comparison of Machine Learning Algorithms for Brand Engagement Insight through Comment Volume Prediction Analysis

Hazel Mae P. Casey[1], Audrey Lauren G. Chua[1], Ashley Mikhaela B. Tan[1], Adrienne Joyce W. Kosca[1], Neil Oliver Velasco[1], Arvin Fernando[1], and Marielet Guillermo[1,*]
[1] *Gokongwei College of Engineering, De La Salle University*
**Corresponding Author: marielet.guillermo@dlsu.edu.ph*

Abstract: Social media engagement has gained popularity among businesses in its products and services campaign in pursuit of generating brand perception and gaining wider and targeted customer reach. Given the tight competition among long-established companies and start-up businesses, especially in its online presence, there is a necessity to synthesize strategies that would entice and warrant the target audience to connect and interact with posts in order to land a better position than the rest. Hence, targeting the most active social networking platform where this principle is applied to (Facebook's Pages), user interaction through comments were examined and analyzed through supervised regression algorithms, namely Linear Regression, Random Forest, Decision Tree, and Gradient Boosting. With the proposed machine learning computations, attributes or features that influence a user's likelihood of commenting on a post were identified in addition to the determination of the most accurate model to predict comment volume. As a result, an appropriate strategy may be created in order to fully maximize audience engagement with the brand's posts.

Key Words: brand engagement; machine learning; prediction analysis; supervised regression algorithms; user interaction insights

1. INTRODUCTION

Significant development and shift in social media presence paved new means for companies and brands to communicate with current and potential clients from its targeted market. According to the University of Massachusetts Dartmouth's Center for Marketing Research (Barnes et al, 2020), Inc., Magazine's top 500 companies use at least one online platform to interact with their customers to be able to further enhance service quality and brand perception thereafter (Merdiaty et al, 2022). To optimize engagement with the audience given its limited time and energy (Dai et al, 2021), it is indispensable to carry out appropriate strategies such as knowing its demographics, and narrowing down to the interests of relevant market. Generating posts that would warrant target audience interaction poses difficulty for both the brand and the business. There must be some mechanism to predict a client's perceived action to a corresponding post.

A social media platform wherein this may be applied to is Facebook, and more specifically, through its Pages feature since it is the most used feature within Facebook for brands and businesses. The researchers proposed the use of machine learning to predict how a certain client would engage with a post under various factors such as the popularity and category of the page, and the length of the post. The engagement was limited to comments a user would make on a post. Based on this, the following research questions were formed: (1) which features greatly affect audience engagement, and (2) which machine learning approach yields higher accuracy in making predictions. In this study, the volume of comments on a post is predicted using supervised regression. The specific algorithms utilized are Linear Regression (LR), Decision Tree (DT), Random Forest (RF), and Gradient Boosting.

1.1 Social Media Marketing Strategies

In the past few years, social media marketing, along with social media influencers, have become more prominent. It is especially advantageous for small businesses as social media marketing is more affordable compared to the traditional forms of digital advertising. Regardless of the platform to be used, it is important to first establish the business intention, its corresponding social media goal, and target audience before marketing (Newberry et al, 2024), (Siani et al, 2020). Afterwards, several marketing strategies can be applied to ensure a successful social media branding. The first strategy is to utilize chat bots since it can directly communicate with the user to answer questions or resolve problems. The next strategy is to personalize the user experience based on their past inputs. The third strategy is to provide quality content without sacrificing audience impact. While sticking to social media goals is important, having a variety of content is crucial to keep the audience interested. The last strategy is to engage and interact with the users through community posts and live streaming (Newlands, 2017).

1.2 User Interaction

User interaction can be any form of interaction between the user and a website or application. For websites, the conversation is made through the clicks that lead to their designated results, which may be other related pages or links. Similarly, applications such as social media provide notifications to inform users when a button is clicked or when an action is done (Perez, 2020). In other words, the user dictates the flow of the conversation through the provided interface in the web or mobile applications.

User engagement is particularly vital to social media platforms because its primary goal is to provide a convenient way for people to build connections around the world. Social media applications have feedback avenues such as the like or reaction buttons, and the comments section. This way, the company can gain insight on how the application is received by its users. Although negative feedback is valuable for the improvement of the application, positive feedback is also essential as it can establish assurance and loyalty between the users and the company (Perez, 2020).

1.3 Machine Learning in Social Media

Machine learning is a form of data analysis algorithm that trains computers to make predictions based on the given data (Tobias et al, 2020)(Fernando et al, 2021). It is often used as a marketing analytics tool in social media. Popular social media applications such as Facebook, Instagram, and Twitter use built-in machine learning algorithms to analyze (Yasay et al, 2023), monitor, and enhance the users' experience. One feature that affects user experience is the use of chatbots, from which the algorithm will organize a feed that will mostly display the favored information (Gotter, 2021).

Other social media platforms also use machine learning algorithms to determine what type of content has more audience impact through the number of likes and the comments. Furthermore, the algorithm is beneficial to companies since it can filter out unwanted contents or alert them when there is negative feedback. With that said, any social media platform equipped with machine learning will be able to determine which posts will be more relevant by predicting the number of comments for each post.

2. METHODOLOGY

This study dealt with a multivariate dataset consisting of 53 independent variables and 1 dependent variable which is the Target Variable (Singh, 2016). These are summarized as attributes in Table 1. Jupyter Notebook was used as the coding and computing platform to implement supervised regression. The software application was chosen to be used for its compatibility across different programming languages and operating systems. It also has its independent runtime environment making it more fit for extensive experiments without affecting the system processes of the physical computing machine.

Table 1. Attributes in the dataset

No.	Name	Description
1	Page popularity/ likes	This defines the popularity or support for the source of the document.
2	Page check-ins	This describes how many individuals so far visited this place.
3	Page talking about	This defines the daily interest of individuals towards a post. Examples of this interest are comments, likes, etc.
4	Page category	This defines the category of the source of the post (ex. place, institution, business, etc.)
5-29	Derived	These features are aggregated by page, by calculating min, max, average, median and standard deviation of essential features.

30	CC1	The total number of comments before selected base date/time.
31	CC2	The number of comments in the last 24 hours, relative to base date/time.
32	CC3	The number of comments in the last 48 to last 24 hours relative to base date/time.
33	CC4	The number of comments in the first 24 hours after the publication of a post but before the base date/time.
34	CC5	The difference between CC2 and CC3.
35	Base time	Selected time to simulate the scenario.
36	Post length	The character count in the post.
37	Post share count	The number of shares of the post
38	Post promotion status	This defines whether the post is promoted (1) or not (0). In other words, whether an individual has shared the page to his/her page.
39	H local	This describes the H hrs, for which we have the target variable/ comments received.
40-46	Post published weekday	This represents the day (Sunday – Saturday) on which the post was published.
47-53	Base date/ time weekday	This represents the day (Sunday – Saturday) on a selected base date/time.
54	Target variable	The number of comments in the next H hours.

2.1 Experimental Design

The Facebook pages collected through the source underwent crawling to create entries as the data sets. The subsequent steps to implement the prediction modelling using supervised regression are as follows: (1) data cleaning and handling to check for null values, outliers and duplicates, (2) data processing to check the correlation between variables, (3) data allocation to split the data into train and test sets, (4) feature selection to isolate and find the most relevant features for the construction of the model, (5) supervised regression model training, (6) optimization of hyperparameters turning to obtain the most accurate results possible, and (7) testing and evaluation. Each of these steps is explained further in the next subsections.

2.2 Modules and Libraries

In order to utilize various functions that would aid in the next stages of predicting the comment volume through multiple regression algorithms, the following modules and libraries were imported:

1. OS - This module is one of Python's standard utility modules. This was used to get functions for interacting with the operating system such as the file system. This is characterized by the portable usage of an operating system-dependent functionality. More specifically, OS is commonly used for handling current working directory, creating directories, listing out files, as well as deleting files and directories (Doorwar, 2024).
2. NumPy - Numerical Python is a free, open-source Python library, however, it is also partially written in C or C++ for areas in need of fast computation. It is used for handling arrays and for providing functions in the area of linear algebra, Fourier transform, and matrices. The significance of this library is the alternative it provides to the traditional python lists which are processed slowly. NumPy offers an array object that can be processed 50x faster. Arrays are very useful, especially in data science, where the resources must be manipulated with speed.
3. Pandas - A powerful and flexible quantitative data analysis tool. This library is built from and acts as a wrapper of two other libraries namely NumPy for mathematical operations and matplotlib for data visualization. This allows methods to be done with less code. Pandas provides two new object types for storing data – Series and DataFrames with list-like and tabular structures respectively.
4. Sklearn - An open-source library that is mainly characterized by its machine learning package that supports both supervised and unsupervised learning. Its other functions include tools for model development, selection, evaluation, and data pre-processing. It is built on other libraries namely NumPy, SciPy and matplotlib.
5. Math - This module provides mathematical functions that conforms with the C standard. However, the functions will not work on complex numbers. It comes with the standard Python release and is often used for its predefined constants as well as calculating or solving combinations, permutations, trigonometric functions, exponential functions, hyperbolic functions, and quadratic equations. This also provides functions for simulating periodic functions using trigonometric functions (Myrianthous, 2021).
6. Matplotlib - This library enables the creation of dynamic or animated, static, and interactive plots or visualizations necessary for the study. Its pyplot submodule allows the visualization of 2D plots of arrays that makes data digestible (Liyanapathirana, L. (n.d.)).

7. Seaborn - This is much like Matplotlib which is a visualization library that is closely integrated with Pandas and is built on top of the former library. This also helps in visualizing data such as heatmaps.
8. Collections - This module provides various types of containers which are objects with the purpose of storing other objects that can be accessed and iterated over. Tuple, List and Dictionary are a few of its built-in containers. This module provides other containers namely Counters, OrderedDict, DefaultDict, ChainMap, NamedTuple, DeQue, UserDict, UserList, and UserString (Katari, 2020).
9. Time - This is a Python module that offers multiple ways of representing time in code. It also provides functions for delays in executing code and measuring code efficiency. Python time can be in seconds as a floating-point number or as a string representing local time (Aggarwal, 2024). In this study, this module was used in the optimization part where the total duration of each regression algorithm was provided.

2.3 Dataset Cleaning and Handling

The dataset used in predicting the comment volume is adapted from (Singh, 2016). It contains different files for the training and testing dataset. The training dataset contains 5 data while the testing dataset has 10. After importing the datasets to the program, data were subjected to cleaning. This was to make sure that there are no missing values or outliers in the dataset. However, this step could be skipped because based on the team's observation, there were little to no outliers in the dataset as seen in Fig. 1 and Fig. 2. In Fig. 1, only duplicating values were removed. On the other hand, Fig 2 presents the data in a box plot where the duplicates and outliers were removed from the dataset.

2.4 Data Processing

Fig. 3 depicts the correlation of the variables with one another. To establish the correlation better, attributes with a high correlation value against the target variable (53) were plotted using a scatter matrix in Fig. 4 and the whole dataset is represented in the form of a heatmap. Only, the last column, 53, was considered as the target variable. From this, it can be seen that the top attributes with the highest correlation values are 30, 33, 11, 6, 16, 21, and 7. Its corresponding descriptions are summarized in Table 2.

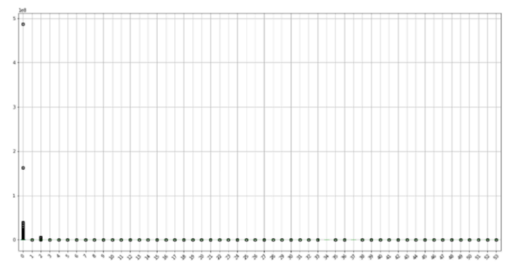

Fig. 1. Box plot of the dataset

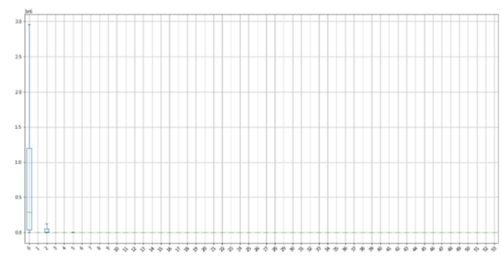

Fig. 2. Box plot of the dataset (cleaned version)

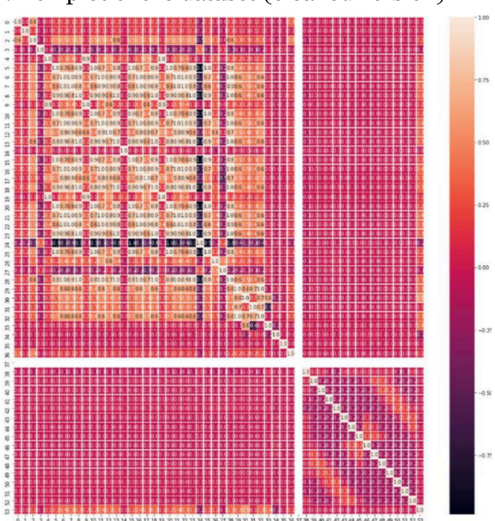

Fig. 3. Correlation heatmap for the entire dataset

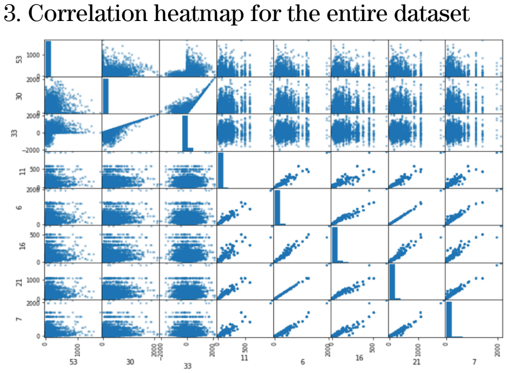

Fig. 4. Scatter plots for the top correlated attributes

Table 2. Top correlated attributes

No.	Name	Description	Correlation Value
30	CC2	The number of comments in the last 24 hours, relative to base date/time.	0.545260
33	CC5	The difference between CC2 and CC3.	0.375318
11	Derived	These features are aggregated by page, by calculating min, max, average, median and standard deviation of essential features.	0.368003
6			0.359536
16			0.356985
21			0.356769
7			0.349452

2.5 Train/Test

The training and testing sets contain more than one file. To easily call a corresponding set, the five training datasets were concatenated and split into the variable X_train_list and y_train_list. Likewise, the other 10 testing datasets were split into the X_test and y_test.

2.6 Feature Selection

Feature selection was executed by removing the features that have the same value in the training set. When running the feature selection cell, it would show that the number of features used was 52 while the number of features ignored was only 1.

2.7 Supervised Regression Algorithms

Each of the algorithms were implemented using 6 variants – variants 1-5 and variants overall. The models were imported using the sklearn library. Root mean squared error for every variant-algorithm pair was generated.

1. Linear regression (LR) - Since there is more than one independent variable, multiple linear regression was implemented. The estimated regression function is f(x1, x2) = b0 + b1x1 + ⋯ + brxr where the bn variables are the values of the weights and when the number of inputs is r, there are r + 1 weights to be determined (Fernando et al, 2015). LinearRegression() was used to store the model into a variable. The parameters were left to its default settings.

2. Decision Tree (DT) - Decision trees, which can also be used in classification, is a non-parametric supervised learning method. This aims to generate a model which can perform target variable value prediction through learning decision rules from the independent variables. The fit of the model increases when the depth of the tree increases (Stojiljković, M. (n.d.). DecisionTreeRegressor() was used to store the model into a variable. The parameters that were specified and not left to default settings were max_depth (12) and random_state (42).

3. Random Forest (RF) - Random forest is a meta estimator that fits a number of classifying decision trees on multiple sub-samples of the dataset. To improve the predictive accuracy and control-overfitting, it uses the averaging method. If bootstrap is in default (True), the max_samples parameters determine the sub-sample size. If bootstrap is set to False, the whole dataset is utilized to generate each tree (Pedregosa et al, 2011). RandomForestregressor() was used to store the model into a variable. The parameters were set to the following values: max_depth = 8, random_state = 0, and n_estimators = 100.

4. Gradient Boosting - This algorithm is used to build an additive model through a forward stage-wise function which permits the optimization of arbitrary differentiable loss functions. For each stage, the negative gradient of the specified loss function has a fitted regression tree (Pedregosa et al, 2011). GradientBoostingRegressor() was used to store the model into a variable. The parameters not left to its default settings are max_depth (8), n_estimators (3), learning_rate (1.0), and random_state (42).

2.8 Cross Validation and Hold Out

This process was done to evaluate the algorithms performed in the study. This was performed to test the accuracy of the classification models built. According to (Gil et al, 2000), cross validation evaluates and compares the learning algorithm that had been divided into two. The first divided part will be used to train the model while the other part will be used to validate the model. This will be done in successive rounds so that each data point will have the chance to be validated against. On the other hand, holdout is used when the dataset is split into a training and testing set. The testing set can also be called the holdout set and is the out-of-sample data. This is because holdout has the approach of an out-of-sample evaluation.

For this study, the function grid_search was implemented so that it is easier to show the cross-validation results for each algorithm. Holdout was also implemented in the study. It was done in order to evaluate how well the models performed. It would give the results for the mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), explained variance score, and the R2 error. For explained variance score and R2 error, the ideal result computed for the algorithms is 1.

2.9 Optimization

The researchers also did optimization on the four algorithms to see which gives the least RMSE result along with the time duration that it took the software to produce the result. The least RMSE was obtained because having a low RMSE means that the model has a better fit. The lower the value, the better fit the model is (Martin, (n.d.)). In this section, the optimization was only done using the overall variant of X_train_list and y_train_list. Additionally, each of the regression models were limited to two parameters not left to the default settings. Linear regression parameters were normalized and n_jobs. Decision Trees parameters were max_depth and max_features. Random Forest parameters were n_estimattos and max_depth. Lastly, Gradient Boosting parameters were n_estimators and learning_rate. Each of the parameters were given three possible settings except for those that require True or False.

3. RESULTS AND DISCUSSION

The results obtained from running the four regression models are summarized in Tables 3 to 5. The proponents focused on obtaining the RMSE from each algorithm because it is a good way to evaluate how accurate the algorithm's prediction is since the residuals would provide a good measurement of how far the data points are from the regression line. This is the most critical criterion to take note of for fit especially if the key objective of the algorithm is the prediction (Martin, (n.d.)), (Statistics How To, (n.d.)). Table 3 contains the RMSE result obtained through cross validation by calling the function grid_search(model). It could be observed that Random Forest or RF (57.804) had the best result while Gradient Boosting or GB (82.901) had the worst result. On the other hand, the RMSE of RF (57.804) and Decision Tree or DT (60.529) are somewhat close to each other. This means that the algorithm DT had the second-best result while Linear Regression or LR (69.009) had the third best result. The researchers note, however, that the time of completion for RF was substantially longer than that of other tested models.

Table 3. Cross validation (Grid search)

Model	RMSE
LR	69.009
DT	60.529
RF	57.804
GB	82.901

Table 4 consists of the results (MAE, MSE, RMSE, Variance, and R2 Error) computed for each algorithm which was obtained from holdout validation. It could be noticed again that for RMSE, it had the same results as in Table 3. This means that RF had the best result while GB had the worst result. This could be further proven by the other metric results such as variance and R2 Error. The ideal result for variance and R2 Error is 1 which means the nearer the result is to 1, the better. From the table, it could be seen that RF has, once again, the best results among the algorithms as it had the highest value for variance and R2 Error. On the other hand, DT had the second-best result, LR had the third best result, and GB had the worst result.

Table 4. Hold-out validation

Model	MAE	MSE	RMSE	Variance	R2 Error
LR	26.6616	4762.2815	69.0093	0.6789	0.6675
DT	23.7825	3663.7876	60.5292	0.7478	0.7442
RF	21.9991	3341.3025	57.804	0.7712	0.7667
GB	28.8127	6872.5391	82.9008	0.5277	0.5201

Table 5 contains the least RMSE obtained through optimization and its corresponding parameter settings. As mentioned in the optimization part of the paper, the lower the value of RMSE, the better. From this, it could be observed again that RF had the best result while LR had the worst. DT had the second best while GB had the third best.

Table 5. Cross validation (Optimized)

Model	Least RMSE	Parameter Settings
LR	28.099541	n_jobs: 3, normalize: True
DT	21.438601	max_depth: 7, max_features: auto
RF	17.395831	max_depth: 12, n_estimators: 20

GB	22.778222	learning_rate: 0.6, n_estimators: 3

In summary, it could be observed that RF (57.804 and 17.396 for optimized) had the most accurate prediction among the rest of the algorithms implemented based on the RMSE value while GB (82.901) had the least accurate prediction both in cross and hold-out validation. in Tables 3 and 4. In the cross validation of optimized form, LR (28.1) had the least accurate prediction. Relating this back to marketing strategy and the posed questions in the beginning, the target variable is most affected by the attributes listed in Table 2, where the number of comments in the last 24 hours, relative to base date/time influences the audience's probability of engaging with the post the most. Furthermore, correlation was low with the post promotion status, post published weekday and base date/time weekday attributes. Hence, it can be deduced that to increase comment engagement, attributes such as the post's and the page's features as well as time that it is posted should be given attention in promoting a post. Once rectified, it is best to use the Random Forest regression model in the prediction of the volume of comments given its level of accuracy among the other three models tested.

4. CONCLUSIONS

The identification of the attributes that influence the behavior of consumers, specifically their comment activity was successfully identified through data processing. In addition, the determination of which model would be able to provide the most accurate prediction of comment volume was also accomplished through the usage of supervised regression algorithms. Based on the results, Random Forest was the most accurate model given that it had the lowest RMSE, which is a significant component for prediction, compared to Linear Regression, Decision Tree and Gradient Boosting for both cross validation and holdout methods. This was further proven by hyperparameters tuning, where RF retained its position as the most accurate model with lower RMSE than that of former methods despite the long execution time. Overall, this study can be utilized to create strategies to ensure engagement with current and future consumers of a brand. As a result, it can increase brand experience, perception, and loyalty.

5. ACKNOWLEDGMENTS

The researchers would like to thank Engr. Sandy Lauguico, Dr. Renann Baldovino, and Dr. Marielet Guillermo for their guidance as mentors in LBYMF3B - Computational Intelligence in Manufacturing System Laboratory subject. De La Salle University is hereby acknowledged as well for being instrumental in providing the necessary resources for the conduct of the study.

6. REFERENCES

Aggarwal, N. (2024, Mar 27). *Python Collections Module*. From https://www.geeksforgeeks.org/python-collections-module/

Barnes, N. G., Mazzola, A., Killeen, M., & Khalil, N. (2020, Feb. 11). *Increases in Use of Instagram and Paid Social But Not Social Planning: Social Media Among the 2019 Inc. 500.* Retrieved from https://www.umassd.edu/cmr/research/2019-inc-500.html

Dai, Y., & Wang, T. (2021). Prediction of customer engagement behavior response to marketing posts based on machine learning. *Connection Science 33*(4), 891–910. https://doi.org/10.1080/09540091.2021.1912710

Doorwar, P. (2024, Mar. 11). *OS Module in Python with Examples*. From https://www.geeksforgeeks.org/os-module-python-examples/

Fernando, A., & GanLim, L. (2021, March). Velocity analysis of a six wheel modular mobile robot using MATLAB-Simulink. In IOP Conference Series: Materials Science and Engineering (Vol. 1109, No. 1, p. 012037). IOP Publishing.

Fernando, A. H., Marfori, I. A. V., & Maglaya, A. B. (2015, December). A comparative study between artificial neural network and linear regression for optimizing a hinged blade cross axis turbine. In 2015 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM) (pp. 1-4). IEEE.

Gil', M. I., & Ailon, A. (2000). On input-output stability of nonlinear retarded systems. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 47*(5), 761-763. https://doi.org/10.1109/81.847882

Gotter, A. (2021, July 22). *Should I Be Using Social Media Chatbots?*. From https://www.agorapulse.com/blog/social-media-chatbots/

Grace-Martin, K. (n.d.). *beyond r-squared: assessing the fit of regression models*. From https://www.theanalysisfactor.com/assessing-the-fit-of-regression-models/

Katari, Kaushnik (2020, Aug. 12). *Seaborn: Python*. From https://towardsdatascience.com/seaborn-python-8563c3d0ad41

Liyanapathirana, L. (n.d.). *The Python math Module: Everything You Need to Know*. From https://realpython.com/python-math-module/

Merdiaty, N., & Aldrin, N. (2022). Effect of Brand Experience on Customer Engagement Through Quality Services of Online Sellers to Students in Bekasi. *Frontiers in Psychology, 12*, 6073. https://doi.org/10.3389/FPSYG.2021.801439

Myrianthous, G. (2021, Aug 29). *Is There Any Difference Between Scikit-Learn and Sklearn?*. From https://towardsdatascience.com/scikit-learn-vs-sklearn-6944b9dc1736

Newberry, C., & Wood, A. (2024, March 13). *How to Create a Social Media Strategy in 9 Steps* (Free Template). From https://blog.hootsuite.com/how-to-create-a-social-media-marketing-plan/

Newlands, M. (2017, May 1).*10 Social-Media Marketing Strategies for Companies*. Retrieved February 22, 2022, from https://www.entrepreneur.com/article/293321

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12,* 2825–2830.

Perez, M. (2020, December 11). *4 Ways to Increase The User Interaction on Your Website*. Retrieved February 22, 2022, from https://www.momentcrm.com/blog/increase-user-interaction-website/

Siani, T. (2020, Dec. 10). *Social Media Marketing: Be Smarter By Using Machine Learning*. Retrieved February 22, 2022, from https://www.clicdata.com/blog/social-media-marketing-using-machine-learning/

Singh, Kamaljot. (2016). Facebook Comment Volume. *UCI Machine Learning Repository*. https://doi.org/10.24432/C5Q886.

Statistics How To, (n.d.). *RMSE: Root Mean Square Error*. From https://www.statisticshowto.com/probability-and-statistics/regression-analysis/rmse-root-mean-square-error/

Stojiljković, M. (n.d.). *Linear Regression in Python*. From https://realpython.com/linear-regression-in-python/#multiple-linear-regression

Tobias, R. R., De Jesus, L. C., Mital, M. E., Lauguico, S., Guillermo, M., Sybingco, E., ... & Dadios, E. (2020, March). Android Application for Chest X-ray Health Classification From a CNN Deep Learning TensorFlow Model. In 2020 IEEE 2nd Global Conference on Life Sciences and Technologies (LifeTech) (pp. 255-259). IEEE.

Yasay, B., Francisco, K., Velasco, N. O., Guillermo, M., & Bandala, A. (2023, May). Fuzzy Logic Adaptive Bus Route Determination for Intelligent Transportation System of Buses along Epifanio Delos Santos Avenue (EDSA). In 2023 8th International Conference on Business and Industrial Research (ICBIR) (pp. 1014-1019). IEEE.