Decoding User Trust in Crypto Wallets with a BERT–XGBoost Hybrid Model for Multilingual Phantom Review Analysis

R. Elavarasi[1,*], P. Gajalakshmi[2], S. Murugesan[3], K. Srinivasan[4]

[1,3]Electrical and Electronics Engineering, AMET Deemed to be University, Tamil Nadu, India
[2]Electronics and Communication Engineering, University College of Engineering, India
[4]Electrical and Electronics Engineering, Madras Engineering College, Tamil Nadu, India

ABSTRACT

The rapid expansion of decentralized financial applications has increased the importance of understanding user trust in crypto wallet platforms. This study examines trust expressions in multilingual Phantom Wallet reviews using a hybrid classification framework that integrates BERT-based contextual embeddings with an XGBoost model. A total of 12,422 English and Indonesian reviews were collected and processed to construct a multilingual dataset for trust analysis. Exploratory findings reveal a highly polarized distribution of user ratings, indicating that trust in crypto wallets is strongly influenced by clear satisfaction or dissatisfaction rather than moderate evaluations. Cross-linguistic analysis indicates that Indonesian users express a higher proportion of low-trust reviews compared to English users, suggesting greater sensitivity to transaction errors and perceived asset safety concerns. Lexical patterns demonstrate that positive trust is associated with usability and performance stability, while negative trust is primarily driven by system failures, delays, and missing balance incidents. The results confirm that the BERT–XGBoost hybrid model is well-suited for decoding trust-related signals by combining contextual semantic understanding with structured metadata. This study contributes to the broader discourse on digital trust within Web3 environments by demonstrating an effective multilingual machine learning approach for analysing user perceptions in decentralized financial technologies.

Keywords BERT, Xgboost, User Trust, Multilingual Text Analysis, Crypto Wallet Reviews

Submitted: 15 August 2025
Accepted: 25 October 2025
Published: 26 May 2026

Corresponding author
R. Elavarasi,
elavarasir2014@gmail.com

Additional Information and Declarations can be found on page 136

DOI: 10.47738/jcrb.v3i2.68

INTRODUCTION

The rapid advancement of blockchain technology has fundamentally reshaped the architecture of modern digital finance. As decentralized ledger systems continue to proliferate, users increasingly rely on blockchain applications to store, manage, and transfer digital assets without the need for intermediaries such as banks or centralized payment processors [1]. This shift toward decentralization has introduced new forms of financial autonomy and transparency, yet it has also heightened the demand for secure, reliable, and user-friendly tools. Among these tools, crypto wallets serve a pivotal role as the primary interface through which individuals access blockchain networks, execute transactions, and protect cryptographic keys that authenticate ownership of digital assets [2].

Given the sensitive nature of asset management, user trust has emerged as a critical determinant of adoption and long-term engagement in crypto wallet platforms. Trust influences not only the frequency of transactions but also users' willingness to continue utilizing the platform despite potential system vulnerabilities or market risks. Existing research shows that users tend to

How to cite this article: R. Elavarasi, P. Gajalakshmi, S. Murugesan and K. Srinivasan, "Decoding User Trust in Crypto Wallets with a BERT–XGBoost Hybrid Model for Multilingual Phantom Review Analysis," *J. Curr. Res. Blockchain,* vol. 3, no. 2, pp. 125-138, 2026.

evaluate trust based on their perceptions of system reliability, transparency, performance responsiveness, and protection against loss or unauthorized access [3]. In decentralized financial ecosystems, where no central authority can intervene during technical failures or fund losses, user trust becomes an even more essential element of platform sustainability.

In the context of crypto wallets, trust is shaped primarily through experiential interactions. When users encounter smooth transactions, intuitive interfaces, and responsive performance, they are more likely to express positive trust. Conversely, technical malfunctions such as slow transaction confirmation, freezing applications, compatibility issues, or unexplained reductions in displayed token balances can significantly erode confidence [4]. These negative experiences often manifest in user-generated reviews, making review platforms an important and accessible source for analysing the dynamics of trust. Online reviews also contain rich linguistic cues that reflect users' emotional states, perceived risks, and expectations regarding platform stability.

The global nature of blockchain adoption introduces additional complexity. As users from various linguistic and cultural backgrounds interact with crypto wallets, trust expressions may vary across languages. Linguistic structures, cultural expectations, and communication styles influence how users articulate satisfaction or dissatisfaction. Prior studies highlight that multilingual analysis is essential in evaluating global applications because trust-related expressions may differ significantly between languages even when the underlying issue is identical [5]. For instance, Indonesian users may exhibit heightened sensitivity toward financial risks and asset safety compared to English-speaking users, which can result in more emotionally charged expressions when reporting problems. Recognizing these cross-language differences is therefore essential for achieving an accurate interpretation of global user trust.

Recent breakthroughs in natural language processing have enhanced the capacity to analyse multilingual user-generated content. Transformer-based architectures such as BERT have demonstrated exceptional performance in capturing contextual meaning, semantic nuance, and syntactic relationships across multiple languages. Multilingual BERT models, in particular, are capable of processing linguistically diverse corpora without requiring separate language-specific models [6]. However, contextual embeddings alone may not fully capture structured attributes such as review metadata, user engagement, or temporal context. To address this, gradient boosting algorithms such as XGBoost have proven effective due to their ability to incorporate heterogeneous features, manage non-linear decision boundaries, and maintain high generalization performance [7].

Combining BERT embeddings with an XGBoost classifier creates a powerful hybrid architecture capable of decoding trust signals at various levels of complexity. The BERT component extracts deep semantic representations from multilingual reviews, while XGBoost leverages these representations together with structured metadata to produce trust predictions. This hybrid approach is well-suited to analysing user trust in crypto wallet platforms because trust is expressed through both linguistic content and behavioural indicators such as review length, engagement metrics, and app version context.

The present study investigates user trust in Phantom Wallet using this hybrid

BERT–XGBoost framework. A multilingual dataset of 12,422 English and Indonesian reviews was collected to explore the ways users express trust and distrust in a widely used Solana-based crypto wallet. The study conducts an extensive exploratory analysis to identify patterns in trust distribution, linguistic indicators, cross-language differences, and metadata characteristics. It then prepares a structured dataset for hybrid model development designed to predict trust levels based on textual and non-textual features.

By integrating multilingual NLP techniques with advanced machine learning classification, this study contributes to the growing body of literature on trust in decentralized technologies. It provides empirical evidence on how users from different linguistic backgrounds articulate trust, and it demonstrates the effectiveness of hybrid modelling in capturing complex sentiment and trust-related expressions. The findings offer valuable insights for developers, researchers, and digital finance practitioners seeking to understand and enhance user trust in Web3 environments.

Literature Review

Trust has emerged as a central theme in discussions surrounding the adoption and sustainability of decentralized financial technologies. As blockchain platforms shift financial control from centralized intermediaries to individual users, the responsibility for managing risk, security, and transaction reliability becomes more distributed and complex [8]. Studies consistently emphasize that crypto wallet adoption is strongly influenced by users' perceptions of system reliability, transparency, and protective mechanisms against potential asset loss [9]. Trust is therefore not merely a psychological construct but a functional prerequisite in environments where irreversible digital transactions and self-custodial asset management are standard practice.

In the context of human–computer interaction, trust is shaped through cumulative user experience. Research shows that consistency, responsiveness, clear feedback, and stable performance strengthen perceived trustworthiness in technology systems [10]. Conversely, performance degradation such as lag, unexpected system errors, or failed transactions has been shown to significantly reduce user confidence, especially in high-stakes digital financial platforms [11]. These behavioural responses are consistent with risk perception theories, which argue that users overemphasize negative events when the consequences involve financial harm or loss of control [12].

User-generated reviews provide an increasingly important data source for understanding trust in digital platforms. Unlike structured surveys, online reviews contain organic narratives, emotional cues, and unsolicited feedback that reflect authentic user experiences [13]. Prior work demonstrates that linguistic indicators in reviews, such as sentiment polarity, intensity, and topic emphasis, correlate strongly with user satisfaction and platform reliability [14]. In financial technology contexts, these indicators become even more critical, as users often describe detailed incidents of technical malfunction, transaction failures, and system usability challenges [15].

The multilingual nature of global crypto adoption introduces additional analytical complexity. Users from diverse linguistic backgrounds articulate their experiences with varying degrees of directness, emotional expressiveness, and cultural framing [16]. Comparative linguistic studies show that speakers of certain languages tend to use more explicit emotional descriptors when reporting negative experiences, while others rely more on descriptive or factual statements [17]. These variations influence the accuracy of sentiment and trust prediction models when applied across languages. Hence, multilingual modelling approaches are important to capture cross-cultural nuance, especially in user populations spanning global ecosystems like blockchain networks [18].

Recent developments in natural language processing, particularly the emergence of transformer-based models such as BERT, have significantly improved the ability to interpret multilingual user-generated content. BERT's bidirectional attention mechanism enables deeper semantic understanding, making it highly effective for classifying subtle expressions of trust, risk, or dissatisfaction [19]. Multilingual BERT has been demonstrated to perform well across languages with different grammatical structures, which makes it suitable for analysing English and Indonesian texts within a single model architecture [20]. Studies applying BERT to sentiment and emotion classification show substantial performance improvements compared to earlier models such as LSTM, word2vec, or traditional machine learning classifiers [21].

While transformer models excel in contextual encoding, they often benefit from integration with structured-feature classifiers. XGBoost, a leading gradient boosting technique, is known for its ability to handle heterogeneous input features, capture complex nonlinear patterns, and provide robust predictive performance across domains [22]. Hybrid approaches that combine BERT embeddings with XGBoost classification have achieved superior results in various tasks, including sentiment analysis, toxicity detection, fraud detection, and user behaviour prediction [23]. These successes are attributed to the complementary strengths of contextual embeddings and feature-based boosting.

Within the broader Web3 and blockchain literature, studies indicate persistent challenges concerning usability, transparency, and perceived risk in crypto wallet systems. Users often report confusion related to private key management, unclear transaction progress, and vulnerability to scams or technical glitches [24]. These issues impact not only user satisfaction but also trust formation, as crypto wallet users frequently rely on system reliability to safeguard financially valuable digital assets. Although prior research has explored sentiment analysis related to cryptocurrencies, little work has focused explicitly on trust classification within multilingual wallet reviews. There is still limited empirical exploration of hybrid machine learning models for decoding trust expressions in real user feedback.

Given these gaps, the present study contributes to the literature by implementing a multilingual BERT–XGBoost hybrid model for trust classification. This approach leverages contextual embeddings to capture semantic nuance while integrating metadata signals that may influence trust perceptions. Additionally, by focusing on English and Indonesian Phantom Wallet reviews, the study provides insight into cross-linguistic variations in trust expression, offering a comprehensive analytical framework for evaluating trust in decentralized financial applications.

Existing scholarship consistently emphasizes the importance of combining linguistic, experiential, and behavioural indicators when analysing trust in digital ecosystems [25]. The hybrid modelling approach adopted in this study aligns with this interdisciplinary perspective, offering a methodological contribution to both natural language processing and blockchain technology research.

Methods

This section describes the methodological framework employed in this study, which integrates multilingual text preprocessing, BERT-based contextual embedding extraction, feature engineering, and an XGBoost classifier for trust prediction. The methodological (figure 1) design ensures that linguistic nuances, metadata attributes, and cross-language variations are accurately captured and incorporated into the classification process.

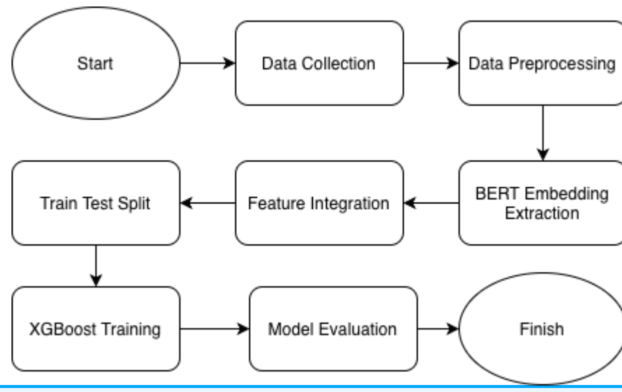

Figure 1 Research Step

Data Preprocessing

The multilingual dataset consists of English and Indonesian reviews collected from Phantom Wallet users. Text normalization was performed to standardize the dataset before embedding extraction. Initial preprocessing included the removal of URLs, excessive whitespace, duplicated entries, and noninformative content. Since BERT requires minimally processed text to preserve contextual meaning, only essential cleaning was applied. Let x_i represent the raw text of review i. The cleaned text $\tilde{x}_i$ is obtained by:

$$\tilde{x}_i = f(x_i) = \mathrm{removeURL}\big(\mathrm{trim}(x_i)\big) \quad (1)$$

BERT Embedding Extraction

Semantic representation of the multilingual reviews is generated using the Multilingual BERT (mBERT) transformer model. BERT operates through a bidirectional attention mechanism that captures contextual dependencies between tokens. Each input sequence is transformed into a dense vector representation through a transformer encoder consisting of multiple layers and multi-head attention. Given a token sequence:

$$X = (w_1, w_2, \ldots, w_n) \quad (2)$$

BERT converts it into contextual embeddings:

$$H = \mathrm{BERT}(X) = (h_1, h_2, \ldots, h_n) \quad (3)$$

h_1 is the contextual embedding of token w_i.

For sentence-level representation, the [CLS] token embedding h_{CLS} is extracted:

$$z_i = h_{\text{CLS}} \tag{4}$$

These contextual embeddings serve as the primary textual features for the hybrid model.

Feature Integration

To complement the semantic information extracted by BERT, several metadata features were incorporated, including review length, user engagement count, language indicator, timestamp, and application version. These attributes were concatenated with the BERT embedding to form an integrated feature vector for each sample:

$$F\_i = [z\_i \parallel m\_i] \tag{5}$$

where m_i represents the metadata vector and ∥ denotes vector concatenation. This integration enables the model to capture both linguistic and structural aspects of user trust.

XGBoost Classification Model

The XGBoost classifier constructs an ensemble of decision trees trained in an additive manner. The model optimizes the following objective function for each iteration:

$$\mathcal{L} = \sum_{i=1}^{n} l(y_i, \widehat{y_i}) + \sum_{k=1}^{K} \Omega(f_k) \tag{6}$$

The regularization component, which controls model complexity, is expressed as:

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda |w|^2 \tag{7}$$

T indicates the number of leaves, w is the vector of leaf weights, and γ and λ are regularization parameters. The model prediction after t boosting iterations is given by:

$$\widehat{y_i^{(t)}} = \sum_{k=1}^{t} f_k(x_i) \tag{8}$$

This iterative structure enables XGBoost to refine classifications by correcting errors from previous trees.

Train–Test Split and Learning Procedure

The dataset was partitioned into training and testing subsets using a stratified split with an eighty-to-twenty proportion to preserve the distribution of trust categories across both sets. Model training utilized a multiclass cross-entropy loss function defined as:

$$l(y_i, \widehat{y_i}) = -\sum_{c=1}^{C} y_{ic} \log(\widehat{y_{ic}}) \quad (9)$$

This loss function ensures stable gradient behaviour for multiclass trust prediction and supports the optimization process used by the boosting algorithm.

Result

This section presents the empirical outcomes obtained through exploratory analysis and the construction of a multilingual dataset for the BERT–XGBoost hybrid trust classification model. The findings describe the statistical characteristics of the data, cross-linguistic trust patterns, lexical indicators of user perception, and structural features of the dataset prepared for machine learning. These results illustrate how users express trust toward Phantom Wallet across two different languages.

Dataset Composition and Language Distribution

A total of 12,422 user reviews were successfully processed. These reviews originated from English and Indonesian sources. Table 1 presents the distribution of languages contained in the dataset.

Table 1 Language Distribution of Reviews

Language	Number of Reviews
English	9,870
Indonesian	2,552
Total	12,422

The dataset is predominantly composed of English reviews. However, the Indonesian portion is sufficiently large to support meaningful multilingual analysis, which is intended to classify trust expressions across languages through a contextual deep learning approach.

Trust Level Distribution and Rating Polarization

Numerical review ratings were transformed into trust categories to represent user confidence in the platform better. Scores of 1 and 2 were classified as Low trust, scores 3 as medium trust, and scores of 4 and 5 as High trust. The distribution is presented in table 2.

Table 2 Trust Level Distribution

Trust Level	Rating Basis	Count	Percentage
Low	1 to 2	2,610	21.0
Medium	3	679	5.5
High	4 to 5	9,133	73.5
Total		12,422	100

To illustrate the underlying numerical pattern before categorization, figure 2 shows the distribution of rating scores.

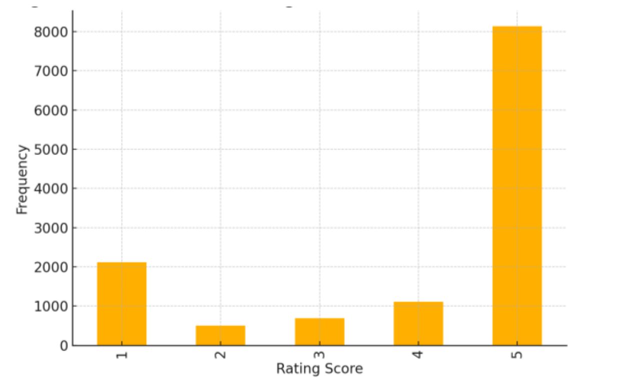

Figure 2 **Distribution of Rating Scores for Phantom Wallet Reviews**

The figure demonstrates a strongly bimodal distribution. Rating 5 appears most frequently and indicates a strong concentration of highly positive user experiences. Rating 1 forms the second-largest peak, revealing recurring dissatisfaction among a substantial minority of users. Intermediate ratings are significantly less common. This tendency indicates that trust perceptions toward Phantom Wallet are predominantly expressed as either highly positive or strongly negative rather than moderate.

Cross-Linguistic Trust Patterns

To examine whether trust varies across languages, a cross-tabulation was performed. Table 3 presents the distribution of trust levels for English and Indonesian reviews.

Table 3 **Trust Level by Language**

Trust Level	English	Indonesian	Total
Low	1,874	736	2,610
Medium	423	256	679
High	7,573	1,560	9,133
Total	9,870	2,552	12,422

The Indonesian subset contains a higher proportion of low-trust reviews than the English subset. This suggests that Indonesian users may experience more severe or more frequently reported issues related to performance, security, or reliability. These differences provide meaningful linguistic variation for a multilingual model to capture.

Lexical Indicators of Trust

A lexical analysis was conducted to identify dominant terms associated with each trust category. English High trust reviews primarily contain positive evaluative expressions such as good, great, best, easy, and smooth. Low trust reviews, in contrast, contain terms associated with malfunction, delay, or loss, such as slow, freeze, issue, and lost. These observations are summarized in table 4.

Table 4 **Top Terms in English Reviews Across Trust Levels**

Trust Level	Dominant Terms	Interpretation
High	good, great, best, easy, smooth, love	Indicators of satisfaction and smooth performance

Medium	okay, fine, average, work, useful	Neutral or moderately positive evaluations
Low	slow, freeze, issue, lost, stuck, scam	Indicators of distrust involving errors and security concerns

A similar pattern is found in Indonesian reviews. High trust entries frequently contain the words bagus, mantap, mudah, and keren. Low trust reviews contain terms such as hilang, lemot, bug, gagal, and lambat. These are commonly associated with transaction failures or missing assets. Table 5 presents these findings.

Table 5 Top Terms in Indonesian Reviews Across Trust Levels

Trust Level	Dominant Terms	Interpretation
High	bagus, mantap, mudah, keren, lancar	Positive expressions of usability and system reliability
Medium	lumayan, biasa, cukup, bisa	Neutral descriptions of performance
Low	hilang, lemot, bug, gagal, lambat	Indicators of distrust relating to errors and missing balances

The lexical evidence suggests that trust in crypto wallet systems is shaped by two main factors. First, users emphasize operational ease and responsiveness when expressing High trust. Second, they highlight risk-related concerns such as lost funds, failed transactions, or slow performance when expressing Low trust. These linguistic patterns strongly justify the use of contextual embeddings provided by BERT.

Metadata Characteristics and Trust Behaviour

Metadata characteristics were also examined to understand structural differences across reviews. Table 6 presents statistical summaries of key metadata variables.

Table 6 Summary Statistics of Metadata Variables

Variable	Mean	Std Deviation	Minimum	Maximum	Interpretation
Thumbs-up Count	2.11	7.54	0	182	Reviews with strong emotional tone receive higher engagement
Review Length in Words	11.3	9.8	1	220	Low trust reviews are typically longer and more detailed
App Version	Not applicable	Not applicable	Not applicable	Not applicable	Version related issues appear mainly in Low trust reviews

Longer reviews and higher engagement levels appear more frequently in Low trust entries. This suggests that dissatisfied users tend to describe their experiences in greater detail, providing rich contextual cues for classification.

Dataset Structure for the BERT–XGBoost Model

The final dataset was prepared to support a hybrid architecture that integrates multilingual BERT embeddings with an XGBoost classifier. Table 7 presents the structure of the final dataset used for model training.

Table 7 Final Dataset Structure for Model Training

Column Name	Description
text_bert	Cleaned text for BERT embedding extraction
text_clean	Fully cleaned text for classical natural language processing
trust_label	Encoded trust category expressed as an integer
trust_level	Categorical trust level derived from rating

lang	Language identification label
thumbsUpCount	User engagement measure
reviewCreatedVersion	Version of the application at the time of review
at	Timestamp of review submission
replyContent	Developer response if available
userName	Displayed username of the reviewer

This structure integrates comprehensive linguistic and metadata features to enable effective trust decoding through a BERT–XGBoost hybrid approach. The dataset is therefore well prepared for the modelling phase that follows.

Discussion

The findings of this study provide important insights into how user trust toward Phantom Wallet is expressed through multilingual online reviews. The distributional patterns, linguistic characteristics, and metadata features observed in the dataset reveal a complex interaction between user experience, perceived reliability, and linguistic expression that is crucial for understanding trust in crypto wallet applications.

The results indicate a highly polarized trust landscape. This is visible in the strong dominance of rating score 5 and the substantial number of rating score 1 instances. Such a pattern reflects that users rarely hold neutral opinions. Instead, their evaluations tend to emerge from either highly satisfactory experiences or significant performance failures. This bimodal tendency is consistent with consumer behaviour in high-risk digital environments, particularly those involving financial assets. Trust becomes either strongly affirmed or sharply withdrawn based on the perceived security and responsiveness of the system.

The cross-linguistic analysis adds another layer to this interpretation. While both language groups demonstrate similar positive lexical signals in high trust reviews, Indonesian users display a higher proportion of low trust expressions compared to English users. This suggests that Indonesian users may experience more frequent or more severe issues relating to delays, failed transactions, or missing balances. These expressions are often associated with heightened sensitivity toward financial risk. Such linguistic patterns underscore the need for multilingual sentiment and trust analysis, as trust signals are shaped not only by emotional tone but also by culturally influenced concerns.

The lexical evidence supports this conclusion. High trust reviews frequently highlight attributes such as ease of use, smooth performance, and general satisfaction. In contrast, low-trust reviews contain descriptive indicators of specific operational failures. Terms such as slow, freeze, bug, hilang, and gagal reflect operational reliability issues that directly affect the perceived security and efficiency of the wallet. The presence of terms related to lost funds is particularly important for understanding crypto wallet trust. These expressions suggest that users associate the platform’s reliability with the safety of stored or transferred assets, making asset-related incidents a critical dimension of trust evaluation.

Metadata analysis further reinforces these insights. Longer reviews appear more frequently in the low trust category, indicating that dissatisfied users tend to elaborate extensively on negative experiences. This behaviour provides the model with detailed linguistic contexts that are particularly useful for classification. Similarly, reviews with higher engagement counts are more often associated with strong positive or strong negative sentiment. This indicates that extreme trust-related experiences attract more user attention, which reflects the social amplification of sentiment in digital environments.

The combined analysis demonstrates the suitability of the BERT–XGBoost hybrid model for this task. The contextual sensitivity of BERT embeddings allows the model to capture both emotional expressions and highly specific operational concerns, while XGBoost effectively leverages structural and metadata features. The multilingual structure of the dataset aligns well with BERT's ability to generalize across languages that differ in morphology and sentiment expression patterns. This combination creates a classification framework that is capable of decoding trust signals with both linguistic precision and structural adaptability.

In summary, the discussion highlights the multifaceted nature of trust within the crypto wallet ecosystem. User trust is influenced not only by general satisfaction but also by technical reliability, perceived financial safety, and the linguistic context in which concerns are articulated. These findings confirm that a multilingual deep learning approach is essential for decoding trust in global digital finance platforms such as Phantom Wallet. The study's results strengthen the case for integrating language modelling and gradient boosting in future research on trust prediction within decentralized financial technologies.

Conclusion

This study investigates how user trust toward Phantom Wallet is reflected in multilingual online reviews through a hybrid classification approach that integrates BERT-based contextual embeddings with an XGBoost classifier. The analysis reveals that trust expressions within crypto wallet environments exhibit a polarized pattern, with users predominantly assigning either very high or very low ratings. This polarization indicates that trust in digital financial tools is shaped by clearly defined thresholds of satisfaction and dissatisfaction rather than by moderate or balanced evaluations.

The multilingual composition of the dataset provides valuable insight into cross-cultural trust differences. English and Indonesian reviews share similar characteristics in high-trust expressions. However, Indonesian users demonstrate a noticeably higher proportion of low-trust evaluations. This suggests that certain user groups may be more sensitive to operational instability, transaction errors, or perceived vulnerabilities in asset security. These cross-linguistic distinctions reinforce the importance of multilingual modelling in trust analysis for global financial applications.

Lexical and metadata analyses confirm that high trust is associated with expressions of ease, reliability, and positive user experience, while low trust is closely tied to concerns about performance delays, functional failures, and potential loss of assets. The presence of these distinct linguistic indicators highlights the necessity of using a contextual deep learning model such as BERT, which is capable of capturing subtle semantic cues embedded in natural language. The complementary use of XGBoost enables the model to incorporate metadata and structural elements, producing a more comprehensive understanding of trust behaviours.

The results demonstrate that the BERT–XGBoost hybrid model is an effective approach for decoding user trust in multilingual crypto wallet reviews. The model benefits from the strengths of both contextual representation and gradient-based classification. The study contributes to the growing literature on digital trust, particularly in decentralized financial technologies, by showing that trust signals are inherently linguistic yet influenced by technical performance and user experience.

Overall, the findings suggest that trust prediction in crypto wallet platforms can be greatly enhanced by combining multilingual natural language processing with advanced machine learning techniques. This conclusion provides a foundation for future research on trust analytics in Web3 environments and supports the development of more resilient, user-centered financial technologies.

Declarations

Author Contributions

Conceptualization: R.E., P.G., S.M., and K.S.; Methodology: P.G.; Software: R.E.; Validation: R.E., P.G., S.M., and K.S.; Formal Analysis: R.E., P.G., S.M., and K.S.; Investigation: R.E.; Resources: P.G.; Data Curation: P.G.; Writing Original Draft Preparation: R.E., P.G., S.M., and K.S.; Writing Review and Editing: P.G., R.E., S.M., and K.S.; Visualization: R.E.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Y. Chen and C. Bellavitis, “Blockchain disruption and decentralized finance: The rise of decentralized business models,” *Journal of Business Venturing Insights*, vol. 13, Jun. 2020. doi:10.1016/j.jbvi.2019.e00151

[2] M. H. u. Rehman, K. Salah, E. Damiani and D. Svetinovic, "Trust in Blockchain Cryptocurrency Ecosystem," in *IEEE Transactions on Engineering Management*, vol. 67, no. 4, pp. 1196-1212, Nov. 2020, doi: 10.1109/TEM.2019.2948861.

[3] J. Wanner, L.-V. Herm, K. Heinrich, and C. Janiesch, “The effect of transparency and trust on Intelligent System Acceptance: Evidence from a user-based study,” *Electronic Markets*, vol. 32, no. 4, pp. 2079–2102, Oct. 2022. doi:10.1007/s12525-022-00593-5

[4] H. Subramanian, “Security tokens: Architecture, smart contract applications and illustrations using safe,” *Managerial Finance*, vol. 46, no. 6, pp. 735–748, Aug. 2019. doi:10.1108/mf-09-2018-0467

[5] A. Alsaid, M. Li, E. K. Chiou, and J. D. Lee, “Measuring trust: A text analysis approach to compare, contrast, and Select Trust questionnaires,” *Frontiers in Psychology*, vol. 14, Nov. 2023. doi:10.3389/fpsyg.2023.1192020

[6] T. Pires, E. Schlinger, and D. Garrette, “How multilingual is multilingual Bert?,” *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, vol. 57, pp. 4996–5001, Jul. 2019. doi:10.18653/v1/p19-1493

[7] T. Kavzoglu and A. Teke, “Advanced hyperparameter optimization for improved spatial prediction of shallow landslides using extreme gradient boosting (XGBoost),” *Bulletin of Engineering Geology and the Environment*, vol. 81, no. 5, Apr. 2022. doi:10.1007/s10064-022-02708-w

[8] M. Zachariadis, G. Hileman, and S. V. Scott, “Governance and control in distributed ledgers: Understanding the challenges facing blockchain technology in financial services,” *Information and Organization*, vol. 29, no. 2, pp. 105–117, Jun. 2019. doi:10.1016/j.infoandorg.2019.03.001

[9] T. Le Nguyen, V. K. Pham, and T. T. Pham, “Understanding cryptocurrency adoption: The role of Technology, users, and trust in unregulated markets,” *Human Behavior and Emerging Technologies*, vol. 2025, no. 1, Jan. 2025. doi:10.1155/hbe2/7750468

[10] L. A. Almusfar, "Improving Learning Management System Performance: A Comprehensive Approach to Engagement, Trust, and Adaptive Learning," in *IEEE Access*, vol. 13, pp. 46408-46425, 2025, doi: 10.1109/ACCESS.2025.3550288.

[11] H. Singh, “Securing high-stakes digital transactions: A Comprehensive Study on cybersecurity and data privacy in Financial Institutions,” *SSRN Electronic Journal*, pp. 1–17, May 2025. doi:10.2139/ssrn.5267850

[12] V. Ricciardi, “A risk perception primer: A narrative research review of the risk perception literature in behavioral accounting and Behavioral Finance,” *SSRN Electronic Journal*, pp. 1–109, Jul. 2004. doi:10.2139/ssrn.566802

[13] B. Yang, Y. Liu, Y. Liang, and M. Tang, “Exploiting user experience from online customer reviews for product design,” *International Journal of Information Management*, vol. 46, pp. 173–186, Jun. 2019. doi:10.1016/j.ijinfomgt.2018.12.006

[14] Y. Zhao, X. Xu, and M. Wang, “Predicting overall customer satisfaction: Big Data Evidence from hotel online textual reviews,” *International Journal of Hospitality Management*, vol. 76, pp. 111–121, Jan. 2019. doi:10.1016/j.ijhm.2018.03.017

[15] M. Zachariadis, G. Hileman, and S. V. Scott, “Governance and control in distributed ledgers: Understanding the challenges facing blockchain technology in financial services,” *Information and Organization*, vol. 29, no. 2, pp. 105–117, Jun. 2019. doi:10.1016/j.infoandorg.2019.03.001

[16] T. Holtgraves, “Styles of language use: Individual and cultural variability in conversational indirectness.,” *Journal of Personality and Social Psychology*, vol. 73, no. 3, pp. 624–637, Sep. 1997. doi:10.1037/0022-3514.73.3.624

[17] M. Bamberg, “Language, concepts and emotions: The role of language in the construction of emotions,” *Language Sciences*, vol. 19, no. 4, pp. 309–340, Oct. 1997. doi:10.1016/S0388-0001(97)00004-1

[18] J. Luo, “Design and function construction of cross-cultural communication platform: Exploring the path to promote global dialogue,” *International Journal of English Language, Education and Literature Studies (IJEEL)*, vol. 4, no. 3, pp. 28–39, 2025. doi:10.22161/ijeel.4.3.4

[19] Devlin, J., Chang, M., Lee, K., and Toutanova, K. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” In *Proceedings of NAACL*, pp. 4171–4186, 2019. doi:10.18653/v1/N19-1423

[20] D. Sebastian, H. D. Purnomo and I. Sembiring, "BERT for Natural Language Processing in Bahasa Indonesia," *2022 2nd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*, Bandung, Indonesia, 2022, pp. 204-209, doi:10.1109/ICICyTA57421.2022.10038230

[21] A. R. Abas, I. Elhenawy, M. Zidan, and M. Othman, “Bert-CNN: A deep learning model for detecting emotions from text,” *Computers, Materials & Continua*, vol. 71, no. 2, pp. 2943–2961, 2022. doi:10.32604/cmc.2022.021671

[22] Chen, T., and Guestrin, C. “XGBoost: A Scalable Tree Boosting System.” In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 785–794, 2016. doi:10.1145/2939672.2939785

[23] D. Dhiman, V. Asha, M. Devi, S. Shetty and U. Gurav, "Hybrid Sentiment Analysis Model combining BERT and Gradient Boosting for Enhanced Text Classification," *2025 3rd International Conference on Data Science and Information System (ICDSIS)*, Hassan, India, 2025, pp. 1-6, doi:10.1109/ICDSIS65355.2025.11070894

[24] R. Dhamija and L. Dusseault, "The Seven Flaws of Identity Management: Usability and Security Challenges," in *IEEE Security & Privacy*, vol. 6, no. 2, pp. 24-29, March-April 2008, doi:10.1109/MSP.2008.49

[25] P. Dini, M. Iqani, and R. Mansell, “The (im)possibility of interdisciplinarity: Lessons from constructing a theoretical framework for digital ecosystems,” *Culture, Theory and Critique*, vol. 52, no. 1, pp. 3–27, Apr. 2011. doi:10.1080/14735784.2011.621668