A BERT-based Method for Assessing Script-Actor Consistency

Siyong Wu

Beijing Haidian Kaiwen Academy, Beijing, China

17304773235@163.com

Keywords: BERT, NLP, Acting stars, Script selection, Consistency assessment

Abstract: In the modern entertainment industry, the career of an acting star is closely related to his/her public image and reputation, and the consistency between the two plays a crucial role in the star's market performance and career development. However, current script selection and characterization rely mainly on the subjective judgments of directors, casting agents, and producers, and there is a lack of systematic tools to scientifically assess the fit between a star's image and a scripted character. This approach is not only susceptible to bias, but also suffers from decision-making risk and uncertainty. In order to solve this problem, this study proposes a BERT (Bidirectional Encoder Representations from Transformers) based model for assessing the consistency between scripts and stars' images, aiming to analyze the similarity between script texts and stars' established public images through natural language processing techniques. In this study, a large amount of script data and star image description data were collected, and the semantic features of the text were extracted by using the BERT model through preprocessing steps such as data cleaning, word splitting and context embedding. Then, the cosine similarity is used to calculate the similarity between the script and the star image, and the model is trained and consistency evaluated by supervised learning. The experimental results show that the model performs well in several indexes such as accuracy, recall and F1 value, and is able to effectively capture the complex semantic relations in the text and accurately judge the consistency between the script and the star's image, with strong generalization ability. In summary, this study provides a scientific decision support tool for drama selection in the entertainment industry, effectively improves the efficiency and success rate of drama selection, and reduces the risk caused by the inconsistency between the script and the star's image, which is an important application value and theoretical contribution to the management of the star's image and the selection of scripts in the entertainment industry.

1. Introduction

In the modern entertainment industry, the careers of acting stars are closely linked to their public image and reputation. Public image is not only a demonstration of a star's personal charisma, but also an important basis for them to gain a foothold in the marketplace. As the entertainment industry becomes increasingly saturated and competitive, top actors and actresses must build a unique image and reputation through their work roles, variety activities, and interactions with their fans. This image is not built overnight; it requires long-term effort and accumulation. A consistent and distinctive public image can guide the public's expectations and increase the buzz around a star's work, which in turn boosts the star's own popularity and commercial value. On the contrary, if there is inconsistency between the star's established image and the character of the new work, it often leads to negative reactions and emotions from the audience, and may even seriously affect the actor's career and the market performance of the work. Therefore, it can be argued that the consistency between a script and the public image that an acting star displays to the public over time, i.e., the fit between the actor's persona and the script's persona, is crucial to the success of a film or television production as well as to the star's own stardom.

However, the current script selection process relies heavily on the intuition and experience of directors, casting agents and producers, an approach that is often subjective and susceptible to bias. It is difficult to avoid the uncertainty and risk involved in the process of making script selection

DOI: 10.25236/iwmecs.2025.056

decisions. In order to reduce the risk of miscalculation associated with this subjectivity, it is necessary to develop a scientific and systematic tool to help assess whether a script is in line with an acting star's public image and reputation. The development of artificial intelligence, especially the advancement of natural language processing technology, provides new possibilities to solve this problem.

In this study, we propose a Bidirectional Encoder Representations from Transformers (BERT) based model for evaluating the consistency between a script and a star's image for assessing whether a given script conforms to the established public image and reputation of an acting star. BERT, as a state-of-the-art natural language processing model, is capable of capturing the complexity of a text and capturing the complexity of a text's image and reputation. BERT, as an advanced natural language processing model, is capable of capturing complex semantic relationships in text and has the ability to deeply understand the context. The BERT-based model is able to establish complex semantic associations between the script text and the star's image description, and then assess the consistency between the two. Therefore, the evaluation tool based on the BERT model can deeply understand the compatibility between the script text and the star image description, and help the entertainment company and the star team to make more informed choices in the decision-making process of selecting dramas.

This study provides new research perspectives and methods for star image management and script selection by introducing artificial intelligence and natural language processing technologies. By constructing a practically usable evaluation model, this study provides the entertainment industry with a scientific decision support tool for drama selection, which improves the efficiency and success rate of drama selection. At the same time, the method can also effectively reduce the risk due to the inconsistency between the script and the star's image, and protect and enhance the star's public image and reputation.

2. Related Work

2.1. Public image and reputation management

Public image and reputation management are crucial in the entertainment industry, directly affecting a star's career development, market value, and public perception and acceptance of them [1]. Public image can be understood as the overall impression of a star in the eyes of the public, which is usually shaped by a combination of factors such as the star's behavior and speech, past works, social media activities, and news reports. Reputation, on the other hand, involves a wider range of social perceptions, including the public's evaluation of a star's moral character, professionalism, and sense of public responsibility.

It has been shown that a star's public image and reputation have a direct impact on his or her professional success. A positive public image can enhance fans' emotional commitment and loyalty, and increase the commercial value and market competitiveness of stars [2]. On the contrary, a negative public image may lead to fan loss and even seriously affect a star's career and market performance. For example, studies have pointed out that when the image of stars is inconsistent with the image of the characters they portray in film and television works, viewers tend to develop negative emotions, which affects the popularity and box office performance of the works [3].

In practice, the public image management of stars often relies on experienced public relations teams, who maintain and enhance the public image of stars by planning media activities, managing social media accounts, and controlling the winds of public opinion [4]. Meanwhile, crisis management is also an important part of image management. In today's extremely fast information dissemination, stars' performance and response strategies in the face of crisis directly affect the speed and effect of their public image recovery. Although these studies provide important references for celebrity image management, most of them rely on qualitative analysis and lack systematic and quantitative assessment tools to provide sufficient data support in script selection or brand cooperation decisions.

2.2. Natural language processing in script analysis

The application of Natural Language Processing (NLP) techniques in the field of text analysis has made significant progress [5], especially in the entertainment industry, where NLP is used to analyze the plot development, character traits, and dialogue patterns of scripts. In recent years, with the development of deep learning techniques, models based on the Transformer [6] architecture (e.g., BERT [7], GPT [8], etc.) have performed particularly well in natural language processing tasks. These models are able to capture complex semantic relationships in text and show strong capabilities in contextual understanding.

In the specific application of script analysis, NLP techniques are used for a variety of tasks. For example, plot analysis of screenplays usually involves identifying and understanding plot turning points, climaxes, and emotional ups and downs in the text, which can help writers and directors better understand the story structure and make improvements. Character analysis is another important application; by analyzing characters' dialogues, behaviors, and descriptions, researchers can automatically identify characters' personality traits and developmental trajectories, which is important for casting decisions and characterization [9]. In addition, NLP techniques have been used for sentiment analysis and thematic categorization of screenplays, and these analyses can help filmmaking teams better understand the overall style and tone of a script [10]. Through sentiment analysis techniques, researchers can identify emotional expressions and mood changes in a screenplay to understand the psychological state of the characters and the emotional curve of the plot development. This kind of analysis not only helps to reveal the emotional level of the script, but also provides valuable references for directors and actors to better interpret the characters. In addition, by analyzing the dialogues and interactions between the characters in the script, the researcher can automatically extract and analyze the relationship patterns between the characters. This approach helps to understand the social relationships between characters and the character traits of the characters, and also provides data support for scriptwriting in theater and film.

Although these applications have made great progress in script analysis, most studies have focused on single-dimensional analysis of scripts (e.g., plot or character analysis) and lacked a comprehensive assessment of the consistency of scripts with the public image of stars. This lack of comprehensive assessment hinders the scientific decision-making of celebrity teams and entertainment companies in the script selection process.

3. Relevant Technology Base

3.1. Natural language processing

Prior to BERT, traditional natural language processing models mainly relied on Word Embeddings techniques, such as Word2Vec and GloVe. These models represent the semantic information of words by mapping them into low-dimensional vectors. However, a major limitation of these word embedding models is that they are static, i.e., the vector representation of each word is fixed and cannot be adapted to the context. For example, the word “bank” has different meanings in “river bank” and “financial bank”, but the static word embedding cannot capture this kind of information. Static word embeddings cannot capture this polysemy.

To overcome this limitation, researchers have introduced deep learning-based sequence models such as LSTM (Long Short-Term Memory Network) and GRU (Gated Recurrent Unit), which are capable of generating dynamic context-sensitive word representations based on the input sequence. However, LSTMs and GRUs suffer from the problem of gradient vanishing when processing long texts, as well as low training efficiency due to their sequential processing nature.

In 2017, Vaswani et al. proposed the Transformer model, which is completely based on Attention Mechanism and is free from the dependence of RNN. The advantage of Transformer is that it can process data in parallel, which greatly improves the training efficiency, while its Self-Attention Mechanism) can capture long-range dependencies, enabling it to better understand contextual information. The emergence of Transformer laid the foundation for BERT, but the initial Transformer applications (e.g., GPT, Generative Pre-trained Transformer) still were unidirectional and could only generate text from left to right, limiting the depth of contextual understanding.

3.2. BERT

BERT (Bidirectional Encoder Representations from Transformers) is proposed to address the limitation that unidirectional models are unable to fully understand the context, as shown in Figure 1. BERT, by adopting the structure of a Bidirectional Encoder, allows the model to be trained to By using a Bidirectional Encoder structure, BERT allows the model to focus on the context of a word during training, thus providing a deeper understanding of the overall semantics of the sentence. This bi-directionality allows BERT to learn the representation of a word in the context surrounding that word, which is important in many language understanding tasks.

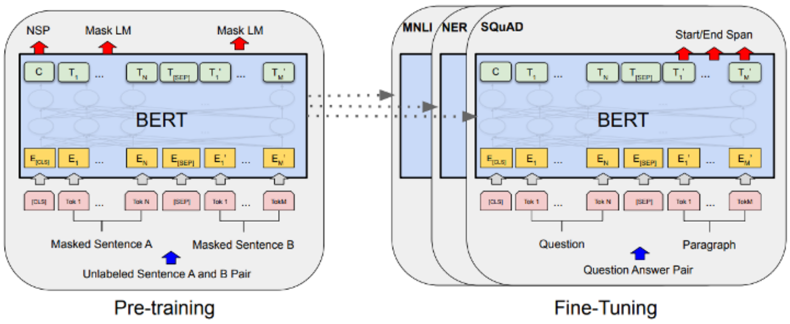

Figure 1 Overall pre-training and fine-tuning procedures for BERT.

BERT employs a new pre-training strategy that includes Masked Language Model (MLM) and Next Sentence Prediction (NSP). In the MLM task, BERT masks a number of words in a sentence by randomly masking them and then allowing the model to predict these masked words. This approach allows the model to capture the contextual relationships of individual words in the sentence, rather than just predicting the next word. In the NSP task, BERT helps the model understand longer contexts and logical relationships between sentences by, given pairs of sentences, having the model predict whether the second sentence is a natural successor to the first, which helps the model understand longer contexts and logical relationships between sentences.

These pre-training tasks enable BERT to learn language patterns from large amounts of unlabeled data in an unsupervised manner. By pre-training on large-scale corpora (e.g., Wikipedia and book datasets), BERT models capture rich semantic information and linguistic features. After completing the pre-training, BERT can be quickly adapted to various downstream tasks, such as text categorization, named entity recognition, sentiment analysis, and question-answer systems, through Fine-tuning. This pre-training plus fine-tuning model greatly reduces the need for labeled data and significantly improves the model's performance on a variety of NLP tasks. BERT has achieved unprecedented results in multiple benchmarks, which triggered a wave of research in the field of pre-training language models in NLP.

4. A BERT-based Model for Assessing Script-Actor Consistency

4.1. Data collection and pre-processing

In this study, in order to construct an effective BERT-based model for assessing the consistency of scripts and actors, we first carried out systematic data collection and preprocessing. The goal of this process is to obtain rich and diverse textual data that can accurately reflect the plot development, characterization of different types of scripts, as well as the public's general perceptions and evaluations of acting stars.

Taking actor Wang He Di as an example, in order to analyze whether Wang He Di's role as Dongfang Qingcang in the TV series Love Between Fairy and Devil is consistent with his consistent sunny and positive image, as shown in Figure 2 and to explore the impact of the character performance on enhancing Wang He Di's popularity, we conducted systematic data collection and pre-processing. The core goal of this process was to obtain rich textual data that could reflect Wang

He Di's public image and the effects of his character performance, as well as to assess the impact of the performance on public awareness and discussion through social media data.

Figure 2 Actor Wang Hedi and the role of Dongfang Qingcang in "Love Between Fairy and Devil".

First, we collected detailed profile data on Wang He Di, emphasizing his sunny and positive image in the public eye. This data comes from multiple sources, including mainstream media reports, Wang He Di's personal social media postings (e.g., Weibo, Instagram, etc.), fan community discussions, and movie reviews of his past works. We pay special attention to textual sources that reflect Wang He Di's character traits and public perception, such as the content of his interactions with fans on social media, media reports when he attends public events, and fans' comments on his image. Through these diverse sources, we were able to paint a comprehensive picture of Wang He Di's public image and ensure the richness and representativeness of the data.

Next, we conducted an in-depth analysis of the script text for the character of Dongfang Qingcang in Love Between Fairy and Devil. The script text mainly comes from the official scripts of the TV series and secondary creative content (such as drama reviews, homoerotic creations, etc.). In the process of data collection, we focused on collecting lines, scene descriptions, character analysis and plot development related to the character of Dongfang Qingcang. These data provide the model with contextual information about the character, which helps the model understand the character traits and emotional expression of the character Dongfang Qingcang, and further analyze whether it is consistent with Wang He Di's consistent sunny and positive image.

In order to assess the impact of Wang He Di's performance in Cang Lan Jue on his popularity, we also collected a large amount of social media data, especially indicators such as the frequency of hot searches and discussions from Weibo. By using web crawler tools (e.g., Scrapy and Selenium), we automated the collection of microblogging data about Wang He Di and The Secret of Cang Lan, including the frequency of hot search lists on microblogs, the number of discussions on related topics, and the content. These data not only include the amount of fan comments and retweets, but also cover secondary communication content such as emoticons, video clips, and fan-created content, which can be used as indirect indicators to measure the effectiveness of the performance and the influence of the star. By analyzing the Weibo hot search data, we can quantify the degree of public discussion and exposure caused by Wang Hedi's role as Dongfang Qingcang, and thus assess the effect of the character performance on his popularity.

After the data collection was completed, we conducted a comprehensive preprocessing exercise. First, we cleaned and standardized all the collected text data to remove irrelevant content and noise data.

Through the above data collection and preprocessing steps, this study not only constructed a comprehensive textual dataset for analyzing Wang Hedi's performance and image consistency, but also assessed the actual impact of his performance on popularity through social media data. Such an approach allows us to more precisely understand the relationship between celebrity image and character selection, and provides a scientific and quantitative basis to support the decision-making process in the entertainment industry.

4.2. Text processing and feature extraction

The core of text processing lies in how to convert unstructured text data into structured data suitable for processing by machine learning models, especially high-dimensional semantic vector representations. This process mainly includes the steps of disambiguation, embedding extraction, and feature representation, which are combined with the powerful semantic understanding of the BERT model, which can effectively capture the complex semantic relationships in the text.

In the text processing and feature extraction part of this study, we utilize advanced natural language processing techniques to ensure that rich semantic features are extracted from the scripts and celebrity image data to provide high-quality inputs for the subsequent consistency assessment model. The core of text processing lies in how to convert unstructured text data into structured data suitable for processing by machine learning models, especially high-dimensional semantic vector representations. This process, which mainly includes steps such as disambiguation, embedding extraction and feature representation, combines the powerful semantic understanding of the BERT model, which is able to effectively capture the complex semantic relationships in the text.

First, in order to process the collected script texts and celebrity image descriptions, we use the Tokenizer that comes with the BERT model to perform lexical segmentation.The BERT's Tokenizer adopts the WordPiece segmentation method, which is able to break down the vocabulary into finer-grained subword units. For example, for less common or compound words, WordPiece breaks them down into smaller, more common subword fragments, which greatly reduces the size of the vocabulary and performs well when dealing with rare or polysemous words. Suppose that given a sequence of text $T = \{t_1, t_2, \ldots, t_n\}$, where t_i denotes the i-th word in the text sequence, the word splitter converts the text into a sequence of subword units $S = \{s_1, s_2, \ldots, s_m\}$, where m ≥ n and each s_j is a subword fragment of t_j. This fine-grained representation helps to capture richer lexical information while preserving the original semantics of the text.

After completing the segmentation, we feed the preprocessed text sequences into the BERT model to extract contextual embeddings. The BERT model is based on the Transformer architecture, which uses a multilayer Bidirectional Encoder to generate deep semantic representations of the text. The BERT model is based on the Transformer architecture. When each input sequence passes through the encoder of the BERT model, a vector $\mathbf{H} = \{\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_m\}$ of contextual representations of the sequence is generated, where $\mathbf{h}_i \in \mathbb{R}^d$ denotes the contextual embeddings of the i-th subword unit and d is the hidden layer dimension of the model. In this process, BERT captures the semantic relationship of each subword in the whole sequence through the self-attention mechanism (self-attention mechanism), which is calculated as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

where Q,K,V represent the query, key and value matrices, respectively, and are the dimensions of the key vectors. Through the self-attention mechanism, BERT is able to capture inter-word dependencies globally, such that the representation of each word does not only depend on its local context, but combines the contextual information of the whole sentence.

To further represent the text as feature vectors that can be used for similarity computation, we perform an aggregation operation on the context embeddings generated by BERT. Specifically, we employ Pooling techniques such as mean pooling or max pooling to obtain a fixed dimensional representation of the whole text. Set the overall representation of the text as $\mathbf{v} \in \mathbb{R}^d$, for the mean pooling method, it can be expressed as:

$$\mathbf{v} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{h}_i \tag{2}$$

This approach averages the contextual embeddings of each subword to generate an overall semantic representation of that text. Accordingly, for the maximum pooling approach, each dimension of $\mathbf{v}$ depends on the maximum value of all subwords in that dimension. Through these aggregation operations, we convert the variable-length sequence of subword embeddings into a fixed-dimension feature vector, which facilitates the subsequent similarity computation.

These preprocessing and feature extraction steps not only ensure that the feature representations between the script and star image texts are rich in semantic information, but also enable these texts to be compared and computed in a unified vector space. This processing method fully utilizes the advantages of BERT in understanding complex semantic and contextual relationships, providing strong support for the model in assessing the consistency between scripts and star images.

4.3. Similarity calculation

In order to assess whether Wang Hedi's role as Dongfang Qingcang in Love Between Fairy and Devil is consistent with his usual sunny and positive image, we do this by calculating the similarity between the text of the script and the description of the star's image. The core of this similarity calculation lies in how to transform the textual representations of the script and the star's image into high-dimensional feature vectors, and measure the degree of similarity between the two in the same vector space. In order to ensure the accuracy and validity of the similarity calculation, we adopt the contextual embedding extracted based on the BERT model, and Cosine Similarity as the similarity metric.

First, we use the BERT model to extract semantic embeddings from script text and star image descriptions. Specifically, given a script text T_s and a celebrity image description T_p, we obtain their contextual representation vectors $\mathbf{H}_s = \{\mathbf{h}_{s1}, \mathbf{h}_{s2}, \ldots, \mathbf{h}_{sm}\}$ and $\mathbf{H}_p = \{\mathbf{h}_{p1}, \mathbf{h}_{p2}, \ldots, \mathbf{h}_{pn}\}$, where $\mathbf{h}_{si}$ and $\mathbf{h}_{pj}$ denote the embeddings of the i-th and j-th subwords in the script and the celebrity description text, respectively, and are the high-dimensional vectors output by the BERT model, and $\mathbf{h}_{si}, \mathbf{h}_{pj} \in \mathbb{R}^d$ is the dimension of the hidden layer. In order to transform these embedding representations into fixed-length feature vectors, we perform an aggregation operation, such as mean pooling or max pooling, on the embedding representation vectors to obtain the global representation vectors of the script and the star image $\mathbf{v}_s$ and $\mathbf{v}_p$.

After obtaining the feature vector representation of the text, we use cosine similarity to measure the similarity between the script and the star image. Cosine similarity is a commonly used similarity measure that evaluates the degree of similarity between two vectors by calculating the cosine of the angle between them, with values ranging from -1 to 1, where 1 means that the two vectors are completely similar and -1 means that the two vectors are completely opposite. Given two eigenvectors and , the cosine similarity between them can be expressed as:

$$\text{Cosine Similarity}(\mathbf{v}_s, \mathbf{v}_p) = \frac{\mathbf{v}_s \cdot \mathbf{v}_p}{\|\mathbf{v}_s\| \ \|\mathbf{v}_p\|} = \frac{\sum_{i=1}^{d} v_{si} \cdot v_{pi}}{\sqrt{\sum_{i=1}^{d} v_{si}^2} \cdot \sqrt{\sum_{i=1}^{d} v_{pi}^2}} \tag{3}$$

where $\mathbf{v}_s \cdot \mathbf{v}_p$ denotes the dot product of the vector, $\|\mathbf{v}_s\|$ and $\|\mathbf{v}_p\|$ denote the Euclidean norm of the vector, respectively. By calculating the cosine similarity, we can quantify the degree of semantic consistency between the scripted character and the star's public image. A high similarity score indicates a high degree of similarity between the scripted character's features and the star's established public image, while the opposite indicates a low degree of congruence.

In order to further improve the robustness and accuracy of the similarity computation, we also considered the use of multi-dimensional feature representation for the computation. Specifically, instead of just relying on the global representation of a single text, we combine the embedded representations of different text passages or plot segments to construct a more fine-grained framework for similarity computation. For example, for script text, we can extract multiple embedding vectors by scene or character dialog segments, and then calculate the similarity between each segment and the star's image description separately, and finally obtain the overall similarity score by weighted average. The weighted average can be expressed as:

$$\text{Overall Similarity} = \frac{\sum_{k=1}^{K} w_k \cdot \text{Cosine Similarity}(\mathbf{v}_{sk}, \mathbf{v}_p)}{\sum_{k=1}^{K} w_k} \tag{4}$$

where K denotes the number of text segments, and w_k is the weight coefficient of each segment, which can be set according to the importance of the segment or the length of the text. This multi-granularity similarity calculation method can better capture the semantic differences between different text passages and improve the model's ability to consistently evaluate in different contexts.

4.4. BERT model training

After completing the text processing and similarity calculation, the key step in this study is to optimize the parameters by training the BERT model to automate the assessment of the consistency between the script and the star's image. The core objective of model training is to learn to effectively extract features from the input text and accurately predict the degree of consistency between the scripted character and the established image of the star based on these features. To this end, we design a classification task that transforms the consistency assessment problem into a supervised learning problem, enabling the BERT model to learn how to make classification predictions based on the feature representations of the script text and the star's image, guided by the labeled data.

The training process starts with the construction of a labeled dataset containing a large number of script and star profile pairs, each of which is manually labeled as “matching” or “not matching”. Assume that given a script text T_s and a star profile T_p, the corresponding labels $y \in \{0,1\}$ indicate whether or not they match. The label distribution of the samples in the dataset should be as balanced as possible to avoid bias in the training process. In addition, in order to improve the generalization ability and robustness of the model, we perform data augmentation during the data preparation process, e.g., generating additional training samples by replacing synonyms, reorganizing sentence structures, and so on. Through these measures, we ensure the diversity and representativeness of the training data.

In training the BERT model, we feed each text pair (T_s, T_p) into the model's bi-directional encoder to extract the respective contextual embedding representations. In order to compute the similarity between the script and the star's image description, we pass the feature representations of the text pairs through the output layer of BERT (usually the [CLS] tagged output) to a fully connected layer, which outputs a scalar representation of the consistency score. This score is normalized by a sigmoid activation function to ensure that its value is between 0 and 1. The predictions of the model can be expressed as follows:

$$\hat{y} = \sigma(\mathbf{W} \cdot [\mathbf{v}_s ; \mathbf{v}_p] + b) \tag{5}$$

where, $\mathbf{W}$ and b are the weights and bias parameters of the fully connected layer, $[\mathbf{v}_s ; \mathbf{v}_p]$ denote the merged representation after splicing the feature vectors of the script and the star image, and $\sigma(x) = \frac{1}{1+e^{-x}}$ is the sigmoid activation function. The training objective of the model is to minimize the binary cross-entropy loss function, which is given by:

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^{N} \left(y_i \log(\hat{y}_i) + (1-y_i) \log(1-\hat{y}_i) \right) \tag{6}$$

where N denotes the total number of training samples, y_i and $\hat{y}_i$ are the true label and prediction probability of the ith sample, respectively. By minimizing the loss function, the model adjusts its internal parameters so that the output consistent prediction results are as close as possible to the true label.

4.5. Coherence assessment

After completing the training of the BERT model, we perform consistency evaluation, which aims to use the trained model to make consistency judgments on new scripts and star profiles. In the consistency evaluation phase, given a new script text T_s and a star profile T_p, we first perform the same preprocessing operations on the input text as in the training phase to extract contextual embedding representations. Then, these embedded representations are fed into the trained BERT model, which outputs a consistency score $\hat{y}$ that reflects the model's prediction probability of the consistency between the two. To facilitate interpretation and application, we convert the score into discrete categories, and set a threshold (e.g., 0.5) to categorize samples with scores greater than this threshold as “consistent”, otherwise as “inconsistent”.

In order to further evaluate the performance and validity of the model, we combined social media metrics, such as the frequency of Weibo hot searches and the degree of discussion, as an external validation criterion for the model's effectiveness. By monitoring the Weibo hot search data related to Wang Hedi and his role as Dongfang Qingcang in Love Between Fairy and Devil, we can assess whether the character performance is consistent with the star's established image and whether it has contributed to his popularity. Specifically, we collected data on metrics related to the frequency of hot searches, fan discussions, and sentiment analysis of related topics, and used these data to validate the accuracy and practical effectiveness of the BERT model's prediction results. Through this method of combining external data validation, we ensure that the model not only has good classification ability in theory, but also can show significant effects in practical application, thus providing an important basis for star teams and entertainment companies to make scientific decisions on script selection and character positioning.

5. Experiments and results

5.1. Experimental setup

In the experimental setup of this study, we used a high-performance computing environment and advanced deep learning tools to ensure that the model training and inference process is efficient and accurate. Specifically, our experiments were conducted on a server equipped with one NVIDIA RTX 3090 Ti, a hardware configuration that provides the necessary computational power and parallel processing capabilities for the training of large-scale neural network models. In terms of the software environment,, we mainly use the PyTorch framework to build the BERT model and take advantage of its flexibility and dynamic computational graph to facilitate fine-tuning and parameter optimization of the model.

For the training of the BERT model, we set several key hyperparameters to optimize the model performance and ensure its effectiveness on a given task. First, the learning rate is one of the most important hyperparameters during the model training process, determining the step size of the model parameters to be updated at each iteration. In order to balance the model convergence speed and stability, we set the initial learning rate to . In addition, we adopt a learning rate decay strategy to gradually reduce the learning rate during the training process based on the performance of the validation set to prevent the model from falling into a local optimum and to improve the final generalization ability.

5.2. Results and analysis

We conducted a comprehensive evaluation of the performance of the BERT-based script-actor consistency assessment model, using a variety of performance metrics to measure the model's performance on different dimensions. The main performance metrics of the model include Accuracy, Recall, and F1-score, as shown in Figure 3 which provide us with an all-encompassing understanding of the model's classification capabilities.

First, accuracy rate is one of the basic metrics for evaluating the performance of a classification model, indicating the proportion of samples correctly classified by the model on the test set. In this experiment, the accuracy rate is used to measure the overall classification ability of the model in recognizing the consistency between scripts and star images. By making predictions on the test set, we calculated the accuracy rate of the model, and the results show that the model exhibits high classification accuracy on different types of script and star image combinations, which indicates that the model is able to effectively capture the semantic relationships and features implied in the text, and thus accurately determine the consistency between the two.

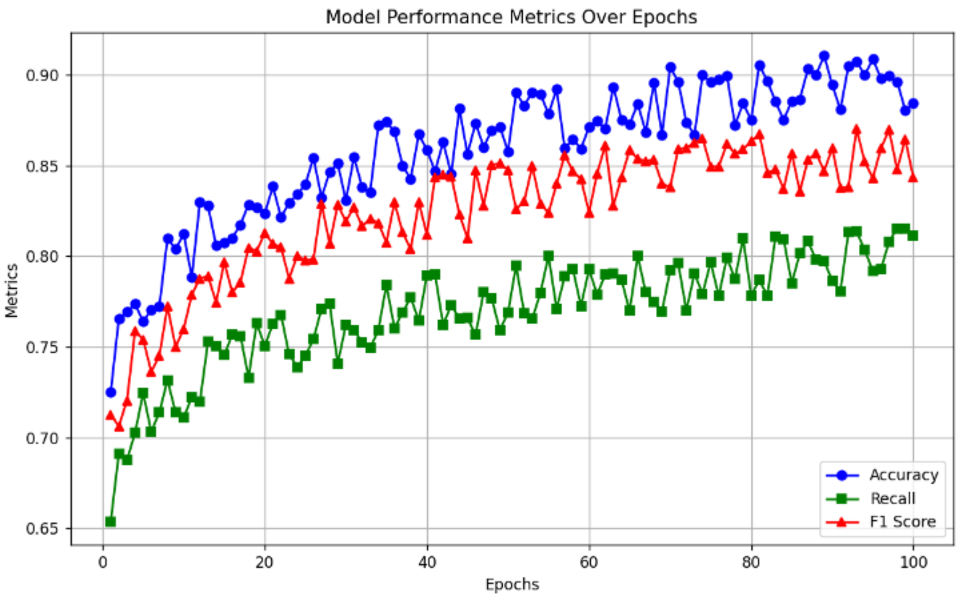

Figure 3 Experimental result.

Second, recall is another important performance metric to assess the model's ability to recognize consistent samples. Specifically, recall measures the proportion of all actual consistent samples that can be correctly recognized by the model. A high recall means that the model is able to identify consistent samples more comprehensively and reduce the occurrence of missed judgments. In this study, the model's recall performance is more satisfactory in most scenarios, especially in those script clips involving obvious emotional characteristics and clear characterization, the model is able to identify the parts that are consistent with the established image of the star very well. This indicates that the model has a strong ability in capturing obvious semantic features in the text of scripts and star images.

In order to comprehensively evaluate the accuracy and recall of the model, we also calculated the F1 value, which is the reconciled average of the accuracy and recall, and is a metric that can comprehensively measure the performance of the model. By calculating the F1 value, we are able to better understand the overall performance of the model in handling the consistency assessment task. In the experiments, the F1 value of the model stays at a high level, which indicates that the model achieves a good balance between accuracy and recall ability, and is able to perform evenly on different types of test samples. This balanced performance makes the model more robust and reliable in practical applications.

The main advantage of the model is its high accuracy and strong generalization ability. Since the BERT model is able to capture the complex semantic relationships in the text, it shows good adaptability when dealing with different types of scripts and star images. Especially in scripts with clear plot structure and emotional clues, the model is able to accurately identify the match between character traits and star images, which makes it of great practical value in decision support for drama selection. In addition, the model maintains stable performance on different types of test datasets, which further validates its generalization ability and gives it the potential to be generalized in practical applications.

References

[1] Yang X. TV Sitcom's Construction To The Image of Citizens--Take" Home With Kids "as an Example[J]. China Media Report Overseas, 2013, 9(2).

[2] Wei Y A N, Haojie F. The Genre Characteristics and Artistic Innovation of Character Creation of Television Dramas During Major Publicity Periods in Recent Years[J]. Journal of Chongqing Jiaotong University Social Sciences Edition, 2021, 21(2): 75.

[3] Duo L A N. Innovative Path of Chinese Cinema’s Overseas Online Communication in New Era[J]. Journal of Beijing University of Posts and Telecommunications (Social Sciences Edition), 2022, 24(3): 45.

[4] Ting-yu O U. The Negative Influence of" Fan Circle" Disorder on Teenagers' Values and Its Countermeasures[J]. Journal of Xihua University (Philosophy & Social Sciences), 2023, 42(3): 80-88.

[5] Keliang Z, Wanghao G U O. Revisit the Concept in Language Philosophy and Linguistics[J]. Contemporary Foreign Languages Studies, 2020, 20(6): 41.

[6] Vaswani A. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017.

[7] Devlin J. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint arXiv:1810.04805, 2018.

[8] Radford A, Narasimhan K, Salimans T, et al. Improving language understanding by generative pre-training. preprint arXiv 2018.

[9] Chen C. The Feature of Animation Role Mold in Animation Film from Monkey King: Hero Is Back[C]//2016 International Conference on Education, Management and Computer Science. Atlantis Press, 2016: 937-940.

[10] Wang Y, Zhu J, Wang Z, et al. Review of applications of natural language processing in text sentiment analysis[J]. Journal of Computer Applications, 2022, 42(4): 1011.