

TRABALHO DA DISCIPLINA IAA 003 – LINGUAGEM R

Equipe 16

Ana Beatriz Kindinger

Daniel Victor Andrade

Igor Buess Atala Y Mansour

Marcelo da Cruz Salvador

Marlon Mateus Prudente de Oliveira

Ronaldo Santana da Silva Moco

1 - Pesquisa com Dados de Satélite (Satellite)

Base de dados Satellite da biblioteca mlbench.

Treino 80% e Teste 20%

Colunas usadas no treino: *x17, x18, x19, x20* e *classes*.

Seed = 7

Resultados obtidos

Matriz de confusão – Random Forest

Predição	red soil	cotton crop	grey soil	damp grey soil	vegetation stubble	very damp grey soil
red soil	296	1	5	2	8	0
cotton crop	0	123	0	0	3	1
grey soil	3	0	239	28	1	10
damp grey soil	1	0	20	60	1	40
vegetation stubble	6	10	0	0	118	8
very damp grey soil	0	6	7	35	10	242

Acurácia: 0,8396

Matriz de confusão – SVM

Predição	red soil	cotton crop	grey soil	damp grey soil	vegetation stubble	very damp grey soil
red soil	299	1	4	2	7	0
cotton crop	1	120	0	0	4	0
grey soil	4	0	260	29	1	12
damp grey soil	0	1	7	69	2	32
vegetation stubble	2	14	0	2	117	3

Acurácia: 0,8715

Matriz de confusão RNA

Predição	red soil	cotton crop	grey soil	damp grey soil	vegetation stubble	very damp grey soil
red soil	291	2	1	0	10	0
cotton crop	0	118	0	0	0	0
grey soil	6	0	252	85	4	69
damp grey soil	0	0	0	0	0	0
vegetation stubble	0	1	0	0	3	0

Acurácia: 0,6978

O melhor modelo de predição foi obtido a partir do método SVM, pois é o que apresenta maior acurácia na base de teste.

2 - Estimativa de Volumes de Árvores

Base de dados utilizada disponível no link <http://www.razer.net.br/datasets/Volumes.csv>, carregamento realizado por meio da função “read.csv2” indicando os parâmetros de separador “;” e símbolo decimal “,”.

Coluna NR eliminada usando a função “subset” indicando os parâmetros do *dataframe* original e “-” a coluna indesejada.

Treino 80% e Teste 20%

Resultados obtidos com as funções UDF:

Random Forest

- R^2 : 0,8538019
- S_{xy} : 0,1445527
- $S_{xy}\%$: 11,07658
- RMSE: 0,1371347

SVM

- R^2 : 0,849634
- S_{xy} : 0,1470481
- $S_{xy}\%$: 11,2678
- RMSE: 0,1395021

Redes Neurais

- R^2 : 0,8825988
- S_{xy} : 0,1295415
- $S_{xy}\%$: 9,926323
- RMSE: 0,1228938

Modelo alométrico de Spurr

- R^2 : 0,8271824
- S_{xy} : 0,1574296
- $S_{xy}\%$: 12,0633
- RMSE: 0,1493509

O modelo escolhido é o de Redes Neurais, pois apresenta os melhores resultados em todas as métricas propostas.