INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

OPEN ACCESS
Received: 13/09/2025
Accepted: 27/09/2025
Published: 20/10/2025

Citation: Sowmiya R, Lavanya B (2025) Authorship Verification with AI-Generated Mimicry: A Study with Bronte-Style Poems. Indian Journal of Science and Technology 18(37): 3073-3082. https://doi.org/10.17485/IJST/v18i37.1579

* **Corresponding author**.

sowmiyar99r@gmail.com

Funding: University Grants Commission (UGC), India

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Authorship Verification with AI-Generated Mimicry: A Study with Bronte-Style Poems

R Sowmiya[1]*, **B Lavanya**[1]

1 Department of Computer Science, University of Madras, Tamil Nadu, India

Abstract

Objectives: Determining authorship is crucial in digital humanities, forensics, and research paper attribution. Authorship Verification (AV), an open-set problem, determines whether the same individual authored the provided pair of texts. **Methods:** This study examines the poetry of the Brontë sisters, which demonstrates considerable thematic and linguistic similarity, therefore presenting a critical problem for AV. The difficulty of the challenge is further exacerbated by the inclusion of Artificial Intelligence (AI)-generated poems emulating the writing style of the Brontë sisters. This study examines the optimal representative feature aspect when trained using both complete poems and excerpts. **Findings:** The proposed AV framework utilizes a Siamese network that uses Manhattan distance-based similarity assessment with optimal parameters. In the absence of a comprehensive context, the semantic and structural features of poetry significantly enhance AV performance, surpassing quantitative lexical components by 36%. The overall ability of the baseline and deployed models to distinguish between original and AI-generated poems is examined using the sub-task of closed-set Authorship Attribution (AA). This study establishes a basis for future research on AI plagiarism and academic integrity, offering significant insights for further investigation. Also, AV with AI-generated poems highlights that AI can mimic the authors only to some extent, and that is well depicted with the accuracy of the demonstrated method. **Novelty:** A computationally and literarily challenging AV task, involving poems, has been carried out (including AI-generated poems). To the best of our knowledge, a few works on poem authorship exist, but none systematically on AV on poem excerpts.

Keywords: Poetry verification; AI poems; Authorship verification; Siamese networks; Stylistics

1 Introduction

Poetic language is intricately stylized, serving as a prime subject for examining personal writing styles, linguistic innovation, and stylistic development across time. It frequently exhibits non-standard grammar, vocabulary, and syntax. Examining authorship in this context enhances the capabilities of Natural Language Processing (NLP) models, extending their robustness in dealing with varied linguistic styles. In computational linguistics, Authorship Verification (AV) seeks to ascertain if a specific text is composed by a designated author, typically through the analysis of distinctive language patterns. The works of the Brontë sisters: Anne, Charlotte, and Emily, offer an expansive ground for this objective. Their texts demonstrate significant stylistic uniformity stemming from their common literary and cultural contexts; yet, their individual works vary in tone, structure, and topic focus, albeit with certain similarities or overlaps. The shared experience of cohabitation and collaboration in writing enabled them to claim the formation of a cohesive female narrative characterized by thematic inclinations while maintaining distinct artistic styles. This is evident at various levels of textual coherence: individual primary or secondary narrative elements, character development, intersecting themes and topics, poetic devices, and linguistic constructions, etc.[1],[2],[3]. Utilizing the Siamese neural networks and exploring various features, the primary characteristics that differentiate their unique voices and confirm authorship are identified. On testing with Artificial Intelligence (AI)-generated poems in the styles of Brontë, which are unseen by the model during the training, this research also investigates how well the model differentiates between original and generated poems of the same author, especially when the intra-stylistic difference, per se, is subtle and complex. Other interesting NLP questions bordering Natural Language Generation (NLG) analyzed in this study are "When the poems - both the original and AI-generated ones - are broken and only a stanza is provided to observe the authorial style, how well do the models perform?" and "Which feature helps it to learn better in such a scenario?". Thus, the main contributions of this paper are summarized below:

- Authorship Verification of Victorian Poets: The study verifies the authorship of Victorian-era poets who exhibit significant thematic and textual similarities, highlighting the effectiveness of the proposed AV framework.
- Evaluation on Fragmented Texts: The proposed approach is assessed using incomplete poems lacking full text, establishing a robust framework applicable to fragmentology, philological analysis, textual forensics, and related domains.
- Comparative Analysis of Feature Representations: The performance of trainable embeddings is analyzed alongside feature-engineered stylistic and linguistic attributes in the context of Authorship Attribution (AA) and AV.
- Dataset Construction and Availability: A new dataset has been curated and will be made available upon mail request for further research.
- Investigation of model's understanding in AA and AV tasks: This study examines the model's errors in AA and AV of original and AI-generated poems attributed to the Brontë sisters

Accelerated research in the area of differentiating generative AI content from human content is spreading across multiple media, such as images[4],[5], videos[6], etc. While attribution of generated text[7] is unfolding its possibilities to mitigate[8] or debate about plagiarism at academic and research levels[9], among other issues, the creative level of producing poetry is also being challenged. Recent research highlights how people find AI-generated poetry indistinguishable from human-written ones[10].

Largely, the Authorship Attribution (AA) task of determining the author of the given text from one of the known authors, has been popular in poetry research. Sentiment polarity features have been found to work well for poetry attribution[11], both as a standalone feature and when used along with other features in an ensemble classifier. Using the weighted and modified version of the term frequency-inverse document frequency, authors of[12] attempted to also attribute poems to the authors, but that method did not yield as high a result as novel attribution. Approaching poetry as sound-sensitive material, authors of[13] experimented with phoneme-level and character-level n-grams, along with function words for attributing authors of English poems, employing the Support Vector classifier.

From the standpoint of the deployed model in this study (Siamese neural network utilizing Manhattan distance-based computation), Manhattan-Long Short-Term Memory (MaLSTM) has been mildly popular in a few poetry verification tasks, including the work of[14], which verified the authorship of Indonesian poetry using MaLSTM. Besides, Siamese neural networks in general have been employed in text authorship verification as in[15], where graph-based components are also integrated into the framework.

As briefly introduced at the beginning of this Section, Brontë's works, utilized in this study, despite being an impetus in the Victorian literary revolution and adding to the female literary wealth, are controversial with claims of co-authorships. Charlotte, elder to the other two sisters, is suspected of correcting and updating (adding and deleting some portions of) the works of Emily to ameliorate the critics of that time[16]. In order to maintain the scope of this work to verify the authorship of a single author attributing a work, despite multiple suspicious poems spotted by various literary analysts[17],[18], two poems suspected of an evident higher rate of co-authorship (as hinted[16]) have been excluded from this study, namely: 'Encouragement' and 'Visionary' by Emily. There is a boost in the accuracy of the baseline models after deleting those poems; especially perceptron, while performing AA on the original poems with StyloMetrix – an open-source tool to extract a set of linguistic features[19] in two variations: on broken poems/excerpts attribution and complete poems attribution yielded a boost of 7% and 26%

respectively after the two poems are removed. This may well be credited to the clearer pattern of authorial styles that is observable by the models, after deletion.

The rest of the paper is organized into the following sections: Section 2 presents the dataset and prompt details, and the methodology. Section 3 presents the results and error analysis, and Section 4 summarizes the work.

2 Methodology

2.1 Dataset and Prompt details

The Brontë poems are scraped from the Free E-book available at the Project Gutenberg site[20]. Untitled poems or those including only a few lines without distinct stanzaic organization are likewise regarded as distinct components and incorporated into the study. Altogether, this set of poems is loosely referred to as 'orig_set' or 'original' in the rest of the work. The details about the original dataset are tabulated in Table 1, where LC represents Line count, SC represents Stanza count, Std is the standard deviation, and Min and Max are the minimum and maximum, respectively. As for short notation for author names, the first letters of the authors in the orig_set are used. (e.g. Emily in the orig_set is 'E'). The next step of data curation required AI-generated poems for which five different open-source generative AI models (GenAI)[21],[22],[23],[24],[25] are used.

Table 1. Full_orig Dataset details

Author	Instance	Mean LC	Std LC	Min LC	Max LC	Mean LC
A	31	42.03	36.94	12	216	8.903
C	25	86.64	68.15	12	258	12.36
E	36	37.86	17.48	12	72	8.083

LC stands for Line count and Author A – Anne, C – Charlotte and E - Emily

Initially, ChatGPT[22] is prompted to explain its understanding of stylistic differences and the uniqueness of Brontë's works. Its idea about their thematic overlaps and stylistic analysis is highly agreeable with the introductory opinions of[3]. Later, the prompt for the AV task in hand is also enhanced by ChatGPT, which is finalized and used to prompt the other models as well. Prominently, batch-wise data elicitation is done in a batch of three (one for each author; each model is prompted once, except ChatGPT, which is prompted twice), and thus a total of eighteen poems is procured successfully. The prompt for GenAI models is as follows:

'Generate three poems, each in the distinct styles of the Brontë sisters.

- Anne Brontë: Write a 12-line poem that is introspective and serene, with themes of nature and spirituality. Use a gentle tone and rich vocabulary.
- Charlotte Brontë: Write a 12-line poem that is dramatic and emotional, exploring themes of resilience, human struggles, and triumphs. Use vivid imagery and strong, structured language.
- Emily Brontë: Write a 12-line poem that is wild, untamed, and passionate, with a strong connection to rugged nature and themes of freedom and intensity. Use raw, evocative language.

Format the output in a CSV-ready style:

Poem, Author
"Poem 1 in Anne Brontë's style", Anne
"Poem 2 in Charlotte Brontë's style", Charlotte
"Poem 3 in Emily Brontë's style", Emily.'

Noticeably, a minimum line constraint of twelve lines, the shortest length recorded in the orig_set, is enforced, and a total of six poems per author are collected. In the rest of the study, this set is referred to as the 'gen_set'. Later, the problem of verification is branched into two:

- Presence of full text
- Absence of full text

For the latter, both the orig_set and gen_set poems are broken at an interval of four lines. The set of broken/excerpts of poems is denoted with the prefix 'broken', e.g., broken_orig and broken_gen, whereas the complete poems are prefixed by 'full' or 'Full'. The broken_gen set has eighteen instances for each poetess. The unique word ratio (uwr) of a poem is considered the ratio of

Table 2. Mean Unique Word Ratio (UWR) details

Set	Orig_set	A	C	E	G_A	G_C	G_E
Full	92	0.73	0.72	0.77	0.82	0.835	0.832
Broken	1181	0.93	0.95	0.94	0.93	0.933	0.922

total unique words to the total number of words in the poem, and its details are in Table 2, along with other information. The first letters prefixed with 'G' for the gen_set are used to represent the authors (e.g. Emily in the gen_set is 'G_E').

It is important to understand illustrated in the prompt above, the model is prompted along with the stylistic specifications that are themselves extracted from the model; thus, this approach embeds the model's knowledge of Brontë's styles in the final prompt. This is to limit the outputs within the intended scope and prevent extraneous generation or hallucinations.

On moving further, the dataset appropriate for the AV task is built. With random sampling and pairing, two types of AV datasets: primary (P) and secondary (S), are prepared for each of the full and broken cases. The training data only consists of the original poems, whereas the test set has both the orig_set and gen_set in substantial amounts. The splits of AV datasets are tabulated in Table 3.

Table 3. Splits of AV datasets

Datasets/splits	Full (P)	Broken (P)	Full (S)	Broken (S)
0 (both the texts in a pair are not written by the same author)	3268	4565	4978	2299
1 (both the texts in a pair are written by the same author)	3268	4565	810	1637

Redundant pairs in the same set are avoided. The train–validation–test split is in the ratio of 70:10:20. The secondary datasets are designed to test the model on generative instances in pairs. Due to the limited number of instances, the test set from the primary set, and a separate randomly paired subset are combined and put to use.

Following the quantification of the predictive power of the model, to enhance the explainability of the paired original authorial distinction of the model, during the error analysis, a subset consisting of 1008 pairs with 28 instances for each author pair is built, called the 'rearranged' set.

2.2 Experiment

In the preliminary stage, the StyloMetrix features (a total of 196) are extracted and popular baseline Machine Learning models (ML): Support Vector Machine variants – SVM with Radial Basis Function (RBF) kernel (SVM) and Linear SVM (L_SVC), Decision Tree Classifier (DT) and Logistic Regression (LR) and Perceptron (PP) are trained and tested on AA task with 80-20 split of individual dataset respectively. This is carried out to test the efficiency of the quantitative lexical features in representing the authorial styles. Table 4 organizes the accuracy of the baseline models with different datasets. Given the smaller dataset size of full_gen and broken_gen, LR, L_SVC, SVC and PP could manage to distinguish the styles with the available features, and DT could do it only with broken_gen.

Table 4. Accuracy % on AA task on three authors with 196 StyloMetrix features

Models/Datasets	L_SVC	SVC	DT	LR	PP
Full_orig	58	63	63	68	63
Full_gen	50	50	0	50	50
Broken_orig	53	54	38	52	49
Broken_gen	91	73	73	82	73

Meanwhile, in the full_orig set LR and in the broken_orig, L_SVC tops the list. However, from orig_set, it is evident that the deficit of context in the original set increases the complexity of the problem, and that the extracted lexical features are inadequate in representing the authorial styles. Thus, with an understanding of the effect of the present lexical features on style classification, the main AV framework is developed.

The Siamese MaLSTM network has been popular in sentiment analysis [26] and duplicated question detection [27] among many others. The similarity measure incorporated in MaLSTM is given in **Equation (1)** where h1 and h2 are outputs of hidden states one and two from two different LSTMs, respectively, learning the representation of two different pieces of text simultaneously. **Equation (1)** helps compute the similarity of the tensors using the Manhattan-based similarity measure.

$$e^{(-\|h_1-h_2\|)} \tag{1}$$

In this study, the Manhattan distance itself, as in **Equation (2)**, is employed. It does not serve as a direct similarity measure, yet indirectly, the dissimilarity score passed on to the final dense layer with sigmoid activation would perform the desired computation.

$$\|h_1-h_2\| \tag{2}$$

The exponentiation function in **Equation (1)** is computationally expensive, whereas **Equation (2)** is simpler to compute and robust, and also serves as a proxy for a similarity measure. i.e. the model will tend to associate greater Manhattan distance with less similarity relationship (style-wise) and vice versa. Since the poetry data at hand, when vectorized is still less complicated, **Equation (2)** is chosen to serve the purpose. The framework of the model is illustrated in Figure 1, where 'Lambda' calls the function to compute **Equation (2)**.

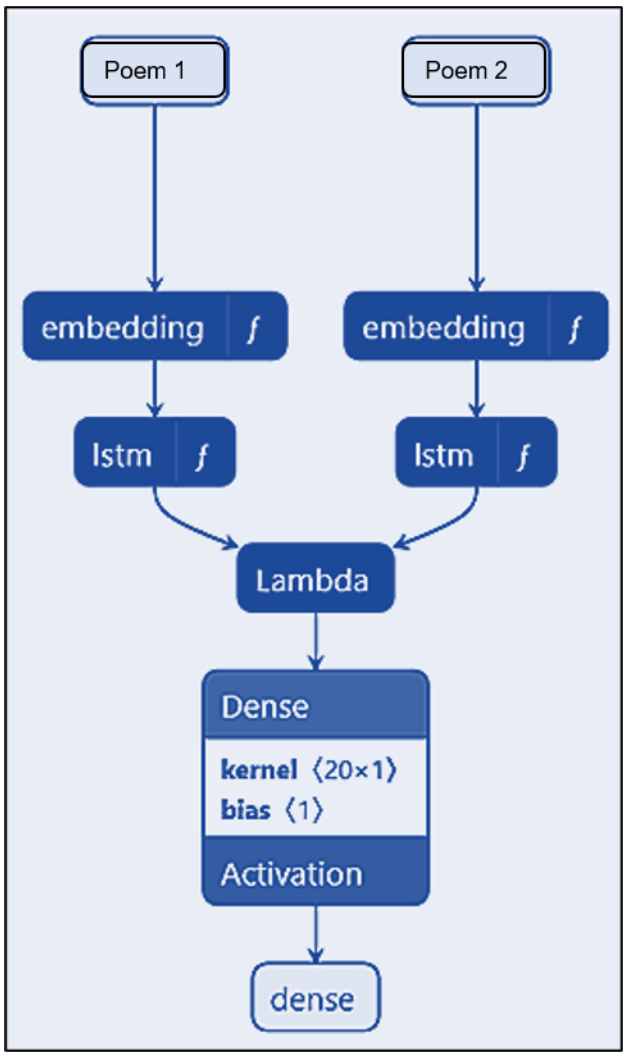

Fig 1. Siamese LSTM-AV Framework employed

As for the embeddings, the Keras default embeddings[28] are employed spanning 100 dimensions. These fixed embeddings, when trained, are subject to adjustment, thereby promoting the learning process and enabling the model to implicitly capture the structural and semantic relationships between the text.

Another aspect of this study shifts the focus on the selection of appropriate feature set i.e. The quantitative lexical features versus syntactic and semantic embeddings. While training the model with Keras embeddings, the LSTM is provided with 20 hidden layers, whereas for StyloMetrix features input, 128 hidden layers are found to be optimal. The other optimal parameters and hyperparameters are tabulated in Table 5.

Table 5. Optimal parameters in Siamese LSTM for AV

Datasets/parameters	Full	Broken
Vocabulary length	4277	4243
Max length	1451	40
Embedding size	100	50
Hidden layer	20	128
Batch size	512	512
Epochs	10	10
Learning rate	0.001	0.001
Optimizer	Adam	Adam
Loss	Binary cross-entropy	Binary cross-entropy
Last layer activation	Sigmoid	Sigmoid
Optimized based on	Accuracy	Accuracy

2.2.1. Preprocessing

Before encoding the text to proceed with the derivation of embeddings, a few preprocessing steps are followed, as listed below:

- Contraction mapping is carried out. E.g. "I've" is converted to "I have" and "we'll" is converted to "we will".
- Possessive apostrophes are replaced with spaces. E.g. "Man's coat" became "Man coat"
- Exclamations and colons are preserved.
- Finally, double spaces are turned into single spaces.

The encoded text is padded to the maximum length. Lastly, it is sent to the input layer, and then Keras randomly initializes the value of its vector. They are updated during the training.

For StyloMetrix embeddings, no preprocessing is involved; however, during training with baseline models, feature selection is performed using an LR model, as feature engineering is generally less critical in deep learning architectures compared to traditional machine learning models, where it often constitutes a significant overhead. The final selected features for baseline evaluation were approximately 55 and 75 for the full and broken sets, respectively. The training and validation losses and accuracies for both the full and broken cases are presented in Figure 2 and Figure 3, respectively.

2.2.2. Training

The training conducted over 10 epochs yielded the training and validation loss values documented in the last epoch, presented in Table 6, where ** represents the statistical significance with p-value<0.01 in a one-tailed t-test (favouring the models with the Keras embeddings) on comparing the results of both the full and broken sets.

Table 6. Losses and accuracies with the Siamese LSTM AV model

Params/Datasets	Train loss	Val loss	Train accuracy	Val accuracy	Test accuracy
Full (with Keras embedding)	0.265	0.263	0.998	1.00	**83****
Broken (with Keras embedding)	0.203	0.252	0.964	0.939	**91****
Full (with StyloMetrix)	0.693	0.693	0.510	0.506	74
Broken (with StyloMetrix)	0.693	0.693	0.505	0.499	53

**statistical significance with p-value<0.01 in a one-tailed t-test against models trained with Keras embeddings

It is clear that StyloMetrix quantitative features are producing inferior results in the Siamese LSTM model when compared to the Keras embeddings, highlighting the importance of syntactic and semantic relationships in poetry, which hold relatively

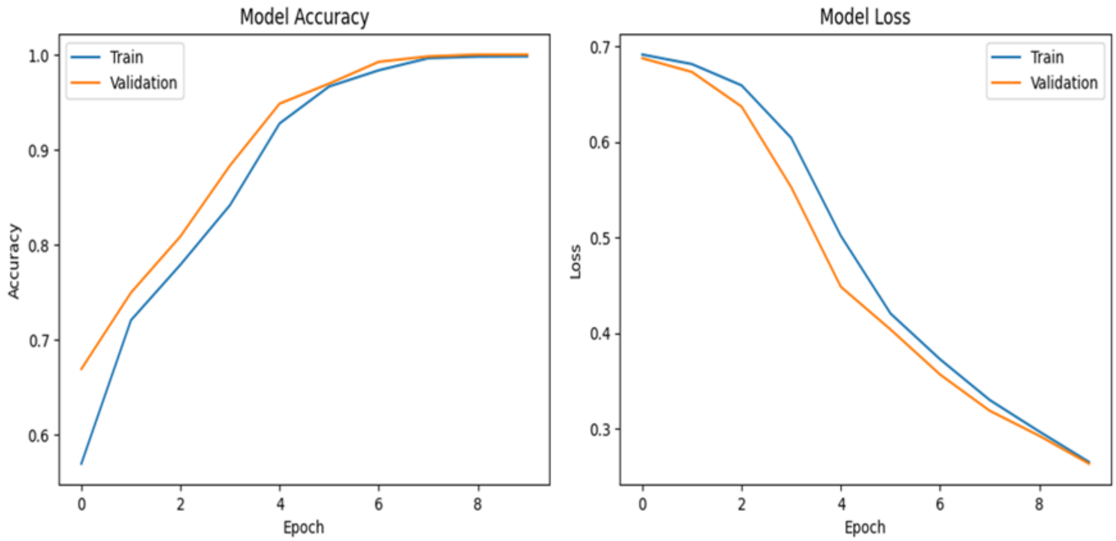

Fig 2. Full set – Siamese LSTM AV model – Accuracy and Loss curves

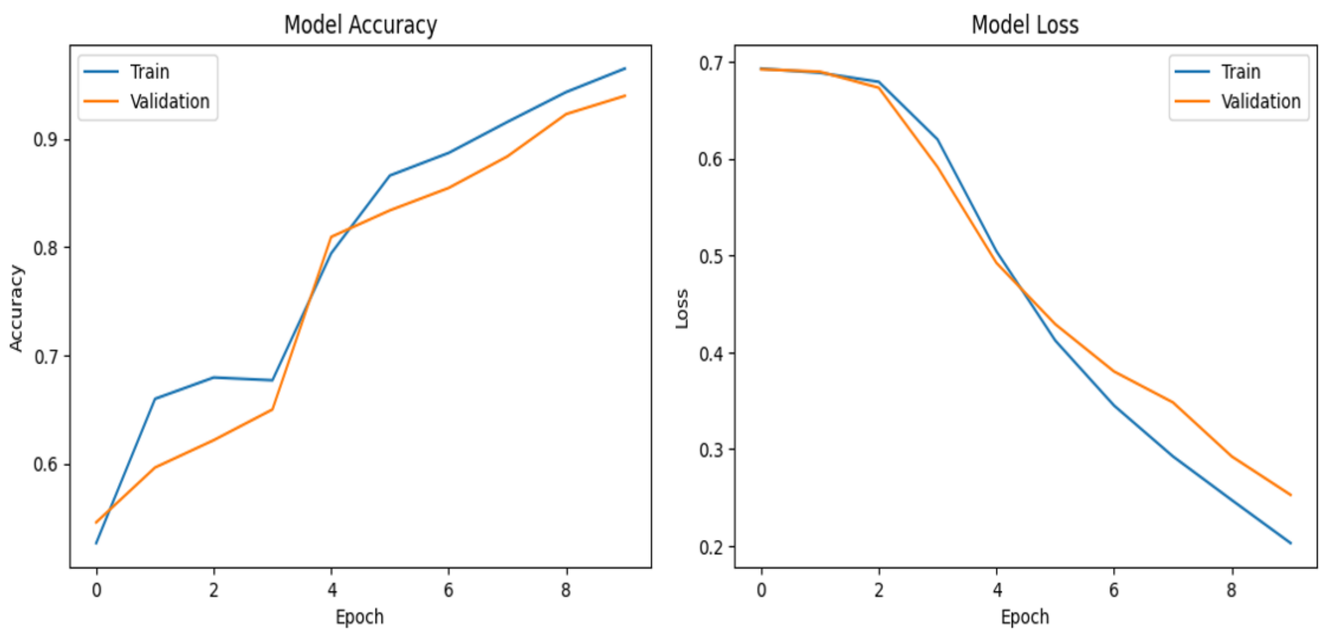

Fig 3. Broken set – Siamese LSTM AV model – Accuracy and Loss curves

higher authorial style information. Meanwhile, training baseline models with StyloMetrix features involved subtracting the LR-selected features from one another and then fed to the classifiers. Their accuracies are presented in Table 7, which highlights the comparative and superior results in full and broken cases, respectively (with StyloMetrix).

Table 7. Accuracy % of selected StyloMetrix with baseline models in the AV task

Datasets/Models	L_SVC	SVC	DT	LR	PP
Full	79	**83**	72	79	58
Broken	**55**	54	53	**55**	52

The highest accuracy score in each dataset is given in bold

3 Results and Discussion

3.1 Feature comparison

Both in Table 6 and Table 7, the highest values in the set are given in bold. From there, it can be deduced that the trainable Keras embeddings outperform StyloMetrix embeddings in both the full and broken cases, which implicitly denote the rich stylistic cues hidden among syntactic and semantic features.

On comparing the StyloMetrix cases in Table 6 and Table 7, the baseline models, however, yield competing results in the full case but superior results in the broken case with processed StyloMetrix, which denotes the presence of richer stylistic details hidden in these processed features and that it could be unveiled only through careful processing and feature selection. Also, the results illustrate that subtracting the selected lexical features from the other could be considered a better compact way of representing stylistic differences of poems for the task of AV (as shown in Table 7). However, the inferior performance of a variant of Siamese LSTM trained with StyloMetrix (in Table 6) could also hint at further hyperparameter tuning required.

3.2 Error Analysis

This subsection aims to answer one of the main questions of this study: whether the model could differentiate the writings of the original and generated authors in both full and broken conditions. Hence, the rearranged dataset introduced in Subsection 2.1 is utilized for this purpose. The author-wise report of the full set is given in Table 8, followed by the report of the broken set in Table 9.

Table 8. Siamese LSTM with the Keras embeddings – Full set – Accuracy comparison against Authors (AV)

Author vs Author	E	A	C
E	100	100	100
A	100	100	100
C	100	100	100
G_E	54	79	96
G_A	82	100	71
G_C	54	96	100

Batch-wise prediction is made on 28 instances of one author, say A against another author say, B (with unique pairs). If both instances are from the same author, all 28 predictions are expected to be 1, otherwise if A≠B, then all 28 are expected to output 0, in case of accurate predictions. The overall accuracy of 80% and 74% are observed in the rearranged dataset of full and broken sets respectively. Table 8 and Table 9 report the model's incapacity in distinguishing between Emily from the generated style of Emily. This might be subtly because of a few unsuspected poems with co-authorship yet to be discarded or generative models' inadequate knowledge about true Emily's style.

Table 9. Accuracy % of selected StyloMetrix with baseline models in the AV task

Author vs Author	E	A	C
E	79	82	96
A	82	96	100
C	96	100	100
G_E	54	64	89
G_A	79	86	43
G_C	57	64	93

Nevertheless, Charlotte's style could be well-differentiated from G_E, while there is yet more scope for improvement in differentiating G_E from Anne. Another notable observation is that the model can differentiate between the original authors very well, particularly in comparison to the original versus the generated authors. Thus, generative models are capable of mimicking the original authorial styles only to a moderate degree. At the core of the algorithm, the Manhattan distance in **Equation (2)** inversely teaches the similarity, serving the purpose well. On the whole, full poems have better scores compared to broken poems in the rearranged set. Finally, with all six authors (in orig_set and gen_set), the LR-based feature selected–StyloMetrix features is employed in the AA task (as reported in Table 10) with baseline models to ensure whether this feature selection method, which succeeded in AV, also boosts the accuracy in AA.

Table 10. Baseline Accuracy % in AA with selected StyloMetrix features on six authors

Models/Datasets	L_SVC	SVC	DT	LR	PP
Full	79	**83**	72	79	58
Broken	**55**	54	53	**55**	52

The highest accuracy score in each dataset is given in bold

4 Conclusion

This study attempts to unveil answers to attributing and verifying poetry authorship of the challenging case of the Brontë sisters, whose writings are per se said to have overlapping commonalities, hence with a higher challenge of classification; on taking this one step further, generative AI poems mimicking their styles are also procured and included in the test sets. Adding to the complexity, poem excerpt-authorship verification is also carried out, and the best representative feature (of style) among syntactic, semantic, and lexical aspects is explored. The latter task could be particularly helpful when a part of an author's writing is only available, in literary analysis or forensics, which needs to be verified. It is found that carefully processed lexical features could achieve better and comparative results in the broken/excerpts set and the full poems set, respectively. The research reports superior performance of the Siamese LSTM built with Manhattan distance to indirectly teach stylistic similarities between poem pairs to the model, using trainable embeddings that outperformed quantitative lexical features. The errors of the model are analyzed to provide a better understanding of the behaviour of the model. Notably, it is found that the Keras embeddings performed superior to the selected style-related features (especially, in the case of broken context of poems), and the generative models showed elementary mimicking of the authorial styles in poetry generation. This research enhances the field of authorship studies and illustrates the potential of NLP in literary analysis, emphasizing its application in connecting computational science and the humanities. Working towards this direction of research could also provide valuable solutions to the problem of authorship (including, in the domains of fragmentology, philology and forensics) and AI-plagiarism in academia. Although this work pioneered in exploring interesting NLP questions with a greater literary importance, the generalizability of the findings is limited by the small size of the dataset and the non-exhaustive list of style-related features employed in the comparison. In the future, authors with a higher degree of stylistic commonalities and prolific works can be verified using transformer models.

5 Acknowledgements

The research was **funded by the University Grants Commission (UGC), India**, under the Junior Research Fellowship (JRF) scheme.

References

1) Polishchuk N. The Hermeneutical circle of Literary work: The Brontës Context. 2024. https://doi.org/10.30525/978-9934-26-436-8-15.
2) Newman H. Some Common Features in the Brontë Sisters' Novels. *Brontë Studies*. 2019;44(2):186–203. https://doi.org/10.1080/14748932.2019.1567166.
3) Bronte A, Bronte C, Bronte EJ. The Bronte Sisters. Routledge. 2012. https://doi.org/10.4324/9780203446058.
4) Lingesh B, Monesh MS, Kannaiah SK, Pandi SS. Detecting AI Face Fraud Detection Using CNN Based Deep Learning Algorithm. *Proceedings of the Second International Conference on Advances in Information Technology (ICAIT)*. 2024;p. 1–7.
5) Khoo B, Phan RW, Lim C. Deepfake attribution: On the source identification of artificially generated images. *WIREs Data Mining and Knowledge Discovery*. 2021;12(3). https://doi.org/10.1002/widm.1438.
6) Chen B, Li T, Ding W. Detecting deepfake videos based on spatiotemporal attention and convolutional LSTM. *Information Sciences*. 2022;601:58–70. https://doi.org/10.1016/j.ins.2022.04.014.
7) Uchendu A, Le T, Lee D. TOPFORMER: Topology-Aware Authorship Attribution of Deepfake Texts with Diverse Writing Styles. *arXivorg*. 2023. https://arxiv.org/abs/2309.12934.
8) Aburass S, Rumman MA. Authenticity in authorship: the Writer's Integrity framework for verifying human-generated text. *Ethics and Information Technology*. 2024;26(3). https://doi.org/10.1007/s10676-024-09797-z.
9) Soos C, Haroutunian L. On the Question of Authorship in Large Language Models (LLMs). *NASKO*. 2023;9(1):1–17. https://doi.org/10.7152/nasko.v9i1.16299.
10) Köbis N, Mossink LD. Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Computers in Human Behavior*. 2021;114(114).
11) Ivanov L, Perez F. Authorship Attribution of English Poetry using Sentiment Analysis. *The International FLAIRS Conference Proceedings*. 2024;37(1). https://doi.org/10.32473/flairs.37.1.135272.
12) Abedzadeh A, Ramezani R, Fatemi A. A Weighted TF-IDF-based Approach for Authorship Attribution. *11th International Conference on Computer Engineering and Knowledge (ICCKE)*. 2021;p. 188–193. https://doi.org/10.1109/iccke54056.2021.9721474.
13) Forstall C, Scheirer W. Features from Frequency: Authorship and Stylistic Analysis Using Repetitive Sound. *Journal of the Chicago Colloquium on Digital Humanities and Computer Science*. 2024;1(2):1–23. https://doi.org/10.6082/M1TM789G.

14) Anggriawan E, Farid F, Fitri R. Sentences Similarity Detection in Indonesian Poetry Comparison using Siamese MALSTM. *ICIC Express Letters*. 2023;17(4):389–396. https://doi.org/10.24507/icicel.17.04.389.
15) Embarcadero-Ruiz D, Gómez-Adorno H, Embarcadero-Ruiz A, Sierra G. Graph-Based Siamese Network for Authorship Verification. *Mathematics*. 2022;10(2):277–277. https://doi.org/10.3390/math10020277.
16) Malfait O, Demoor M. Sibling Collaboration and Literary After-life: The Case of the Brontës. *Brontë Studies*. 2015;40(3):187–200. https://doi.org/10.1179/1474893215z.000000000148.
17) Melani L. Emily Bronte's Poems. 2024. Available from: https://academic.brooklyn.cuny.edu/english/melani/novel_19c/wuthering/poetry.html.
18) Brownson SB. Emily Brontë. 2019. Available from: https://www.poetryfoundation.org/poets/emily-bronte.
19) Okulska I, Stetsenko D, Kołos A, Karlińska A, Głąbińska K, Nowakowski A. StyloMetrix: An Open-Source Multilingual Tool for Representing Stylometric Vectors. *arXivorg*. 2023. https://arxiv.org/abs/2309.12810.
20) Google. The Project Gutenberg eBook of Poems, by Currer, Ellis, and Acton Bell (Brontë Sisters). 2025. Available from: https://www.gutenberg.org/files/1019/1019-h/1019-h.htm.
21) Inc DA. AI Poem Generator. Available from: https://deepai.org/chat/poet.
22) OpenAI. ChatGPT. 2025. Available from: https://chatgpt.com/.
23) Microsoft. Copilot | Microsoft Edge. 2025. Available from: https://www.microsoft.com/en-in/edge/features/copilot.
24) Anthropic. Claude. 2025. Available from: https://claude.ai/.
25) AI J. Julius AI | Your AI Data Analyst. 2025. Available from: https://julius.ai/ai.
26) Srinidhi H, Siddesh GM, Srinivasa KG. A hybrid model using MaLSTM based on recurrent neural networks with support vector machines for sentiment analysis. *Engineering and Applied Science Research*. 2020;47(3):232–240.
27) Altamimi A, Umer M, Hanif D, Alsubai S, Kim TH, Ashraf I. Employing Siamese MaLSTM Model and ELMO Word Embedding for Quora Duplicate Questions Detection. *IEEE access*. 2024;12. https://doi.org/10.1109/access.2024.3367978.
28) Gulli A, Pal S. Deep Learning with Keras. Packt Publishing. 2017.