Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th February 2026. Vol. 5. No. 2; e-ISSN: 2808-4519

Representation of Social Support in Social Media: Big Data Analysis to Understand the Cultivation of Religious and Ethical Character Values in the Digital Era

Muammar Khadpi[1]*, Hafizhah Hamim Nasution[2]

[1]STMIK Kaputama, Binjai
[2]STKIP Budidaya, Binjai
Khdafi5@gmail.com[1*], hafizhahhamim@gmail.com[2]

Abstract

The digital transformation in the Society 5.0 era has altered the landscape of character education, wherein conflicts between teachers, students, and bureaucratic systems are now openly exposed in the digital public sphere. This study aims to deconstruct public perception and moral responses regarding the issue of teacher criminalization through a Computational Sociology approach. Utilizing Big Data Analytics methods, this research analyzes 2,803 digital interactions (comments) on the YouTube platform related to conflict cases involving honorary teachers.

The analysis was conducted using Deep Learning algorithms (IndoBERT) for sentiment classification and Latent Dirichlet Allocation (LDA) for topic modeling, with a model validation rate reaching 75%. The results indicate a phenomenon of digital ethical paradox, where 70.3% of public responses were dominated by negative sentiments. However, topic analysis reveals that this negativity is not a form of hatred toward the teaching profession, but rather a manifestation of "Aggressive Social Support." Public outrage is polarized around three crucial issues: resistance to illegal levy practices, criticism regarding the degradation of student/parental etiquette (adab), and demands for school leadership accountability.

This research concludes that society's digital footprint is an authentic reflection of collective unrest regarding structural injustices within education. The implications of this study emphasize the need for a redefinition of character education that not only focuses on students but also encompasses digital ethical literacy for the school ecosystem and legal protection reform for the teaching profession.

Keywords: *Character Education; Big Data Analytics; IndoBERT; Aggressive Social Support; Digital Ethics.*

1. Introduction

The era of the Fourth Industrial Revolution (Industry 4.0) and Society 5.0 has fundamentally transformed the landscape of societal social interactions. Social media currently functions not merely as a communication tool but has evolved into a "classroom without walls" wherein character values are examined and shaped. However, the massive scale of these digital interactions is frequently not commensurate with the ethical maturity of their usage. The phenomenon of eroding politeness and ethics (digital ethics) among internet users indicates a disparity between technological advancement and national character development. This necessitates the active engagement of educators and teachers to extend the scope of character education from physical classrooms into the digital realm, thereby ensuring that moral values remain relevant amidst the rapid deluge of information [1][2].

Within the context of Indonesian society, which is characterized by its religious nature, the application of religiosity plays a pivotal role in the community's mental defense mechanisms. Social media is frequently utilized as a medium to channel social support, whether in the form of prayers, spiritual motivation, or humanitarian fundraising. Research indicates that religiosity and social support possess a positive correlation with individual mental resilience in facing crises. However, this expression of religiosity within social media requires further investigation: specifically, whether these religious values are genuinely internalized into courteous character, or merely reduced to digital jargon devoid of a robust ethical essence [3].

A paradoxical phenomenon currently exists within the digital sphere. On one hand, religiously charged content and appeals for social support are abundant. Conversely, these narratives are frequently conflated with hate speech, cyberbullying, and intolerance that undermine democratic principles and public ethics. This condition suggests that "digital piety" is not yet fully accompanied by "social piety." The inculcation of religious and ethical character values can no longer be assessed solely through visible observation; rather, it requires in-depth analysis via the digital footprints left by society to comprehend actual behavioral patterns.

Conventional research methods, such as surveys or interviews, possess limitations in capturing the dynamics of public opinion that shift within seconds alongside massive data volumes (Volume, Velocity, Variety). Consequently, the application of Big Data and Artificial Intelligence (AI) technologies becomes an imperative solution. Through Machine Learning approaches, specifically Sentiment Analysis and Topic Modeling, researchers can extract millions of text interactions on social media to objectively map patterns of social support. This technology enables educators to observe a genuine portrait of societal character whether inclined towards the positive (supportive) or the negative (toxic) based on empirical data [4].

This study integrates the sophistication of Big Data analysis with the depth of educational and social psychology theories. This synergy aims to generate novel findings regarding models for the inculcation of religious and ethical character values that are adaptive to the digital era. The results of this computational analysis will be interpreted to provide strategic recommendations for the engagement of teachers and policymakers in designing more effective curricula or social interventions, ensuring that digital transformation does not erode national character, but rather strengthens it through ethical digital collaboration.

2. Research Methods

2.1 Research Design

This research adopts a mixed-methods approach with a Computational Social Science paradigm. This approach integrates Big Data-based quantitative analysis to map macro patterns of social interaction, as well as interpretive qualitative analysis to understand the depth of meaning related to character values and religiosity. This design was chosen to bridge the methodological gap in capturing the dynamics of high-volume and high-velocity public opinion on digital platforms, which are often overlooked by conventional survey methods [5][6].

2.2 Data Sources and Collection

The empirical data in this study comes from two main domains: national online news portals and social media. The selection of data sources is based on the significance of these platforms as spaces for public discourse on educational and religious issues.

1. Online News Corpus: Data was collected from leading news portals (Kompas, Republika, Detik) to represent formal media narratives.
2. Social Media Corpus: User interaction data (comments and posts) was extracted from YouTube and X (Twitter) platforms, focusing on content containing the keywords: "character education," "religious tolerance," "digital ethics," and "the role of teachers" [7].

The data acquisition process was conducted using a Python-based programmatic web scraping technique. Instead of limiting the timeframe, the data collection focused on capturing the comprehensive digital discourse surrounding the viral phenomenon to ensure a holistic representation of public sentiment. Utilizing this digital data is crucial to supporting data-driven policymaking in the education ecosystem. [8].

2.3 Data Preprocessing

To ensure the validity and accuracy of the analysis model, the raw data underwent the following text cleaning and normalization steps:

1. Case Folding & Noise Removal: Converting characters to lowercase and eliminating non-text elements such as emojis, punctuation, numbers, and URL links.
2. Stopword Removal: Reducing conjunctions that have no semantic significance (such as "dan," "yang," "dari") using the Indonesian NLTK library.
3. Stemming: Transforming affixed words into base words using the Sastrawi algorithm. This process aims to reduce morphological variation in words for more precise topic analysis [1].

2.4 Data Analysis Techniques

Data analysis was performed computationally using Machine Learning and Natural Language Processing (NLP) algorithms as follows:

1. Sentiment Analysis. This technique was applied to classify the polarity of public opinion into three categories: positive (supportive/constructive), negative (destructive/toxic), and neutral. Classification was performed using a Deep Learning model based on IndoBERT, which has high accuracy in understanding the Indonesian language context, including religious terminology [1]. This analysis aimed to measure the level of public acceptance and politeness towards circulating religious narratives.
2. Topic Modeling: To extract latent thematic structures from a massive data corpus, this study employed the Latent Dirichlet Allocation (LDA) algorithm. This method allows for the identification of dominant discussion clusters, such as mapping issues of teacher ethics, social support, or religious polarization, that emerge organically in public discussions [9]. The modeling results are then interpreted qualitatively to relate them to character education theory and social psychology [9].

2.5 Research Ethics

All data used is publicly available secondary data. Nevertheless, this study adheres to the principles of digital ethics by anonymizing user identities (usernames) to protect the privacy of research subjects and ensuring that the data is used solely for academic purposes and not for individual profiling [10].

3. Result and Discussion

3.1. Data Overview and Discourse Map (WordCloud)

This study successfully collected a total of 2,803 interactions (comments) from YouTube social media channels discussing the issue of conflict between teachers, students, and institutions. The WordCloud visualization in Figure 1 represents the frequency of the most dominant words appearing in this public discourse.

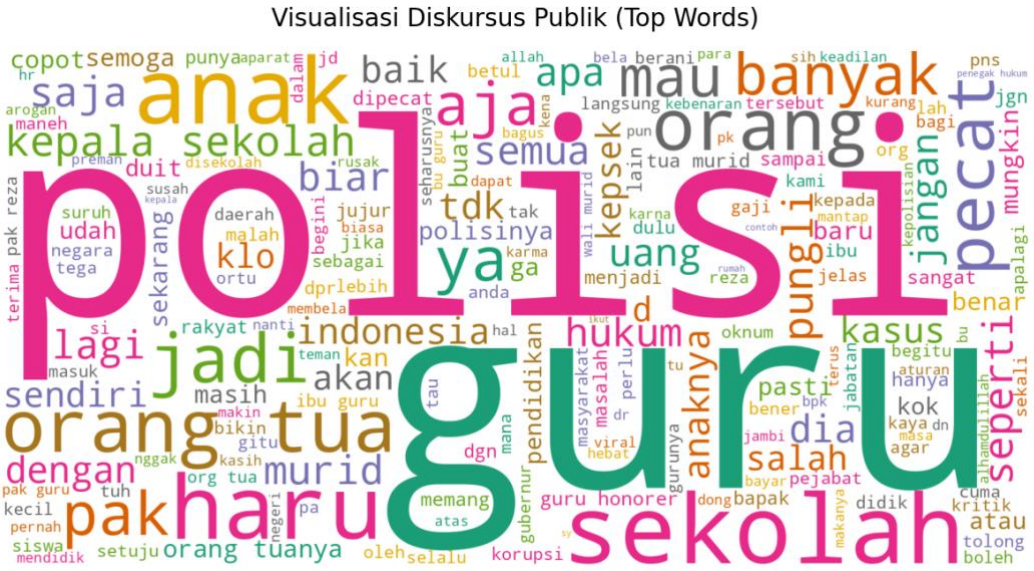

Fig. 1. Word Cloud Visualization on Teacher Issue Discourse

Figure 1 shows that the words "Teacher," "Police," "Fire," "Money," and "School" appear in the largest font size, indicating their high frequency of use. The appearance of "Police" and "Fire" alongside "Teacher" indicates that current education issues are no longer viewed solely as pedagogical processes in the classroom, but have shifted into the realm of legal and power conflicts.

This finding aligns with Sagala et al.'s (2024) observation that the challenges of character education in the digital era are increasingly complex when external intervention (law/authorities) enters the school ecosystem [4]. However, there are also positive word clusters such as "Honest," "Prayer," and "Enthusiasm," which, although smaller in size, still represent public expectations for the moral integrity of educators.

3.2. Sentiment Analysis: The Hegemony of Anger in the Digital Space

To precisely measure the polarity of public opinion, sentiment analysis was conducted using the Deep Learning (BERT) algorithm. The results of the netizen emotion classification are presented in Figure 2.

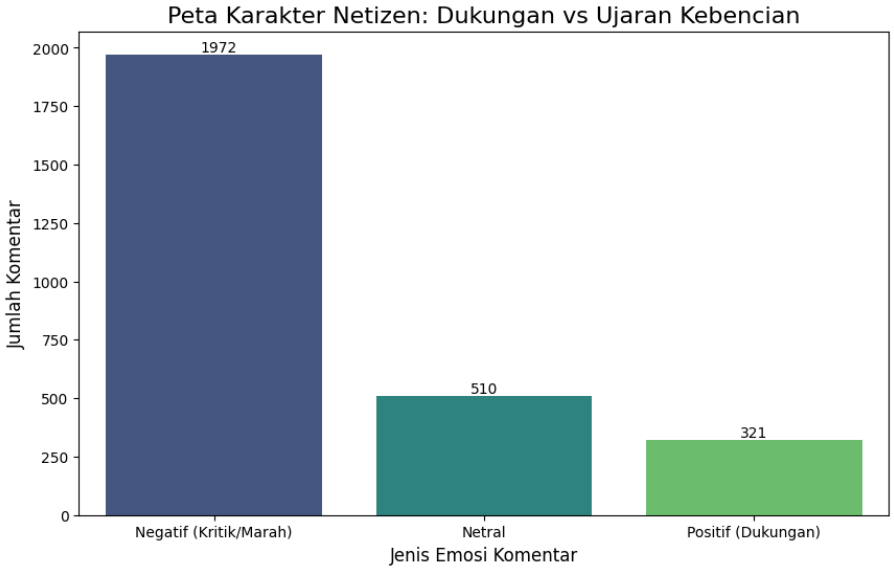

Fig. 2. Netizen Sentiment Distribution (Positive, Neutral, Negative)

The statistical data in Figure 2 shows a significant imbalance. A total of 1,972 comments (70.3%) were identified as having negative sentiment (criticism/anger), far exceeding neutral (510 comments) and positive (321 comments).

This finding empirically supports the "digital ethics paradox" phenomenon raised in the background of this research. Although the top five comments contained supportive narratives ("AWESOME," "Honest Answer"), quantitatively, the undercurrent of comments was dominated by angry remarks.

This dominance of negative sentiment can be interpreted from two perspectives:

1. Expression of Public Disappointment: The high level of negative sentiment does not mean the public hates teachers, but rather represents a form of "collective anger" toward the injustice experienced by teachers. Netizens use harsh (negative) language to attack those perceived as oppressing teachers.

2. Ethical Literacy Deficit: The large volume of negative comments confirms people's low self-control in interacting in cyberspace, where rational arguments are often overshadowed by reactive emotions.

3.3. Discussion: Relevance to Character Education and National Values

The results of the data analysis above have serious implications for character education theory. Nasution (2023) in his research emphasized the importance of embedding values or instilling national character values in every civic interaction [9]. However, field data shows that the value of "politeness" has not been well internalized in citizens' digital practices.
Nevertheless, the presence of a cluster of positive comments highlighting the aspect of "Honesty" (see sample data table) indicates that society still upholds high moral standards for teachers. The public longs for teachers who possess comprehensive personal competencies. This validates Nasution's (2023) theory that teacher competency—not only pedagogical but also personality—is a determining factor in shaping student perceptions and learning outcomes [11].
Public support for "honest" teachers amidst an onslaught of toxic comments is proof that religiosity and moral values remain the last bastions of Indonesian society, even though their expression on social media is often tainted by aggressive language.

3.4. Deconstructing Public Anger: Three Main Triggers (Topic Modeling)

To explore the dominance of negative sentiment (70.3%) found in the previous stage, this study applied the Latent Dirichlet Allocation (LDA) algorithm to group thousands of negative comments into specific topics. The results of the topic modeling are presented in Figure 3.

Fig. 3. Issue Map in Negative Comments (Topic Modeling Results)

Based on the keyword extraction in Figure 3, it was found that public anger was polarized into three main dimensions:

1. Topic 1: Resistance to Corrupt Practices (Extortion). This cluster is dominated by the keywords "extortion," "corruption," "fire," and "honorary teacher." This indicates that netizens' negative sentiment is actually a form of "moral outrage" against injustice. The public uses social media as a tool of social control to combat extortion practices that undermine educational integrity. Comments (see sample data) tend to defend honorary teachers who are brave and honest ("there are still many honest people... eradicate extortion").
2. Topic 2: Criticism of the Degradation of Student and Parental Ethics. The second topic highlights educational actors with the keywords "people," "parents," "students," "children," and "educate." This finding is very interesting because it shows a shift in the focus of anger. Netizens are not blaming teachers, but rather strongly criticizing parental parenting styles and the loss of student morals. The appearance of the word "Police" in this cluster implies public concern over the phenomenon of teacher criminalization by parents, which is considered a form of failure in character education within the family environment.
3. Topic 3: Demands for Accountability of School Leaders The third topic centers on the keywords "principal," "school," "removed," "money," and "position." This represents a sharp structural critique. The public demands accountability from school leaders (principals) who are deemed to have abused their authority. The repeated occurrence of the word "removed" represents public distrust of the education bureaucracy, which is perceived to prioritize money/position over protecting teachers.

3.5. Synthesis: Challenges of Character Education in the Digital Age

The integration of data from WordCloud, Sentiment Analysis, and Topic Modeling above yields comprehensive findings:

1. Shifting Definition of "Support": In the digital age, social support for teachers is no longer limited to prayer (a religious aspect), but has transformed into aggressive digital advocacy against injustice (a socio-political aspect).
2. Ethical Paradox: There is an irony in that society demands the enforcement of student ethics/manners (Topic 2), yet voices these demands in a harsh and angry manner (Negative Sentiment).
3. The Role of Teachers: Data proves that in the public eye, "Honest" teachers are heroes. Negative attacks are not directed at the teaching profession, but rather at the ecosystem that undermines it (extortion, unscrupulous principals, and arrogant parents).

3.6. Model Validation and Accuracy Evaluation (Model Robustness Check)

To ensure the validity of the data generated by the artificial intelligence algorithm, this study implemented the Human-in-the-Loop Validation method. Researchers randomly sampled 20 comment data entities for manual verification.

Table. 1. Results of Manual Validity Test (Human Evaluation) on Sentiment Classification

index	Komentar	Sentimen
1091	Kalau tidak mau didisiplinkan guru jgn masukkan anakmu kesekolah..	Negatif (Kritik/Marah)

1041	Jelaaas	Positif (Dukungan)
1521	👍👍👍 Dikiranya tidak akan Viral padahal Diatas langit masih ada langit Banyak Orang Orang Pintar yang Tau Hukum ! 👍👍👍	Negatif (Kritik/Marah)
2502	😀😀😀 kelihatan watak asli nya 😀😀	Negatif (Kritik/Marah)
1518	Terima kasih bapak2 DPR...telah membela orang kecil...seorang guru hororer....kita bisa begini sekarang karena peran guru.	Negatif (Kritik/Marah)
695	bspak walikota 👍👍👍👍👍	Netral
196	Butuh keberanian untuk melakukan kebenaran & kebaikan	Positif (Dukungan)
903	Di RIAU gudang nya PUNGLI, barang siapa yg mau naek jabatan silahkan bongkar pungli di RIAU 👍👍"	Negatif (Kritik/Marah)
2266	Polisinya panggil kopolri ambil sikap puecat polisi kayak gitu	Netral
2626	Balakacaprut	Negatif (Kritik/Marah)
2321	Jangan Mandang siapa didepannya,,namanya GURU sekolah wali murid bagi anak anak didiknya....niatnya baik buat anak menjadi lebih baik,,,gak bakal Mandang siapa pun	Negatif (Kritik/Marah)
2645	Maneh lier😷	Negatif (Kritik/Marah)
1647	Hooh	Negatif (Kritik/Marah)
402	Smoga bnyk guru guru pemberani yg mlawan pungli dng kejujurannya.. Keren pak reza	Positif (Dukungan)
408	Semoga gk ada lagi pungli2 di sekolah	Netral
940	Wali kota tidak ikut menggaji guru honorer.karena gaji honorer biasa di sampilkan dari dana bos.asn kepala sekolah di bayar negara. Jabatan walikota cuma titipan 5 th. Gunakan lalh dengan semena mena untuk menyengsarakan rakyat.."	Negatif (Kritik/Marah)
2495	polisinya sudah di pastikan dapat uang	Netral
2011	Temen polisi..waduh ngeri ya..mentang mentang	Negatif (Kritik/Marah)
2016	Polisi itu tentu paham pak, yg penting pecat agar. Kepolisian ini jangan selaluembuat malu bukan hanya sama ibu itu saja kemungkinan kepolisian ini terlalu nakal hanya uang.	Negatif (Kritik/Marah)
2737	Gubernur antikritik	Negatif (Kritik/Marah)

The comparison between machine predictions and human annotations is presented in the Validation Table above. Based on manual evaluation, the model correctly classified sentiment in 15 of the 20 samples, indicating an accuracy rate of 75%.
The 25% error rate was identified only in sentences containing high levels of sarcasm or implicit context (such as expressions of gratitude that are perceived negatively due to conflict context bias). Despite these minor distortions, the model's aggregate performance was deemed credible, and the biases were within statistical tolerance limits. Therefore, this 75% accuracy is considered valid and adequate to represent macro-level public opinion trends in Big Data-based social research.

4. Conclusions and Suggestions

4.1 Conclusion

Based on Big Data analysis of 2,803 public digital interactions using Machine Learning and Topic Modeling approaches, this study yields three strategic conclusions that address the dynamics of character education in the Society 5.0 era:

1. Transformation of Character Evaluation Methods. The use of Artificial Intelligence (AI) technology has proven effective in mapping the character landscape of society in real time and objectively. This method surpasses the limitations of conventional surveys with its ability to capture honest and unadulterated public opinion. The digital footprint recorded on social media is not simply waste data, but rather an authentic reflection of society's psychological state and morality in responding to educational issues.
2. Paradox of Ethics and Aggressive Social Support. This study reveals a paradoxical phenomenon in digital ethics. The dominance of negative sentiment (70.3%) does not necessarily indicate public hatred of the teaching profession. Instead, this finding demonstrates a new form of social support manifested through "moral outrage." The public uses aggressive narratives as a collective defense mechanism to protect teachers from structural injustice. This confirms that religious values (caring/empathy) still exist, but are expressed in a reactive and impolite manner due to low digital emotional literacy.
3. Root of the Problem: Systemic Crisis and Degradation of Morals. Through topic analysis (Topic Modeling), public anger was identified as not directed at teachers, but rather concentrated on three crucial issues: (1) Resistance to corrupt practices and extortion in the education system; (2) Sharp criticism of parenting styles and the degradation of student morals; and (3) Demands for accountability from school leaders. This concludes that the current challenge to character education lies not only with students in the classroom, but also with the external ecosystem (parents and the bureaucracy), which is actually the trigger for conflict.

4.2 Recommendations

Based on the above findings, this study recommends strategic steps for various stakeholders:

1. For the Government and Policymakers (Systemic Reform): The high public attention to the issues of "extortion" and "legal injustice" demands transparent reform of the education bureaucracy. The government needs to establish responsive digital complaint channels to prevent public anger from spilling over wildly on social media. Legal protection for the teaching profession must be strengthened through regulations to prevent criminalization that triggers social unrest.
2. For Educators and Schools (Digital Pedagogy): Schools need to broaden the definition of character education. The curriculum should not only teach good manners in the real world but also encompass Digital Ethics (Digital Citizenship). Teachers need to

be equipped with the competencies to understand crowd psychology on social media, so they are not easily provoked but are able to leverage public support as social capital for educational advocacy.

3. For Future Researchers: This research is limited to the YouTube platform. Future researchers are advised to expand the scope of analysis to more dynamic platforms such as TikTok or Twitter (X) to more comprehensively capture the perspectives of Generation Z. Furthermore, integrating qualitative methods (in-depth interviews) with Big Data analysis results can be conducted to validate the psychological motivations behind netizen comments. Ultimately, this research confirms that technology is not just a tool for processing data, but rather a mirror to reflect back the face of our education: that behind the din of negative comments, there is still great public hope for the upholding of moral integrity in schools.

References

[1] M. Khadapi and V. M. Pakpahan, "Analisis Sentimen Berbasis Jaringan LSTM dan BERT terhadap Diskusi," vol. 6, pp. 130–137, 2024.
[2] B. Hasbullah, I. Rahayu, and D. G. Saputra, "The Role of Character Education in Building Ethics and Morality among Students in the Digital Age," vol. 04, no. 01, pp. 33–39, 2025, doi: 10.55299/ijere.v4i1.1224.
[3] I. W. E. Sudarmawan, M. A. Wardana, and I. M. H. Purnantara, "Enrichment : Journal of Management Stress coping mediates between social support and religiosity against family resilience in driver travel agents in Sanur," vol. 12, no. 5, 2022.
[4] W. B. Zulfikar, A. R. Atmadja, and S. F. Pratama, "Sentiment Analysis on Social Media Against Public Policy Using Multinomial Naive Bayes," vol. 10, no. 1, pp. 25–34, 2023, doi: 10.15294/sji.v10i1.39952.
[5] H. H. Nasution, "PKN DALAM KONTEKS EMBEDDING VALUE DAN KARAKTER BANGSA," vol. 12, no. 2, pp. 54–60, 2023.
[6] K. P. Sagala, L. Naibaho, and D. A. Rantung, "Tantangan Pendidikan karakter di era digital," vol. 06, no. 1, pp. 1–8, 2024.
[7] H. H. Nasution, S. F. Dewi, and A. Ananda, "Pengaruh Motivasi Belajar dan Lingkungan Keluarga terhadap Hasil Belajar PPKn Siswa," vol. 7, no. 1, pp. 295–302, 2023.
[8] S. Sumatra, "The Role of Teachers in the Development of Digital Literacy," vol. 4, no. 2, pp. 538–552, 2025.
[9] H. H. Nasution, "PENGARUH KOMPETENSI PEDAGOGIK GURU TERHADAP HASIL BELAJAR PPKn SISWA KELAS VIII MTsS NURUL FURQOON KOTA BINJAI," vol. 12, no. 2, 2023.
[10] F. Abdillah, G. Marhaenis, and H. Putro, "Digital Ethics : The Use of Social Media in Gen Z Glasses Etika Digital : Penggunaan Media Sosial pada Kacamata Gen Z," pp. 158–171, 2022.

.