SMART BRAIN-DRIVEN DATA KNOWLEDGE GRAPH CONSTRUCTION METHOD AND SYSTEM FOR COMPLEX ECOLOGY

TECHNICAL FIELD

[0001] The present invention relates to a complex ecological technology, and specifically relates to a smart brain-driven data knowledge graph construction method and system for complex ecology.

BACKGROUND

[0002] Knowledge graphs, as a structured form of knowledge representation, play an increasingly important role in various fields of artificial intelligence, such as natural language processing, recommendation systems, and knowledge inference. Traditional knowledge graph construction methods usually rely on manual construction or rule-based automation methods, which have problems such as low efficiency, poor scalability, and difficulty in adapting to complex and dynamic real-world data. In recent years, with the rise of deep learning and graph neural networks, researchers have begun to explore the use of these technologies to automatically construct and infer the knowledge graphs.

[0003] Existing methods often struggle to effectively handle multi-modal data. In the real world, data often contains information from various modalities such as text, images, and audio, while most existing methods primarily focus on single-modal data, making it difficult to fully utilize the rich semantic information in the multi-modal data.

[0004] The existing methods still have limitations in knowledge inference. Knowledge inference is one of the important functions of the knowledge graphs, but the inference by the existing methods is often restricted to shallow inference paths, making it difficult to perform complex multi-hop inference.

[0005] The existing methods lack effective optimization mechanisms. The construction of the knowledge graphs is an iterative optimization process that requires continuous correction and refinement for graph structures and knowledge. However, the existing methods often lack the effective optimization mechanisms, making it difficult to ensure the quality and accuracy of the graphs.

SUMMARY OF THE INVENTION

[0006] The embodiments of the present invention provide a smart brain-driven data

knowledge graph construction method and system for complex ecology, which can solve the problems in related technology.

[0007] In a first aspect,

[0008] the smart brain-driven data knowledge graph construction method for complex ecology provided by the embodiments of the present invention includes:

[0009] collecting scenario data by a deep feature sensing network, acquiring an original multi-dimensional sequence containing a service process, a processing rule and asset association, performing semantic parsing on the original multi-dimensional sequence to generate an initial knowledge vector, constructing an initial graph structure containing entity nodes and relational edges based on the initial knowledge vector, and encoding the initial graph structure by a graph neural network to obtain a graph feature fingerprint used for feature distribution representation and a feature set used for knowledge discovery;

[0010] computing a similarity matrix among entities in the initial graph structure based on the graph feature fingerprint, screening entity pairs with relevancy greater than a preset similarity threshold from the similarity matrix by the feature set, executing multi-hop path inference on the screened entity pairs to generate an inference path set, inputting the inference path set into a recurrent neural network to generate an entity association confidence index, performing multi-modal feature extraction on data in the initial graph structure and the inference path set based on the entity association confidence index, the multi-modal features including text features, image features and audio features, and aligning and fusing the multi-modal features by a cross-modal attention mechanism to generate a unified multi-modal knowledge representation vector; and

[0011] inputting the multi-modal knowledge representation vector and the entity association confidence index into a multi-level feature network, generating a multi-modal knowledge representation vector, the multi-modal knowledge representation vector including an entity feature vector, a relational feature vector and a sub-graph feature vector; based on the multi-modal knowledge representation vector, identifying influence features by an SHAP method, inputting the influence features into a problem evolution graph, computing a transmission probability by causal inference, and performing comparative analysis on the transmission probability and the graph feature fingerprint to generate an optimization target; extracting gradient information in a feature space according to the optimization target, and inputting the gradient information into an Adam optimizer to generate an error correction strategy, generating an adversarial sample for verification based on the error correction strategy, dynamically adjusting an inference depth according to the verification result by a dual-DQN network, and finally, performing comparative

verification on optimized graph features and the graph feature fingerprint to obtain an optimization effect to form a complete optimization closed loop.

[0012] The computing a similarity matrix among entities in the initial graph structure based on the graph feature fingerprint, screening entity pairs with relevancy greater than a preset similarity threshold from the similarity matrix by the feature set includes:

[0013] representing the graph feature fingerprint to be a high-dimensional feature vector, the high-dimensional feature vector containing a graph structure feature, a semantic distribution feature and a topological attribute feature, and the knowledge feature containing an attribute feature matrix of the entity nodes;

[0014] performing feature transformation on the attribute feature matrix by a linear transformation unit to generate a query matrix, a key matrix and a value matrix, the query matrix being obtained by multiplying a query weight matrix and the attribute feature matrix, the key matrix being obtained by multiplying a key weight matrix and the attribute feature matrix, and the value matrix being obtained by multiplying a value weight matrix and the attribute feature matrix;

[0015] dividing the query matrix, the key matrix and the value matrix into a plurality of attention heads, each attention head independently computing an attention score, performing transposition multiplication on the query matrix and the key matrix and carrying out dimension normalization to obtain an attention weight of each attention head, multiplying the attention weights by the value matrix to obtain head outputs, and splicing a plurality of head outputs to form a multi-head attention output;

[0016] based on the multi-head attention output, computing a similarity relationship among the entity pairs through cosine similarity, and generating a similarity matrix, each element of the similarity matrix representing a similarity score of the corresponding entity pair;

[0017] setting a dynamic similarity threshold according to distribution features in the graph feature fingerprint, and performing bidirectional screening verification on the similarity matrix through the dynamic similarity threshold, the bidirectional screening verification requiring that forward similarity and reverse similarity of the entity pairs both meet the dynamic similarity threshold; and

[0018] filtering the entity pairs subjected to the bidirectional verification in combination with constraints in the knowledge feature, so as to generate final related entity pairs.

[0019] The performing multi-modal feature extraction on data in the initial graph structure and the inference path set based on the entity association confidence index by a

BERT-Large model, a Vision Transformer model and a Wav2Vec model, the multi-modal features including text features, image features and audio features, and aligning and fusing the multi-modal features by a cross-modal attention mechanism to generate a unified multi-modal knowledge representation vector includes:

[0020] encoding the inputted text data by a BERT-Large pre-training model, extracting context semantic features of the text data through a multi-level Transformer structure, the context semantic features being represented as a text feature matrix, and generating a text pre-training feature vector based on the text feature matrix;

[0021] dividing the inputted image data into an image block sequence with the size of 16 * 16 pixels by the Vision Transformer model, inputting the image block sequence to a self-attention layer after position encoding, extracting a global visual feature of the image block sequence, and jointly encoding the global visual features and the text pre-training feature vector to obtain an image-text joint feature vector;

[0022] constructing an attention guidance matrix based on the image-text joint feature vector, the attention guidance matrix being used for processing the inputted audio data, inputting the audio data into the Wav2Vec model, extracting the audio features through a multi-level convolutional network and the Transformer structure, and performing feature alignment on the audio features and the image-text joint feature vector to generate a multi-modal alignment feature;

[0023] computing an inter-modal similarity matrix for the multi-modal alignment feature, constructing a cross-modal attention network based on the similarity matrix, and outputting attention weight distribution through the cross-modal attention network; and

[0024] inputting the attention weight distribution into a linear projection layer, and mapping different modal features to the unified semantic space through the linear projection layer to generate the unified multi-modal knowledge representation vector.

[0025] The based on the multi-modal knowledge representation vector, identifying influence features by an SHAP method, inputting the influence features into a problem evolution graph, computing a transmission probability by causal inference, and performing comparative analysis on the transmission probability and the graph feature fingerprint to generate an optimization target includes:

[0026] inputting the multi-modal knowledge representation vector into an improved SHAP model, and computing a synergistic effect among the features by the improved SHAP model based on a feature combination assessment mechanism to obtain a feature importance score; screening the multi-modal knowledge representation vector according to the feature importance score, and selecting the features with the importance score

exceeding a preset feature threshold as influence features;

[0027] constructing the problem evolution graph by taking the influence features as nodes, node attributes of the problem evolution graph containing the feature value and the importance score, and constructing a directed edge based on a time-order relationship among the features; inputting the problem evolution graph into a causal inference network, computing a conditional probability among the nodes by the causal inference network using a variational inference method, and performing Monte Carlo sampling to obtain the probability distribution of the transmission path; and

[0028] comparing the probability distribution of the transmission path with a pre-stored graph feature fingerprint, and computing KL divergence to obtain state difference distribution; and constructing an optimization target function based on the state difference distribution, the optimization target function containing a distribution difference item and an integrity constraint item, and performing weighted summation to obtain an optimization target of the knowledge graph.

[0029] The inputting the multi-modal knowledge representation vector into an improved SHAP model, and computing a synergistic effect among the features by the improved SHAP model based on a feature combination assessment mechanism to obtain a feature importance score; screening the multi-modal knowledge representation vector according to the feature importance score, and selecting the features with the importance score exceeding a preset feature threshold as influence features includes:

[0030] inputting the multi-modal knowledge representation vector into the improved SHAP model, the improved SHAP model adopting a sliding window to generate a feature combination pool, and computing a feature combination weight based on mutual information, relevancy and time-order relevancy of features in the feature combination pool; mapping the feature combination weight by the improved SHAP model using a nonlinear mapping layer, and computing the synergistic effect among the features based on the feature combination assessment mechanism;

[0031] constructing a feature contribution matrix based on the synergistic effect, matrix elements of the feature contribution matrix representing the contribution of the features to a sample; utilizing time-order attenuation factors to perform time-order weighting on the feature contribution matrix to obtain a time-order contribution, and computing an initial SHAP value based on the time-order contribution; computing a modal adaptive weight according to variances of different modal features, and utilizing the modal adaptive weights to perform weighted fusion on the initial SHAP value to obtain the feature importance score; and

[0032] screening the multi-modal knowledge representation vector according to the feature importance score, computing the preset feature threshold based on a mean value and a standard deviation of the feature importance score, and dynamically adjusting the preset feature threshold in combination with a change rate of the feature importance score; and constructing a feature dependence graph to analyze an association relationship among the features, selecting the features with the importance score exceeding the preset feature threshold as the influence features based on a feature group effect and stability constraints, the influence features being used for guiding subsequent feature optimization.

[0033] The extracting gradient information in a feature space according to the optimization target, and inputting the gradient information into an Adam optimizer to generate an error correction strategy, generating an adversarial sample for verification based on the error correction strategy, dynamically adjusting an inference depth according to the verification result by a dual-DQN network, and finally, performing comparative verification on optimized graph features and the graph feature fingerprint to obtain an optimization effect to form a complete optimization closed loop includes:

[0034] extracting the gradient information in the feature space according to the optimization target, computing a local gradient through a multi-scale window, and obtaining a smooth gradient by a gradient cutting mechanism and index moving average; inputting the smooth gradient into the Adam optimizer, the Adam optimizer adaptively adjusting a learning rate based on a gradient norm and generating the error correction strategy through second-order momentum correction;

[0035] generating the adversarial sample based on the error correction strategy, the adversarial sample ensuring validity by disturbance boundary constraints and being prevented from being degraded through diversity constraints; inputting the adversarial sample into a multi-level verification network for performing quality assessment to obtain a verification result;

[0036] inputting the verification result into the dual-DQN network, the dual-DQN network constructing a state space based on a current inference depth, the verification result and resource occupation, constructing an action space based on a depth adjustment step length and a jump connection parameter, and dynamically adjusting the inference depth through the state space and the action space to obtain an adjusted inference depth; and

[0037] optimizing the knowledge graph according to the adjusted inference depth to obtain optimized graph features; comparing the optimized graph features with the graph feature fingerprint, and computing the KL divergence and the topological structure

similarity of feature distribution to serve as assessment indexes; and feeding back the assessment indexes to the Adam optimizer for optimizing parameter update, so as to form the complete optimization closed loop.

[0038] The optimizing the knowledge graph according to the adjusted inference depth to obtain optimized graph features; comparing the optimized graph features with the graph feature fingerprint, and computing the KL divergence and the topological structure similarity of feature distribution to serve as assessment indexes; and feeding back the assessment indexes to the Adam optimizer for optimizing parameter update, so as to form the complete optimization closed loop includes:

[0039] constructing a knowledge graph optimization model according to the adjusted inference depth, the knowledge graph optimization model computing a node attention weight and an edge relationship attention weight through a hierarchical attention mechanism, and performing feature recombination on the knowledge graph based on the node attention weight and the edge relationship attention weight to obtain optimized graph features;

[0040] comparing the optimized graph features with the pre-constructed graph feature fingerprint, and computing the KL divergence of feature distribution by a kernel function estimator, the kernel function estimator adding an adaptive smoothing factor to prevent a zero probability problem; constructing a multi-scale structure descriptor to describe graph structure features, and computing topological structure similarity based on the multi-scale structure descriptor; performing weighted fusion on the KL divergence and the topological structure similarity to obtain an assessment index; and

[0041] feeding back the assessment index to the Adam optimizer, the Adam optimizer updating parameters by an adaptive learning rate and reducing parameter update fluctuation by gradient accumulation; constructing a convergence criterion based on a change value of the assessment index and a parameter change value, and judging whether optimization is converged according to the convergence criterion; and if the optimization is not converged, feeding back the updated parameters to the knowledge graph optimization model for next round of optimization until convergence is finished and the closed loop is optimized.

[0042] In a second aspect,

[0043] the smart brain-driven data knowledge graph construction system for complex ecology provided by the embodiments of the present invention includes:

[0044] a first unit, configured to collect scenario data by a deep feature sensing network, acquire an original multi-dimensional sequence containing a service process, a processing rule and asset association, perform semantic parsing on the original

multi-dimensional sequence to generate an initial knowledge vector, construct an initial graph structure containing entity nodes and relational edges based on the initial knowledge vector, and encode the initial graph structure by a graph neural network to obtain a graph feature fingerprint used for feature distribution representation and a feature set used for knowledge discovery;

[0045] a second unit, configured to compute a similarity matrix among entities in the initial graph structure based on the graph feature fingerprint, screen entity pairs with relevancy greater than a preset similarity threshold from the similarity matrix by the feature set, execute multi-hop path inference on the screened entity pairs to generate an inference path set, input the inference path set into a recurrent neural network to generate an entity association confidence index, perform multi-modal feature extraction on data in the initial graph structure and the inference path set based on the entity association confidence index, the multi-modal features including text features, image features and audio features, and align and fuse the multi-modal features by a cross-modal attention mechanism to generate a unified multi-modal knowledge representation vector; and

[0046] a third unit, configured to input the multi-modal knowledge representation vector and the entity association confidence index into a multi-level feature network, generate a multi-modal knowledge representation vector, the multi-modal knowledge representation vector including an entity feature vector, a relational feature vector and a sub-graph feature vector; based on the multi-modal knowledge representation vector, identify influence features by an SHAP method, input the influence features into a problem evolution graph, compute a transmission probability by causal inference, and perform comparative analysis on the transmission probability and the graph feature fingerprint to generate an optimization target; extract gradient information in a feature space according to the optimization target, and input the gradient information into an Adam optimizer to generate an error correction strategy, generate an adversarial sample for verification based on the error correction strategy, dynamically adjust an inference depth according to the verification result by a dual-DQN network, and finally, perform comparative verification on optimized graph features and the graph feature fingerprint to obtain an optimization effect to form a complete optimization closed loop.

[0047] In a third aspect,

[0048] an electronic device provided by the embodiment of the present invention includes:

[0049] a processor; and

[0050] a memory configured to store processor executable instructions; and

[0051] the processor is configured to call instructions stored in the memory to execute the abovementioned method.

[0052] In a fourth aspect,

[0053] the embodiment of the present invention provides a computer-readable storage medium having computer program instructions stored therein, when executed by a processor, the computer program instructions execute the abovementioned method.

[0054] The present application have the beneficial effects that:

[0055] 1. the knowledge graph construction efficiency and accuracy can be improved: by performing data collection and semantic parsing through the deep feature sensing network, and in combination with the graph neural network encoding, the initial graph structure can be efficiently constructed and the key features can be extracted. By multi-hop path inference and the recurrent neural network, the entity association can be identified more accurately, and more comprehensive knowledge representation is constructed through the multi-modal feature extraction and the cross-modal attention mechanism.

[0056] 2. The knowledge discovery and inference capabilities can be enhanced: the similarity computation and entity screening are performed based on the graph feature fingerprint, and the potential entity association and knowledge can be discovered in combination with the multi-hop path inference. The influence features are identified by the SHAP method, and the transmission probability is computed by causal inference, so that the causal relationship between knowledge can be better understood, and more accurate inference is performed.

[0057] 3. Continuous optimization and self-adaptation of the knowledge graph can be realized: by comparing and analyzing the transmission probability and the graph feature fingerprints, the optimization target is generated, and the error correction strategy is generated by the Adam optimizer. The inference depth is dynamically adjusted through adversarial sample verification and the dual-DQN network to form the complete optimization closed loop, so that the knowledge graph can be continuously optimized and adapted, and the robustness and generalization capability of the knowledge graph are improved.

BRIEF DESCRIPTION OF THE DRAWINGS

[0058] FIG. 1 is a flowchart of a smart brain-driven data knowledge graph construction method for complex ecology according to an embodiment of the present invention; and

[0059] FIG. 2 is a schematic structural diagram of a smart brain-driven data

knowledge graph construction system for complex ecology according to an embodiment of the present invention.

DETAILED DESCRIPTION

[0060] In order to make the purpose, technical solution and advantages of the embodiment of the present invention clearer, the technical solutions in the embodiments of the present invention will be clearly and completely described in combination with the accompanying drawings in the embodiments of the present invention. Definitely, the described embodiments are only a part of the embodiments of the present invention, not all of them. Based on the embodiments in the present invention, all other embodiments obtained by those of ordinary skill in the art without carrying out creative labor fall within the scope of protection of the present invention.

[0061] The technical solutions of the present invention are described in detail by specific embodiments below. The following specific embodiments can be combined with each other, and the same or similar concepts or processes may not be repeated in some embodiments.

[0062] FIG. 1 is a flowchart of a smart brain-driven data knowledge graph construction method for complex ecology according to an embodiment of the present invention, as shown in FIG. 1, the method includes:

[0063] S11: collecting scenario data by a deep feature sensing network, acquiring an original multi-dimensional sequence containing a service process, a processing rule and asset association, performing semantic parsing on the original multi-dimensional sequence to generate an initial knowledge vector, constructing an initial graph structure containing entity nodes and relational edges based on the initial knowledge vector, and encoding the initial graph structure by a graph neural network to obtain a graph feature fingerprint used for feature distribution representation and a feature set used for knowledge discovery;

[0064] S12: computing a similarity matrix among entities in the initial graph structure based on the graph feature fingerprint, screening entity pairs with relevancy greater than a preset similarity threshold from the similarity matrix by the feature set, executing multi-hop path inference on the screened entity pairs to generate an inference path set, inputting the inference path set into a recurrent neural network to generate an entity association confidence index, performing multi-modal feature extraction on data in the initial graph structure and the inference path set based on the entity association confidence index, the multi-modal features including text features, image features and audio features, and aligning and fusing the multi-modal features by a cross-modal attention mechanism to

generate a unified multi-modal knowledge representation vector; and

[0065] S13: inputting the multi-modal knowledge representation vector and the entity association confidence index into a multi-level feature network, generating a multi-modal knowledge representation vector, the multi-modal knowledge representation vector including an entity feature vector, a relational feature vector and a sub-graph feature vector; based on the multi-modal knowledge representation vector, identifying influence features by an SHAP method, inputting the influence features into a problem evolution graph, computing a transmission probability by causal inference, and performing comparative analysis on the transmission probability and the graph feature fingerprint to generate an optimization target; extracting gradient information in a feature space according to the optimization target, and inputting the gradient information into an Adam optimizer to generate an error correction strategy, generating an adversarial sample for verification based on the error correction strategy, dynamically adjusting an inference depth according to the verification result by a dual-DQN network, and finally, performing comparative verification on optimized graph features and the graph feature fingerprint to obtain an optimization effect to form a complete optimization closed loop.

[0066] In an optional embodiment, the computing a similarity matrix among entities in the initial graph structure based on the graph feature fingerprint, screening entity pairs with relevancy greater than a preset similarity threshold from the similarity matrix by the feature set includes:

[0067] representing the graph feature fingerprint to be a high-dimensional feature vector, the high-dimensional feature vector containing a graph structure feature, a semantic distribution feature and a topological attribute feature, and the knowledge feature containing an attribute feature matrix of the entity nodes;

[0068] performing feature transformation on the attribute feature matrix by a linear transformation unit to generate a query matrix, a key matrix and a value matrix, the query matrix being obtained by multiplying a query weight matrix and the attribute feature matrix, the key matrix being obtained by multiplying a key weight matrix and the attribute feature matrix, and the value matrix being obtained by multiplying a value weight matrix and the attribute feature matrix;

[0069] dividing the query matrix, the key matrix and the value matrix into a plurality of attention heads, each attention head independently computing an attention score, performing transposition multiplication on the query matrix and the key matrix and carrying out dimension normalization to obtain an attention weight of each attention head, multiplying the attention weights by the value matrix to obtain head outputs, and splicing a

plurality of head outputs to form a multi-head attention output;

[0070] based on the multi-head attention output, computing a similarity relationship among the entity pairs through cosine similarity, and generating a similarity matrix, each element of the similarity matrix representing a similarity score of the corresponding entity pair;

[0071] setting a dynamic similarity threshold according to distribution features in the graph feature fingerprint, and performing bidirectional screening verification on the similarity matrix through the dynamic similarity threshold, the bidirectional screening verification requiring that forward similarity and reverse similarity of the entity pairs both meet the dynamic similarity threshold; and

[0072] filtering the entity pairs subjected to the bidirectional verification in combination with constraints in the knowledge feature, so as to generate final related entity pairs.

[0073] A graph feature fingerprint based entity relevancy analysis method is used for computing the similarity among the entities in the initial graph structure and screening entity pairs with relevancy greater than the preset threshold.

[0074] Firstly, the graph feature fingerprint is constructed. The graph feature fingerprint is represented as the high-dimensional feature vector and contains the graph structure feature, the semantic distribution feature and the topological attribute feature. For example, the graph structure feature can be a degree, centrality and the like of the entity nodes; the semantic distribution feature can be the probability distribution of the entity nodes under different themes; and the topological attribute feature can be information such as positions of the entity nodes in the graph and neighbor nodes. Moreover, the knowledge features are extracted, containing the attribute feature matrix of the entity nodes. For example, the attributes of a film entity can include directors, actors, type, release date and the like.

[0075] Secondly, feature transformation is performed on the attribute feature matrix by the linear transformation unit to generate the query matrix, the key matrix and the value matrix. Specifically, three different weight matrixes are multiplied by the attribute feature matrix to obtain the query matrix, the key matrix and the value matrix. For example, if the query weight matrix is [[0.1, 0.2, 0.3, 0.4], and [0.2, 0.3, 0.4, 0.1]], a first row of [Zhang san, Li si, story, 2022] of the attribute feature matrix is multiplied by the query weight matrix to obtain a first row of the query matrix.

[0076] Thirdly, the query matrix, the key matrix and the value matrix are divided into a plurality of attention heads. The attention score is independently computed for each

attention head. For each attention head, the query matrix and the key matrix are subjected to transformation multiplication and dimension normalization to obtain the attention weight. The attention weight is multiplied by the value matrix to obtain the head output. The plurality of head outputs are spliced to form the multi-head attention output. For example, it is assumed that the query matrix, the key matrix and the value matrix are all divided into two attention heads. The computed attention weight of the first attention head is multiplied by the value matrix to obtain a first head output. The computed attention weight of the second attention head is multiplied by the value matrix to obtain a second head output. The two head outputs are spliced to form the multi-head attention output.

[0077] The similarity relationship among the entity pairs is computed through the cosine similarity on the basis of the multi-head attention output so as to generate the similarity matrix. Each element of the similarity matrix represents the similarity score of the corresponding entity pair. For example, if the multi-head attention outputs of a film A and a film B are correspondingly [0.1, 0.2, 0.3] and [0.2, 0.3, 0.4], their cosine similarity is a product obtained by dividing a dot product of the two vectors by moduli thereof.

[0078] A dynamic similarity threshold is set according to the distribution features in the graph feature fingerprint. Bidirectional verification screening is performed on the similarity matrix by the dynamic similarity threshold. The bidirectional verification screening is performed on the premise that forward similarity and reverse similarity of the entity pairs both meet the dynamic similarity threshold. For example, if the similarity between the film A and the film B is 0.8, the similarity between the film B and the film A is 0.9 and the dynamic similarity threshold is 0.7, it will be determined that both the film A and the film B meet the bidirectional verification.

[0079] The entity pairs passing the bidirectional verification are filtered in combination with the semantic constraints in the knowledge features to generate the final related entity pairs. For example, if the semantic constraints are that the film types are ensured to be the same, the film A is related to a film C because they are both feature films. Although the film A and the film B pass the bidirectional verification, the film A and the film B are finally filtered out due to different types.

[0080] Through the solution of the present invention:

[0081] the accuracy of entity relevancy analysis can be improved: by combining the graph structure features, the semantic distribution features and the topological attribute features, the graph feature fingerprint is constructed, which can more comprehensively describe the relevancy among the entities, thereby improving the accuracy of entity relevancy analysis. The robustness of entity relevancy analysis can be enhanced: the

dynamic similarity threshold can be adaptively adjusted according to the distribution features of different entities, so that the robustness of entity relevancy analysis is enhanced, and the limitation caused by a fixed threshold is avoided. The efficiency of entity relevancy analysis can be improved: by the multi-head attention mechanism, the similarity among the entities can be computed in parallel, thus the efficiency of entity relevancy analysis is improved, and the performance is more outstanding especially on large-scale graph data.

[0082] In an optional embodiment, the performing multi-modal feature extraction on data in the initial graph structure and the inference path set based on the entity association confidence index by a BERT-Large model, a Vision Transformer model and a Wav2Vec model, the multi-modal features including text features, image features and audio features, and aligning and fusing the multi-modal features by a cross-modal attention mechanism to generate a unified multi-modal knowledge representation vector includes:

[0083] encoding the inputted text data by a BERT-Large pre-training model, extracting context semantic features of the text data through a multi-level Transformer structure, the context semantic features being represented as a text feature matrix, and generating a text pre-training feature vector based on the text feature matrix;

[0084] dividing the inputted image data into an image block sequence with the size of 16 * 16 pixels by the Vision Transformer model, inputting the image block sequence to a self-attention layer after position encoding, extracting a global visual feature of the image block sequence, and jointly encoding the global visual features and the text pre-training feature vector to obtain an image-text joint feature vector;

[0085] constructing an attention guidance matrix based on the image-text joint feature vector, the attention guidance matrix being used for processing the inputted audio data, inputting the audio data into the Wav2Vec model, extracting the audio features through a multi-level convolutional network and the Transformer structure, and performing feature alignment on the audio features and the image-text joint feature vector to generate a multi-modal alignment feature;

[0086] computing an inter-modal similarity matrix for the multi-modal alignment feature, constructing a cross-modal attention network based on the similarity matrix, and outputting attention weight distribution through the cross-modal attention network; and

[0087] inputting the attention weight distribution into a linear projection layer, and mapping different modal features to the unified semantic space through the linear projection layer to generate the unified multi-modal knowledge representation vector.

[0088] A knowledge graph inference method based on multi-modal knowledge representation is used for improving the accuracy and efficiency of the knowledge graph

inference. The core of the method is to extract the text, image and audio features by the BERT-Large model, the Vision Transformer model and the Wav2Vec model respectively, and fuse the multi-modal features through the cross-modal attention mechanism to generate the unified multi-modal knowledge representation vector for subsequent knowledge inference.

[0089] Firstly, the entities and relationship in the knowledge graph are initialized to construct the initial graph structure. For example, if the graph contains entities “cat” and “animal” and the relationship “belong to”, an initial triple (cat, belong to, animal) can be constructed. Moreover, an inference path set is generated according to the requirements of an inference task, for example, if it is to infer “what sound does a cat make”, an inference path of “cat -> belongs to -> animal -> makes sound -> meow”. The data in the initial graph structure and the inference path set will be associated with a confidence index to represent the reliability level of the data in the inference process. For example, according to an existing knowledge base, the confidence of “cat belongs to animal” may be very high, and the confidence of “cat makes sound” is relatively low.

[0090] Secondly, the multi-modal feature extraction is performed on the data in the initial graph structure and the inference path set. Taking the entity “cat” as an example, textual description of “a small mammal that is usually kept as a pet by humans” is extracted, and the text is encoded by the BERT-Large model to obtain a text feature vector. Moreover, an image of “cat” is extracted, such as a picture of the cat, and the image features are extracted by the Vision Transformer model. In addition, the audio of “cat” is extracted, such as a piece of audio of meowing of the cat, and the audio feature is extracted by the Wav2Vec model.

[0091] Thirdly, the multi-modal feature fusion is performed. Joint encoding is performed on the text feature vector generated by the BER-Large model and the image feature extracted by the Vision Transformer model so as to obtain an image-text joint feature vector. The audio data is inputted into the Wav2Vec model for extracting the audio features, and feature alignment is performed on the audio features and the image-text joint feature vector so as to generate the multi-modal alignment features. The inter-modal similarity matrix of the multi-modal alignment features is computed, and the cross-modal attention network is constructed on the basis of the similarity matrix so as to obtain the attention weight distribution. Finally, the attention weight distribution is inputted into the linear projection layer, the features of different modals are mapped to a uniform semantic space to generate the uniform multi-modal knowledge representation vector. For example, the finally obtained multi-modal knowledge representation vector of “cat” fuses the text,

image and audio information, so that the concept of “cat” can be represented more comprehensively.

[0092] Fourthly, the knowledge inference is performed on the basis of the generated uniform multi-modal knowledge representation vector. For example, if it is to infer “what sound does a cat make”, the multi-modal knowledge representation vector of the “cat” can be compared with the multi-modal knowledge representation vectors of other related entities (such as “animal”, and “sound”) and computed, so that the sound of the “cat” is finally deduced to be “meow”.

[0093] Through the solution of the present invention:

[0094] the comprehensiveness of knowledge representation can be improved: by fusing the multi-modal information such as texts, images and audios, the knowledge can be represented more comprehensively, and deviation caused by single-modal information is avoided, and therefore, the accuracy and integrity of knowledge representation are improved. For example, the picture, text description and sound of the cart are fused, which can express the concept of “cat” more completely than using text description alone. The accuracy of knowledge inference can be enhanced: the multi-modal knowledge representation vector contains richer semantic information, and the association between entities and relationship can be better captured, so that the accuracy of knowledge inference is improved. For example, when inferring “what sound does a cat make”, the multi-modal representation vector fusing the audio of the cat can more accurately deduce the answer. The application range of the knowledge graph can be expanded: the multi-modal knowledge representation can better handle multimedia data, thus the application of the knowledge graph in the fields of image recognition, voice recognition and the like is expanded. For example, the multi-modal knowledge graph containing the images, sounds and texts can be constructed by the method, and it can be applied to applications such as intelligent question and answer, and information retrieval.

[0095] In an optional embodiment, the based on the multi-modal knowledge representation vector, identifying influence features by an SHAP method, inputting the influence features into a problem evolution graph, computing a transmission probability by causal inference, and performing comparative analysis on the transmission probability and the graph feature fingerprint to generate an optimization target includes:

[0096] inputting the multi-modal knowledge representation vector into an improved SHAP model, and computing a synergistic effect among the features by the improved SHAP model based on a feature combination assessment mechanism to obtain a feature importance score; screening the multi-modal knowledge representation vector according to

the feature importance score, and selecting the features with the importance score exceeding a preset feature threshold as influence features;

[0097] constructing the problem evolution graph by taking the influence features as nodes, node attributes of the problem evolution graph containing the feature value and the importance score, and constructing a directed edge based on a time-order relationship among the features; inputting the problem evolution graph into a causal inference network, computing a conditional probability among the nodes by the causal inference network using a variational inference method, and performing Monte Carlo sampling to obtain the probability distribution of the transmission path; and

[0098] comparing the probability distribution of the transmission path with a pre-stored graph feature fingerprint, and computing KL divergence to obtain state difference distribution; and constructing an optimization target function based on the state difference distribution, the optimization target function containing a distribution difference item and an integrity constraint item, and performing weighted summation to obtain an optimization target of the knowledge graph.

[0099] The multi-modal knowledge graph optimization method is used for enhancing the completeness and accuracy of the knowledge graph. The core idea of the method is to analyze the transmission probability of influence features in a problem evolution graph and compare them with the existing graph feature fingerprints, thereby identifying the parts to be optimized.

[0100] Firstly, the feature selection is performed on the multi-modal knowledge representation vector. The multi-modal knowledge representation vector is inputted into the improved SHAP model. The improved SHAP model is not a conventional SHAP, but takes into account the assessment mechanism of feature combination, and it can capture the synergistic effect among the features, so that the importance of each feature can be assessed more accurately. For example, an image has the features of “cat” and the features of “sofa”, and the combination of the two features possibly hints higher-level semantic information of “cat is on sofa”, thus the improved SHAP model can capture the synergistic effect. Through this model, the importance score of each feature is obtained. A preset feature threshold, such as 0.05, is set, and the features with the importance score exceeding the threshold are screened as the influence features. If the multi-modal vector contains four features of “cat”, “sofa”, “table”, and “dog”, and the importance scores of the four features are 0.1, 0.08, 0.02, and 0.01 respectively, the “cat” and the “sofa” are selected as the influence features.

[0101] Secondly, the problem evolution graph is constructed. The influence features screened in the previous step are taken as the nodes of the graph. Each node contains two

attributes: the feature value and the importance score. The nodes are connected by directed edges, and the directions of the edges represent a time-order relationship among the features. For example, the node of “cat” points to the node of “sofa”, which represents a state that an action of “cat” is prior to a state of “cat is on sofa”. If the importance score of “cat” is 0.1 and the importance score of “sofa” is 0.08, the edges from the node of “cat” to the node of “sofa” can be represented as (“cat”, 0.1) → (“sofa”, 0.08).

[0102] Thirdly, the causal inference and transmission probability computation are performed. The constructed problem evolution graph is inputted into the causal inference network. The network computes the conditional probability among the nodes by the variational inference method, and simulates the transmission path of the features in the graph by a Monte Carlo sampling method to obtain the probability distribution of the transmission path. For example, if 1000 times of sampling are simulated, and the results of 800 times of sampling show that “cat” will jump to “sofa”, the transmission probability from “cat” to “sofa” will be 0.8.

[0103] The computed transmission path probability distribution is compared with the pre-stored graph feature fingerprint. The graph feature fingerprint is pre-stored and represents the feature transmission probability distribution of common modes in the knowledge graph. The KL divergence between the transmission probability distribution and the graph feature fingerprint is computed to obtain the state difference distribution. The KL divergence is used for measuring the difference degree between the two probability distributions.

[0104] Fourthly, the optimization target of the knowledge graph is generated. The optimization target function is constructed based on the state difference distribution. The function contains two items: the distribution difference item and the integrity constraint item. The distribution difference item is used for measuring the difference between the current knowledge graph and an ideal state, and the integrity constraint item is used for ensuring the integrity and consistency of the knowledge graph. The weighted summation is performed on the two items to obtain the final optimization target.

[0105] Through the solution of the present invention:

[0106] the accuracy of the knowledge graph can be improved: by identifying and correcting the feature transmission path inconsistent with a preset mode, error information in the knowledge graph can be effectively corrected, thereby improving the accuracy of the knowledge graph. The integrity of the knowledge graph can be enhanced: by the method, the missing information in the knowledge graph can be identified and supplemented completely, so that the coverage range and the information amount of the knowledge graph

are enhanced. The interpretability of the knowledge graph can be improved: by analyzing the causal relationship and the transmission path among the features, internal logic and evolution laws of the knowledge graph can be better understood, thus its interpretability is improved.

[0107] In an optional embodiment, the inputting the multi-modal knowledge representation vector into an improved SHAP model, and computing a synergistic effect among the features by the improved SHAP model based on a feature combination assessment mechanism to obtain a feature importance score; screening the multi-modal knowledge representation vector according to the feature importance score, and selecting the features with the importance score exceeding a preset feature threshold as influence features includes:

[0108] inputting the multi-modal knowledge representation vector into the improved SHAP model, the improved SHAP model adopting a sliding window to generate a feature combination pool, and computing a feature combination weight based on mutual information, relevancy and time-order relevancy of features in the feature combination pool; mapping the feature combination weight by the improved SHAP model using a nonlinear mapping layer, and computing the synergistic effect among the features based on the feature combination assessment mechanism;

[0109] constructing a feature contribution matrix based on the synergistic effect, matrix elements of the feature contribution matrix representing the contribution of the features to a sample; utilizing time-order attenuation factors to perform time-order weighting on the feature contribution matrix to obtain a time-order contribution, and computing an initial SHAP value based on the time-order contribution; computing a modal adaptive weight according to variances of different modal features, and utilizing the modal adaptive weights to perform weighted fusion on the initial SHAP value to obtain the feature importance score; and

[0110] screening the multi-modal knowledge representation vector according to the feature importance score, computing the preset feature threshold based on a mean value and a standard deviation of the feature importance score, and dynamically adjusting the preset feature threshold in combination with a change rate of the feature importance score; and constructing a feature dependence graph to analyze an association relationship among the features, selecting the features with the importance score exceeding the preset feature threshold as the influence features based on a feature group effect and stability constraints, the influence features being used for guiding subsequent feature optimization.

[0111] Firstly, the multi-modal knowledge representation vector is preprocessed. The

data of different modals, such as images, texts and audios, is transformed into a unified vector representation form. For example, the features of image data can be extracted by the convolutional neural network, the text data can be transformed into vectors by a word embedding model, and the audio data can be transformed into vectors by spectrogram feature extraction. These vectors are spliced together to form the multi-modal knowledge representation vector. If a multi-modal sample contains three modals of the images, and the texts and the audios, the vector dimensions of the modals are 128, 64 and 32 respectively, the final multi-modal knowledge representation vector dimension will be 224.

[0112] Secondly, the preprocessed multi-modal knowledge representation vector is inputted into the improved SHAP model. Firstly, the feature combination pool is generated by the model using the sliding window mechanism. For example, if the dimension of the multi-modal knowledge representation vector is 224 and the size of the sliding window is set to be 3, 222 feature combinations will be generated, and each combination contains 3 continuous features. Then, the weight of the feature combination is computed based on the mutual information, relevancy and time-order relevancy of the features in the feature combination pool. For example, the higher the occurrence frequency of the two features in time is, the higher the time-order relevancy between the two features is, and the higher the weight of the corresponding feature combination is. If the mutual information of a certain feature combination is 0.8, the relevancy is 0.9 and the time-order relevancy is 0.7, the weight of the feature combination can be computed according to a preset weight computation method, such as weighted average. If the weights of the mutual information, the relevancy and the time-order relevancy are 0.3, 0.4 and 0.3 respectively, the weight of the feature combination will be 0.3 * 0.8 + 0.4 * 0.9 + 0.3 * 0.7 = 0.78. The improved SHAP model maps the feature combination weights by the nonlinear mapping layer, for example, the weights are mapped to be between 0 and 1 by the Sigmoid function. And then, the synergistic effect among the features is computed based on the feature combination assessment mechanism, namely, the influence of the combined action of the plurality of features on the result is computed. For example, if the combination weight of the image features and the text features is high, it indicates that there is a strong synergistic effect between the two features.

[0113] Thirdly, the feature contribution matrix is constructed on the basis of the synergistic effect. Each element of the matrix represents the contribution of the features for the samples. For example, if a certain image feature has a great influence on a classification result of the samples, the contribution value corresponding to the feature will be relatively high. The time-order weighting is performed on the feature contribution matrix through the

time-order attenuation factors so as to obtain the time-order contribution. For example, for the time-order data, the influence of the nearest feature for the current result is relatively high, thus the corresponding time-order attenuation factor is relatively high. If the contribution of a certain feature at a moment t is 0.8 and the time-order attenuation factor is 0.9, the time-order contribution of the feature at a moment t-1 will be 0.8 * 0.9 = 0.72. The initial SHAP value is computed on the basis of the time-order contribution. The modal adaptive weight is computed according to variances of different modal features. For example, if the variance of the image feature is relatively high, it indicates that the distinction degree of the image feature for the sample is relatively high and the corresponding modal adaptive weight is relatively high. The weighted fusion is performed on the initial SHAP value by the modal adaptive weight so as to obtain the feature importance score.

[0114] Fourthly, the multi-modal knowledge representation vector is screened according to the feature importance score. The preset feature threshold is computed based on the mean value and the standard deviation of the feature importance score, for example, the mean value is added with the standard deviation to serve as the threshold. The preset feature threshold is dynamically adjusted in combination with the change rate of the feature importance score. For example, if the change rate of the feature importance scores is relatively high, it indicates that the importance of features is continuously changed, and it is needed to dynamically adjust the threshold. The feature dependency graph is constructed to analyze the association relationship among the features, for example, if two features frequently occur at the same time, there will be a relatively high association relationship between the two features. Based on the feature group effect and the stability constraint, the features with the importance score exceeding the preset feature threshold are selected as the influence features. For example, if a certain feature group has a very great influence on the classification result of the sample, the features in this group are more possibly selected. If the preset feature threshold is 0.8 and the importance score of a certain feature is 0.9, the feature will be selected as the influence feature. The influence features are used for guiding subsequent feature optimization.

[0115] Through the solution of the present invention:

[0116] the accuracy of feature screening can be improved: by considering the synergistic effect, time-order information and modal difference among the features, the features with important influence on the result can be identified more accurately, and key information is prevented from being missed, so that the accuracy of feature screening is improved. The interpretability of the model can be enhanced: by constructing the feature

dependency graph, the association relationship among the features can be analyzed, and an influence mechanism of the features on the result can be better understood, so that the interpretability of the model is enhanced. The performance of the model can be optimized: by screening the important influence features, redundant information can be reduced, the computation complexity can be reduced, and the efficiency and performance of the model can be improved.

[0117] In an optional embodiment, the extracting gradient information in a feature space according to the optimization target, and inputting the gradient information into an Adam optimizer to generate an error correction strategy, generating an adversarial sample for verification based on the error correction strategy, dynamically adjusting an inference depth according to the verification result by a dual-DQN network, and finally, performing comparative verification on optimized graph features and the graph feature fingerprint to obtain an optimization effect to form a complete optimization closed loop includes:

[0118] extracting the gradient information in the feature space according to the optimization target, computing a local gradient through a multi-scale window, and obtaining a smooth gradient by a gradient cutting mechanism and index moving average; inputting the smooth gradient into the Adam optimizer, the Adam optimizer adaptively adjusting a learning rate based on a gradient norm and generating the error correction strategy through second-order momentum correction;

[0119] generating the adversarial sample based on the error correction strategy, the adversarial sample ensuring validity by disturbance boundary constraints and being prevented from being degraded through diversity constraints; inputting the adversarial sample into a multi-level verification network for performing quality assessment to obtain a verification result;

[0120] inputting the verification result into the dual-DQN network, the dual-DQN network constructing a state space based on a current inference depth, the verification result and resource occupation, constructing an action space based on a depth adjustment step length and a jump connection parameter, and dynamically adjusting the inference depth through the state space and the action space to obtain an adjusted inference depth; and

[0121] optimizing the knowledge graph according to the adjusted inference depth to obtain optimized graph features; comparing the optimized graph features with the graph feature fingerprint, and computing the KL divergence and the topological structure similarity of feature distribution to serve as assessment indexes; and feeding back the assessment indexes to the Adam optimizer for optimizing parameter update, so as to form

the complete optimization closed loop.

[0122] The target is to optimize the knowledge graph features to be closer to preset graph feature fingerprint. According to the method, adversarial training and deep dynamic adjustment strategy are adopted, and the closed-loop feedback mechanism is utilized to continue optimization.

[0123] The feature gradient extraction and smoothing are performed. Firstly, the gradient information is extracted in the feature space according to the optimization target, for example, improving the performance of the knowledge graph in a downstream task. In order to acquire more comprehensive gradient information, the local gradient is computed by the multi-scale window. Specifically, windows with different sizes slide on the feature graph to compute the gradient in each window, and then the gradients are aggregated. In order to avoid gradient explosion or disappearance, a gradient cutting mechanism is adopted to limit gradient amplitudes within a certain range. Finally, in order to reduce gradient fluctuation, an exponential moving average method is adopted to perform smoothing processing on the gradients, so as to obtain smooth gradients. For example, the sizes of three windows are set to be 3x3, 5x5 and 7x7 respectively, weighted average is performed after the local gradient is computed respectively, and the weights are 0.2, 0.3 and 0.5 respectively. Then the gradient amplitudes are limited between [-1, 1]. Finally, exponential moving average is adopted, the attenuation rate is set to be 0.9, and thus the smooth gradients are obtained.

[0124] The error correction strategy is generated. The smooth gradients are inputted into the Adam optimizer. The Adam optimizer adaptively adjusts the learning rate according to the gradient norm and corrects the gradient direction by the second-order momentum so as to generate the error correction strategy. For example, the initial learning rate is set to be 0.001, the first-order momentum attenuation rate is set to be 0.9, and the second-order momentum attenuation rate is set to be 0.999.

[0125] The adversarial samples are generated and verified. The adversarial samples are generated based on the error correction strategy. The generation process of the adversarial samples needs to meet the disturbance boundary constraints and the diversity constraints. The disturbance boundary constraints limit the difference between the adversarial samples and the original sample, thereby ensuring the validity of the adversarial samples. The diversity constraints support the generation of different adversarial samples to prevent sample degradation. For example, the disturbance boundary constraints are set to be 0.1, and random noises are added in the gradient direction to realize the diversity constraints. The generated adversarial samples are inputted to the multi-level verification

network for performing quality assessment so as to obtain the verification result, for example, a loss value change caused by the adversarial samples.

[0126] The inference depth is dynamically adjusted. The verification result is inputted to the dual-DQN network. The state space of the dual-DQN network is composed of the current inference depth, the verification result and resource occupation. The action space consists of the depth adjustment step length and the jump connection parameter. The dual-DQN network selects the optimal action according to the current state so as to dynamically adjust the inference depth. For example, the state space includes the current inference depth (for example, 5), the verification result (for example, the loss value is increased by 0.05), and the resource occupation (for example, an internal memory usage amount). The action space includes the depth adjustment step length (for example, one layer is added or reduced) and the jump connection parameter (for example, whether jump connection is enabled or not).

[0127] Graph feature optimization and verification are performed. The knowledge graph is optimized according to the adjusted inference depth so as to obtain the optimized graph features. The optimized graph features are compared with preset graph feature fingerprint, and the KL divergence of feature distribution and the topological structure similarity are computed as the assessment indexes. For example, the KL divergence between the optimized graph features and fingerprint features is computed to be 0.02, and the topological structure similarity is computed to be 0.95.

[0128] Closed-loop feedback and optimization are performed. The assessment indexes are fed back to the Adam optimizer for adjusting the parameters of the optimizer, such as learning rate and momentum, so as to form the complete optimized closed loop. For example, if KL divergence is relatively large, the learning rate will be reduced; and if topological structure similarity is relatively low, a gradient computation mode will be adjusted.

[0129] Through the solution of the present invention:

[0130] the quality of the knowledge graph can be improved: by adversarial training and dynamic depth adjustment, the features of the knowledge graph are closer to the preset fingerprint, so that the quality of the knowledge graph and the performance in the downstream task are improved. The adaptive allocation of resources can be realized: the dual-DQN network dynamically adjusts inference depth according to the verification result and resource occupation, thus the adaptive allocation of the resources is realized and the resource waste is avoided. The complete optimized closed loop can be constructed: the assessment indexes are fed back to the optimizer to form the complete optimized closed

loop, therefore, continuous optimization is realized and the quality of the knowledge graph is continuously improved.

[0131] In an optional embodiment, the optimizing the knowledge graph according to the adjusted inference depth to obtain optimized graph features; comparing the optimized graph features with the graph feature fingerprint, and computing the KL divergence and the topological structure similarity of feature distribution to serve as assessment indexes; and feeding back the assessment indexes to the Adam optimizer for optimizing parameter update, so as to form the complete optimization closed loop includes:

[0132] constructing a knowledge graph optimization model according to the adjusted inference depth, the knowledge graph optimization model computing a node attention weight and an edge relationship attention weight through a hierarchical attention mechanism, and performing feature recombination on the knowledge graph based on the node attention weight and the edge relationship attention weight to obtain optimized graph features;

[0133] comparing the optimized graph features with the pre-constructed graph feature fingerprint, and computing the KL divergence of feature distribution by a kernel function estimator, the kernel function estimator adding an adaptive smoothing factor to prevent a zero probability problem; constructing a multi-scale structure descriptor to describe graph structure features, and computing topological structure similarity based on the multi-scale structure descriptor; performing weighted fusion on the KL divergence and the topological structure similarity to obtain an assessment index; and

[0134] feeding back the assessment index to the Adam optimizer, the Adam optimizer updating parameters by an adaptive learning rate and reducing parameter update fluctuation by gradient accumulation; constructing a convergence criterion based on a change value of the assessment index and a parameter change value, and judging whether optimization is converged according to the convergence criterion; and if the optimization is not converged, feeding back the updated parameters to the knowledge graph optimization model for performing next round of optimization until convergence is finished and a closed loop is optimized.

[0135] Firstly, it is needed to construct a knowledge graph and the corresponding graph feature fingerprint in the preparation stage. The graph feature fingerprint can be understood as the description for ideal graph features, which can be pre-defined and extracted from the high-quality knowledge graph or constructed by an expert according to domain knowledge. For example, if the target is to construct a knowledge graph about the film, the graph feature fingerprint can contain information such as the film type, actors,

directors, year of release and the like and the relationship among them. Moreover, it is needed to set the initial inference depth, for example, it is set to be 3, which represents that it takes into account the relationship within three hops between the nodes in the graph.

[0136] Secondly, the iterative optimization stage is performed. According to the current inference depth, the knowledge graph optimization model is constructed. The model adopts the hierarchical attention mechanism. Specifically, the model will compute the attention weight of each node, for example, different nodes are endowed with different importance according to the degree, centrality and other indexes of the nodes. Then, the model will compute the attention weight of the edge relationship among the nodes, for example, different edges are endowed with different importance according to the type, strength and other indexes of the edges. Based on the computed node and edge attention weights, the knowledge graph is subjected to feature recombination to obtain the optimized graph features. For example, if there is a "common director" relationship between the film A and the film B, and the attention weight of the relationship is high, the feature representation of the film A and the film B will be more similar in the optimized graph features.

[0137] Thirdly, the optimized graph features are compared with the pre-constructed graph feature fingerprint. In order to assess the difference of feature distribution, the kernel function estimator is utilized to compute the KL divergence of feature distribution. The kernel function estimator introduces an adaptive smoothing factor to prevent zero probability problem, that is, to avoid computation errors caused by some feature values not appearing in the sample. For example, if the fingerprint contains a certain common film type, and there is no film of such type in the current graph, the adaptive smoothing factor can avoid the situation that the KL divergence computation is extremely rare. Moreover, in order to assess the similarity of the graph topology, multi-scale structure descriptors are constructed to depict the graph structure features, such as node degree distribution, and clustering coefficients. Based on these descriptors, the topological structure similarity between the optimized graph and the fingerprint graph is computed. Finally, the KL divergence and the topological structure similarity are subjected to weight fusion to obtain a comprehensive assessment index. The weight can be adjusted according to actual needs.

[0138] The assessment indexes are fed back to the Adam optimizer. The model parameters are updated by the Adam optimizer using the adaptive learning rate, and the fluctuation of parameter update is reduced through the gradient accumulation, so that the optimization process is more stable. The convergence criterion is established according to the change value of the assessment indexes and the parameter change value. For example, a

threshold can be set, and when the change value of the assessment indexes and the parameter change value are both less than the threshold, it can be considered that the optimization process is converged. If the optimization process is not converged, the updated parameters are fed back to the knowledge graph optimization model, the inference depth is adjusted, for example, the inference depth is increased to 4, the above steps are repeated, and then next round of optimization is performed until the convergence condition is met.

[0139] Through the solution of the present invention:

[0140] the quality of the knowledge graph can be improved. By comparing and optimizing the knowledge graph with the preset graph feature fingerprint, the integrity, accuracy and consistency of the knowledge graph can be effectively improved, and thus the knowledge graph better conforms to expectation and better supports downstream application. The controllability can be enhanced. By adjusting the inference depth and introducing the attention mechanism, the optimization process of the knowledge graph can be controlled more finely, so that the optimization result better conforms to the requirements of a specific field. Automatic optimization can be realized. In the method, the Adam optimizer and the adaptive learning rate are utilized, so that the parameters can be automatically adjusted without manual intervention, the optimization efficiency is improved, and the use threshold is reduced.

[0141] FIG. 2 is a schematic structural diagram of a smart brain-driven data knowledge graph construction system for complex ecology according to an embodiment of the present invention, as shown in FIG. 2, the system includes:

[0142] a first unit, configured to collect scenario data by a deep feature sensing network, acquire an original multi-dimensional sequence containing a service process, a processing rule and asset association, perform semantic parsing on the original multi-dimensional sequence to generate an initial knowledge vector, construct an initial graph structure containing entity nodes and relational edges based on the initial knowledge vector, and encode the initial graph structure by a graph neural network to obtain a graph feature fingerprint used for feature distribution representation and a feature set used for knowledge discovery;

[0143] a second unit, configured to compute a similarity matrix among entities in the initial graph structure based on the graph feature fingerprint, screen entity pairs with relevancy greater than a preset similarity threshold from the similarity matrix by the feature set, execute multi-hop path inference on the screened entity pairs to generate an inference path set, input the inference path set into a recurrent neural network to generate an entity

association confidence index, perform multi-modal feature extraction on data in the initial graph structure and the inference path set based on the entity association confidence index, the multi-modal features including text features, image features and audio features, and align and fuse the multi-modal features by a cross-modal attention mechanism to generate a unified multi-modal knowledge representation vector; and

[0144] a third unit, configured to input the multi-modal knowledge representation vector and the entity association confidence index into a multi-level feature network, generate a multi-modal knowledge representation vector, the multi-modal knowledge representation vector including an entity feature vector, a relational feature vector and a sub-graph feature vector; based on the multi-modal knowledge representation vector, identify influence features by an SHAP method, input the influence features into a problem evolution graph, compute a transmission probability by causal inference, and perform comparative analysis on the transmission probability and the graph feature fingerprint to generate an optimization target; extract gradient information in a feature space according to the optimization target, and input the gradient information into an Adam optimizer to generate an error correction strategy, generate an adversarial sample for verification based on the error correction strategy, dynamically adjust an inference depth according to the verification result by a dual-DQN network, and finally, perform comparative verification on optimized graph features and the graph feature fingerprint to obtain an optimization effect to form a complete optimization closed loop.

[0145] In a third aspect,

[0146] an electronic device provided by the embodiment of the present invention includes:

[0147] a processor; and

[0148] a memory configured to store processor executable instructions;

[0149] the processor is configured to call instructions stored in the memory to execute the abovementioned method.

[0150] In a fourth aspect,

[0151] the embodiment of the present invention provides a computer-readable storage medium having computer program instructions stored therein, when executed by a processor, the computer program instructions execute the abovementioned method.

[0152] The present invention can be a method, an apparatus, a system, and/or a computer program product. The computer program product can include a computer-readable storage medium that contains computer-readable program instructions for executing the present invention in various aspects.

[0153] Finally, it is to be noted that: the above embodiments are only used for illustrating the technical solution of the present invention, not to limit it. Although the present invention is described in detail with reference to the aforementioned embodiments, those of ordinary skill in the art should understand that it may still modify the technical solutions recorded in the aforementioned embodiments, or replace some or all of the technical features thereof. These modifications or substitutions do not make the essence of the corresponding technical solution out of the scope of the technical solutions of each embodiment of the present invention.

CLAIMS

We claim:

1. A smart brain-driven data knowledge graph construction method for complex ecology, comprising:

collecting scenario data by a deep feature sensing network, acquiring an original multi-dimensional sequence containing a service process, a processing rule and asset association, performing semantic parsing on the original multi-dimensional sequence to generate an initial knowledge vector, constructing an initial graph structure containing entity nodes and relational edges based on the initial knowledge vector, and encoding the initial graph structure by a graph neural network to obtain a graph feature fingerprint used for feature distribution representation and a feature set used for knowledge discovery;

computing a similarity matrix among entities in the initial graph structure based on the graph feature fingerprint, screening entity pairs with relevancy greater than a preset similarity threshold from the similarity matrix by the feature set, executing multi-hop path inference on the screened entity pairs to generate an inference path set, inputting the inference path set into a recurrent neural network to generate an entity association confidence index, performing multi-modal feature extraction on data in the initial graph structure and the inference path set based on the entity association confidence index, the multi-modal features comprising text features, image features and audio features, and aligning and fusing the multi-modal features by a cross-modal attention mechanism to generate a unified multi-modal knowledge representation vector; and

inputting the multi-modal knowledge representation vector and the entity association confidence index into a multi-level feature network, generating a multi-modal knowledge representation vector, the multi-modal knowledge representation vector comprising an entity feature vector, a relational feature vector and a sub-graph feature vector; based on the multi-modal knowledge representation vector, identifying influence features by an SHAP method, inputting the influence features into a problem evolution graph, computing a transmission probability by causal inference, and performing comparative analysis on the transmission probability and the graph feature fingerprint to generate an optimization target; extracting gradient information in a feature space according to the optimization target, and inputting the gradient information into an Adam optimizer to generate an error correction strategy, generating an adversarial sample for verification based on the error correction strategy, dynamically adjusting an inference depth according to the verification result by a dual-DQN network, and finally, performing comparative verification on optimized graph

features and the graph feature fingerprint to obtain an optimization effect to form a complete optimization closed loop.

2. The method according to claim 1, wherein the computing a similarity matrix among entities in the initial graph structure based on the graph feature fingerprint, screening entity pairs with relevancy greater than a preset similarity threshold from the similarity matrix by the feature set comprises:

representing the graph feature fingerprint to be a high-dimensional feature vector, the high-dimensional feature vector containing a graph structure feature, a semantic distribution feature and a topological attribute feature, and the knowledge feature containing an attribute feature matrix of the entity nodes;

performing feature transformation on the attribute feature matrix by a linear transformation unit to generate a query matrix, a key matrix and a value matrix, the query matrix being obtained by multiplying a query weight matrix and the attribute feature matrix, the key matrix being obtained by multiplying a key weight matrix and the attribute feature matrix, and the value matrix being obtained by multiplying a value weight matrix and the attribute feature matrix;

dividing the query matrix, the key matrix and the value matrix into a plurality of attention heads, each attention head independently computing an attention score, performing transposition multiplication on the query matrix and the key matrix and carrying out dimension normalization to obtain an attention weight of each attention head, multiplying the attention weights by the value matrix to obtain head outputs, and splicing a plurality of head outputs to form a multi-head attention output;

based on the multi-head attention output, computing a similarity relationship among the entity pairs through cosine similarity, and generating a similarity matrix, each element of the similarity matrix representing a similarity score of the corresponding entity pair;

setting a dynamic similarity threshold according to distribution features in the graph feature fingerprint, and performing bidirectional screening verification on the similarity matrix through the dynamic similarity threshold, the bidirectional screening verification requiring that forward similarity and reverse similarity of the entity pairs both meet the dynamic similarity threshold; and

filtering the entity pairs subjected to the bidirectional verification in combination with constraints in the knowledge feature, so as to generate final related entity pairs.

3. The method according to claim 1, wherein the performing multi-modal feature extraction on data in the initial graph structure and the inference path set based on the

entity association confidence index by a BERT-Large model, a Vision Transformer model and a Wav2Vec model, the multi-modal features comprising text features, image features and audio features, and aligning and fusing the multi-modal features by a cross-modal attention mechanism to generate a unified multi-modal knowledge representation vector comprises:

encoding the inputted text data by a BERT-Large pre-training model, extracting context semantic features of the text data through a multi-level Transformer structure, the context semantic features being represented as a text feature matrix, and generating a text pre-training feature vector based on the text feature matrix;

dividing the inputted image data into an image block sequence with the size of 16 * 16 pixels by the Vision Transformer model, inputting the image block sequence to a self-attention layer after position encoding, extracting a global visual feature of the image block sequence, and jointly encoding the global visual features and the text pre-training feature vector to obtain an image-text joint feature vector;

constructing an attention guidance matrix based on the image-text joint feature vector, the attention guidance matrix being used for processing the inputted audio data, inputting the audio data into the Wav2Vec model, extracting the audio features through a multi-level convolutional network and the Transformer structure, and performing feature alignment on the audio features and the image-text joint feature vector to generate a multi-modal alignment feature;

computing an inter-modal similarity matrix for the multi-modal alignment feature, constructing a cross-modal attention network based on the similarity matrix, and outputting attention weight distribution through the cross-modal attention network; and

inputting the attention weight distribution into a linear projection layer, and mapping different modal features to the unified semantic space through the linear projection layer to generate the unified multi-modal knowledge representation vector.

4. The method according to claim 1, wherein the based on the multi-modal knowledge representation vector, identifying influence features by an SHAP method, inputting the influence features into a problem evolution graph, computing a transmission probability by causal inference, and performing comparative analysis on the transmission probability and the graph feature fingerprint to generate an optimization target comprises:

inputting the multi-modal knowledge representation vector into an improved SHAP model, and computing a synergistic effect among the features by the improved SHAP model based on a feature combination assessment mechanism to obtain a feature

importance score; screening the multi-modal knowledge representation vector according to the feature importance score, and selecting the features with the importance score exceeding a preset feature threshold as influence features;

constructing the problem evolution graph by taking the influence features as nodes, node attributes of the problem evolution graph containing the feature value and the importance score, and constructing a directed edge based on a time-order relationship among the features; inputting the problem evolution graph into a causal inference network, computing a conditional probability among the nodes by the causal inference network using a variational inference method, and performing Monte Carlo sampling to obtain the probability distribution of the transmission path; and

comparing the probability distribution of the transmission path with a pre-stored graph feature fingerprint, and computing KL divergence to obtain state difference distribution; and constructing an optimization target function based on the state difference distribution, the optimization target function containing a distribution difference item and an integrity constraint item, and performing weighted summation to obtain an optimization target of the knowledge graph.

5. The method according to claim 4, wherein the inputting the multi-modal knowledge representation vector into an improved SHAP model, and computing a synergistic effect among the features by the improved SHAP model based on a feature combination assessment mechanism to obtain a feature importance score; screening the multi-modal knowledge representation vector according to the feature importance score, and selecting the features with the importance score exceeding a preset feature threshold as influence features comprises:

inputting the multi-modal knowledge representation vector into the improved SHAP model, the improved SHAP model adopting a sliding window to generate a feature combination pool, and computing a feature combination weight based on mutual information, relevancy and time-order relevancy of features in the feature combination pool; mapping the feature combination weight by the improved SHAP model using a nonlinear mapping layer, and computing the synergistic effect among the features based on the feature combination assessment mechanism;

constructing a feature contribution matrix based on the synergistic effect, matrix elements of the feature contribution matrix representing the contribution of the features to a sample; utilizing time-order attenuation factors to perform time-order weighting on the feature contribution matrix to obtain a time-order contribution, and computing an initial

SHAP value based on the time-order contribution; computing a modal adaptive weight according to variances of different modal features, and utilizing the modal adaptive weights to perform weighted fusion on the initial SHAP value to obtain the feature importance score; and

screening the multi-modal knowledge representation vector according to the feature importance score, computing the preset feature threshold based on a mean value and a standard deviation of the feature importance score, and dynamically adjusting the preset feature threshold in combination with a change rate of the feature importance score; and constructing a feature dependence graph to analyze an association relationship among the features, selecting the features with the importance score exceeding the preset feature threshold as the influence features based on a feature group effect and stability constraints, the influence features being used for guiding subsequent feature optimization.

6. The method according to claim 1, wherein the extracting gradient information in a feature space according to the optimization target, and inputting the gradient information into an Adam optimizer to generate an error correction strategy, generating an adversarial sample for verification based on the error correction strategy, dynamically adjusting an inference depth according to the verification result by a dual-DQN network, and finally, performing comparative verification on optimized graph features and the graph feature fingerprint to obtain an optimization effect to form a complete optimization closed loop comprises:

extracting the gradient information in the feature space according to the optimization target, computing a local gradient through a multi-scale window, and obtaining a smooth gradient by a gradient cutting mechanism and index moving average; inputting the smooth gradient into the Adam optimizer, the Adam optimizer adaptively adjusting a learning rate based on a gradient norm and generating the error correction strategy through second-order momentum correction;

generating the adversarial sample based on the error correction strategy, the adversarial sample ensuring validity by disturbance boundary constraints and being prevented from being degraded through diversity constraints; inputting the adversarial sample into a multi-level verification network for performing quality assessment to obtain a verification result;

inputting the verification result into the dual-DQN network, the dual-DQN network constructing a state space based on a current inference depth, the verification result and resource occupation, constructing an action space based on a depth adjustment step length

and a jump connection parameter, and dynamically adjusting the inference depth through the state space and the action space to obtain an adjusted inference depth; and

optimizing the knowledge graph according to the adjusted inference depth to obtain optimized graph features; comparing the optimized graph features with the graph feature fingerprint, and computing the KL divergence and the topological structure similarity of feature distribution to serve as assessment indexes; and feeding back the assessment indexes to the Adam optimizer for optimizing parameter update, so as to form the complete optimization closed loop.

7. The method according to claim 6, wherein the optimizing the knowledge graph according to the adjusted inference depth to obtain optimized graph features; comparing the optimized graph features with the graph feature fingerprint, and computing the KL divergence and the topological structure similarity of feature distribution to serve as assessment indexes; and feeding back the assessment indexes to the Adam optimizer for optimizing parameter update, so as to form the complete optimization closed loop comprises:

constructing a knowledge graph optimization model according to the adjusted inference depth, the knowledge graph optimization model computing a node attention weight and an edge relationship attention weight through a hierarchical attention mechanism, and performing feature recombination on the knowledge graph based on the node attention weight and the edge relationship attention weight to obtain optimized graph features;

comparing the optimized graph features with the pre-constructed graph feature fingerprint, and computing the KL divergence of feature distribution by a kernel function estimator, the kernel function estimator adding an adaptive smoothing factor to prevent a zero probability problem; constructing a multi-scale structure descriptor to describe graph structure features, and computing topological structure similarity based on the multi-scale structure descriptor; performing weighted fusion on the KL divergence and the topological structure similarity to obtain an assessment index; and

feeding back the assessment index to the Adam optimizer, the Adam optimizer updating parameters by an adaptive learning rate and reducing parameter update fluctuation by gradient accumulation; constructing a convergence criterion based on a change value of the assessment index and a parameter change value, and judging whether optimization is converged according to the convergence criterion; and if the optimization is not converged, feeding back the updated parameters to the knowledge graph optimization model for

performing next round of optimization until convergence is finished and a closed loop is optimized.

8. A smart brain-driven data knowledge graph construction system for complex ecology, applied to implement the method according to any one of claims 1 to 7, comprising:

a first unit, configured to collect scenario data by a deep feature sensing network, acquire an original multi-dimensional sequence containing a service process, a processing rule and asset association, perform semantic parsing on the original multi-dimensional sequence to generate an initial knowledge vector, construct an initial graph structure containing entity nodes and relational edges based on the initial knowledge vector, and encode the initial graph structure by a graph neural network to obtain a graph feature fingerprint used for feature distribution representation and a feature set used for knowledge discovery;

a second unit, configured to compute a similarity matrix among entities in the initial graph structure based on the graph feature fingerprint, screen entity pairs with relevancy greater than a preset similarity threshold from the similarity matrix by the feature set, execute multi-hop path inference on the screened entity pairs to generate an inference path set, input the inference path set into a recurrent neural network to generate an entity association confidence index, perform multi-modal feature extraction on data in the initial graph structure and the inference path set based on the entity association confidence index, the multi-modal features comprising text features, image features and audio features, and align and fuse the multi-modal features by a cross-modal attention mechanism to generate a unified multi-modal knowledge representation vector; and

a third unit, configured to input the multi-modal knowledge representation vector and the entity association confidence index into a multi-level feature network, generate a multi-modal knowledge representation vector, the multi-modal knowledge representation vector comprising an entity feature vector, a relational feature vector and a sub-graph feature vector; based on the multi-modal knowledge representation vector, identify influence features by an SHAP method, input the influence features into a problem evolution graph, compute a transmission probability by causal inference, and perform comparative analysis on the transmission probability and the graph feature fingerprint to generate an optimization target; extract gradient information in a feature space according to the optimization target, and input the gradient information into an Adam optimizer to generate an error correction strategy, generate an adversarial sample for verification based

on the error correction strategy, dynamically adjust an inference depth according to the verification result by a dual-DQN network, and finally, perform comparative verification on optimized graph features and the graph feature fingerprint to obtain an optimization effect to form a complete optimization closed loop.

9. An electronic device, comprising a processor; and

a memory configured to store processor executable instructions; and

the processor is configured to call instructions stored in the memory to execute the method according to any one of claims 1 to 7.

10. A computer-readable storage medium having computer program instructions stored therein, wherein when executed by a processor, the computer program instructions execute the method according to any one of claims 1 to 7.