



Original Article

Enhancing Academic Integrity for Bangladesh's Educational Landscape

Rifat Al Mamun Rudro¹ , Md. Faruk Abdullah Al Sohan² , Afroza Nahar³ 



<https://doi.org/10.62865/bjbio.v15i2.90>

Abstract: Integrating Artificial Intelligence (AI) tools in Bangladesh's academic landscape has ignited concerns about potentially eroding students' creative writing and critical thinking abilities. While AI offers efficient and personalized learning, there is a looming risk of students using it as a shortcut to success. Educators and policymakers must emphasize the cultivation of writing skills and critical thinking while guiding students to recognize the limitations of AI. While plagiarism checking is crucial for academic integrity, it often falls short in acknowledging students' originality. Natural Language Processing (NLP) emerges as a promising solution, augmenting plagiarism detection by comprehending context, nuances, and novel expressions. This paper proposed a model based on NLP that can incorporate a robust token identification mechanism into AI tools used by the students. A unique identification token is generated at the time of system login and stored in a publicly accessible resource, acting as a distinct marker for user activity. These identification tokens support a plagiarism detection system that continually reviews and compares activity records. It also enables a thorough examination of user involvement across various academic supplementary tools, such as academic writing, presentation creation, image generation, etc.

Keywords: NLP, token generalization, plagiarism, examination, presentation

Introduction: In the dynamic landscape of modern education, integrating Artificial Intelligence (AI) tools has ushered in a wave of

transformative possibilities for learning and academic support. Nevertheless, the growing reliance on AI tools within academia has

1. MSc Student, Dept. of Computer Science, American International University-Bangladesh (AIUB), Email: rifat.rudro138964@gmail.com ORCID ID: <http://orcid.org/0000-0002-0608-0369>
2. MSc (Computer Science), Lecturer, Dept. of Computer Science, American International University-Bangladesh (AIUB). Email: faruk.sohan@aiub.edu ORCID ID: <https://orcid.org/0000-0002-8291-7431>
3. PhD (Energy Engineering), Associate Professor, Dept. of Computer Science, American International University-Bangladesh (AIUB). Email: afroza@aiub.edu ORCID ID <https://orcid.org/0000-0001-8039-0296>

Corresponding Author: Rifat Al Mamun Rudro, Email: rifat.rudro138964@gmail.com



raised a pressing concern, the potential erosion of students' creative writing skills and critical thinking abilities¹. While AI offers undeniable advantages in terms of efficiency, accessibility, and personalized learning experiences, there is a looming risk that students may increasingly resort to these tools as shortcuts, potentially diminishing their engagement in independent creative processes². This paper explores the balance between utilizing AI technology's advantages and protecting students' fundamental skills. This article also highlights the crucial responsibility of educators and policymakers in promoting the growth of students' inherent writing skills and analytical thinking while also presenting AI as an additional tool that enhances rather than replaces cognitive efforts. The paper explores the evolving paradigm of academic integrity, particularly in plagiarism detection. Traditional plagiarism checkers may not accurately identify the originality and distinct perspectives in students' work³. The paper proposes integrating natural language processing (NLP) into plagiarism detection tools to better understand academic content's context, nuances, and inventive expressions. This integration can analyze the semantic meaning behind text, allowing for a more comprehensive assessment of originality. This approach can potentially revolutionize plagiarism detection and address it in academic settings, promoting a deeper understanding and appreciation of student's unique contributions. The research examines novel approaches to leverage the capabilities of AI while maintaining the originality of student's intellectual input, in a time when collaboration between technology and education is crucial⁴. It emphasizes the crucial significance of achieving a harmonious equilibrium between technological expertise and fundamental human abilities in pursuing a comprehensive and forward-thinking educational experience⁵. As AI continues to transform the educational environment, it is important to acknowledge that technology should function as a tool to augment, rather than supplant, human connection and creativity. By cultivating a pedagogical atmosphere that promotes the development

of critical thinking, cooperation, and emotional intelligence in conjunction with the integration of artificial intelligence, educators can enable students to become flexible problem solvers well-prepared to tackle future challenges.

Methodology: This paper employs a mixed-methods approach, combining qualitative and quantitative research methods to provide a comprehensive understanding of the impact of AI in education and the effectiveness of NLP in plagiarism detection⁶. The duration of the research was from September to November 2023. In the quantitative aspect of the study, data were collected through surveys on 196 participants administered to a diverse sample of educators and students across various educational institutions. The surveys were designed to gather quantitative data on AI's perceived benefits and educational challenges. Additionally, data related to the effectiveness of NLP in plagiarism detection were collected through quantitative measures, including the analysis of plagiarism detection accuracy rates using NLP algorithms.

Results:

Used Dataset: The used dataset presents the results of a survey that aimed to explore different aspects of the educational landscape. The dataset includes 196 responses, with male participants accounting for 152 (77.6%) and female participants representing 43 (21.9%) and prefer not to say their gender representing 1 (0.5%). Most participants, 178 (90.8%), identified themselves as students. At the same time, the remainder belonged to different educational backgrounds, including SSC (Secondary School Certificate), HSC (Higher Secondary Certificate), Postgraduate, PhD, Above PhD, and Below SSC. The dataset also provides information regarding the roles of the participants, including students, academicians, and other classifications, which provides valuable insights into the diverse perspectives and backgrounds of the respondents. These insights contribute to the comprehensiveness of the dataset.

Proposed Model Architecture: The proposed architecture for integrating AI tools in an academic environment follows a systematic ten-step process, as shown in Figure 1.

At first, students or users log into AI tools designed for academic purposes which prompts the real-time generation of unique identification tokens. This token is stored securely for subsequent reference. The stored tokens are used for diverse purposes including plagiarism detection and academic assessment.

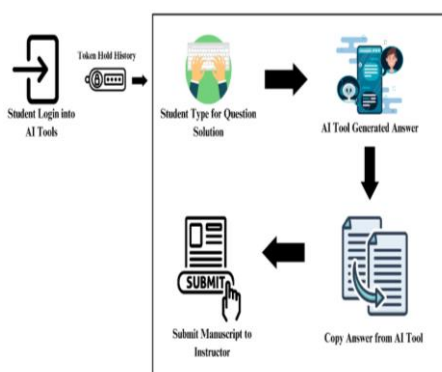


Figure 1 Token Identification from Student End

The architecture integrates a plagiarism detection system wherein a dedicated system compares activity records associated with each stored token⁷ including user actions within various academic tools and identify potential user engagement patterns and trends to locate any potential misconduct or irregularities. The architecture establishes a direct relationship between the plagiarism detection system and AI tools enabling the classification of students' academic writing and activity and checking for originality⁸.

The system's capacity to discern patterns and trends in user engagement provides a nuanced understanding of academic improvement over time, aiding in tracking student progress and offering personalized feedback⁹. Figure 2 represents the NLP pattern recognition concepts in different categories of users where the students' population is 178 out of 196, academicians is 8 and rest of the other categories. The token-based identification system ensures unique

identification and tracking of each user's activity¹⁰, enhancing the overall efficacy of plagiarism detection and academic assessment within the educational framework.

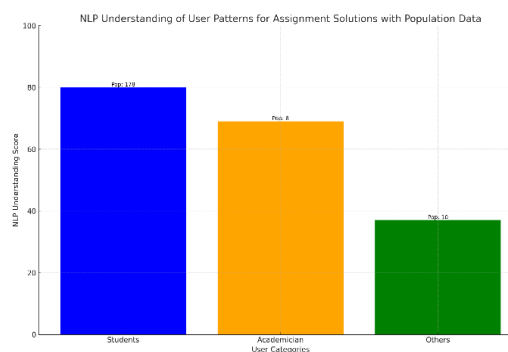


Figure 2 NLP pattern recognition in different category segments

Proposed Model Outcome: Integrating Artificial Intelligence (AI) tools in Bangladesh's academic landscape has introduced transformative possibilities for education¹¹. It has also sparked concerns about potentially erasing students' creative writing and critical thinking abilities, as shown in Figure 3. Despite these concerns, the proposed system offers several key outcomes that can significantly enhance the educational experience.

The robust plagiarism detection system powered by unique identification tokens enhances the identification accuracy of potential plagiarism or suspicious activity¹². Moreover, the Natural language processing (NLP) enables comprehending the context, nuances, and novel expressions within academic content that in turn improves academic integrity. Apart from that, the system offers personalized feedback by analyzing user's engagement patterns and content creation¹³. This also allows the educators to track students' progress over time¹⁴.

The token-based identification system uniquely identifies each user and tracks their activity enhancing security and accountability. The good thing with this architecture is that it underscores the importance of AI as a tool to augment rather than replace human creativity and critical thinking.¹⁵

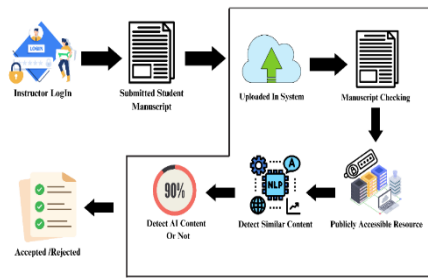


Figure 3 Assessment form Instructor End

Survey Data Analysis: Based on the survey response received from a sample of 196 individuals, ChatGPT has predominantly attracted young adults aged between 18-30 years, with a majority being male (152 respondents) and undergraduate students (174 respondents), particularly in the field of Engineering (138 respondents). These individuals primarily use ChatGPT as students (178 respondents) and are particularly interested in receiving help with homework or assignments around 118 respondents, as shown in Figure 4. The most reported frequency of use is daily (71 out of 196). They used it for doing their homework or assignments (50 out of 71). A subtle improvement regarding the impact on the educational experience has been noted, and it promotes independent learning (65 out of 196). Plagiarism has emerged as a prominent ethical concern (61 out of 196), as shown in Figure 5. In general, satisfaction with ChatGPT for educational support is high, and opinions on its integration into formal education curricula are neutral representing 57 (29.1%) respondents shown in Figure 6. This data underscores the significant role of AI tools like ChatGPT in educational settings and highlights concerns and potential growth areas.

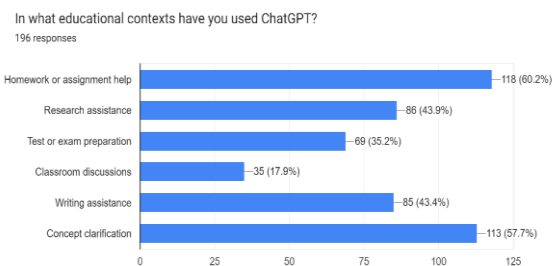


Figure 4 The chart represents a survey on the educational

contexts in which respondents have used ChatGPT, based on 196 responses.

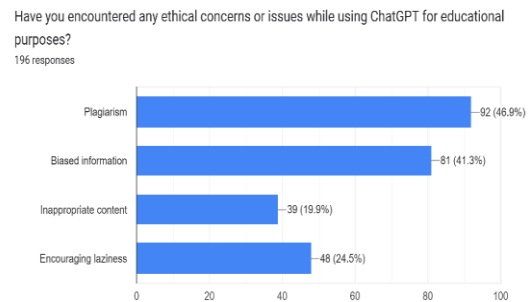


Figure 5 The chart displays the ethical concerns or issues encountered by respondents while using ChatGPT for educational purposes. N = 196.

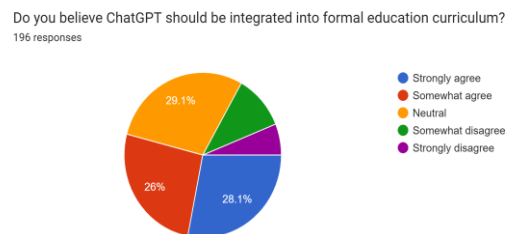


Figure 6 The chart represents the decision for integrating ChatGPT into formal education curriculum

Figure 7 represents the developed system user input contents where the first half represents the contents, and the second half represents the user asking for the content solutions.

Input Content

A blockchain consists of programs called scripts that conduct the tasks you usually would in a database: Entering and accessing information and saving and storing it somewhere. A blockchain is distributed, which means multiple copies are saved on many machines, and they must all match for it to be valid. → Content

What Is the Blockchain Architecture Gaps? Make the findings in point

→ User asking for the solution

User Input 1

Input Content

DeepLabv3+ is a semantic segmentation architecture that builds on DeepLabv3 by adding a simple yet effective decoder module to enhance segmentation results.

Multiple downsampling of a CNN will lead the feature map resolution to become smaller, resulting in lower prediction accuracy and loss of boundary information in semantic segmentation.

How does the DeeplabV3+ are more suitable from U-Net Architecture?

User Input 2

Figure 7 User Inputs

System Interface: We have conducted qualitative research on around 196 population. Our targeted audience was mainly students in different education segments. Figure 8 depicted that AL generated content reports based on Figure 6 user inputs.

AI-Generated Content Detection Report:

Report Item	Details
User	Rifat Rudro
Submission Date	19:12:2023
IP Address	192.168.1.100
Analysis	
Detected AI-Generated Content	62%
Identified AI Model	GPT-3.5
Token Identification	b4bdf874-8c03-5bd8-8fd7-5e409dfdd82c0
AI-Generated Content	AI-powered tools offer numerous benefits, such as personalized learning experiences and efficient content creation. However, it's crucial to strike a balance between utilizing AI and preserving the unique abilities of human educators and students.
Feedback	Our system has detected that a significant portion of your submission appears to be generated by an AI model (GPT-3.5). While AI-generated content can be a valuable asset, it's essential to provide original insights and critical thinking in your academic work. Please review and enhance your submission with your unique perspectives and analysis. If you have any questions or concerns regarding this report, please feel free to reach out.

Figure 8 System Generated Report

Conclusion: Artificial Intelligence technologies in education have many benefits. Though they may weaken students' creativity and critical thinking. There are some innovative approaches like Natural Language Processing (NLP) for plagiarism detection that can help us appreciate and safeguard originality in academic work. Artificial Intelligence should serve as a tool to enhance, rather than replace, human connection and creativity. The research shows how important it is to find a method for AI and human behaviors to coexist peacefully. It stresses that technology should be used to improve things, not replace them. Encourages educators to increase critical thinking, teamwork, and emotional intelligence while integrating AI to prepare students for future difficulties. In real life, we have implemented a plagiarism detection system by NLP that can give approximately 97-98% AI-generated content.

Recommendation: The recommendations revolve around responsible and balanced AI integration, focusing on skills development, ethical use, and ongoing evaluation to ensure that AI technology enhances the educational experience without compromising students' fundamental abilities.

References

1. Leung CH, Chan YY. A natural language processing approach to automatic plagiarism detection. In Proceedings of the 8th ACM SIGITE conference on Information technology education 2007 (pp. 213-218). <https://dl.acm.org/doi/abs/10.1145/1324302.1324348> (Accessed: 13 December 2023)
2. Chong MY. A study on plagiarism detection and plagiarism direction identification using natural language processing techniques 2023. Ph. D. Thesis. University of Wolverhampton. Available at: <http://hdl.handle.net/2436/298219> (Accessed: 15 December 2023)
3. Ali, W., Ahmad, T., Rehman, Z., Rehman, A., Shah, M., Abbas, A., & Dustgeer, G. (2018). A Novel Framework for Plagiarism Detection: A Case Study for Urdu Language. In 2018 24th International Conference on Automation and Computing (ICAC) (pp. 1-6). <https://ieeexplore.ieee.org/abstract/document/874912> (Accessed: 17 December 2023)
4. Hong, Y, Yang, J, Chen, Y, Dong, H Research on the Development of Online Education in the Age of AI and 5G. In 2021 2nd International Conference on Education, Knowledge and Information Management (ICEKIM) 2021 (pp. 233-237). <https://ieeexplore.ieee.org/abstract/document/947952> (Accessed: 17 December 2023)
5. He, Z, Niu, X, Applying Artificial Intelligence to Primary and Secondary School Physical Education. In 2021 2nd International Conference on Information Science and Education (ICISE-IE) 2021 (pp. 1577-1581). <https://ieeexplore.ieee.org/abstract/document/974254> (Accessed: 18 December 2023)
6. Neumann, M, Rauschenberger, M, Schön, EM "We Need to Talk About ChatGPT": The Future of AI and Higher Education. In 2023 IEEE/ACM 5th International Workshop on Software Engineering Education for the Next Generation (SEENG) 2023 (pp. 29-32). <https://serwiss.bib.hs-hannover.de/frontdoor/index/index/docId/2467> (Accessed: 20 December 2023)
7. Alargami, A, Eljazzar, MImam: Word Embedding Model for Islamic Arabic NLP. In 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES) 2020 (pp. 520-524). DOI:<http://doi.org/10.1109/NILES50944.2020.925793>
8. Rudro RA, Al Sohan MF. Utilization of Machine Learning Strategies in the Investigation of Suspected Credit Card Fraud. Int. J. Advanced Networking and Applications. 2023;15(02):5869-74. DOI: <http://doi.org/10.35444/IJANA.2023.15205>
9. Lertpiya, A, Chaiwachirasak, T, Maharattanamalai, N, Lapjaturapit, T, Chalothorn, T, Tirasaroj, N, Chuangsuwanich, EA Preliminary Study on Fundamental Thai NLP Tasks for User-generated Web Content. In 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAI-NLP) 2018 (pp. 1-8).

- <https://ieeexplore.ieee.org/abstract/document/8692946> (Accessed: 20 December 2023)
10. Rudro RA, Al Sohan MF, Chaity SK, Reya RI. Enhancing DDoS Attack Detection Using Machine Learning: A Framework with Feature Selection and Comparative Analysis of Algorithms. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*. 2023;14(03):1185-92. DOI: <https://doi.org/10.61841/turcomat.v14i03.14086>
 11. Lasker, S. P., Begum, M., Hossain, A., Matin, M. A., Islam, S., & Macer, D. (2021). Understanding of Authorship by the Post Graduate Medical Students at a Center in Bangladesh. *Bangladesh Journal of Bioethics*, 12(1), 25–34. <https://doi.org/10.3329/bioethics.v12i1.5197>
 12. Jiang, N, Yuan, C On Current Situation, and Improvement Strategies of Family-Centered Early Education in Urban Areas of Chongqing. In 2020 International Conference on Big Data and Informatization Education (ICBDIE) 2020 (pp. 1-5). <https://ieeexplore.ieee.org/abstract/document/9150275>. (Accessed: 21 December 2023)
 13. Lu, J, Shi, Y, Research on teaching mode innovation in higher education under the guidance of distance education concept. In 2020 International Conference on Modern Education and Information Management (ICMEIM) 2020 (pp. 171-174). <https://ieeexplore.ieee.org/abstract/document/9384797> (Accessed: 21 December 2023)
 14. Ma, R, Yang, X, Gao, F, Non-Fungible Token forecast based on LSTM. In 2022 4th International Conference on Applied Machine Learning (ICAML) 2022 (pp. 81-84). <https://ieeexplore.ieee.org/abstract/document/10056583> (Accessed: 21 December 2023)
 15. Lu, X, Research Hotspots and Development Trends of AI in Education: A Study Based on Knowledge Map and Co-Word Analysis. In 2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE) 2020 (pp. 212-21) <https://ieeexplore.ieee.org/abstract/document/9131380> (Accessed: 21 December 2023)

Author contribution: Rifat Al Mamun Rudro; Conceptualization the idea, wrote the methodology, developed Rudro software, done the analysis, and supervise the research, Md. Faruk Abdullah Al Sohan and Afroza Nahar; investigate the survey, wrote the original draft, and reviewed and edited the manuscript. Md. Faruk Abdullah Al Sohan, Afroza Nahar and Rifat Al Mamun Rudro; visualization, all authors have read and agreed to the published version of the manuscript.

Conflict of interest: The authors declare no conflict of interest.

Funding: This research did not receive funding from any sector.

Data availability: The used data are available at <https://github.com/RifatRudro/Data-Science-Work> .