International Journal of Global Perspective in Academic Research

Journal homepage: https://ijgpar.org/index.php/journal/index

Predicting Audience Awareness on Instagram by Using Linear Regression

Sonam Jha, Vikas Chaurasia,*

VBS Purvanchal University, India

Abstract: As far as audiences are concerned, Instagram is quickly rising to the top of the social media ranks. The primary cause of these important elements is the brand's engagement rate, creativity, user-friendliness, interactiveness, and visualization. In order to predict an Instagram user's engagement percentage, a variety of user interaction-related data points must be analyzed. These include collecting information about interacting users, posting time, hashtags, and captions; computing the user's engagement rate; segmenting the data; descriptive statistics; predictive modeling; machine learning algorithms; feature selection; model validation; metrics; and continuous improvement. Depending on their objectives, Instagram users can profit from a range of tools that the platform offers. A metric called engagement rate is used to gauge how much audience involvement a piece of information gets, usually on social media sites. To compute the percentage, divide the total engagements (likes, comments, shares, etc.) by the total reach or impressions and multiply the result by 100. For example, if a post has 100 engagements and a reach of 1,000, the engagement rate would be (100/1000) * 100 = 10%. Brands and content producers can evaluate how well their audience is engaging with their material with the use of this statistic. Increased engagement rates typically signify that the audience finds the content to be compelling. The most effective method for predicting actual value on Instagram datasets is the linear regression model. Interpreting and understanding it is simple. It is simple to visualize the relationship between the variables because it is represented as a straight line. It is faster to train and forecast than more sophisticated models since it uses less processing power. Certain assumptions made by linear regression—such as linearity, independence, and error normality—can be verified and checked with ease.It offers a useful starting point for forecasting, particularly in cases where correlations are roughly linear. The relationship between changes in the predictor variables and the response variable is directly revealed by the coefficients. In order to simplify the model, feature selection might assist in determining which variables are significant predictors. When its assumptions are not fully met, linear regression can often perform remarkably well—as long as the breaches are not too serious.

Keywords: Instagram, Linear Regression, Instagram Tools, Engagement Rate.

1 Introduction

As is well known, Instagram is one of the most widely used social media platforms, with users communicating with one another more than on other networks, particularly among younger users[1]. Instagram measurement is the process of calculating user engagement based on interactive material uploaded to Instagram. It takes into account information like the number of our followers as well as interactions like likes, comments, saves, and shares. This number gives us insight into how well our content connects with readers[2]. Like, remark, share, answer, and save the story. Next, compute using the following formula:

$$Engagement\ rate = (Total\ interactions/Impressions) \times 100$$

The study concludes that various factors influence Instagram engagement, including the audience, high-performing posts, generating engagement through stories, short videos, and reels, promoting on multiple networks, managing Instagram ads, captions, and hashtags strategically, publishing content at the optimal moment, and

boosting Instagram engagement. Additionally, the Instagram algorithm can expand the platform's reach, draw in new users, and boost reputation. Since the follower mill offers "fame for sale," it is particularly problematic for Instagram because it is possible to purchase likes, comments, and followers[3]. In line with the query of whether or not we can rely on engagement analytics and social media data. Companies are motivated to create a thorough social media strategy by the low cost and ease of use of social media platforms, which offer a creative yet accessible means of interacting and connecting with their target audience[4]. Even more powerful are non-traditional superstars like blggers, YouTube stars, and "Instafamous" profiles[5]. Engagement indicators, such as likes, comments, and views, have become addicting since receiving praise from others is so fulfilling, especially for those with poor self-esteem[6]. Nonetheless, the quantity of people reacting—liking and commenting on postings on social media, for instance—is frequently used to gauge engagement[7]. Given that Instagram is becoming more and more popular than Facebook, we also wanted to find out what the public thought of some of the inspiring and instructional postings. We noticed that young people tended to gravitate toward inspirational content. Even so, measures like likes, comments, and shares influence the audience engagement rate. Thus, in order to learn about the actual prediction, we used machine learning algorithms and Python code.

Instagram's Popularity among audience can be attributed to several key factors:

- Visual Content: The platform emphasizes photos and videos, which resonate well with users who prefer visual story telling.
- User-Friendly Interface: Its user-friendly layout makes it simple to navigate and interact with content.
- Social Interaction: Instagram facilitates connections through likes, comments and direct messaging, fostering a sense of community.
- Influencer Culture: Many users are drawn to influencers who shares relatable content, trends and lifestyles, impacting their purchasing decisions and social behaviours.
- Trends and Fads: Features like stories and reels keep the content fresh and encourage users to engage regularly with new trends.
- Creative Expression:The platform allows users to express themselves through various creative tools, filters and editing options.
- Brand Engagement: Brands often engage with audiences through targeted marketing and interactive content, making the platform a space for discovering new products.
- Data Description: There are four fields that is used to identify response of each audience on each post.

When people see an Instagram post they like, they usually interact with it by liking and commenting[8]. Some may even share it with others or save it to their collection. These engagements can tell us if our Instagram strategy is working. As a result, we need to closely monitor our Instagram engagement rate to see how we are performing.

2 How We Evaluated?

We gathered a dataset comprising independent variables as input features and dependent variables as output features. Subsequently, the data will be split into training and test sets so that the model's performance can be assessed afterward. We utilized Python resources and selected a linear regression model.

The Instagram engagement rate is a measure that provides insight into how users engage with our content. It considers information such as the quantity of followers in addition to interactions like Like, Accounts Reached, and Non Followers[9]. This indicator aids in assessing how well our material connects with the intended audience. As shown in table 1, we also require the following two details in order to compute the Instagram engagement rate: "Post_Type" and "Post_Date".

Table 1 Instagram engagement metrics

	Followers	Non_Followers	Accounts_Reached	Like
count	367.000000	367.000000	367.000000	367.000000
mean	34.623978	34.038147	48.893733	17.746594
std	48.244089	40.008724	69.939882	10.919156
min	0.000000	0.000000	0.000000	0.000000
25%	0.000000	3.000000	2.000000	10.000000
50%	4.000000	14.000000	6.000000	15.000000
75%	74.000000	63.000000	109.500000	23.000000
max	177.000000	205.000000	328.000000	67.000000

	Post_Type	Post_Date
0	Post_Type	Post_Date
1	Motivational Reel	May 29, 2024
2	Motivational Reel	May 8, 2024
3	Motivational Reel	May 5, 2024
4	Motivational Quotes	May 3, 2024
..	...	...
363	Educational Topic	September 18, 2019
364	Educational Topic	September 9, 2019
365	Educational Topic	September 5,2019
366	Educational Topic	September 1, 2019
367	Educational Topic	August 28, 2019

3 Instagram Analytics Tools

3.1 Sprout Social

Businesses and organizations can manage their social media presence with the aid of Sprout Social, a well-known social media management tool. These are some of the main attributes and advantages of Sprout Social:

1. Social Media Scheduling: Posts can be planned for several social media networks, enabling users to distribute material regularly and on time.

2. Analytics and Reporting: Sprout Social offers comprehensive data to gauge a social media campaign's effectiveness. Users are able to monitor reach, engagement, and other crucial metrics.

3. Social Listening: The platform allows users to monitor social conversations and trends relevant to their brand, helping to inform content strategy and customer engagement.

4. Customer Engagement: Sprout Social facilitates interaction with followers through its unified inbox, enabling users to manage messages, comments, and mentions from different platforms in one place.

5. Collaboration Tools: Teams can collaborate on social media strategies, content creation, and approvals, making it easier to work together efficiently.

6. Content Planning: The platform includes features for content planning and management, helping teams to streamline their content calendar and ensure alignment with marketing goals.

7. Integration: Sprout Social integrates with various other tools and platforms, including CRM systems, to enhance overall marketing efforts.

8. Audience Targeting: Users can analyze their audience demographics and tailor content to better resonate with their target market.

9. Campaign Management: Sprout Social allows for the creation and management of social media campaigns, tracking their performance and adjusting strategies as needed.

3.2 Iconosquare

Iconosquare is a tool for social media management and analytics that focuses mostly on Facebook and Instagram. These are some of the main attributes and advantages of Iconosquare:

1. In-Depth Analytics: Iconosquare offers comprehensive analysis of your social media success, encompassing post-performance analysis, follower growth, and engagement data. This aids companies in determining what kind of information appeals to their target market.

2. Scheduling and Publishing: Users can schedule and publish posts directly to Instagram and Facebook, ensuring a consistent posting schedule without needing to be online at all times.

3. Content Planning: The platform includes tools for visual content planning, allowing users to see how their feed will look before posts go live. This is especially useful for maintaining a cohesive brand aesthetic.

4. Hashtag Tracking: Iconosquare offers hashtag analytics to help users identify effective hashtags for increasing reach and engagement.

5. Competitor Analysis: Users can monitor competitor performance and strategies, providing insights that can inform their own social media tactics.

6. Engagement Tools: The platform includes features to help users respond to comments and messages efficiently, fostering better engagement with their audience.

7. Reports and Dashboards: Iconosquare allows users to create customizable reports and dashboards, making it easy to share insights with team members or stakeholders.

8. Instagram Stories Analytics: Users can track the performance of Instagram Stories, including reach and completion rates, to better understand how these temporary posts engage followers.

9. Mobile App: With the help of Iconosquare's mobile app, users can handle their social media accounts while on the go.

10. Team Collaboration: The tool facilitates group collaboration, enabling users to collaborate on social media content and strategy.

3.3 Union Metrics

Union Metrics is a platform for social media analytics and

monitoring that helps businesses and brands gauge their involvement and performance on social media[10]. These are some of Union Metrics' main attributes and advantages:

1. Social Media Analytics: Union Metrics provides in-depth analytics for platforms like Twitter, Instagram, and Facebook. Users can track engagement, reach, and follower growth, helping them understand the effectiveness of their social media strategies.

2. Real-Time Monitoring: The tool helps organizations monitor discussions about their brand or sector by providing real-time tracking of brand mentions and related keywords.

3. Competitive Analysis: Users can analyze competitors' social media performance, identifying strengths and weaknesses that can inform their own strategies.

4. Hashtag Analytics: Union Metrics provides insights into hashtag performance, helping users find the best hashtags to increase reach and engagement.

5. Content Performance: The platform allows users to assess which types of content perform best, enabling data-driven decisions for future posts.

6. Reporting Features: Union Metrics includes customizable reporting tools that allow users to create and share reports on their social media performance with team members or stakeholders.

7. Campaign Tracking: Users can track specific campaigns or initiatives, measuring their impact and effectiveness over time.

8. Engagement Metrics: The platform helps measure various engagement metrics, including likes, shares, retweets, and comments, providing a comprehensive view of audience interaction.

9. Social Listening: Union Metrics enables brands to listen to their audience by monitoring sentiment and trends, which can be crucial for shaping content and marketing strategies.

10. User-Friendly Interface: Users without substantial expertise in analytics can easily utilize the software due to its easy design.

3.4 Bitly

Bitly is a popular tool for shortening URLs so that users can share them on other platforms[11]. The following are some of Bitly's main attributes and advantages:

1. URL Shortening: With Bitly, users can easily share long URLs by shortening them into manageable links, especially on Twitter and other platforms where character limits apply.

2. Custom Links: By creating unique, branded short links, users may improve click-through rates and brand recognition.

3. Tracking and Analytics: For every shortened link, Bitly offers comprehensive data that include click-through rates, user location, and sources of referral traffic. Users benefit from having a better understanding of their audience and the success of their initiatives.

4. Link Management: Bitly provides tools for managing and organizing links, which simplifies the tracking of many campaigns and their outcomes in a single location.

5. Integration with Other Tools: Bitly may be easily tracked and managed inside current workflows because it interfaces with a number of marketing and analytics systems.

6. QR Code Generation: Bitly can generate QR codes for shortened links, providing a quick and easy way for users to access content via mobile devices.

7. Link Retargeting: Bitly provides options for retargeting users who click on links, enabling businesses to reconnect with their audience through targeted advertising.

8. Branded Domains: Users can set up their own branded domain for link shortening, which can enhance brand trust and recognition.

9. API Access: Bitly offers an API for developers, allowing for custom integrations and advanced functionality in applications.

10. User-Friendly Interface: Because of the platform's ease of use, users with different degrees of technical ability can use it. Marketers, social media managers, and anybody else wishing to gain insight and expedite their link-sharing process will find Bitly especially helpful.

3.5 Keyhole

Keyhole is a social media analytics and monitoring platform made to assist companies and brands in keeping tabs on their performance and online presence[12]. These are a few of its main attributes and advantages:

1. Social Media Monitoring: Keyhole gives consumers the ability to track brand mentions, hashtags, and keywords in real time on a variety of social media networks. This aids companies in remaining up to date on discussions pertaining to their industry or brand.

2. Hashtag Analytics: Users can examine the

effectiveness of particular hashtags, as well as their reach, engagement metrics, and user demographics when dealing with them.

3. Campaign Tracking: Keyhole gives users the ability to monitor the effectiveness of particular marketing initiatives over time, giving them insights into metrics like reach and engagement.

4. Competitor Analysis: Users can compare their social media performance against competitors, gaining insights into strengths and weaknesses that can inform their strategies.

5. Customizable Reports: Keyhole provides the ability to create and share customizable reports, making it easy for teams to present data and insights to stakeholders.

6. Influencer Tracking: The platform can help identify key influencers in your industry, allowing brands to engage with them effectively and potentially collaborate on marketing efforts.

7. Sentiment Analysis: Keyhole offers sentiment analysis tools to assess the overall sentiment of conversations about your brand, helping to gauge public perception and customer sentiment.

8. Visual Analytics: The platform includes visual dashboards that present data in an easily digestible format, allowing users to quickly grasp key insights and trends.

9. Integration: Keyhole can integrate with other marketing tools and platforms, allowing for a more comprehensive view of marketing performance.

10. User-Friendly Interface: Because of its straightforward design, people with different degrees of technical competence can utilize the platform. Keyhole is particularly useful for marketers, social media managers, and brands looking to enhance their social media strategies through data-driven insights[13].

4 Instagram Insights

Within the Instagram app for corporate accounts is a potent analytics tool called Instagram Insights. It provides useful analytics and statistics that customers can utilize to better understand their audience and develop their content strategy[14]. These are some of Instagram Insights' main attributes and advantages:

4.1 Key Features:

1. Audience Insights:

Demographics: View information about your followers, including age, gender, location, and language.

Active Times: See when your audience is most active, helping you choose optimal posting times.

2. Content Performance:

Post Insights: Analyze the performance of individual posts, including likes, comments, shares, saves, and reach.

Story Insights: Track metrics for Instagram Stories, such as views, replies, and exits, to gauge engagement.

3. Engagement Metrics:

Measure overall engagement rates, including interactions and engagement relative to follower count.

4. Impressions and Reach:

Monitor the quantity of impressions (the number of times your posts were shown) and reach (the quantity of unique users that read your posts).

5. Website Clicks:

If you have a link in your bio, you can see how many users clicked on it through Insights, helping to measure traffic driven from your Instagram account.

6. Promotions:

Analyze the performance of any paid promotions or ads, tracking engagement, reach, and conversion metrics.

7. Content Type Comparison:

Evaluate how well various content formats (pictures, videos, carousels) perform to determine which ones your audience responds to the most.

4.2 Benefits:

1. Data-Driven Decisions: Insights help users make informed decisions about content strategies, posting times, and audience engagement tactics.

2. Improved Engagement: By understanding what content performs best, users can tailor their posts to maximize engagement and reach.

3. Audience Understanding: Knowing your audience's demographics and behaviors allows for more targeted marketing efforts and content creation.

4. Campaign Tracking: Insights enable users to track the performance of specific campaigns, making it easier to evaluate success and make adjustments.

5. Competitive Edge: By regularly reviewing Insights, users can stay ahead of trends and adapt their strategies to meet audience expectations.

A vital tool for companies and artists trying to maximize their visibility on Instagram is Instagram

Insights[15].

5 Procedures

To fit model – The model was fitted using training data, and it will eventually determine how the independent and dependent variables relate to one another.

To make predictions - When the model is trained, add fresh data points to make it predict values.

Evaluate the model – Measuring the model's performance via metrics such as Mean Absolute Error, Mean Squared Error, or R-squared.

Figure 1 shows the date-wise postings on various themes on Instagram. Users who post about non-followers, accounts reached, according to likes, and according to followers are shown in Figures 1(a), 1(b), 1(c), and 1(d), respectively.

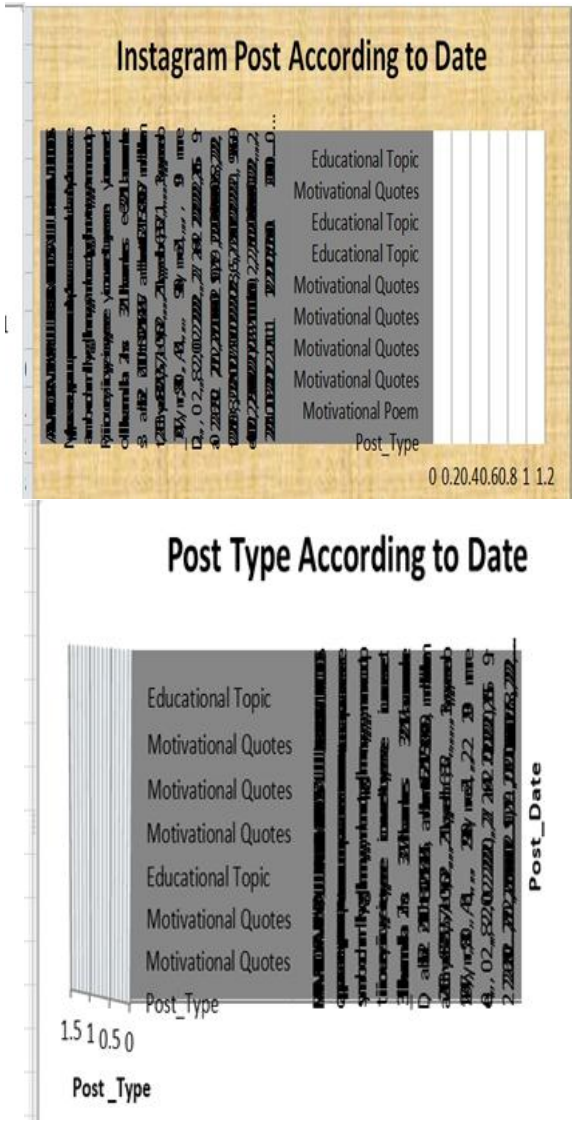

Fig.1 Date-wise posting in different topics

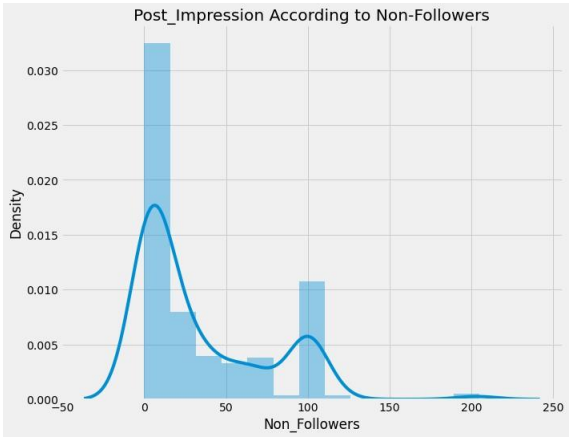

Fig.1 (a) Non-Followers post impression

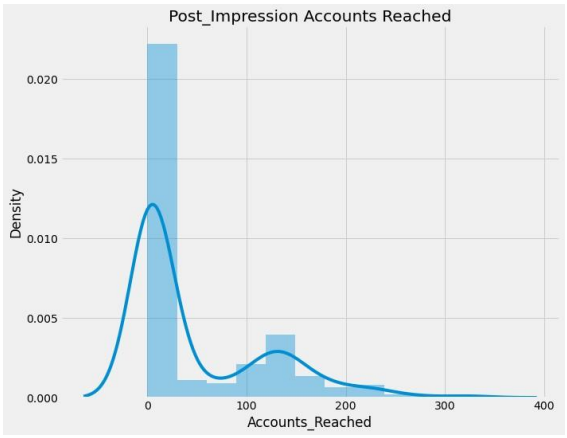

Fig.1 (b) Accounts Reached post impression

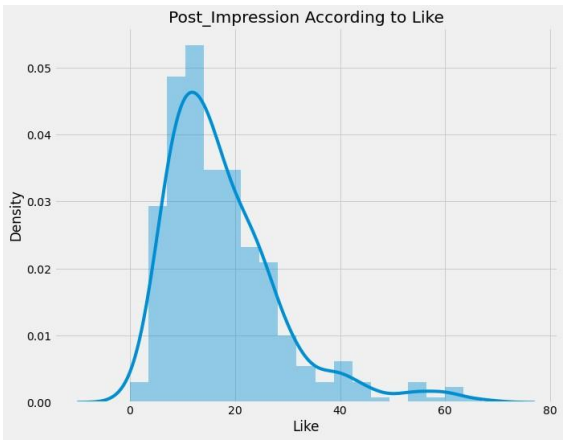

Fig.1(c) Post_Impression according to Like

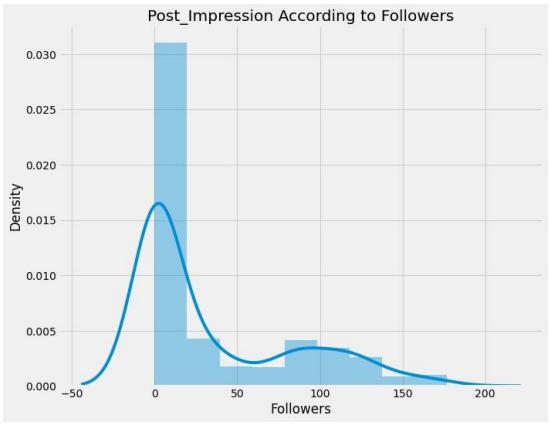

Fig.1 (d) Followers according to Post Impression

6 Discussion

A continuous dependent variable (target variable) can be predicted using the statistical technique of linear regression by taking into account one or more independent variables (predictors*)[16]*. According to this method, the dependent and independent variables have a linear connection, meaning that changes in one will cause a corresponding change in the other. In other words, the purpose of linear regression is to ascertain the degree to which the value of the dependent variable may be predicted by one or more variables. If we are trying to figure out how close our estimates are to the real numbers, we could look at things like this: Collect information, divided data, Select a model, and Adapt the model, Make Forecasts in addition to Assess the performance of the data model using measures such as mean absolute error, mean squared error, or R-squared:

R-squared: Represents the percentage of the dependent variable's variance that can be predicted based on the independent variables. Higher numbers indicate a better fit [17]. Numbers range from 0 to 1.

Mean Absolute Percentage Error: The absolute percentage error is expressed as a percentage of the actual number using the mean absolute percentage error [18]. The % error that MAPE provides can be used to assess how accurate a forecast is.

Based on the input factors, such as likes increasing in proportion to audiences reaching the accounts, we obtain anticipated values. Total Engagement (likes + comments + shares) = Engagement Rate Total Followers * 100. This provides us with a percentage that shows us the audience's level of engagement with the content. Descriptive statistics are used in data analysis to compile engagement data and spot trends and patterns [19]. Thus, analyze data according to time, content type, or demographics to understand how various elements affect engagement. We employed the regression model, which is the best algorithm for utilizing historical data to forecast future interaction. Divide the data into sets for testing and training to assess the performance of the model [20]. Update the model often with fresh information to improve forecasts and take shifting user behaviour patterns into consideration. Predictive modelling and data manipulation are done with Python tools such as Pandas, Scikit-learn, and Statsmodels[21]. With the assistance of influencers and brands in optimizing their Instagram post strategies, we were able to estimate engagement percentages with more accuracy.

Let us have a look at a dataset where each feature x has a response value of y. Figure 2 below lists the fundamental presumptions a linear regression model makes about the dataset it is applied to:

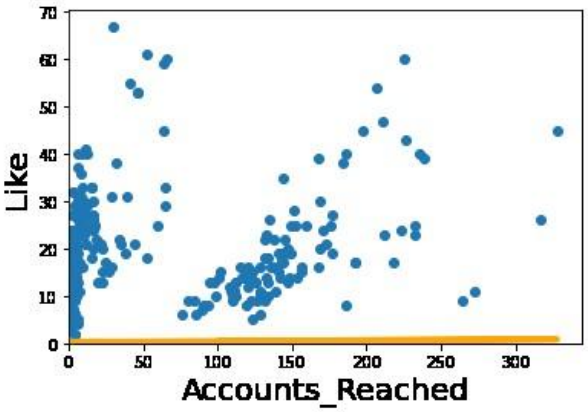

Fig. 2 Linear regression model according to Like and Accounts Reached

7 Conclusion

Interpreting and understanding linear regression is simple. Even someone without a solid background in statistics can understand it because of the simple nature of the relationship between the variables. It makes it possible to forecast a dependent variable's value using one or more independent variables. Making decisions is aided by the use of linear regression, which finds and measures the direction and strength of correlations between variables. You may ascertain how much a change in the independent variable affects the dependent variable by looking at the model's coefficients. In order to improve model specifications, linear regression offers a framework for validating data assumptions

such as linearity, normalcy, and homoscedasticity.It provides the framework for more intricate models. The concepts of linear regression are expanded upon by variations such as multiple regression, polynomial regression, and regularization strategies.It makes it possible to test hypotheses to see if the relationships found are statistically significant. Finding outliers that may require additional research can be facilitated by analyzing the residuals from a linear regression model. Quick analysis is possible because linear regression typically involves less computing than more sophisticated models. It is simple to visualize, which facilitates sharing findings and insights with stakeholders. All things considered, linear regression is a flexible technique that can offer insightful information and guide judgments in a variety of fields. Linear regression is a fundamental machine learning model that relies on the assumption of a linear relationship between two variables: the independent variable and the dependent variable. As a result, we search for a linear function that, given the feature or independent variable (X), can as correctly predict the response value (y).Using Instagram datasets as a base; we applied a linear regression model to predict an accurate value.

References

[1] Szeto, M. D., Mamo, A., Afrin, A., Militello, M., & Barber, C. (2021). Social media in dermatology and an overview of popular social media platforms. *Current dermatology reports*, 1-8.

[2] Arriagada, A., & Ibáñez, F. (2020). "You need at least one picture daily, if not, you're dead": content creators and platform evolution in the social media ecology. *Social Media+ Society*, 6(3), 2056305120944624.

[3] Rogers, R. (2021). Visual media analysis for Instagram and other online platforms. *Big Data & Society*, 8(1), 20539517211022370.

[4] Yost, E., Zhang, T., & Qi, R. (2021). The power of engagement: Understanding active social media engagement and the impact on sales in the hospitality industry. *Journal of Hospitality and Tourism Management*, 46, 83-95.

[5] Rejeb, A., Rejeb, K., Abdollahi, A., & Treiblmaier, H. (2022). The big picture on Instagram research: Insights from a bibliometric analysis. *Telematics and Informatics*, 73, 101876.

[6.Leong, L. Y., Hew, T. S., Ooi, K. B., Hajli, N., & Tan, G. W. H. (2024). Revisiting the social commerce paradigm: The social commerce (SC) framework and a research agenda. *Internet Research*, 34(4), 1346-1393.

[7] Kim, C., & Yang, S. U. (2017). Like, comment, and share on Facebook: How each behavior differs from the other. *Public relations review*, 43(2), 441-449.

[8.] Ahmadi, A., Fakhimi, S., & Ahmadi, Y. (2022). Instagram celebrities and positive user responses. The mediating role of user "like". *Journal of Contemporary Marketing Science*, 5(1), 65-80.

[9] Riquelme, F., & González-Cantergiani, P. (2016). Measuring user influence on Twitter: A survey. *Information processing & management*, 52(5), 949-975.

[10] Garg, P., Gupta, B., Dzever, S., Sivarajah, U., & Kumar, V. (2020). Examining the relationship between social media analytics practices and business performance in the Indian retail and IT industries: The mediation role of customer engagement. *International journal of information management*, 52, 102069.

[11] Bell, S., & Komisarczuk, P. (2020, February). Measuring the effectiveness of twitter's url shortener (t. co) at protecting users from phishing and malware attacks. *In Proceedings of the Australasian Computer Science Week Multiconference* (pp. 1-11).

[12.] Al Kurdi, B., Alshurideh, M., Alshurideh, H., & Al-Gasaymeh, A. (2022). The role of business intelligence in social media marketing and its impact on firm performance. *International Journal of Theory of Organization and Practice (IJTOP)*, 2(1), 16-36.

[13] Schwaiger, J., Hammerl, T., Florian, J., & Leist, S. (2021). UR: SMART–A tool for analyzing social media content. *Information Systems and e-Business Management*, 19(4), 1275-1320.

[14] Arora, A., Bansal, S., Kandpal, C., Aswani, R., & Dwivedi, Y. (2019). Measuring social media influencer index-insights from Facebook, Twitter and Instagram. *Journal of retailing and consumer services*, 49, 86-101.

[15] Maares, P., Banjac, S., & Hanusch, F. (2021). The labour of visual authenticity on social media: Exploring producers' and audiences' perceptions on Instagram. *Poetics*, 84, 101502.

[16] Chaurasia, V., & Pal, S. (2022). An ensemble

framework-stacking and feature selection technique for detection of breast cancer. *International Journal of Medical Engineering and Informatics*, 14(3), 240-251.

[17] Arora, A., Bansal, S., Kandpal, C., Aswani, R., & Dwivedi, Y. (2019). Measuring social media influencer index-insights from Facebook, Twitter and Instagram. *Journal of retailing and consumer services*, 49, 86-101.

[18] Purba, K. R., Asirvatham, D., & Murugesan, R. K. (2021). Instagram post popularity trend analysis and prediction using hashtag, image assessment, and user history features. *Int. Arab J. Inf. Technol.*, 18(1), 85-94.

[19] Howard, S. K., Ma, J., & Yang, J. (2016). Student rules: Exploring patterns of students' computer-efficacy and engagement with digital technologies in learning. *Computers & Education*, 101, 29-42.

[20] Salazar, J. J., Garland, L., Ochoa, J., & Pyrcz, M. J. (2022). Fair train-test split in machine learning: Mitigating spatial autocorrelation for improved prediction accuracy. *Journal of Petroleum Science and Engineering*, 209, 109885.

[21] Molin, S. (2019). *Hands-On Data Analysis with Pandas: Efficiently perform data collection, wrangling, analysis, and visualization using Python*. Packt Publishing Ltd.