

T. Mani sai Lokesh

192224105

Artificial Intelligence & Data  
science  
Guide: Manikandan

Verified  
Dr. B. Manikandan

Hint Paper: 1

Title: Improving social media impact prediction accuracy for marketing strategy an analysis of Random forest and Logistic Regression.

Paragraph 1:-

Defination:-

- \* The application of machine Learning techniques to forecast the potential impact or user content may have social media Platforms.
- \* Analysis of Engagement, interactions these Predictive models.
- \* To provide insights a user content audience, aiding in decision making for marketing purpose.
- \* The goal provide marketers insights, influential users and tailor marketing strategies for optimal reach impact dynamic landscape of social media.

Importance:-

- 1) Business and marketing : It servers a powerful Platforms for business to reach and target audience. It is cost effective way to promote product & services.
- 2) Awareness and Activism: Plays a crucial role raising awareness about issues, promoting activism.
- 3) Global connectivity:- break down barriers connectivity People diverse backgrounds clusters globalized digital community.

## Applications:-

- \* The Professional network connected with family, friends & colleagues.
- \* To share information content, articles, videos and creative expression for brand awareness
- \* In real time news & information dissemination, global events & trends.
- \* customer support through social media channels feedback & resolve issues.

## Paragraph - 2:-

### Number of Articles:-

- \* sciencedirect - 501
- \* IEEE - 1850
- \* Google scholar - 1355

### Most cited Articles:-

- [1] Li, Jingxuan, et al. "social network user influence sense-making and dynamics Prediction." Expert system with applications 41.11 (2014): 5115-5124.
- [2] Alizadeh, Meysam, et al. "content-based features Predict social media influence operations." science advances 6.30(2020): eabb5824.
- [3] Al Marouf, Hasan, (2020). comparative Analysis of Feature selection algorithm for computational Personality Prediction for social media. IEEE Transactions on computational social systems, 7(3) 587-599.



[4] Yu, sheng, and subhash kat. "A survey of Prediction using social Media." arxiv Preprint arxiv: 1203.1647 (2012).

### Best study:-

- \* Illustrating the latent capacity of machine learning techniques in anticipation of social media in similar setting.
- \* Business Promotion Powerful for marketing strategy and advertising are high income.
- \* They increased demand & supply occurrences in future.

### Paragraph 3:-

#### Existing Algorithm:-

- \* In the existing algorithm, the accuracy is less because of the comparison of short period of data.
- \* The machine trained used the existing can predict the social media influence in marketing strategy to owner.

#### Aim of the Algorithm:-

utilizing the machine learning Algorithm and social media influence analysis to develop robust models and predict influence of users on social media content. Enhancing market strategies provide insights, ultimately decision making process for business and organizations.

## Materials And Methods:-

### Paragraph 1:-

study setting: SIMATS school of Engineering

No of Groups: 2

Group 1: Random forest

Group 2: Logistic Regression

sample size: 20

Dataset: Predicting the accuracy occurrences of social media influences.

Power: 80%.

### Paragraph 2:-

Data sample Preparation

Group 1: Random forest

~~Group 2: Ada Boost~~

\* Information: Dataset

\* Apply Random Forest Algorithm

\* calculate the total no of users see in the Post.

\* calculate the accuracy of social media <sup>influence</sup> Prediction.

### Paragraph 3:-

\* Data sample Preparation Group 2:

Logistic Regression

- \* Information : Dataset
- \* Applying Logistic Regression Algorithm
- \* calculate the total no of Peoples can see Post
- \* calculate the accuracy of social media influence Prediction.

#### Paragraph 4:-

##### Testing setup

- \* Google colab
- \* i7 Intel 8th Gen
- \* 8 GB RAM
- \* windows 10 os

#### Paragraph 3:-

##### statistical Analysis:-

- \* Utilizing version 26.0 of IBM SPSS software, computation were performed.
- \* Provided values are → Mean  
→ standard deviation  
→ standard error mean.

Independent variable: <sup>QOL, IAT</sup> Date, time, ~~cost~~, settlement, index, Age, Education, Field,

Dependent variable :- social media influence time in hours.



- \* Analysis done : yes
- \* Result: Random forest has better prediction than the Logistic Regression.
- \* Notably independent variable of interest is rough score.
- \* Dependent variable of Research size and recorded data utilized of T-test outcomes.

#### Limitations:-

- \* In order to maximize the fetch time and increase the accuracy.
- \* Bulk data analysis is complicated.

#### Future scope:-

- \* Accuracy increased using Random Forest Algorithm.

#### Testing Procedure:-

- \* Preparing the dataset
- \* Train 70% of the Dataset
- \* Test 30% of the Dataset
- \* create Embedded model using Random Forest algorithm and Logistic regression.

#### Results and Discussions:

- \* Improving Accuracy in Prediction of social media influence among analysis in marketing strategies by minimizing false datasets.

\* using Random Forest Algorithm and support vector Machine.

### Paragraph 1:-

In this study we defined that the Random Forest Algorithm has better Prediction than the Logistic Regression.

### Paragraph 2:-

#### Data collection:-

- \* Data is trained
- \* saved Embedded model
- \* Input sample Dataset.
  - i) Age: Above 28
  - ii) Field: Engineering
  - iii) Education: Bachelor's Degree
  - iv) settlement: Urban.

### conclusion:

- \* The analysis of Random Forest and Logistic Regression models forecasting social media's influence on marketing strategy, both techniques enhance accuracy.
- \* Models enhances Predictive capabilities, empowering marketers with more Precise insights.

\*

Group Statistics				
	ALGORITHM	N	Mean	Std. Deviation
ACCURACY	RF	20	93.90	1.744
	LR	20	84.35	2.641

Independent Samples Test										
Levene's Test for Equality of Variances					t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Y	Equal variances assumed	5.950	.020	13.493	38	.025	9.550	.708	8.117	10.983
	Equal variances not assumed			13.493	32.923	.025	9.550	.708	8.110	10.990

