

Exploring Audience Perspectives on Self-Development and Career Growth Through YouTube Comments: Evidence from the SUARA BERKELAS Podcast

Original Article

Andi Nurfiana Haz^{1*}, Chairunnisaa M Z Huzaen²

^{1,2}Department of Management, Faculty Economics and Business, Universitas Negeri Makassar

Email: ¹⁾ andinurfiahaz@unm.ac.id, ²⁾ chairunnisaa@unm.ac.id

Received : 25 March 2026

Accepted : 31 March 2026

Published online : 21 April 2026

Abstract

The growth of digital media has transformed the way individuals access information, share experiences, and engage in discussions related to self-development and career growth. Among various digital platforms, YouTube podcasts have emerged as popular sources of informal learning and personal reflection. This study explored audience perspectives on self-development and career growth through comments posted on the SUARA BERKELAS podcast on YouTube. The data consisted of 704 audience comments collected using a web scraping technique. Text preprocessing was conducted through case folding, cleaning, tokenization, and stopword removal. Subsequently, bigram analysis and Latent Dirichlet Allocation (LDA) topic modeling were applied to identify dominant discussion themes. The findings reveal five major themes: social reflection and national issues, life aspirations and future goals, youth self-development, motivation and life struggles, and concerns about the future. These themes indicate that audiences perceive the podcast not only as a source of entertainment but also as a medium for learning, self-reflection, and career development. The coherence score of 0.44 suggests that the topic model demonstrates acceptable consistency in representing audience discussions. This study highlights the role of digital podcasts in shaping audience perspectives on personal growth and career-related issues in contemporary society.

Keywords: self-development, career development, YouTube comments, topic modeling, digital media

1. Introduction

The development of human resources has become increasingly important in today's rapidly changing social and professional environment. Individuals are expected to continuously improve their knowledge, skills, and competencies to remain competitive and adapt to evolving challenges (Bandura, 2001). From a human capital perspective, investments in self-development and career growth contribute to enhancing individual capabilities and future opportunities. Therefore, access to information, learning resources, and personal development opportunities has become an essential aspect of contemporary human resource development (Altalhi, 2021).

The rapid growth of digital media has transformed the way individuals access information, acquire knowledge, and develop new competencies. Digital platforms are no longer used solely for entertainment purposes but have also become important sources of learning, motivation, and career-related information. Among various digital platforms, YouTube has emerged as one of the most widely used channels for educational content,



allowing users not only to consume information but also to participate in discussions through interactive features such as comment sections (Elprida et al., 2026). (Manning et al., 2008)

One form of digital content that has gained increasing popularity is podcast-based video content on YouTube. Podcasts provide opportunities for audiences to engage with discussions related to education, personal growth, career development, and social issues. Unlike traditional media, YouTube podcasts encourage audience participation through comments, enabling viewers to express opinions, share experiences, and reflect on the topics being discussed. As a result, audience comments have become a valuable form of user-generated content that can reveal how individuals interpret issues related to self-development and career growth (Chunqiong et al., 2025a).

One YouTube podcast that has attracted considerable audience attention is the *SUARA BERKELAS* podcast. The episode featuring Jerome Polin accumulated approximately 828,000 views at the time of data collection, indicating substantial audience interest in discussions related to education, self-development, and future opportunities. The high level of engagement reflected in the comment section suggests that viewers actively respond to topics associated with personal competencies, career aspirations, and life challenges. These comments provide a valuable source of information for understanding how audiences perceive self-development and career growth in contemporary society (Chunqiong et al., 2025a).

Previous studies have extensively examined social media content through sentiment analysis, opinion mining, and audience engagement approaches. However, research exploring how audiences interpret self-development and career-related issues through YouTube podcast discussions remains relatively limited. Existing studies often focus on emotional polarity rather than identifying the underlying themes and meanings expressed by audiences. Consequently, there is a need to further explore audience discussions to understand how individuals construct perspectives on self-development and career growth through digital media interactions (“Podcasting in the 21st Century,” n.d.).

Therefore, this study aims to explore audience perspectives on self-development and career growth through comments posted on the *SUARA BERKELAS* podcast on YouTube. Using a text mining approach and Latent Dirichlet Allocation (LDA) topic modeling, this study identifies dominant themes emerging from audience discussions. The findings are expected to contribute to the literature on human resource development, digital media, and audience engagement by providing insights into how digital audiences perceive self-development and career-related issues in the contemporary digital era (Altalhi, 2021).

2. Literature Review

2.1. Self Development

Self-development is a process of enhancing an individual's abilities, potential, and quality of life in order to achieve a better life. From a human resource perspective, self-development is associated with an individual's efforts to improve competencies, mindset, and adaptability to environmental changes. Self-development is not only acquired through formal education but can also be achieved through social experiences and digital media. (Greenhow & Lewin, 2016a)

The development of digital media has made it easier for people to access information and engage in self-directed learning. Digital content such as podcasts, educational videos, and social media has become one of the informal means through which individuals gain motivation, insights, and knowledge related to life and career development. This indicates that digital media plays a role in supporting the self-development process of society, particularly among younger generations.

2.2. Career Development

Career development is the process of enhancing an individual's capabilities to achieve job-related goals and future career aspirations. In the field of human resource management, career development is associated with career planning, competency improvement, and an individual's readiness to adapt to changes in the world of work. Today's younger generation faces increasingly complex career challenges due to technological advancements and the dynamic nature of the work environment. (Chunqiong et al., 2025b)

Digital media has become one of the sources of information that influences how individuals understand the world of work and their future career prospects. Content discussing motivation, life experiences, and career journeys often serves as a reference for audiences in shaping their perspectives on employment and future life. Therefore, audience interactions within digital media can represent aspirations, anxieties, and expectations related to career development.

2.3. Previous Research

Digital media enables individuals not only to act as consumers of information but also as content producers through various forms of interaction, one of which is commenting on social media platforms and YouTube. Audience comments represent a form of user-generated content that reflects public opinions, experiences, and perceptions regarding a particular piece of content (Saputra et al., 2025)

Comments on digital platforms can be utilized as a valuable data source for understanding communication patterns and identifying dominant themes in public discourse. In the context of YouTube podcasts, audience comments can reveal how specific topics are

understood, interpreted, and responded to by the public. Therefore, the analysis of digital comments is highly relevant for exploring audience perceptions of self-development and career-related issues.

2.4. Text Mining and Modeling

Text mining is a text data analysis method used to identify patterns, word relationships, and specific information from large collections of unstructured data. In digital media research, text mining is widely employed to analyze social media comments, user reviews, and public opinions. (Maier et al., 2018a)

One of the commonly used techniques in text mining is topic modeling through the Latent Dirichlet Allocation (LDA) method. This approach is utilized to identify clusters of topics based on the occurrence of dominant words within a text corpus. Topic modeling enables researchers to automatically discover underlying themes without the need for manual labeling.

In this study, topic modeling was applied to identify the dominant themes emerging from audience comments on the *SUARA BERKELAS* podcast. This approach provides insights into audience perceptions regarding self-development and career-related

3. Methods

This study employed an exploratory qualitative approach using text mining techniques to analyze audience comments on the *SUARA BERKELAS* podcast on YouTube. The overall research procedure is presented in Figure 1.

3.1. Research Design

This study employed an exploratory qualitative approach using text mining techniques to analyze audience comments on the *SUARA BERKELAS* podcast on YouTube. The overall research procedure is presented in Figure 1.

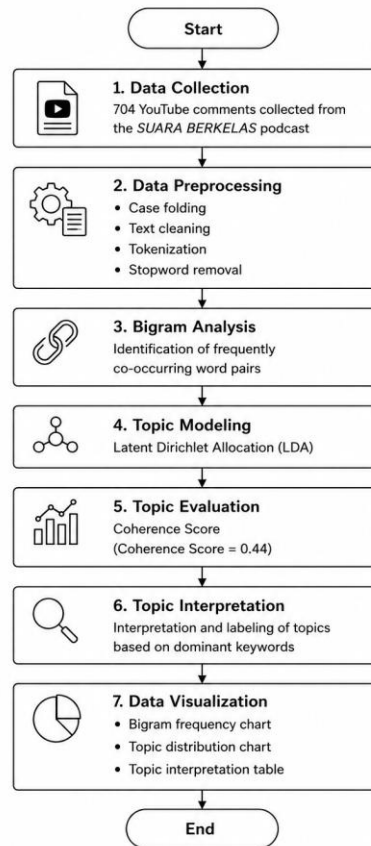


Figure 1. Research Framework

The data consisted of 704 comments collected from the *SUARA BERKELAS* podcast on YouTube through a web scraping technique using Python and Google Colab. The collected comments were used as the primary dataset because they reflected audience opinions, experiences, and perspectives regarding self-development and career-related issues.

3.2. Data Preprocessing and Topic Modeling Procedure

Following data collection, a preprocessing stage was conducted to improve text quality before analysis. The preprocessing process included case folding, text cleaning, tokenization, and stopwords removal. These procedures were performed to eliminate irrelevant characters and reduce textual noise in the dataset.

After preprocessing, bigram analysis was applied to identify pairs of words that frequently appeared together in audience comments. This analysis provided an initial overview of recurring concepts and discussion patterns related to self-development and career growth.

Subsequently, topic modeling was performed using the Latent Dirichlet Allocation (LDA) method to identify dominant themes within the comments. The resulting topics were interpreted and labeled according to their dominant keywords and contextual meanings (Anonim, 2026).

To evaluate topic quality, coherence score analysis was employed. The model achieved a coherence score of 0.44, indicating an acceptable level of topic consistency and interpretability. Finally, the findings were visualized through bigram frequency charts, topic distribution graphs, and topic interpretation tables to facilitate the analysis and discussion of audience perspectives (Mimno et al., 2011; Rosner et al., 2014; Tess, 2015a).

4. Results and Discussion

4.1. Data Collection

A total of 704 comments were collected from the *SUARA BERKELAS* podcast on YouTube using a web scraping technique. The data collection process was conducted on May 25, 2026, using Google Colab and the *youtube-comment-downloader* library. The collected comments served as the primary dataset for analysis because they represented audience opinions and experiences related to self-development and career-related issues discussed in the podcast. Figure 2 presents a sample of the collected comments, including comment text, author information, number of likes, and publication time

	text	author	likes	published
0	Kata kata jerom di akhir vidio ngena ke gua lagi 🤔	@wildan-n3r	0	2 days ago
1	Bisa ngak sih kementerian liat2 gituu saran2 p...	@Uswacuaa	0	2 days ago
2	hehe	@romina6907	0	4 days ago
3	Salam inspirasi sukses selalu suka \n\nAaaadd...	@AhmedOmr-t2g	0	6 days ago
4	ini beneran sih, banyak orang yg mau jadi guru...	@bangyarzz777	0	6 days ago

Figure 2. Sample of Collected YouTube Comments Dataset

4.2. Data Preprocessing

Before conducting the text mining analysis, the collected comments were processed through several preprocessing stages to improve data quality and reduce textual noise. The preprocessing procedures included case folding, text cleaning, tokenization, and stopword removal. These steps were performed to standardize the textual data and ensure that only relevant words were retained for further analysis.

index	text	clean_comment
0	Kata kata jerom di akhir vidio ngena ke gua lagi 🤔	kata kata jerom akhir vidio ngena gua
1	Bisa ngak sih kementerian liat2 gituu saran2 podcast kek gini	ngak sih kementerian liat gituu saran podcast kek gini
2	hehe	hehe
3	Salam inspirasi sukses selalu suka Aaaadduuuuuhhhh gemes anak muda yg selalu ingin bertumbuh netizen menyimak sukses selalu yah 🤔🤔🤔🤔🤔🤔🤔🤔🤔	salam inspirasi sukses selalu suka aaaadduuuuuhhhh gemes anak muda yg selalu bertumbuh netizen menyimak sukses selalu yah
4	ini beneran sih, banyak orang yg mau jadi guru cuman gajinya kecil, terus orang bilang lagi kenapa kalau gaji kecil ga ikhlas berarti ngajarnya? bukan ga ikhlas cuman ga make sense guru capek belajar nyusun ini itu tapi ketika pulang harus urusin kontrakan rumah, kpr rumah biaya anak yg gajinya si guru itu mungkin sekitaran 3jtan, jadi menurut gw tingkatin dulu kualitas guru dengan menaikkan gajinya yg sangat tinggi, karena kalau tinggi guru" pada semangat dan akan bisa fokus ke ngajar murid jadi dia ga perlu lagi ngurusin berlebihan tentang sewa rumah, kontrakan, kpr, dll karena dia udah nyaman disitu	beneran sih banyak orang yg mau jadi guru cuman gajinya kecil terus orang bilang kalau gaji kecil ga ikhlas berarti ngajarnya bukan ga ikhlas cuman ga make sense guru capek belajar nyusun pulang urusin kontrakan rumah kpr rumah biaya anak yg gajinya si guru mungkin sekitaran jtan jadi gw tingkatin dulu kualitas guru menaikkan gajinya yg sangat tinggi kalau tinggi guru semangat fokus ngajar murid jadi ga perlu ngurusin berlebihan sewa rumah kontrakan kpr udah nyaman disitu

Figure 3. Example of Text Preprocessing Result

Figure 3 illustrates the transformation of raw comments into a cleaner textual format. The preprocessing stage successfully removed unnecessary characters, symbols, and irrelevant expressions while preserving the semantic meaning of the comments. As a result, the processed data became more suitable for bigram analysis and topic modelling.

Figure 5 shows that words such as *orang*, *sendiri*, *nonton*, *pemerintah*, *diri*, *bangsa*, *DPR*, and *mimpi* appeared frequently in audience discussions. These terms suggest that viewers discussed not only personal development and self-reflection but also broader social and institutional issues. The word cloud provides an initial overview of audience discussions, while a more detailed thematic interpretation is presented through topic modeling in the following section.

4.5. Topic Modelling Result

To identify the dominant themes discussed by audiences, topic modeling was performed using the Latent Dirichlet Allocation (LDA) method. The analysis grouped audience comments into several latent topics based on patterns of word co-occurrence. The resulting topic distribution is presented in Figure 6.

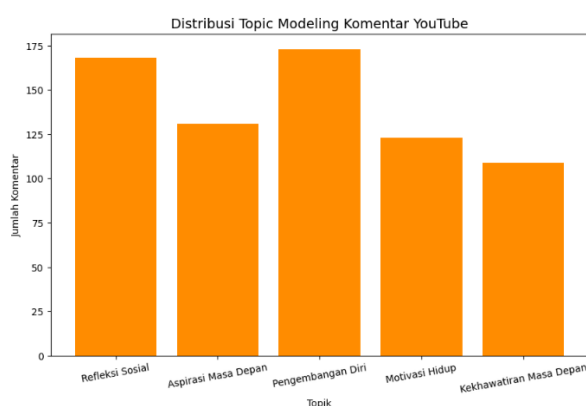


Figure 6. Topic Distribution of Audience Comments

Figure 6 shows that audience discussions were distributed across several major themes. The results indicate that viewers discussed not only self-development and career-related issues but also broader social, educational, and motivational topics. This finding suggests that the podcast generated diverse audience responses and encouraged discussions extending beyond the primary content of the video.

4.6. Topic Interpretation

To facilitate thematic interpretation, the topics generated by the LDA model were examined based on their dominant keywords and contextual meanings. The interpretation process resulted in five major themes representing audience perspectives on self-development, career growth, and broader social issues. The topic interpretation results are presented in Table 1.

Table 1. Topic Interpretations Result

Topik	Kata Dominan	Interpretasi
Topik 1	presiden, masalah, dpr, paham	Refleksi sosial dan isu nasional
Topik 2	luar, mimpi, fokus, pemerintah	Aspirasi hidup dan masa depan

Topik	Kata Dominan	Interpretasi
Topik 3	diri, dunia, sistem, muda	Pengembangan diri generasi muda
Topik 4	semangat, uang, diri, akhirnya	Motivasi dan perjuangan hidup
Topik 5	bermimpi, pintar, tinggi, kecil	Kekhawatiran terhadap masa depan

Topic 1: Social Reflection and National Issues

The first topic contained keywords such as *presiden*, *masalah*, *DPR*, and *paham*. This topic reflects audience discussions regarding social conditions, governance, and national issues. Many viewers associated the podcast content with broader societal challenges and expressed concerns about public policies and institutional performance. The findings suggest that audience discussions extended beyond personal development and frequently addressed issues affecting society as a whole.

Topic 2: Life Aspirations and Future Goals

The second topic included keywords such as *luar*, *mimpi*, *fokus*, and *pemerintah*. This theme represents audience aspirations regarding education, career opportunities, and future success. Many comments reflected hopes for achieving a better quality of life and overcoming challenges through continuous learning and personal improvement.

Topic 3: Youth Self-Development

The third topic consisted of keywords such as *diri*, *dunia*, *sistem*, and *muda*. This topic highlights discussions related to self-development among younger generations. Audience comments emphasized the importance of improving personal capabilities, adapting to changing environments, and preparing for future challenges in both personal and professional contexts.

Topic 4: Motivation and Life Struggles

The fourth topic contained keywords such as *semangat*, *uang*, *diri*, and *akhirnya*. This theme reflects audience discussions about perseverance, motivation, and personal struggles. Many viewers shared experiences related to overcoming obstacles, maintaining optimism, and pursuing long-term goals despite various difficulties.

Topic 5: Concerns About the Future

The fifth topic included keywords such as *bermimpi*, *pinter*, *tinggi*, and *kecil*. This topic illustrates audience concerns and uncertainties regarding future opportunities and personal achievement. The comments suggest that many viewers were aware of the challenges associated with achieving success and expressed both optimism and anxiety regarding their future prospects.

To evaluate the quality of the generated topics, coherence score analysis was conducted. The model achieved a coherence score of 0.44, indicating an acceptable level of semantic consistency and interpretability. This result suggests that the identified topics are sufficiently representative of audience discussions and can be used to explain audience perspectives on self-development and career growth.

4.7. Discussion

The findings indicate that audiences perceive the *SUARA BERKELAS* podcast as more than a source of entertainment. The dominant themes identified through topic modeling suggest that viewers actively engage in discussions related to self-development, future aspirations, motivation, and career-related challenges (Maier et al., 2018b). This finding demonstrates that YouTube podcasts can function as spaces for reflection and informal learning, where audiences exchange perspectives on both personal growth and societal realities. (Tess, 2015b).

One notable finding is the prominence of the *Youth Self-Development* theme, which recorded the highest number of comments among all identified topics. This result suggests that audiences are highly interested in discussions related to personal growth, continuous learning, and future preparation. The emergence of themes such as *Life Aspirations and Future Goals* and *Motivation and Life Struggles* further reinforces the importance of self-improvement and long-term planning in audience discussions. Many viewers viewed success as a process that requires persistence, learning, and personal commitment. (Service, 2024).

At the same time, the strong presence of the *Social Reflection and National Issues* theme indicates that audiences frequently connected self-development with broader social conditions. Discussions involving social issues, public concerns, and national realities suggest that viewers do not perceive personal development as an entirely individual process. Instead, many comments reflected the belief that educational opportunities, social environments, and external circumstances may influence individual growth and future prospects.

In addition to personal development themes, audience discussions also reflected awareness of broader social conditions that may influence individual opportunities and career growth (Greenhow & Lewin, 2016b). The appearance of themes related to social reflection suggests that audiences recognize the importance of external environments in supporting personal development. However, the dominant focus of audience discussions remained on motivation, self-improvement, and future aspirations rather than on social criticism itself (Morgeson & Humphrey, 2014).

Overall, the findings demonstrate that the *SUARA BERKELAS* podcast serves as a platform for learning, reflection, and audience engagement. Through the comment section, viewers shared perspectives on self-development, future aspirations, and social realities,

highlighting the growing role of digital media in supporting human capital development and encouraging meaningful public discussion in contemporary society (Jarman, 2018).

5. Conclusion

This study explored audience perspectives on self-development and career growth through comments on the *SUARA BERKELAS* podcast on YouTube using a topic modeling approach. The findings reveal that audience discussions were dominated by themes related to social reflection, future aspirations, self-development, motivation, and concerns about the future. These results indicate that audiences perceive the podcast not only as a source of entertainment but also as a medium for learning, reflection, and personal growth. The study suggests that content creators can further develop educational and motivational content that encourages meaningful audience engagement, while future researchers may examine similar discussions across different digital platforms and podcast genres to gain a broader understanding of audience perspectives in the digital era.

6. References

- Altalhi, O. A. (2021). A Systematic Literature Review of the Impact of Human Resources Management Strategies on Employee's Performance: Role of Job Embeddedness and Organizational Commitment as Mediator Variables. *European Journal of Business and Management*, 13(20), 56–61. <https://www.iiste.org>
- Anonim. (2026). Audience Perspectives on Self-Development and Career Growth: A Text Mining, Bigram, and Latent Dirichlet Allocation (LDA) Topic Modeling on the SUARA BERKELAS Podcast on YouTube. *Draft Penelitian / Jurnal Mandiri*, 1–10.
- Bandura, A. (2001). Social Cognitive Theory: An Agentic Perspective. *Annual Review of Psychology*, 52(1), 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Chunqiong, W., Shan, J., Jianhong, S., & Yingqi, L. (2025a). Impact of YouTube User-Generated Content on News Dissemination and Youth Information Reception. *Health Expectations*. <https://doi.org/10.1111/hex.14036>
- Chunqiong, W., Shan, J., Jianhong, S., & Yingqi, L. (2025b). Impact of YouTube User-Generated Content on News Dissemination and Youth Information Reception. *Health Expectations*. <https://doi.org/10.1111/hex.14036>
- Elprida, E., Yulianty, D., Roswinna, W., & Anggraen, A. F. (2026). STRATEGI PENGEMBANGAN KOMPETENSI DIGITAL SUMBER DAYA MANUSIA DI ERA DIGITAL. *Journal of Innovative and Creativity*, 6(1), 19083–19090. <https://joecy.org/index.php/joecy>
- Greenhow, C., & Lewin, C. (2016a). Social media and education: reconceptualizing the boundaries of formal and informal learning. *Learning, Media and Technology*, 41(1), 6–30. <https://doi.org/10.1080/17439884.2015.1064954>
- Greenhow, C., & Lewin, C. (2016b). Social Media and Education: Reconceptualizing the Boundaries of Formal and Informal Learning. *Learning, Media and Technology*, 41(1), 6–30. <https://doi.org/10.1080/17439884.2015.1064954>

- Jarman, B. (2018). Learning to Be: Humanism, Education, and the Philosophy of Erich Fromm. *PhilPapers*, 12(3), 145–162. <https://doi.org/10.5840/philpapers.JARLTB>
- Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., Schmid-Petri, H., & Adam, S. (2018a). Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology. *Communication Methods and Measures*, 12(2–3), 93–118. <https://doi.org/10.1080/19312458.2018.1430754>
- Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, T., Reber, U., Häussler, T., Schmid-Petri, J., & Adam, S. (2018b). Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology. *Communication Methods and Measures*, 12(2–3), 93–118. <https://doi.org/10.1080/19312458.2018.1430754>
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *An Introduction to Information Retrieval*. Cambridge University Press. <http://www.informationretrieval.org/>
- Mimno, D., Wallach, H. M., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing Semantic Coherence in Topic Models. *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, 262–272. <https://aclanthology.org/D11-1024>
- Morgeson, F. P., & Humphrey, S. E. (2014). Work Design: A Review and Theoretical Framework for Next-Generation Research. *Annual Review of Organizational Psychology and Organizational Behavior*, 1(1), 55–79. <https://doi.org/10.1146/annurev-orgpsych-031413-091321>
- Podcasting in the 21st Century. (n.d.). In *User-Generated Content in Education*. Wikibooks. Retrieved https://en.wikibooks.org/wiki/User-Generated_Content_in_Education/Podcasting_in_the_21st_Century
- Rosner, F., Hinneburg, A., Röder, M., Nettling, M., & Both, A. (2014). Evaluating topic coherence measures. *ArXiv Preprint ArXiv:1403.6397*. <https://arxiv.org/abs/1403.6397>
- Saputra, M. A., Rasdi, Muh. A., Resa, & Nur, S. W. (2025). Pengaruh Penggunaan Teknologi Artificial Intelligence (Ai) Terhadap Pola Belajar Mahasiswa Di Era Digital: Antara Efisiensi Atau Ketergantungan. *Laa Maghrib: Journal of Development in Islamic Economics and Business*.
- Service, S. R. (2024). Academic Reference Entry ID 4040136. *Scientific Research Publishing Reference Database*, 2024(1), 101–115. <https://www.scirp.org/reference/referencespapers?referenceid=4040136>
- Tess, P. A. (2015a). The role of social media in higher education classes (real-world applications) – A literature review. *Learning, Media and Technology*, 40(1), 60–80. <https://doi.org/10.1080/17439884.2015.1064954>
- Tess, P. A. (2015b). The role of social media in higher education classes (real-world applications) – A literature review. *Learning, Media and Technology*, 40(1), 60–80. <https://doi.org/10.1080/17439884.2015.1064954>