

RULE MINING METHOD AND APPARATUS

TECHNICAL FIELD

[0001] Embodiments of this specification relate to the field of machine learning technologies, and in particular, to a rule mining method and apparatus.

BACKGROUND

[0002] With development of Internet technologies, online services are applied to increasingly more scenarios, to effectively provide more convenient services for users. In the online services, graph-structured data such as social data, risk control data, code quality data, or check association data includes a large amount of node-edge interaction topology information. The information can be effectively modeled and used in service scenarios such as risk fraud, code vulnerability scanning, and check rule association mining. In a security risk control scenario, numerical association rules (a rule body includes more than two columns) are mainly consistency rules. However, there is still a scenario (a linear scenario or an inequality scenario) with more complex computational association. If only a single-column rule and a consistency rule included in the rule body are used, it is difficult to completely cover all types of numerical columns. In addition, considering that rule relationships between different numerical columns are unclear, more rules are configured based on manual experience after service code logic is manually sorted. In this case, not only human and material resources are wasted, but also efficiency is low, and comprehensiveness of rule configuration cannot be ensured. Therefore, an effective solution is urgently needed to resolve the foregoing problem.

SUMMARY

[0003] In view of this, embodiments of this specification provide a rule mining method. One or more embodiments of this specification further relate to a rule mining apparatus, a computing device, and a computer-readable storage medium, to resolve a technical disadvantage in the existing technology.

[0004] According to a first aspect of the embodiments of this specification, a rule mining method is provided and includes:

obtaining service data, and selecting a numerical computable column associated with a target service scenario from the service data;

constructing a plurality of graph nodes based on a computation operator and the numerical computable column;

5 inputting the plurality of graph nodes to a rule prediction model, and processing the plurality of graph nodes through a fusion unit and a link prediction unit in the rule prediction model, to obtain a graph node link relationship; and

generating a numerical column computation rule corresponding to the target service scenario based on the graph node link relationship.

10 **[0005]** According to a second aspect of the embodiments of this specification, a rule mining apparatus is provided and includes:

an obtaining module, configured to obtain service data, and select a numerical computable column associated with a target service scenario from the service data;

15 a construction module, configured to construct a plurality of graph nodes based on a computation operator and the numerical computable column;

a prediction module, configured to input the plurality of graph nodes to a rule prediction model, and process the plurality of graph nodes through a fusion unit and a link prediction unit in the rule prediction model, to obtain a graph node link relationship; and

20 a generation module, configured to generate a numerical column computation rule corresponding to the target service scenario based on the graph node link relationship.

[0006] According to a third aspect of the embodiments of this specification, a computing device is provided, including:

a memory and a processor.

25 **[0007]** The memory is configured to store computer-executable instructions. When the processor is configured to execute the computer-executable instructions, the steps of the rule mining method are implemented.

[0008] According to a fourth aspect of the embodiments of this specification, a computer-readable storage medium is provided. The computer-readable storage medium stores computer-executable instructions. When the instructions are executed by a processor, the steps of the rule mining method are implemented.

30 **[0009]** In the rule mining method provided in the embodiments, to improve rule mining efficiency and ensure comprehensiveness of rule mining, after the full service data is obtained, the numerical computable column associated with the target service scenario can be selected from the service data, to mine the computation rule corresponding to the target service scenario. Based on this, considering

that the computation rule needs to be mined, the plurality of graph nodes can be first constructed based on the computation operator and the numerical computable column, to implement mining processing based on a graph structure. That is, the plurality of graph nodes are input to the rule prediction model, so that the plurality of graph nodes are processed through the fusion unit and the link prediction unit in the rule prediction model, to obtain the graph node link relationship based on a processing result, so as to mine all possible node link relationships by using the rule prediction model. Based on this, by splicing the graph node link relationships, the numerical column computation rule associated with the target service scenario can be obtained based on a splicing result. In this way, all link relationships can be accurately and efficiently predicted by using the rule prediction model without manual participation, and the computation rule can be constructed. This greatly reduces searchable space of the computation rule of the numerical computable column, and improves a generalization capability of rule mining.

BRIEF DESCRIPTION OF DRAWINGS

- [0010] FIG. 1 is a schematic diagram of a rule mining method according to an embodiment of this specification;
- [0011] FIG. 2 is a flowchart of a rule mining method according to an embodiment of this specification;
- [0012] FIG. 3 is a schematic diagram of feature generation in a rule mining method according to an embodiment of this specification;
- [0013] FIG. 4 is a schematic diagram of constructing a graph node in a rule mining method according to an embodiment of this specification;
- [0014] FIG. 5 is a flowchart of a processing process of a rule mining method according to an embodiment of this specification;
- [0015] FIG. 6 is a schematic diagram of a structure of a rule mining apparatus according to an embodiment of this specification; and
- [0016] FIG. 7 is a block diagram of a structure of a computing device according to an embodiment of this specification.

DESCRIPTION OF EMBODIMENTS

- [0017] Many specific details are described in the following descriptions to facilitate full understanding of this specification. However, this specification can be implemented in many different manners from those described herein. A person skilled in the art can make similar promotion without

departing from the content of this specification. Therefore, this specification is not limited to the specific implementations disclosed below.

[0018] Terms used in one or more embodiments of this specification are merely used to describe specific embodiments, and are not intended to limit the one or more embodiments of this specification. The terms "a" and "the" of singular forms used in one or more embodiments of this specification and the appended claims are also intended to include plural forms, unless otherwise specified in the context clearly. It should be further understood that the term "and/or" used in one or more embodiments of this specification refers to and includes any or all possible combinations of one or more associated listed items.

[0019] It should be understood that although terms "first", "second", and the like can be used in one or more embodiments of this specification to describe various types of information, the information is not limited to these terms. These terms are merely used to differentiate between information of the same type. For example, "first" can also be referred to as "second" without departing from the scope of one or more embodiments of this specification, and similarly, "second" can also be referred to as "first". Depending on the context, for example, the word "if" used herein can be explained as "while", "when", or "in response to determining".

[0020] First, terms used in one or more embodiments of this specification are explained.

[0021] Graph-structured data (Graph Data) describes a complex non-linear data storage structure that cannot be described in Euclidean space. The graph-structured data can quickly describe concepts and a relationship between the concepts in a physical world. The graph-structured data can be modeled to form a corresponding node-edge-node relationship, to aggregate a large amount of knowledge information. A modeling method designed by using the graph-structured data has been widely applied to fields such as intelligent search, intelligent question answering, personalized recommendation, intelligence analysis, and anti-fraud. Graph data with a plurality of edge-node relationships is referred to as a heterogeneous graph.

[0022] A graph neural network (GNN) is a class of deep learning-based methods for processing graph domain information, and is a deep learning architecture that can run on a social network or other graph-structured data-based networks, and therefore is defined as a graph structure-based general neural network. Because of good performance and interpretability of the graph neural network, the GNN has recently become a widely applied graph analysis method.

[0023] Graph representation learning is a process of performing embedding learning on a relationship between a node and an edge in graph-structured data based on a graph neural network, and is also a process of mapping graph data (which is usually a high-dimensional dense matrix) to a low-dimensional dense vector. In graph representation learning, a topology structure of a graph, a

connection relationship between different types of vertices (at least one of a node and an edge has a plurality of types), and other information (for example, a sub-graph or a connection edge) need to be captured. Finally, low-dimensional vector representations at a node level, an edge level, and a graph level can be separately obtained.

5 [0024] This specification provides a rule mining method. This specification further relates to a rule mining apparatus, a computing device, and a computer-readable storage medium, which are described in detail in the following embodiments.

[0025] In actual applications, for computational rule mining in a service scenario, linear regression modeling is currently mainly used to generate a check rule for $ax+by=cz$, where a, b, and
10 c are respectively coefficients, and x, y, and z are selection of a computed column. This method has obvious disadvantages and limitations: 1. A large quantity of computing resources need to be consumed to determine the x, y, and z columns through Brute-force search. When there is a combination of n numerical columns and m pieces of data, complexity is higher. 2. For each selected computed column combination, a linear model needs to be trained. As a quantity of filtering fields
15 increases, a quantity of models that need to be trained linearly increases. 3. Potential rules in actual rules cannot be effectively mined. Therefore, an effective solution is urgently needed to resolve the foregoing problem.

[0026] With reference to a schematic diagram shown in FIG. 1, in the rule mining method provided in this embodiment, to improve rule mining efficiency and ensure comprehensiveness of
20 rule mining, after full service data is obtained, a numerical computable column associated with a target service scenario can be selected from the service data, to mine a computation rule corresponding to the target service scenario. Based on this, considering that the computation rule needs to be mined, a plurality of graph nodes can be first constructed based on a computation operator and the numerical computable column, to implement mining processing based on a graph structure.
25 That is, the plurality of graph nodes are input to a rule prediction model, so that the plurality of graph nodes are processed through a fusion unit and a link prediction unit in the rule prediction model, to obtain a graph node link relationship based on a processing result, so as to mine all possible node link relationships by using the rule prediction model. Based on this, by splicing the graph node link relationships, the numerical column computation rule associated with the target service scenario can
30 be obtained based on a splicing result. In this way, all link relationships can be accurately and efficiently predicted by using the rule prediction model without manual participation, and the computation rule can be constructed. This greatly reduces searchable space of the computation rule of the numerical computable column, and improves a generalization capability of rule mining.

[0027] FIG. 2 is a flowchart of a rule mining method according to an embodiment of this specification. The method specifically includes the following steps.

[0028] Step S202: Obtain service data, and select a numerical computable column associated with a target service scenario from the service data.

5 [0029] The rule mining method provided in this embodiment can implement rule mining, that is, mine a computational rule, for data that has a computing requirement in data in any service scenario. The rule mining method can be applied to a resource transfer scenario, a fund payment scenario, a fund security management scenario, and the like. This is not limited herein in this embodiment. For
10 computation rule mining of a numerical computable column in any service scenario, refer to the same or corresponding descriptions in this embodiment. Details are not described herein in this embodiment.

[0030] Specifically, the service data is specifically full service data provided by a platform that provides a service. In actual applications, a service platform may simultaneously provide a plurality of services for a user, and data related to all of the services may be stored together because of a user
15 feature. When a computation rule is mined, the computation rule is usually a rule of a specific service scenario, and in a cross-service scenario, there may be no rule that can be mined. Therefore, after the service data is read, the numerical computable column associated with the target service scenario further needs to be selected from the service data, Correspondingly, the target service scenario is specifically a service scenario in which rule mining needs to be performed. The target service scenario
20 can be a resource transfer scenario, a fund payment scenario, a fund security management scenario, or the like. Correspondingly, the numerical computable column is specifically a numerical column whose result needs to be computed in the target service scenario, and is used to subsequently mine a computation rule based on the numerical computable column, so that after the computation rule is configured in the target service scenario, security management can be implemented on a resource or
25 a fund.

[0031] In actual applications, in a daily running process of the service platform, a large quantity of service scenario-based data such as user data, product data, fund data, code data, and data in another dimension is generated. The generated data is usually organized by using a common data cleaning procedure. For the massive data generated from a data source, to obtain an internal computation rule
30 based on this, the numerical computable column associated with the target service scenario can be first selected from the service data, so that subsequently computation rule mining can be performed for the target service scenario.

[0032] Further, in a process of selecting the numerical computable column associated with the corresponding target service scenario from the service data, considering that the service data includes

a relatively large quantity of data dimensions, and different dimensions corresponding to different data, to ensure that computation rule mining can be performed for the target service scenario, a numerical column of a specified type needs to be selected for subsequently determining the numerical computable column. In this embodiment, a specific implementation is as follows:

5 **[0033]** The service data is extracted from a target database; a numerical column of an enumeration column type is selected from a numerical column included in the service data as a conditional numerical column corresponding to the target service scenario; and the numerical computable column is selected from the conditional numerical column based on computation distribution information of the conditional numerical column.

10 **[0034]** Specifically, the target database is specifically a database that records all data of the service platform. The numerical computable column that needs to be used needs to be selected from the database. Therefore, a corresponding rule needs to be used for implementation. Correspondingly, the enumeration column type is specifically a numerical column whose value needs to be determined through enumeration. Correspondingly, the computation distribution information is specifically
15 distribution information corresponding to the numerical column of the enumeration column type, and is used to select the numerical computable column.

[0035] Based on this, to improve precision of selecting the numerical computable column, the service data can be first extracted from the target database. In this case, the read service data is full service data, and content included in the service data is more complex. Therefore, to select the
20 numerical computable column corresponding to the target service scenario for use, the numerical column of the enumeration column type can be selected as the conditional numerical column corresponding to the target service scenario from the numerical column included in the service data; and then the numerical computable column can be selected for use from the conditional numerical column based on the computation distribution information of the conditional numerical column.

25 **[0036]** For example, in a fund security computation scenario, to ensure security of user funds, all service data can be first read from the target database of the platform, then a numerical column whose value needs to be computed through enumeration can be selected from all the service data as a conditional numerical column, and then a data distribution status of each conditional numerical column can be determined based on computation distribution information of the conditional
30 numerical column, so that a numerical computable column used for rule mining can be selected based on this, to facilitate subsequent use.

[0037] In conclusion, the numerical computable column corresponding to the target service scenario is selected from the service data with reference to a numerical column type, so that all unrelated numerical columns can be initially filtered out from the full service data, and then the

numerical computable column is selected based on this and the computation distribution information, to ensure that all numerical columns in the target service scenario are included, so as to facilitate subsequent use in rule mining.

[0038] Based on this, in a process of selecting the conditional numerical column, to implement selection for the target service scenario, a computation rule can be mined in the numerical column corresponding to the target service scenario in a traversal manner. In this embodiment, a specific implementation is as follows:

[0039] The numerical column included in the service data is traversed based on a preset selection policy; and the numerical column of the enumeration column type is selected as the conditional numerical column corresponding to the target service scenario based on a traversal result.

[0040] Specifically, the selection policy is specifically a policy that can be used to subsequently determine and select the numerical computable column from the numerical column in the service data. The policy is used to traverse the numerical column, and determine, based on the traversal result, whether each numerical column is available, to select the conditional numerical column that satisfies a use requirement.

[0041] Based on this, when the conditional numerical column corresponding to the target service scenario is selected, considering that the service data includes a relatively large quantity of numerical columns, and different numerical columns correspond to different service scenarios, to ensure selection precision, the preset selection policy can be used for completion, so that after the numerical column is traversed based on the preset selection policy, the numerical column of the enumeration column type is selected as the conditional numerical column corresponding to the target service scenario based on the traversal result, to facilitate subsequent use.

[0042] In actual applications, considering that in the target service scenario, a size of a data source for rule mining is relatively large, all numerical column types included in the service data can be detected based on a statistical indicator or an NLP model that is set in advance, to select the numerical column of the enumeration column type from the service data as the conditional column, so as to facilitate subsequent use and determine the target service scenario in which rule mining needs to be performed. After the conditional column is obtained, whether each numerical column is a numerical computable column can be determined by using a distribution status of this type of numerical column, to select all numerical computable columns to form a numerical computable column set, so as to facilitate subsequent mining of all computation rules in the target service scenario based on this.

[0043] In the foregoing example, when the conditional numerical column is determined, all the service data can be detected based on the statistical indicator preset by the platform or the NLP model, to quickly select the conditional numerical column from the service data for use.

[0044] In conclusion, the numerical computable column is determined by using the preset selection policy, to select the numerical computable column corresponding to the target service scenario for use, so that more internal computation rules can be subsequently mined based on this, to save more human and material resources.

5 [0045] Step S204: Construct a plurality of graph nodes based on a computation operator and the numerical computable column.

[0046] Specifically, after the numerical computable column associated with the target service scenario is obtained, further, to fully mine all computation rules based on the numerical computable column, and ensure comprehensiveness and accuracy of rule mining, an edge between nodes in a graph structure can be determined by analyzing the graph structure, so that a link relationship is mapped based on the edge, and the computation rule is organized based on the link relationship, to improve rule mining efficiency and comprehensiveness. Before this, the graph structure needs to be first constructed based on the numerical computable column. Considering that in the target service scenario, a rule that needs to be mined is a computation rule, and the computation rule definitely involves a computation operator, the graph nodes in the graph structure can be first constructed based on the computation operator and the numerical computable column, so that a link relationship between the graph nodes is subsequently mined by using a graph neural network, to map a rule.

[0047] The computation operator is specifically an operator that is used when the numerical column is computed, and includes but is not limited to an operator such as +, −, *, or /. Correspondingly, the graph node is specifically a feature construction node formed after initial embedding is performed on the computation operator and the numerical computable column. The graph node corresponds to a node in graph-structured data, and is used to subsequently mine a computation rule, namely, the computation rule that can be mined from edge mapping between the graph nodes, based on the graph structure.

25 [0048] Further, in a process of constructing the graph node, considering that the rule that needs to be mined is a computation rule between the numerical computable columns, the graph nodes need to be constructed simultaneously with reference to the computation operator and the numerical computable column, to subsequently predict a node edge based on the graph structure, and reflect association between nodes based on the predicted edge, so as to map the computation rule between the numerical computable columns. In this embodiment, a specific implementation is as follows:

[0049] An operator feature corresponding to the computation operator and a numerical column feature corresponding to the numerical computable column are constructed; graphical processing is separately performed on the operator feature and the numerical column feature, to obtain an operator

feature node and a numerical column feature node; and the plurality of graph nodes are generated based on the operator feature node and the numerical column feature node.

[0050] Specifically, the operator feature is specifically a vector representation corresponding to the computation operator, and different computation operators correspond to different operator features. Correspondingly, the numerical column feature is specifically a vector representation corresponding to the numerical computable column. Correspondingly, the graphical processing is specifically a processing operation of converting a feature representation into a graph structure, and is used to convert the computation operator and the numerical computable column into nodes in a graph.

[0051] Based on this, when the graph node is constructed, to mine the computation rule based on the graph structure, the computation operator and the numerical computable column need to be converted into representations corresponding to the graph structure. Therefore, the operator feature corresponding to the computation operator and the numerical column feature corresponding to the numerical computable column can be first constructed. In this case, the feature representations corresponding to the computation operator and the numerical computable column are obtained. Then, graphical processing can be separately performed on the operator feature and the numerical column feature, to obtain the representations, that is, the operator feature node corresponding to the computation operator and the numerical column feature node corresponding to the numerical computable column, corresponding to the graph structure based on a graphical processing result. Finally, the plurality of graph nodes can be generated based on the operator feature node and the numerical column feature node, to form a graph structure that does not include an edge connection relationship, so as to facilitate subsequent prediction of the link relationship and mine the computation rule.

[0052] With reference to schematic diagrams shown in (a) and (b) in FIG. 3, after a plurality of numerical computable columns in the target service scenario are obtained, an initialization embedding vector corresponding to each numerical computable column can be first constructed, as shown in (a) in FIG. 3. A to-be-selected metric column is the numerical computable column. After the initialization embedding vector corresponding to each numerical computable column is obtained, because the feature representation is not an expression form corresponding to the graph structure, to fully mine an internal relationship between the nodes based on the graph structure, and map the computation rule, the feature representation corresponding to each numerical computable column can be graphically processed, that is, converted into an expression form corresponding to the graph structure. In this case, a node-edge topology graph corresponding to all of the numerical computable columns and the computation operator can be formed, as shown in (b) in FIG. 3. In the graph structure, there

are two types of node types: a to-be-selected metric column type corresponding to the numerical computable column and a computation operator type corresponding to the computation operator. With reference to (b) in FIG. 3, B, N, Q, X, M, T, and P are graph nodes corresponding to the numerical computable column, and +, −, *, and / are graph nodes corresponding to the computation operator.

5 This helps subsequently perform edge prediction based on the graph structure, to mine the computation rule between the numerical computable columns.

[0053] In conclusion, the computation operator and the numerical computable column are converted into representations corresponding to the graph structure, to form the graph structure, so that edge prediction is subsequently performed by using the nodes in the graph structure, and the
10 computation rule is mapped based on the predicted link relationship, to fully mine the computation rule for an internal relationship between the numerical computable columns, and ensure adequacy of rule mining.

[0054] Step S206: Input the plurality of graph nodes to a rule prediction model, and process the plurality of graph nodes through a fusion unit and a link prediction unit in the rule prediction model,
15 to obtain a graph node link relationship.

[0055] Specifically, after the plurality of graph nodes corresponding to the graph structure are obtained, further, to fully mine a link relationship between the graph nodes based on the graph structure to mine the computation rule between the numerical computable columns corresponding to the graph nodes, so as to map a more comprehensive computation rule, the node link relationship can
20 be predicted with reference to the graph neural network corresponding to the graph structure. That is, after the plurality of graph nodes corresponding to the graph structure are obtained, the plurality of graph nodes can be input to the pre-trained rule prediction model, to predict an edge between the plurality of graph nodes through the fusion unit and the link prediction unit in the rule prediction model, that is, predict a link relationship between the plurality of graph nodes, so that after the graph
25 node link relationship is obtained, computation rule integration is subsequently performed based on the graph node link relationship.

[0056] The rule prediction model is specifically a graph neural network, for example, a GNN model, that predicts an edge between graph nodes in a graph structure. Correspondingly, the link prediction unit is specifically a model for predicting an edge between graph nodes, that is, predicting
30 a link relationship between the nodes. Correspondingly, the fusion unit is specifically a unit that performs fusion on the link relationship predicted through the link prediction unit in the rule prediction model, to fuse link relationships in association with each other. The link prediction unit is configured to predict a link relationship between a graph node corresponding to the numerical computable column in the graph nodes and the computation operator. The fusion unit is configured

to perform fusion on the link relationship. Both prediction and fusion are prediction processing processes, and are used to form a graph node link relationship among the numerical computable column, the computation operator, and the numerical computable column. That is, the graph node link relationship is a link relationship formed among the numerical computable column, the computation operator, and the numerical computable column in the graph nodes. For example, the model predicts a link relationship between a graph node a and +, and a link relationship between b and +, and then performs fusion to obtain a graph node link relationship of a+b, to facilitate subsequent mining of a computation rule in this structure.

[0057] Further, when the graph node link relationship is predicted by using the rule prediction model, to ensure prediction precision, the fusion unit and the link prediction unit can be combined to predict an edge between nodes based on the graph structure, so as to map the graph node link relationship. In this embodiment, a specific implementation is as follows:

[0058] The plurality of graph nodes are input to the rule prediction model; a sub-link relationship between an operator node and a numerical column node in the plurality of graph nodes is computed through the link prediction unit in the rule prediction model; fusion processing is performed on the sub-link relationship through the fusion unit in the rule prediction model; and the graph node link relationship is generated based on a fusion processing result.

[0059] Specifically, the sub-link relationship is specifically a relationship between the computation operator and the numerical computable column, and is determined by using the computation operator node and the numerical column node. When fusion processing is performed on the sub-link relationship, fusion processing is performed on a sub-link relationship formed between the operator node and the numerical column node. This is used to mine sub-link relationships that can be fused, to splice at least one computation operator and at least two numerical computable columns together, so as to obtain the graph node link relationship, which is subsequently organized as a rule for use.

[0060] Based on this, when the graph node link relationship is predicted by using the rule prediction model, the plurality of graph nodes are actually input to the rule prediction model; and the sub-link relationship between the operator node and the numerical column node in the plurality of graph nodes is computed through the link prediction unit in the rule prediction model. In this case, a single-layer link relationship between the nodes is obtained by using the rule prediction model, and then fusion processing is performed on the sub-link relationship through the fusion unit in the rule prediction model, to generate the graph node link relationship based on the fusion processing result, so as to facilitate subsequent use.

[0061] For example, in the plurality of graph nodes, a, b, c, and d respectively represent operator nodes, and x_1 , x_2 , ..., and x_n respectively represent numerical column nodes. After the plurality of graph nodes are input to the rule prediction model, a link relationship between each numerical column node and each of the operator nodes a, b, c, and d is computed through the link prediction unit in the rule prediction model. After sub-link relationships $\{x_1a; x_1b; x_1c; \dots; x_nd\}$ between the operator nodes and the numerical column nodes are obtained, any two sub-link relationships can be fused through the fusion unit in the model, to obtain a graph node link relationship, for example, x_1ax_2 , based on a fusion result. The graph node link relationship is subsequently converted into a rule.

[0062] In conclusion, the single-layer link relationship between the nodes is mined through the link prediction unit in the rule prediction model, and then the single-layer link relationship is integrated through the fusion unit, to obtain the graph node link relationship between the corresponding graph nodes, to help subsequently mine the computation rule based on this.

[0063] Before this, to ensure that the rule prediction model can have relatively high prediction precision, training can be completed with reference to a sample set. In this embodiment, a specific implementation is as follows:

[0064] A sample set is obtained, where the sample set includes a labeled sample pair and an unlabeled sample; the labeled sample pair is processed by using an initial rule prediction model, and a relationship parameter is determined based on a processing result; the initial rule prediction model is updated based on the relationship parameter, to obtain an intermediate rule prediction model; and the unlabeled sample is processed by using the intermediate rule prediction model, and a parameter of the intermediate rule prediction model is adjusted based on a processing result until the rule prediction model that satisfies a training stop condition is obtained.

[0065] Specifically, the sample set is specifically a sample set used to train the rule prediction model, and includes a large quantity of samples. Correspondingly, a labeled sample is specifically a sample with a label. The unlabeled sample is specifically a sample without a label. To obtain, through training, a model that satisfies a use requirement, the labeled sample can be processed by using the initial rule prediction model, and the relationship parameter is determined and then fixed. The unlabeled sample is processed by using the model with the fixed relationship parameter, to optimize another parameter, so as to obtain the intermediate rule prediction model. Correspondingly, the training stop condition is specifically a condition for stopping training on the model, including but not limited to an iteration quantity condition, a loss value comparison condition, and the like. This is not limited herein in this embodiment.

[0066] Based on this, in a model training phase, to obtain a model with high enough precision through training, the sample set including the labeled sample pair and the unlabeled sample can be

first obtained. Then, the labeled sample pair is processed by using the initial rule prediction model, and the relationship parameter is determined based on the processing result, to further predict the unlabeled sample based on the relationship parameter, and update the initial rule prediction model based on the relationship parameter, to obtain the intermediate rule prediction model. Finally, the unlabeled sample is processed by using the intermediate rule prediction model, and the parameter of the intermediate rule prediction model is adjusted based on the processing result until the rule prediction model that satisfies the training stop condition is obtained.

[0067] Further, when supervised training is performed, to implement that the model can have relatively high prediction precision, prediction processing in each phase can be completed in the following manner. In this embodiment, a specific implementation is as follows:

[0068] The unlabeled sample is input to the intermediate rule prediction model; the unlabeled sample is processed through an intermediate fusion unit and an intermediate link prediction unit in the intermediate rule prediction model, to obtain a predicted link relationship; a predicted numerical column computation rule is generated based on the predicted link relationship; and when it is determined, based on the predicted numerical column computation rule, that the intermediate rule prediction model satisfies the training stop condition, the intermediate rule prediction model is used as the rule prediction model.

[0069] Based on this, when model training is performed, specific processing is as follows: The unlabeled sample is input to the intermediate rule prediction model; the unlabeled sample is processed through the intermediate fusion unit and the intermediate link prediction unit in the intermediate rule prediction model, to obtain the predicted link relationship; then the predicted numerical column computation rule is generated based on the predicted link relationship; and when it is determined, based on the predicted numerical column computation rule, that the intermediate rule prediction model satisfies the training stop condition, the intermediate rule prediction model can be used as the rule prediction model.

[0070] In specific implementation, in the graph structure, a rule that exists can be actually defined as an edge prediction task, that is, the edge between the nodes is predicted by using the rule prediction model. In this case, preliminary model training can be performed by using a small quantity of positive and negative samples, for example, 1% of positive and negative samples, and in this process, an embedding vector of the related graph node is updated. Therefore, inference is performed in a large quantity of unlabeled samples, to enable the model to learn an edge prediction capability. For example, with reference to a schematic diagram shown in (a) in FIG. 4, a metric column set $\{m|m \in M\}$ and an operator set $\{o|o \in O\}$ are defined, t metric columns and operator columns $\{s_1, s_2, \dots, s_t\}$ are given, initial embeddings of all nodes are given, and a hyper-parameter threshold

threshold_o/threshold_m is given. Pseudo code for predicting the edge between the nodes is shown in (b) FIG. 4.

[0071] Through iterative supervised learning, W in a fusion function and a LinkPred link prediction function is continuously updated to determine whether there is a link relationship between two nodes. In actual applications, the fusion function and the LinkPred link prediction function can use a leading-edge model structure, or can use a simple MLP. This is not limited herein in this embodiment. As shown in (a) in FIG. 4, a link relationship between a node N and a node $+$ is $s_m = \text{LinkPred}(W \cdot o_+, W \cdot m_b)$, and a link relationship between a node B and the node $+$ is $s_o = \text{LinkPred}(W \cdot m_B, W \cdot o_+)$, where $o_+ = \text{fusion}(W \cdot m_B, W \cdot o_+)$.

[0072] After a graph node link relationship among the graph node N , the graph node $+$, and the graph node B is obtained, the graph node link relationship can be output to the model, and a parameter of the model is adjusted with reference to a pre-generated label until the rule prediction model that satisfies a use requirement is obtained for use.

[0073] In conclusion, by adequately training the rule prediction model, it can be ensured that the graph link relationship between the graph nodes is adequately predicted in an application phase, so that a computation relationship can be subsequently mined based on this.

[0074] Step S208: Generate a numerical column computation rule corresponding to the target service scenario based on the graph node link relationship.

[0075] Specifically, after the graph node link relationship is obtained, all computation relationships in the numerical computable column in the target service scenario are mined. To obtain a computation rule through integration based on the computation relationships and configure the computation rule in the target service scenario, the graph node link relationship can be integrated to obtain the numerical column computation rule in the target service scenario. The numerical column computation rule is used to compute the numerical column. The numerical column computation rule is a rule used to compute the numerical column. For example, $ax+by=z$. This computation relationship is the numerical column computation rule.

[0076] Further, when the numerical column computation rule is constructed based on the graph node link relationship, to ensure comprehensiveness of rule mining, the graph node link relationship can be written into a set until the set satisfies a specified condition. Generation is performed based on all relationships included in the set. In this embodiment, a specific implementation is as follows:

[0077] The graph node link relationship is written into a target relationship set; and when the target relationship set satisfies a rule generation condition, the numerical column computation rule corresponding to the target service scenario is generated based on the target relationship set.

- [0078] Specifically, the target relationship set is specifically a set for storing a mined graph node link relationship. When the rule generation condition is satisfied, all graph node link relationships included in the set can be sorted, to obtain the numerical column computation rule that can be used to compute the numerical computable column.
- 5 [0079] Based on this, to fully mine all the computation rules, the graph node link relationship can be written into the target relationship set. When the target relationship set satisfies the rule generation condition, the numerical column computation rule corresponding to the target service scenario can be generated based on the target relationship set.
- 10 [0080] For example, after the graph node link relationship is obtained, the graph node link relationship can be written into the target relationship set, and the foregoing process can be repeated until it is determined, in a phase, that the graph node link relationship included in the target relationship set satisfies the rule generation condition. In this case, all the graph node link relationships included in the set can be fused, so that all relationships that can be spliced together are fused to obtain all the computation rules.
- 15 [0081] In conclusion, the graph node link relationship output by the model is integrated by using the target relationship set, so that after enough relationships are mined, the computation rule can be generated, to ensure comprehensiveness of rule mining.
- [0082] Based on this, when rule generation is performed, a linear computation relationship between relationships is actually established by using an optimizer, to complete mining of the computation rule. In this embodiment, a specific implementation is as follows:
- 20 [0083] Linear combination processing is performed on the graph node link relationship included in the target relationship set through a generation optimizer, to obtain a graph node link; and the numerical column computation rule corresponding to the target service scenario is generated based on the graph node link.
- 25 [0084] Based on this, when the computation rule is constructed, linear combination processing can be performed on the graph node link relationship included in the target relationship set through the generation optimizer, to obtain the graph node link. Then, the numerical column computation rule corresponding to the target service scenario is generated based on the graph node link.
- 30 [0085] With reference to a schematic diagram shown in (c) in FIG. 4, after the graph node link relationship predicted by the rule prediction model is obtained, the graph node link relationship can be placed in a set, and then a linear combination relationship is obtained by using a generation optimizer, to generate a final rule. An optimized objective function can be $f_a = \alpha_1 m_1 + \alpha_2 m_2 \dots \alpha_n m_n$. Through rule sorting, a rule 1 $T = w_1 X + w_1 P$ and a rule 2 $T = -w_1 N + w_2 M$ can be obtained, to facilitate subsequent use in the target service scenario.

[0086] In addition, after rule mining is completed, to avoid a case in which service running logic in the target service scenario is affected because a rule that includes noise is mined, the rule can be verified. In this embodiment, a specific implementation is as follows:

[0087] The numerical column computation rule is deployed and run on a data processing node corresponding to the target service scenario; and when it is determined, based on a running result, that the numerical column computation rule satisfies a service condition, the numerical column computation rule is input to a rule base corresponding to the target service scenario.

[0088] Based on this, the numerical column computation rule can be deployed and run on the data processing node corresponding to the target service scenario; and when it is determined, based on the running result, that the numerical column computation rule satisfies the service condition, the numerical column computation rule is input to the rule base corresponding to the target service scenario.

[0089] In the rule mining method provided in this embodiment, to improve rule mining efficiency and ensure comprehensiveness of rule mining, after the full service data is obtained, the numerical computable column associated with the target service scenario can be selected from the service data, to mine the computation rule corresponding to the target service scenario. Based on this, considering that the computation rule needs to be mined, the plurality of graph nodes can be first constructed based on the computation operator and the numerical computable column, to implement mining processing based on a graph structure. That is, the plurality of graph nodes are input to the rule prediction model, so that the plurality of graph nodes are processed through the fusion unit and the link prediction unit in the rule prediction model, to obtain the graph node link relationship based on a processing result, so as to mine all possible node link relationships by using the rule prediction model. Based on this, by splicing the graph node link relationships, the numerical column computation rule associated with the target service scenario can be obtained based on a splicing result. In this way, all link relationships can be accurately and efficiently predicted by using the rule prediction model without manual participation, and the computation rule can be constructed. This greatly reduces searchable space of the computation rule of the numerical computable column, and improves a generalization capability of rule mining.

[0090] With reference to FIG. 5, the following further describes the rule mining method by using an application of the rule mining method provided in this specification to a fund security scenario as an example. FIG. 5 is a flowchart of a processing process of a rule mining method according to an embodiment of this specification. The method specifically includes the following steps.

[0091] Step S502: Extract service data from a target database.

- [0092] Step S504: Traverse a numerical column included in the service data based on a preset selection policy.
- [0093] Step S506: Select a numerical column of an enumeration column type as a conditional numerical column corresponding to a target service scenario based on a traversal result.
- 5 [0094] Step S508: Select a numerical computable column from the conditional numerical column based on computation distribution information of the conditional numerical column.
- [0095] Step S510: Construct an operator feature corresponding to a computation operator and a numerical column feature corresponding to the numerical computable column.
- [0096] Step S512: Separately perform graphical processing on the operator feature and the
10 numerical column feature, to obtain an operator feature node and a numerical column feature node.
- [0097] Step S514: Generate a plurality of graph nodes based on the operator feature node and the numerical column feature node.
- [0098] Step S516: Input the plurality of graph nodes to a rule prediction model.
- [0099] Step S518: Compute a sub-link relationship between an operator node and a numerical
15 column node in the plurality of graph nodes through a link prediction unit in the rule prediction model.
- [0100] Step S520: Perform fusion processing on the sub-link relationship through a fusion unit in the rule prediction model.
- [0101] Step S522: Generate a graph node link relationship based on a fusion processing result.
- [0102] Step S524: Write the graph node link relationship into a target relationship set.
- 20 [0103] Step S526: When the target relationship set satisfies a rule generation condition, perform linear combination processing on a graph node link relationship included in the target relationship set through a generation optimizer, to obtain a graph node link.
- [0104] Step S528: Generate a numerical column computation rule corresponding to the target service scenario based on the graph node link.
- 25 [0105] In the rule mining method provided in this embodiment, to improve rule mining efficiency and ensure comprehensiveness of rule mining, after the full service data is obtained, the numerical computable column associated with the target service scenario can be selected from the service data, to mine the computation rule corresponding to the target service scenario. Based on this, considering that the computation rule needs to be mined, the plurality of graph nodes can be first constructed
30 based on the computation operator and the numerical computable column, to implement mining processing based on a graph structure. That is, the plurality of graph nodes are input to the rule prediction model, so that the plurality of graph nodes are processed through the fusion unit and the link prediction unit in the rule prediction model, to obtain the graph node link relationship based on a processing result, so as to mine all possible node link relationships by using the rule prediction

model. Based on this, by splicing the graph node link relationships, the numerical column computation rule associated with the target service scenario can be obtained based on a splicing result. In this way, all link relationships can be accurately and efficiently predicted by using the rule prediction model without manual participation, and the computation rule can be constructed. This greatly reduces searchable space of the computation rule of the numerical computable column, and improves a generalization capability of rule mining.

[0106] Corresponding to the foregoing method embodiments, this specification further provides an embodiment of a rule mining apparatus. FIG. 6 is a schematic diagram of a structure of a rule mining apparatus according to an embodiment of this specification. As shown in FIG. 6, the apparatus includes:

an obtaining module 602, configured to obtain service data, and select a numerical computable column associated with a target service scenario from the service data;

a construction module 604, configured to construct a plurality of graph nodes based on a computation operator and the numerical computable column;

a prediction module 606, configured to input the plurality of graph nodes to a rule prediction model, and process the plurality of graph nodes through a fusion unit and a link prediction unit in the rule prediction model, to obtain a graph node link relationship; and

a generation module 608, configured to generate a numerical column computation rule corresponding to the target service scenario based on the graph node link relationship.

[0107] In an optional embodiment, the obtaining module 602 is further configured to:

extract the service data from a target database; select a numerical column of an enumeration column type from a numerical column included in the service data as a conditional numerical column corresponding to the target service scenario; and select the numerical computable column from the conditional numerical column based on computation distribution information of the conditional numerical column.

[0108] In an optional embodiment, the obtaining module 602 is further configured to:

traverse the numerical column included in the service data based on a preset selection policy; and select the numerical column of the enumeration column type as the conditional numerical column corresponding to the target service scenario based on a traversal result.

[0109] In an optional embodiment, the construction module 604 is further configured to:

construct an operator feature corresponding to the computation operator and a numerical column feature corresponding to the numerical computable column; separately perform graphical processing on the operator feature and the numerical column feature, to obtain an operator feature

node and a numerical column feature node; and generate the plurality of graph nodes based on the operator feature node and the numerical column feature node.

[0110] In an optional embodiment, the prediction module 606 is further configured to:

input the plurality of graph nodes to the rule prediction model; compute a sub-link relationship between an operator node and a numerical column node in the plurality of graph nodes through the link prediction unit in the rule prediction model; perform fusion processing on the sub-link relationship through the fusion unit in the rule prediction model; and generate the graph node link relationship based on a fusion processing result.

[0111] In an optional embodiment, the generation module 608 is further configured to:

write the graph node link relationship into a target relationship set; and when the target relationship set satisfies a rule generation condition, generate the numerical column computation rule corresponding to the target service scenario based on the target relationship set.

[0112] In an optional embodiment, the generation module 608 is further configured to:

perform linear combination processing on a graph node link relationship included in the target relationship set through a generation optimizer, to obtain a graph node link; and generate the numerical column computation rule corresponding to the target service scenario based on the graph node link.

[0113] In an optional embodiment, training of the rule prediction model includes:

obtaining a sample set, where the sample set includes a labeled sample pair and an unlabeled sample; processing the labeled sample pair by using an initial rule prediction model, and determining a relationship parameter based on a processing result; updating the initial rule prediction model based on the relationship parameter, to obtain an intermediate rule prediction model; and processing the unlabeled sample by using the intermediate rule prediction model, and adjusting a parameter of the intermediate rule prediction model based on a processing result until the rule prediction model that satisfies a training stop condition is obtained.

[0114] In an optional embodiment, the processing the unlabeled sample by using the intermediate rule prediction model, and adjusting a parameter of the intermediate rule prediction model based on a processing result until the rule prediction model that satisfies a training stop condition is obtained includes:

inputting the unlabeled sample to the intermediate rule prediction model; processing the unlabeled sample through an intermediate fusion unit and an intermediate link prediction unit in the intermediate rule prediction model, to obtain a predicted link relationship; generating a predicted numerical column computation rule based on the predicted link relationship; and when determining, based on the predicted numerical column computation rule, that the intermediate rule prediction

model satisfies the training stop condition, using the intermediate rule prediction model as the rule prediction model.

[0115] In an optional embodiment, the apparatus further includes:

an input module, configured to deploy and run the numerical column computation rule on a data processing node corresponding to the target service scenario; and when determining, based on a running result, that the numerical column computation rule satisfies a service condition, input the numerical column computation rule to a rule base corresponding to the target service scenario.

[0116] In the rule mining apparatus provided in this embodiment, to improve rule mining efficiency and ensure comprehensiveness of rule mining, after the full service data is obtained, the numerical computable column associated with the target service scenario can be selected from the service data, to mine the computation rule corresponding to the target service scenario. Based on this, considering that the computation rule needs to be mined, the plurality of graph nodes can be first constructed based on the computation operator and the numerical computable column, to implement mining processing based on a graph structure. That is, the plurality of graph nodes are input to the rule prediction model, so that the plurality of graph nodes are processed through the fusion unit and the link prediction unit in the rule prediction model, to obtain the graph node link relationship based on a processing result, so as to mine all possible node link relationships by using the rule prediction model. Based on this, by splicing the graph node link relationships, the numerical column computation rule associated with the target service scenario can be obtained based on a splicing result. In this way, all link relationships can be accurately and efficiently predicted by using the rule prediction model without manual participation, and the computation rule can be constructed. This greatly reduces searchable space of the computation rule of the numerical computable column, and improves a generalization capability of rule mining.

[0117] The schematic solution for the rule mining apparatus in the embodiments is described above. It should be noted that the technical solutions of the rule mining apparatus are of the same concept as the technical solutions of the foregoing rule mining method. For details that are not described in detail in the technical solutions of the rule mining apparatus, refer to the descriptions of the technical solutions of the foregoing rule mining method.

[0118] FIG. 7 is a block diagram of a structure of a computing device 700 according to an embodiment of this specification. Components of the computing device 700 include but are not limited to a memory 710 and a processor 720. The processor 720 and the memory 710 are connected by using a bus 730, and a database 750 is configured to store data.

[0119] The computing device 700 further includes an access device 740. The access device 740 enables the computing device 700 to communicate through one or more networks 760. Examples of

these networks include a public switched telephone network (PSTN), a local area network (LAN), a wide area network (WAN), a personal area network (PAN), or a combination of communication networks such as the Internet. The access device 740 can include one or more of any type of wired or wireless network interfaces (for example, a network interface controller (NIC)), for example, an IEEE 802.11 wireless local area network (WLAN) wireless interface, a worldwide interoperability for microwave access (Wi-MAX) interface, an Ethernet interface, a universal serial bus (USB) interface, a cellular network interface, a Bluetooth interface, and a near field communication (NFC) interface.

[0120] In an embodiment of this application, the foregoing components of the computing device 700 and other components not shown in FIG. 7 can alternatively be connected to each other, for example, by using a bus. It should be understood that the block diagram of the structure of the computing device shown in FIG. 7 is for an illustrative purpose only and is not intended to limit the scope of this application. A person skilled in the art can add or replace other components as needed.

[0121] The computing device 700 can be any type of stationary or mobile computing device, including a mobile computer or a mobile computing device (for example, a tablet computer, a personal digital assistant, a laptop computer, a notebook computer, or a netbook), a mobile phone (for example, a smartphone), a wearable computing device (for example, a smartwatch or smart glasses), or another type of mobile device, or a stationary computing device such as a desktop computer or a personal computer (PC). The computing device 700 can alternatively be a mobile or stationary server.

[0122] The processor 720 is configured to execute the following computer-executable instructions. When the computer-executable instructions are executed by the processor, the steps of the foregoing rule mining method are implemented.

[0123] The schematic solution for the computing device in the embodiments is described above. It should be noted that the technical solutions of the computing device are of the same concept as the technical solutions of the foregoing rule mining method. For details that are not described in detail in the technical solutions of the computing device, refer to the descriptions of the technical solutions of the foregoing rule mining method.

[0124] An embodiment of this specification further provides a computer-readable storage medium. The computer-readable storage medium stores computer-executable instructions. When the computer-executable instructions are executed by a processor, the steps of the foregoing rule mining method are implemented.

[0125] The schematic solution for the computer-readable storage medium in the embodiments is described above. It should be noted that the technical solutions of the storage medium are of the same concept as the technical solutions of the foregoing rule mining method. For details that are not

described in detail in the technical solutions of the storage medium, refer to the descriptions of the technical solutions of the foregoing rule mining method.

[0126] Specific embodiments of this specification are described above. Other embodiments fall within the scope of the appended claims. In some cases, the actions or steps described in the claims can be performed in a sequence different from that in the embodiments and desired results can still be achieved. In addition, the process depicted in the accompanying drawings does not necessarily need a particular sequence or a consecutive sequence to achieve the desired results. In some implementations, multi-tasking and parallel processing are feasible or may be advantageous.

[0127] The computer instructions include computer program code, and the computer program code can be in a source code form, an object code form, an executable file form, some intermediate forms, or the like. The computer-readable medium can include any entity or apparatus that can carry the computer program code, a recording medium, a USB flash drive, a removable hard disk, a magnetic disk, an optical disc, a computer memory, a read-only memory (ROM), a random access memory (RAM), an electrical carrier signal, a telecommunications signal, a software distribution medium, or the like. It should be noted that the content included in the computer-readable medium can be appropriately increased or decreased in accordance with the requirements of the legislation and patent practice in the jurisdiction, for example, in some jurisdictions, the computer-readable medium does not include electrical carrier signals or telecommunications signals in accordance with the legislation and patent practice.

[0128] It should be noted that for brief description, the foregoing method embodiments are represented as a series of actions. However, a person skilled in the art should understand that the embodiments of this specification are not limited to the described sequence of the actions, because according to the embodiments of this specification, some steps can be performed in another sequence or simultaneously. In addition, a person skilled in the art should also understand that all the embodiments described in this specification are example embodiments, and the related actions and modules are not necessarily mandatory to the embodiments of this specification.

[0129] In the foregoing embodiments, the descriptions of the embodiments have respective focuses. For a part that is not described in detail in an embodiment, refer to related descriptions in other embodiments.

[0130] The example embodiments of this specification disclosed above are intended merely to help describe this specification. The optional embodiments do not describe all the details in detail, and do not limit the present invention to the specific implementations described. Clearly, many modifications and changes can be made based on the content of the embodiments of this specification. These embodiments are selected and described in detail in this specification to better explain the

principles and practical applications of the embodiments of this specification, so that a person skilled in the art can well understand and use this specification. This specification is limited only to the claims and the full scope and equivalents thereof.

CLAIMS

What is claimed is:

1. A rule mining method, comprising:

obtaining service data, and selecting a numerical computable column associated with a target
5 service scenario from the service data;

constructing a plurality of graph nodes based on a computation operator and the numerical
computable column;

inputting the plurality of graph nodes to a rule prediction model, and processing the plurality of
graph nodes through a fusion unit and a link prediction unit in the rule prediction model, to obtain a
10 graph node link relationship; and

generating a numerical column computation rule corresponding to the target service scenario
based on the graph node link relationship.

2. The method according to claim 1, wherein the obtaining service data, and selecting a
numerical computable column associated with a target service scenario from the service data
15 comprises:

extracting the service data from a target database;

selecting a numerical column of an enumeration column type from a numerical column
comprised in the service data as a conditional numerical column corresponding to the target service
scenario; and

20 selecting the numerical computable column from the conditional numerical column based on
computation distribution information of the conditional numerical column.

3. The method according to claim 2, wherein the selecting a numerical column of an enumeration
column type from a numerical column comprised in the service data as a conditional numerical
column corresponding to the target service scenario comprises:

25 traversing the numerical column comprised in the service data based on a preset selection policy;
and

selecting the numerical column of the enumeration column type as the conditional numerical
column corresponding to the target service scenario based on a traversal result.

4. The method according to claim 1, wherein the constructing a plurality of graph nodes based
30 on a computation operator and the numerical computable column comprises:

constructing an operator feature corresponding to the computation operator and a numerical
column feature corresponding to the numerical computable column;

separately performing graphical processing on the operator feature and the numerical column
feature, to obtain an operator feature node and a numerical column feature node; and

generating the plurality of graph nodes based on the operator feature node and the numerical column feature node.

5. The method according to claim 1, wherein the inputting the plurality of graph nodes to a rule prediction model, and processing the plurality of graph nodes through a fusion unit and a link prediction unit in the rule prediction model, to obtain a graph node link relationship comprises:

inputting the plurality of graph nodes to the rule prediction model;

computing a sub-link relationship between an operator node and a numerical column node in the plurality of graph nodes through the link prediction unit in the rule prediction model;

performing fusion processing on the sub-link relationship through the fusion unit in the rule prediction model; and

generating the graph node link relationship based on a fusion processing result.

6. The method according to claim 1, wherein the generating a numerical column computation rule corresponding to the target service scenario based on the graph node link relationship comprises:

writing the graph node link relationship into a target relationship set; and

when the target relationship set satisfies a rule generation condition, generating the numerical column computation rule corresponding to the target service scenario based on the target relationship set.

7. The method according to claim 6, wherein the generating the numerical column computation rule corresponding to the target service scenario based on the target relationship set comprises:

performing linear combination processing on a graph node link relationship comprised in the target relationship set through a generation optimizer, to obtain a graph node link; and

generating the numerical column computation rule corresponding to the target service scenario based on the graph node link.

8. The method according to any one of claims 1 to 7, wherein training of the rule prediction model comprises:

obtaining a sample set, wherein the sample set comprises a labeled sample pair and an unlabeled sample;

processing the labeled sample pair by using an initial rule prediction model, and determining a relationship parameter based on a processing result;

updating the initial rule prediction model based on the relationship parameter, to obtain an intermediate rule prediction model; and

processing the unlabeled sample by using the intermediate rule prediction model, and adjusting a parameter of the intermediate rule prediction model based on a processing result until the rule prediction model that satisfies a training stop condition is obtained.

9. The method according to claim 8, wherein the processing the unlabeled sample by using the intermediate rule prediction model, and adjusting a parameter of the intermediate rule prediction model based on a processing result until the rule prediction model that satisfies a training stop condition is obtained comprises:

- 5 inputting the unlabeled sample to the intermediate rule prediction model;
 processing the unlabeled sample through an intermediate fusion unit and an intermediate link prediction unit in the intermediate rule prediction model, to obtain a predicted link relationship;
 generating a predicted numerical column computation rule based on the predicted link relationship; and
- 10 when determining, based on the predicted numerical column computation rule, that the intermediate rule prediction model satisfies the training stop condition, using the intermediate rule prediction model as the rule prediction model.

10. The method according to any one of claims 1 to 7, wherein after the step of generating a numerical column computation rule corresponding to the target service scenario based on the graph node link relationship is performed, the method further comprises:

- 15 deploying and running the numerical column computation rule on a data processing node corresponding to the target service scenario; and
 when determining, based on a running result, that the numerical column computation rule satisfies a service condition, inputting the numerical column computation rule to a rule base
- 20 corresponding to the target service scenario.

11. A rule mining apparatus, comprising:

- an obtaining module, configured to obtain service data, and select a numerical computable column associated with a target service scenario from the service data;
- a construction module, configured to construct a plurality of graph nodes based on a computation
- 25 operator and the numerical computable column;
- a prediction module, configured to input the plurality of graph nodes to a rule prediction model, and process the plurality of graph nodes through a fusion unit and a link prediction unit in the rule prediction model, to obtain a graph node link relationship; and
- a generation module, configured to generate a numerical column computation rule
- 30 corresponding to the target service scenario based on the graph node link relationship.

12. A computing device, comprising:

- a memory and a processor, wherein
- the memory is configured to store computer-executable instructions, the processor is configured to execute the computer-executable instructions, and when the computer-executable instructions are

executed by the processor, the steps of the method according to any one of claims 1 to 10 are implemented.

13. A computer-readable storage medium, wherein the computer-readable storage medium stores computer-executable instructions, and when the computer-executable instructions are executed
5 by a processor, the steps of the method according to any one of claims 1 to 10 are implemented.