

CLAIMS

What is claimed is:

1. A method for classifying samples to be assessed, comprising:
 - obtaining sample T to be assessed and sample feature F_t of sample T to be
5 assessed;
 - selecting the first quantity N of example samples from a classification sample
database;
 - obtaining feature similarity SIM_i between sample T to be assessed and each of
the N example samples i, wherein feature similarity SIM_i is determined based on
10 sample feature F_t of sample T to be assessed and sample feature F_i of each example
sample i;
 - obtaining sample quality Q_i of each example sample i, wherein for each example
sample i, the sample quality Q_i is indicative of a generalization ability of the example
sample i when compared against different candidate samples;
 - 15 determining comprehensive similarities S_i between sample T to be assessed and
each example sample i based on at least a difference values r_i between feature
similarity SIM_i and sample quality Q_i ;
 - determining, based on comprehensive similarity S_i , whether sample T to be
assessed falls in the category of the classification sample database.
 - 20
2. The method according to claim 1, wherein the selecting the first quantity N of
example samples from a classification samples comprises:
 - calculating feature similarities between sample T to be assessed and each of a
second quantity M of example samples based on sample feature F_t of sample T to be
assessed and sample features of the second quantity M of example samples in the
25 classification sample database, wherein the second quantity M is greater than the first
quantity N; and
 - selecting the first quantity N of example samples from the second quantity M of
example samples based on feature similarity between the sample to be assessed and
each of the second quantity M of example samples.
- 30
3. The method according to claim 2, wherein the selecting the first quantity N of
example samples from the second quantity M of example samples comprises:
selecting, from the second quantity M of example samples, the first quantity N of
example samples with the highest feature similarities to sample T to be assessed.

4. The method according to claim 1, wherein the selecting the first quantity N of example samples from a classification sample database comprises: selecting the first quantity N of example samples from the classification sample database based on the sorting result of sample quality of each sample in the classification sample database.

5 5. The method according to claim 1, wherein feature similarity SIM_i is determined by normalizing the distance between sample feature F_t of sample T to be assessed and sample feature F_i of each example sample i.

6. The method according to claim 1, wherein the determining comprehensive similarities S_i between sample T to be assessed and each example sample i comprises:
10 determining comprehensive similarity S_i as $S_i = a + b * r_i * c$, wherein $a + b = 1$ and c is a coefficient associated with sample quality Q_i .

7. The method according to claim 6, wherein in the case of $r_i \geq 0$, $c = 1/(1-Q_i)$ and in the case of $r_i < 0$, $c = 1/Q_i$.

8. The method according to claim 1, wherein the determining whether sample T
15 to be assessed falls in the category of the classification sample database based on comprehensive similarity S_i comprises:

 determining, based on comprehensive similarities S_i between sample T to be assessed and each example sample i, a total similarity score of the sample to be assessed; and

20 determining, based on the total similarity score, whether sample T to be assessed falls in the category of the classification sample database.

9. The method according to claim 8, wherein the determining the total similarity score of the sample to be assessed comprises:

 if at least one $r_i \geq 0$, determining the total similarity score as the maximum
25 value among comprehensive similarities S_i between sample T to be assessed and each example sample i; or

 otherwise, determining the total similarity score as the minimum value among comprehensive similarities S_i between sample T to be assessed and each example sample i.

30 10. The method according to claim 8, wherein the determining a total similarity score of the sample to be assessed comprises: determining the total similarity score as an average value of comprehensive similarities S_i between the sample to be assessed and each example sample i.

11. A device for classifying samples to be assessed, comprising:

a sample acquisition unit, configured to obtain sample T to be assessed and sample feature F_t of sample T to be assessed;

a selection unit, configured to select the first quantity N of example samples from a classification sample database;

5 a first acquisition unit, configured to obtain feature similarity SIM_i between sample T to be assessed and each of the N example samples i, wherein feature similarity SIM_i is determined based on sample feature F_t of sample T to be assessed and sample feature F_i of each example sample i;

10 a second acquisition unit, configured to obtain sample quality Q_i of each example sample i, wherein for each example sample i, the sample quality Q_i is indicative of a generalization ability of the example sample i when compared against different candidate samples;

15 a processing unit, configured to determine comprehensive similarities S_i between sample T to be assessed and each example sample i based on at least a difference value r_i between feature similarity SIM_i and sample quality Q_i ; and

a classification unit, configured to determine, based on comprehensive similarity S_i , whether sample T to be assessed falls in the category of the classification sample database.

12. The device according to claim 11, wherein the selection unit comprises:

20 a calculation subunit, configured to calculate, based on sample feature F_t of sample T to be assessed and sample features of a second quantity M of example samples, feature similarities between each of the second quantity M of example samples and sample T to be assessed, wherein the second quantity M is greater than the second quantity N.

25 a selection subunit, configured to select the first quantity N of example samples from the second quantity M of example samples based on the feature similarities between the sample to be assessed and each of the second quantity M of example samples.

30 13. The device according to claim 12, wherein the selection subunit is configured to select, from the second quantity M of example samples, the first quantity N of example samples with the highest feature similarities to sample T to be assessed.

14. The device according to claim 11, wherein the selection unit is configured to select the first quantity N of example samples from the classification sample database based on the sorting result of the sample quality of each sample in the classification

sample database.

15. The device according to claim 11, wherein feature similarity SIM_i is determined by normalizing the distance between sample feature F_t of sample T to be assessed and sample feature F_i of each example sample i.

5 16. The device according to claim 11, wherein the processing unit is configured to determine comprehensive similarity S_i as $S_i = a + b * r_i * c$, wherein $a + b = 1$, and c is a coefficient associated with sample quality Q_i .

17. The device according to claim 16, wherein in the case of $r_i \geq 0$, $c = 1/(1-Q_i)$ and in the case of $r_i < 0$, $c = 1/Q_i$.

10 18. The device according to claim 11, wherein the classification unit is configured to:

 determine, based on comprehensive similarities S_i between sample T to be assessed and each example sample i, a total similarity score of the sample to be assessed; and

15 determine, based on the total similarity score, whether sample T to be assessed falls in the category of the classification sample database.

19. The device according to claim 18, wherein the classification unit is configured to:

 if at least one $r_i \geq 0$, determine the total similarity score as the maximum value
20 among comprehensive similarities S_i between sample T to be assessed and each example sample i; or

 otherwise, determine the total similarity score as the minimum value among comprehensive similarities S_i between sample T to be assessed and each example sample i.

25 20. The device according to claim 18, wherein the classification unit is configured to determine the total similarity score as the average value of comprehensive similarities S_i of sample T to be assessed and each example sample i.

21. A computer readable storage medium, wherein the computer readable storage medium stores a computer program, and when the computer program is executed in a
30 computer, the computer is enabled to perform the method according to any one of claims 1 to 10.

22. A computing device, comprising a memory and a processor, wherein the memory stores executable code, and when executing the executable code, the processor implements the method according to any one of claims 1 to 10.