

CLAIMS

WHAT IS CLAIMED IS:

1. A computer-implemented method for anti-money laundering (AML) analysis, comprising:
 - (a) obtaining, by a computer, a dataset comprising a plurality of accounts, each of the plurality of accounts corresponding to an account holder among a plurality of account holders, wherein each account of the plurality of accounts is defined by a plurality of account variables, wherein the plurality of account variables comprises financial transactions;
 - (b) applying, by the computer, a trained algorithm to the dataset to generate a money laundering risk score for each of the plurality of account holders and one or more key risk drivers associated with the money laundering risk score; and
 - (c) outputting, by the computer, at least a subset of the plurality of account holders based at least on the money laundering risk scores of the plurality of account holders.
2. The method of claim 1, wherein obtaining the dataset comprises obtaining and aggregating datasets from a plurality of disparate sources.
3. The method of claim 2, wherein the datasets comprise internal datasets and external datasets.
4. The method of claim 2, wherein the plurality of disparate sources comprises one or more of the following: online and retail transactions, account and account holder characteristics in a pre-selected time window, trading surveillance platforms, PEP lists, sanction and regulatory catalogs, terror and criminal watch lists, currency exchange history, or cross-border transaction information.
5. The method of claim 2, further comprising generating at least a portion of the plurality of account variables based on the aggregated datasets.
6. The method of claim 1, wherein the trained algorithm comprises a machine learning algorithm.
7. The method of claim 6, wherein the machine learning algorithm comprises one or more of the following: a support vector machine (SVM), a naïve Bayes classification, a linear regression, a quantile regression, a logistic regression, a random forest, a neural network, gradient-boosted classifier or regressor, or another supervised or unsupervised machine learning algorithm.
8. The method of claim 6, wherein generating the money laundering risk score for a given account holder comprises processing the plurality of account variables of the accounts corresponding to the given account holder using the trained algorithm.
9. The method of claim 1, further comprising storing, by the computer, the plurality of money laundering risk scores in a database.

10. The method of claim 1, further comprising sorting the plurality of account holders based at least on the money laundering risk scores for each of the plurality of account holders.
11. The method of claim 1, wherein obtaining at least a subset of the dataset is performed through a cloud-based network.
12. The method of claim 1, further comprising outputting the subset of the plurality of account holders when the money laundering risk score of each account holder in the subset satisfies a predetermined criterion.
13. The method of claim 1, wherein each of the plurality of money laundering risk scores is indicative of a probability that the account holder corresponding to the money laundering risk score has one or more accounts comprising one or more financial transactions corresponding to money laundering activity.
14. The method of claim 13, wherein the predetermined criterion is a money laundering risk score being at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, or at least about 99%.
15. The method of claim 1, further comprising generating a weighted priority score for each of the plurality of account holders based at least on the money laundering risk score of the account holder and a quantitative measure of the account holder or of a transaction of the account holder.
16. The method of claim 15, wherein the quantitative measure comprises one or more of the following: a quantity of at-risk assets, a quantity of total assets, a net worth, a number or a total value of suspicious transactions, a length of time of a suspicious transaction or activity, or a quantitative measure related to the account holder's relationship to a set of accounts.
17. The method of claim 15, further comprising sorting the plurality of account holders based at least on the weighted priority scores for each of the plurality of account holders.
18. The method of claim 15, further comprising storing, by the computer, the plurality of weighted priority scores in a database.
19. The method of claim 15, further comprising outputting the subset of the plurality of account holders when the weighted priority score of each account holder in the subset a predetermined criterion.
20. The method of claim 19, wherein the predetermined criterion is a weighted priority score being a dollar amount of at least about 10 thousand dollars, at least about 25 thousand dollars, at least about 50 thousand dollars, at least about 75 thousand dollars, at least about 100 thousand dollars, at least about 250 thousand dollars, at least about 500 thousand dollars, at least about 750 thousand dollars, or at least about 1 million dollars.

21. The method of claim 19, wherein the predetermined criterion is a weighted priority score being a number of suspicious transactions of at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or more than 20.
22. The method of claim 19, wherein the predetermined criterion is a weighted priority score being a length of time of at least about 1, 2, 3, 4, 5, 6, or 7 days, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 months, or about 1, 2, 3, 4, 5, or more years.
23. The method of claim 1, further comprising generating, by the computer, one or more recommended decisions based on the identified subset of the plurality of account holders.
24. The method of claim 1, further comprising generating an alert when a money laundering risk score for one or more account holders among a plurality of watch list account holders satisfies a predetermined criterion.
25. The method of claim 15, further comprising generating an alert when a weighted priority score for one or more of a set of watch list account holders satisfies a predetermined criterion.
26. The method of claim 1, wherein the plurality of money laundering risk scores is generated in real time.
27. The method of claim 1, further comprising processing the identified subset of the plurality of account holders to generate an analytics chart, wherein the analytics chart comprises a visualization and analytic information of each of the identified subset of the plurality of account holders; and displaying the analytics chart to a user.
28. The method of claim 27, wherein the visualization comprises a geospatial visualization.
29. The method of claim 27, wherein the analytics chart is generated and displayed to the user in real time.
30. A computer system comprising a digital processing device comprising at least one processor, an operating system configured to perform executable instructions, a memory, and a computer program including instructions executable by the digital processing device to create an application for anti-money laundering (AML) analysis, the application comprising:
 - a first module programmed to obtain a dataset comprising a plurality of accounts, each of the plurality of accounts corresponding to an account holder among a plurality of account holders, wherein each account of the plurality of accounts is defined by a plurality of account variables, wherein the plurality of account variables comprises financial transactions;
 - a scoring module programmed to apply a trained algorithm to the dataset to generate a money laundering risk score for each of the plurality of account holders and one or more key risk drivers associated with the money laundering risk score; and

a presentation module programmed to present at least a subset of the plurality of account holders in a graphical user interface based at least on the money laundering risk scores of the plurality of account holders.

31. The system of claim 30, wherein the application further comprises an aggregation module programmed to obtain and aggregate datasets from a plurality of disparate sources.

32. The system of claim 31, wherein the datasets comprise internal datasets and external datasets.

33. The system of claim 31, wherein the plurality of disparate sources comprises one or more of the following: online and retail transactions, account and account holder characteristics, trading surveillance platforms, PEP lists, sanction and regulatory catalogs, and terror and criminal watch lists, currency exchange history, or cross-border transaction information.

34. The system of claim 31, wherein the aggregation module is programmed to further generate at least a portion of the plurality of account variables based on the aggregated datasets.

35. The system of claim 30, wherein the trained algorithm comprises a machine learning algorithm.

36. The system of claim 35, wherein the machine learning algorithm comprises one or more of the following: a support vector machine (SVM), a naïve Bayes classification, a linear regression, a quantile regression, a logistic regression, a random forest, a neural network, a gradient-boosted classifier, or another supervised or unsupervised machine learning algorithm.

37. The system of claim 35, wherein the scoring module is programmed to generate the money laundering risk score for a given account holder by processing the plurality of account variables of the accounts corresponding to the given account holder using the trained algorithm.

38. The system of claim 37, wherein the scoring module is programmed to process the plurality of account variables of the accounts corresponding to the given account holder by fitting a statistical distribution to the plurality of account variables.

39. The system of claim 30, wherein the application further comprises a storage module programmed to store, by the computer, the plurality of money laundering risk scores in a database.

40. The system of claim 30, wherein the application further comprises a sorting module programmed to sort the plurality of account holders based at least on the money laundering risk scores for each of the plurality of account holders.

41. The system of claim 30, wherein the first module is programmed to obtain at least a subset of the dataset through a cloud-based network.

42. The system of claim 30, wherein the presentation module is programmed to present the subset of the plurality of account holders for investigation when the money laundering risk score of each account holder in the subset satisfies a predetermined criterion.
43. The system of claim 30, wherein each of the plurality of money laundering risk scores is indicative of a probability that the account holder corresponding to the money laundering risk score has one or more accounts comprising one or more account variables comprising one or more financial transactions corresponding to money laundering activity.
44. The system of claim 43, wherein the predetermined criterion is a money laundering risk score being at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, or at least about 99%.
45. The system of claim 30, wherein the scoring module is programmed to further generate a weighted priority score for each of the plurality of account holders based at least on the money laundering risk score of the account holder and a quantitative measure of the account holder or of a transaction of the account holder.
46. The system of claim 45, wherein the quantitative measure comprises one or more of the following: a quantity of at-risk assets, a quantity of total assets, a net worth, a number or a total value of suspicious transactions, a length of time of a suspicious transaction or activity, or a quantitative measure related to the account holder's relationship to a set of accounts.
47. The system of claim 45, wherein the application further comprises a sorting module programmed to sort the plurality of account holders based at least on the weighted priority scores for each of the plurality of account holders.
48. The system of claim 45, wherein the application further comprises a storage module programmed to store, by the computer, the plurality of weighted priority scores in a database.
49. The system of claim 45, wherein the presentation module is programmed to present the subset of the plurality of account holders for investigation when the weighted priority score of each account holder in the subset satisfies a predetermined criterion.
50. The system of claim 49, wherein the predetermined criterion is a weighted priority score being at least about 10 thousand dollars, at least about 25 thousand dollars, at least about 50 thousand dollars, at least about 75 thousand dollars, or at least about 100 thousand dollars.
51. The system of claim 49, wherein the predetermined criterion is a weighted priority score being a number of suspicious transactions of at least about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or more than 20.

52. The system of claim 49, wherein the predetermined criterion is a weighted priority score being a length of time of at least about 1, 2, 3, 4, 5, 6, or 7 days, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 months, or about 1, 2, 3, 4, 5, or more years.
53. The system of claim 30, wherein the application further comprises a recommendation module programmed to generate, by the computer, one or more recommended decisions based on the identified subset of the plurality of account holders.
54. The system of claim 30, wherein the application further comprises an alert module programmed to generate an alert when a money laundering risk score for one or more account holders among a plurality of watch list account holders satisfies a predetermined criterion.
55. The system of claim 45, wherein the application further comprises an alert module programmed to generate an alert when a weighted priority score for one or more of a set of watch list account holders satisfies a predetermined criterion.
56. The system of claim 30, wherein the scoring module is programmed to generate the plurality of money laundering risk scores in real time.
57. The system of claim 30, wherein the application further comprises an analytics module programmed to process the identified subset of the plurality of account holders to generate an analytics chart, wherein the analytics chart comprises a visualization and analytic information of each of the identified subset of the plurality of account holders; and display the analytics chart to a user.
58. The system of claim 57, wherein the visualization comprises a geospatial visualization.
59. The system of claim 57, wherein the analytics module is programmed to generate the analytics chart and display the analytics chart to the user in real time.
60. A non-transitory computer readable medium comprising machine-executable code that, upon execution by one or more computer processors, implements a method for anti-money laundering (AML) analysis, the method comprising:
- (a) obtaining, by the computer, a dataset comprising a plurality of accounts, each of the plurality of accounts corresponding to an account holder among a plurality of account holders, wherein each account of the plurality of accounts is defined by a plurality of account variables, wherein the plurality of account variables comprises financial transactions;
 - (b) applying, by the computer, a trained algorithm to the dataset to generate a money laundering risk score for each of the plurality of account holders and one or more key risk drivers associated with the money laundering risk score; and
 - (c) outputting, by the computer, at least a subset of the plurality of account holders based at least on the money laundering risk scores of the plurality of account holders.

61. The method of claim 1, wherein the trained algorithm comprises a natural language processing algorithm configured determine a similarity score between two or more accounts of the plurality of accounts based at least in part on text-based information associated with the two or more accounts.
62. The method of claim 61, wherein the natural language processing algorithm is an n-gram model.
63. The method of claim 8, wherein the plurality of account variables comprises one or more of a cash ratio, a layering risk, a structuring risk, a credit risk, a total balance, an address change, a transaction frequency, and a transaction spacing.
64. The method of claim 1, wherein each of the plurality of money laundering risk scores is a probability that a corresponding account is being used for money laundering.
65. The method of claim 1, wherein the trained algorithm comprises layering analytics configured to detect transactions indicative of layering.
66. The method of claim 1, wherein the trained algorithm comprises foreign exchange analytics configured to identify one or more of (i) transactions associated with known terrorist organizations, (ii) transactions without FOREX discrimination, (iii) SWIFT-identified suspicious transactions, and (iv) inconsistent currency exchange transactions.
67. The method of claim 1, wherein the trained algorithm comprises suspicious actors analytics configured to target and flag transactions with recipients or senders that have known connections to sanctioned entities or criminal or terrorist databases.
68. The method of claim 1, wherein the trained algorithm comprises transaction analytics configured to identify cash ratios, identify multiple deposits of similar size, or flag usage of non-conventional monetary instruments.
69. The method of claim 1, wherein the trained algorithm comprises account analytics configured to analyze connections of accounts and account holders to correlate criminal or suspicious activity based on historical cases.
70. The method of claim 1, wherein the trained algorithm comprises structuring analytics configured to identify structured transactions.
71. The method of claim 1, wherein the trained algorithm comprises graph technology configured to analyze connections between accounts.
72. The method of claim 1, wherein (c) comprises presenting the subset of the plurality of account holders in the graphical user interface in descending order of money laundering risk scores.
73. The method of claim 1, wherein the subset of the plurality of account holders comprises account holders flagged for investigation for money laundering.

74. A system for detecting fraudulent activity, comprising:
- a user interface;
 - one or more computer processors; and
 - memory comprising machine-executable instructions that, upon execution by said the or more computer processors, cause the one or more computer processors to perform operations comprising:
 - (a) obtaining account and transaction data associated with a plurality of accounts;
 - (b) applying a trained algorithm to the account and transaction data to generate a money laundering risk score for each of the plurality of accounts; and
 - (c) presenting at least a subset of the plurality of accounts in the user interface based at least on the money laundering risk scores of the plurality of accounts.

CLAIMS

WHAT IS CLAIMED IS:

1. A computer-implemented method for anti-money laundering (AML) analysis, comprising:
 - (a) obtaining, by a computer, a dataset comprising a plurality of accounts, each of the plurality of accounts corresponding to an account holder among a plurality of account holders, wherein each account of the plurality of accounts is defined by a plurality of account variables, wherein the plurality of account variables comprises variables about financial transactions associated with the account;
 - (b) extracting, by the computer, a set of features from the dataset, wherein the set of features are associated with at least one of the plurality of account holders and at least one of the plurality of account variables;
 - (c) applying, by the computer, a trained machine learning algorithm to the set of features to determine (i) a money laundering risk score for each of the plurality of account holders and (ii) a feature importance value for each of the set of features associated with the determined money laundering risk score;
 - (d) ranking, by the computer, at least a subset of the plurality of account holders for investigation for money laundering, based at least in part on money laundering risk scores of account holders of the subset; and
 - (e) displaying, by the computer via a graphical user interface, (i) at least the subset of the plurality of account holders ranked in (d) and (ii) for each of the subset of the plurality of account holders, at least a subset of the set of features that contribute most to the money laundering risk score of the account holder.
2. The method of claim 1, wherein obtaining the dataset comprises aggregating datasets from a plurality of disparate sources selected from the group consisting of: online and retail transactions, account and account holder characteristics in a pre-selected time window, trading surveillance platforms, PEP lists, sanction and regulatory catalogs, terror and criminal watch lists, currency exchange history, and cross-border transaction information.
3. The method of claim 2, further comprising generating at least a subset of the plurality of account variables based at least in part on the aggregated datasets.
4. The method of claim 1, wherein the trained machine learning algorithm is selected from

the group consisting of: a support vector machine (SVM), a naïve Bayes classification, a linear regression, a quantile regression, a logistic regression, a random forest, a neural network, and a gradient-boosted classifier or regressor.

5. The method of claim 1, wherein each of the plurality of money laundering risk scores is indicative of a probability that the account holder corresponding to the money laundering risk score has one or more accounts comprising one or more financial transactions corresponding to money laundering activity.

6. The method of claim 1, further comprising generating a weighted priority score for each of the plurality of account holders based at least in part on the money laundering risk score of the account holder and a quantitative measure of the account holder or of a transaction of the account holder.

7. The method of claim 6, wherein the quantitative measure is selected from the group consisting of: a quantity of at-risk assets, a quantity of total assets, a net worth, a number or a total value of suspicious transactions, a length of time of a suspicious transaction or activity, and a quantitative measure related to the account holder's relationship to a set of accounts.

8. The method of claim 6, further comprising sorting the plurality of account holders based at least in part on the weighted priority scores for each of the plurality of account holders.

9. The method of claim 6, further comprising ranking an account holder of the subset of the plurality of account holders for investigation for money laundering when the weighted priority score of the account holder of the subset meets a pre-determined criterion.

10. The method of claim 9, wherein the pre-determined criterion is a dollar amount of a suspicious transaction on the account of the account holder being at least about 10 thousand dollars, at least about 25 thousand dollars, at least about 50 thousand dollars, at least about 75 thousand dollars, at least about 100 thousand dollars, at least about 250 thousand dollars, at least about 500 thousand dollars, at least about 750 thousand dollars, or at least about 1 million dollars.

11. The method of claim 9, wherein the pre-determined criterion is a number of suspicious transactions associated with the account of the account holder being at least about all 2, 3, 4, 5, 6,

7, 8, 9, 10, 15, 20, or more than 20.

12. The method of claim 6, further comprising generating an alert when a weighted priority score for one or more of a set of watch list account holders satisfies a pre-determined threshold.

13. The method of claim 1, further comprising generating, by the computer, a recommended decision based at least in part on the identified subset of the plurality of account holders.

14. The method of claim 1, further comprising generating an alert when a money laundering risk score for one or more account holders among a plurality of watch list account holders satisfies a pre-determined threshold.

15. The method of claim 1, further comprising processing the identified subset of the plurality of account holders to generate an analytics chart, wherein the analytics chart comprises a visualization and analytic information of each of the ranked subset of the plurality of account holders; and displaying the analytics chart to a user.

16. The method of claim 15, wherein the visualization comprises a geospatial visualization.

17. The method of claim 1, wherein the trained machine learning algorithm comprises a natural language processing algorithm comprising an n-gram model, wherein the natural language processing algorithm is configured to determine a similarity score between two or more accounts of the plurality of accounts based at least in part on text-based information associated with the two or more accounts.

18. The method of claim 1, wherein the plurality of account variables comprises one or more of a cash ratio, a layering risk, a structuring risk, a credit risk, a total balance, an address change, a transaction frequency, and a transaction spacing.

19. The method of claim 1, wherein the trained algorithm detects transactions indicative of layering.

20. The method of claim 1, wherein the trained algorithm detects one or more of (i) transactions associated with sanctioned entities or known criminal or terrorist organizations, (ii) transactions without FOREX discrimination, (iii) SWIFT-identified suspicious transactions, (iv)

inconsistent currency exchange transactions, and (v) structured transactions.

21. The method of claim 1, wherein the trained algorithm determines cash ratios, detects multiple deposits of similar size, or detects usage of non-conventional monetary instruments.
22. The method of claim 1, wherein the trained algorithm analyzes connections of accounts and account holders to correlate criminal or suspicious activity based on historical cases.
23. The method of claim 1, wherein (f) comprises displaying at least the subset of the plurality of account holders via a graphical user interface in descending order of money laundering risk scores.
24. The method of claim 1, wherein the variables about financial transactions comprise at least a transaction type, a transaction amount, and an aggregate transaction amount over time.
25. The method of claim 1, wherein the set of features comprises at least one of: party attributes, party behaviors, anomalies, associations, and segmentation.
26. The method of claim 1, wherein the feature importance values are indicative of relative contributions of each of the set of features used by the trained machine learning algorithm to determine the money laundering risk scores for the plurality of account holders.
27. The method of claim 1, wherein (f) further comprises displaying, for each of the subset of the plurality of account holders, feature importance values of the subset of the set of features that contribute most to the money laundering risk score of the account holder.
28. The method of claim 1, wherein (f) further comprises displaying, for each of the subset of the plurality of account holders, the subset of the set of features that contribute most to the money laundering risk score of the account holder grouped by risk typologies.
29. The method of claim 1, wherein the method further comprises, prior to (c), identifying, by the computer, at least the subset of the set of features that contribute most to the money laundering risk score based at least in part on feature importance values for the set of features, and that satisfy an objective function for training and optimizing performance of the trained machine learning algorithm.

30. The method of claim 2, wherein the datasets comprise internal datasets and external datasets.
31. The method of claim 1, wherein at least a subset of the dataset is obtained via a cloud-based network.
32. The method of claim 6, further comprising storing, by the computer, the weighted priority scores in a database.
33. The method of claim 9, wherein the predetermined criterion is a length of time of at least about 1, 2, 3, 4, 5, 6, or 7 days, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 months, or about 1, 2, 3, 4, 5, or more years.
34. The method of claim 1, wherein the plurality of money laundering risk scores is generated in real time.
35. The method of claim 15, wherein the analytics chart is generated and displayed to the user in real time.
36. The method of claim 1, wherein the trained machine learning algorithm is configured to target and flag transactions with recipients or senders that have known connections to sanctioned entities or criminal or terrorist databases.
37. The method of claim 1, wherein (e) further comprises outputting, for each of the subset of the plurality of account holders, feature importance values of the subset of the set of features that contribute most to the money laundering risk score of the account holder.
38. The method of claim 1, wherein (e) further comprises outputting, for each of the subset of the plurality of account holders, the subset of the set of features that contribute most to the money laundering risk score of the account holder grouped by risk typologies.