IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Vol.20, No.2, April 2026, pp. 199~210
ISSN (print): 1978-1520, ISSN (online): 2460-7258
DOI: 10.22146/ijccs.116764

Evaluating Latent Emotional Structures through Unsupervised Semantic Text Clustering

Edy*[1], Junaedi[2], Aditiya Hermawan[3], Yusuf Kurnia[4], Ardiane Rossi Kurniawan Maranto[5]
[1,2,3,4]Department Of Sains and Technology, Buddhi Dharma University, Tangerang, Indonesia
e-mail: ***[1]edy.edy@ubd.ac.id**, [2]junaedi.junaedi@ubd.ac.id, [3]aditiya.hermawan@ubd.ac.id,
[4]yusuf.kurnia@ubd.ac.id, [5]ardiane.rossi@ubd.ac.id

Abstrak

Analisis emosi pada teks merupakan topik penting dalam pemrosesan bahasa alami karena emosi berperan besar dalam memahami opini publik, kondisi psikologis, dan dinamika interaksi digital. Namun, sebagian besar penelitian sebelumnya masih bergantung pada pendekatan klasifikasi terawasi berbasis label emosi, yang berpotensi mengabaikan struktur semantik laten dan tumpang tindih antar emosi dalam bahasa alami. Penelitian ini bertujuan untuk mengevaluasi struktur emosi laten dalam teks melalui pendekatan klasterisasi semantik tanpa pengawasan. Metode yang digunakan meliputi prapemrosesan teks, representasi fitur menggunakan Term Frequency–Inverse Document Frequency (TF–IDF), reduksi dimensi dengan Singular Value Decomposition (SVD), serta klasterisasi menggunakan algoritma K-Means dan Hierarchical Agglomerative Clustering. Evaluasi dilakukan menggunakan metrik internal dan eksternal secara pasca-hoc untuk menganalisis kualitas klaster dan keterkaitannya dengan label emosi yang tersedia. Hasil penelitian menunjukkan bahwa K-Means menghasilkan klaster yang lebih baik dan interpretatif dibandingkan pendekatan hierarkis, sementara kedua metode sama-sama memperlihatkan tingkat tumpang tindih emosi yang tinggi. Temuan ini mengindikasikan bahwa ekspresi emosi dalam teks bersifat kontinu dan tidak sepenuhnya terwakili oleh kategori diskret. Penelitian ini menegaskan pentingnya pendekatan klasterisasi semantik sebagai alat analitis untuk memahami struktur emosi laten dalam teks secara lebih mendalam.

Kata kunci*—Analisis Emosi; Unsupervised Learning; Klasterisasi Semantik; Penambangan Teks; Pemrosesan Bahasa Alami*

Abstract

Emotion analysis in textual data is an important topic in natural language processing, as emotions play a crucial role in understanding public opinion, psychological states, and dynamics of digital interaction. However, most existing studies rely heavily on supervised classification approaches based on predefined emotion labels, which may overlook latent semantic structures and emotional overlap inherent in natural language. This study aims to evaluate latent emotional structures in text using an unsupervised semantic clustering approach. The proposed method involves text preprocessing, feature representation using Term Frequency–Inverse Document Frequency (TF–IDF), dimensionality reduction through Singular Value Decomposition (SVD), and clustering using K-Means and Hierarchical Agglomerative algorithms. Both internal and post-hoc external evaluation metrics are employed to assess cluster quality and examine their correspondence with available emotion labels. The results indicate that K-Means clustering produces more compact and interpretable clusters than the hierarchical approach, while both methods reveal substantial emotional overlap across clusters. These findings suggest that emotional expressions in text exhibit a continuous semantic structure rather than discrete categorical boundaries. This study highlights the importance of unsupervised semantic clustering as an analytical tool for gaining deeper insight into latent emotional patterns in textual data.

Keywords— *Emotion Analysis, Unsupervised Learning, Semantic Clustering, Text Mining, Natural Language Processing*

Received January 7th,2026; Revised March 20th, 2026; Accepted April 28th, 2026

1. INTRODUCTION

The automatic analysis of emotions expressed in textual data has become a central research problem in natural language processing (NLP) and affective computing, driven by the exponential growth of user-generated text across digital platforms. Emotional information embedded in text plays a crucial role in diverse applications, including public opinion monitoring, mental health assessment, crisis detection, marketing analytics, and policy evaluation. Despite significant progress in computational methods, the question of how emotions are structurally organized in language, and whether existing computational models can faithfully capture this organization, remains poorly understood, particularly given the context-sensitive and multidimensional nature of affective expression [1].

Most contemporary approaches to emotion analysis conceptualize the task as a supervised classification problem, where textual inputs are mapped onto predefined emotion categories using labeled datasets. Recent advances in deep learning, particularly transformer-based architectures, have led to substantial performance gains under this paradigm, reinforcing the dominance of label-driven emotion modeling in the literature [2], [3]. However, growing empirical evidence indicates that high classification accuracy does not necessarily equate to a semantically accurate or theoretically grounded representation of emotional meaning. Emotion labels are often heterogeneous, overlapping, and influenced by annotation guidelines or cultural assumptions, raising concerns about whether supervised models capture genuine affective structure or merely replicate annotation conventions [4], [5].

This tension has led to an emerging contradiction in the field. On the one hand, supervised emotion classifiers continue to improve quantitatively on benchmark datasets. On the other hand, qualitative analyses increasingly suggest that emotional expressions in text are continuous, context-dependent, and semantically intersecting, rather than cleanly separable into discrete categories [6]. Studies in affective computing and cognitive modeling further indicate that emotional states often manifest as blended or transitional forms rather than isolated classes, challenging rigid taxonomic representations [7]. As a result, the reliance on fixed emotion taxonomies may obscure latent patterns of emotional expression that do not align with categorical boundaries. This contradiction highlights a critical gap in current knowledge: the lack of systematic investigation into the latent semantic organization of emotions beyond label-based frameworks.

In response to these limitations, unsupervised and exploratory approaches have gained renewed attention as complementary strategies for emotion analysis. Rather than predicting predefined labels, unsupervised methods aim to discover intrinsic structures in textual data based on semantic similarity. Among these approaches, semantic text clustering offers a principled way to group texts based on shared linguistic and affective characteristics, potentially revealing latent emotional patterns that emerge naturally from the data. Prior research has demonstrated that clustering techniques can uncover meaningful groupings in large text corpora, including affective and sentiment-related patterns, without relying on annotated labels [8]. Recent work further suggests that clustering based on semantic representations can expose affective regularities that are not readily observable through supervised classification alone [9].

Nonetheless, existing clustering-based studies on emotion analysis remain fragmented and conceptually underdeveloped. Many works employ clustering primarily as a technical preprocessing step such as reducing class cardinality or supporting downstream classification rather than as an analytical lens for interrogating emotional structure itself. For example, recent studies applying clustering to emotion-labeled text often emphasize visualization or efficiency gains while offering limited theoretical interpretation of what the resulting clusters signify in affective terms [10]. As demonstrated in prior work, clustering outcomes frequently reveal substantial overlap among emotion categories, suggesting that emotional expressions may share common semantic foundations rather than forming isolated groups [10]. Similar observations have been reported in studies leveraging large-scale semantic embeddings, where clusters tend to reflect gradual affective continua rather than discrete emotional partitions [11].

Furthermore, methodological critiques have highlighted that interpreting clustering results is far from straightforward. Internal validation metrics commonly used to assess clustering quality can behave inconsistently when applied to high-dimensional semantic data, and recent evidence shows that some widely adopted external validation measures may exhibit biased behavior under certain configurations [12]. This concern is compounded in emotion analysis, where semantic proximity does not always translate into clear cluster boundaries due to inherent affective ambiguity. Consequently, low cluster separability should not automatically be interpreted as methodological failure; instead, it may reflect fundamental properties of emotional language, such as gradual transitions and semantic overlap between affective states, as also noted in recent affective computing surveys [1].

Despite these insights, a clear research gap persists. There is a limited systematic understanding of whether unsupervised semantic clustering can meaningfully capture latent emotional structures in text and how such structures relate to, diverge from, or challenge predefined emotion labels. In particular, few studies explicitly frame clustering as a tool for evaluating the adequacy of categorical emotion representations or for examining the continuity of emotional meaning in language. Moreover, existing research rarely articulates clustering results within broader theoretical debates on emotion representation, leaving their epistemic implications underexplored despite calls for theory-aware affective modeling [7].

The present study addresses this gap by reframing unsupervised semantic text clustering as an analytical instrument for evaluating emotional structure rather than as a purely technical procedure. Instead of optimizing predictive performance, the study treats clustering outcomes as empirical evidence of how emotional expressions are organized in semantic space. Building on prior findings that show logical but overlapping emotion groupings emerging from clustering analyses [10], this work advances a more theoretically informed interpretation of clusters as manifestations of latent affective organization rather than approximations of discrete emotion classes. Recent advances in semantic representation learning further motivate this perspective by enabling richer, more flexible modeling of affective similarity [9], [11].

The innovation of this study lies in its conceptual repositioning of clustering-based emotion analysis as an evaluative framework rather than a predictive task. Instead of focusing on label prediction accuracy, this research examines whether unsupervised semantic text clustering can uncover coherent latent emotional structures independent of predefined categories and what such structures reveal about the semantic organization of emotions in language. By emphasizing interpretability, post-hoc validation, and alignment with contemporary theories that conceptualize emotion as a continuous and context-sensitive phenomenon, this study addresses a critical gap at the intersection of NLP methodology and emotion theory. It contributes a theory-aware analytical perspective to affective computing research beyond performance-driven evaluation [5].

2. METHODS

This study employs a structured methodological framework to investigate latent emotional structures in textual data through unsupervised semantic clustering. As illustrated in Figure 1, the workflow comprises data understanding and exploratory analysis, text preprocessing, semantic representation and dimensionality reduction, unsupervised clustering, and post-hoc evaluation to analyze latent emotional organization based on semantic similarity rather than predefined labels.

2.1 Data Source and Characteristics

This study uses the “Emotion Detection from Text” dataset [13], which is publicly available on Kaggle and widely used in emotion analysis research. The dataset consists of short user-generated text entries originally collected from social media and is structured to support large-scale emotion detection tasks. In its commonly distributed form, the dataset contains approximately 40,000 records organized into four primary fields, including a unique identifier (e.g., tweet ID), an emotion or sentiment label, and the raw textual content. The dataset schema, as documented in independent technical reports, confirms that the key attributes comprise *tweet_id*, *sentiment* (emotion label), and *content* (unfiltered text), reflecting its explicit design for mapping short textual instances to emotion categories [14].

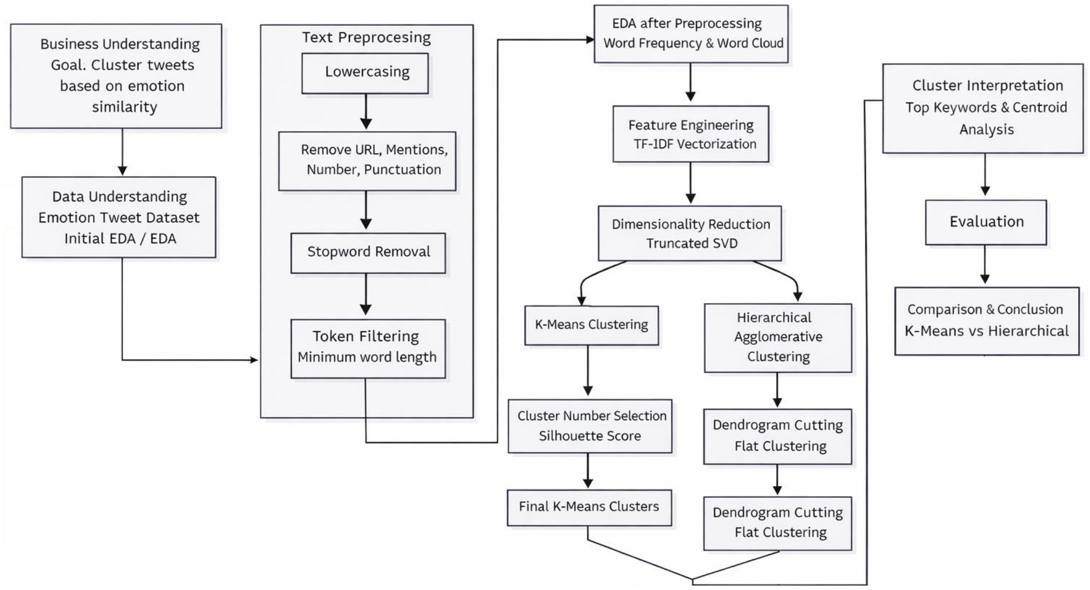

Figure 1 Research Method

These characteristics make the dataset particularly suitable for the present study, which aims to evaluate latent affective structures in short-form text. Short social-media texts exhibit high lexical variability, contextual ambiguity, and substantial semantic overlap across emotion categories, posing challenges for rigid label-based representations. Although emotion labels are provided, this study deliberately treats them as post-hoc reference signals for external comparison and interpretive validation rather than as supervision targets during model construction. This design choice ensures that the clustering process remains fully unsupervised and that the resulting groupings reflect intrinsic semantic similarity patterns in the text, enabling a principled examination of the alignment—or divergence—between categorical emotion annotations and latent semantic structures in emotional language.

2.2 Data Understanding and Exploratory Analysis

An initial data understanding phase is conducted to examine the distribution of emotion labels and the intrinsic characteristics of the textual corpus before any transformation or modeling. The dataset consists of short-form user-generated texts annotated with multiple emotion categories, a structure known to introduce high lexical variability, contextual ambiguity, and informal language usage. Exploratory data analysis (EDA) is therefore performed on the raw text to assess text length distribution, vocabulary richness, and the prevalence of noise elements such as URLs, user mentions, hashtags, numerical tokens, and punctuation, which are particularly relevant for social media–derived corpora where surface-level noise can substantially distort semantic representations if left unaddressed [15]. The EDA results reveal considerable lexical and semantic overlap among emotion labels, reflected in the frequent co-occurrence of similar terms across categories, a phenomenon widely reported in recent emotion-focused analyses of social media text that highlight annotation subjectivity and reduced categorical distinction in applied schemes [16]. Related syntheses further emphasize the prevalence of mixed or overlapping affective states and the need to treat taxonomy differences as substantive methodological issues rather than noise [7], while social media studies explicitly motivate moving beyond single-label assumptions because texts often encode complex and mixed emotions [17]. Consequently, these observations motivate adopting unsupervised clustering as an exploratory analytical framework, enabling investigation of latent emotional structures based on semantic similarity without reliance on predefined labels and supporting a more nuanced, structure-oriented analysis of emotional language.

2.3 Text Preprocessing and Representation

Text preprocessing is conducted to reduce noise while preserving semantically meaningful information relevant to emotional expression in short textual data. The preprocessing pipeline includes handling missing values, removing non-linguistic elements such as URLs, user mentions, and hashtags, eliminating stopwords, converting text to lowercase for lexical consistency, and inspecting post-preprocessing text length to ensure that emotionally informative content is not excessively truncated. These steps are particularly critical for social media–derived text, where informal language use and platform-specific artifacts can obscure genuine semantic similarity patterns, and the preprocessing strategy is therefore intentionally conservative to reduce superficial variability while preserving affective nuance and lexical diversity. Following preprocessing, textual data are transformed into numerical representations using the Term Frequency–Inverse Document Frequency (TF-IDF) scheme, which emphasizes terms salient within individual documents while down-weighting ubiquitous corpus-level terms, making it well-suited to short, emotion-laden texts where emotionally meaningful words may appear infrequently [18]. In addition to its effectiveness, TF-IDF offers high interpretability, as cluster characteristics can be examined through representative keywords derived from term weights, enabling transparent semantic inspection and qualitative analysis of latent themes. Although more complex embedding techniques are available, TF-IDF is deliberately employed to maintain methodological transparency and interpretability, as it supports subsequent clustering and thematic interpretation through keyword inspection [19].

2.4 Dimensionality Reduction via Singular Value Decomposition

To address the high dimensionality and sparsity of TF-IDF representations, truncated Singular Value Decomposition (SVD) is applied to obtain a low-dimensional semantic space, commonly referred to as Latent Semantic Analysis/Indexing (LSA/LSI) in textual applications, which captures latent semantic structure through dominant term–document co-occurrence patterns [20]. By decomposing the TF-IDF matrix into orthogonal components and retaining the most significant singular vectors, SVD filters out noise and weakly informative features, thereby enhancing the interpretability, stability, and computational efficiency of subsequent clustering operations while preserving critical semantic relationships in high-variability short texts [21]. In this study, dimensionality is reduced from 5,000 features to 200 latent components, yielding a dense representation that mitigates sparsity effects and improves the robustness of distance-based clustering. The transformation is applied in a fully unsupervised manner without introducing emotion-specific priors, ensuring that the resulting latent space reflects inherent semantic structure rather than predefined emotional assumptions.

2.5 Clustering Algorithms and Interpretation

Two unsupervised clustering algorithms, K-Means and Hierarchical Agglomerative clustering, are employed to identify latent emotional groupings within the reduced semantic space. These methods are selected for their complementary analytical characteristics and for their widespread use in exploratory text clustering, enabling comparisons between partition-based optimization and hierarchical data aggregation (Petukhova et al., 2025). K-Means clustering partitions a dataset into a predefined number of clusters by minimizing the within-cluster sum of squares, which is commonly used as a measure of within-cluster variance, over vector representations in the latent semantic space [22], with the number of clusters determined using both the Silhouette coefficient and the Elbow method to reflect different structural assumptions. Silhouette-based model selection emphasizes cluster compactness and inter-cluster separation. In contrast, the Elbow criterion selects the number of clusters by inspecting diminishing returns in the reduction of within-cluster dispersion. Consequently, the two criteria can yield different k solutions (e.g., 12 vs. 7), enabling a systematic comparison of clustering outcomes under distinct, yet theoretically motivated, validity assumptions [23]. Importantly, internal cluster validity indices (including Silhouette) are treated as exploratory heuristics rather than definitive evidence of intrinsic cluster structure, because different indices emphasize different notions of cohesion/separation and can behave differently under high-dimensional settings [24].

Hierarchical Agglomerative Clustering is a contrastive clustering approach that constructs a dendrogram via a bottom-up procedure, starting from singleton clusters and progressively merging the most similar clusters based on a linkage criterion, thereby enabling exploratory analysis without requiring an a priori specification of cluster centroids [25]. Flat cluster partitions are extracted from the hierarchical structure to enable direct comparison with K-Means results and to assess whether the data naturally form balanced groupings or converge toward dominant clusters. To support interpretive analysis, representative top keywords are extracted from the TF-IDF feature space for each identified cluster by examining terms with the highest average TF-IDF weights, serving as proxies for dominant semantic and affective themes. Visualization techniques are employed to complement qualitative interpretation rather than to assert strict separability, including two-dimensional scatter plots based on SVD projections to illustrate cluster overlap and heatmaps to examine post-hoc correspondence between clustering outcomes and original emotion labels. Together, these analytical components provide a multifaceted perspective on latent emotional structure, emphasizing semantic continuity and overlap rather than categorical separation.

2.6 Evaluation Strategy

Clustering performance is evaluated using a combination of internal and external (post-hoc) validation metrics to assess cluster quality from complementary perspectives, as commonly recommended in unsupervised learning studies [26]. Internal evaluation employs the Silhouette coefficient, which measures both cluster compactness and inter-cluster separation in the latent semantic space and is widely regarded as a robust intrinsic validity index for clustering analysis. Based on this criterion, K-Means demonstrates a higher Silhouette score than Hierarchical Agglomerative clustering, indicating more compact and structurally stable partitions under centroid-based optimization [27]. External evaluation is conducted post hoc using Normalized Mutual Information (NMI), Adjusted Rand Index (ARI), and Purity with respect to the original emotion labels; these measures are not used for model selection or optimization, but are solely used to support interpretive analysis of clustering outcomes against reference annotations. The generally low NMI and ARI values observed across both clustering methods reflect limited correspondence between induced clusters and predefined labels, a typical and theoretically expected outcome in unsupervised settings that reflects semantic overlap and the continuous nature of emotional expression in text [28]. The relatively higher Purity achieved by K-Means suggests the presence of dominant emotion tendencies within certain clusters; however, the absence of a one-to-one correspondence with emotion categories reinforces the interpretation that the clustering results primarily capture latent semantic similarity rather than discrete emotions.

3. RESULTS AND DISCUSSION

This section presents the experimental results of unsupervised semantic clustering applied to the emotion text dataset using the methodological pipeline described previously, including TF-IDF vectorization, dimensionality reduction via Singular Value Decomposition (SVD), and clustering with K-Means and Hierarchical Agglomerative approaches.

Dimensionality reduction from 5,000 terms to 200 latent components substantially reduces sparsity while preserving the dominant semantic variance, enabling stable distance-based clustering; however, visualization of the SVD-projected space reveals dense, overlapping regions that do not align with predefined emotion labels. Quantitative evaluation metrics and qualitative analysis of representative keywords indicate that K-Means produces more compact, interpretable flat clusters than the hierarchical approach. At the same time, a post hoc comparison with the original emotion labels shows weak correspondence between the two methods. These findings are theoretically consistent with the semantic overlap and continuity of emotional expressions and suggest that categorical emotion labels capture only part of the underlying organization of emotional language, thereby reinforcing the value of unsupervised semantic clustering as an exploratory framework for uncovering latent affective structures beyond rigid classification schemes.

3.1 K-Means Clustering

K-Means clustering was applied in two configurations to examine cluster structure at different levels of granularity, with the number of clusters set to 12 based on the Silhouette coefficient and to 7 based on the Elbow method. In the twelve-cluster configuration, internal evaluation yielded a Silhouette score of 0.0391, indicating higher cluster compactness and separation than in the hierarchical approach. The corresponding heatmap (Figure 2) illustrates the distribution of original emotion labels across clusters and shows that Cluster 4 emerges as the largest, dominated by the *worry*, *sadness*, and *neutral* labels. This pattern suggests that expressions of general negative affect and emotionally neutral statements are the most prevalent semantic structures in the dataset, outweighing more specific or intense emotional categories. It reinforces the presence of overlapping emotional expressions in textual data.

Color intensity represents the frequency of text instances per cluster–label combination. The semantic characteristics of the clusters are further elucidated through analysis of the top keywords extracted from each cluster (Table 1). The keyword distributions reveal that clusters are defined by recurring lexical patterns related to concern, uncertainty, and everyday discourse, rather than by exclusive emotion-specific vocabularies. While some clusters show partial dominance of particular emotion labels, none exhibit complete homogeneity, underscoring the semantic overlap among emotional expressions.

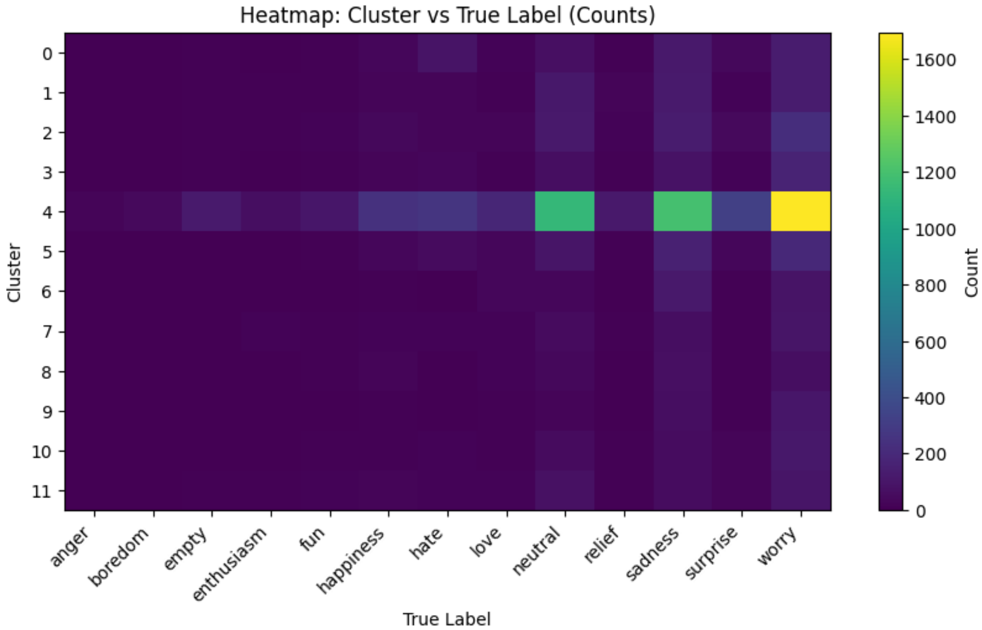

Figure 2 Heatmap of K-Means clustering results (12 clusters) versus original emotion labels.

Table 1 Representative top keywords for selected K-Means clusters in the 12-cluster configuration.

Cluster ID	Cluster Size	Representative Keywords (TF–IDF weight)
Cluster 0	510	*quot* (0.1281), *hate* (0.1171), *last* (0.1054), *night* (0.0655), *last night* (0.0569), *last day* (0.0298), *day* (0.0261), *today* (0.0129), *one* (0.0120), *much* (0.0082),
Cluster 1	497	*work* (0.2753), *day* (0.0286), *back work* (0.0258), *today* (0.0248), *back* (0.0217), *get* (0.0201), *work today* (0.0178), *going* (0.0169), *weekend* (0.0151),
Cluster 2	652	*get* (0.1434), *going* (0.1178), *tonight* (0.0140), *see* (0.0134), *back* (0.0127), *today* (0.0118), *home* (0.0115), *day* (0.0112), *get see* (0.0096), *well* (0.0095),

When the number of clusters was reduced to 7 using the Elbow method, the overall clustering structure remained consistent. Cluster 0 became the dominant cluster, again primarily associated with *worry*, *sadness*, and *neutral* labels. The distribution of data points in the SVD space (Figure 3) demonstrates that merging clusters does not substantially alter the degree of overlap among emotional expressions, indicating that the underlying semantic continuity persists regardless of cluster granularity.

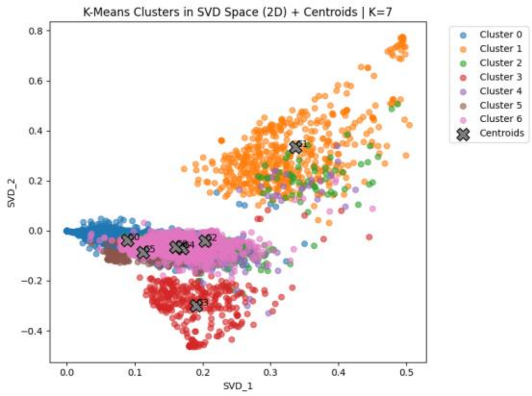

Figure 3 Distribution of text instances in the SVD space based on K-Means clusters (7-cluster configuration)

3.2 Visualization Based on Original Emotion Labels

To further examine the relationship between latent semantic structure and annotation schemes, the data were visualized in the SVD-projected space according to their original emotion labels (Figure 4). The visualization reveals extensive overlap among labels, with no clear boundaries separating individual emotional categories. However, several dense regions are observed; these regions do not correspond to specific labels, indicating shared lexical and contextual features across emotions. This finding provides empirical evidence that emotional expressions in text are distributed along a semantic continuum rather than forming isolated clusters, thereby explaining the limited alignment between unsupervised clustering results and predefined emotion labels.

Figure 4 SVD-based visualization of text instances colored by original emotion labels.

3.3 Hierarchical Agglomerative Clustering

Hierarchical Agglomerative clustering was applied as a contrastive method to explore alternative structural assumptions. The heatmap presented in Figure 5 shows that a substantial majority of data points are aggregated into a single dominant cluster (Cluster 0), which is again primarily composed of worry, sadness, and neutral labels. The remaining clusters contain only a small number of instances and do not form semantically coherent or balanced groupings.

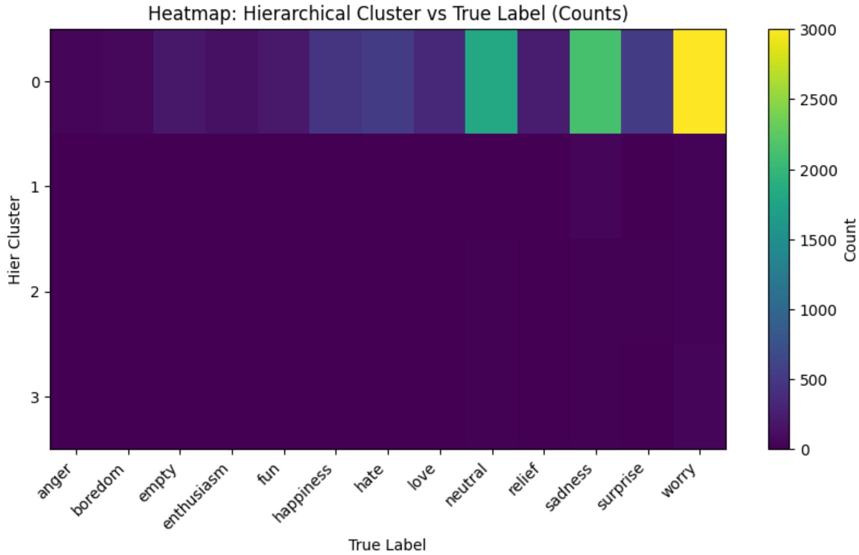

Figure 5 Heatmap of Hierarchical Agglomerative clustering results versus original emotion labels.

Top keyword analysis for the hierarchical clusters (Table 2) reveals minimal differentiation among clusters, confirming that the flat clustering solution produced by this method is less effective for separating emotional patterns in this dataset. Nonetheless, the hierarchical approach remains informative for exploratory analysis, particularly for examining nested relationships through dendrogram structures.

Table 2 Representative keywords for selected clusters obtained via Hierarchical Agglomerative clustering.

Hierarchical Cluster ID	Cluster Size (n)	Representative Keywords (TF–IDF weight)
Cluster 0	9,656	*work* (0.0165), *get* (0.0130), *day* (0.0119), *today* (0.0113), *like* (0.0110), *got* (0.0109), *going* (0.0108), *want* (0.0102), *know* (0.0099), *back*
Cluster 1	97	*miss* (0.4721), *going miss* (0.0358), *going* (0.0286), *really miss* (0.0204), *miss place* (0.0186), *back* (0.0153), *want* (0.0153), *miss puppy* (0.0148),
Cluster 3	103	*amp* (0.4654), *today amp* (0.0197), *going* (0.0191), *back amp* (0.0187), *know* (0.0178), *good* (0.0149), *ben* (0.0137), *thought* (0.0133), *day amp*

3.4 Discussion

The comparative analysis reveals substantive differences in cluster quality and interpretability between K-Means and Hierarchical Agglomerative clustering. Internal evaluation using the Silhouette coefficient indicates that K-Means produces more compact, better-separated clusters than the hierarchical approach (0.0391 versus 0.0119). While these absolute values remain low, the consistent advantage of K-Means suggests that partition-based optimization is more effective for capturing localized semantic coherence in short emotional texts than bottom-up aggregation under the present experimental conditions. This observation aligns with recent clustering studies showing that centroid-based methods tend to outperform hierarchical strategies when applied to high-dimensional semantic representations, particularly in scenarios characterized by gradual transitions rather than well-separated groups [23], [27]. The notably low Silhouette score obtained for Hierarchical clustering further indicates substantial overlap among clusters and highlights its limited suitability for producing balanced flat partitions in reduced semantic spaces dominated by overlapping affective expressions.

Post-hoc external evaluation using Normalized Mutual Information (NMI), Adjusted Rand Index (ARI), and Purity further demonstrates weak correspondence between the discovered clusters and the original emotion labels across both clustering methods. Importantly, this limited alignment should not be interpreted as methodological inadequacy; rather, it is theoretically consistent with the unsupervised nature of the analysis and with known properties of emotional language, where semantic proximity does not necessarily imply categorical equivalence. Recent empirical work has shown that external agreement measures such as NMI and ARI can exhibit biased or misleading behavior when applied to clustering solutions with uneven cluster sizes or overlapping structures, reinforcing the need for cautious interpretation in semantic applications [12]. Within this context, the slightly higher Purity achieved by K-Means suggests the presence of dominant emotional tendencies within certain clusters; however, the absence of one-to-one mappings between clusters and emotion labels underscores the inherent limitations of categorical annotation schemes when confronted with context-dependent and blended emotional expressions, as also reported in recent short-text clustering studies based on semantic similarity [9].

From a theoretical perspective, these findings provide empirical support for conceptualizing emotions in text as semantically overlapping and continuous phenomena rather than as discrete and mutually exclusive categories. The observed clustering patterns demonstrate that unsupervised semantic grouping can reveal latent affective similarity structures that transcend predefined annotation schemes, thereby complementing and extending insights from recent affective computing surveys that emphasize emotional fluidity, contextual dependence, and mixed affective states in natural language [1], [7]. Taken together, the results suggest that while K-Means clustering is more effective than Hierarchical Agglomerative clustering for producing interpretable flat clusters in this dataset, both approaches offer complementary analytical perspectives on latent emotional organization. More broadly, the findings reinforce the study's central claim that emotional meaning in language is best understood as a continuum of semantic expression rather than as a collection of fixed categorical states. That unsupervised clustering can serve as a theoretically informative tool for interrogating the adequacy of label-based emotion representations.

4. CONCLUSIONS

This study demonstrates that unsupervised semantic text clustering constitutes an effective analytical framework for exploring latent emotional structures in short textual data beyond predefined categorical labels. By combining TF–IDF representation, dimensionality reduction via Singular Value Decomposition, and clustering using K-Means and Hierarchical Agglomerative methods, the analysis reveals that emotional expressions in text exhibit substantial semantic overlap and continuity. The experimental results show that K-Means clustering produces more compact and interpretable flat clusters than Hierarchical Agglomerative clustering, as indicated by higher internal validation scores and clearer dominant lexical themes within clusters. At the same time, the weak correspondence between clusters and original emotion labels, as reflected by low external validation scores, provides empirical evidence that categorical emotion annotations only partially capture the underlying semantic organization of emotional language.

Despite these strengths, the proposed approach also presents limitations that must be acknowledged. The reliance on sparse lexical representations restricts the modeling of deeper contextual and compositional semantics. At the same time, hierarchical clustering's effectiveness for flat partitioning is limited when highly overlapping emotional expressions are present. Nevertheless, these limitations do not undermine the study's central contribution; rather, they reinforce the conclusion that emotional meaning in text is inherently fluid, context-dependent, and better conceptualized as a continuum rather than as a set of discrete categories. By reframing clustering outcomes as analytically informative signals rather than classification failures, this work contributes to ongoing theoretical debates in affective computing. It offers a foundation for future research aimed at developing more theory-aware, flexible, and semantically grounded approaches to emotion analysis in natural language processing.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to Universitas Buddhi Dharma for providing institutional and financial support that made this research possible. The support received from the university played a crucial role in facilitating data processing, computational resources, and the overall completion of this study.

REFERENCES

[1] J. Deng and F. Ren, "A Survey of Textual Emotion Recognition and Its Challenges," *IEEE Trans. Affect. Comput.*, vol. 14, no. 1, pp. 49–67, Jan. 2023, doi: 10.1109/TAFFC.2021.3053275.

[2] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, "Transformer models for text-based emotion detection: a review of BERT-based approaches," *Artif. Intell. Rev.*, vol. 54, no. 8, pp. 5789–5829, Dec. 2021, doi: 10.1007/s10462-021-09958-2.

[3] K. Machová, M. Szabóova, J. Paralič, and J. Mičko, "Detection of emotion by text analysis using machine learning," *Front. Psychol.*, vol. 14, p. 1190326, Sep. 2023, doi: 10.3389/fpsyg.2023.1190326.

[4] H. Aka Uymaz and S. Kumova Metin, "Vector based sentiment and emotion analysis from text: A survey," *Eng. Appl. Artif. Intell.*, vol. 113, p. 104922, Aug. 2022, doi: 10.1016/j.engappai.2022.104922.

[5] S. Peng *et al.*, "A survey on deep learning for textual emotion analysis in social networks," *Digit. Commun. Netw.*, vol. 8, no. 5, pp. 745–762, Oct. 2022, doi: 10.1016/j.dcan.2021.10.003.

[6] A. Koufakou and E. Nieves, "Review of recent emotion-annotated text corpora and resources," *Lang. Resour. Eval.*, vol. 59, no. 4, pp. 4313–4347, Dec. 2025, doi: 10.1007/s10579-025-09828-1.

[7] P. Pereira, H. Moniz, and J. P. Carvalho, "Deep emotion recognition in textual conversations: a survey," *Artif. Intell. Rev.*, vol. 58, no. 1, p. 10, Nov. 2024, doi: 10.1007/s10462-024-11010-y.

[8] A. Subakti, H. Murfi, and N. Hariadi, "The performance of BERT as data representation of text clustering," *J. Big Data*, vol. 9, no. 1, p. 15, Dec. 2022, doi: 10.1186/s40537-022-00564-9.

[9] K. Abdalgader, A. A. Matroud, and K. Hossin, "Experimental study on short-text clustering using transformer-based semantic similarity measure," *PeerJ Comput. Sci.*, vol. 10, p. e2078, May 2024, doi: 10.7717/peerj-cs.2078.

[10] S. Salloum, K. Alhumaid, A. Salloum, and K. Shaalan, "K-means Clustering of Tweet Emotions: A 2D PCA Visualization Approach," *Procedia Comput. Sci.*, vol. 244, pp. 30–36, 2024, doi: 10.1016/j.procs.2024.10.175.

[11] A. Petukhova, J. P. Matos-Carvalho, and N. Fachada, "Text clustering with large language model embeddings," *Int. J. Cogn. Comput. Eng.*, vol. 6, pp. 100–108, Dec. 2025, doi: 10.1016/j.ijcce.2024.11.004.

[12] A. Mahmoudi and D. Jemielniak, “Proof of biased behavior of Normalized Mutual Information,” *Sci. Rep.*, vol. 14, no. 1, p. 9021, Apr. 2024, doi: 10.1038/s41598-024-59073-9.
[13] P. Gupta, “Emotion Detection from Text.” [Online]. Available: https://www.kaggle.com/datasets/pashupatigupta/emotion-detection-from-text
[14] Y. Mei and M. Abisado, “BERT-based two-channel neural network model text emotion analysis,” *Salud Cienc. Tecnol.*, vol. 5, p. 1444, Mar. 2025, doi: 10.56294/saludcyt20251444.
[15] F. Barbieri, J. Camacho-Collados, L. Neves, and L. Espinosa-Anke, “TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification,” 2020, *arXiv*. doi: 10.48550/ARXIV.2010.12421.
[16] F. Zhang, J. Chen, Q. Tang, and Y. Tian, “Evaluation of emotion classification schemes in social media text: an annotation-based approach,” *BMC Psychol.*, vol. 12, no. 1, p. 503, Sep. 2024, doi: 10.1186/s40359-024-02008-w.
[17] Y. Li, J. Chan, G. Peko, and D. Sundaram, “Mixed emotion extraction analysis and visualisation of social media text,” *Data Knowl. Eng.*, vol. 148, p. 102220, Nov. 2023, doi: 10.1016/j.datak.2023.102220.
[18] M. Jamalian and H. Vahdat-Nejad, “Effective clustering of big text data: Empirical evaluation and comparative analysis,” *Egypt. Inform. J.*, vol. 32, p. 100812, Dec. 2025, doi: 10.1016/j.eij.2025.100812.
[19] J. Tussupov *et al.*, “Analysis of Short Texts Using Intelligent Clustering Methods,” *Algorithms*, vol. 18, no. 5, p. 289, May 2025, doi: 10.3390/a18050289.
[20] C.-F. Méndez-Cruz, J. Rodríguez-Herrera, A. Varela-Vega, V. Mateo-Estrada, and S. Castillo-Ramírez, “Unsupervised learning and natural language processing highlight research trends in a superbug,” *Front. Artif. Intell.*, vol. 7, p. 1336071, Mar. 2024, doi: 10.3389/frai.2024.1336071.
[21] E. Di Corso, S. Proto, B. Vacchetti, P. Bethaz, and T. Cerquitelli, “Simplifying Text Mining Activities: Scalable and Self-Tuning Methodology for Topic Detection and Characterization,” *Appl. Sci.*, vol. 12, no. 10, p. 5125, May 2022, doi: 10.3390/app12105125.
[22] E. Detrinidad and V.-R. López-Ruiz, “The Interplay of Happiness and Sustainability: A Multidimensional Scaling and K-Means Cluster Approach,” *Sustainability*, vol. 16, no. 22, p. 10068, Nov. 2024, doi: 10.3390/su162210068.
[23] A. M. Bagirov, R. M. Aliguliyev, and N. Sultanova, “Finding compact and well-separated clusters: Clustering using silhouette coefficients,” *Pattern Recognit.*, vol. 135, p. 109144, Mar. 2023, doi: 10.1016/j.patcog.2022.109144.
[24] A. M. Ikotun, F. Habyarimana, and A. E. Ezugwu, “Cluster validity indices for automatic clustering: A comprehensive review,” *Heliyon*, vol. 11, no. 2, p. e41953, Jan. 2025, doi: 10.1016/j.heliyon.2025.e41953.
[25] C. P. Ufeli, M. U. Sattar, R. Hasan, and S. Mahmood, “Enhancing Customer Segmentation Through Factor Analysis of Mixed Data (FAMD)-Based Approach Using K-Means and Hierarchical Clustering Algorithms,” *Information*, vol. 16, no. 6, p. 441, May 2025, doi: 10.3390/info16060441.
[26] M. Alshammari, J. Stavrakakis, and M. Takatsuka, “Refining a k -nearest neighbor graph for a computationally efficient spectral clustering,” *Pattern Recognit.*, vol. 114, p. 107869, Jun. 2021, doi: 10.1016/j.patcog.2021.107869.
[27] E. Schubert, “Stop using the elbow criterion for k-means and how to choose the number of clusters instead,” *ACM SIGKDD Explor. Newsl.*, vol. 25, no. 1, pp. 36–42, Jun. 2023, doi: 10.1145/3606274.3606278.
[28] Y. Ortakci, “Revolutionary text clustering: Investigating transfer learning capacity of SBERT models through pooling techniques,” *Eng. Sci. Technol. Int. J.*, vol. 55, p. 101730, Jul. 2024, doi: 10.1016/j.jestch.2024.101730.

Bukti Korespondensi - IJCCS

IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Accredited Sinta 2 by the Ministry of Research, Technology and Higher Education of the Republic of Indonesia
Publish jointly by The Department of Computer Science and Electronics, FMIPA UGM and IndoCEISS
ISSN 1978-1520 (print); ISSN 2460-7258 (online)

#116764 Summary

SUMMARY | REVIEW | EDITING

Submission

Authors	Edy Edy, Junaedi Junaedi, Aditiya Hermawan, Yusuf Kurnia, Ardiane Rossi Kurniawan Maranto
Title	Evaluating Latent Emotional Structures through Unsupervised Semantic Text Clustering
Original file	116764-432136-1-SM.DOC 2026-02-12
Supp. files	None
Submitter	Edy Edy
Date submitted	February 12, 2026 - 10:30 AM
Section	Articles
Editor	Paholo Prakoso (Editing)
Abstract Views	72

Status

Status	Published Vol 20, No 2 (2026): April
Initiated	2026-04-30
Last modified	2026-05-08

Submission Metadata

Authors

Name	Edy Edy
Affiliation	Universitas Buddhi Dharma
Country	Indonesia
Bio Statement	—

Principal contact for editorial correspondence.

Name	Junaedi Junaedi
Affiliation	Buddhi Dharma University
Country	Indonesia
Bio Statement	—
Name	Aditiya Hermawan
Affiliation	Buddhi Dharma University
Country	Indonesia
Bio Statement	—
Name	Yusuf Kurnia
Affiliation	Buddhi Dharma University
Country	Indonesia
Bio Statement	—
Name	Ardiane Rossi Kurniawan Maranto
Affiliation	Buddhi Dharma University
Country	Indonesia
Bio Statement	—

Editorial Board
Focus & Scope
Ethics & Malpractice Statement
Author Guidelines
Peer Review Process
Reviewer Guidelines
Screening Plagiarism
Online Submission
Author Fee
Statement of Originality
Copyright Transfer Form
Peer Reviewers
The College's Commitment
MoU with IndoCEISS
Decree of Accreditation
New Membership IndoCEISS
Membership Update IndoCEISS
Visitor Statistics

CITATION ANALYSIS
- SCOPUS
- Google Scholar

NOTIFICATIONS
- View (6 new)
- Manage

USER
You are logged in as...
edy_ubd

Bukti Korespondensi - IJCCS

My Journals
My Profile
Log Out

DOWNLOAD

Article template

TOOLS REFERENCE

MENDELEY
grammarly
EndNote ...Bibliographies Made Easy
Zotero

Title and Abstract

Title	Evaluating Latent Emotional Structures through Unsupervised Semantic Text Clustering
Abstract	*Emotion analysis in textual data is an important topic in natural language processing, as emotions play a crucial role in understanding public opinion, psychological states, and dynamics of digital interaction. However, most existing studies rely heavily on supervised classification approaches based on predefined emotion labels, which may overlook latent semantic structures and emotional overlap inherent in natural language. This study aims to evaluate latent emotional structures in text using an unsupervised semantic clustering approach. The proposed method involves text preprocessing, feature representation using Term Frequency–Inverse Document Frequency (TF–IDF), dimensionality reduction through Singular Value Decomposition (SVD), and clustering using K-Means and Hierarchical Agglomerative algorithms. Both internal and post-hoc external evaluation metrics are employed to assess cluster quality and examine their correspondence with available emotion labels. The results indicate that K-Means clustering produces more compact and interpretable clusters than the hierarchical approach, while both methods reveal substantial emotional overlap across clusters. These findings suggest that emotional expressions in text exhibit a continuous semantic structure rather than discrete categorical boundaries. This study highlights the importance of unsupervised semantic clustering as an analytical tool for gaining deeper insight into latent emotional patterns in textual data.*

Indexing

Academic discipline and sub-disciplines	—
Keywords	Emotion Analysis; Unsupervised Learning;Semantic Clustering;Text Mining;Natural Language Processing
Language	en

Supporting Agencies

Agencies	—

OpenAIRE Specific Metadata

ProjectID	—

IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Accredited Sinta 2 by the Ministry of Research, Technology and Higher Education of the Republic of Indonesia
Publish jointly by The Department of Computer Science and Electronics, FMIPA UGM and IndoCEISS
ISSN 1978-1520 (print); ISSN 2460-7258 (online)

Home | About | User Home | Search | Current | Archives | Announcements | Statistics | Indexing | History | Contact | Journal Help

Home > User > Author > Submissions > #116764 > **Review**

#116764 Review

SUMMARY | REVIEW | EDITING

Submission

Authors	Edy Edy, Junaedi Junaedi, Aditiya Hermawan, Yusuf Kurnia, Ardiane Rossi Kurniawan Maranto
Title	Evaluating Latent Emotional Structures through Unsupervised Semantic Text Clustering
Section	Articles
Editor	Paholo Prakoso (Editing)

Peer Review

Round 1

Review Version	116764-432140-2-RV.DOC 2026-03-07
Initiated	2026-03-07
Last modified	2026-04-24
Uploaded file	Reviewer A 116764-440237-1-RV.DOCX 2026-04-24
Editor Version	116764-435159-1-ED.DOC 2026-03-07
Author Version	116764-439539-1-ED.DOC 2026-04-20 116764-439539-2-ED.DOC 2026-04-27

Editorial Board
Focus & Scope
Ethics & Malpractice Statement
Author Guidelines
Peer Review Process
Reviewer Guidelines
Screening Plagiarism
Online Submission
Author Fee
Statement of Originality
Copyright Transfer Form
Peer Reviewers
The College's Commitment
MoU with IndoCEISS

Round 2

Review Version	116764-432140-3-RV.DOC 2026-04-27
Initiated	2026-04-27
Last modified	2026-04-30
Uploaded file	None

Editor Decision

Decision	Accept Submission 2026-04-30
Notify Editor	Editor/Author Email Record 2026-04-30
Editor Version	116764-435159-2-ED.DOC 2026-04-27
Author Version	116764-439539-3-ED.DOC 2026-04-28 DELETE
Upload Author Version	Choose file No file chosen Upload

Decree of Acreditation
New Membership IndoCEISS
Membership Update IndoCEISS
Visitor Statistics

CITATION ANALYSIS

- SCOPUS
- Google Scholar

NOTIFICATIONS

- View (6 new)
- Manage

USER

You are logged in as...
edy_ubd

- My Journals

IN ENGLISH

Title : Evaluating Latent Emotional Structures through Unsupervised Semantic Text Clustering

Revision :

1. **ABSTRACT** :

    - Add 4 mandatory elements: Problem gap (e.g., latent emotions are difficult to capture without labels), Main method, Dataset/domain, and Quantitative results.
    - No mention of semantic representation (e.g., Word2Vec, GloVe, BERT, Sentence Transformer).
    - Lack of mention of evaluation metrics (e.g., Silhouette Score, Davies-Bouldin Index, or topic coherence).

2. **INTRODUCTION :**
    - Explicitly state the research gap.
    - Add state-of-the-art abbreviations (3–5 recent papers).
    - Description and sources of unpublished data.
    - Close with contributions:
      - Semantic clustering framework
      - Evaluation of latent emotion structure
      - Cluster interpretation insights

3. **METHODS**
    - The discussion is still general, lacking a discussion of modern emotional modeling and semantic embedding.

- Apply the following literature to the distribution of ideas that describe the method :
  - Emotion Detection (Supervised)
  - Unsupervised Text Clustering
  - Semantic Embeddings (BERT, SBERT, etc.)
  - Research Gaps
- Structure the idea for writing the research methodology, as follows:
  - Dataset: The data source (e.g., Twitter, Reddit, etc.) has not been clearly stated.
  - Preprocessing: Tokenization, Stopword Removal, and Normalization (adjust accordingly).
  - Text Representation: Word Embeddings (explain why they were chosen).
  - Clustering Method: Explain the algorithm and rationale for the selection.
  - Evaluation: Internal metrics (e.g., Silhouette, DBI) and cluster interpretation (top keywords/terms)
- Explain why unsupervised learning is suitable for latent emotions
- Add a flowchart of the applied approach/method

4. **RESULTS AND DISCUSSION**:
   - Present a table of clustering evaluation results and visualizations (e.g., PCA/t-SNE/UMAP)
   - Add interpretations: Cluster 1 → anger, Cluster 2 → sadness, Cluster 3 → joy (although not labeled). Or provide additional information about which emotion each cluster represents.
   - Compare with different baseline or embedding methods.
   - Discussion: Separate and discuss the following :
     - Implications for computational psychology and social media analysis.
     - Advantages (no labels required).
     - Limitations (subjective interpretation and sensitivity to clustering parameters).

5. **CONCLUSION**

   Additional suggestions for the future (Future work) :
   - Hybrid supervised-unsupervised
   - Deep clustering
   - Multilingual emotion detection

Bukti Korespondensi - IJCCS

Editor/Author Correspondence - Google Chrome

journal.ugm.ac.id/ijccs/author/viewEditorDecisionComments/116764#83975

Subject: [IJCCS] Editor Decision DE

Edy Edy:

We have reached a decision regarding your submission to IJCCS (Indonesian Journal of Computing and Cybernetics Systems), "Evaluating Latent Emotional Structures through Unsupervised Semanti Text Clustering".

Our decision is: Revisions Required

Editor IJCCS
Universitas Gadjah mada
ijccs.mipa@ugm.ac.id

Reviewer A:

Originality:
Fair

Technical quality:
Fair

Importance in its field:
Fair

Style & Overall representation:
Fair

Strong Focus on Targets:
Fair

Editor/Author Correspondence - Google Chrome

journal.ugm.ac.id/ijccs/author/viewEditorDecisionComments/116764#83975

Adequate & Quality of figures:
Fair

Recommendation:
Accepted, revisions required before publishing

Comment:
Title : Evaluating Latent Emotional Structures through Unsupervised Semantic Text Clustering
Revision :
1. ABSTRACT :
• Add 4 mandatory elements: Problem gap (e.g., latent emotions are difficult to capture without labels), Main method, Dataset/domain, and Quantitative results.
• No mention of semantic representation (e.g., Word2Vec, GloVe, BERT, Sentence Transformer).
• Lack of mention of evaluation metrics (e.g., Silhouette Score, Davies-Bouldin Index, or topic coherence).
2. INTRODUCTION :
• Explicitly state the research gap.
• Add state-of-the-art abbreviations (3–5 recent papers).
• Description and sources of unpublished data.
• Close with contributions:
- Semantic clustering framework
- Evaluation of latent emotion structure
- Cluster interpretation insights

3. METHODS
• The discussion is still general, lacking a discussion of modern emotional modeling and semantic embedding.
• Apply the following literature to the distribution of ideas that describe the method :
- Emotion Detection (Supervised)
- Unsupervised Text Clustering
- Semantic Embeddings (BERT, SBERT, etc.)
- Research Gaps

Bukti Korespondensi - IJCCS

• Structure the idea for writing the research methodology, as follows:
- Dataset: The data source (e.g., Twitter, Reddit, etc.) has not been clearly stated.
- Preprocessing: Tokenization, Stopword Removal, and Normalization (adjust accordingly).
- Text Representation: Word Embeddings (explain why they were chosen).
- Clustering Method: Explain the algorithm and rationale for the selection.
- Evaluation: Internal metrics (e.g., Silhouette, DBI) and cluster interpretation (top keywords/terms
• Explain why unsupervised learning is suitable for latent emotions
• Add a flowchart of the applied approach/method

4. RESULTS AND DISCUSSION:
• Present a table of clustering evaluation results and visualizations (e.g., PCA/t-SNE/UMAP)
• Add interpretations: Cluster 1 → anger, Cluster 2 → sadness, Cluster 3 → joy (although not label
Or provide additional information about which emotion each cluster represents.
• Compare with different baseline or embedding methods.
• Discussion: Separate and discuss the following :
- Implications for computational psychology and social media analysis.
- Advantages (no labels required).
- Limitations (subjective interpretation and sensitivity to clustering parameters).

5. CONCLUSION
Additional suggestions for the future (Future work) :
☐ Hybrid supervised-unsupervised
☐ Deep clustering
☐ Multilingual emotion detection

____

IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
http://journal.ugm.ac.id/ijccs
Email: ijccs.mipa@ugm.ac.id
ISSN 1978-1520 (print)
ISSN 2460-7258 (online)

Editor/Author Correspondence - Google Chrome

journal.ugm.ac.id/ijccs/author/viewEditorDecisionComments/116764#83975

Subject: [IJCCS] Editor Decision DE

Edy Edy:

We have reached a decision regarding your submission to IJCCS (Indonesian Journal of Computing and Cybernetics Systems), "Evaluating Latent Emotional Structures through Unsupervised Semanti Text Clustering".

Our decision is: Resubmit

Editor IJCCS
Universitas Gadjah Mada
ijccs.mipa@ugm.ac.id

____

IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
http://journal.ugm.ac.id/ijccs
Email: ijccs.mipa@ugm.ac.id
ISSN 1978-1520 (print)
ISSN 2460-7258 (online)

Subject: [IJCCS] Editor Decision DE

Edy Edy:

We have reached a decision regarding your submission to IJCCS (Indonesian Journal of Computing and Cybernetics Systems), "Evaluating Latent Emotional Structures through Unsupervised Semanti Text Clustering".

Our decision is to: Accept Submission

Editor IJCCS
Universitas Gadjah mada

Bukti Korespondensi - IJCCS

IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Accredited Sinta 2 by the Ministry of Research, Technology and Higher Education of the Republic of Indonesia
Publish jointly by The Department of Computer Science and Electronics, FMIPA UGM and IndoCEISS
ISSN 1978-1520 (print); ISSN 2460-7258 (online)

#116764 Editing

SUMMARY | REVIEW | **EDITING**

Submission

Authors	Edy Edy, Junaedi Junaedi, Aditiya Hermawan, Yusuf Kurnia, Ardiane Rossi Kurniawan Maranto
Title	Evaluating Latent Emotional Structures through Unsupervised Semantic Text Clustering
Section	Articles
Editor	Paholo Prakoso (Editing)

Copyediting

COPYEDIT INSTRUCTIONS

REVIEW METADATA	REQUEST	UNDERWAY	COMPLETE
1. Initial Copyedit	2026-04-30	—	2026-04-30
File: 116764-441052-1-CE.DOC 2026-04-30			
2. Author Copyedit	2026-04-30	2026-05-04	
File: None			
Choose file No file chosen Upload			
3. Final Copyedit	—	—	—
File: None			

Copyedit Comments No Comments

Layout

Galley Format	FILE	
1. PDF VIEW PROOF	116764-441073-2-PB.PDF 2026-05-08	201

Supplementary Files	FILE
	None

Layout Comments No Comments

Editorial Board
Focus & Scope
Ethics & Malpractice Statement
Author Guidelines
Peer Review Process
Reviewer Guidelines
Screening Plagiarism
Online Submission
Author Fee
Statement of Originality
Copyright Transfer Form
Peer Reviewers
The College's Commitment
MoU with IndoCEISS
Decree of Acreditation
New Membership IndoCEISS
Membership Update IndoCEISS
Visitor Statistics

CITATION ANALYSIS

- **SCOPUS**
- **Google Scholar**

Bukti Korespondensi - IJCCS

Proofreading

REVIEW METADATA

		REQUEST	UNDERWAY	COMPLETE
1.	Author	—	—	
2.	Proofreader	—	—	—
3.	Layout Editor	—	—	—

Proofreading Corrections No Comments PROOFING INSTRUCTIONS

NOTIFICATIONS

- View (6 new)
- Manage

USER

You are logged in as...
edy_ubd

- My Journals
- My Profile
- Log Out

DOWNLOAD

IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Accredited Sinta 2 by the Ministry of Research, Technology and Higher Education of the Republic of Indonesia
Publish jointly by The Department of Computer Science and Electronics, FMIPA UGM and IndoCEISS
ISSN 1978-1520 (print); ISSN 2460-7258 (online)

Home | About | User Home | Search | Current | Archives | Announcements | Statistics | Indexing | History | Contact | Journal Help

Home > User > Author > **Archive**

Archive

ACTIVE ARCHIVE

ID	MM-DD SUBMIT	SEC	AUTHORS	TITLE	VIEWS	STATUS
116764	02-12	ART	Edy, Junaedi, Hermawan, Kurnia, Maranto	EVALUATING LATENT EMOTIONAL STRUCTURES THROUGH...	201	Vol 20, No 2 (2026): April

- Editorial Board
- Focus & Scope
- Ethics & Malpractice Statement
- Author Guidelines
- Peer Review Process
- Reviewer Guidelines
- Screening Plagiarism