International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: XII Month of publication: December 2025

DOI: https://doi.org/10.22214/ijraset.2025.76429

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XII Dec 2025- Available at www.ijraset.com

A Comparative Survey of Personality Trait Prediction Using Digital Footprints with Emphasis on Spending Behaviour

Mrs. Arpitha C. N.[1], Dr. Sunitha M. R.[2]
[1]*Department of Information Science and Engineering, Adichunchanagiri Institute of Technology, Chikkamagaluru, Karnataka, India,* Visvesvaraya Technological University, Belagavi-590018
[2]*Professor and Head, Department of AIML, Adichunchanagiri Institute of Technology, Chikkamagaluru, Karnataka, India, Visvesvaraya Technological University, Belagavi-590018*

Abstract: The rapid growth of digital footprints has enabled large-scale analysis of human psychological characteristics using computational techniques. Among various digital traces, spending behaviour captured through transaction data provides an objective and continuous representation of real-world decision-making. This paper presents a survey and comparative analysis of personality prediction using digital footprints, with a primary focus on expenditure-based behavioural data. Existing studies leveraging questionnaire responses, social media activity, smartphone sensing, and multimodal data sources are systematically reviewed and compared with transaction-based approaches. Machine learning and deep learning models are analysed for their effectiveness in predicting Big Five personality traits along with additional traits such as self-control and materialism. The findings indicate that while predictive accuracy for broad personality dimensions remains moderate, more specific traits exhibit stronger and more stable associations with spending patterns. Furthermore, spending-based personality predictions demonstrate relative robustness across different socioeconomic groups when compared to other digital behavioural sources. The survey highlights the potential of transaction data as a scalable and ecologically valid alternative for personality assessment, while also emphasizing the ethical and privacy considerations associated with large-scale psychological profiling.
Keywords: Personality Trait Prediction, Digital Footprints, Spending Behaviour, Big Five Personality Traits, Machine Learning, Transaction Data.

I. INTRODUCTION

Personality refers to combination of behaviors and emotions, which have a unique set of qualities depending on the environment and biological causes. This will be varying from one person to another person depending on many conditions based on their thinking and feelings [1]. The person quality will construct in nature based on their activities, thinking, feelings, and overall behavior in different conditions. Generally, this activity of any human is based on his nature and knowledge. This varies from person to person and is studied in personality psychology. To measure different behavior for different individuals is that the consistency in many situations with stability. Recently, social sites are increasing abruptly for different communication purposes. People are using those platforms to share their thoughts, expectations, and feelings. Based on their likes and unlike activities, they can be categorized easily. This information is collected as a data set by social media [2]. Initially, this information is used for banking sectors to identify the person's location and position. The extracted data has been explored year by year in many sectors. Besides, improving the quality of the services and products is done by this phenomenon.

II. BIG FIVE PERSONALITY TRAITS

There are five basic personality aspects, according to many research and analyses in psychology. The 1949 principal theory served as the foundation for the evidence supporting this notion. The theory names extraversion, sometimes known as extroversion, conscientiousness, agreeableness, openness, and neuroticism as the five fundamental personality behaviors or qualities. D. W. Fiske created the five fundamental personality traits and underlying theory in 1949; later, academics like Norman, Smith, Goldberg, McCrae, and Costa added to it in 1967, 1981, and 1987[3]. The two more personality traits considered in this survey are self-control and materialism.

A. Openness

Openness reflects an individual's tendency toward imagination, curiosity, and intellectual exploration. People with high openness actively seek new knowledge, experiences, and ideas, and they are generally receptive to change and innovation. Such individuals enjoy creative thinking, problem-solving from multiple perspectives, and engaging in activities that expand their understanding of the world. They are often comfortable with abstract reasoning and conceptual thought, which supports adaptability in decision-making situations. Examples include individuals who enjoy experimenting with new cuisines, traveling to unfamiliar destinations, or exploring diverse cultures and ideas. In contrast, individuals with low openness prefer familiarity, routine, and traditional approaches to life. They may feel uncomfortable when faced with novel situations or problems that require thinking beyond their established knowledge base.

B. Conscientiousness

Conscientiousness is associated with self-discipline, responsibility, and a strong sense of goal orientation. Individuals who score high on conscientiousness tend to be organized, methodical, and dependable in both personal and professional contexts. They plan ahead, adhere to schedules, and carefully evaluate their actions to ensure they align with long-term objectives. This trait is commonly observed in professions that demand precision, structure, and accountability, such as project management, finance, research, and administrative roles. Highly conscientious individuals are also mindful of how their behavior affects others and consistently strive for efficiency and quality. Conversely, individuals with low conscientiousness may struggle with time management, resist structured environments, and frequently leave tasks incomplete, which can negatively impact performance and reliability.

C. Extraversion

Extraversion, also referred to as extroversion, describes a personality trait characterized by sociability, enthusiasm, and high levels of external energy. Extroverted individuals feel motivated and energized through social interactions and are often expressive, talkative, and confident in group settings. They enjoy meeting new people, participating in social activities, and taking leadership roles in public or collaborative environments. Due to these qualities, extroverts are frequently found in professions that involve communication and public engagement, such as sales, marketing, education, and politics. In contrast, individuals low in extraversion tend to prefer solitude, exhibit reserved behavior, and experience lower levels of social stimulation. These individuals may communicate less frequently but often excel in independent or reflective roles.

D. Agreeableness

Agreeableness represents an individual's inclination toward compassion, cooperation, and social harmony. People with high agreeableness demonstrate kindness, empathy, trust, and a genuine concern for the well-being of others. They are supportive in difficult situations and often prioritize collaboration, sharing, and conflict resolution. Such individuals are naturally drawn to careers that emphasize helping and caregiving, including healthcare, counseling, social work, and charitable services. On the other hand, low agreeableness may manifest as uncooperative behavior, lack of empathy, and a tendency toward interpersonal conflict. Individuals with lower agreeableness may find it challenging to maintain harmonious relationships and may show limited interest in addressing the needs or concerns of others.

E. Neuroticism

Neuroticism is associated with emotional sensitivity, mood instability, and heightened responses to stress. Individuals high in neuroticism are more likely to experience emotions such as anxiety, irritability, sadness, and emotional tension, particularly in demanding or uncertain situations. This trait reflects how a person perceives and reacts to stress, perceived threats, or pressure in daily life. Highly neurotic individuals may over analyze situations, struggle with emotional regulation, and find it difficult to cope with change. In contrast, individuals with low neuroticism demonstrate emotional stability, resilience, and a calm approach to stress. They are better equipped to manage pressure and recover quickly from challenging experiences.

F. Self-Control

Self-control refers to an individual's ability to regulate impulses, emotions, and behaviors in pursuit of long-term goals. People with strong self-control can resist immediate temptations and make thoughtful decisions aligned with their personal or professional objectives.

This trait plays a crucial role in financial discipline, healthy lifestyle choices, and consistent performance. Individuals with high self-control are more likely to manage their spending responsibly, maintain focus, and avoid impulsive actions. In contrast, low self-control may result in impulsive decision-making, difficulty in delaying gratification, and inconsistent behavior patterns, which can negatively influence both financial stability and personal development.

G. Materialism

Materialism reflects the importance an individual places on material possessions and financial success as indicators of happiness and achievement. Highly materialistic individuals often associate self-worth and life satisfaction with wealth, luxury goods, and visible symbols of success. Their consumption patterns tend to emphasize branded products, high-value purchases, and status-driven spending. While materialism can motivate ambition and economic activity, excessive material focus may reduce satisfaction derived from non-material aspects of life, such as relationships or personal growth. Individuals with low materialism place less emphasis on possessions and are more likely to value experiences, emotional well-being, and social connections over financial display.

III.LITERATURE SURVEY

Personality trait prediction has attracted significant attention in recent years due to its applications in personalized systems, behavioral analytics, and human–computer interaction. With the availability of large-scale digital data, researchers have explored multiple data modalities for inferring psychological traits using machine learning and deep learning techniques. Existing studies can be broadly categorized based on the type of data source employed for personality inference, including questionnaire-based data, social media data, smartphone and sensor data, and multimodal behavioural data. Comparative overview of data sources and methods for personality trait prediction is represented in table 1:

A. Questionnaire and Psychometric Test Data

Questionnaire-based personality assessment remains the most established and theoretically grounded approach for personality trait classification. Standard psychometric instruments such as the Big Five Inventory (BFI) and NEO Personality Inventory-Revised (NEO-PI-R) have been widely used to generate reliable personality labels for supervised learning models. Shenavi *et al.* developed a personality classification framework using structured questionnaire responses and applied machine learning algorithms such as Random Forest, Logistic Regression, and Support Vector Machine for Big Five trait prediction. Their results demonstrated that Random Forest achieved superior accuracy due to its capability to model non-linear relationships among behavioural indicators derived from questionnaire responses [1]. However, the study highlighted limitations related to scalability and respondent bias inherent in self-reported data.Chi *et al.* conducted a large-scale personality analysis using more than one million Big Five questionnaire responses. By combining k-means clustering, discriminant analysis, and multilayer perceptron neural networks, the study achieved high classification accuracy and revealed that Neuroticism, Conscientiousness, and Openness were dominant traits in personality differentiation [2]. Despite strong predictive performance, the reliance on questionnaire data limited real-world deployment for continuous personality assessment.

B. Social Media Data

Social media platforms provide rich behavioural and linguistic information that reflects users' psychological characteristics. Textual content, posting frequency, emotional expression, and interaction patterns have been extensively explored for personality prediction. Karanatsiou *et al.* proposed a multi-output regression framework using Twitter data to predict Big Five personality traits. The model integrated linguistic features, emotional indicators, and behavioural activity patterns, achieving improved accuracy by jointly modelling correlations between traits such as Openness and Extraversion [3]. Their work demonstrated the effectiveness of combining heterogeneous social media features for personality inference. Billy *et al.* applied BERT-based contextual embeddings to Twitter posts and combined them with classical classifiers for personality prediction. Their hybrid model outperformed traditional TF–IDF-based approaches, particularly in predicting Extraversion and Openness, highlighting the importance of deep semantic representations in social media-based personality classification [4].

C. Smartphone and Sensor Data

Smartphone-based personality prediction leverages passive sensing data to capture real-world behavioral patterns without active user input. Sensor streams such as GPS location, call logs, messaging behavior, and application usage have been shown to correlate with personality traits.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XII Dec 2025- Available at www.ijraset.com

Lane *et al.* introduced a personality sensing framework that utilized smartphone sensor data to infer Big Five traits. Their study demonstrated strong associations between mobility patterns and Extraversion, as well as regular behavioral routines and Conscientiousness [5]. The framework enabled continuous and unobtrusive personality assessment but raised concerns related to privacy and long-term data collection. Wang *et al.* explored temporal behavioral patterns from smartphone sensors using machine learning techniques and showed that incorporating sequential activity information significantly improves personality prediction accuracy. Their findings emphasized the importance of time-dependent behavioral modelling for psychological inference [6].

D. Multimodal and Behavioural Data (Including Spending Behaviour)

Multimodal personality prediction combines multiple data sources to capture complementary psychological cues and improve robustness. Recently, spending behavior has gained attention as an objective behavioural signal directly linked to real-world decision-making. Wang *et al.* proposed a multimodal framework that fused questionnaire responses with textual behavioural data to predict proactive personality traits. Their results confirmed that integrating structured psychometric data with unstructured behavioural features significantly enhanced classification performance compared to single-modality approaches [7].

Matz *et al.* investigated the relationship between consumer spending behaviour and psychological traits, demonstrating that transactional purchase patterns can reliably predict traits such as Self-Control and Materialism. Their study established financial spending data as a powerful yet underutilized modality for personality profiling, motivating further exploration using advanced machine learning and deep learning techniques [8].

Table 1: Comparative Overview of Data Sources and Methods for Personality Trait Prediction

Data Source	Commonly Used Algorithms	Key Findings	Limitations
Questionnaire & Psychometric Data	Random Forest, Logistic Regression, Support Vector Machine (SVM), k-means clustering, Discriminant Analysis, Multilayer Perceptron (MLP).	High predictive accuracy for Big Five traits due to structured and validated inputs; strong theoretical grounding and label reliability.	Limited scalability, respondent fatigue, self-report bias, not suitable for continuous or real-time assessment.
Social Media Data	Multi-output Regression, Support Vector Machine, Logistic Regression, BERT-based deep learning models, Neural Networks.	Linguistic, emotional, and behavioral cues enable effective prediction of traits such as Openness and Extraversion; deep semantic models improve performance.	Platform dependency, noisy and sparse data, privacy concerns, language and cultural bias.
Smartphone & Sensor Data	Random Forest, Support Vector Machine, Sequential ML models, Temporal pattern analysis.	Passive sensing captures real-world behaviour; mobility and routine patterns correlate strongly with Extraversion and Conscientiousness.	High privacy risk, long-term data collection challenges, sensor availability and user consent issues.
Multimodal & Behavioural Data (Including Spending)	Random Forest, Gradient Boosting, Neural Networks, Deep Learning fusion models.	Combining multiple data sources improves robustness; spending behaviour reliably predicts focused traits such as Self-Control and Materialism.	Increased model complexity, ethical concerns.

IV. METHODOLOGY

Fig. 1 presents the end-to-end framework for predicting human psychological traits using transaction data. The process begins with data collection, where individual spending records are gathered, followed by data pre-processing to clean, normalize, and prepare the data for analysis. Next, feature extraction is performed to derive meaningful spending behavior indicators from the transactional data. These extracted features are refined through feature selection, where relevant attributes are identified using statistical methods such as Pearson correlation. The selected features are then provided as input to prediction models, which include both machine learning and deep learning algorithms. Finally, the trained models analyze spending behavior patterns and produce the output in the form of predicted psychological traits. [9][10].

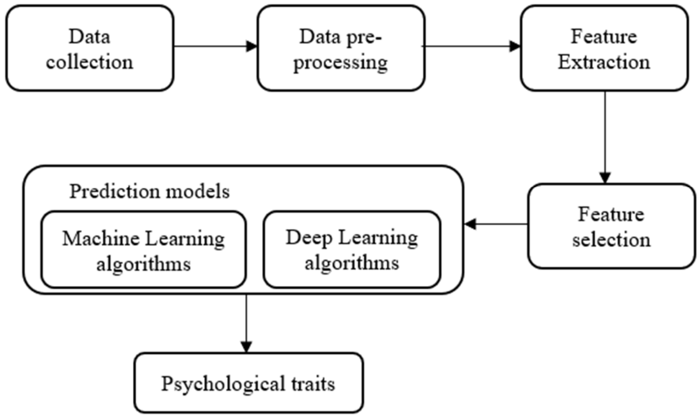


Fig. 1. Architecture diagram for human psychological behavior analysis from spending categories

A. Impact of spending behavior on psychological behavior.

The detailed records of every purchase completed with a person's aid were included in the transaction records obtained from their bank account's credit and debit details. The expenditure list is produced only by using the debit data. Spending lists were automatically divided into groups such as household, supermarkets, kids, transport, furniture stores, and insurance policies.[11][12]. The positive and negative traits that correspond with each personality trait are listed in Table 2.

Table 2: Personality traits and their positive, negative correlation

Psychological Trait	Most Positively Correlated Spending Categories	Most Negatively Correlated Spending Categories
Openness	Travel & Tourism, Entertainment, Education, Lifestyle & Hobbies, Dining & Exotic Food	Utilities, Insurance, Routine Household Expenses
Conscientiousness	Groceries, Utilities, Insurance, Healthcare, Savings & Investments	Entertainment, Luxury Goods, Impulsive Online Shopping
Extraversion	Dining & Restaurants, Entertainment, Travel, Social Events, Fashion	Utilities, Insurance, Solitary Home-based Expenses
Agreeableness	Charity & Donations, Healthcare, Family & Household Needs, Education	Luxury Goods, Gambling, Self-centered High-End Purchases
Neuroticism	Online Shopping, Entertainment Subscriptions, Health-related Expenses, Instant Purchases	Long-term Investments, Savings, Planned Financial Products
Self-Control	Savings, Investments, Essential Groceries, Utilities	Impulsive Purchases, Luxury Goods, Entertainment & Gaming
Materialism	Luxury Goods, Electronics, Fashion & Accessories, Branded Products	Charity & Donations, Savings, Basic Household Expenses

B. Illustration of using machine learning and deep learning models to analyse human behaviour

Machine learning and deep learning models are designed to emulate human reasoning by learning patterns from data and progressively reducing uncertainty in predictions [2][13]. Similar to how a psychologist evaluates a patient by considering multiple symptoms before arriving at a diagnosis, these models analyze a combination of behavioral indicators rather than relying on a single feature. Through iterative learning, the models identify relationships between inputs and outcomes, enabling accurate classification of behavioral states. When applied to transaction data, the models examine spending behaviors as observable manifestations of psychological traits. For instance, higher expenditure on dining, entertainment, and social activities—categories commonly linked to extraversion—can increase the likelihood of predicting an extroverted personality [1].

Similarly, materialism can be inferred by assessing whether an individual prioritizes spending on shopping, dining, and mobile devices over charitable contributions and other socially driven expenses [14][15]. By integrating evidence from multiple spending features, machine learning and deep learning algorithms effectively narrow the range of possible psychological outcomes. This data-driven approach enables robust and scalable analysis of human behavior, making it suitable for predicting psychological traits based on transactional spending patterns.

V. RESULTS AND DISCUSSIONS

The survey results indicate that digital expenditure data can be used to predict personality traits at a scale that was previously difficult to achieve, contributing to advancements in automated psychological trait prediction. Although some prediction error is present, more specific psychological traits—particularly materialism and self-control—show higher predictive accuracy. The predictions remain largely consistent across respondents with different levels of financial resources, with only minor variations observed based on the level of deprivation in the participant's local area.

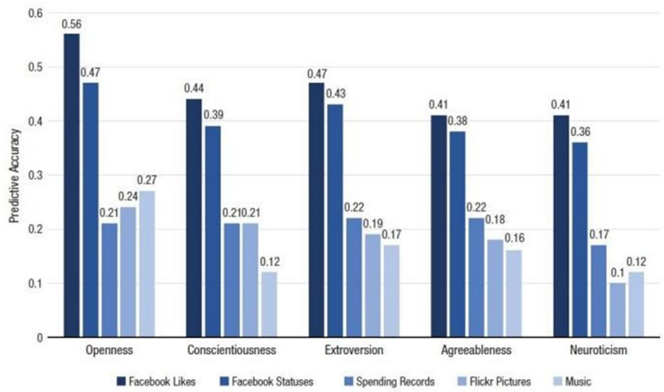


Fig. 2. Predictive accuracies of the Big Five traits,

Predictive accuracy of the main five qualities is depicted in Fig. 2 and compares Facebook Likes, Facebook Statuses, Spending History, Instagram Pictures, and Music Preferences. Bars show relationships that adjust for attenuation. Twitter Followers 2014 Facebook status updates from Youyou, Kosinski, and Stillwell Gladstone et al. 2021, Flickr images Park et al., 2014, expenditure records for this study preferences for music, Segalin, Perina, Cristani, and Vinciarelli, 2017, 2018 [16] Nave et al.

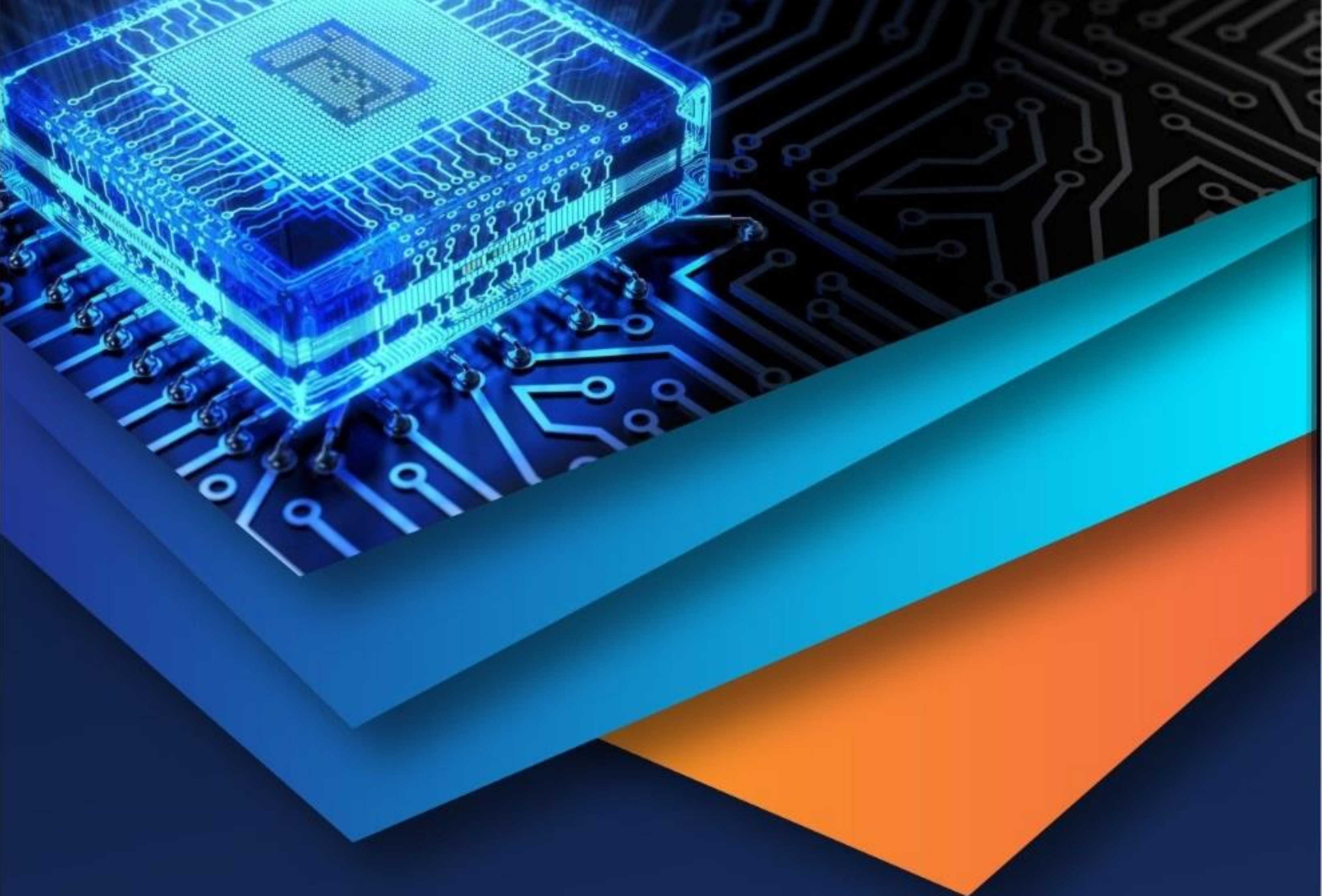
VI. CONCLUSION

This paper presented a comprehensive survey and comparative analysis of human psychological personality prediction using digital footprints, with a particular focus on spending behaviour derived from transaction data. By examining existing research across questionnaire-based methods, social media data, smartphone sensing, and multimodal approaches, the study highlighted the growing role of machine learning and deep learning techniques in large-scale personality inference. The analysis demonstrated that spending behaviour serves as a reliable and objective digital footprint for predicting both Big Five personality traits and additional traits such as self-control and materialism. Compared to other digital sources, spending-based predictions exhibit relatively stable performance across different socioeconomic groups and over time. While predictive accuracies remain moderate for broader personality dimensions, more focused traits show stronger associations with expenditure patterns. Overall, this survey underscores the potential of transaction data as a scalable alternative for psychological assessment, while also emphasizing the need for responsible use and ethical safeguards as digital personality prediction technologies continue to evolve.

REFERENCES

[1] S. Shenavi, R. K. Rao, and P. R. Kumar, "Personality trait classification using machine learning techniques based on questionnaire data," International Journal of Advanced Computer Science and Applications, vol. 10, no. 6, pp. 412–419, 2019.
[2] L. Chi, J. Li, and X. Fu, "Large-scale personality trait analysis using Big Five questionnaire and neural networks," IEEE Access, vol. 7, pp. 176589–176600, 2019.
[3] A. Karanatsiou, M. Thelwall, and K. Kappas, "Predicting Big Five personality traits from Twitter data using multi-output regression," Computers in Human Behavior, vol. 86, pp. 1–10, 2018.
[4] J. Billy, S. Guntuku, and L. Ungar, "Personality prediction from social media text using contextual word embeddings," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 5, pp. 697–704, 2020.

[5] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, "A survey of mobile phone sensing," IEEE Communications Magazine, vol. 48, no. 9, pp. 140–150, 2010.
[6] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. T. Campbell, and A. T. Campbell, "StudentLife: Assessing mental health, academic performance and behavioral trends of college students using smartphones," Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 3–14, 2014.
[7] Y. Wang, J. Li, and S. Zhao, "Multimodal personality prediction based on questionnaire and behavioral data fusion," Expert Systems with Applications, vol. 113, pp. 456–465, 2018.
[8] S. C. Matz, M. Kosinski, G. Nave, and D. J. Stillwell, "Psychological targeting as an effective approach to digital mass persuasion," Proceedings of the National Academy of Sciences (PNAS), vol. 114, no. 48, pp. 12714–12719, 2017.
[9] Pooja Sharma, Sachin Kumar, Analyzing Customer Behavior on Online Products Using Decision Tree Machine Learning 2022 International Conference on Advances in Computing, Communication and Materials (ICACCM), ISBN:978-1-6654- 7439-9, Dehradun,India,10-11 Nov. 2022
[10] R. Valanarasu, , Comparative Analysis for Personality Prediction by Digital Footprints in Social Media, Journal of Information Technologyand Digital World (2021) Vol. 03/ No. 02 Pages: 77-91
[11] Manil Vaidhya; Bikash Shrestha; Bishal Sainju; Kiran Khaniya; Aman Shakya, Personality Traits Analysis from Facebook Data,,21st International Computer Science and Engineering Conference (ICSEC) ,DOI: 10.1109/ICSEC41384.2017,15-18 Nov. 2017
[12] Kaan Sonmezoz; Ozgur Ugur; Banu Diri,MBTI Personality Prediction With Machine Learning,2020 28th Signal Process- ing and Communications Applications Conference (SIU),DOI: 10.1109/SIU49456.20205-7 Oct. 2020
[13] Avishka Jaysundara; Dimanthinie De Silva; Priyantha Kumarawadu, Personality Prediction of Social Network Users using LSTM based Sentiment Analysis,2022 International Conference on Smart Tech- nologies and Systems for Next Generation Computing (ICSTSN), DOI: 10.1109/ICSTSN53084.2022.
[14] Natkamon Tovanich, Simone Centellegher, Nacera Bennacer Seghouani, Joe Gladstone, Sandra Matz and Bruno Lepri, Inferring psychological traits from spending categories and dynamic consumption patterns, EPJ Data Science volume 10, Article number: 24 (2021)
[15] Sanjit Kumar R, Shrivatson R G, Rishi Priyan S, Padmavathy T, Personality Prediction using Twitter Data, InternationalResearch Journal of Engineering and Technology (IRJET) e-ISSN: 2395- 0056, Volume: 07, Issue: 07 — July 2020, ISSN: 2395-007.
[16] Gladstone, J. J., Matz, S. C., Lemaire, A. Can psychological traitsbe inferred from spending? Evidence from transaction data.Evidence from transaction data. Psychological Science, 30(7), 1087–1096,2019



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY