

What is claimed is:

1. A method of constructing a decision model using one or more processors, the method comprising:

obtaining a rule template data and extracting each variable object and template sample from the rule template data;

clustering and analyzing the variable objects to obtain a clustering result;

matching the clustering result with each template sample according to the rule template data, and serving the matched clustering result as a first feature;

calculating a black sample probability for each variable object and serving the black sample probability of each variable object as a second feature; and

constructing the decision model according to the first feature and the second feature.

2. The method of claim 1, wherein after calculating the black sample probability for each variable object and serving the black sample probability of each variable object as the second feature, the method further comprises:

mapping each variable object to a predefined label according to a preset algorithm; and

matching the label with each template sample according to the rule template data, and serving the matched label as a third feature;

wherein the constructing the decision model according to the first feature and the second feature comprises:

constructing the decision model according to the first feature, the second feature, and the third feature.

3. The method of claim 2, wherein the constructing the decision model according to the first feature, the second feature, and the third feature comprises:

establishing an original node;

obtaining a result type of each template sample according to the rule template data;

traversing and reading the first feature, the second feature, and the third feature respectively to generate a reading record;

calculating a division purity of each reading record according to the result type of

each template sample, and determining a division point according to the division purity; and
obtaining a feature corresponding to the division point, and establishing a new node.

4. The method of claim 1, wherein the clustering and analyzing the variable objects to obtain the clustering result comprises:
selecting a plurality of variable objects randomly from the variable objects as a first cluster center of one cluster, each first cluster center corresponding to one of the cluster;
calculating a distance from each variable object to each first cluster center, respectively;
dividing each variable object according to the calculation result, and dividing each variable object into a cluster corresponding to a first cluster center wherein the distance from the each variable object to the first cluster center is shortest;
calculating a second cluster center of each cluster respectively after dividing the variable objects; and
determining whether the distance between the first cluster center and the second cluster center in each cluster is less than a preset threshold value; and if yes, outputting each cluster as the clustering result; or else, replacing the first cluster center of the corresponding cluster with the second cluster center, and continuing to calculate the distance from each variable object to each first cluster center, respectively.

5. The method of claim 4, wherein the calculating the distance from each variable object to each first cluster center respectively comprises:
obtaining a multidimensional data of each variable object according to the rule template data; and
calculating a distance from each variable object to each first cluster center respectively according to the multidimensional data of each variable object.

6. A device of constructing a decision model, comprising:

an extraction module configured to obtain a rule template data and extract each variable object and each template sample from the rule template data;
a cluster module configured to cluster and analyze the variable objects to obtain a clustering result;
a first feature module configured to match the clustering result with each template sample according to the rule template data, and serve the matched clustering result as a first feature;
a second feature module configured to calculate a black sample probability for each variable object and serve the black sample probability of each variable object as a second feature; and
a construction module configured to construct the decision model according to the first feature and the second feature.

7. The device of claim 6, wherein the device further comprises:

a mapping module configured to map each variable object to a predefined label according to a preset algorithm; and
a third feature module configured to match the label with each template sample according to the rule template data, and serve the matched label as a third feature;
the construction module is further configured to construct the decision model according to the first feature, the second feature, and the third feature.

8. The device of claim 7, wherein the construction module comprises:

an establishment unit configured to establish an original node;
an obtainment unit configured to obtain a result type of each template sample according to the rule template data;
a traversal unit configured to traverse and read the first feature, the second feature, and the third feature respectively to generate a reading record; and
a purity calculation unit configured to calculate a division purity of each reading record according to the result type of each template sample, and determine a division point according to the division purity;
wherein the establishment unit is further configured to obtain a feature corresponding to the division point, and establish a new node.

9. The device of claim 6, wherein the cluster module comprises:

a selection unit configured to select a plurality of variable objects randomly from the variable objects as a first cluster center of one cluster, wherein each first cluster center corresponding to one of the cluster;

a distance calculation unit configured to calculate a distance from each variable object to each first cluster center respectively;

a division unit configured to divide each variable object into a cluster corresponding to a first cluster center wherein the distance from the each variable object to the first cluster center is shortest;

a center calculation unit configured to calculate a second cluster center of each cluster respectively after dividing the variable objects; and

a determination unit configured to determine whether the distance between the first cluster center and the second cluster center in each cluster is less than a preset threshold value; and if yes, each cluster is output as the clustering result; or else, the first cluster center of the corresponding cluster is replaced with the second cluster center, and the distance calculation unit continues to calculate the distance from each variable object to each first cluster center respectively.

10. The device of claim 9, wherein the distance calculation unit comprises:

an obtainment subunit configured to obtain a multidimensional data of each variable object according to the rule template data; and

a calculation subunit configured to calculate a distance from each variable object to each first cluster center according to the multidimensional data of each variable object respectively.

11. A computer apparatus, comprising a processor and a memory storing computer executable instructions stored that, when executed by the processor, cause the processor to perform operations comprising:

obtaining a rule template data and extracting each variable object and each template sample from the rule template data;

clustering and analyze the variable objects to obtain a clustering result;

matching the clustering result with each template sample according to the rule template data, and serving the matched clustering result as a first feature;
calculating a black sample probability for each variable object and serving the black sample probability of each variable object as a second feature; and
constructing the decision model according to the first feature and the second feature.

12. The computer apparatus of claim 11, wherein after the step of calculating the black sample probability for each variable object and serving the black sample probability of each variable object as the second feature, the computer executable instructions, when executed by the processor, further cause the processor to perform operations comprising:

mapping each variable object to a predefined label according to a preset algorithm;
and

matching the label with each template sample according to the rule template data, and serving the matched label as a third feature;

the constructing the decision model according to the first feature and the second feature comprises:

constructing the decision model according to the first feature, the second feature, and the third feature.

13. The computer apparatus of claim 12, wherein the constructing the decision model according to the first feature, the second feature, and the third feature comprises:

establishing an original node;

obtaining a result type of each template sample according to the rule template data;
traversing and reading the first feature, the second feature, and the third feature respectively to generate a reading record;

calculating a division purity of each reading record according to the result type of each template sample, and determining a division point according to the division purity; and

obtaining a feature corresponding to the division point, and establishing a new

node.

14. The computer apparatus of claim 11, wherein the clustering and analyzing the variable objects to obtain the clustering result comprises:

selecting a plurality of variable objects randomly from the variable objects as a first cluster center of one cluster, wherein each first cluster center corresponding to one of the cluster;

calculating a distance from each variable object to each first cluster center respectively;

dividing each variable object according to the calculation result, and dividing each variable object into a cluster corresponding to a first cluster center wherein the distance from the each variable object to the first cluster center is shortest;

calculating a second cluster center of each cluster respectively after dividing the variable objects; and

determining whether the distance between the first cluster center and the second cluster center in each cluster is less than a preset threshold value; and if yes, outputting each cluster as the clustering result; or else, replacing the first cluster center of the corresponding cluster with the second cluster center, and continuing to calculate the distance from each variable object to each first cluster center respectively.

15. The computer apparatus of claim 14, wherein the calculating the distance from each variable object to each first cluster center respectively comprises:

obtaining a multidimensional data of each variable object according to the rule template data; and

calculating a distance from each variable object to each first cluster center according to the multidimensional data of each variable object respectively.

16. One or more storage apparatus storing computer executable instructions that, when executed by the one or more processors, cause the one or more processors to perform the steps of:

obtaining a rule template data and extracting each variable object and each

template sample from the rule template data;
clustering and analyzing the variable objects to obtain a clustering result;
matching the clustering result with each template sample according to the rule template data, and serving the matched clustering result as a first feature;
calculating a black sample probability for each variable object and serving the black sample probability of each variable object as a second feature; and
constructing the decision model according to the first feature and the second feature.

17. The storage apparatus of claim 16, wherein the computer executable instructions, when executed by the one or more processors, further cause the one or more processors, after the step of calculating the black sample probability for each variable object and serving the black sample probability of each variable object as a second feature, to further perform operations the steps of:
mapping each variable object to a predefined label according to a preset algorithm;
and
matching the label with each template sample according to the rule template data, and serving the matched label as a third feature;
the constructing the decision model according to the first feature and the second feature comprises:
constructing the decision model according to the first feature, the second feature, and the third feature.

18. The storage apparatus of claim 17, wherein the constructing the decision model according to the first feature, the second feature, and the third feature comprises:
establishing an original node;
obtaining a result type of each template sample according to the rule template data;
traversing and reading the first feature, the second feature, and the third feature respectively to generate a reading record;
calculating a division purity of each reading record according to the result type of each template sample, and determining a division point according to the division purity; and

obtaining a feature corresponding to the division point, and establishing a new node.

19. The storage apparatus of claim 16, wherein the clustering and analyzing the variable objects to obtain the clustering result comprises:

selecting a plurality of variable objects randomly from the variable objects as a first cluster center of one cluster, wherein each first cluster center corresponding to one of the cluster;

calculating a distance from each variable object to each first cluster center respectively;

dividing each variable object according to the calculation result, and dividing each variable object into a cluster corresponding to a first cluster center wherein the distance from the each variable object to the first cluster center is shortest;

calculating a second cluster center of each cluster respectively after dividing the variable objects; and

determining whether the distance between the first cluster center and the second cluster center in each cluster is less than a preset threshold value; and if yes, outputting each cluster as the clustering result; or else, replacing the first cluster center of the corresponding cluster with the second cluster center, and continuing to calculate the distance from each variable object to each first cluster center respectively.

20. The storage apparatus of claim 19, wherein the calculating the distance from each variable object to each first cluster center respectively comprises:

obtaining a multidimensional data of each variable object according to the rule template data; and

calculating a distance from each variable object to each first cluster center according to the multidimensional data of each variable object respectively.