

Sistemas Inteligentes Aplicados

Trabalho

Parte 1: implemente comitês de classificadores (5 pontos)

1.1 Você pode criar seu comitê em C, C++, Java, Python ou Ruby.

1.2 O sistema deve ser portátil. Não utilize bibliotecas dependentes de plataforma, por exemplo, conio.h presente em compiladores C da Borland. Não utilize interface gráfica. Falhar neste requisito acarretará na perda de 2.0 pontos da nota.

1.3 Pré-processar o dataset Iris no Weka rodando três classificadores diferentes.

1.3.1 Ao rodar o classificador, clicar em “more options” (aba classify)

1.3.2 Clicar em “choose” ao lado de “output predictions”

1.3.3 Selecionar CSV

1.3.4 Clicar na palavra “CSV” ao lado de “output predictions”

1.3.5 Output predictions = true

1.3.6 Output file = “resultado.csv” (clicar no hífen ao lado de “output file”)

1.3.7 Executar o classificador escolhido

1.4 Exemplo de execução:

```
> <nome-do-program> resultado1.csv resultado2.csv resultado3.csv
```

1.5 Exemplo de entradas:

1.5.1 resultado1.csv

```
inst#,actual,predicted,error,distribution,,
1,3:Iris-virginica,3:Iris-virginica,,0,*0.9,0.1
2,3:Iris-virginica,3:Iris-virginica,,0,0.062,*0.938
3,3:Iris-virginica,3:Iris-virginica,,0,0,*1
4,3:Iris-virginica,3:Iris-virginica,,0,0.001,*0.999
5,3:Iris-virginica,3:Iris-virginica,,0,0,*1
[...]
```

1.5.2 resultado2.csv

```
inst#,actual,predicted,error,distribution,,
1,3:Iris-virginica,3:Iris-virginica,,0,*0.667,0.333
2,3:Iris-virginica,3:Iris-virginica,,0,0.333,*0.667
3,3:Iris-virginica,3:Iris-virginica,,0,0.333,*0.667
4,3:Iris-virginica,3:Iris-virginica,,0,0.333,*0.667
5,3:Iris-virginica,3:Iris-virginica,,0,0.333,*0.667
[...]
```

1.5.3 resultado3.csv

```
inst#,actual,predicted,error,distribution,,
1,3:Iris-virginica,3:Iris-virginica,,0.007,*0.986,0.007
2,3:Iris-virginica,3:Iris-virginica,,0.007,0.007,*0.986
3,3:Iris-virginica,3:Iris-virginica,,0.007,0.007,*0.986
4,3:Iris-virginica,3:Iris-virginica,,0.007,0.007,*0.986
5,3:Iris-virginica,3:Iris-virginica,,0.007,0.007,*0.986
[...]
```

1.5.4 Note que as três últimas colunas forma quase um vetor one-hot.

1.6 Para cada instância, seu classificador deve selecionar a classe mais votada por: soma das probabildiades e produto das probabilidades

1.7 Exemplo de saída:

s1	s2	s3	cs	p1	p2	p3	cp
0,007	2,553	0,440	2	0,000	0,591	0,002	2
0,007	0,402	2,591	3	0,000	0,002	0,616	3
0,007	0,340	2,653	3	0,000	0,000	0,657	3
0,007	0,341	2,652	3	0,000	0,000	0,657	3
0,007	0,340	2,653	3	0,000	0,000	0,657	3

[...]

1.7 s1 são as somas dos três classificadores para a classe setosa; s2 para versicolor e s3 para virginica; cs é a classe escolhida (1 para setosa, 2 para versicolor, 3 para vîrginica)

1.8 Os valores p1, p2, p3 e cp são análogos, mas para o produto das probabilidades

1.9 Você pode usar bibliotecas prontas para importar o CSV, mas não para fazer os cálculos do comitê.

Parte 2: faça o relatório (5 pontos)

2.1 Faça um relatório contendo: ao menos 400 palavras (usar “contar palavras” do Word ou similar).

Obs1: códigos fontes, pseudo-códigos e tabelas não contam como palavras.

Obs2: excertos de arquivos de entradas e saídas do seu sistema não contam como palavras.

2.2 Faça uma explicação resumida, alto nível em pseudo-código do seu sistema.

2.3 Faça uma explicação sobre comitê de classificadores (também chamados de ensembles ou combinação de classificadores).

2.4 Mostre os resultados para instâncias obtidas pelo seu comitê.

Considerações finais

3.1 A atividade deve ser feita individualmente.

3.2 Uma explicação sobre como funciona as medidas de impureza de conjunto e de qualidade da divisão.

3.3 Seu código será submetido a uma bateria de testes. Ou ele passa ou reprova em um teste.

3.4 Código não compilável/interpretável **anulará o trabalho**

3.5 Detecção de plágio, mesmo que em pequenas porções de texto ou código **anularão o trabalho**.
Se estiver em dúvida sobre o que caracteriza plágio, consulte o professor.