

SELEÇÃO DO ALGORITMO DE APRENDIZAGEM DE MÁQUINA

Wendley, 06/11/2018

Algumas linhas do arquivo CSV:

```
,rssi,pr,pr2,snr,txentrega,relacao,latencia-ms,potencia
61,-109.160458782,0.5,0.5,-3.4127756946,44.4444444444,7.875,999.0,45.0
62,-109.144513482,0.5,0.5,-3.94957215834,42.1052631579,8.0,999.0,45.0
63,-109.1324661,0.5,0.5,-1.94352365376,42.1052631579,8.125,999.0,45.0
64,-109.090254156,0.5,0.5,-3.53613334798,42.1052631579,8.25,999.0,45.0
65,-109.194556812,0.5,0.5,-2.91647114115,40.0,8.375,999.0,45.0
66,-109.240579826,0.5,0.5,-3.56539245306,40.0,8.5,999.0,45.0
```

Disponível em: <https://github.com/wendley/sdr/blob/master/traces-Cen1a4.csv> (71472 linhas)

Métodos utilizados nos testes:

- Validação cruzada (Cross-validation) com Shuffle Split
- Validação cruzada com kFold
- Proporção 80%/20%

Inputs: ['rssi','pr','snr']

Target: 'txentrega'

	Acurácia Cross Validation (Shuffle split=5)	Acurácia Cross Validation (Kfold=5)	80% / 20% RMSE	80% / 20% Explained variance score *	80% / 20% MAE
Linear Regression	0.93 (+/- 0.00)	0.93 (+/- 0.00)	8.6708	0.9286	5.6977
SVR	0.92 (+/- 0.00)	0.92 (+/- 0.00)	9.0880	0.9217	4.8781
LinearSVR	0.90 (+/- 0.03)	0.90 (+/- 0.04)	10.2189	0.9240	6.6723
NuSVR	0.78 (+/- 0.00)	0.78 (+/- 0.01)	14.9823	0.8152	13.5685
KNN	0.94 (+/- 0.00)	0.94 (+/- 0.00)	7.6541	0.9444	3.9763
Decision Tree Reg.	0.97 (+/- 0.00)	0.97 (+/- 0.00)	5.7607	0.9685	2.3567
NNet	0.95 (+/- 0.03)	0.95 (+/- 0.02)	8.4909	0.9327	5.9932
GBE	0.97 (+/- 0.00)	0.97 (+/- 0.00)	5.7263	0.9688	3.1882

* Explained variance regression score function: Best possible score is 1.0, lower values are worse.

A seguir, mais detalhes sobre o tempo de execução e outras métricas (R^2 e MSE)

Cross validation com shuffle split

Scores 1 - LinearRegression:

[0.92444975 0.92863816 0.92941018 0.92749279 0.92708773]

Accuracy: 0.93 (+/- 0.00)

Tempo de cross validation:

0:00:00.339168

Scores 2 - SVR:

[0.91664057 0.92272254 0.92365913 0.92075292 0.92030868]

Accuracy: 0.92 (+/- 0.00)

Tempo de cross validation:

0:27:34.783880

Scores 3 - LinearSVR:

[0.89745628 0.90176684 0.92292723 0.91546824 0.88434001]

Accuracy: 0.90 (+/- 0.03)

Tempo de cross validation:

0:00:40.894546

Scores 4 - NuSVR:

[0.78263038 0.78525376 0.78479349 0.7813724 0.7831869]

Accuracy: 0.78 (+/- 0.00)

Tempo de cross validation:

0:06:54.158864

Scores 5 - KNN:

[0.93984595 0.94454279 0.94497517 0.94028477 0.94371393]

Accuracy: 0.94 (+/- 0.00)

Tempo de cross validation:

0:00:00.595204

Scores 6 - DecisionTreeRegressor:

[0.9653723 0.96882996 0.96701018 0.96831844 0.96401301]

Accuracy: 0.97 (+/- 0.00)

Tempo de cross validation:

0:00:02.027357

Scores 7 - NNet:

[0.92669126 0.962206 0.93311013 0.9617979 0.95679611]

Accuracy: 0.95 (+/- 0.03)

Tempo de cross validation:

0:03:58.837893

Scores 8 - GBE:

[0.96717125 0.96908217 0.96976121 0.96851752 0.9675402]

Accuracy: 0.97 (+/- 0.00)

Tempo de cross validation:

0:00:10.941112

Cross validation com KFold = 5

Scores 1 - LinearRegression:

[0.92753751 0.92527324 0.92846216 0.92717315 0.92865424]

Accuracy: 0.93 (+/- 0.00)

Tempo de cross validation:

0:00:00.332766

Scores 2 - SVR:

[0.92064519 0.91984772 0.92141502 0.92077003 0.92162819]

Accuracy: 0.92 (+/- 0.00)

Tempo de cross validation:

0:27:18.859683

Scores 3 - LinearSVR:

[0.8728417 0.91592413 0.92217188 0.91792192 0.89289644]

Accuracy: 0.90 (+/- 0.04)

Tempo de cross validation:

0:00:41.007015

Scores 4 - NuSVR:

[0.7844322 0.78437582 0.78252651 0.77847571 0.78698498]

Accuracy: 0.78 (+/- 0.01)

Tempo de cross validation:

0:06:45.907892

Scores 5 - KNN:

[0.9403039 0.94406872 0.94369181 0.94056877 0.94440997]

Accuracy: 0.94 (+/- 0.00)

Tempo de cross validation:

0:00:00.908110

Scores 6 - DecisionTreeRegressor:

[0.96636489 0.965202 0.96550781 0.96478963 0.96852649]

Accuracy: 0.97 (+/- 0.00)

Tempo de cross validation:

0:00:02.253831

Scores 7 - NNet:

[0.93067327 0.95056222 0.96184397 0.95997811 0.95111198]

Accuracy: 0.95 (+/- 0.02)

Tempo de cross validation:

0:03:51.281845

Scores 8 - GBE:

[0.9683126 0.9679189 0.96849253 0.96884281 0.9691194]

Accuracy: 0.97 (+/- 0.00)

Tempo de cross validation:

0:00:11.042854

Treino com 80% e teste com 20%

LinearRegression -----

Tempo de treinamento/carregamento:

0:00:00.005924

Tempo de predicao:

0:00:00.000426

r2 score: 0.928660

mse: 75.183444

rmse: 8.670838714

mae: 5.697742

explained variance score: 0.928660

SVR -----

Tempo de treinamento/carregamento:

0:03:49.874539

Tempo de predicao:

0:00:18.930311

r2 score: 0.921630

mse: 82.592285

rmse: 9.088029764

mae: 4.878133

explained variance score: 0.921751

Linear SVR -----

Tempo de treinamento/carregamento:

0:00:06.986434

Tempo de predicao:

0:00:00.000340

r2 score: 0.900911

mse: 104.427458

rmse: 10.218975389

mae: 6.672393

explained variance score: 0.924046

NuSVR -----

Tempo de treinamento/carregamento:

0:01:20.147258

Tempo de predicao:

0:00:02.507026

r2 score: 0.787005

mse: 224.470249

rmse: 14.982331227

mae: 13.568541

explained variance score: 0.815299

KNN -----

Tempo de treinamento/carregamento:

0:00:00.056518

Tempo de predicao:

0:00:00.074330

r2 score: 0.944409

mse: 58.586560

rmse: 7.654185783

mae: 3.976390

explained variance score: 0.944413

DecisionTreeRegressor -----

Tempo de treinamento/carregamento:

0:00:00.444863

Tempo de predicao:

0:00:00.007448

r2 score: 0.968511

mse: 33.185952

rmse: 5.760724954

mae: 2.356724

explained variance score: 0.968531

NNet -----

Tempo de treinamento/carregamento:
0:00:11.260512
Tempo de predicao:
0:00:00.036005
r2 score: 0.931589
mse: 72.096447
rmse: 8.490962666
mae: 5.993289
explained variance score: 0.932753

GBE -----

Tempo de treinamento/carregamento:
0:00:02.238957
Tempo de predicao:
0:00:00.021074
r2 score: 0.968885
mse: 32.791239
rmse: 5.726363506
mae: 3.188219
explained variance score: 0.968893
