Journal of Artificial Intelligence and Data Science
Vol. 1, No. 1, 2025, pp. 19-30
DOI:

Classifier-Chain SVM for TAM-Driven Multilabel Sentiment Analysis in Large-Scale User Acceptance Evaluation

Nafla Aurellia, Wiyli Yustanti
Program Study of Information System, Faculty of Engineering, Universitas Negeri Surabaya, Indonesia
e-mail address: naflaaurellia.22038@mhs.unesa.ac.id (corresponding author)

Hanim Maria Astuti
School of Information, College of Communication and Information, Florida State University, USA
Information Systems Department, Faculty of Intelligent Electrical and Informatics Technology,
Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia.

Received:30 September 2025 | Revised: 10 Oktober 2025 | Accepted: 20 Oktober 2025

ABSTRACT

In this work we suggest an alternative and novel integrated method which combines TAM with multilabel sentiment analysis by a SVM classifier. In contrast to literature on traditional TAM research that often comprises survey-based measures, this study utilises user-generated comments written in large numbers as a close substitute factor for perceived usefulness, perceived ease of use, attitude and behavioural intention. A set of 22,366 reviews were harvested from the Google Play Store for the video-editing application CapCut and handled within the Knowledge Discovery in Databases (KDD) framework that comprised data selection, preprocessing, labeling, transformation, mining and evaluation. It was observed that the multilabel SVM model trains 80% of Corpus19 and tested on remaining 20% achieves F1-score: 0.95 (Train) and 0.91 (Test). Stronger associations between both Perceived Usefulness and Perceived Ease of Use, and Attitude Toward Using (r = 0.912, 0.816 respectively) were again shown by Pearson correlation analysis. Meanwhile, Attitude Toward Using was positively related to Behavioral Intention of Use (r = 0.510). These results suggest that the TAM dimensions of comments can be efficiently modeled as multilabel sentiments, and usefulness and ease of use continue to play a significant role in shaping user acceptance towards CapCut.

Keywords-Sentiment Analysis; Multilable Classification, Support Vector Machine (SVM), Technology Acceptance Model (TAM); Video Editing Application

I. INTRODUCTION

The fast-paced growth of information technologies has rendered increasingly necessary digital literacy, especially concerning the capability to produce and understand multimedia-related contents. Video as one of the dominant communication channels In digital age, video is ubiquitously adopted for storytelling, advertisement, education and information propagation [1]. Its increasing dependence on video content is one of the reasons for th popularization of mobile video editing applications that can be found in online stores like Google Play Store. And other editing tools, such as Motion Ninja, VivaVideo, KineMaster, PowerDirector, FilmoraGo and CapCut etc., provide users with rich functions to satisfy for different quality creation requirement. The more video editing apps rise to stardom, the more reviews are tented to appear on deferent platforms. These reviews are a goldmine of rich and voluntary feedback from users, expressing their thoughts about usefulness, ease of use, quality of functionality etc. along with overall impression. Nevertheless, given their size and unstructured nature it is unrealistic to manually analyse reviews. Sentiment analysis, an important task in the field of Natural Language Processing (NLP), provides us with automatic mechanisms for extracting users' opinions and grouping them according to polarity or even topic aspects [2].

The Technology Acceptance Model (TAM) is one of the most frequently used theoretical frameworks to explain psychological antecedents of technology adoption. History of TAM is based on predefined survey tools to measure the constructs as Perceived Usefulness (PU), Perceived Ease of Use (PEU), Attitude Toward Using (AT), and Behavioral Intention to Use (BIU), Actual Use of System (AUS) [3-5]. While survey data allow controlled measures, they are not entirely equivalent to the natural and spontaneous perceptions of users in real life. With the advent of user-generated content (UGC), there are opportunities to operationalize TAM from unstructured text: data. The user review contents usually contain multiple dimensions such as a comment which is used for various purposes (ease of use or usefulness etc.), hence it is an ideal candidate for multi-label classification, in which each review can be connected to one or more TAM constructs [6, 7]. For this, SVM multilabel classification is a useful method to assign multiple class labels to a single document [8]. TAM has previously been used in a wide range of studies (e.g., e-learning satisfaction [9], interactive website evaluation [4]), however survey data is predominantly used in these studies. On the other hand, work in multilabel classification has focused in different textual domains (e.g., dangerous speech detection [10], religious text classification [11]) not including TAM. Furthermore, the sentiment analysis studies of mobile apps mainly use lexicon-based approaches or generic polarization categorization and do not connect user sentiments to theoretical acceptance aspects [12]. This gap underscores the necessity for a methodological link between TAM and actual user-generated data. In contrast to traditional TAM research based on structured questionnaires, we aim at mapping user comments directly to TAM dimensions using multilabel sentiment analysis. Through the use of multilabel SVM model, this method allows for TAM constructs to be operationalized from natural language user reviews offering a more recital and organic representation on users perceptions, and behavioral intentions. Additionally, this study provides evidence that TAM is still valid on unstructured data and can be extended to current digital media.

II. METHOD

A review of previous research is necessary to build the theoretical and methodological bases for this study. Related Work Prior work related to this study covers three main categories: TAM-based studies using structured survey data, advanced technology acceptance research which includes extended models and multilabel text classification for unstructured textual data. The next sub-sections review related work in these fields, which can be used to put the contribution of this paper in place in a larger academic context.

A. TAM-Based Studies using Structured Survey Data

Technology Acceptance Model (TAM) of Davis (1986) [3] is a cornerstone theoretical model for explaining user acceptance and use of information systems. The model posits that Perceived Usefulness (PU) and Perceived Ease of Use (PEU) are the key cognitive determinants affecting an individual's attitude to using a technology, this in turn will determine Behavioral Intention and hence actual system usage. 1.1 Technology Acceptance Model The TAM is based on the Theory of Reasoned Action (TRA), which postulates that technology adoption is a reasoned process leading to predictable user behavior including the evaluation of expected performance benefits and effort associated with system use. Conceptually, the TAM model is shown in Figure 1.

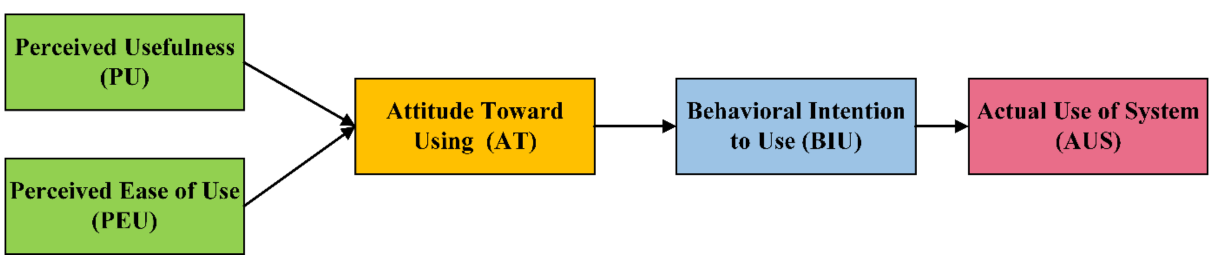

Figure. 1. The Technology Acceptance Model (TAM) Model

A substantial amount of research has adopted the TAM via formal survey approaches. A hybrid TAM– ELQ model is constructed for assessing e- learning user satisfaction from 145 questionnaire samples [9]. The study employed SmartPLS, G*Power, EFA and PLS-SEM to investigate 12 factors and it tested 14 hypotheses of which 11 were accepted. Service Quality, Appearance Quality and Information Quality demonstrated significantly effective for user satisfaction, well-validated the structure model ($R^2 = 0.858$). This study contributes a holistic perspective of user acceptance, albeit solely relying on structured survey data disregarding the valuable insights from users' natural feedback (e.g., open-ended responses). Likewise, an IoT cloud platform acceptance is assessed through a UTAUT–TAM questionnaire answered by 129 university students. SEM was conducted through AMOS and SPSS, only two out of eight generic technology acceptance models constructs failed to have an impact on behavioral intention: Effort expectancy and social influence emerged as the two non-significant constructs, although the latter became significant subsequent to model re-specification [6]. This was a methodologically strong study with good goodness-of-fit indices (CFI = 0.998, RMSEA < 0.08), nevertheless it is still restricted by the limitations of both Likert-type questionnaires and small homogeneous samples. A further research focusing on acceptance of technology through interactive sites based on 456 valid responses. Extended TAM framework indicated that SIMPLICITY, readability and the aesthetic appeal that reached statistical significance with respect to PU and PEU The other two factors: translational accessibility and individual experience showed no significant impact. [4] Even though the research employs both quantitative SEM and qualitative interviews, it further propagates the traditional TAM survey-based paradigm. These structured-forum studies corroborated TAM's efficacy, but are themselves limited with the reliance on only questionnaire-based instruments for data collection rather than user-generated content.

B. Text Classification and Sentiment Analysis

Text classification and sentiment analysis A number of advance computational methods have been proposed recently for the text classification and sentimental analysis that are relevant to this study. Alzanin et al. [11], an optimized Arabic multilabel text classification framework was also suggested that combines Genetic Algorithms for feature and hyperparameter optimization and ensemble learning to improve predictive performance. This work shows that optimization methods are crucial in dealing with the texts which have multiple semantics labels. Similarly, Mutinda et al. [12] proposed LeBERT, based on BERT contextual embeddings and CNN layer for the extraction of local features over tweets, to benefit both semantics and syntax from pre-trained lexicon-oriented knowledge in terms of sentiment analy-sis task, outperforming typical deep learning models. Some lexicon-based sentiment methods in the context of application reviews have also been considered. Asri et al. [13] demonstrated that the lexicon based approach works well on providing sentiment analysis of unstructured text, from Google Play reviews in particular. The lexicon-based approach is efficient if the review text is short, because direct, informal, with no deep knowledge of context. In line with this, Abdullah and Abdulazeez [14] extensively reviewed applications of SVM in machine learning tasks and stressed on its adaptability and robustness especially for text classification problems and forecasting. Luo [15] also examined several classical machine learning methods, including Naïve Bayes, SVM, and Random Forests, and highlighted that both feature engineering strategies (e.g., TF-IDF, Bag-of-Words) and dataset characteristics significantly influence classification performance.

In addition to classification algorithms, a few work about computational environment and infrastructure for supporting ML workflows could also be found. Carneiro et al. [16] studied the GPU and TPU capabilities in Google Colaboratory and found that they are very efficient for deep learning experiment. Wang et al. [17] studied the failure characteristics of ML workloads on public Jupyter notebooks and found that out-of-memory and dependency conflict were common types. The results have implications on the requirement of strong computational settings in large-scale text mining studies. Dimitri and Schraivogel (2020) have also proposed state-of-the art deep learning, and hybrid approaches for complex multimodal as well as multilingual tasks. Mosleh et al. [18] designed a deep learning and fuzzy matching-based combined model to translate Arabic Sign Language in real-time, and found that the combination of heterogeneous learning methods was able to improve accuracy and inclusivity in communication technology. Akkem et al. [19] introduced ensemble learning methods to a Streamlit based recommendation system, paving the way for real-world applications of high performing ML models in interactive systems.

Emotion analysis is popular topic in the area of business and social analytics. Ahmed et al. [20], merged sentiment analysis and machine learning for predicting customer sentiments on social media platforms, and SVM, Naïve Bayes, and Random Forest were found to be good predictive techniques. Umarani et al. [21] additionally contrasted traditional ML models with deep learning models and found that RNN, LSTM and CNN variants achieved the best sentiment classification across domains. Samyuktha et al. [22] also evaluated machine learning (ML) based sentiment analysis pipelines, and re-validated SVM as well-suited for the role of sentiment polarity classifier. In more specific conditions, Costola et al. [23] demonstrated that sentiment machine learning techniques by leveraging the sentiments from appraisals of COVID-19 news can forecast stock market response, whereas Obagbuwa et al. [24] used supervised ML algorithms to recognize depressive sentiment, which reflects the general applicability of sentiment analysis in behavioral and social apps. Finally, Elgeldawi et al. [25] investigated the importance of optimizing parameters in a machine learning algorithm for Arabic sentiment analysis where they showed that use of optimized parameter settings can lead to significant boost in model performance for complex linguistic scenarios. Together, these papers are testimony to exciting advances in multilabel text classification, deep learning sentiment analysis and machine learning optimization. However, despite advances in processing unstructured textual data, none of the existing works integrate multilabel sentiment classification with the TAM or attempt to operationalize TAM constructs using real-world user-generated reviews. This underscores a critical research gap that the present study addresses by combining multilabel SVM classification with TAM-based acceptance evaluation.

III. PROPOSED METHOD

The general outline of the methodology applied in this work is shown in Figure 2, which integrates three complementary components: the Knowledge Discovery in Databases (KDD) process, confirmatory analysis of TAM constructs, and a Prototype-based system development approach. The work starts with a literature survey and then explains the KDD stages for developing multilabel SVM model to predict TAM dimensions using user generated reviews. The generated labeled data will then be used to analyze the theoretical TAM via hypothesis testing, correlation analysis, and decision validation. Lastly, the paper continues with implementation a prediction application by using the Prototype framework. Such combined approach allows both an analytical robustness for model construction and a practical feasibility for system development.

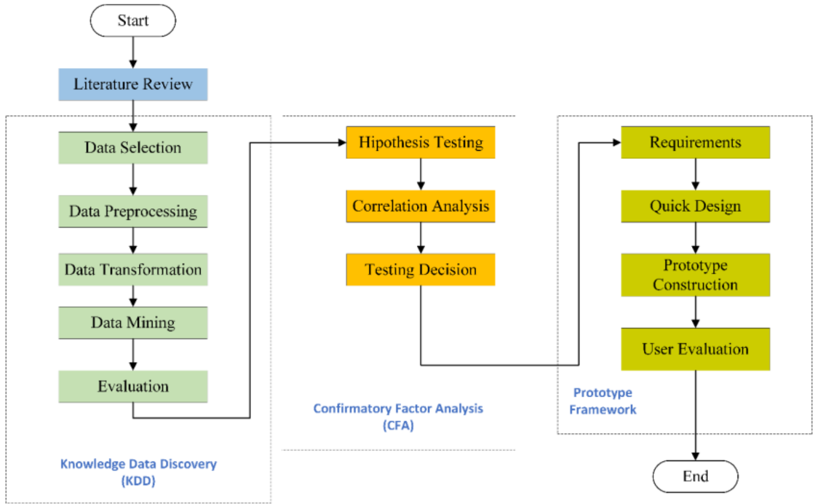

Figure. 2. Research Methodology Framework

A. Knowledge Data Discovery (KDD)

The Knowledge Discovery in Databases (KDD) framework is used in this work to build a processing pipeline that converts raw user-generated reviews into structured analytical outcomes for multilabel classification. KDD processes forward a structured procedure through which data are extraction, cleaning, transformation and analysis properly with discipline from an algorithmic standpoint to develop the most accurate and dependable predictive model. The KDD process aims to provide end-to-end preparation of textual data for the purpose of mapping user comments into TAM constructs through a Multilabel classifier based on SVM, by means of its core steps that include data selection and preprocessing, transformation, data mining and evaluation. The stages are described in details in the following sub-sections.

1) Data Selection

In this paper, the data under consideration are user comments in CapCut application from Google Play Store. The data were obtained by web scraping, performed from August-October 2025, including the immediate release after the most recent application update date on 20 August 2025. This timing allows the reviews to capture user experiences with the latest available version of the app. In total 22,367 Indonesian-language reviews were crawled including the text of the review and its rating. These records serve as a starting corpus for further pre-processing, transformation and multi-label annotation in the KDD pipeline.

2) Data Preprocessing

The final goal of the preprocessing step is to convert (or refine) this raw text, acquired through scraping explicitly, into a form that is clean and structured for processing and modeling. A sequence of operations is used to perform the process. It starts with data cleaning by removing noise including HTML tags, emojis, URLs, punctuation, and duplicate entries. The cleaned text is normalized in the sense that informal language, slang, and diverse spellings are standardized. The text is then tokenized to extract a set of tokens for the sentence, based on which each sentence is further divided into word level. Stopword removal is performed after tokenization t eliminate excessively common, semantically irrelevant words that do not provide any useful information to the classification task. The stemming, the last step keep each token in its root form and then reduces vocabulary diversity and makes the features representation more consistent. Open image in new window After all of these preprocessing steps, the processed data are suitable for the next sentiment labeling and TAM labeling to construct a multilabel classifier.

3) Data Transformation

The transformation phase concentrates on the conversion of the preprocessed textual data into a structured form that can be used and extended directly for classification or other analytical sub-processes. In this paper, the target transformation includes two key sub-tasks: sentiment labeling and TAM labeling with lexicon-based methods. The first step, sentiment labeling, is to label the polarity category of each user review. We use a lexicon-based approach with Inset Lexicon, an Indonesian sentiment lexicon containing predefined list of positive and negative words [26]. A polarity score is two generated for each review, which consists of the number of positive and negative terms found in the lexicon. Reviews are classified as positive, negative and neutral sentiment categories according to the combined polarity score. This method achieves an automatic sentiment discovery without the use of labeled data and offers a confident sentiment baseline for TAM projection. The sentiment classification achieved is presented in Table 1.

TABLE 1. A SAMPLE OF SENTIMENT CLASSIFICATION USING INSET LEXICON

Post Preprocessing Review	Positive Score	Negative Score	Label
['oke', 'sangat', 'bagus', 'tanya']	4	2	positive
['bagus', 'sebab', 'edit', 'puas', 'lalu', 'efek', 'edit', 'sangan', 'familiar']	6	1	positive
['aplikasi', 'bagus']	1	0	positive
['update']	0	0	neutral
['apk', 'kikir', 'dikitdikit', 'premium', 'dikitdikit', 'bayar', 'gue', 'masa', 'pengin', 'nambahin', 'audio', 'biar', 'jelas', 'premium', 'jiers']	2	4	negative

Secondly, TAM labeling is applied to map sentiment-labeled reviews to the predefined constructs of the TAM. This method is based on lexicon with Fuzzy String matching to recognize variation of user expressions [27, 28]. Keywords of all reviews are contrasted against a keyword dictionary specifically built for the five TAM dimensions: PU, PEU, AT, BIU and AUS. If the similarity score between a review and a TAM dimension's keywords is higher than the predetermined threshold, then that dimension will be given value of 1 (detected), otherwise, it will be assigned with value of 0 (not detected). This binarized multilabel coding enables a review to be labeled with multiple TAM factors at the same time. The TAM dimension labeling results are shown in the example of Table 2.

TABLE 2. A SAMPLE OF TAM DIMENSION MULTILABELING USING FUZZY STRING MATCHING

Post Preprocessing Review	Sentiment Label	PEU	PU	AT	BIU	AUS
['oke', 'sangat', 'bagus', 'tanya']	positive	1	1	1	1	0
['bagus', 'sebab', 'edit', 'puas', 'lalu', 'efek', 'edit', 'sangan', 'familiar']	positive	1	1	1	0	0
['aplikasi', 'bagus']	positive	0	1	1	0	0
['update']	neutral	0	1	0	1	0
['apk', 'kikir', 'dikitdikit', 'premium', 'dikitdikit', 'bayar', 'gue', 'masa', 'pengin', 'nambahin', 'audio', 'biar', 'jelas', 'premium', 'jiers']	negative	0	1	0	0	0

Finally, the results of sentiment and TAM are combined to obtain a single perception score for all dimensions of TAM. Sentiment labels are represented as integers (1 for negative, 2 for neutral and 3 for positive). If a review contains a TAM aspect, the sentiment rating weight for the corresponding aspect is given. If no TAM label is encountered, a default value of the dimension is given. This transformation results in a structured dataset that reflects both the presence of TAM constructs as well as sentiment orientation towards each construct, derived on which we build and test further models. A part of the user perception values obtained by the combination Sentiment and TAM labels are listed in Table 3.

TABLE 3. A SAMPLE OF THE INTEGRATED TAM AND SENTIMENT MULTILABEL OUTPUT

Post Preprocessing Review	Sentiment Label	PEU	PU	AT	BIU	AUS
['oke', 'sangat', 'bagus', 'tanya']	positive	3	3	3	3	2
['bagus', 'sebab', 'edit', 'puas', 'lalu', 'efek', 'edit', 'sangan', 'familiar']	positive	3	3	3	2	2
['aplikasi', 'bagus']	positive	2	3	3	2	2
['update']	neutral	2	2	2	2	2
['apk', 'kikir', 'dikitdikit', 'premium', 'dikitdikit', 'bayar', 'gue', 'masa', 'pengin', 'nambahin', 'audio', 'biar', 'jelas', 'premium', 'jiers']	negative	2	1	2	2	2

4) Data Mining

In the data mining phase a predictive model is developed, which categorizes user reviews according to the TAM dimensions. In this work, we employ the multilabel classification algorithms of SVM algorithm since it is effective and robust to handle high-dimensional textual data. As the TF-IDF representation results in sparse and high-dimensional feature vectors, kernel function is used to enable efficient computation along with excellent classification performance. Support Vector Machine (SVM) is a supervised learning algorithm for regression and classification. It is based on Structural Risk Minimization (SRM) principle to minimize the generalization upper bound, not just reduce agility, which stands in contrast with Empirical Risk Minimization that focuses on error during the training. For SVM, the learning task is converted to a convex optimization problem and solved using quadratic programming in its dual space. The classifier searches for the best hyperplane, the dividing line that maximizes the margin, or distance, between it and data points from all classes. These closest points (called support vectors) are essential for the final decision function [29]. Despite that, linear and non-linear SVM can be applied for complex classification tasks such as high dimensional text data like TF-IDF presentations. The best separator field search formula or destination function as (1).

$$\min_{\boldsymbol{w},\xi}\left(\frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{n}\xi_i\right) \qquad (1)$$

Where w is the weighting vector and C is a parameter that determines the amount of cost due to misclassification of training data during the learning process. When the value of C is large, the margin will be smaller, indicating that fault tolerance will be lower when an error occurs. Conversely, when the value of C is small, the degree of error tolerance will be greater. Thus, the separating hyperplane is expressed by (2), where b is the bias and the slack variable (ξ_i) for the-i data point aims to soften constraints by providing a tolerance for data that may not be fully classified. This tolerance is called a soft margin hyperplane.

$$y_i\left(\boldsymbol{x}_i^T\boldsymbol{w} + b\right) \geq 1 - \xi_i \qquad (2)$$

Under the theorem advantages of SVM and sequential dependencies representation mechanism from classifier chain, a novel method is proposed to extract TAM-related constructs from user reviews. Finally, at each review we got a multilabel output which reflects the detected TAM dimensions and that allows for structured interpretation of technology acceptance indicators reported by the users. In order to consider the potential dependence of TAM dimensions, MLC process is applied using Classifier Chain technique [30], whereby multiple binary SVMs are linked in a chain-like form and in which predicted labels from previous procedures impact on next ones. By so doing, the model can learn the interplay between TAM variables, and it can predict more accurately than independent binary models [31].

5) Evaluation

The performance and robustness of the multilabel SVM classification model proposed in this paper are evaluated. Upon data preprocessing, transformation, and modeling phases finishes the dataset is splitted as 80:20 for a balanced assessment of model generalization. The evaluation method measures how well the classifier predict several TAM dimensions at once from user comments. Since the problem is multilabel, various evaluation metrics are used to evaluate the performance of model in a comprehensive manner. These are accuracy, precision, recall and F1-score, all calculated as macro-averaged to control for class imbalance due to TAM labels. The parameters of the final model are also cross-validated with hyperparameter optimization using GridSearchCV to confirm that the classifier used in deployment is tuned on the most optimal settings. The results of this assessment are indicative that the classifier is capable of encoding linguistic patterns related with the TAM constructs, and to maintain its predicate over multiple arguments with a suitable level of accuracy. This phase finally confirms the utility of the proposed multilabel SVM model in analyzing user perceptions and technology acceptance based on textual reviews.

B. Confirmatory Factor Analysis (CFA)

This step seeks to confirm the construct validity of the Technology Acceptance Model (TAM) constructs based on the TAM–Sentiment labeled dataset from user-generated reviews. The endorsement of multi-label TAM dimensions and their corresponding perception based scores on each review are then combined: PU, PEU, AT, BIU, AUS into an empirical evidence to justify the validity of the structure that represents TAM. Such a study applies Confirmatory Factor Analysis (CFA) to test the degree to which latent constructs inferred from textual responses adhere to the theory-established model presented in Davis (1986). Factor loadings, construct correlations and structural paths are evaluated in this process to assess how much user-generated reviews reflect the hypothesized causal relationships in TAM. The strength and significance of relationships between the constructs are evaluated through correlation analysis, and it is guaranteed that the model represents a valid underlying structure. Based on the TAM model illustrated in Figure 1, the following hypotheses are evaluated:

1. H1: PU has a positive and significant effect on AT.
2. H2: PEU has a positive and significant effect on AT.
3. H3: AT has a positive and significant effect on BIU.
4. H4: BIU has a positive and significant effect on AUS.

A hypothesis is confirmed when the sign of the correlation or structural path coefficient is positive and statistically significant, as well as congruent with TAM hypotheses. By combining sentiment-based TAM labeling with hypothesis testing in CFA, this paper also extends the validation of TAM from traditional survey based data to unstructured textual reviews, thus providing evidence for the relation between users praising rhetoric and technology acceptance pattern.

C. Model Deployment based Prototype Framework

Prototype-based software development paradigm is employed for the implementation of the multilabel TAM prediction model. It certainly allows for an iterative development of the system and ensures that the implemented application meets user requirements and functional expectations. It starts with specifying the system requirements, covering core functions: text input processing, automatic preprocessing and classifying by trained Classifier-Chain SVM model are realized as well as visualization of TAM prediction outputs. Once the requirement is specified, a quick design phase is given to sketch an initial view of how the system interface and flow will look like. In this design focuses on essential components, including data submission modules, model execution pipelines, and output panels, allowing stakeholders to obtain an early conceptual view of the application. Based on this design, a functional prototype is constructed by integrating the trained multilabel SVM model with the backend processing modules responsible for text cleaning, vectorization, and label inference.

IV. RESULT AND DISCUSSION

In this section, we will show the results and analysis of Multilable TAM Classification using Classifier-Chain SVM. Also show the Graphical User Interface (GUI) of the website that use for predict the TAM aspects.

A. Results of Multilabel TAM Classification

The CC-SVM-based multilabel classification model shows very strong predictive performance for all TAM dimensions. Precision, recall and F1-score during the training and testing phases are shown in Table 4. Overall, the model achieves consistently high scores in identifying PU, PEU, AT, BIU and AUS. For the training set, the classifier is perfect with F1-scores from 0.88 to 0.98, that suggests the model is learning effectively from underlying linguistic rules for each TAM construct. The best training performance is obtained on the AU S dimension (F1 =, 0.97–0.98), followed by PU and PEU with F1-scores slightly above 0.95, indicating that model can well distinguish user statements about usefulness and ease of use. While the training F1-score of BIU is slightly decreased (0.88), it is also within an acceptable range in textual multilabel categorization.

TABLE 4. THE TRAINING-TESTING PERFORMANCE OF CLASSIFIER-CHAIN SVM FOR MULTILABEL CLASSIFICATION

Dimension	Dataset	Precision	Recall	F1-Score
PU	Training	0.96	0.99	0.97
	Testing	0.91	0.97	0.94
PEU	Training	0.97	0.98	0.98
	Testing	0.96	0.95	0.95
AT	Training	0.96	0.98	0.97
	Testing	0.92	0.96	0.94
BIU	Training	0.85	0.92	0.88
	Testing	0.78	0.82	0.80
AUS	Training	0.97	0.98	0.97
	Testing	0.94	0.93	0.93

To clarify the performance comparison of the modeling results for each TAM dimension in both the training and testing datasets, Figure 3 presents the corresponding visualization.

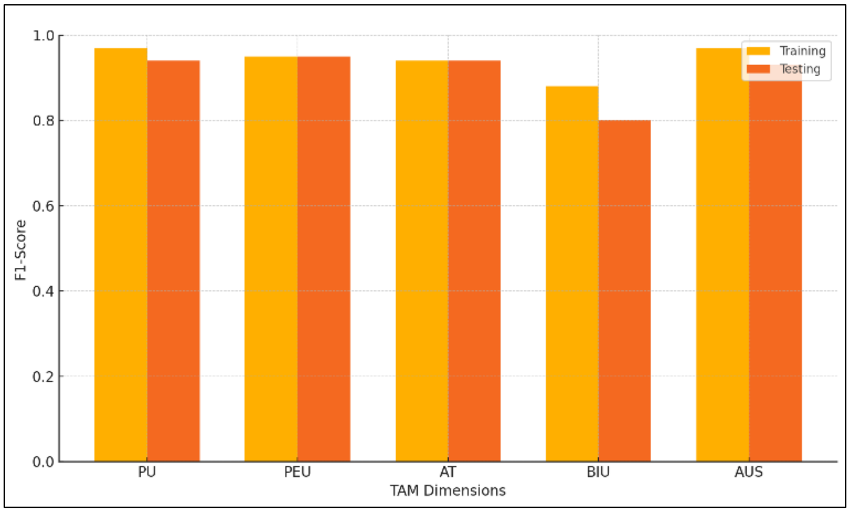

Figure. 3. The Performance of Classifier Chain-SVM in Training and Testing Datasets

The generalization performance of the model is verified by testing and the F1-scores vary from 0.80 to 0.94. For AUS dimension, once again getting the highest testing F1-score (0.93), which indicates that our model properly grasps behavioural expressions connected with actual system's utilization. PU, PEU, and AT also perform very well (F1-scores between 0.90 and 0.94), reflecting stable predictive tendencies over these constructs as well. BIU features the least testing F1-score (0.80), which is not surprising because users communicate intention with a more varied language than other TAM dimensions.

TABLE 5. THE TRAINING-TESTING FOR OVERALL PERFORMANCE OF CLASSIFIER-CHAIN SVM

Dataset	Precision	Recall	F1-Score
Training	0.942	0.97	0.954
Testing	0.902	0.926	0.912

The overall prediction results from Table 5 demonstrate that the proposed multilabel classifier can perform well and effective. On the train set, the classifier achieves precision of 0.942, recall of 0.970 and F1-score of 0.954 which indicates that it is very good in its ability to classify the relevant TAM labels rightly with low rate of misclassification. The high recall in particular suggests that the model has the potential to identify a large portion of true presence samples based on TAM dimension. The model obtains a precision of 0.902, recall of 0.926 and F1-score of 0.912 on the test set. Slightly less performance compared with training is the natural consequence of generalization effects, and the performance is stable. For classification accuracies the low difference between training and testing reflects that model is not over fitted and it generalizes well on unseen data. In summary, the conclusions regarding to the robustness and effectiveness of design for multilabel TAM classification are affirmed that indicators can accurately identify user perceptions, acceptance with high accuracy over large-scale user-generated reviews.

B. Evaluation of TAM Structural Hypotheses

The structural relationships among the TAM constructs were evaluated using Pearson correlation analysis based on aggregated TAM–Sentiment labels. Table 6 summarizes the correlation coefficients, p-values, and hypothesis testing results for the four proposed relationships within the TAM framework.

TABLE 6. THE HYPOTHESES TESTING FOR TAM STRUCTURAL MODEL

Code	Hypotheses	Pearson Correlation	p-value	Decison
H1	PU has a positive and significant effect on AT	0.912	0.00	H_0 Rejected
H2	PEU has a positive and significant effect on AT	0.816	0.00	H_0 Rejected
H3	AT has a positive and significant effect on BIU	0.510	0.0.0	H_0 Rejected
H4	BIU has a positive and significant effect on AUS	0.468	0.00	H_0 Rejected

The first hypothesis (H1), PU does affect AT, the relationship is supported with strong positive 0.912 correlation (p = 0.000), it means that the relation between them is very significance. This finding implies that users consider CapCut to be relevant and inclined towards having a positive attitude to use the app. Likewise, the second hypothesis (H2) that PEU influence AT is supported. A strong and significant association is indicated with a correlation coefficient of r = 0.816 (p<0.000). This suggests that users who perceive the low intensive to use the application tend to have positive attitude toward its use. The third hypothesis (H3) states that AT has a positive effect on the BIU. The correlation coefficient between these two variables is 0.510, a little lower but still significant (value-p = 0.000). It indicates that optimistic attitude does play the role of facilitating intention to use the app, even though it is not as great. According to H4 (H4: Behavioral intention to use predicts actual use of system), BIU is supposed to predict AUS positively.

The correlation coefficient between these vectors is 0.468 (p=0.000), indicating that this relationship is statistically significant. The outcome suggests that product users who have the intention to continue or escalate use of CapCut certainly will improve those intentions into real usage. All hypotheses are positively numbered and have significant relationships causing non-rejection of the null hypothesis (H0). This indicates that the proposed TAM model continues to enjoy validity and utility despite being developed from unstructured UGCs as opposes to classic survey-based measures. The findings corroborate that perceived usefulness and perceived ease of use are strong determinants of user attitude. Whereas, attitude and behavioural intention still have significant predictive effects on usage behaviour.

C. Deployment

As a part of deployment, we developed a GUI to allow users to predict TAM aspects real time from unstructured user-generated reviews. The manual input prediction capability featured in the prototype is described in Figure 4. By enter text review to input field, the system then handles automatic pre-processing and apply trained CC-SVM model for multilabel prediction. The prediction result is printed in a formatted, JSON-like fashion by showing the identified TAM dimensions (PU, PEU, AT, BIU, AUS). Each dimension is mapped to a binary value (1 = detected, 0 = not detected) so that users are able to Figure 4. Example review showing three positive and two negative acceptance constructs identify which acceptance constructs the review contains as soon as it is displayed. In the demonstrated example, the system well detects PU, PEU, AT, and BIU, here implying that we have strong positive perceptions about usefulness of application ease of use continuation with use respectively from review. Other than single-review manual analysis, the system also allows batch traffic processing by uploading a CSV file. This mode makes it possible for you to examine many reviews simultaneously and can be helpful to organizations or academics with large datasets. Once a CSV file is submitted, the system feeds every record through an identical preprocessing pipeline (e.g., text cleaning, tokenization, stopword removal, stemming and TF–IDF transformation) before multilabel classification with the running SVM model is carried out. We provide the results in a downloadble output file with predicted TAM labels of each review for large-scale analysis.

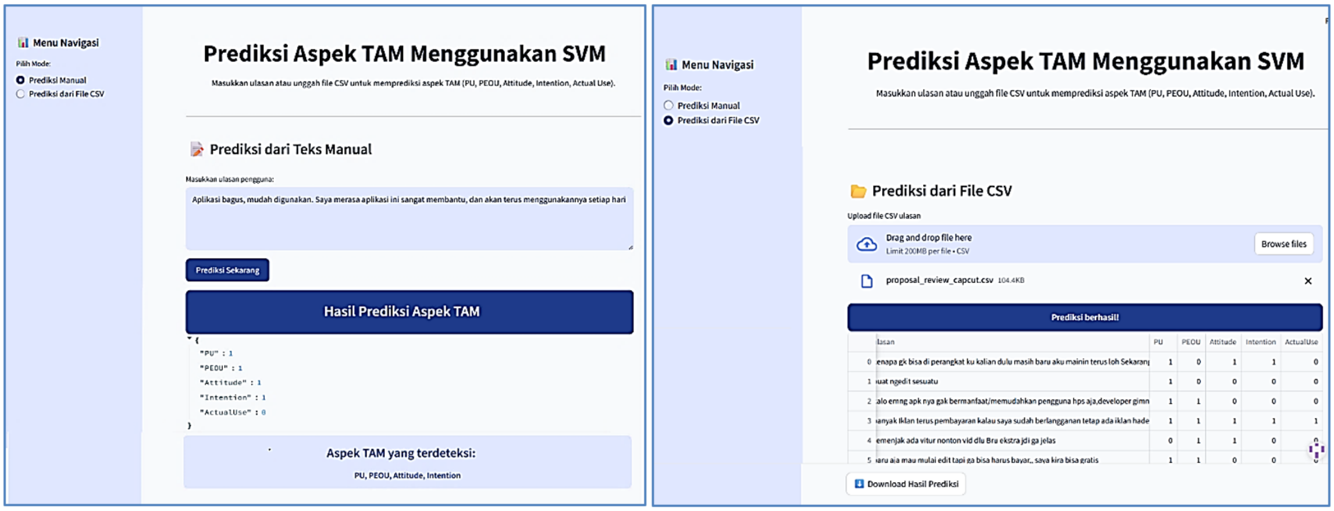

Figure. 4. The Web Application for TAM-Sentiment Multilabel Classification

D. Discussion

The findings of this study is validate that multilabel learning can be used to successfully identify the TAM constructs using unstructured user reviews. The Classifier-Chain SVM model demonstrated strong generalization capability, suggesting that a linear decision boundary in TF–IDF space is sufficient to capture the linguistic markers associated with PU, PEU, AT, BIU, and AUS. Notably, PU and PEU were consistently detected with higher reliability, which aligns with prior TAM literature indicating that usefulness and ease-of-use cues are more explicitly expressed in user-generated text. In contrast, BIU and AUS exhibited lower detection performance, reflecting the implicit and context-dependent nature of behavioral expressions, a challenge also reported in multilabel text classification studies and intention-mining research. The validation of TAM structural relationships through correlation analysis further reinforces the theoretical consistency of TAM when derived from natural language data. The strong PU→AT and PEU→AT correlations are consistent with results from conventional survey-based TAM research (e.g. utilised in the interactive platform and learning system). Simultaneously, moderate relationships of AT→BIU and BIU→AUS are in line with acceptance literature, where behavioral constructs tend to have less pronounced but still significant predictive strength. Overall, these results show that TAM can be adequately re-constructed from opinion-annotated text and may help free us more from our reliance on questionnaires to evaluate acceptance in natural settings.

Pragmatically, it offers an efficient way to automatically derive TAM constructs in support of a scale-wide (comprehensive) continuous user acceptance monitoring. The provided prototype facilitates on-the-fly TAM dimension extraction, which can be used to monitor changes of user perceived usability by developers or service providers without the need for sending out surveys. In contrast to previous work's lexicon-only sentiment systems and deep learning classifiers, the approach is highly efficient in terms of accuracy as well as interpretability and runs in near-linear time complexity compared with the size of the training data. But there are several caveats. The first attempt TAM labeling is lexicon-based and uses fuzzy matching, that may not capture deeper semantic patterns, in particular in BIU and even more so for AUS expressions. The TF–IDF representation is also less capable of understanding context, compared to the transformer-based embeddings. Moreover, correlation-based validation cannot establish causal relationships between TAM variables. A prototype system was further deployed to support real-time TAM prediction, highlighting the practical value of the proposed framework. Future work will explore contextual language models, broader lexical resources, and advanced multilabel architectures to enhance semantic representation and classification accuracy further.

V. CONCLUSION

This study presented a multilabel sentiment-based approach for evaluating technology acceptance using user-generated reviews, integrating the TAM with a Classifier-Chain SVM classifier. The proposed model achieved strong predictive performance across all TAM dimensions, demonstrating the suitability of linear SVM and TF–IDF features for high-dimensional multilabel text classification. The structural relationships within TAM were also confirmed using correlations derived from the automatically generated TAM–Sentiment labels, indicating that PU and PEU strongly influence AT, and that AT and BIU remain meaningful predictors of AUS. These findings show that TAM can be effectively operationalized using unstructured review data, extending its applicability beyond conventional survey-based methods. A prototype system was further deployed to support real-time TAM prediction, highlighting the practical value of the proposed framework. Future work will explore contextual language models, broader lexical resources, and advanced multilabel architectures to enhance semantic representation and classification accuracy further.

REFERENCES

[1] N. Patel, L. P. Ramisetty, A. S. Parmar, J. Li, and Y. Zhang, "4EV: Adaptive Video Editing with Spatial Temporal Dynamics and Motion Pathways," *IEEE Access*, 2025, doi: 10.1109/ACCESS.2025.3624055.

[2] J. K. Adarsh, V. T. Sreedevi, and D. Thangavelusamy, "Product Review System With BERT for Sentiment Analysis and Implementation of Administrative Privileges on Node-RED," *IEEE Access*, vol. 11, pp. 65968–65976, 2023, doi: 10.1109/ACCESS.2023.3275738.

[3] P. Silva, "Davis' Technology Acceptance Model (TAM) (1989)," in *Information Seeking Behavior and Technology Adoption: Theories and Trends*, 2015. doi: 10.4018/978-1-4666-8156-9.ch013.

[4] L. Lun, D. Zetian, T. W. Hoe, X. Juan, D. Jiaxin, and W. Fulai, "Factors Influencing User Intentions on Interactive Websites: Insights from the Technology Acceptance Model," *IEEE Access*, vol. 12, pp. 122735–122756, 2024, doi: 10.1109/ACCESS.2024.3437418.

[5] B. Nakisa, F. Ansarizadeh, P. Oommen, and S. Shrestha, "Technology Acceptance Model: A Case Study of Palm Vein Authentication Technology," *IEEE Access*, vol. 10, pp. 120436–120449, 2022, doi: 10.1109/ACCESS.2022.3221413.

[6] L. Tobarra, J. M. Haut, R. Hernández, R. Pastor-Vargas, and A. Robles-Gómez, "Analyzing the users' acceptance of an IoT cloud platform using the UTAUT/TAM model," *IEEE Access*, vol. 9, pp. 150004–150020, 2021, doi: 10.1109/ACCESS.2021.3125497.

[7] J. Bai and X. Li, "Chinese Multilabel Short Text Classification Method Based on GAN and Pinyin Embedding," *IEEE Access*, vol. 12, pp. 83323–83329, 2024, doi: 10.1109/ACCESS.2024.3412649.

[8] K. I. Gunawan and J. Santoso, "Multilabel Text Classification Menggunakan SVM dan Doc2Vec Classification Pada Dokumen Berita Bahasa Indonesia," *Journal of Information System,Graphics, Hospitality and Technology*, vol. 3, no. 01, pp. 29–38, Apr. 2021, doi: 10.37823/insight.v3i01.126.

[9] S. Sayardoost Tabrizi, A. Sabzian, and A. Moeini, "A Hybrid Model for Evaluation of E-Learning User Satisfaction with TAM and ELQ Approach," *International Journal of Information and Communication Technology Research*, vol. 16, no. 2, pp. 45–54, Apr. 2024, doi: 10.61186/itrc.16.2.45.

[10] Y. Findawati, D. Purwitasari, and A. B. Raharjo, "Dangerous Speech Classification on Twitter using Multilabel Aspects and Structural Features," *IEEE Access*, p. 1, 2025, doi: 10.1109/ACCESS.2025.3627932.

[11] S. M. Alzanin, A. Gumaei, M. A. Haque, and A. Y. Muaad, "An Optimized Arabic Multilabel Text Classification Approach Using Genetic Algorithm and Ensemble Learning," *Applied Sciences (Switzerland)*, vol. 13, no. 18, 2023, doi: 10.3390/app131810264.

[12] J. Mutinda, W. Mwangi, and G. Okeyo, "Sentiment Analysis of Text Reviews Using Lexicon-Enhanced Bert Embedding (LeBERT) Model with Convolutional Neural Network," *Applied Sciences (Switzerland)*, vol. 13, no. 3, 2023, doi: 10.3390/app13031445.

[13] Y. Asri, W. N. Suliyanti, D. Kuswardani, M. Fajri, "Pelabelan Otomatis Lexicon Vader dan Klasifikasi Naïve Bayes dalam Menganalisis Sentimen Data Ulasan PLN Mobile," *Jurnal Pengkajian dan Penerapan Teknik Informatika,* vol. 15 No. 2, 2022, doi: 10.33322/petir.v15i2.1733.

[14] D. M. Abdullah and A. M. Abdulazeez, "Machine Learning Applications based on SVM Classification: A Review," *Qubahan Academic Journal*, vol. Vol.1 No.2, 2021, doi: 10.48161/qaj.v1n2a50.

[15] X. Luo, "Efficient English text classification using selected Machine Learning Techniques," *Alexandria Engineering Journal*, vol. 60, no. 3, 2021, doi: 10.1016/j.aej.2021.02.009.

[16] T. Carneiro, R. V. M. Da Nobrega, T. Nepomuceno, G. Bin Bian, V. H. C. De Albuquerque, and P. P. R. Filho, "Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications," *IEEE Access*, vol. 6, pp. 61677–61685, 2018, doi: 10.1109/ACCESS.2018.2874767.

[17] Y. Wang, W. Meijer, J. Antonio Hernández López, U. Nilsson, and D. Varró, "Why Do Machine Learning Notebooks Crash? An Empirical Study on Public Python Jupyter Notebooks," *IEEE Transactions on Software Engineering*, vol. 51, no. 7, pp. 2181–2196, 2025, doi: 10.1109/TSE.2025.3574500.

[18] M. A. A. Mosleh, A. A. A. Mohammed, E. E. A. Esmail, R. A. A. Mohammed, and B. Almuhaya, "Hybrid Deep Learning and Fuzzy Matching for Real-Time Bidirectional Arabic Sign Language Translation: Toward Inclusive Communication Technologies," *IEEE Access*, vol. 13, pp. 94118–94136, 2025, doi: 10.1109/ACCESS.2025.3574103.

[19] Y. Akkem, B. S. Kumar, and A. Varanasi, "Streamlit Application for Advanced Ensemble Learning Methods in Crop Recommendation Systems – A Review and Implementation," *Indian J Sci Technol*, vol. 16, no. 48, pp. 4688–4702, Dec. 2023, doi: 10.17485/IJST/v16i48.2850.

[20] C. Ahmed, A. ElKorany, and E. ElSayed, "Prediction of customer's perception in social networks by integrating sentiment analysis and machine learning," *J Intell Inf Syst*, vol. 60, no. 3, 2023, doi: 10.1007/s10844-022-00756-y.

[21] V. Umarani, A. Julian, and J. Deepa, "Sentiment Analysis using various Machine Learning and Deep Learning Techniques," *Journal of the Nigerian Society of Physical Sciences*, vol. 3, no. 4, 2021, doi: 10.46481/jnsps.2021.308.

[22] A. Samyuktha, M. Pallavi, L. Jagan, D. . Y. Srinivasulu, and Mr. M. Rakesh, "Sentiment Analysis Using Machine Learning," *Int J Res Appl Sci Eng Technol*, vol. 11, no. 1, 2023, doi: 10.22214/ijraset.2023.48706.

[23] M. Costola, O. Hinz, M. Nofer, and L. Pelizzon, "Machine learning sentiment analysis, COVID-19 news and stock market reactions," *Res Int Bus Finance*, vol. 64, 2023, doi: 10.1016/j.ribaf.2023.101881.

[24] I. C. Obagbuwa, S. Danster, and O. C. Chibaya, "Supervised machine learning models for depression sentiment analysis," *Front Artif Intell*, vol. 6, 2023, doi: 10.3389/frai.2023.1230649.

[25] E. Elgeldawi, A. Sayed, A. R. Galal, and A. M. Zaki, "Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis," *Informatics*, vol. 8, no. 4, 2021, doi: 10.3390/informatics8040079.

[26] R. Wijayanti and A. Arisal, "Automatic Indonesian Sentiment Lexicon Curation with Sentiment Valence Tuning for Social Media Sentiment Analysis," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 20, no. 1, 2021, doi: 10.1145/3425632.

[27] H. R. Bosker, "Using fuzzy string matching for automated assessment of listener transcripts in speech intelligibility studies," *Behav Res Methods*, vol. 53, no. 5, 2021, doi: 10.3758/s13428-021-01542-4.

[28] W. Suwarningsih and N. Nuryani, "Generate fuzzy string-matching to build self attention on Indonesian medical-chatbot," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 1, 2024, doi: 10.11591/ijece.v14i1.pp819-829.

[29] W. K. Härdle, D. D. Prastyo, and C. M. Hafner, "Support Vector Machines with Evolutionary Feature Selection for Default Prediction," *SSRN Electronic Journal*, 2017, doi: 10.2139/ssrn.2894201.

[30] J. Read, B. Pfahringer, G. Holmes, and E. Frank, "Classifier chains for multi-label classification," *Mach Learn*, vol. 85, no. 3, 2011, doi: 10.1007/s10994-011-5256-5.

[31] A. Abdullahi, N. A. Samsudin, S. K. A. Khalid, and Z. A. Othman, "An Improved Multi-label Classifier Chain Method for Automated Text Classification," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 3, 2021, doi: 10.14569/IJACSA.2021.0120352.