CLAIMS

What is claimed is:

1. A model integration method, the method comprising:

obtaining an integrated model, the integrated model having one integrated output value and a plurality of input values, the plurality of input values corresponding to a plurality of output values of a plurality of independent models;

performing one or more iterations of an optimizing process until a preset iteration stop condition is satisfied, wherein the optimizing process comprises:

acquiring a prediction output by the integrated model based on a preset test event set;

determining an index value of the integrated model based on the prediction output, the index value indicating a performance evaluation of the integrated model;

if the index value fails to meet a preset performance requirement, optimizing the integrated model based on a preset optimization event set different from the preset test event set by performing at least the following adjustments:

refitting the integrated model by adjusting a plurality of integration weights in the integrated model respectively corresponding to the plurality of independent models;

separately refitting one or more of the plurality of independent models by performing operations, comprising:

adjusting one or more inputs associated with each of the one or more independent models;

sorting the plurality of independent models based on the plurality of corresponding integration weights;

selecting for refitting a first independent model of the sorted plurality of independent models that has not been refitted and has a highest integration weight;

after selecting the first independent model for refitting, refitting the first independent model; and

after refitting the first independent model, retraining the first independent model; and

after the preset iteration stop condition is satisfied, determining the integrated model as acceptable.

2. The method according to claim 1, wherein the preset iteration stop condition comprises a requirement that the index value satisfies the preset performance requirement, or

a number of the one or more iterations reaches a preset threshold.

3. The method according to claim 1, wherein optimizing the integrated model comprises:

refitting the integrated model in a first iteration by adjusting the plurality of integration weights respectively corresponding to the plurality of independent models, and

if the preset performance requirement is not satisfied:

separately refitting one or more of the plurality of independent models in one or more subsequent iterations after the first iteration by adjusting the one or more inputs associated with each of the one or more independent models; and

retraining one or more of the plurality of independent models in one or more subsequent iterations after the first iteration to acquire one or more new independent models after removing input features from or adding input features to each of the one or more independent models.

4. The method according to claim 1, wherein optimizing the integrated model further comprises::

retraining the integrated model to acquire a new integrated model after removing one or more independent models from the plurality of independent models to provide inputs to the integrated model, or after adding one or more new independent models to the plurality of independent models to provide inputs to the integrated model.

5. The method according to claim 1, wherein retraining the first independent model comprises:

removing input features from the first independent model or adding input features to the first independent model.

6. The method according to claim 1, wherein obtaining the integrated model comprises:

obtaining a set of historical event data records, each comprising a plurality of predicted values generated by the plurality of independent models and an integrated label value, wherein each of the plurality of predicted values is within a predicted value range associated with the corresponding independent model;

dividing the predicted value range associated with the corresponding independent model into a plurality of subintervals, wherein each of the plurality of predicted values falls into one of the plurality of subintervals;

converting, for each of the set of historical event data records, the plurality of predicted values respectively to a plurality of encoded values;

obtaining a training data set based on the set of historical event data records, wherein

each training data of the training data set is associated with a plurality of eigenvalues determined based on the plurality of corresponding encoded values; and

training the integrated model by a supervised learning algorithm based on the training data set.

7. The method according to claim 6, wherein converting, for each of the set of historical event data records, the plurality of predicted values respectively to a plurality of encoded values comprises:

converting the plurality of predicted values respectively to a plurality of One-Hot encoded values, wherein each of the plurality of One-Hot encoded values comprises a plurality of bits, and a quantity of the plurality of bits equals to a quantity of the plurality of subintervals.

8. The method according to claim 6, wherein converting, for each of the set of historical event data records, the plurality of predicted values respectively to a plurality of encoded values comprises:

for each independent model of the plurality of independent models, determining a plurality of weight of evidence (WOE) scores respectively for the plurality of subintervals; and

converting each predicted value from the independent model to a WOE score corresponding to a subinterval of the plurality of subintervals in which the each predicted value belongs.

9. The method of claim 1, wherein the integrated model is a risk control model, and each of the plurality of independent models is a risk control model corresponding to a risk type.

10. A model integration system, comprising one or more processors and one or more non-transitory computer-readable memories coupled to the one or more processors and configured with instructions executable by the one or more processors to cause the system to perform operations comprising:

obtaining an integrated model, the integrated model having one integrated output value and a plurality of input values, the plurality of input values corresponding to a plurality of output values of a plurality of independent models;

performing one or more iterations of an optimizing process until a preset iteration stop condition is satisfied, wherein the optimizing process comprises:

acquiring a prediction output by the integrated model based on a preset test event set;

determining an index value of the integrated model based on the prediction output, the

index value indicating a performance evaluation of the integrated model;

if the index value fails to meet a preset performance requirement, optimizing the integrated model based on a preset optimization event set different from the preset test event set by performing at least the following adjustments:

refitting the integrated model by adjusting a plurality of integration weights in the integrated model respectively corresponding to the plurality of independent models;

separately refitting one or more of the plurality of independent models, wherein the separately refitting comprises:

adjusting one or more inputs associated with each of the one or more independent models;

sorting the plurality of independent models based on the plurality of corresponding integration weights;

selecting for refitting a first independent model of the sorted plurality of independent models that has not been refitted and has a highest integration weight; and

after selecting the first independent model for refitting, refitting the first independent model; and

after refitting the first independent model, retraining the first independent model;

after the preset iteration stop condition is satisfied, determining the integrated model as acceptable.

11. The system according to claim 10, wherein the preset iteration stop condition comprises a requirement that the index value satisfies the preset performance requirement, or a number of the one or more iterations reaches a preset threshold.

12. The system according to claim 10, wherein optimizing the integrated model comprises:

refitting the integrated model in a first iteration by adjusting the plurality of integration weights respectively corresponding to the plurality of independent models; and

if the preset performance requirement is not satisfied:

separately refitting one or more of the plurality of independent models in one or more subsequent iterations after the first iteration by adjusting the one or more inputs associated with each of the one or more independent models; and

retraining one or more of the plurality of independent models in one or more subsequent iterations after the first iteration to acquire one or more new independent models after removing input features from or adding input features to each of the one or more independent models.

13. The system according to claim 10, wherein obtaining the integrated model comprises:

obtaining a set of historical event data records, each comprising a plurality of predicted values generated by the plurality of independent models and an integrated label value, wherein each of the plurality of predicted values is within a predicted value range associated with the corresponding independent model;

dividing the predicted value range associated with the corresponding independent model into a plurality of subintervals, wherein each of the plurality of predicted values falls into one of the plurality of subintervals;

converting, for each of the set of historical event data records, the plurality of predicted values respectively to a plurality of encoded values;

obtaining a training data set based on the set of historical event data records, wherein each training data of the training data set is associated with a plurality of eigenvalues determined based on the plurality of corresponding encoded values; and

training the integrated model by a supervised learning algorithm based on the training data set.

14. The system of claim 10, wherein the integrated model is a risk control model, and each of the plurality of independent models is a risk control model corresponding to a risk type.

15. A non-transitory computer-readable storage medium for model integration configured with instructions executable by one or more processors to cause the one or more processors to perform operations comprising:

obtaining an integrated model, the integrated model having one integrated output value and a plurality of input values, the plurality of input values corresponding to a plurality of output values of a plurality of independent models;

performing one or more iterations of an optimizing process until a preset iteration stop condition is satisfied, wherein the optimizing process comprises:

acquiring a prediction output by the integrated model based on a preset test event set;

determining an index value of the integrated model based on the prediction output, the index value indicating a performance evaluation of the integrated model;

if the index value fails to meet a preset performance requirement, optimizing the integrated model based on a preset optimization event set different from the preset test event set by performing at least the following adjustments:

refitting the integrated model by adjusting a plurality of integration weights in the

integrated model respectively corresponding to the plurality of independent models;

separately refitting one or more of the plurality of independent models, wherein the separately refitting comprises:

adjusting one or more inputs associated with each of the one or more independent models;

sorting the plurality of independent models based on the plurality of corresponding integration weights;

selecting for refitting a first independent model of the sorted plurality of independent models that has not been refitted and has a highest integration weight; and

after selecting the first independent model for refitting, refitting the first independent model; and

after refitting the first independent model, retraining the first independent model; and

after the preset iteration stop condition is satisfied, determining the integrated model as acceptable.

16. The non-transitory computer-readable storage medium according to claim 15, wherein the preset iteration stop condition comprises a requirement that the index value satisfies the preset performance requirement, or a number of the one or more iterations reaches a preset threshold.

17. The non-transitory computer-readable storage medium according to claim 15, wherein optimizing the integrated model comprises:

refitting the integrated model in a first iteration by adjusting the plurality of integration weights respectively corresponding to the plurality of independent models; and

if the preset performance requirement is not satisfied:

separately refitting one or more of the plurality of independent models in one or more subsequent iterations after the first iteration by adjusting the one or more inputs associated with each of the one or more independent models; and

retraining one or more of the plurality of independent models in one or more subsequent iterations after the first iteration to acquire one or more new independent models after removing input features from or adding input features to each of the one or more independent models.

18. The non-transitory computer-readable storage medium according to claim 15, wherein retraining the first independent model comprises:

removing input features from the first independent model or adding input features to the first independent model.

19. The non-transitory computer-readable storage medium according to claim 15, wherein obtaining the integrated model comprises:

obtaining a set of historical event data records, each comprising a plurality of predicted values generated by the plurality of independent models and an integrated label value, wherein each of the plurality of predicted values is within a predicted value range associated with the corresponding independent model;

dividing the predicted value range associated with the corresponding independent model into a plurality of subintervals, wherein each of the plurality of predicted values falls into one of the plurality of subintervals;

converting, for each of the set of historical event data records, the plurality of predicted values respectively to a plurality of encoded values;

obtaining a training data set based on the set of historical event data records, wherein each training data of the training data set is associated with a plurality of eigenvalues determined based on the plurality of corresponding encoded values; and

training the integrated model by a supervised learning algorithm based on the training data set.

20. The non-transitory computer-readable storage medium of claim 15, wherein the integrated model is a risk control model, and each of the plurality of independent models is a risk control model corresponding to a risk type.