ISIA Theme Report

Database: Cylinder Bands - <https://archive.ics.uci.edu/dataset/32/cylinder+bands>

Cylindrical Bands are used in decision tree induction for process delay mitigation.

Instances: 512

Attributes: 39

Character attributes: Categorical, Integer, Real

As libraries we used: sklearn,numpy,random

We split the train data and the test data, 75% - train 25% - test.

We calculated performance using SVM - Support Vector Machine

We varied the Cost to calculate the accuracy.

We started by reading the data. I transformed the matrix into a line. I replaced the following in the database:

39. Band = 1 Noband = 0

1. I removed the first column (timestamp)

2. I removed the second column (cylinder number)

3. I removed the third line (client)

4.I removed the fourth line (order number)

5. grain screened: yes = 1, no = 0

6. ink colour: nominal; key = 1, type = 0

7. ctd ink proof: nominal; yes = 1 , no = 0

8. Blade factory: nominal; benton = 0, daetwyler = 1, uddeholm = 2

9. cylinder division: nominal; gallatin = 0, warsaw = 1, mattoon = 2

10. paper type: nominal; uncoated = 0, coated = 1, super = 2

11. ink type: nominal; uncoated = 0, coated = 1, coated = 3

12. direct steam: nominal; use; yes = 1, no = 0\*

13. solvent type: nominal; xylol = 0, lactol = 1, naphtha = 2, line = 3

14. type per cylinder: nominal; yes = 1, no = 0

15. press type: nominal; use; 70 wood hoe = 0, 70 motter = 2, 70 albert = 3, 94 motter =1

18. cylinder size: nominal; catalogue=0, Spiegel=1, tabloid=2

19. paper mill location: nominal; northern US=0, southern US=4, Canadian=1, Scandinavian=2, mid-European=3

The following are replaced the 1000 "?" with the respective values of each attribute.

I have formatted x - attributes (without the first 4 columns and without the last one) , y - labels.(last column)

After that I have separated the train and test data. To further calculate the accuracy, vary the cost.

|  |  |
| --- | --- |
| Cost | Accuracy |
| 2-5 | 0.5234375 |
| 2-4 | 0.53125 |
| 2-3 | 0.5 |
| 2-2 | 0.484375 |
| 2-1 | 0.5 |
| 20 | 0.4609375 |
| 21 | 0.4921875 |
| 22 | 0.46875 |
| 23 | 0.453125 |
| 24 | 0.4296875 |
| 25 | 0.453125 |
| 26 | 0.4453125 |
| 27 | 0.515625 |