WHAT IS CLAIMED IS:

1. A method for payment risk analysis, comprising:

(i) for each baseline machine learning model of a set of baseline machine learning models, with at least one processor:

training that baseline machine learning model based on a plurality of first training samples associated with a plurality of first payment transactions configured to be processed in an electronic payment processing network, wherein the plurality of first training samples includes a plurality of different data types, and wherein training that baseline machine learning model generates a plurality of first predictions for the plurality of first training samples;

generating, for that baseline machine learning model, based on the plurality of first predictions for the plurality of first training samples, a rejection region associated with at least one data type of the plurality of different data types; and

processing, with that baseline machine learning model, a subset of second training samples of a plurality of second training samples outside the rejection region of that baseline machine learning model, to generate a subset of second predictions for the subset of second training samples of the plurality of second training samples outside the rejection region of that baseline machine learning model, wherein the plurality of second training samples is associated with a plurality of second payment transactions configured to be processed in the electronic payment network, wherein a baseline model predictive performance metric for that baseline machine learning model is determined based on the subset of second predictions of that baseline machine learning model, and wherein the plurality of second training samples is associated with a plurality of rejection flags for that baseline machine learning model, wherein each rejection flag of the plurality of rejection flags indicates whether a corresponding second sample of the plurality of second samples is within the rejection region of that baseline machine learning model;

(ii) generating, with the at least one processor, a global rejection region associated with one or more data types of the plurality of different data types based on the rejection region associated with each baseline machine learning model;

(iii) training, with the at least one processor, an ensemble machine learning model ensembled based on the set of baseline machine learning models, based on (a) a further subset of second training samples of the plurality of second training samples outside the global rejection region, (b) the plurality of rejection flags for the plurality of second samples associated with each baseline machine learning model, and (c) the subset of second predictions for the subset of second training samples generated for each baseline machine learning model, wherein training the ensemble machine learning model generates a subset of ensemble predictions for the further subset of second training samples of the plurality of second training samples outside the global rejection region, and wherein an ensemble model predictive performance metric is determined based on the subset of ensemble predictions;

(iv) updating, with the at least one processor, based on the baseline model predictive performance metric for each baseline machine learning model, the set of baseline machine learning models; [[and]]

(v) repeating, with the at least one processor, (i) – (iv) until there is a single baseline machine learning model in the set of baseline machine learning models or at least one of the ensemble model predictive performance metric satisfies a threshold ensemble model predictive performance, a ratio of the plurality of second training samples outside the global rejection region satisfies a threshold ratio, or any combination thereof;

(vi) obtaining, with the at least one processor, a current sample associated with a current payment transaction currently being processed in the electronic payment network;

(vii) determining, with the at least one processor, whether the current sample is within the global rejection region;

(viii) in response to determining that the current sample is outside the global rejection region, automatically processing, with the at least one processor, using the ensemble machine learning model, the current sample to generate a current prediction for the current sample of a payment risk associated with the current payment transaction; and

(ix) in response to determining that the current sample is within the global rejection region, automatically flagging, with the at least one processor, the current sample as unable to receive, from the ensemble classifier, a credible prediction of the payment risk associated with the current payment transaction.

2. The method of claim 1, wherein, for each baseline machine learning model of the set of baseline machine learning models, generating, for that baseline machine learning model, based on the plurality of first predictions for the plurality of first training samples, the rejection region associated with the at least one data type of the plurality of different data types includes optimizing an objective function defined according to the following equation:

$$argmax_{(t_{lb})}\, T_{acc}(t_{lb})$$

with $0 < t_{lb} < 1$; $\frac{N_{rej}}{N} \leq a\%$;

and any other constraints of interest;

where t_{lb} is a lower bound of the rejection region; T_{acc} is a model performance evaluation metric of that baseline machine learning model calculated using first samples of the plurality of samples outside the rejection region; N_{rej} is a number of the plurality of first samples rejected with scores within the rejection region; N is a number of the plurality of first samples in total; and $a\%$ is a user-defined rejection ratio limit for that baseline machine learning model.

3. The method of claim 2, wherein an optimal solution of t_{lb} is obtained using at last one of the following searching or optimization algorithms: a Grid search, a Bayesian optimization, a Simulated annealing, a Genetic algorithm, a Particle swarm optimization, or any combination thereof.

4. The method of claim 1, wherein, for each baseline machine learning model of the set of baseline machine learning models, the plurality of second training samples is associated with a plurality of distance measures, wherein each distance measure of the plurality of distance measures indicates a distance of a corresponding second sample from at least one boundary of the rejection region of that baseline machine learning model, and wherein training, with the at least one processor, the ensemble machine learning model is further based on (d) the plurality of distance measures for the plurality of second samples associated with each baseline machine learning model.

5. The method of claim 1, wherein each baseline machine learning model of the set of baseline machine learning models includes a multi-class classification model for predicting one of a number of classes q, where q is more than two classes, and wherein the rejection region of each baseline machine learning model includes a number $q - 1$ bounds defining the rejection region.

6. The method of claim 1, wherein a meta-model is used to ensemble the set of baseline machine learning models into the ensemble machine learning model.

7. A system for payment risk analysis, comprising:

at least one processor configured to:

(i) for each baseline machine learning model of a set of baseline machine learning models:

train that baseline machine learning model based on a plurality of first training samples associated with a plurality of first payment transactions configured to be processed in an electronic payment processing network, wherein the plurality of first training samples includes a plurality of different data types, and wherein training that baseline machine learning model generates a plurality of first predictions for the plurality of first training samples;

generate, for that baseline machine learning model, based on the plurality of first predictions for the plurality of first training samples, a rejection region associated with at least one data type of the plurality of different data types; and

process, with that baseline machine learning model, a subset of second training samples of a plurality of second training samples outside the rejection region of that baseline machine learning model, to generate a subset of second predictions for the subset of second training samples of the plurality of second training samples outside the rejection region of that baseline machine learning model, wherein the plurality of second training samples is associated with a plurality of second payment transactions configured to be processed in the electronic payment network, wherein a baseline model predictive performance metric for that baseline machine

learning model is determined based on the subset of second predictions of that baseline machine learning model, and wherein the plurality of second training samples is associated with a plurality of rejection flags for that baseline machine learning model, wherein each rejection flag of the plurality of rejection flags indicates whether a corresponding second sample of the plurality of second samples is within the rejection region of that baseline machine learning model;

(ii) generate a global rejection region associated with one or more data types of the plurality of different data types based on the rejection region associated with each baseline machine learning model;

(iii) train an ensemble machine learning model ensembled based on the set of baseline machine learning models, based on (a) a further subset of second training samples of the plurality of second training samples outside the global rejection region, (b) the plurality of rejection flags for the plurality of second samples associated with each baseline machine learning model, and (c) the subset of second predictions for the subset of second training samples generated for each baseline machine learning model, wherein training the ensemble machine learning model generates a subset of ensemble predictions for the further subset of second training samples of the plurality of second training samples outside the global rejection region, and wherein an ensemble model predictive performance metric is determined based on the subset of ensemble predictions;

(iv) update, based on the baseline model predictive performance metric for each baseline machine learning model, the set of baseline machine learning models; [[and]]

(v) repeat (i) – (iv) until there is a single baseline machine learning model in the set of baseline machine learning models or at least one of the ensemble model predictive performance metric satisfies a threshold ensemble model predictive performance, a ratio of the plurality of second training samples outside the global rejection region satisfies a threshold ratio, or any combination thereof;

(vi) obtain a current sample associated with a current payment transaction currently being processed in the electronic payment network;

(vii) determine whether the current sample is within the global rejection region;

(viii) in response to determining that the current sample is outside the global rejection region, automatically process, using the ensemble machine learning model, the current sample to generate a current prediction for the current sample of a payment risk associated with the current payment transaction; and

(ix) in response to determining that the current sample is within the global rejection region, automatically flag the current sample as unable to receive, from the ensemble classifier, a credible prediction of the payment risk associated with the current payment transaction.

8. The system of claim 7, wherein the at least one processor is configured to, for each baseline machine learning model of the set of baseline machine learning models, generate, for that baseline machine learning model, based on the plurality of first predictions for the plurality of first training samples, the rejection region associated with the at least one data type of the plurality of different data types by optimizing an objective function defined according to the following equation:

$$argmax_{(t_{lb})}\, T_{acc}(t_{lb})$$

with $0 < t_{lb} < 1$; $\frac{N_{rej}}{N} \leq a\%$;

and any other constraints of interest;

where t_{lb} is a lower bound of the rejection region; T_{acc} is a model performance evaluation metric of that baseline machine learning model calculated using first samples of the plurality of samples outside the rejection region; N_{rej} is a number of the plurality of first samples rejected with scores within the rejection region; N is a number of the plurality of first samples in total; and $a\%$ is a user-defined rejection ratio limit for that baseline machine learning model.

9. The system of claim 8, wherein an optimal solution of t_{lb} is obtained using at last one of the following searching or optimization algorithms: a Grid search, a Bayesian optimization, a Simulated annealing, a Genetic algorithm, a Particle swarm optimization, or any combination thereof.

10. The system of claim 7, wherein, for each baseline machine learning model of the set of baseline machine learning models, the plurality of second training samples is associated with a plurality of distance measures, wherein each distance measure of the plurality of distance measures indicates a distance of a corresponding second sample from at least one boundary of the rejection region of that baseline machine learning model, and wherein the at least one processor is further configured to train the ensemble machine learning model based on (d) the plurality of distance measures for the plurality of second samples associated with each baseline machine learning model.

11. The system of claim7, wherein each baseline machine learning model of the set of baseline machine learning models includes a multi-class classification model for predicting one of a number of classes *q*, where *q* is more than two classes, and wherein the rejection region of each baseline machine learning model includes a number $q - 1$ bounds defining the rejection region.

12. The system of claim 7, wherein a meta-model is used to ensemble the set of baseline machine learning models into the ensemble machine learning model.

13. A computer program product comprising at least one non-transitory computer-readable medium including program instructions for payment risk analysis that, when executed by at least one processor, cause the at least one processor to:

(i) for each baseline machine learning model of a set of baseline machine learning models:

train that baseline machine learning model based on a plurality of first training samples associated with a plurality of first payment transactions configured to be processed in an electronic payment processing network, wherein the plurality of first training samples includes a plurality of different data types, and wherein training that baseline machine learning model generates a plurality of first predictions for the plurality of first training samples;

generate, for that baseline machine learning model, based on the plurality of first predictions for the plurality of first training samples, a rejection region associated with at least one data type of the plurality of different data types; and

process, with that baseline machine learning model, a subset of second training samples of a plurality of second training samples outside the rejection region of that baseline machine learning model, to generate a subset of second predictions for the subset of second training samples of the plurality of second training samples outside the rejection region of that baseline machine learning model, wherein the plurality of second training samples is associated with a plurality of second payment transactions configured to be processed in the electronic payment network, wherein a baseline model predictive performance metric for that baseline machine learning model is determined based on the subset of second predictions of that baseline machine learning model, and wherein the plurality of second training samples is associated with a plurality of rejection flags for that baseline machine learning model, wherein each rejection flag of the plurality of rejection flags indicates whether a corresponding second sample of the plurality of second samples is within the rejection region of that baseline machine learning model;

(ii) generate a global rejection region associated with one or more data types of the plurality of different data types based on the rejection region associated with each baseline machine learning model;

(iii) train an ensemble machine learning model ensembled based on the set of baseline machine learning models, based on (a) a further subset of second training samples of the plurality of second training samples outside the global rejection region, (b) the plurality of rejection flags for the plurality of second samples associated with each baseline machine learning model, and (c) the subset of second predictions for the subset of second training samples generated for each baseline machine learning model, wherein training the ensemble machine learning model generates a subset of ensemble predictions for the further subset of second training samples of the plurality of second training samples outside the global rejection region, and wherein an ensemble model predictive performance metric is determined based on the subset of ensemble predictions;

(iv) update, based on the baseline model predictive performance metric for each baseline machine learning model, the set of baseline machine learning models; [[and]]

(v) repeat (i) – (iv) until there is a single baseline machine learning model in the set of baseline machine learning models or at least one of the ensemble model predictive performance metric satisfies a threshold ensemble model predictive performance,

a ratio of the plurality of second training samples outside the global rejection region satisfies a threshold ratio, or any combination thereof;

(vi) obtain a current sample associated with a current payment transaction currently being processed in the electronic payment network;

(vii) determine whether the current sample is within the global rejection region;

(viii) in response to determining that the current sample is outside the global rejection region, automatically process, using the ensemble machine learning model, the current sample to generate a current prediction for the current sample of a payment risk associated with the current payment transaction; and

(ix) in response to determining that the current sample is within the global rejection region, automatically flag the current sample as unable to receive, from the ensemble classifier, a credible prediction of the payment risk associated with the current payment transaction.

14. The computer program product of claim 13, wherein the program instructions, when executed by the at least one processor, cause the at least one processor to, for each baseline machine learning model of the set of baseline machine learning models, generate, for that baseline machine learning model, based on the plurality of first predictions for the plurality of first training samples, the rejection region associated with the at least one data type of the plurality of different data types by optimizing an objective function defined according to the following equation:

$$argmax_{(t_{lb})}\, T_{acc}(t_{lb})$$

with $0 < t_{lb} < 1$; $\frac{N_{rej}}{N} \leq a\%$;

and any other constraints of interest;

where t_{lb} is a lower bound of the rejection region; T_{acc} is a model performance evaluation metric of that baseline machine learning model calculated using first samples of the plurality of samples outside the rejection region; N_{rej} is a number of the plurality of first samples rejected with scores within the rejection region; N is a number of the plurality of first

samples in total; and $a\%$ is a user-defined rejection ratio limit for that baseline machine learning model.

15. The computer program product of claim 14, wherein an optimal solution of t_{lb} is obtained using at last one of the following searching or optimization algorithms: a Grid search, a Bayesian optimization, a Simulated annealing, a Genetic algorithm, a Particle swarm optimization, or any combination thereof.

16. The computer program product of claim 17, wherein, for each baseline machine learning model of the set of baseline machine learning models, the plurality of second training samples is associated with a plurality of distance measures, wherein each distance measure of the plurality of distance measures indicates a distance of a corresponding second sample from at least one boundary of the rejection region of that baseline machine learning model, and wherein the program instructions, when executed by the at least one processor, further cause the at least one processor to train the ensemble machine learning model based on (d) the plurality of distance measures for the plurality of second samples associated with each baseline machine learning model.

17. The computer program product of claim 13, wherein a meta-model is used to ensemble the set of baseline machine learning models into the ensemble machine learning model.