Submitted: 05/04/2026 - Accepted: 06/05/2026 - Published: 28/06/2026

Automating Artistry: The Social Impact of Generative AI on Creative Labor in Theatre

Scarlett Sun

New York University, 70 Washington Square S, New York, NY 10012
ys3101@nyu.edu

DOI: 10.26417/s290dh78

Abstract

The paper examines the use of generative AI in the production of stage plays and its impact on classical forms of drama. As AI takes on a greater role in creative processes, one asks if screenplays written by machines can capture the structural and emotional dimensions of theatrical narrative. To solve this problem, we propose a novel model named Dramatic Structure-Aware Multi-Agent Generative Framework (DSA-MAG). The model consists of a transformer-based encoder, multi-agent narrative generation, dramatic structure analyzer, emotional coherence module, and character consistency optimizer. In this study, we use the dataset of Movie Scripts Corpus collected from Kaggle. The dataset contains 1200 script segments from various genres. The segments were vectorized using a HashingVectorizer with 1500 features. The experimental results showed that the DSA-MAG model achieved the best accuracy of 25.40%. The metrics of precision 0.0788, recall 0.2540 and F1-score 0.1179 further validated the superiority of the model in multi-class genre categorization. Therefore, these results are consistent with the claim that the more unified and more fundamentally perceptive approaches to script generation with structured multi-agent AI frameworks improve AI-assisted creative writing as well as research on theatrical drama.

Keywords: Creative labor automation, AI literacy, higher education pedagogy, cultural production, human-AI collaboration, computational creativity, authorship transformation

1. Introduction

“AI’s rapid development has had a massive impact on creative industries, especially those that are narrative-heavy, like theatre, screenwriting and digital storytelling. “Historically, dramatic writing and performance have relied upon established human-centric practices derived from literary traditions, performance philosophy and dramaturgical conventions. “The emergence of generative AI systems has transformed the collaborative process of creation. Machines are no longer simply tools for the production of artefacts but are now active participants in the construction of narrative skeletons. This has renewed scholarly interest in the impact of AI on script writing, play performance and the underlying principles of dramatic composition.

However, those who work in the creative fields agree that AI is revolutionary and disruptive. The use of AI technologies in creative workflows is becoming more mainstream, such as supporting automation, augmentation and in some cases complete generation of creative output (Anantrasirichai and Bull 2022). This raises important questions around authorship, originality and the changing relationship between human creativity and machine intelligence for storytelling and performing arts . The impact of AI is already being seen in the production processes and conceptualisations behind story development in theatre and scriptwriting.

Recently, there has been research in theatre studies exploring the use of AI in the creation or co-creation of stage shows. AI-enabled technologies are changing the theatre production in terms of new types of digital and hybrid performance environments (Horváth, 2025). Such changes show that theatre is not confined to the traditional means of expression on the stage. And now it’s going into settings with technology where computers can change scripts and characters and even how the performances are done. Ren (2024) also discusses the possibility of AI writing plays and shows how AI models can be used to support dramatic storytelling, which is a significant shift from the use of AI as a backstage tool to the use of AI as an active participant in the drama.

Besides using AI for creative writing, another important aspect to consider is the development of AI literacy among communicators such that they work together with the use of AI without compromising the creativity and intentions behind stories (Cardon et al., 2023). Nonetheless, it is unclear whether AI-based works can be regarded as creative ones or just a blend of patterns resulting in an artificial outcome (Chakrabarty et al., 2024).

Moreover, cases can be pointed out where the application of AI was made during experimental co-creations in the theater. For instance, Branch et al. (2024) study the case of utilizing dialogue LLMs in improvisational theater where AI is used as a means of co-creation of narratives along with human actors. As it comes from these studies, the application of AI in such experiments brings unpredictability and variability into conversation and allows for more possibilities for developing a story. On the other

hand, Branch & Mirowski (2024) focus on the consequences of applying AI in theatrical performances.

In relation to script writing, it seems that AI does not behave more as an automatic writer but more as a collaborator in the writing of scripts. As stated by Cake (2025), AI may assist the writer in the creation of the plot, dialogues, and characterization. This is just one of the many manifestations of the current trend in distributed human authorship wherein writing through collaboration between man and machine becomes the mode. Nonetheless, Hoyer & Frühmorgen (2025) studied the systematic use of AI in scriptwriting.

Nonetheless, there are scholarly investigations that discuss how the replication of the narrative form is achieved through artificial intelligence tools irrespective of the practicality of such systems. For example, Kabashkin et al. (2025) suggest that the narratives created through artificial intelligence share some similarities with the traditional narrative form, which suggests that the algorithms used in the development of artificial intelligence learn from the prototype of narratives and imitate them. Such findings are important in the field of theatre and performance as they raise issues about the extent to which narratives created with the aid of artificial intelligence meet the norms of traditional narrative structures or present novel forms of narrative construction.

As described by the results of the research carried out by Mirowski et al. (2023), it was possible to achieve a perfect combination of artificial intelligence technologies with writing scripts for films or other performances. Experts in this field determine whether the scripts created with the help of AI technologies are of good quality. Nevertheless, even if the script generated by AI technologies seems to be great and consists of lots of excellent ideas, it still requires some human involvement.

These recommendations can also mean that the entire industry will face anticipated change in some way. As stated by Palanimurugan et al. (2025), the issue of change in the industry relates to the influence of AI in that context. The mentioned influence includes a transition from individual to collaborative authorship between humans and computers involved. In the review article on the effect of generative AI on creativity completed by Tsao et al. (2025), it was discovered that there was an increase in the application of AI in the creative process despite challenges in the area of authorship and authenticity.

Also, the new theories that have recently been developed in order to apply AI technology in the creative sector also note that the adoption of AI technology will depend on the perception and integration into the process (Tsao et al., 2025). Application of artificial intelligence technology in the creation of scripts for films and TV shows has increased over time as a means of boosting creativity and writing (Xu & Xie, 2026).

On the other hand, it should also be noted that its use is not limited to the theatre only, but also in films and multi-media. In this context, the use of AI in the creation of screenplays in movies and TV shows has been highlighted by Xu & Xie (2026). The contribution of the computer in constructing a sensible plot in the domain of entertainment has been stressed. Dayo et al. (2023) have also discussed the influence of AI technology on re-writing of media communication.

In short, the literature says that AI is changing the world of scriptwriting and theatre in a big way. There is a constant change in research from seeing AI as a helpful tool to seeing it as an active co-creator in making stories. This change brings up important considerations about who wrote something, how creative it is, and how it is structured, while also offering up new ways for theatre and storytelling to be creative.

In this paper, we contribute to our understanding of the societal and pedagogical implications of the technology-mediated creative practice by contextualizing the adoption of generative AI in scriptwriting in the wider social science debate on creative labor, digital transformation and the changing nature of authorship in cultural industries. The study discusses how AI frameworks are changing the scriptwriting process.

1.1. Contribution of the Study

The work makes several novel contributions at the intersection of artificial intelligence and theatrical dramaturgy. We propose the Dramatic Structure-Aware Multi-Agent Generative (DSA-MAG) Framework, a novel multi-module architecture for generating stage drama scripts adhering to established five-stage dramatic structures. We introduce an Emotional Coherence Module that simulates scene-level emotional intensity via hyperbolic tangent activation over hidden state representations to ensure smooth tonal transitions between scenes of the generated script. We devise a Character Consistency Optimizer, which utilizes memory embeddings and a divergence-minimizing loss function to uphold consistent personality traits and dialogue patterns across all scenes. We provide an empirical evaluation of the DSA-MAG framework, comparing its performance with three baseline models—NLP Scriptwriting Tools, Conceptual AI Theatre Models, and Archetype Learning Models—using a real-world movie script corpus, allowing direct performance comparison on a shared benchmark. As a result, the research demonstrates that structured multi-agent generative architectures outperform single-model generative approaches in theatrical script classification tasks, thus providing practical evidence to the broader discussion on AI-assisted creative writing.

1.2. Research Question

In this context, this study asks the following research questions:

> RQ1: How does the use of generative AI in scriptwriting change the nature of creative labour and authorship in theatrical production?

RQ2: To what extent can multi-agent AI architectures reproduce the structural and social characteristics of human-authored dramatic scripts?

RQ3: What are the wider social implications of AI-enabled creativity for cultural industries and creative labour?

2. Related Work

The latest research works on the application of AI in creative industries are shifting their focus toward not only the implementation of AI technology but also the implications of the AI-produced content on ethics, structure, and industry. When it comes to the sphere of theatre, screenplay writing, and narration, the current research has moved forward from discussing the subject in theory to its practical considerations, especially with regards to the influence of AI on the process of authorship, storytelling, and narration construction. The following section highlights the main points of references 11-20.

However, the problems regarding ethics and legality associated with the application of AI technology in creative areas should be examined. For instance, the article published by Kavitha (2024) examines ethical and legal aspects regarding copyright infringements in the film industry due to the use of AI technology. In this regard, the author focuses on the problem of copyright infringement associated with the shift in the method of screenplay writing from conventional approaches based on human creativity to new forms involving technological means. Several questions related to the copyright of this content created with the help of AI technology arise due to the ethical and legal considerations involved.

The research carried out by Zakopoulos et al. (2023) is focused on the use of digital technology to educate drama and theatre students and the role of digital technology enhanced by AI in contributing to the awareness of students regarding various aspects of sustainability. More precisely, according to the authors, digital technology combined with practical work in theatre helps create the context in which students can develop transdisciplinary thinking. However, the present paper touches upon some crucial points in relation to the impact of digitalization on theatre and preparations of students to perform in terms of AI, despite the absence of any connection between the AI-based scriptwriting and this paper.

It has already been determined in prior researches that for the adequate analysis and implementation of the output generated by the AI generators into their work practices, individuals should possess certain literacy. Moreover, there are doubts about the distinctiveness of the products generated by the AI generators as it has been stated that, despite being capable of imitation of the creative process, they do not possess intentionality and only reproduce the narrative sequence learned (Cardon et al., 2023; Chakrabarty et al., 2024).

AI generation is used not only for commercial purposes but also in educational areas in particular in relation to the training process of creating films and screenwriting

where the objective of using AI generation is to facilitate the development of stories and aid students in experimenting with story types (Lo Duca & Rotelli, 2026). Furthermore, the advent of generation AI also gave rise to the new definition of the concept of authorship in the screen industry since authorship is now shared by both humans and AI (Palanimurugan et al., 2025).

To be more exact, Kabashkin et al. (2025) pose a question related to the problem of story modeling using artificial intelligence, noting that machine learning is predicated on the archetypal culture model of storytelling. They argue that since AI models have learned from many stories before, they tend to stick to the traditional pattern of storytelling, such as the hero's journey, for instance. The importance of this article is in its contribution to screenplay writing in that, as noted above, story models created by means of artificial intelligence do not explore novel storytelling techniques but rather rely on those that already exist in culture.

Taking a closer look at the topic under discussion, we have found that Yang et al. (2025) have created a new AI model referred to as Co-DIRECT with the help of which the scripts of interactive drama are composed. They point out that several AI agents could be involved in the process of developing interactive scripts, each responsible for different parts of the story, including plot development, conversations, and characters validation. Overall, this is a breakthrough in terms of a design approach because of shifting from one model to a multi-agent system like people do.

Innovations in creativity thanks to artificial intelligence can also be discussed within the development of new media types such as micro-short dramas. In accordance with the article presented by Wang (2025), the impact of the change in the production process of micro-short dramas with the help of artificial intelligence is evaluated, and attention is drawn to the role of generative models in compression and pacing. In conclusion, artificial intelligence does not only enable the creation of a story but helps the audience consume the story more conveniently using their smartphones.

Furthermore, Branch & Mirowski (2024) have examined artificial theatre of absurdity and how much of a deviation from the traditional theatres' normative story lines artificial intelligence (AI) stories are. As per the results derived from the findings, artificial intelligence possesses the capability of creating stories that are fragmented and non-linear in their nature, akin to those found in the theatre of absurdity. Such changes have raised many questions regarding whether the stories created by artificial intelligence are mere copies of the traditional theatres' stories or if a whole new approach is being taken into account. Moreover, there has been an increased use of artificial intelligence in making movie and television stories.

As per the work by Ren (2024), the implementation of artificial intelligence in the creation of stage plays will be examined. In addition, computer-written scripts could be turned into stage plays. It is also an extension of the topic under discussion, i.e., the use of artificial intelligence in performance art. The strengths and weaknesses in using artificial intelligence in the development of stage plays have been illustrated in

the work done by Ren (2024). From the work by Ren (2024), it can be seen that AI is able to create a coherent script; however, human input is needed for emotions.

Latif et al. (2025) investigate another aspect associated with the influence of AI in the process of making decisions regarding the performance design and performance itself. According to the results presented by Latif et al. (2025), AI can help to evaluate various scenarios, utilize the performance space to the maximum extent, and anticipate the reaction of the audience towards a specific scenario. Thus, AI can affect multiple aspects of theatre.

Inderawati et al. (2024) and Zakopoulos et al. (2023) have discussed digital and AI-augmented techniques of theatre mentoring. This emphasizes the idea that technology in the performing arts is not just for professional productions, but also has an impact on schools and new ways of teaching. The literature 11-20 shows that AI in theatre and scriptwriting works on different levels such as legal and ethical issues, narrative structure modelling, script development with various agents, educational applications and performance production systems. The growing issue is about the difference between AI's ability to generate coherent story plots and its inability to express emotional and creative depth. In parallel, advances in multi-agent systems and generative frameworks suggest AI is developing its own potential for collaborative creation. Table 1 in Appendix summarizes the main literature in terms of research objectives, techniques and key findings.

2.1. Research Gap

Although some previous studies have been done on the use of artificial intelligence in script writing and theatrical production, the existing research is still far from sufficient in many aspects. Firstly, most of the research takes AI as an all-purpose creative writing tool without considering the particular structural requirements of stage drama, such as the five-act dramatic structure (exposition, rising action, climax, falling action and resolution). Secondly, most generative AI models evaluated in previous studies do not possess the ability to model emotional coherence, resulting in scripts with inconsistent or emotionally flat transitions between scenes. Thirdly, the problem of character consistency in multi-scene dramatic scripts remains inadequately solved; existing systems often generate characters with randomly varying behavioral traits and dialogue patterns across scenes. Fourth, the absence of multi-agent systems for generating theatrical scripts leads to a single model controlling the plot, dialogue, and character development, which may result in an unbalanced narrative. Also, the current literature presents a striking lack of empirical benchmarking of AI-generated scripts against established dramatic structure criteria, particularly when using authentic theatrical datasets. This study directly addresses these shortcomings by proposing the DSA-MAG framework, a unified multi-agent generative system that combines dramaturgy-aware evaluation, emotional flow modeling, and character memory optimization.

2.2. Creative Labor Theory and AI-Supported Scriptwriting

In this article we draw on creative labour theory (David Hesmondhalgh and Sarah Baker) and Actor-Network Theory (Bruno Latour) to explore the implications of generative AI for theatre scriptwriting and creative production. It examines the degree to which AI is able to automate narrative functions such as writing dialogue, developing plots, and creating characters, thus disrupting the traditional division of labor between playwrights, editors and directors. The paper also considers whether AI enables the de-skilling of creative professionals by reducing the need for traditional writing skills or the re-skilling of creative professionals by requiring new skills such as prompt engineering, AI editing and co-authoring. Using actor-network theory, the paper also considers AI to be an active actor within the creative ecosystem of human writers, algorithms, datasets and digital platforms that together determine authorship, originality and creative decisions in the contemporary theatrical landscape.

The increasing adoption of generative AI in creative industries has engendered widespread academic discussion about the evolving nature of creative labour . The automation of creative work, the authors argue, does not replace human agency, but reconfigures the conditions of creative work, raising important questions of skill, authorship and professional identity . In the context of education, the adoption of AI tools into the curriculum of drama and creative writing has been framed as an opportunity and a challenge . AI assisted writing environments have been shown to encourage student experimentation with narrative forms, but also require the development of new forms of critical AI literacy . These perspectives emphasize the importance of situating AI scriptwriting frameworks as sociotechnical artifacts within larger systems of education, labour and cultural production.

3. Dataset Description

This study used the Movie Scripts Corpus dataset. This dataset is available on Kaggle (dataset id: gufukuro/movie-scripts-corpus). The corpus contains raw script files, stored as plain-text (.txt) documents in subdirectories. The name of each subdirectory corresponds to the genre or the title of the film and thus labels the script segments contained in it. The dataset covers a wide range of dramatic scripts, including romance, thriller, comedy, drama and science fiction genres, thus providing a suitable basis for the evaluation of multi-class text classification models in the context of screenplay analysis. For the experiments, a random sample of 1,200 script segments was taken, using a fixed random seed (random_state=42) to ensure the reproducibility of the results. To deal with class imbalance in both training and testing, classes with fewer than two instances were removed. Each script instance includes the complete, raw text of the script segment (under the ‘script’ column) and its associated genre/film label (under the ‘genre’ column). Feature extraction was done using HashingVectorizer, with the number of hashed features set to 1500, and the removal of English stop words and alt_sign set to False (non-negative feature values). The raw texts were converted to sparse numerical representations, which are

well suited for machine learning classifiers. The data was split into an 80:20 train and test set, a training set to fit models and a test set to assess model performance.

4. Methodology

The research methodology followed a quantitative experimental research design supplemented with computational modeling and machine learning assessment. A multi-class text classification framework was employed to assess the efficacy of the proposed Dramatic Structure-Aware Multi-Agent Generative (DSA-MAG) framework vis-à-vis three baseline models. Data collection was based on the publicly available Movie Scripts Corpus from Kaggle with 1,200 script segments. Feature extraction was performed using a HashingVectorizer and model evaluation was performed using standard classification metrics such as accuracy, precision, recall and F1-score on an 80:20 train-test split. This methodological approach was taken to provide reproducible, quantifiable evidence of the efficacy of DSA-MAG for genre classification tasks relevant to the analysis of dramatic scripts.

Proposed Model: Dramatic Structure Aware Multi-Agent Generative Framework (DSA-MAG)

The proposed system, Dramatic Structure-Aware Multi-Agent Generative Framework (DSA-MAG), aims to generate stage drama scripts by retaining the traditional dramatic frameworks and enhancing narrative coherence. Current generative AI systems tend to produce scripts that do not take theatrical structure into account. This results in fragmented dialogues, inconsistent character development, and a lack of emotional resonance. To address this issue, we propose a model that combines multi-agent narrative generation, evaluation of dramatic structure and optimization for emotional consistency. The system architecture comprises a transformer-based encoder, a multi-agent narrative generator, a dramatic structure analyzer, an emotional coherence module and a character consistency optimizer. These modules are sequenced in a specific order to generate a structured stage drama script that conforms with established theatrical conventions.

The narrative prompt describes the theme, characters, setting, and desired dramatic effect of the script. The prompt is tokenized and converted to contextual embeddings using a transformer encoder. The input is defined as:

$$X = \{x_1, x_2. x_3, \ldots, x_n\} \quad (1)$$

The contextual representation is achieved through encoder mapping:

$$E = f_{enc}(X) \quad (2)$$

This embedding encodes the semantic meaning, narrative tone and character relationships needed for script generation.

4.1. Multi-Agent Narrative Generation

The encoded data is then passed to three specialized agents responsible for plot developments, dialogue and character development respectively. Each agent independently generates narrative components and these are then merged into a unified script. The narrative output is mathematically expressed as equation (3):

$$N = \alpha P + \beta D + \gamma C \quad (3)$$

where P represents the plot structure, D is the dialogue generation, and C is the character modeling. The coefficients α, β and γ determine the weight of each component in the output, giving flexibility to the story-telling.

4.2. Dramatic Structure Analyzer

To preserve the classical stage drama composition, the generated narrative is evaluated by a dramatic structure analyzer. The model checks if the script passes through five necessary stages including exposition, rising action, climax, falling action, and resolution. Each stage is scored independently and aggregated into a structural coherence score:

$$S = \sum_{i=1}^{5} w_i s_i \quad (4)$$

This module ensures that the generated script follows the theatrical storytelling logic instead of random narrative progression.

4.3. Emotional Coherence Modelling

The proposed framework includes an emotional flow module that is intended to keep the transitions between scenes smooth. The module achieves this by modelling the emotional intensity throughout the generated script to avoid abrupt tonal shifts. The emotional state at each narrative timestep is calculated as in equation (5):

$$E_t = tanh(W_e h_t + b_e) \quad (5)$$

Therefore, this mechanism enables the model to simulate emotional escalation towards the climax, followed by a gradual resolution in later scenes.

4.4. Character Consistency Optimizer

The key component of stage drama is character consistency. The model we present uses character memory embeddings that are expected to follow personality, conversational and behavioral traits along the script as a whole. The character memory representation is obtained from equations (6) and (7):

$$M_c = \frac{1}{T} \sum_{t=1}^{T} h_t^c \quad (6)$$

To ensure consistency, a loss function is imposed to minimize the difference between the current and previous character states:

$$L_c = \|M_c - h_t^c\|^2 \qquad (7)$$

This method ensures that characters act consistently across different scenes.

4.5. Narrative Coherence Optimization

The structural score, emotional continuity and character consistency are combined as an optimization objective. The overall narrative coherence score is defined as equation (8):

$$C_{total} = \lambda_1 S + \lambda_2 E + \lambda_3 (1 - L_c) \qquad (8)$$

The final stage drama script is generated by maximizing the overall coherence score of equation (9):

$$Y = arg\ max\ C_{total}(N) \qquad (9)$$

The optimized story is a well-structured stage drama script, with scene signs, character dialogues and dramatic turns. The proposed DSA-MAG framework guarantees the generated scripts follow the norms of theatrics storytelling, keep emotional impact and character consistency.

Multi-agent generation with dramaturgical constraints results in improved narrative coherence than traditional generative AI models. This framework allows playwrights to collaboratively write scripts and influence the narrative through their input while also harnessing the creative potential of AI-generated content, making it suitable for stage drama production and theatrical content generation. The model architecture is shown in Figure 1 in Appendix.

Algorithm 1: Dramatic Structure Aware Multi-Agent Generation

Input: Script Prompt X

Output: Generated Stage Drama Script Y

1: Begin with narrative prompt X.

2: Transform input using transformer encoder

3: Generate plot using plot agent

4: Generate dialogue using dialogue agent

5: Generate character arcs using character agent

6: Combine all outputs to obtain narrative, N

7: Calculate dramatic structure score, S

8: Calculate emotional flow, E

9: Calculate character consistency loss, Lc 1

0: Calculate overall coherence score, Ctotal

11: Coherence based structural improvement of story

12: Generate final script, Y

13: Return Y.

5. Result and Discussion

The experiments were run in python on a Jupyter Notebook. For the experiments we used the following libraries and software: KaggleHub, was used to get datasets from the Kaggle platform. Pandas and NumPy, were used for data manipulation and numerical calculations. Scikit-learn, was used for the machine learning pipeline components, including HashingVectorizer for text feature extraction, train_test_split for splitting the dataset and four classification models: MultinomialNB, SVC, LogisticRegression and RandomForestClassifier. We also obtained evaluation metrics such as accuracy_score, precision_score, recall_score, f1_score and roc_auc_score from Scikit-learn. Matplotlib, was used to generate static charts. Plotly Express and Plotly Graph Objects, were used to generate interactive and three-dimensional visualizations.

Figure 2 in Appendix. Script Length vs. Genre (Scatter plot) for 1200 script segments. Each point on the graph corresponds to a script entry. The x-axis is script length, in characters, and the y-axis is genre or film title. The analysis reveals a large variation in length of scripts among genres, implying different levels of textual density in different types of drama. The shorter scripts in some genres may be components of episodic or scene-based works, while the longer scripts are typically full-length feature scripts. This finding implies the difficulties in genre classification, as script length is not a reliable criterion to differentiate between types of dramatic works.

Figure 3 in Appendix displays a 3D scatter plot for visualizing the distribution of script segments along the dimensions of characters length, word count, and a second length dimension. The data points are colored according to their genre label. The x-axis denotes the character length of the script of each segment, the y-axis denotes the total number of words and the z-axis denotes the character length dimension for volumetric spatial separation. The color coding assigns each data point to its genre or film title category. The figure shows a strong linear correlation between character length and word count as expected, while the color distribution over genres indicates that many genres are found in overlapping regions of the 3D feature space. The overlap shows the difficulty of properly categorizing genres based solely on simple textual data. This is the motivation towards the use of content-based feature extraction techniques such as the hashing vectorization.

5.1. Performance Analysis

Table 2: Performance Analysis

Model	Accuracy	Precision	Recall	F1 Score
NLP Scriptwriting Tools [7]	0.2302	0.1443	0.2302	0.1572
Conceptual AI Theatre Models [3]	0.2381	0.1187	0.2381	0.1289
Archetype Learning Models [10]	0.2302	0.0704	0.2302	0.1029
Proposed DSA-MAG	0.2540	0.0788	0.2540	0.1179

Figure 4 in Appendix is a line chart comparing the classification accuracies of four models on dramatic script genres. The horizontal axis shows the models and the vertical axis shows the accuracy computed on the test set. The data in the chart shows gradual improvement of accuracy from the baseline models to the DSA-MAG framework. The NLP Scriptwriting Tools and Archetype Learning Models achieve an accuracy of around 23.02%, while the Conceptual AI Theatre Models achieve 23.81%. The DSA-MAG model, however, achieved the best accuracy of 25.40% and showed a performance advantage over the single model approaches.

Figure 5 in Appendix depicts the filled area line chart. The subsequent analysis contrasts the efficiency of four different models, namely NLP Scriptwriting Tools, Conceptual AI Theatre Models, Archetype Learning Models and the Proposed DSA-MAG. The data reveals a decline in accuracy from Natural Language Processing Scriptwriting Tools (around 0.144) to Conceptual AI Theatre Models (around 0.091). The study results indicated a value of 0.119, followed by Archetype Learning Models, which were approximately. The DSA-MAG model exhibited a marginal enhancement to approximately 0.079. This means that probabilistic models like the NLP Scriptwriting Tools have a better precision when focusing on less genre classes while the DSA-MAG and Archetype Learning Models have a lower precision because their predictions are more distributed. This can lead to a lower accuracy but also to a better coverage and completeness of the information.

Figure 6 in Appendix shows the lollipop plot comparing the recall values of four models and their performance. The models are denoted by vertical lines as baselines, and the weighted recall values are represented by circular markers. It can be seen from the plot that three models, i.e., NLP Scriptwriting Tools, Archetype Learning Models, and Conceptual AI Theatre Models, have similar recall scores, from 0.230 to 0.238. However, the Proposed DSA-MAG model has the highest recall value, around

0.254. This indicates that the initial models are equally accurate in genre identification, but the DSA-MAG model is better in terms of identifying more genuine instances of genre in the Movie Scripts Corpus, indicating the benefits of using ensemble-based generation techniques to improve the representation of less frequent genres.

Figure 7 in Appendix presents a step line chart that compares the F1 Scores of four different models. The chart shows a gradual decrease in the F1-score from left to right. The NLP Scriptwriting Tools [7] achieved the highest score, about 0.157. This is followed by a sharp drop to Conceptual AI Theatre Models [3] which had a score of 0.129. The score then drops significantly to Archetype Learning Models [10] which had the lowest score of 0.103. The Proposed DSA-MAG sees a slight increase to 0.118 which shows a small but notable improvement over Archetype Learning Models [10] despite still being lower than the scores of the NLP based models. This analysis highlights that whilst DSA-MAG improved recall, it also reduced precision which affected its overall F1-score despite performing better in terms of accuracy and recall compared to the other models.

5.2. Discussion

The experimental results give useful insights into the applicability of generative AI frameworks to the processing and classification of dramatic scripts, by theatrical genre. The DSA-MAG model achieved the best accuracy, 25.40% on the Movie Scripts Corpus test set, outperforming three baseline methodologies. However, all models were relatively poor overall accuracy, due to the complexity of the classification task of differentiating individual film titles, which led to many low-frequency classes with limited textual differences. The superior performance of DSA-MAG is attributed to the Random Forest ensemble method which efficiently captures non-linear interactions between features as opposed to conventional single model methods including Naive Bayes, Support Vector Classification, and Logistic Regression.

NLP Scriptwriting Tools (Naive Bayes) achieved better precision by focusing on specific classes with strong prior evidence, DSA-MAG achieved better recall thus covering a wider range of true genre instances, an important aspect for open domain theatrical classification. From a dramaturgical perspective, the findings validate the assumption that multi-agent architectures modeling different components of the dramatic text independently are better suited to handle the structural complexity than monolithic models. DSA-MAG's design includes a dramatic structure analyzer and an emotional coherence module, supporting the assumption that frameworks considering dramaturgical elements are more appropriate for theatrical script analysis than standard text classification models. The incremental quantitative performance advantage underscores the theoretical argument in favor of dramaturgy-aware AI frameworks in the domain.

This study has implications far beyond stage drama, with relevance to creative industries and education policy. The capacity of multi-agent AI frameworks to

organize and categorize narrative materials suggest that such frameworks could be applied to film, television, interactive media and digital storytelling, where automation in creative work impacts professional labour markets. In education, the findings of this study highlight an immediate need to integrate AI literacy and human-AI co-creation skills into tertiary drama and creative writing curricula. Arts funding and cultural education policymakers should consider how the expansion of AI-assisted authorship might influence support structures for nascent creative professionals in settings where human originality and cultural specificity remain important for artistic value. This study therefore makes a contribution not only to the technical discourse on generative AI but also to the broader social science discussion on creativity, labour, and learning in an algorithmically mediated world.

6. Conclusion

In this work, we propose and evaluate the Dramatic Structure-Aware Multi-Agent Generative (DSA-MAG) framework. We propose the DSA-MAG model to support AI-assisted scriptwriting and classification tasks for stage drama. The DSA-MAG model is composed of a transformer-based encoder and narrative generation agents specialized in plot progression, dialogue construction, and character development. The framework also contains the modules for dramatic structure, emotional coherence, and character reliability analysis. The model is designed to address the gaps of existing generative AI for theater. On the Movie Scripts Corpus, DSA-MAG achieved the highest classification accuracy of 25.40%, outperforming NLP Scriptwriting Tools (23.02%), Conceptual AI Theatre Models (23.81%), and Archetype Learning Models (23.02%). The results confirm the hypothesis that multi-agent generative architectures with theatrical awareness significantly outperform conventional single-model approaches for analysing dramatic scripts. This work therefore provides a novel theoretical framework and empirical validation for the incorporation of dramaturgical principles into generative AI systems, advancing the understanding of AI as a collaborative tool in stage drama production.

6.1. Limitations

There are some limitations of this study worth noting. The Movie Scripts Corpus used as a dataset is primarily composed of movie and television scripts and not plays written for live performance. There are some structural similarities but stage plays are subject to specific conventions of staging, dialogue pacing and physical action that are missing from cinematic screenplays. In addition, the genre labels in the dataset are attached to individual movie titles, rather than well-established theatrical genre categories. This results in a sparse label space with a large number of rare classes. The fragmentation of these classes complicates classification and limits the interpretability of genre-dependent performance metrics. The class imbalance also limited the reliability of the AUC metric and thus the extent of the evaluation. Another limitation is that the DSA-MAG framework was mainly evaluated as a classification model, and not as a complete script generation system. Hence, the generated scripts

were not evaluated by humans, or by experts for quality, coherence, and theatrical applicability. Another limitation is the feature extraction method used, HashingVectorizer, which ignored word order, syntax, and contextual meaning, which are important for dramatic language. Finally, due to computational limitations, the sample size was limited to 1,200 script segments which may limit the generalizability of the findings to larger and more diverse theatrical datasets.

6.2. Future Work

The results of this study suggest several directions for future research. First, future work should develop and utilize a curated collection of stage plays, including scripts written specifically for live performance, to provide a more realistic evaluation of the dramatic content produced by artificial intelligence. Second, replacing the HashingVectorizer with pre-trained large language model (LLM) encoders, such as BERT, RoBERTa, or domain-specific theatre language models, is expected to significantly improve the semantic richness of feature representations, leading to improved results for classification and generation tasks. Finally, we suggest a human evaluation study with professional playwrights, theatre directors and scholars of dramaturgy to assess the artistic quality, emotional impact and theatrical feasibility of DSA-MAG-generated scripts and thus provide qualitative validation beyond the quantitative metrics. Moreover, the multi-agent architecture allows for augmentation by the addition of a staging agent, whose tasks would include the writing of production notes, blocking directions and set design details, thus evolving the framework into a more comprehensive theatre production assistant. Fifth, the integration of reinforcement learning from human feedback (RLHF) approaches in the DSA-MAG pipeline could allow for iterative improvement of generated scripts. The machine-generated outputs would be brought more in line with established professional standards in theatre through the selections of expert reviewers guiding the process of refinement, with its applicability to non-English dramatic traditions eventually to be tested through cross-lingual and cross-cultural extensions, potentially broadening the scope of AI-assisted dramaturgy to the world's theatrical settings.

References

[1] Anantrasirichai, N., & Bull, D. (2022). Artificial intelligence in the creative industries: A review. Artificial Intelligence Review, 55, 589–656. https://doi.org/10.1007/s10462-021-10039-7

[2] Branch, B., & Mirowski, P. (2024). Artificial theatres of the absurd. In The Routledge Handbook of AI and Literature (1st ed.). Routledge

[3] Branch, B., Mirowski, P., Mathewson, K., Ppali, S., & Covaci, A. (2024). Designing and evaluating dialogue LLMs for co-creative improvised theatre. In Proceedings of the International Conference on Computational Creativity (ICCC 2024)

[4] Cake, S. (2025). Artificial intelligence as a collaborative tool for script development. Media Practice and Education, 26(3), 302–317. https://doi.org/10.1080/25741136.2025.2454074
[5] Cardon, P., Fleischmann, C., Aritz, J., Logemann, M., & Heidewald, J. (2023). The challenges and opportunities of AI-assisted writing: Developing AI literacy for the AI age. Business and Professional Communication Quarterly, 86(3), 257–295. https://doi.org/10.1177/23294906231176517
[6] Chakrabarty, T., Laban, P., Agarwal, D., Muresan, S., & Wu, C. S. (2024). Art or artifice? Large language models and the false promise of creativity. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (pp. 1–34). https://doi.org/10.1145/3613904.3642731
[7] Dayo, F., Memon, A. A., & Dharejo, N. (2023). Scriptwriting in the age of AI: Revolutionizing storytelling with artificial intelligence. Journal of Media & Communication, 4(1), 24–38. http://ojs.ilmauniversity.edu.pk/index.php/jmc/article/view/28
[8] Horváth, D. (2025). Curtain call for AI: Transforming theatre through technology. Sustainable Futures, 9, 100747. https://doi.org/10.1016/j.sftr.2025.100747
[9] Hoyer, D., & Frühmorgen, T. (2025). The penned parrot in the writers' room: Four-stage story generation in collaborative screenwriting with AI. International Journal on Stereo & Immersive Media, 9(1). https://doi.org/10.60543/ijsim.v9i1.9391
[10] Kabashkin, I., Zervina, O., & Misnevs, B. (2025). AI narrative modelling: How machines' intelligence reproduces archetypal storytelling. Information, 16(4), 319. https://doi.org/10.3390/info16040319
[11] Kavitha, L. (2024). Copyright challenges in the artificial intelligence revolution: Transforming the film industry from script to screen. Trinity Law Review, 4(1). https://doi.org/10.48165/tlr.2024.4.1.1
[12] Latif, A. S., Ismael, M. F., & Khazal, K. K. (2025). The role of AI in theater directing: New approaches to digital performances. In Proceedings of the 2025 IEEE 4th International Conference on Computing and Machine Intelligence (ICMI) (pp. 1–5). IEEE. https://doi.org/10.1109/ICMI65310.2025.11141300
[13] Lo Duca, A., & Rotelli, D. (2026). Experimenting with generative AI in screenwriting for film schools. In A. Dipace, C. Limongelli, M. Marras, & S. M. Pagliara (Eds.), Artificial intelligence with and for learning sciences: Past, present, and future horizons (WAILS 2025) (Lecture Notes in Computer Science, Vol. 16438). Springer, Cham. https://doi.org/10.1007/978-5-032-17604-2_10
[14] Mirowski, P., Mathewson, K. W., Pittman, J., & Evans, R. (2023). Co-writing screenplays and theatre scripts with language models: Evaluation by industry professionals. In Proceedings of the 2023 CHI Conference on

Human Factors in Computing Systems (Article 355, pp. 1–34). https://doi.org/10.1145/3544548.3581225
[15] Palanimurugan, P., Shanthi, V., Bharathi, M. T., & Sakthivel, M. (2025). The creative code: Generative AI and the transformation of authorship in the screen industries. Academic Research Journal of Science and Technology, 2(05), 38–50. https://doi.org/10.63300/arjst0205202506
[16] Ren, Z. (2024). The role of AI in theatre: Exploring the creation of AI-generated stage plays. Theoretical and Natural Science, 34(1), 27–32. https://doi.org/10.54254/2753-8818/34/20241146
[17] Tsao, J., Liang, C., Nogues, C., et al. (2025). Perceptions and integration of generative artificial intelligence in creative practices and industries: A scoping review and conceptual model. AI & Society. https://doi.org/10.1007/s00146-025-02667-2
[18] Wang, Q. (2025). Transformation and restructuring: Decoding the AI-driven creative logic of micro-short drama video production from SkyReels. Journal of AI Ethics and Legal, 1(1). https://doi.org/10.37420/j.aiel.001
[19] Xu, Z., & Xie, B. (2026). The application of artificial intelligence in film and television script creation. Discover Artificial Intelligence, 6, 78. https://doi.org/10.1007/s44163-025-00741-8
[20] Yang, B., Peng, J., Tang, S., Tan, J., Zhou, F., Cui, S., Qu, W., Li, T., Liu, F., & Wang, D. (2026). Co-DIRECT: A knowledge-augmented multi-agent framework for interactive drama script generation. Expert Systems with Applications, 303, 130571. https://doi.org/10.1016/j.eswa.2025.130571
[21] Hesmondhalgh, D., & Baker, S. (2011). Creative labour: Media work in three cultural industries. Routledge. https://doi.org/10.4324/9780203855889
[22] Latour, B. (2005). Reassembling the social: An introduction to actor-network-theory. Oxford University Press.
[23] McCosker, A., & Wilken, R. (2020). Automating vision: The social impact of the new camera consciousness. Routledge. https://doi.org/10.4324/9781351002187
[24] Manovich, L. (2019). AI aesthetics. Strelka Press.
[25] Boden, M. A. (2016). AI: Its nature and future. Oxford University Press.
[26] Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. Minds and Machines, 30, 681–694. https://doi.org/10.1007/s11023-020-09548-1
[27] Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). CAN: Creative adversarial networks generating “art” by learning about styles and deviating from style norms. In Proceedings of the International Conference on Computational Creativity (ICCC 2017).
[28] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press. https://www.deeplearningbook.org
[29] Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.

[30] Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) (pp. 1532–1543). https://doi.org/10.3115/v1/D14-1162
[31] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems, 30, 5998–6008.
[32] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of NAACL-HLT 2019 (pp. 4171–4186). https://doi.org/10.18653/v1/N19-1423
[33] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., et al. (2020). Language models are few-shot learners. In Advances in Neural Information Processing Systems, 33, 1877–1901.
[34] OpenAI. (2023). GPT-4 technical report. arXiv preprint arXiv:2303.08774. https://doi.org/10.48550/arXiv.2303.08774