p-ISSN 2089-9955
e-ISSN 2355-8539
11 (2); 2024; 205-218
KONSELI: Jurnal Bimbingan dan Konseling (E-Journal)

Integrating counseling with technology: An evaluation of the Bicarakan.id application through user review analysis with machine learning

Harun Al Azies[1], Hani Brilianti Rochmanto[2], Cindy Asli Pravesti[2], Fenny Fitriani[2]

[1]Universitas Dian Nuswantoro, Indonesia
[2]Universitas PGRI Adi Buana Surabaya, Indonesia
harun.alazies@dsn.dinus.ac.id

Submitted : 02-11-2024, Revised : 20-11-2024, Accepted : 29-11-2024

Abstract: Online counseling has transformed mental health services by offering a convenient and cost-effective alternative to traditional in-person therapy. This study investigates the role of technology in counseling by analyzing user reviews of the Bicarakan.id app from the Google Play Store. A machine learning approach was employed to identify critical patterns and themes within the reviews. Text pre-processing methods such as tokenization, stop-word removal, and TF-IDF vectorization were applied to a dataset of 125 user reviews. The Elbow method helped determine the optimal number of clusters, which was three. Clustering performance was assessed using the Silhouette score, with three clusters yielding the highest average score of 0.4939, indicating a moderate level of clustering effectiveness. Cluster 1 primarily contained positive reviews, emphasizing user satisfaction with the app's services. Cluster 2 included more specific feedback on users' experiences with counselors and app features. Cluster 3 focused on the app's accessibility and ease of use while raising concerns about data privacy and the lack of offline consultation options. The study underscores the significance of using user feedback to enhance and improve technology-driven mental health solutions.
Keywords: Machine learning; mental health services; online counseling; text clustering; user reviews

Introduction

In recent years, online counseling has significantly reshaped mental health services, offering a convenient and cost-effective alternative to traditional face-to-face therapy (Yamamoto et al., 2021). Technological advancements, mainly through mobile apps and digital platforms, have paved the way for more flexible and accessible counseling services (Lattie et al., 2022). The COVID-19 pandemic, which restricted in-person interactions, further highlighted the importance of online counseling. Research has demonstrated the effectiveness of online counseling in addressing various mental health issues, such as depression and anxiety. For instance, Rahmadiana et al (2021) developed an internet-based intervention for college students struggling with these conditions, yielding promising results. In Indonesia, platforms like Bicarakan.id enhance access to professional counseling, improving mental health outcomes. Integrating technology in mental healthcare presents opportunities and challenges (Figueroa & Aguilera, 2020; Gruber et al., 2021). On one hand, online counseling platforms can reach a broader audience, including those with limited access to traditional services.

On the other hand, the success of these platforms depends heavily on user engagement and the responsiveness of developers to feedback (Q. Chen et al., 2021). User reviews on app stores like Google Play provide valuable insights into users' experiences, satisfaction, and areas of concern, which can guide developers in refining their services (Kaveladze et al., 2022). These reviews often highlight technical issues, unmet needs, and features users appreciate. Analyzing such reviews can offer critical information for enhancing the app's design and functionality.

Copyright © 2024 by All Authors
Correspondence Author: harun.alazies@dsn.dinus.ac.id

This study aims to investigate the role of technology in counseling by analyzing user reviews of the Bicarakan.id application. The research employs a machine learning technique, specifically unsupervised learning with k-means clustering, to uncover recurring themes and patterns in these reviews (Lund & Ma, 2021). Text pre-processing methods, such as tokenization, stop-word removal, and TF-IDF vectorization, were applied to reviews gathered from the Google Play Store (Chai, 2023; Vel, 2021). The Elbow method helped determine the optimal number of clusters, which were then analyzed to identify dominant sentiments, recurring issues, and positive feedback about the app (Hamka & Ramdhoni, 2022; Monica et al., 2021). The k-means clustering method, an unsupervised approach, grouped the reviews into distinct clusters, providing a clearer understanding of user feedback and highlighting strengths and improvement areas (Ashabi et al., 2020). This study ultimately offers detailed insights into the role of user feedback in the ongoing development of technology-driven mental health solutions. Additionally, the findings are expected to provide actionable recommendations for developers and mental health service providers to improve the quality and effectiveness of online counseling, ensuring that user needs and expectations are better met.

Technology integration into mental health services and education has seen significant advancements, mainly through machine learning (ML) and Natural Language Processing (NLP) techniques. Xu et al (2024) studied web-based counseling services with a focus on repeat users, applying hierarchical clustering to group users into three categories based on their behavior and suicide risk. Their findisngs highlighted the importance of tailored interventions to enhance the effectiveness of online mental health services. Rahman et al (2020) reviewed the role of Online Social Networks (OSNs) in mental health detection, emphasizing the value of machine learning and multimethod approaches for early identification and intervention. This study showcased how data-driven social platforms can be crucial in recognizing and addressing mental health concerns (Rahman et al., 2020).

Despite the benefits, online counseling faces several technological challenges, including the need for internet access, devices, and software to conduct sessions (Jena, 2020). These issues are particularly prominent in rural areas, where clients and counselors may need help accessing the necessary technology despite a preference for online counseling over traditional face-to-face options. Nevertheless, studies suggest that online counseling can be as effective as in-person therapy. The accessibility of technology further enhances its effectiveness by providing access to professional services that would otherwise be too costly in person (Amos et al., 2020). Additionally, online counseling removes physical barriers so clients do not need to travel to an office. It makes the support they need more reachable. New strategies, like mobile app-based surveys and social media health initiatives, reflect the growing potential of online mental health services (Ifdil et al., 2020).

Goldberg et al (2020) demonstrated the potential of machine learning and NLP in predicting the therapeutic alliance by analyzing psychotherapy session recordings. This study offered valuable insights into how technology could overcome challenges in therapy. Le Glaz et al (2021) explored trends and techniques in machine learning and NLP within the mental health domain, focusing on symptom extraction, disease severity classification, and therapy comparison. They also highlighted ethical considerations and the need for improvements in

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357

NLP technology to enhance its application in mental health. Chen et al (2020) examined how NLP is used in clinical trial text analysis, employing techniques like performance analysis and topic modeling to visualize research trends and collaborations in clinical (Chen et al., 2020).

Omoregbe et al (2020) evaluated a mental health app by conducting sentiment and thematic analysis of user reviews, identifying both positive and negative aspects, and offering recommendations for improving the app's design. Hiremath & Patil (2022) developed a clinical decision support system using aspect-based sentiment analysis of drug review data, finding that the Support Vector Machine (SVM) algorithm provided the best accuracy in sentiment detection, which could help develop more effective clinical decision tools. Halim et al (2020) introduced a machine-learning framework for emotion recognition in email text, achieving an accuracy of 83%, presenting a new method for identifying emotions from limited text features.

Okoye et al (2022) proposed an EPDM+ML model to analyze student teaching evaluations using text mining and machine learning. This model effectively predicted student recommendations based on sentiments and emotions expressed in feedback. Omoregbe et al (2020) developed a medical diagnosis chatbot using NLP and fuzzy logic to diagnose tropical diseases in Nigeria, with the system showing promising usability scores, indicating its potential for personalized medical diagnoses. Ramos-Lima et al (2020) assessed the use of machine learning techniques in diagnosing trauma-related disorders such as Acute Stress Disorder (ASD) and Post-Traumatic Stress Disorder (PTSD), concluding that while machine learning is beneficial due to the clinical and etiological diversity of patients, its practical application in clinical settings remains a challenge (Ramos-Lima et al., 2020)

Unlike many studies focusing on the broader use of machine learning and NLP in mental health and education, this research explicitly investigates the use of technology in online counseling. This study offers unique insights into how counseling apps can be improved based on real user experiences by employing a clustering approach to analyze user feedback. These findings contribute to developing more adaptable and responsive technology-based mental health solutions, providing a foundation for creating more effective and user-centered counseling applications in the future.

Method

This study employs an exploratory quantitative approach, utilizing text data analysis to uncover hidden patterns and structures through clustering techniques. The primary objective is to classify data based on shared characteristics. This chapter begins with a detailed explanation of the data collection process, followed by a thorough description of the pre-processing steps necessary to prepare the data for analysis. Subsequently, the text transformation process using TF-IDF is outlined, enabling the conversion of textual data into a numerical format suitable for clustering. The chapter then delves into the K-means clustering method, discussing the optimal number of clusters and applying the Manhattan distance metric to enhance clustering accuracy. The entire process is illustrated in a flowchart (Figure 1) to provide a comprehensive overview, clearly depicting the study's methodology.

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357

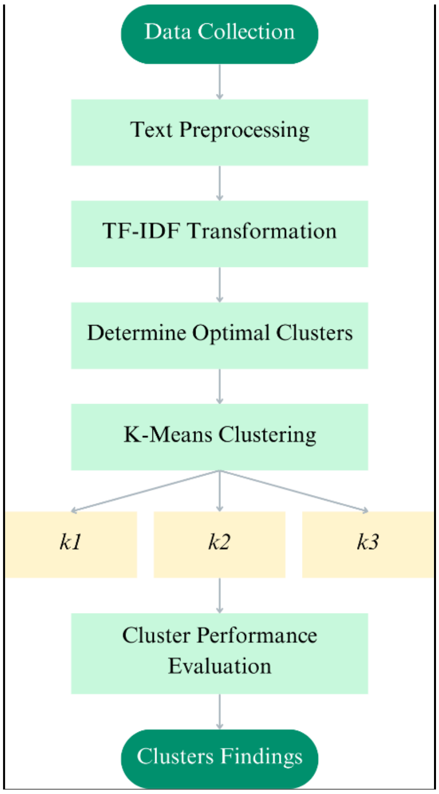

Figure 1. Bicarakan.id User Review Clustering Flowchart

The dataset utilized in this study comprises user comments from the Bicarakan.id online counseling application sourced from the Google Play platform. Data was extracted using the Google-play-scraper module in Google Colaboratory on Monday, July 22, 2024. One hundred twenty-five reviews were gathered, encompassing comment text, user ratings, and posting dates. The choice of Google Play reviews reflects their authenticity, as they capture real user experiences, including satisfaction levels, complaints, suggestions, and expectations regarding the services provided by Bicarakan.id.

Text pre-processing represents a crucial phase that significantly influences the accuracy and effectiveness of NLP models by systematically cleaning and preparing text data for analysis (Siino et al., 2024). This process begins with a cleaning step, which involves removing punctuation, special characters, numerical values, common words (stopwords), and emojis from the review text to eliminate elements that do not contribute to the analysis (Palomino & Aider, 2022). Next, slang words are normalized to standard forms, ensuring clarity and consistency in the extracted tokens (Javed & Kamal, 2018). Case folding is applied to convert all text into lowercase, followed by tokenization, which segments sentences into individual words. Finally, stemming is conducted to reduce inflected words to their root forms (Anreaja et al., 2022). These sequential steps produce a clean and uniform dataset optimized for further analytical processing.

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357

TF-IDF Transformation

Term Frequency-Inverse Document Frequency (TF-IDF) transformation is implemented to quantify the importance of words within a document relative to their frequency across the entire dataset (Cahyani & Patasik, 2021). The formula for TF-IDF is expressed as:

$$w_{ij} = TF_{ij} \times IDF \quad (1)$$

$$IDF = log\left(\frac{N}{df_j}\right) \quad (2)$$

Where w_{ij} denotes the weight of the $j-th$ word in the $i-th$ review, TF_{ij} represents the frequency of the $j-th$ word in the $i-th$ review, N is the total number of reviews, and DF_j indicates the number of reviews containing the $j-th$ word. This transformation enables the representation of textual data in a numeric format, facilitating meaningful analysis in subsequent clustering processes.

K-Means Clustering

K-means clustering is a widely utilized method in data analysis to partition data into distinct groups based on similarity (Oti et al., 2021). This study applied the K-means algorithm to analyze user reviews of the online counseling application Bicarakan.id. The process begins with determining the optimal number of clusters, a critical step to ensure meaningful segmentation. The **Elbow Method** was employed for this purpose, which involves running the K-means algorithm for a range of cluster numbers and plotting the WCSS against the number of clusters (Al Azies et al., 2024). The optimal number of clusters is identified at the "elbow point," where the reduction in WCSS begins to level off, signifying the most appropriate clustering configuration (Naeem & Wumaier, 2018).

The Manhattan Distance Metric was used to calculate the distance between data points and cluster centroids. Also known as the "city block distance", this metric computes the absolute differences across attributes, offering a robust alternative to Euclidean distance by being less sensitive to outliers (Amer & Abdalla, 2020). This feature makes it especially suitable for high-dimensional and sparse text data, where attribute values vary significantly. The formula for Manhattan distance is expressed as:

$$d_{(i,j)} = \sum_{k=1}^{n} \quad |\_ \ y\_k| \quad (3)$$

Here, $d_{i,j}$ represents the distance between data points i and j, with i as the cluster centroid and j as the data point. n denotes the number of attributes or dimensions, while x_k and y_k refer to the $k-th$ attribute values of points x and y, respectively (Faisal & Zamzami, 2020). Once distances are calculated, each data point is assigned to the cluster with the nearest centroid. The centroids are then recalculated based on the mean position of all points within each cluster. This process is repeated iteratively—reassigning points and updating centroids until the centroids stabilize and no significant changes occur in the clustering (Oti et al., 2021). This iterative refinement ensures accurate and meaningful groupings, contributing to a robust dataset analysis.

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357

Silhouette Index

The Silhouette Index is employed in this study to evaluate the effectiveness of the clustering results (Shutaywi & Kachouie, 2021). The Silhouette score for a data point (x_i), denoted as $s(x_i)$, is calculated using the following formula:

$$s(x_i) = \frac{b(x_i) - a(x_i)}{max\{b(x_i), a(x_i)\}} \quad (4)$$

In this equation:

1. $s(x_i)$: the Silhouette score for the ii-th data point.
2. $a(x_i)$: the average distance between the ii-th data point and all other points within the same cluster.
3. $b(x_i)$: the minimum average distance between the ii-th data point and all points in other clusters.

The Silhouette score ranges from -1 to 1:

a. A positive $s(x_i)$ value is achieved when $a(x_i) < b(x_i)$, indicating that the data point is well-clustered.
b. The maximum score of $s(x_i) = 1$ occurs when $a(x_i) = 0$, signifying perfect clustering.
c. Negative values are undesirable, as they occur when $a(x_i) > b(x_i)$, indicating that the dissimilarity within the cluster is greater than the dissimilarity with points in other clusters.

The average $s(x_i)$ across all points in a cluster represents the average Silhouette width for that cluster. Similarly, the mean $s(x_i)$ across all cluster data points provides the average Silhouette width for the entire clustering solution. A higher average Silhouette width indicates better-defined and more cohesive clustering results (Shutaywi & Kachouie, 2021).

Result and Discussion

This chapter presents the process and findings of analyzing user reviews of the Bicarakan.id application utilizing the K-means clustering method. This approach effectively categorized reviews into three primary clusters, each representing distinct themes and focal points. The analytical workflow encompasses several key stages, including text pre-processing, vectorization, identifying the optimal number of clusters, applying the clustering algorithm, and interpreting the clustering outcomes.

Before undergoing analysis, the collected data is subjected to a rigorous pre-processing phase, which involves cleaning, normalizing slang words, converting text to lowercase (case folding), tokenization, removing common words (stopwords), and stemming. This step ensures that the text data is transformed into a structured and consistent format suitable for clustering. A comparative summary of the reviews before and after pre-processing is presented in Table 1, highlighting the transformation of raw text into a more refined and analyzable form.

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357

Table 1 Text Pre-processing Results

Before Pre-processing	After Pre-processing
Thank you. This application is the best and coolest 🏊 🤔	Thank you. This application is the best and incredible.
Thanks to this application, you don't need to take your friends far to the mental hospital.	Thanks to this application, you don't need to take friends to the mental hospital.
This application can reduce crazy people in our environment 😬 👍	This application reduces crazy people in the environment.
For those of you who feel uneasy and stressed, all you have to do is stop by this app; we guarantee it will be beneficial 👍 😬	Feeling uneasy and stressed? Don't worry; visit this application, and you will be helped by professionals.

Text Transformation and Cluster Determination

Following the pre-processing phase, the data is converted into a TF-IDF matrix, where each word in the reviews is assigned a weight based on its frequency and uniqueness across the dataset. This transformation is a critical step in text clustering, as demonstrated in prior research (Wahyuningsih & Chen, 2024) which emphasizes how text vectorization enhances clustering accuracy, particularly with the K-means algorithm. The TF-IDF value reflects the significance of a word in the entire set of reviews; a higher value denotes greater importance, while a value of zero indicates the absence of the word in a specific review.

Determining the optimal number of clusters is a pivotal aspect of cluster analysis, as it directly influences the interpretability and quality of the results. An appropriate number of clusters ensures accurate groupings, where each cluster exhibits internal homogeneity and clear distinctions from other clusters. This study employs the Elbow Method, a widely used technique in cluster analysis to identify the optimal number of clusters.

The Elbow Method involves running the K-means algorithm with varying numbers of clusters and calculating the WCSS for each configuration. For this study, the number of clusters tested ranges from 1 to 10. The WCSS values obtained for each cluster configuration are plotted in Figure 2, with the number of clusters on the x-axis and the corresponding WCSS values on the y-axis. The "elbow point" in the plot indicates the optimal number of clusters, where the rate of WCSS reduction sharply decreases, signifying an ideal balance between cluster compactness and separation.

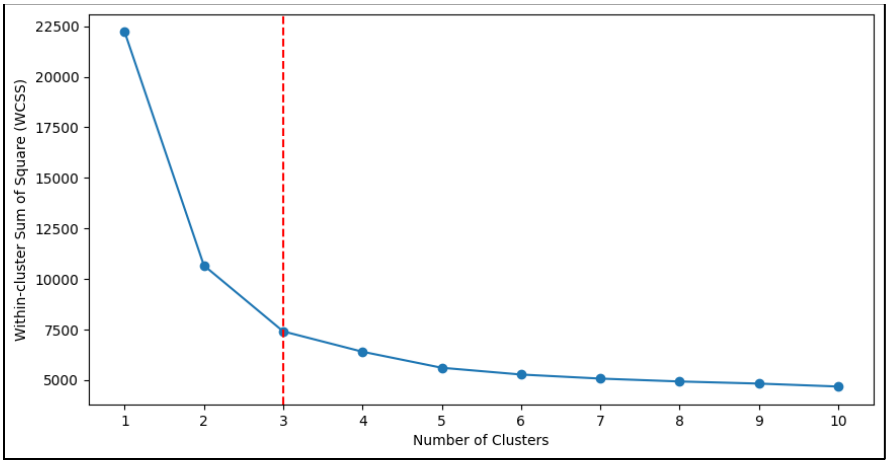

Figure 2. Determining the Optimal Number of Clusters Using the Elbow Method

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357

Determining the Optimal Number of Clusters

The Elbow Point marks the stage where the reduction in the WCSS begins to plateau significantly. At this point, increasing the number of clusters yields minimal improvement in clustering quality. Based on Figure 1, the rate of decline in WCSS noticeably slows after the third cluster, indicating that three clusters represent the optimal configuration for this dataset.

The sharp drop in WCSS values up to the third cluster suggests meaningful reductions in within-cluster variability, while additional clusters beyond this point offer diminishing returns. This conclusion is further supported by the analysis of Silhouette Scores, which provide an alternative measure of cluster quality. The average Silhouette scores for varying numbers of clusters are presented in Table 2, offering critical insights into the cohesion and separation of the clusters formed in this study.

Table 2 Cluster Performance Evaluation Using Silhouette Score

Total Cluster	Average Silhouette Score
$k = 3$	0.494
$k = 4$	0.490
$k = 5$	0.422

Silhouette Score Analysis for Different Cluster Configurations

For the configuration with three clusters, the average Silhouette score is approximately 0.494, indicating moderate cohesion and separation among clusters. This score suggests that most data points are correctly grouped into their respective clusters, though there is room for improvement in cluster clarity. When the number of clusters increases to four, the average Silhouette score decreases slightly to 0.490. This marginal decline indicates that the additional cluster introduces some overlap or ambiguity, slightly reducing the distinctiveness of the groupings.

For five clusters, the average Silhouette score drops significantly to 0.422, signaling a noticeable decline in clustering quality. This reduction suggests increased overlap between clusters and reduced separation, making the groupings less coherent and clear. These findings indicate that increasing the number of clusters beyond three or four can lead to over-segmentation, where data points are unnecessarily divided into too many groups. Such over-segmentation diminishes the overall effectiveness and interpretability of the clustering results. The following sections provide a detailed discussion of the results and characteristics of each cluster.

Figure 3. Word Cloud Cluster 1: Overall Positive Reviews of Bicarakan.id Online Counseling Application

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357

Cluster One: Positive Feedback on User Experience

Cluster one predominantly represents positive feedback from users regarding the Bicarakan.id application. Users in this group expressed satisfaction and gratitude for the services provided, emphasizing that the application effectively facilitated the online counseling process. The experience was described as comfortable and satisfying, aligning with the findings of Mayer et al (2022), which underscore the critical role of user experience in fostering trust in online counseling platforms.

Many users highlighted the application's direct benefits to their mental well-being, such as reduced stress levels and increased self-confidence following counseling sessions. These reviews indicate that the application met users' basic expectations for a practical and effective counseling service.

While some users mentioned minor technical issues, such as logging in or registration challenges, these concerns were relatively independent of their overall positive perception of the platform. The simplicity and clarity of the feedback in this cluster suggest that the application delivered on its promise of ease and effectiveness, providing meaningful support for users' mental health needs.

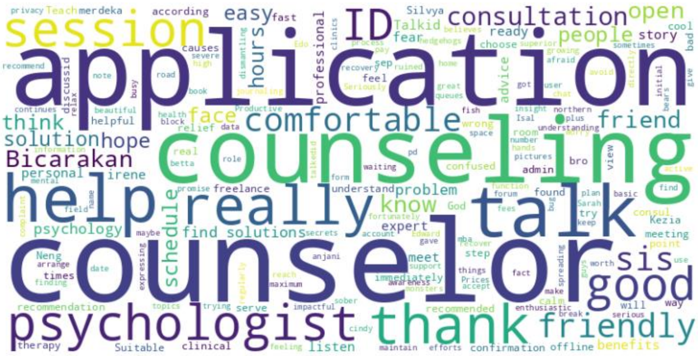

Figure 4. Word Cloud Cluster 2: Recommendations from Users of the Bicarakan.id Online Counseling Application

Cluster Two: Detailed Feedback and Recommendations

The second cluster contains more detailed feedback, often including personal experiences and recommendations related to the Bicarakan.id application. Users in this cluster shared stories about feeling genuinely heard and supported by their counselors, emphasizing how the sessions provided practical solutions to personal challenges.

Positive assessments were frequently directed at the quality of the counselors and app features, such as the journaling tool, which users found particularly helpful. However, alongside these positive comments, some users highlighted issues, such as delays in appointment confirmations and other technical challenges.

This feedback suggests that while users valued the personalized interaction and support they received, they also sought improvements in certain areas to enhance their overall experience. The findings in this cluster align with the research by Fontaine et al (2020), which highlights that personalization and direct engagement with counselors are critical factors in boosting user satisfaction and engagement in online counseling platforms. Cluster two appreciates the application's core features while calling for refinements to improve service delivery.

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357

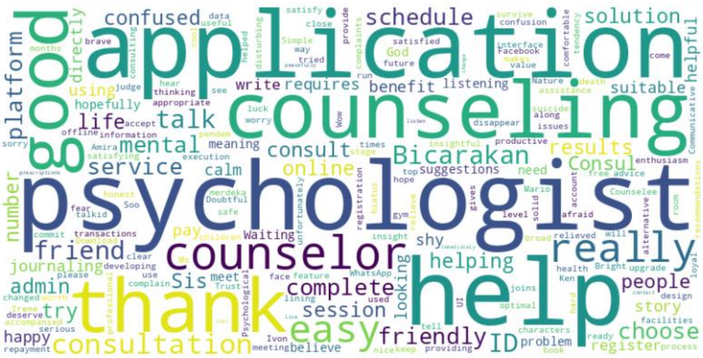

Figure 5. Word Cloud Cluster 3: Review of Features and Usefulness of Bicarakan.id Online Counseling Application

Cluster Three: Features and Usability

The third cluster emphasizes feedback centered on the features and usability of the Bicarakan.id counseling app. Users commended the application's attractive interface design, ease of use, and specific features, such as the journaling tool, which enhanced their overall experience. The positive feedback about the app's interface design and ease of use aligns with established principles of human-computer interaction, where intuitive and aesthetically pleasing designs enhance user satisfaction (Kheder, 2023). Many users highlighted how it streamlined the scheduling and execution of counseling sessions, particularly benefiting those who felt awkward or uncomfortable with in-person counseling.

While the feedback could have been more positive, there were recurring concerns about registration difficulties and pricing. These issues indicate that, although the application effectively delivers its core services, improvements in accessibility and affordability could broaden its appeal. The findings in this cluster are consistent with (Xie et al., 2023), who emphasize that usability and accessibility are critical factors for the success of digital platforms. Users in this cluster valued the application's functionality and convenience, providing a more seamless counseling experience.

Conclusions and Recommendations

The Bicarakan.id online counseling application has received predominantly positive feedback from users on the Google Play Store. Through the K-Means clustering method, three distinct feedback clusters were identified. Cluster One, highlighted overall user satisfaction, with reviews focusing on the positive impact of the application on users' mental well-being. Cluster Two, contained more specific feedback regarding personal experiences with counselors and application features, alongside critiques related to appointment scheduling and technical issues. Cluster Three, focused on the ease of access to mental health services and the user-friendly application interface, concerns about data privacy and the need for offline consultation options were noted.

The findings indicate that the application effectively meets users' needs for online counseling, providing valuable support and a satisfactory experience. Nonetheless, there are opportunities for improvement in technical and operational aspects to elevate the user experience further. Suggested improvements include addressing registration challenges to

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357

enhance accessibility, reducing delays in appointment confirmations to streamline processes, and reviewing service fees to make the platform more affordable and inclusive.

The results of this analysis also underscore the potential for counselors to adopt online counseling applications like Bicarakan.id as part of their services. This technology enables flexible service delivery, particularly for clients or students who face obstacles in accessing face-to-face counseling.

Critical recommendations for maximizing the effectiveness of online counseling platforms include training counselors in digital competencies used to optimize their use of the technology; leveraging features like personal journaling tools for emotional reflection and progress tracking; and maintaining empathetic and responsive communication, both online and offline, to foster trust and engagement with users. By addressing the identified areas for improvement and effectively embracing digital tools, the application and its users can experience enhanced outcomes, further solidifying the role of technology in modern mental health support.

.

References

Abd Rahman, R., Omar, K., Noah, S. A. M., Danuri, M. S. N. M., & Al-Garadi, M. A. (2020). Application of machine learning methods in mental health detection: A systematic review. *Ieee Access*, *8*, 183952–183964.

Al Azies, H., Rohmatullah, F. A., Rochmanto, H. B., & Isnarwaty, D. P. (2024). Towards optimization: A data-driven approach using k-medoids clustering algorithm for regional education quality assessment. *Jurnal Informatika Dan Teknik Elektro Terapan*, *12*(3).

Amer, A. A., & Abdalla, H. I. (2020). A set theory based similarity measure for text clustering and classification. *Journal of Big Data*, *7*(1), 74.

Amos, P. M., Bedu-Addo, P. K. A., & Antwi, T. (2020). Experiences of online counseling among undergraduates in some Ghanaian universities. *Sage Open*, *10*(3), 2158244020941844.

Ashabi, A., Sahibuddin, S. Bin, & Salkhordeh Haghighi, M. (2020). The systematic review of K-means clustering algorithm. *Proceedings of the 2020 9th International Conference on Networks, Communication and Computing*, 13–18.

Cahyani, D. E., & Patasik, I. (2021). Performance comparison of tf-idf and word2vec models for emotion text classification. *Bulletin of Electrical Engineering and Informatics*, *10*(5), 2780–2788.

Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural Language Engineering*, *29*(3), 509–553.

Chen, Q., Chen, C., Hassan, S., Xing, Z., Xia, X., & Hassan, A. E. (2021). How should i improve the ui of my app? a study of user reviews of popular apps in the google play. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, *30*(3), 1–38.

Chen, X., Xie, H., Cheng, G., Poon, L. K. M., Leng, M., & Wang, F. L. (2020). Trends and features of the applications of natural language processing techniques for clinical trials text analysis. *Applied Sciences*, *10*(6), 2157.

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357

Faisal, M., & Zamzami, E. M. (2020). Comparative analysis of inter-centroid K-Means performance using euclidean distance, canberra distance and manhattan distance. *Journal of Physics: Conference Series*, *1566*(1), 012112.

Figueroa, C. A., & Aguilera, A. (2020). The need for a mental health technology revolution in the COVID-19 pandemic. *Frontiers in Psychiatry*, *11*, 523.

Fontaine, G., Cossette, S., Gagnon, M.-P., Dubé, V., & Côté, J. (2020). Effectiveness of a theory-and web-based adaptive implementation intervention on nurses' and nursing students' intentions to provide brief counseling: protocol for a randomized controlled trial. *JMIR Research Protocols*, *9*(7), e18894.

Goldberg, S. B., Flemotomos, N., Martinez, V. R., Tanana, M. J., Kuo, P. B., Pace, B. T., Villatte, J. L., Georgiou, P. G., Van Epps, J., & Imel, Z. E. (2020). Machine learning and natural language processing in psychotherapy research: Alliance as example use case. *Journal of Counseling Psychology*, *67*(4), 438.

Gruber, J., Prinstein, M. J., Clark, L. A., Rottenberg, J., Abramowitz, J. S., Albano, A. M., Aldao, A., Borelli, J. L., Chung, T., & Davila, J. (2021). Mental health and clinical psychological science in the time of COVID-19: Challenges, opportunities, and a call to action. *American Psychologist*, *76*(3), 409.

Halim, Z., Waqar, M., & Tahir, M. (2020). A machine learning-based investigation utilizing the in-text features for the identification of dominant emotion in an email. *Knowledge-Based Systems*, *208*, 106443.

Hamka, M., & Ramdhoni, N. (2022). K-means cluster optimization for potentiality student grouping using elbow method. *AIP Conference Proceedings*, *2578*(1).

Hiremath, B. N., & Patil, M. M. (2022). Enhancing optimized personalized therapy in clinical decision support system using natural language processing. *Journal of King Saud University-Computer and Information Sciences*, *34*(6), 2840–2848.

Ifdil, I., Fadli, R. P., Suranata, K., Zola, N., & Ardi, Z. (2020). Online mental health services in Indonesia during the COVID-19 outbreak. *Asian Journal of Psychiatry*, *51*, 102153.

Javed, M., & Kamal, S. (2018). Normalization of unstructured and informal text in sentiment analysis. *International Journal of Advanced Computer Science and Applications*, *9*(10).

Jena, P. K. (2020). Challenges and Opportunities created by Covid-19 for ODL: A case study of IGNOU. *International Journal for Innovative Research in Multidisciplinary Field (IJIRMF)*, *6*.

Kaveladze, B. T., Wasil, A. R., Bunyi, J. B., Ramirez, V., & Schueller, S. M. (2022). User experience, engagement, and popularity in mental health apps: secondary analysis of app analytics and expert app reviews. *JMIR Human Factors*, *9*(1), e30766.

Kheder, H. A. (2023). Human-Computer Interaction: Enhancing User Experience in Interactive Systems. *Kufa Journal of Engineering*, *14*(4), 23–41.

Lattie, E. G., Stiles-Shields, C., & Graham, A. K. (2022). An overview of and recommendations for more accessible digital mental health services. *Nature Reviews Psychology*, *1*(2), 87–100.

Le Glaz, A., Haralambous, Y., Kim-Dufor, D.-H., Lenca, P., Billot, R., Ryan, T. C., Marsh, J., Devylder, J., Walter, M., & Berrouiguet, S. (2021). Machine learning and natural language processing in mental health: systematic review. *Journal of Medical Internet Research*, *23*(5), e15708.

Lund, B., & Ma, J. (2021). A review of cluster analysis techniques and their uses in library and information science research: k-means and k-medoids clustering. *Performance Measurement and Metrics*, *22*(3), 161–173.

Mayer, G., Hummel, S., Oetjen, N., Gronewold, N., Bubolz, S., Blankenhagel, K., Slawik, M., Zarnekow, R., Hilbel, T., & Schultz, J.-H. (2022). User experience and acceptance of patients and healthy adults testing a personalized self-management app for depression: a non-randomized mixed-methods feasibility study. *Digital Health*, *8*, 20552076221091350.

Monica, M., Ayuningtiyas, N. U., Al Azies, H., Riefky, M., Khusna, H., & Rahayu, S. P. (2021). Unsupervised learning approach for evaluating the impact of COVID-19 on economic growth in Indonesia. *Soft Computing in Data Science: 6th International Conference, SCDS 2021, Virtual Event, November 2–3, 2021, Proceedings 6*, 54–70.

Naeem, S., & Wumaier, A. (2018). Study and implementing K-mean clustering algorithm on English text and techniques to find the optimal value of K. *International Journal of Computer Applications*, *182*(31), 7–14.

Okoye, K., Arrona-Palacios, A., Camacho-Zuñiga, C., Achem, J. A. G., Escamilla, J., & Hosseini, S. (2022). Towards teaching analytics: a contextual model for analysis of students' evaluation of teaching through text mining and machine learning classification. *Education and Information Technologies*, 1–43.

Omoregbe, N. A. I., Ndaman, I. O., Misra, S., Abayomi-Alli, O. O., & Damaševičius, R. (2020). Text Messaging-Based Medical Diagnosis Using Natural Language Processing and Fuzzy Logic. *Journal of Healthcare Engineering*, *2020*(1), 8839524.

Oti, E. U., Olusola, M. O., Eze, F. C., & Enogwe, S. U. (2021). Comprehensive review of K-Means clustering algorithms. *Criterion*, *12*, 22–23.

Palomino, M. A., & Aider, F. (2022). Evaluating the effectiveness of text pre-processing in sentiment analysis. *Applied Sciences*, *12*(17), 8765.

Rahmadiana, M., Karyotaki, E., Schulte, M., Ebert, D. D., Passchier, J., Cuijpers, P., Berger, T., van Ballegooijen, W., Wimbarti, S., & Riper, H. (2021). Transdiagnostic internet intervention for Indonesian university students with depression and anxiety: evaluation of feasibility and acceptability. *JMIR Mental Health*, *8*(3), e20036.

Ramos-Lima, L. F., Waikamp, V., Antonelli-Salgado, T., Passos, I. C., & Freitas, L. H. M. (2020). The use of machine learning techniques in trauma-related disorders: a systematic review. *Journal of Psychiatric Research*, *121*, 159–172.

Robledo Yamamoto, F., Voida, A., & Voida, S. (2021). From therapy to teletherapy: Relocating mental health services online. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–30.

Shutaywi, M., & Kachouie, N. N. (2021). Silhouette analysis for performance evaluation in machine learning with applications to clustering. *Entropy*, *23*(6), 759.

Siino, M., Tinnirello, I., & La Cascia, M. (2024). Is text preprocessing still worth the time? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers. *Information Systems*, *121*, 102342.

Vel, S. S. (2021). Pre-processing techniques of text mining using computational linguistics and

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357

python libraries. *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, 879–884.

Wahyuningsih, T., & Chen, S. C. (2024). Analyzing Sentiment Trends and Patterns in Bitcoin-Related Tweets Using TF-IDF Vectorization and K-Means Clustering. *Journal of Current Research in Blockchain*, *1*(1), 48–69.

Xie, I., Lee, T. H., Lee, H. S., Wang, S., & Babu, R. (2023). Comparison of accessibility and usability of digital libraries in mobile platforms: blind and visually impaired users' assessment. *Information Research an International Electronic Journal*, *28*(3), 59–82.

Xu, Y., Chan, C. S., Chan, E., Chen, J., Cheung, F., Xu, Z., Liu, J., & Yip, P. S. F. (2024). Tracking and profiling repeated users over time in text-based counseling: Longitudinal observational study with hierarchical clustering. *Journal of Medical Internet Research*, *26*, e50976.

https://ejournal.radenintan.ac.id/index.php/konseli
https://doi.org/10.24042/kons.v11i2.24357