-------------------For dataset 1-------------------

accuracy: [0.86512370311253, 0.8677839851024208, 0.87390263367917, 0.8544825751529662, 0.882415536046821, 0.8550146315509444, 0.8627294493216281, 0.86805001330141, 0.8880021282255919, 0.8606012237297154]

---Table of mismatch (3759 samples)---

majority vote | weighted | truthFinder | EM | MV

404 442 628 17 628

Used EM in this model

----baseline ----

Accurary for DT: 0.6396276595744681 ROCAUC = 0.6852836879432624 for RF: 0.7566489361702128 ROCAUC = 0.7154255319148936

----baseline when given accurate tag----

the accurary is: 0.6396276595744681 for DT, 0.7632978723404256 for RF

-------------------For dataset 2-------------------

accuracy: [0.8906624102154829, 0.87975525405693, 0.9063580739558393, 0.88560787443469, 0.8965150305932429, 0.8800212822559191, 0.8922585794094174, 0.9098164405426975, 0.8792231976589518, 0.894120776802341]

---Table of mismatch (3759 samples)---

majority vote | weighted | truthFinder | EM | MV

200 398 412 5 412

Used EM in this model

----baseline ----

Accurary for DT: 0.6396276595744681 ROCAUC = 0.6852836879432624 for RF: 0.7553191489361702 ROCAUC = 0.7021276595744681

-------------------For dataset 3-------------------

accuracy: [0.8786911412609737, 0.8800212822559191, 0.8757648310720937, 0.9111465815376429, 0.888268156424581, 0.8975791433891993, 0.8494280393721735, 0.9140728917265231, 0.8938547486033519, 0.9279063580739558]

---Table of mismatch (3759 samples)---

majority vote | weighted | truthFinder | EM | MV

207 323 405 4 405

Used EM in this model

----baseline ----

Accurary for DT: 0.6343085106382979 ROCAUC = 0.6675531914893617 for RF: 0.761968085106383 ROCAUC = 0.7047872340425532

-------------------For dataset 4-------------------

accuracy: [0.9175312583133812, 0.9366852886405959, 0.9207235967012504, 0.936951316839585, 0.9329608938547486, 0.9507847831870179, 0.9140728917265231, 0.9239159350891194, 0.9239159350891194, 0.9332269220537377]

---Table of mismatch (3759 samples)---

majority vote | weighted | truthFinder | EM | MV

34 251 113 1 113

Used EM in this model

----baseline ----

Accurary for DT: 0.6396276595744681 ROCAUC = 0.6852836879432624 for RF: 0.7553191489361702 ROCAUC = 0.7056737588652482

-------------------For dataset 5-------------------

accuracy: [0.9342910348496941, 0.9422718808193669, 0.9154030327214685, 0.9111465815376429, 0.9116786379356212, 0.9156690609204575, 0.9406757116254323, 0.9241819632881085, 0.9374833732375631, 0.9561053471667997]

---Table of mismatch (3759 samples)---

majority vote | weighted | truthFinder | EM | MV

38 165 117 0 117

Used EM in this model

----baseline ----

Accurary for DT: 0.6396276595744681 ROCAUC = 0.6852836879432624 for RF: 0.7632978723404256 ROCAUC = 0.7127659574468085

-------------------For dataset 6-------------------

accuracy: [0.9473264166001596, 0.939877627028465, 0.9547752061718542, 0.9526469805799415, 0.8922585794094174, 0.9561053471667997, 0.9792498004788508, 0.9058260175578612, 0.9622239957435488, 0.9664804469273743]

---Table of mismatch (3759 samples)---

majority vote | weighted | truthFinder | EM | MV

7 177 32 0 32

Used EM in this model

----baseline ----

Accurary for DT: 0.6396276595744681 ROCAUC = 0.6852836879432624 for RF: 0.7632978723404256 ROCAUC = 0.7127659574468085

-------------------For dataset 7-------------------

accuracy: [0.9377494014365523, 0.9582335727587125, 0.9819100824687417, 0.9430699654163341, 0.9446661346102687, 0.9643522213354616, 0.9510508113860069, 0.9606278265496142, 0.965682362330407, 0.9305666400638468]

---Table of mismatch (3759 samples)---

majority vote | weighted | truthFinder | EM | MV

3 134 14 0 14

Used EM in this model

----baseline ----

Accurary for DT: 0.6396276595744681 ROCAUC = 0.6852836879432624 for RF: 0.7632978723404256 ROCAUC = 0.7127659574468085

-------------------For dataset 8-------------------

accuracy: [0.9590316573556797, 0.9672785315243416, 0.9603617983506252, 0.974195264698058, 0.9760574620909817, 0.9201915403032721, 0.9696727853152434, 0.9292364990689013, 0.9462623038042033, 0.9630220803405161]

---Table of mismatch (3759 samples)---

majority vote | weighted | truthFinder | EM | MV

3 123 15 0 15

Used EM in this model

----baseline ----

Accurary for DT: 0.6396276595744681 ROCAUC = 0.6852836879432624 for RF: 0.7632978723404256 ROCAUC = 0.7127659574468085

-------------------For dataset 9-------------------

accuracy: [0.9523809523809523, 0.9584996009577015, 0.9529130087789306, 0.9672785315243416, 0.9433359936153233, 0.9510508113860069, 0.9523809523809523, 0.9720670391061452, 0.9686086725192871, 0.9606278265496142]

---Table of mismatch (3759 samples)---

majority vote | weighted | truthFinder | EM | MV

6 117 18 0 18

Used EM in this model

----baseline ----

Accurary for DT: 0.6396276595744681 ROCAUC = 0.6852836879432624 for RF: 0.7632978723404256 ROCAUC = 0.7127659574468085

-------------------For dataset 10-------------------

accuracy: [0.9553072625698324, 0.9680766161213089, 0.9725990955041235, 0.9547752061718542, 0.9787177440808725, 0.9696727853152434, 0.9579675445597233, 0.9553072625698324, 0.9467943602021814, 0.9651503059324288]

---Table of mismatch (3759 samples)---

majority vote | weighted | truthFinder | EM | MV

1 111 6 0 6

Used EM in this model

----baseline ----

Accurary for DT: 0.6396276595744681 ROCAUC = 0.6852836879432624 for RF: 0.7632978723404256 ROCAUC = 0.7127659574468085

(len(diag)\*\*abs(temp))

-------------For dataset 1---------

Using the best (iteration, learning rate)= (50, 1),

the accuracy is: 0.7553191489361702

fpr: [0. 0.11347518 1. ] tpr [0. 0.36170213 1. ] AUC area: 0.624113475177305

Originally: Using the best (iteration, learning rate)= (50, 1),

the accuracy is: 0.7792553191489362

------------- This Dataset Ends Here---------

-------------For dataset 2---------

Using the best (iteration, learning rate)= (310, 0.1),

the accuracy is: 0.7632978723404256

fpr: [0. 0.10992908 1. ] tpr [0. 0.38297872 1. ] AUC area: 0.6365248226950354

------------- This Dataset Ends Here---------

-------------For dataset 3---------

Using the best (iteration, learning rate)= (250, 0.01),

the accuracy is: 0.7632978723404256

fpr: [0. 0.12588652 1. ] tpr [0. 0.43085106 1. ] AUC area: 0.6524822695035462

------------- This Dataset Ends Here---------

-------------For dataset 4---------

Using the best (iteration, learning rate)= (270, 0.25),

the accuracy is: 0.773936170212766

fpr: [0. 0.06028369 1. ] tpr [0. 0.27659574 1. ] AUC area: 0.6081560283687943

------------- This Dataset Ends Here---------

-------------For dataset 5---------

Using the best (iteration, learning rate)= (150, 1),

the accuracy is: 0.75

fpr: [0. 0.05851064 1. ] tpr [0. 0.17553191 1. ] AUC area: 0.5585106382978723

------------- This Dataset Ends Here---------

-------------For dataset 6---------

Using the best (iteration, learning rate)= (30, 1),

the accuracy is: 0.7406914893617021

fpr: [0. 0.11524823 1. ] tpr [0. 0.30851064 1. ] AUC area: 0.5966312056737588

------------- This Dataset Ends Here---------

-------------For dataset 7---------

Using the best (iteration, learning rate)= (310, 1),

the accuracy is: 0.7513297872340425

fpr: [0. 0.12234043 1. ] tpr [0. 0.37234043 1. ] AUC area: 0.625

------------- This Dataset Ends Here---------

-------------For dataset 8---------

Using the best (iteration, learning rate)= (270, 0.25),

the accuracy is: 0.7340425531914894

fpr: [0. 0.10283688 1. ] tpr [0. 0.24468085 1. ] AUC area: 0.5709219858156029

------------- This Dataset Ends Here---------

-------------For dataset 9---------

Using the best (iteration, learning rate)= (90, 1),

the accuracy is: 0.7393617021276595

fpr: [0. 0.10815603 1. ] tpr [0. 0.28191489 1. ] AUC area: 0.5868794326241135

------------- This Dataset Ends Here---------

-------------For dataset 10---------

Using the best (iteration, learning rate)= (250, 0.25),

the accuracy is: 0.7632978723404256

fpr: [0. 0.09574468 1. ] tpr [0. 0.34042553 1. ] AUC area: 0.6223404255319149

------------- This Dataset Ends Here---------

-------------For dataset 1---------

Using the best subsample = 0.8

the accuracy is: 0.7805851063829787

Originally: Using the best subsample = 0.8

the accuracy is: 0.7712765957446809

------------- This Dataset Ends Here---------

-------------For dataset 2---------

Using the best subsample = 0.6

the accuracy is: 0.7765957446808511

Originally: Using the best subsample = 0.6

the accuracy is: 0.7752659574468085

------------- This Dataset Ends Here---------

-------------For dataset 3---------

Using the best subsample = 1.0

the accuracy is: 0.7672872340425532

Originally: Using the best subsample = 1.0

the accuracy is: 0.7845744680851063

------------- This Dataset Ends Here---------

-------------For dataset 4---------

Using the best subsample = 0.9

the accuracy is: 0.7779255319148937

Originally: Using the best subsample = 0.9

the accuracy is: 0.7832446808510638

------------- This Dataset Ends Here---------

-------------For dataset 5---------

Using the best subsample = 0.8

the accuracy is: 0.7819148936170213

Originally: Using the best subsample = 0.8

the accuracy is: 0.7779255319148937

------------- This Dataset Ends Here---------

-------------For dataset 6---------

Using the best subsample = 0.8

the accuracy is: 0.7752659574468085

Originally: Using the best subsample = 0.8

the accuracy is: 0.7765957446808511

------------- This Dataset Ends Here---------

-------------For dataset 7---------

Using the best subsample = 0.8

the accuracy is: 0.7699468085106382

Originally: Using the best subsample = 0.8

the accuracy is: 0.7779255319148937

------------- This Dataset Ends Here---------

-------------For dataset 8---------

Using the best subsample = 0.6

the accuracy is: 0.7712765957446809

Originally: Using the best subsample = 0.6

the accuracy is: 0.7805851063829787

------------- This Dataset Ends Here---------

-------------For dataset 9---------

Using the best subsample = 0.6

the accuracy is: 0.7832446808510638

Originally: Using the best subsample = 0.6

the accuracy is: 0.7832446808510638

------------- This Dataset Ends Here---------

-------------For dataset 10---------

Using the best subsample = 0.7

the accuracy is: 0.7659574468085106

Originally: Using the best subsample = 0.7

the accuracy is: 0.7287234042553192

------------- This Dataset Ends Here---------

In [ ]: