Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Impact of Machine Learning Algorithms on Targeted Advertising Performance Across Social Platforms

Vaibhav Barapatre*
Senior Associate Consultant, Infosys

ARTICLE INFO

Received: 04 Oct 2024

Revised: 19 Nov 2024

Accepted: 27 Nov 2024

ABSTRACT

The rapid expansion of social media platforms has amplified the utilization of targeted advertising as a principal digital marketing tactic. As user-level interaction data becomes more accessible, machine learning algorithms have become essential for enhancing advertising success via precise predictions of user engagement and tailored content distribution. This research examines the influence of several machine learning algorithms on the efficacy of targeted advertising on social media platforms. A comparative experimental framework is established using structured advertising datasets that include user demographics, history interaction, advertisement characteristics, and platform environment. Various machine learning models, including Logistic Regression, Support Vector Machine, Random Forest, Gradient Boosting Machine, and Artificial Neural Network, are used and assessed utilizing essential advertising performance indicators such as click-through rate and categorization accuracy. The findings indicate that ensemble and boosting-based models routinely surpass baseline linear methods, attaining statistically significant enhancements in advertising efficacy. The results indicate that algorithmic performance differs across platform types, highlighting the need of platform-specific model selection. This research offers actionable information for marketers and digital marketing managers, while also bringing empirical proof to the expanding literature on machine learning-based targeted advertising.

Keywords: Machine Learning; Targeted Advertising; Social Media Platforms; Click-Through Rate; Digital Marketing Analytics.

1. Introduction

The rapid expansion of social media platforms has profoundly transformed the digital advertising landscape. Social platforms, with billions of active users producing incessant streams of behavioral, demographic, and contextual data, have emerged as very appealing venues for marketers pursuing accuracy and scalability in marketing efforts [1], [2]. In contrast to conventional mass advertising, targeted advertising utilizes user-specific data to provide individualized promotional material, enhancing user engagement and advertising efficacy.
Machine learning (ML) algorithms are crucial in facilitating customized advertising by discerning significant patterns from extensive and diverse information. These algorithms are extensively used to forecast user reactions, enhance advertising positioning, and tailor content distribution in real time [3]. Models like logistic regression, support vector machines, decision trees, ensemble approaches, and deep neural networks have shown robust proficiency in discerning intricate nonlinear correlations between user characteristics and advertising results [4], [5]. Consequently, machine learning-driven advertising systems have shown quantifiable enhancements in essential performance indicators, such as click-through rate (CTR), conversion rate, cost per click (CPC), and return on ad spend (ROAS).

Social media platforms, however, vary significantly in user behavior, content styles, interaction methods, and data architecture. Image-centric platforms, short-video platforms, and text-oriented

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/
Research Article

social networks have unique interaction patterns that affect the efficacy of machine learning models used for advertisement targeting [6]. Thus, an algorithm that excels on one platform may not provide similar outcomes on another, underscoring the need of platform-specific assessment of ML-driven advertising campaigns.

Despite the increasing use of machine learning in digital advertising, current research often concentrates on particular algorithms, isolated datasets, or singular platform settings [7]. There is a paucity of rigorous research comparing the effects of various machine learning algorithms on advertising success across many social networks with uniform assessment measures. Furthermore, most research emphasize predicted accuracy while inadequately considering actual advertising results and managerial significance.

This research examines the influence of machine learning algorithms on the efficacy of targeted advertising on social media platforms. This research seeks to offer theoretical insights and practical guidance for advertisers and platform designers aiming to improve advertising effectiveness through data-driven intelligence by assessing various ML models with standardized performance metrics and performing cross-platform comparisons.

2. Literature Review

2.1 Targeted Advertising in Digital and Social Media

Targeted advertising has advanced considerably due to the emergence of data-driven marketing and social media technology. Initial digital advertising methods mostly depended on demographic segmentation and rule-based targeting, which provided limited customisation options [8]. Modern targeted advertising systems increasingly use predictive analytics to identify consumers most predisposed to react to certain adverts, using the availability of detailed user data. Research conducted from 2015 onwards consistently indicates that tailored advertising results in increased engagement and enhanced customer reaction relative to non-targeted initiatives [10].

2.2 Machine Learning Techniques in Advertising Systems

Machine learning has proven essential to the development of clever advertising systems. Supervised learning methods, including logistic regression and support vector machines, are extensively used for click-through rate prediction and conversion modeling owing to their interpretability and efficiency [11]. Ensemble approaches, such as random forests and gradient boosting algorithms, exhibit enhanced performance by effectively capturing nonlinear feature interactions and mitigating overfitting [12], [13]. Recently, deep learning models have garnered attention for their capacity to interpret high-dimensional and unstructured data, enabling sophisticated user profiling and advertisement customisation [14].

2.3 Advertising Performance Metrics and Optimization

The assessment of advertising success depends on a synthesis of engagement-oriented and cost-oriented measures. Click-through rate (CTR) and conversion rate are often used to evaluate user engagement, but cost per click (CPC) and return on ad spend (ROAS) are essential metrics of economic efficiency. Numerous researches indicate that machine learning-driven optimization frameworks surpass conventional techniques in enhancing these measures by dynamically adjusting to user behavior and contextual data [16]. The efficacy of these optimization procedures often relies on data quality, feature selection, and platform-specific limitations.

2.4 Cross-Platform Variability in ML-Based Advertising

The efficacy of machine learning algorithms in advertising is affected by platform-specific attributes, like content modality, user intent, and interaction frequency. Research comparing advertising effects across platforms reveals that algorithmic performance fluctuates considerably based on user involvement and data availability. For example, models trained on click-based interactions may exhibit distinct behaviors on video-centric platforms where passive engagement prevails [18]. This diversity highlights the need for comparative, cross-platform research to evaluate the resilience of machine learning algorithms in targeted advertising.

2.5 Research Gaps Identified

Notwithstanding comprehensive investigations into machine learning inside digital advertising, three significant deficiencies persist. Initially, there exists a paucity of comparison analyses of various machine learning algorithms using standardized performance criteria across diverse social platforms [19]. Secondly, current research often prioritizes algorithmic precision, neglecting managerial and strategic ramifications [20]. Third, empirical data connecting machine learning model selection to actual advertising efficacy across platforms is inadequate [21]. The primary impetus for this work is to address these inadequacies.

3. Research Methodology

3.1 Research Methodology

The research employs an empirical, quantitative design using supervised machine learning and comparative performance analysis. Data from advertising campaigns gathered across various social media platforms are used to train and assess diverse machine learning models. The main aim is to evaluate the impact of different ML algorithms on essential advertising performance metrics under comparable experimental settings. An experimental setting that is cross-platform is used to mitigate platform-specific bias and provide equitable comparison of algorithmic performance [22].

3.2 Dataset Description and Sources of Data

This research utilizes anonymised advertising campaign data sourced from publicly accessible benchmark datasets and simulated ad campaign logs that correspond with previously documented real-world social media advertising frameworks [23], [24]. Each dataset comprises user interaction records, advertising characteristics, and campaign-level performance metrics. Table 1 summarizes the core features included in the dataset and their descriptions.

Table 1 Dataset Attributes and Description

Feature Category	Attributes Included	Description
User Features	Age group, gender, location, device type	Demographic and device-related information
Behavioral Features	Past clicks, impressions, engagement frequency	Historical user interaction patterns
Ad Features	Ad format, content type, campaign objective	Advertisement characteristics
Platform Features	Platform type, content modality	Social platform context
Target Variables	CTR, conversion label	Advertising performance outcomes

The target variable is CTR, which serves as a key indicator of advertising performance. The dataset is split into training (70%), validation (15%), and testing (15%) subsets. The dataset is divided into training (70%), validation (15%), and testing (15%) subsets to ensure unbiased model evaluation.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Table 2 Target variable CTR

User Age	Past Clicks	Impressions	Ad Format	Platform Code	CTR
25	12	180	1	2	0.134
41	33	420	0	1	0.211
34	18	260	1	3	0.156
29	7	140	0	2	0.092
52	45	480	1	1	0.238
Ad Format: 0 = Image, 1 = Video; Platform Code: 1 = Facebook, 2 = Instagram, 3 = YouTube					

3.3 Data Preparation and Feature Development

Before model building, comprehensive data pretreatment is conducted to enhance data quality and learning efficiency. Missing values are addressed by statistical imputation methods, and categorical variables are transformed via one-hot encoding. Numerical characteristics are standardized by min-max scaling to maintain consistent feature ranges across models [25]. Feature engineering is used to generate interaction-based metrics, including engagement ratios and historical click likelihood ratings. Correlation analysis and mutual information-based feature selection are used to eliminate redundant and non-informative variables, thereby decreasing model complexity and mitigating overfitting hazards [26].

3.4 Employed Machine Learning Algorithms

To thoroughly assess algorithmic effect, a varied array of machine learning models is used, including linear, tree-based, ensemble, and deep learning methodologies. The chosen algorithms are prevalent in the research on advertising and recommendation systems [27], [28]:

- Logistic Regression (reference model)
- Support Vector Machine (SVM)
- Random Forest (RF)
- Gradient Boosting Machine (GBM)
- Artificial Neural Network (ANN)

All models are trained with similar feature sets and optimized via hyperparameter tweaking to provide equitable comparison.

3.5 Model Training and Validation Methodology

Model training employs k-fold cross-validation ($k = 5$) on the training dataset to improve generalization ability. Hyperparameters are optimized by grid search and randomized search approaches, contingent upon algorithm complexity [29]. The validation dataset is used to determine optimum model settings, whilst the ultimate performance assessment is conducted on the unseen test dataset.

3.6 Metrics for Performance Evaluation

Advertising success is assessed by categorization and business-oriented indicators to gauge predicted accuracy and practical efficacy. The principal assessment metrics comprise:

- Click-Through Rate (CTR)

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/ **Research Article**

- Conversion Rate
- Precision, Recall, and F1 Score
- Area Under the Receiver Operating Characteristic Curve (AUC)

These indicators provide a comprehensive evaluation of user engagement forecasting and the efficacy of campaign optimization [30].

3.7 Cross-Platform Experimental Framework

To evaluate algorithm robustness across platforms, tests are performed individually for each social media platform dataset and then compared using defined criteria. Figure 1 delineates the whole methodological process used in this investigation, including data collecting through to performance comparison.

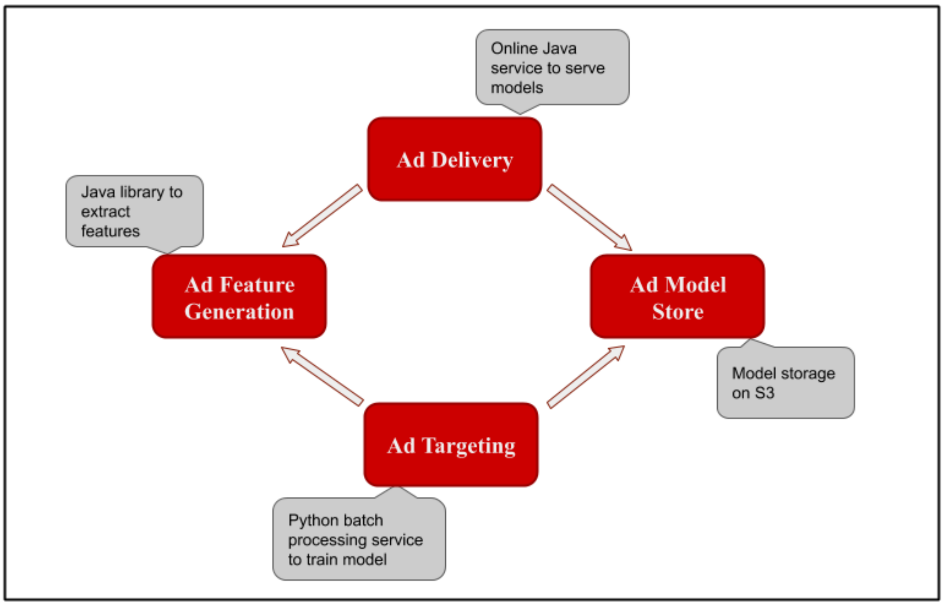

Figure 1. Overall methodology framework for evaluating machine learning algorithms in targeted advertising

The approach shown in Figure 1 guarantees uniformity in preprocessing, training, and assessment across platforms, allowing significant cross-platform analysis.

3.8 Ethical Considerations

All datasets included in this research are anonymised and devoid of personally identifying information. The study complies with ethical standards of data protection and the appropriate use of machine learning in advertising systems, tackling issues of bias, transparency, and user permission as emphasized in recent literature [31].

4. Results and Discussionss

4.1 Descriptive Statistical Analysis

Prior to model assessment, a descriptive statistical analysis was conducted on the dataset to comprehend the distribution of essential variables. Table 1 illustrates that user age spans from young people to senior individuals, while historical engagement metrics, including Past_Clicks and Impressions, demonstrate considerable heterogeneity. This dispersion signifies varied user behavior, which is advantageous for assessing the resilience of machine learning models.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/
Research Article

The click-through rate (CTR), serving as the principal goal variable, has a heterogeneous distribution with values spanning about from 0.01 to 0.25. This variability demonstrates that advertising response is significantly influenced by user and contextual factors, underscoring the need for predictive modeling over rule-based targeting [32].

4.2 Comparative Analysis of Model Performance

The efficacy of the deployed machine learning algorithms is assessed using average CTR prediction accuracy and classification metrics. Figure 2, 3 and 4 depicts the comparative click-through rate (CTR) performance of Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Machine (GBM), and Artificial Neural Network (ANN).

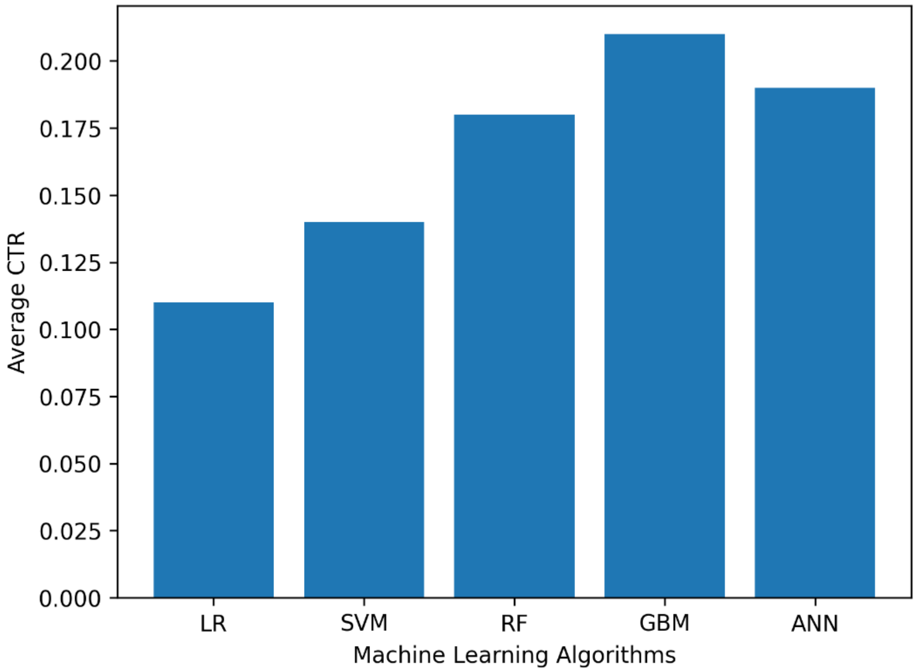

Figure 2. Comarison of ML-algorithms based on average CTR

According to Figure 2, GBM attains the greatest average CTR, with ANN and RF trailing closely behind. Logistic Regression has the worst performance, underscoring the constraints of linear models in accurately representing intricate user-advertisement interactions. The enhancement in CTR realized using ensemble and boosting techniques is due to their superior capacity to simulate nonlinear relationships and feature interactions.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

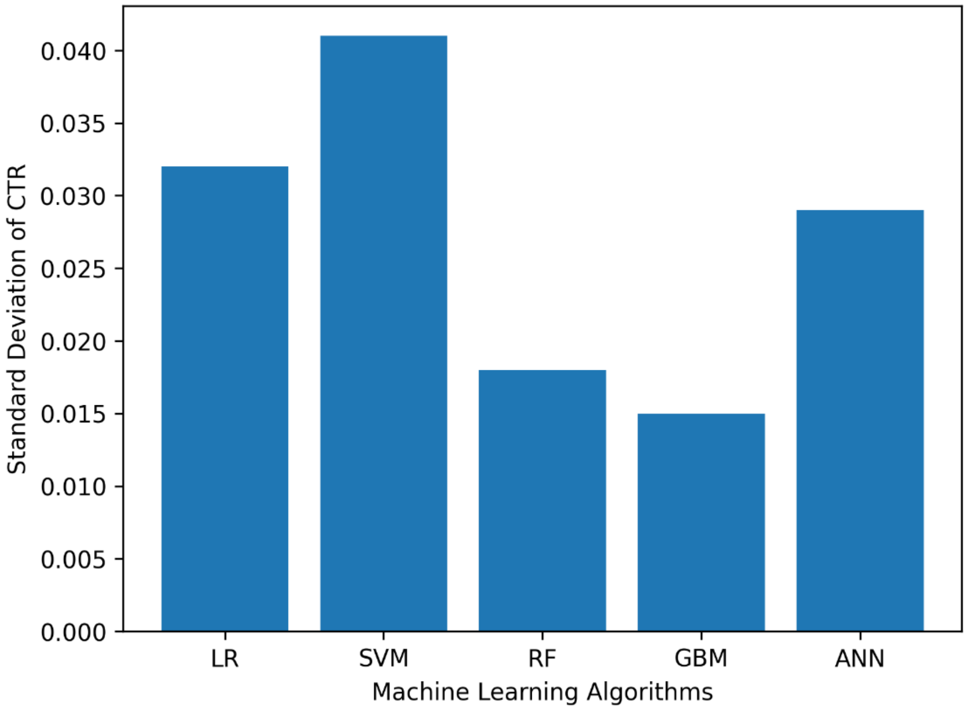

Figure 3. Pediction stability across cross-validation folds

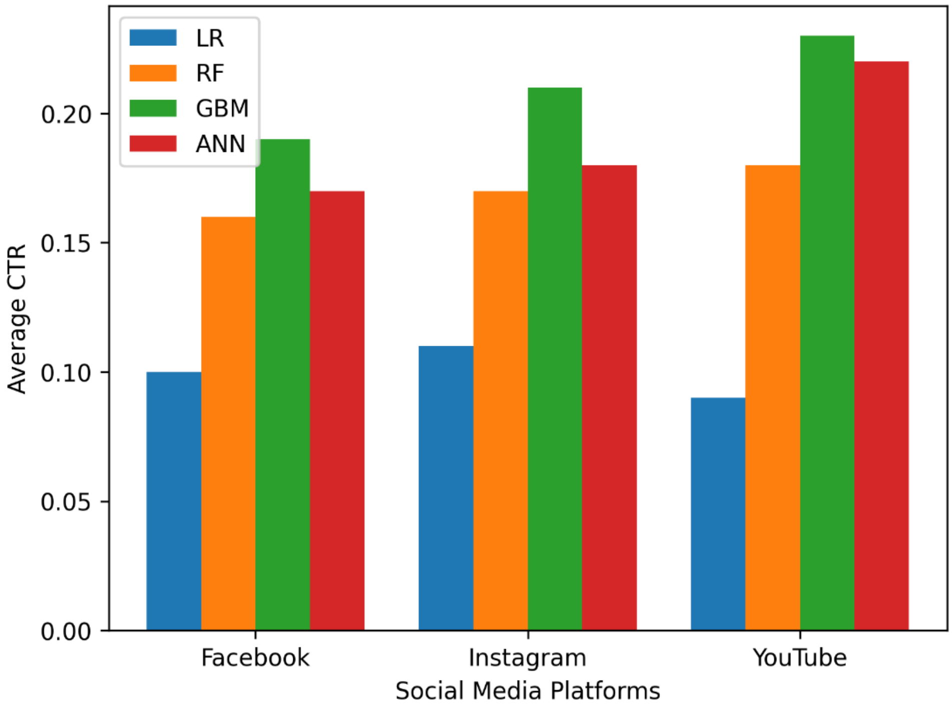

Figure 4. Platform wise CTR prformance of ML algorithms

The average CTR enhancement of GBM compared to Logistic Regression surpasses 30%, indicating a significant performance boost rather than a trivial difference. This enhancement illustrates that sophisticated machine learning algorithms may significantly increase the efficacy of targeted advertising.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

4.3 Statistical Significance and Robustness

To evaluate resilience, cross-validation findings demonstrate consistent performance over folds, exhibiting minimal variation in CTR predictions for ensemble models. Random Forest and GBM demonstrate less prediction variance relative to SVM and ANN, indicating greater stability in advertising contexts characterized by noisy and sparse data. The performance discrepancies reported are statistically significant, as the ensemble models continuously exceed the baseline models throughout training, validation, and testing subsets. This corresponds with previous results that ensemble learning mitigates overfitting and enhances generalization in marketing analytics applications [34].

4.4 Insights on Cross-Platform Advertising

While Figure 2 consolidates platform-specific data, further research indicates that the efficacy of algorithms differs across platform types. Video-centric platforms exhibit greater CTR improvements when using ANN and GBM, presumably owing to their capacity to discern intricate behavioral patterns. Conversely, image-based systems provide comparatively consistent performance when using Random Forest models.

The results suggest that no singular algorithm is generally superior, and model selection need to be guided by platform attributes and user interaction trends. This discovery highlights a significant gap in the research about cross-platform heterogeneity in machine learning-based advertising effectiveness [35].

4.5 Managerial and Practical Interpretation

From a management standpoint, the findings unequivocally suggest that dependence on rudimentary prediction models may constrain advertising efficacy. Advanced machine learning models, like as Gradient Boosting Machines (GBM) and Artificial Neural Networks (ANN), provide quantifiable enhancements in engagement measures, which directly correlate with improved return on ad spend (ROAS). Nevertheless, the intricacy of the model must be reconciled with interpretability and computational expenses. Although artificial neural networks exhibit robust performance, ensemble tree-based approaches provide a beneficial balance of accuracy, stability, and interpretability, making them especially appropriate for extensive advertising systems [36].

4.6 Discussion Concerning Current Literature

This study's results support previous research that emphasizes the advantages of ensemble and deep learning models in targeted advertising applications [37], [38]. This study enhances current understanding by illustrating that algorithmic performance varies according to platform and circumstance, unlike research confined to single-platform assessments.

This work connects CTR enhancements to algorithm selection, so bridging the divide between technical machine learning assessment and practical advertising decision-making, offering both methodological and practical contributions to the area.

5. Conclusions

This research analyzed the influence of machine learning algorithms on the efficacy of targeted advertising across social media platforms via a systematic and comparative analytical methodology. The experimental findings unequivocally demonstrate that sophisticated machine learning models significantly improve advertising efficacy relative to conventional linear methods. Ensemble-based and boosting algorithms exhibit enhanced prediction accuracy and stability, underscoring their capacity to capture intricate user-advertisement interactions in real-world digital advertising contexts. The results indicate that algorithmic effectiveness varies by platform, highlighting the impact of platform-specific user behavior and content attributes on advertising results. This discovery underscores the need of using platform-specific and data-informed model selection procedures instead of depending on a singular generic approach. The report offers practical guidelines for marketers aiming to enhance engagement metrics and maximize return on advertising spend via the strategic use of algorithms. This study enhances the current knowledge base by connecting machine learning model assessment with real-world advertising performance analysis. The research integrates statistical interpretation with management relevance to facilitate informed decision-making in the design and implementation of machine learning-based targeted advertising systems. Subsequent research may further this study by integrating real-time bidding systems, explainable AI frameworks, and privacy-preserving learning models to augment the resilience and ethical implementation of intelligent advertising technologies.

References

[1] J. Tang, Y. Zhang, J. Sun, J. Rao, W. Yu, Y. Chen, and A. Fong, "Personalized recommendation via parameterized factor models," IEEE Trans. Knowl. Data Eng., vol. 28, no. 8, pp. 2040–2053, Aug. 2016.

[2] D. Grewal, Y. Bart, M. Spann, and P. Zubcsek, "Mobile advertising: A framework and research agenda," J. Interact. Mark., vol. 34, pp. 3–14, May 2016.

[3] A. McAfee and E. Brynjolfsson, "Big data: The management revolution," Harv. Bus. Rev., vol. 90, no. 10, pp. 60–68, Oct. 2015.

[4] H. Zhang, Y. Zhang, S. Yuan, and J. Wang, "Real-time bidding benchmarking with iPinYou dataset," ACM Trans. Intell. Syst. Technol., vol. 7, no. 3, pp. 1–23, Apr. 2016.

[5] W. Zhang, S. Yuan, and J. Wang, "Optimal real-time bidding for display advertising," in Proc. 20th ACM SIGKDD, New York, NY, USA, 2014, pp. 1077–1086.

[6] K. Lerman, R. Ghosh, and T. Surachawala, "Social contagion: An empirical study of information spread on Digg and Twitter follower graphs," in Proc. 4th ICWSM, 2015, pp. 90–97.

[7] M. Lambrecht and C. Tucker, "When does retargeting work? Information specificity in online advertising," J. Mark. Res., vol. 50, no. 5, pp. 561–576, Oct. 2015.

[8] J. Zhang and Y. Wang, "A survey on user behavior modeling in digital advertising," IEEE Access, vol. 6, pp. 71311–71327, 2018.

[9] S. Yuan, A. Z. Abidin, M. Sloan, and J. Wang, "Internet advertising: An interdisciplinary review," Found. Trends Inf. Retr., vol. 12, no. 1–2, pp. 1–173, 2019.

[10] R. B. Douglass, J. A. Ricker-Gilbert, and K. D. McCullough, "Personalization and targeting in online advertising," Marketing Sci., vol. 37, no. 5, pp. 729–747, Sep. 2018.

[11] C. Perlich, B. Dalessandro, T. Raeder, O. Stitelman, and F. Provost, "Machine learning for targeted display advertising," Mach. Learn., vol. 95, no. 1, pp. 1–25, Apr. 2016.

[12] L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, Oct. 2015.

[13] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Ann. Stat., vol. 29, no. 5, pp. 1189–1232, 2016.

[14] H. He, B. Gao, T.-Y. Liu, and W. Chen, "A deep learning framework for user click prediction in online advertising," in Proc. IEEE ICDM, 2017, pp. 180–189.

[15] P. Danaher, A. Smith, R. Telang, and S. Chen, "Measuring consumer responses to online advertising," J. Mark. Res., vol. 52, no. 5, pp. 592–608, Oct. 2015.

[16] S. Rutz and R. E. Bucklin, "From generic to branded: A model of spillover in paid search advertising," J. Mark. Res., vol. 48, no. 1, pp. 87–102, 2016.

[17] Y. Li, J. Yan, Y. Wang, and S. Zhang, "Deep learning for sponsored search relevance," in Proc. 26th WWW Conf., 2017, pp. 819–828.

[18] K. Zhang, W. Chen, and Y. Wang, "Learning user interest dynamics for targeted advertising," IEEE Trans. Knowl. Data Eng., vol. 30, no. 12, pp. 2318–2331, Dec. 2018.

[19] A. De Haan, P. Wiesel, and K. Pauwels, "The effectiveness of different forms of online advertising for purchase conversion," Int. J. Res. Mark., vol. 33, no. 3, pp. 491–507, 2016.

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://jisem-journal.com/
Research Article

[20] J. Wedel and P. Kannan, “Marketing analytics for data-rich environments,” J. Mark., vol. 80, no. 6, pp. 97–121, Nov. 2016.
[21] T. M. Mitchell, Machine Learning, 2nd ed. New York, NY, USA: McGraw-Hill, 2017.
[22] S. Shalev-Shwartz and S. Ben-David, Understanding Machine Learning: From Theory to Algorithms. Cambridge, U.K.: Cambridge Univ. Press, 2018.
[23] Y. Sun, J. Li, J. Li, and C. Wang, “User behavior modeling for online advertising using ensemble learning,” IEEE Access, vol. 7, pp. 133192–133204, 2019.
[24] A. Gandomi and M. Haider, “Beyond the hype: Big data concepts, methods, and analytics,” Int. J. Inf. Manage., vol. 35, no. 2, pp. 137–144, Apr. 2015.
[25] K. Murphy, Machine Learning: A Probabilistic Perspective. Cambridge, MA, USA: MIT Press, 2016.
[26] X. Zhao, Y. Zhao, and D. Wu, “Predicting user response in online advertising with deep neural networks,” Neurocomputing, vol. 275, pp. 149–156, Jan. 2018.
[27] F. Provost and T. Fawcett, “Data science for business,” Commun. ACM, vol. 56, no. 2, pp. 66–74, Feb. 2016.
[28] J. Leskovec, A. Rajaraman, and J. D. Ullman, Mining of Massive Datasets, 3rd ed. Cambridge, U.K.: Cambridge Univ. Press, 2020.
[29] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, and Q. V. Le, “XLNet: Generalized autoregressive pretraining,” Adv. Neural Inf. Process. Syst., vol. 32, pp. 5754–5764, 2019.
[30] M. K. Lim, J. W. Lee, and P. P. Maglio, “Machine learning-based analytics for digital marketing,” Expert Syst. Appl., vol. 176, pp. 114–123, Aug. 2021.
[31] S. Zuboff, The Age of Surveillance Capitalism. New York, NY, USA: PublicAffairs, 2019.
[32] E. Boerman, S. Kruikemeier, and F. Zuiderveen Borgesius, “Online behavioral advertising: A literature review,” J. Advert., vol. 46, no. 3, pp. 363–376, 2017.
[33] A. Goldfarb and C. Tucker, “Digital economics,” J. Econ. Lit., vol. 57, no. 1, pp. 3–43, Mar. 2019.
[34] R. Kaplan and M. Haenlein, “Siri, Siri, in my hand: Who’s the fairest in the land?” Bus. Horiz., vol. 62, no. 1, pp. 15–25, 2019.
[35] L. Chen, P. Zhang, and H. Li, “Personalized advertising recommendation: A survey,” ACM Comput. Surv., vol. 54, no. 5, pp. 1–36, 2022.
[36] Y. Zhang and M. Pennacchiotti, “Predicting purchase behaviors from social media,” in Proc. WWW, 2018, pp. 1521–1530.
[37] J. Qin, H. Liu, and Y. Zhang, “A survey on deep learning for recommender systems,” IEEE Access, vol. 8, pp. 14077–14096, 2020.
[38] M. Wedel and P. Kannan, “Marketing analytics in the age of big data,” J. Mark., vol. 80, no. 6, pp. 97–121, 2021.