

Projeto 1 – Otimização de Métodos Monolíticos e Ensembles para Aprendizagem de Máquina

1. Métodos Monolíticos

1.1. Wine Dataset

1.1.1. Decision Tree

Melhores Parâmetros	
max_depth	12

Precisão	
Classe 0	0.89
Classe 1	0.89
Classe 2	0.94
Média/Total	0.91

	Matriz de Confusão		
	0	1	2
0	17	1	0
1	2	16	1
2	0	1	16

1.1.2. Naïve Bayes

Precisão: 0.96 (+/- 0.09)

1.1.3. K-Nearest Neighbors (KNN)

Melhores Parâmetros	
n_neighbors	3
weights	distance

Precisão	
Classe 0	0.84
Classe 1	0.85
Classe 2	0.87
Média/Total	0.85

	Matriz de Confusão		
	0	1	2
0	16	1	1
1	1	17	1
2	2	2	13

1.1.4. Multi-Layer Perceptron (MLP)

Melhores Parâmetros	
alpha	0.01
activation	relu
solver	lbfgs
hidden_layer_sizes	60

Precisão	
Classe 0	0.89
Classe 1	0.94
Classe 2	0.94
Média/Total	0.93

	Matriz de Confusão		
	0	1	2
0	17	1	0
1	2	16	1
2	0	0	17

1.1.5. Support Vector Machine (SVM)

Melhores Parâmetros	
kernel	rbf
C	30
gamma	0.001

Precisão	
Classe 0	1.00
Classe 1	0.86
Classe 2	0.94
Média/Total	0.93

	Matriz de Confusão		
	0	1	2
0	17	1	0
1	0	18	1
2	0	2	15

1.2. Liver Disorders Dataset

1.2.1. Decision Tree

Melhores Parâmetros	
max_depth	2

Precisão	
Mean Absolute Error	2.5240

1.2.2. K-Nearest Neighbors (KNN)

Melhores Parâmetros	
n_neighbors	20
weights	distance

Precisão	
Mean Absolute Error	2.4644

1.2.3. Multi-Layer Perceptron (MLP)

Melhores Parâmetros	
alpha	0.001
activation	tanh
solver	adam
hidden_layer_sizes	55

Precisão	
Mean Absolute Error	2.6744

1.2.4. Multi-Layer Perceptron (MLP)

Melhores Parâmetros	
kernel	rbf
C	100
gamma	0.0001

Precisão	
Mean Absolute Error	2.6875

2. Métodos de Ensembles

2.1. Wine Dataset

2.1.1. Bagging

Melhores Parâmetros	
n_estimators	4
base_estimator	Gaussian NB

Precisão	
Classe 0	1.00
Classe 1	0.91
Classe 2	1.00
Média/Total	0.97

	Matriz de Confusão		
	0	1	2
0	21	2	0
1	0	20	0
2	0	0	11

2.1.2. Boosting – AdaBoost

Melhores Parâmetros	
n_neighbors	80
base_estimator	Gaussian NB
learning_rate	1.0
algorithm	SAMME.R

Precisão	
Classe 0	1.00
Classe 1	0.91
Classe 2	1.00
Média/Total	0.97

	Matriz de Confusão		
	0	1	2
0	22	1	0
1	0	20	0
2	0	1	10

2.1.3. Random Subspaces (RSS)

Melhores Parâmetros	
max_features	0.3
n_estimators	58
base_estimator	Decision Tree
hidden_layer_sizes	60

Precisão	
Classe 0	1.00
Classe 1	0.91
Classe 2	1.00
Média/Total	0.97

	Matriz de Confusão		
	0	1	2
0	21	2	0
1	0	20	0
2	0	0	11

2.1.4. Random Forest (RF)

Melhores Parâmetros	
n_estimators	13
criterion	entropy
max_depth	20

Precisão	
Classe 0	1.00
Classe 1	0.95
Classe 2	1.00
Média/Total	0.98

	Matriz de Confusão		
	0	1	2
0	22	1	0
1	0	20	0
2	0	0	11

2.2. Liver Disorders Dataset

2.2.1. Bagging

Melhores Parâmetros	
n_estimators	93
base_estimator	Decision Tree

Precisão	
<i>Mean Absolute Error</i>	2.3750

2.2.2. Boosting - AdaBoost

Melhores Parâmetros	
n_estimators	9
base_estimator	KNN-3

Precisão	
<i>Mean Absolute Error</i>	2.4070

2.2.3. Random Subspaces (RSS)

Melhores Parâmetros	
n_estimators	45
base_estimator	KNN-3

Precisão	
<i>Mean Absolute Error</i>	2.3602

2.2.4. Random Forest (RF)

Melhores Parâmetros	
n_estimators	23
criterion	Mse
max_depth	3

Precisão	
<i>Mean Absolute Error</i>	2.3796

3. Rankings

3.1. Wine Dataset

Método	Precisão
Random Forest (RF)	0.98
Random Subspaces (RSS)	0.97
Boosting - AdaBoost	0.97
Bagging	0.97
Naive Bayes (Gaussian)	0.96
Support Vector Machine (SVM)	0.93
Multi-Layer Perceptron (MLP)	0.93
Decision Tree	0.91
K-Nearest Neighbors (KNN)	0.85

3.2. Liver Disorders Dataset

Método	MEA
Random Subspaces (RSS)	2.3602
Bagging	2.3750
Random Forest (RF)	2.3796
Boosting - AdaBoost	2.4070
K-Nearest Neighbors (KNN)	2.4644
Decision Tree	2.5240
Multi-Layer Perceptron (MLP)	2.6744
Support Vector Machine (SVM)	2.6874