Email: editor@ijerst.com or editor.ijerst@gmail.com

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025

CLASSIFYING ONLINE USERS THROUGH MACHINE LEARNING AND INFORMATION-SEEKING PATTERNS

[1]*Palle Subbarayudu*, *MCA Student, Department of MCA*

[2] *B. Harish Kumar Reddy, M.Tech, (Ph.D), Assistant Professor, Department of MCA*

[12]*Dr KV Subba Reddy Institute of Technology, Dupadu, Kurnool*

ABSTRACT

The internet is overflowing with unfiltered, impromptu, and continuous material from many sources due to technology in today's environment. In order to successfully deliver information depending on user intent, complex algorithms are created. Users' online experiences include a number of information-seeking behaviours, such as sharing, searching, and information verification. However, a thorough investigation of this complex user behaviour is still pending. This study helps to identify different types of users based on their online engagement, propose a user intent-machine learning model for classifying users based on their online search, share, and verification behaviour, and show that dynamic online interactions can be categorised based on their searching, sharing, and verifying behaviour. Participants from a wide range of age, gender, and vocational backgrounds complete a questionnaire designed to collect user input on online behaviour and practices. After a thorough feature engineering process, K-Mean clustering is used to discover user intent classes or profiles and their attributes based on the key features. To identify these classes, data is subsequently used to train a supervised learning Linear Discriminant Analysis Classifier (LDAC). With 80% accuracy, the suggested framework was able to predict the user intent class. A second user study collects data on users' dynamic interactions, which is used to test the model further. The user profiles that emerge from clustering are used by human raters to label the information search, exchange, and verify activity data after it has been converted to suit the model. While the model predicted the user with 67% accuracy, the study obtains an Inter-rater reliability (IRR) of 60%. According to this study, a user's motivation for looking for information, their propensity to share information on social media, and their propensity to believe information to be reliable can all help to understand their intentions, spot behavioural patterns, and identify intent through dynamic interactions that can be utilised for search engine optimisation and targeted marketing.

I. INTRODUCTION

Humans have been known to have a voracious appetite for information since the dawn of civilisation. They acquired knowledge and disseminated it. Word-of-mouth was the main method used to carry out the information sharing. Greek philosophy is where the word epistemology (the theory of knowing) originated. Plato (429–347 BCE), for instance, was interested in how humans learn. Locke (1632–1704) was interested in how human knowledge worked. Vygotsky believed that we learn via our social culture, but Piaget (1896–1980) created a theory of genetic epistemology or the concept of cognitive development. His socio-cultural theory of cognitive development has earned him widespread recognition [1]. Today's social culture is dominated by the internet and social media, and Vygotsky's adherents encourage digital access to information [2], fostering a global culture. Web channels, online social networks, podcasts, and other easily accessible digital information sources are replacing more conventional distribution channels including newspapers, magazines, books, radio, and television. [3]. Although people are grateful for the new media period, which has made communication easier and given them a forum to

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025

express their views, many are also feeling overpowered by the deluge of information that now bursts onto their screens [3], [4]. It's becoming harder for users to concentrate on a single subject and sort through pertinent data. Information may be efficiently filtered to meet user purpose by taking into account how users interact with the internet to get information, share it on social media, or trust information. Through a survey, the authors in [5] summed up that a method for determining users' online intentions consists of two steps: (1) examining how users view and interact with online platforms and using technology to create a general profile of their online preferences, and (2) creating learning models to identify users based on their preferences for online information.

The following research topics are the focus of this study, which draws inspiration from [5]:

a) Can a user's online information-seeking behaviour be used to categorise their intentions?

b) Can the user intent classifier utilise the dynamic behaviour of the user to anticipate the kind of user intent?

Classifying and forecasting user intention to seek information online is the goal of several research and algorithms [6]. The Theory of Planned Behaviour (TPB) [7], one of the most popular behavioural models, suggests that attitude, subjective norm, and perceived behavioural control all affect user intention. The Theory of Reasoned Action (TRA) [7] is another well-liked concept that contends that attitude and subjective norms influence user intention. Since their behaviour reflects user purpose, these behaviour models correctly establish user intention as an antecedent of behaviour. Therefore, it is crucial to comprehend behaviour in order to forecast user intention, particularly in an online setting where user interactions with the internet serve as a representation of behaviour and may be recorded using a variety of methods and data types.

These models, which use a machine learning methodology, have motivated this study to ascertain user intention by examining user search behaviour, information propagation, and information trust behaviour. There are many phases involved in this process. First, a survey questionnaire is used to gather user data. During the preprocessing stage, behavioural commonalities are found, and user intent groups are defined using K-Mean Clustering. Based on user intents, the clustering method finds five different groupings. The data is then used to train a Linear Discriminant Analysis (LDA) classifier, which is used to categorise user intent clusters. Inter-Rater Reliability is used to the dynamic user interactions recorded by a different user research in order to validate the model. The user profile is used to annotate the user's dynamic interactions and is generalised from the machine learning model's observations.

II. LITERATURE SURVEY

"Shrek meets Vygotsky: Rethinking multimodal literacy practices in schools among adolescents,"

Mills, K. A.

The multimodal and digital character of teenage reading practices has been highlighted by recent studies. These behaviours occur in a variety of social contexts, especially outside of educational institutions. Three cautions that challenge presumptions on the prevalence, significance, and spontaneity of young people's multimodal behaviours in the context of digital communications are produced by this article's reexamination of recent research: (1) Adolescents nowadays are not all "digital natives," as is often believed. More emphasis should be placed on expert scaffolding of these literacies in school settings, even though some new multimodal practices are adopted by adolescents with little instruction in informal contexts. (2) Engaging adolescents in multimodal textual practices should involve more than just adapting the curriculum to their interests and practices; it should expand students' repertoire of skills and genres.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025

"Maintaining honesty in virtual social networks,"

H. Ma, U. Ozertem, P. Pantel, M. Saeidi, F. Silvestri, V. Stoyanov, A. Halevy, and C. Canton-Ferrer

Online social networks provide a forum for free speech and information exchange. Malicious activities including spreading hate speech and false information, selling illicit narcotics, and organising child exploitation or sex trafficking are also carried out via these networks. Known as the integrity challenge, this study examines the current state of the art in protecting online platforms and their users against such damage. This poll is based on the experience of fighting a wide range of integrity infractions at Facebook. We emphasise the methods that have been shown to be effective in real-world settings and that the academic community should pay more attention to. We highlight important elements of the social media ecosystem, each of which is shared by a broad range of violation kinds, rather than going over the many distinct violation types. Additionally, each of these elements represents a research and development field, and the discoveries made have broad applicability.

Utilising the information distribution of red tourist culture as an example, this study examines the impact assessment and the time-series development of public culture's Internet communication against the backdrop of new media.

Zhang Y., Feng X., and Wang X.

The dissemination of public cultural information has been facilitated by the rise of new online media, and a methodical assessment of its efficacy in communication has grown in importance as a study area. In addition to offering recommendations for enhancing the impact of public cultural network communication and creating a positive communication ecology, this study seeks to enhance and improve the evaluation index system for the effect of public cultural network communication. This paper builds the evaluation index system of public culture network communication effect under the backdrop of new media, quantifies and analyses the network communication effect of red public culture in Beijing-Tianjin-Hebei, and crawled 78,140 pieces of multi-source heterogeneous 9 Altmetrics index data from Sina Weibo, Baidu, Sogou, Bilibili, and other platforms based on factors like propagation force, guiding force, and influence. The law of communication is systematically discovered based on the modelling and analysis of the time series development of the communication effect. According to the study, the communicating effect of red public culture varies greatly and has a lot of space for improvement overall. It improves steadily with seasonal changes, and it is better in the summer and autumn than it is in the other two seasons. The dynamic changes of the cultural communication system can be better presented by the ARIMA (Autoregressive Integrated Moving Average) (1,1,1) model after the first-order difference of the time series data. The Beijing-Tianjin-Hebei red culture is creatively incorporated into the public cultural construction and research systems in this article. A thorough communication effect evaluation index system is developed to assess the impact of public culture network communication against the backdrop of new media and examine its dynamic development.

"A review of user profiling: Current trends, obstacles, and solutions,"

L. Shuib, H. F. Nweke, A. A. Norman, and C. I. Eke

The virtual representation of every user has drawn attention from online web users and information and communication technology advancements, since it is essential for efficient service personalisation. In service personalisation, satisfying customer needs and preferences is a constant struggle. By creating a thorough user profile, this problem may be resolved. User profiling is the process of gathering, arranging, and extrapolating user profile data, while a user profile is a synopsis of

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025

the person's interests, traits, behaviours, and preferences. Although there have been several evaluations on user profiling, none have concentrated on the efficient profile modelling procedure. Therefore, the purpose of this paper is to evaluate the most current state-of-the-art method for user profiling. These consist of the user profile's methods, description, attributes, and taxonomy. Performance metrics are also included, along with an analysis of the current user profiling modelling in the areas of data collection, feature extraction, profiling methodologies, and profiling techniques (together with a list of their advantages and disadvantages). Additionally, privacy, datasets, cold start concerns, trust issues, and computational complexity were highlighted in the discussion of the research challenges. Additionally, the report pointed out an area of uncharted study that offers answers to the problems mentioned and encourages more academics to advance user profiling. The results shown that the creation of precise user profiles for service personalisation is improved by an efficient modelling approach.

"A multi-level understanding of web search: Task, information seeking intentions, and user behaviour,"

J. Liu, C. Shah, M. Mitsui, and N. J. Belkin,

A search task may be seen as a series of information-seeking intents that are both motivated by and impacted by search behaviours, according to the cognitive perspective of information retrieval (IR) research. Although a substantial amount of work has addressed the behavioural implications of task aspects, little is known about how various information seeking intents in query segments act as links between task and Web search behaviour. In an effort to address two primary research questions, the authors examined intention and search behaviour data gathered from 693 query segments produced by 40 participants in a controlled laboratory environment in order to create a more thorough, multi-level (i.e., task level, intention level, and behaviour level) understanding of Web search. 1) from task to intention: what effects do various task characteristics have on users' intents to seek information at various points throughout a search session? 2) from intention to behaviour: In the present and subsequent query segments, how do a user's search habits relate to their intents to find information? The findings show that: 1) Task features had a significant impact on the frequency of occurrence of the majority of information seeking intentions, and that these effects gradually diminished as search sessions went on; 2) a variety of intentions in both current and subsequent query segments were linked to and detectable by various behavioural measure subsets. Designing system affordances to support diverse intents and search activities in a range of task stages and settings is impacted by this study's contribution to our knowledge of the relationships between task, intentions in query segments, and search behaviour.

III. SYSTEM ANALYSIS & DESIGN

EXISTING SYSTEM

In order to create a user intent model that can encapsulate the basic traits of users with regard to online behaviour and practices, this study will look at the online actions and preferences that users utilise while seeking, sharing, and confirming information. Figure 1a presents the suggested technique. There are two stages to the procedure. Creating a Machine Learning (ML) model to categorise people according to their online search, sharing, and verification behaviour is the first step. The model builds upon the earlier research covered in [35]. In order to gather data, phase II entails testing the ML model on users' dynamic interactions.

After taking into account relevant work, the machine learning pipeline for Phase 1—shown in Figure 1b—is created, paying special attention to the categorical character of the data. The literature analysis also showed that entire user information-seeking behaviour has not yet been covered by the current models, which do not account for user search and social and verified

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025

information components. Other studies use cutting-edge models like BERT [36] or BART [29] deep net models and include various data kinds, such as textual and vector. However, it is challenging to collect visual data from participants and extract text from social media owing to privacy concerns. This worry has also motivated this study to evaluate the model's prediction accuracy on user dynamic actions using user input on their perceptions of the online experience, while maintaining basic features and easily accessible data.

The study makes use of earlier research [35] that gathered user input on their internet habits, inclinations, and usage of social media to share information. The information has been encoded and preprocessed. Three new qualities are presented: information conscientiousness, online extraversion, and search openness. These characteristics, which indicate ratings for sharing, searching, and behaviour verification, are calculated using the data that is currently available. K-mean To group related data, clustering is used to the new characteristics. Clusters are given labels according to user attributes. After adding to the original data and using these clusters from earlier studies, the study applies several machine learning classifiers to categorise people. While SHAP [37] class-wise scores validation is employed for the model effectiveness and feature contribution in class, other evaluation metrics are used to understand the performance of the models.

In phase II, user interactions are recorded and converted into model characteristics in order to test the ML model. After that, the computer analyses user characteristics to determine user intent. Here, phase II data are labelled using the user profile developed in previous study. Lastly, Inter-Rater Reliability is utilised to assess the model's user prediction findings (with two human raters). Feature engineering, data modelling, and classification are done using PyCaret,1 Python pandas, and scikit-learn2. The Python SciPy3 statistics library is used to perform the Pearson Chi-square Test. Origin-Pro2023.4 and PyCaret are used for visualisation.

Disadvantages

• Data complexity: In order to identify online users by taking advantage of information-seeking behaviour, the majority of machine learning models now in use must be able to correctly understand sizable and intricate datasets.

• Data availability: In order to provide precise predictions, the majority of machine learning models need a lot of data. The accuracy of the model may degrade if data is not accessible in large enough amounts.

• Inaccurate labelling: The accuracy of the machine learning models that are now in use depends on how well the input dataset was used for training. Inaccurate labelling of the data prevents the model from producing reliable predictions.

PROPOSED SYSTEM

The purpose of the suggested system is to record the user's desire to get information and make it simple to display that information on the user interface. User navigation, personal preferences, likes and dislikes, search results, and the relevancy of search queries are all included in the study [8], as is semantic analysis of websites to better tailor search results to the user's purpose [9]. In order to comprehend user behaviour, social media user involvement and preferences are also examined. In order to create a search intent system, the research [10] analyses user interaction in picture searching and content using the physical features of the users. Another research [11] uses deep learning models to predict online purchase intentions by analysing clickstream data. [12] investigates search personalisation based on search and click history, while [13] uses information visualisation and reinforcement learning to model user intent. The

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025

authors of the research [14] have created a taxonomy for computer programming source code searches. Additionally, another paradigm for evaluating user intention in searching is Conversational Information Seeking (CIS), which involves communicating information demands to the search engine. Among other things, authors in [15] and [16] investigated user behaviour and search intent in conversational search.

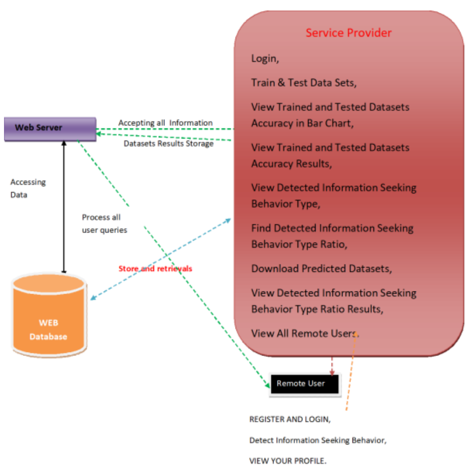
Advantages

• Examined user conduct and methods used to find, share, and validate information online. The three elements of user purpose to seek information have not been discussed, but together they provide a comprehensive picture of the user's goal to learn.

• Created a learning model for user classification.

• The classifier's dynamic behaviour was tested.

SYSTEM ARCHITECTURE



IV. IMPLEMENTATION

Modules

Service Provider

The Service Provider must use a working user name and password to log in to this module. Following a successful login, he may do several tasks including training and testing data sets, See the Accuracy of Trained and Tested Datasets in a Bar Chart View Accuracy Results for Trained and Tested Datasets, See the kind of information-seeking behaviour that was detected, determine the type ratio of that behaviour, Get Predicted Datasets here. View All Remote Users and the Results of the Detected Information Seeking Behaviour Type Ratio.

View and Authorize Users

The administrator may see a list of all registered users in this module. Here, the administrator may see the user's information, like name, email, and address, and they can also grant the user permissions.

Remote User

A total of n users are present in this module. Before beginning any actions, the user needs register. Following registration, the user's information will be entered into the database. Following a successful registration, he must use his password and authorised user name to log in. Following a successful login, the user will be able to see their profile, detect information-seeking behaviour, and register and log in.

ALGORITHMS

Logistic regression Classifiers

The relationship between a collection of independent (explanatory) factors and a categorical dependent variable is examined using logistic regression analysis. When the dependent variable simply has two values, like 0 and 1 or Yes and No, the term logistic regression is used. When the dependent variable contains three or more distinct values, such as married, single, divorced, or widowed, the technique is sometimes referred to as multinomial logistic regression. While the dependent variable's data type differs from multiple regression's, the procedure's practical application is comparable.

When it comes to categorical-response variable analysis, logistic regression and discriminant analysis are competitors. Compared to discriminant analysis, many statisticians believe that logistic regression is more flexible and appropriate for modelling the majority of scenarios. This is due to the fact that, unlike

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025

discriminant analysis, logistic regression does not presume that the independent variables are regularly distributed.

Both binary and multinomial logistic regression are calculated by this software for both category and numerical independent variables. Along with the regression equation, it provides information on likelihood, deviance, odds ratios, confidence limits, and quality of fit. It does a thorough residual analysis that includes diagnostic residual plots and reports. In order to find the optimal regression model with the fewest independent variables, it might conduct an independent variable subset selection search. It offers ROC curves and confidence intervals on expected values to assist in identifying the optimal classification cutoff point. By automatically identifying rows that are not utilised throughout the study, it enables you to confirm your findings.

Naïve Bayes

The supervised learning technique known as the "naive bayes approach" is predicated on the straightforward premise that the existence or lack of a certain class characteristic has no bearing on the existence or nonexistence of any other feature. However, it seems sturdy and effective in spite of this. It performs similarly to other methods of guided learning. Numerous explanations have been put forward in the literature. We emphasise a representation bias-based explanation in this lesson. Along with logistic regression, linear discriminant analysis, and linear SVM (support vector machine), the naive bayes classifier is a linear classifier. The technique used to estimate the classifier's parameters (the learning bias) makes a difference.

Although the Naive Bayes classifier is commonly used in research, practitioners who want to get findings that are useful do not utilise it as often. On the one hand, the researchers discovered that it is very simple to build and apply, that estimating its parameters is simple, that learning occurs quickly even on extremely big datasets, and that, when compared to other methods, its accuracy is rather excellent. The end users, however, do not comprehend the value of such a strategy and do not get a model that is simple to read and implement.

As a consequence, we display the learning process's outcomes in a fresh way. Both the deployment and comprehension of the classifier are simplified. We discuss several theoretical facets of the naive bayes classifier in the first section of this lesson. Next, we use Tanagra to apply the method on a dataset. We contrast the outcomes (the model's parameters) with those from other linear techniques including logistic regression, linear discriminant analysis, and linear support vector machines. We see that the outcomes are quite reliable. This helps to explain why the strategy performs well when compared to others. We employ a variety of tools (Weka 3.6.0, R 2.9.2, Knime 2.1.1, Orange 2.0b, and RapidMiner 4.6.0) on the same dataset in the second section. Above all, we make an effort to comprehend the outcomes.

Random Forest

Random forests, also known as random decision forests, are ensemble learning techniques that build a large number of decision trees during training for tasks like regression and classification. The class chosen by the majority of trees is the random forest's output for classification problems. The mean or average forecast of each individual tree is given back for regression tasks. The tendency of decision trees to overfit to their training set is compensated for by random decision forests. Although random forests are less accurate than gradient enhanced trees, they often perform better than choice trees. However, their performance may be impacted by data peculiarities.

Tin Kam Ho[1] developed the first algorithm for random decision forests in 1995 by using the

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025

random subspace technique, which in Ho's definition is a means of putting Eugene Kleinberg's "stochastic discrimination" approach to classification into practice.

Leo Breiman and Adele Cutler created an algorithm extension and filed for a trademark in 2006 for "Random Forests" (owned by Minitab, Inc. as of 2019). The extension builds a set of decision trees with controlled variance by combining Breiman's "bagging" concept with random feature selection, which was initially proposed by Ho[1] and then separately by Amit and Geman[13].

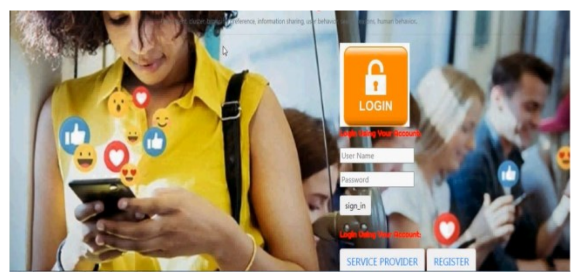
Businesses often employ random forests as "blackbox" models since they need minimal setup and provide accurate forecasts across a variety of inputs.

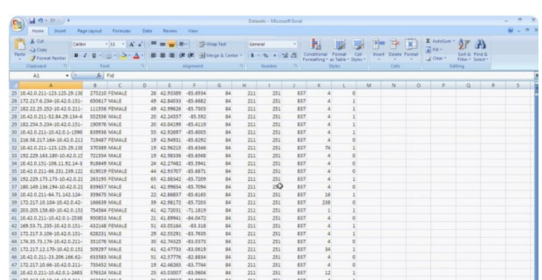
SVM

The goal of a discriminant machine learning approach in classification problems is to identify a discriminant function that can accurately predict labels for newly acquired instances based on an independent and identically distributed (iid) training dataset. A discriminant classification function takes a data point x and assigns it to one of the several classes that are part of the classification job, in contrast to generative machine learning techniques that call for calculations of conditional probability distributions. Discriminant techniques are less effective than generative approaches, which are mostly used when prediction entails the identification of outliers. However, they need less training data and processing resources, particularly when dealing with a multidimensional feature space and when just posterior probabilities are required. Finding the equation for a multidimensional surface that optimally divides the various classes in the feature space is the geometric equivalent of learning a classifier.

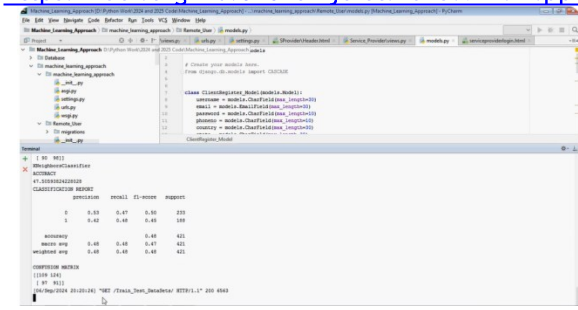
SVM is a discriminant approach that, unlike genetic algorithms (GAs) or perceptrons, which are both often used for classification in machine learning, always returns the same optimum hyperplane value since it solves the convex optimisation issue analytically. The initialisation and termination criteria have a significant impact on the solutions for perceptrons. While the perceptron and GA classifier models are distinct every time training is started, training yields uniquely specified SVM model parameters for a given training set for a certain kernel that converts the data from the input space to the feature space. The only goal of GAs and perceptrons is to reduce training error, which will result in several hyperplanes satisfying this criterion.

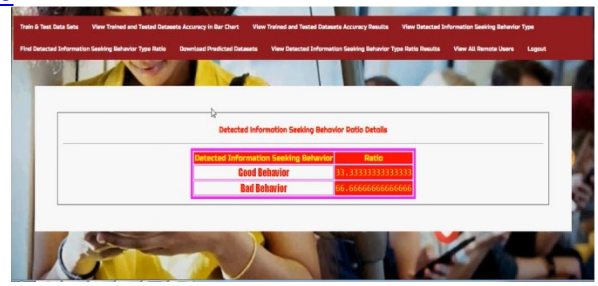
V. SCREEN SHOTS

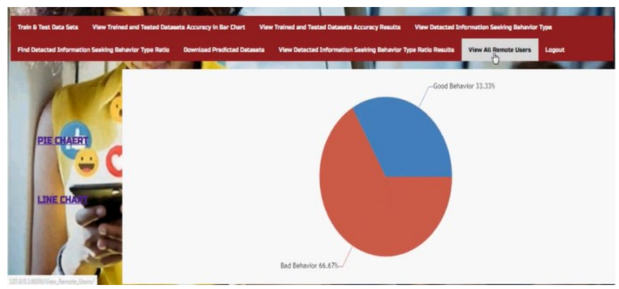


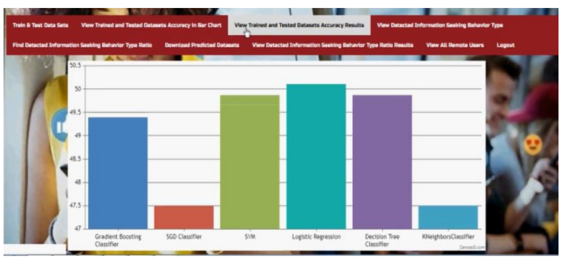


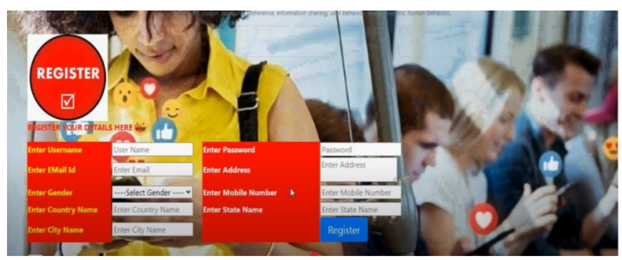
International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025

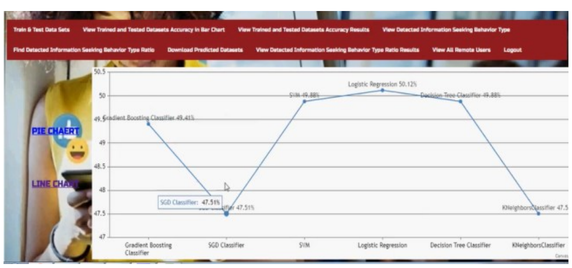


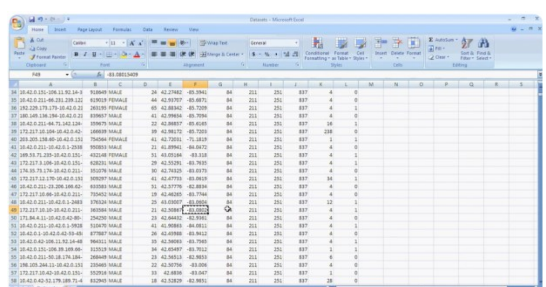


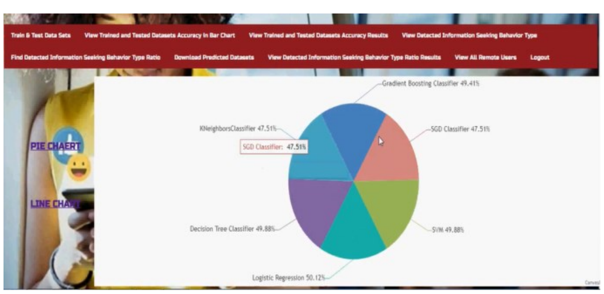




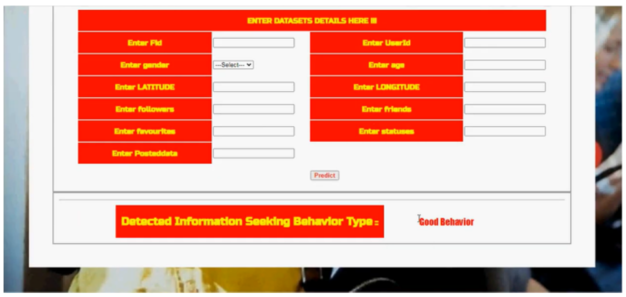


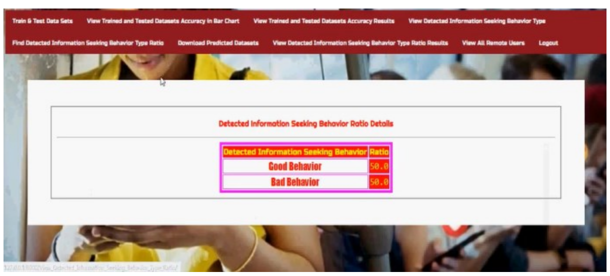


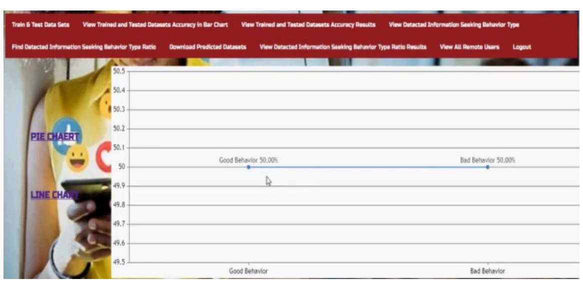




International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025







VI. CONCLUSION

The two issues this study seeks to address are: a) Is it possible to categorise user intentions based on their sharing, searching, and information-verification behaviours and practices? b) How may user interactions be used to evaluate the categorisation model? In order to do this, the study used a hybrid machine learning technique that effectively groups related individuals based on their online search, sharing, and verification habits while also making it easier to categorise new users. While Linear Discriminant Analysis (LDA) is a supervised classifier to forecast user classes, the unsupervised K-Means Clustering approach classifies users with comparable attributes. Five User Intent Classes (UIC) were found using the clustering procedure, and LDA categorised the test data with an accuracy of around 80%. Committed (those who balance work and leisure), Casual Surfer (search and share for leisure with minimal need to authenticate information), NetVenturer (use the internet for all purposes but keep eyes open), Focused (primarily for serious work), and Aware (up-to-date, wise internet usage) are the five distinct user profiles that emerged from the analysis. By supplying the information that users often want, these results may be further investigated to customise online services to employ generic user profiles that can be safely used to improve the user experience. The effectiveness of the UIC model is tested by user online engagement and activities. Inter-rater reliability (IRR) is calculated to verify the annotation, user profile, and UIC model prediction after human raters have annotated the interaction data and LDA model. The model's ability to forecast users on dynamic data was successfully evaluated by the experts, who acknowledged the 60% IRR. The study helps provide a framework for user profiling based on distinct three-dimensional information, such as search intent, information dissemination requirements, and content trustworthiness standards. Although not as a part of user intent, some facets of user behaviour have been studied independently. Since information is accessible across all internet channels in our digital age, information integrity is a crucial consideration. According to the study, search and sharing are

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025

correlated, but verification strengthens the user's motivation to seek the truth. Based on user profile, the study's findings may be used to improve search filter design, SEO, targeted advertising, and the shopping experience.

REFERENCES

[1] L. S. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA, USA: Harvard Univ. Press, 1978.

[2] K. A. Mills, ‘‘Shrek meets vygotsky: Rethinking adolescents’ multimodal literacy practices in schools,’’ *J. Adolescent Adult Literacy*, vol. 54, no. 1,pp. 35–45, Sep. 2010.

[3] A. Halevy, C. Canton-Ferrer, H. Ma, U. Ozertem, P. Pantel, M. Saeidi,F. Silvestri, and V. Stoyanov, ‘‘Preserving integrity in online social networks,’’*Commun. ACM*, vol. 65, no. 2, pp. 92–98, Feb. 2022.

[4] X. Feng, X. Wang, and Y. Zhang, ‘‘Research on the effect evaluation and the time-series evolution of public culture’s Internet communication under the background of new media: Taking the information dissemination of red tourism culture as an example,’’ *J. Comput. Cultural Heritage*, vol. 16,no. 1, pp. 1–15, Mar. 2023.

[5] C. I. Eke, A. A. Norman, L. Shuib, and H. F. Nweke, ‘‘A survey of user profiling: State-of-the-art, challenges, and solutions,’’ *IEEE Access*, vol. 7,pp. 144907–144924, 2019.

[6] J. Liu, M. Mitsui, N. J. Belkin, and C. Shah, ‘‘Task, information seeking intentions, and user behavior: Toward a multi-level understanding of web search,’’ in *Proc. Conf. Human Inf. Interact. Retr.*, Glasgow, U.K.,Mar. 2019, pp. 123–132.

[7] J. L. Hale, B. J. Householder, and K. L. Greene, ‘‘The theory of reasoned action,’’ in *The Persuasion Handbook: Developments in Theory and Practice*,vol. 14. Newbury Park, CA, USA: Sage, 2002, pp. 259–286.

[8] J. Shi, P. Hu, K. K. Lai, and G. Chen, ‘‘Determinants of users’ information dissemination behavior on social networking sites: An elaboration likelihood model perspective,’’ *Internet Res.*, vol. 28, no. 2, pp. 393–418, Apr. 2018.

[9] P. Bedi, S. B. Goyal, A. S. Rajawat, R. N. Shaw, and A. Ghosh, ‘‘A framework for personalizing atypical web search sessions with concept-based user profiles using selective machine learning techniques,’’ in *Advanced Computing and Intelligent Technologies* (Lecture Notes in Networks and Systems), vol. 218. Singapore: Springe, 2022.

[10] M. Soleymani, M. Riegler, and P. Halvorsen, ‘‘Multimodal analysis of user behavior and browsed content under different image search intents,’’ *Int. J. Multimedia Inf. Retr.*, vol. 7, no. 1, pp. 29–41, Mar. 2018, doi: 10.1007/s13735-018-0150-6.

[11] D. Koehn, S. Lessmann, and M. Schaal, ‘‘Predicting online shopping behaviour from clickstream data using deep learning,’’ *Expert Syst. Appl.*,vol. 150, Jul. 2020, Art. no. 113342.

[12] H. Yoganarasimhan, ‘‘Search personalization using machine learning,’’ *Manage. Sci.*, vol. 66, no. 3, pp. 1045–1070, Mar. 2020.

[13] T. Ruotsalo, J. Peltonen, M. J. Eugster, D. Głowacka, P. Floreen,P. Myllymaki, G. Jacucci, and S. Kaski, ‘‘Interactive intent modeling for exploratory search,’’ *ACM Trans. Inf. Syst.*, vol. 36, no. 4, p. 44, Oct. 2018,doi: 10.1145/3231593.

[14] S. K. Shivakumar, ‘‘A survey and taxonomy of intent-based code search,’’ *Int. J. Softw. Innov.*, vol. 9, no. 1, pp. 69–110, Jan. 2021.

[15] P. Ren, Z. Liu, X. Song, H. Tian, Z. Chen, Z. Ren, and M. de Rijke, ‘‘Wizard of search engine: Access to information through conversations with search engines,’’ in *Proc. 44th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*,Jul. 2021.

[16] A. Salle, S. Malmasi, O. Rokhlenko, and E. Agichtein, ‘‘Studying the effectiveness of conversational search refinement through user simulation,’’ in *Advances in Information Retrieval* (Lecture Notes in Computer Science). Cham, Switzerland: Springer, 2021, pp. 587–602, doi: 10.1007/978-3-030-72113-8_39.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1062-1073
Vol. 21, Issue 2, 2025

[17] S. Dixon. *Statista*. Accessed: Jan. 2024. [Online]. Available: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-numberof-users/

[18] S.-F. Tsao, H. Chen, T. Tisseverasinghe, Y. Yang, L. Li, and Z. A. Butt, ''What social media told us in the time of COVID-19: A scoping review,'' *Lancet Digit. Health*, vol. 3, no. 3, pp. e175–e194, Mar. 2021.

[19] Z. Wu, Y. Liu, Q. Zhang, K. Wu, M. Zhang, and S. Ma, ''The influence of image search intents on user behavior and satisfaction,'' in *Proc. 12th ACM Int. Conf.Web Search Data Mining*, Melbourne, VIC, Australia, Jan. 2019, pp. 645–653.

[20] D. Fraszczak, ''Information propagation in online social networks—A simulation case study,'' in *Proc. 38th Int. Bus. Inf. Manag. Assoc.*, Seville, Spain, 2021, pp. 23–24.

[21] X. Lin and X. Wang, ''Examining gender differences in people's information-sharing decisions on social networking sites,'' *Int. J. Inf. Manage.*,vol. 50, pp. 45–56, Feb. 2020.

[22] A. Mohammed and A. Ferraris, ''Factors influencing user participation in social media: Evidence from Twitter usage during COVID-19 pandemic in Saudi Arabia,'' *Technol. Soc.*, vol. 66, Aug. 2021, Art. no. 101651.

[23] S.-H. Liao, R. Widowati, and C.-J. Cheng, ''Investigating Taiwan Instagram users' behaviors for social media and social commerce development,'' *Entertainment Comput.*, vol. 40, Jan. 2022,Art. no. 100461, doi: 10.1016/j.entcom.2021.100461.

[24] X. Wang, F. Chao, G. Yu, and K. Zhang, ''Factors influencing fake news rebuttal acceptance during the COVID-19 pandemic and the moderating effect of cognitive ability,'' *Comput. Hum. Behav.*, vol. 130, May 2022,Art. no. 107174, doi: 10.1016/j.chb.2021.107174.

[25] R. P. Yu, ''How types of Facebook users approach news verification in the mobile media age: Insights from the dual-information-processing model,'' *Mass Commun. Soc.*, vol. 24, no. 2, pp. 233–258, 2021, doi: 10.1080/15205436.2020.1839104.