

Projeto AM 2023-1

Daniel Feitosa
Felipe Alcântara
Matheus Félix
Willian Oliveira

Conjunto de Dados

- Dados extraídos do site “UCI MACHINE LEARNING”
 - Imagens descritas através de 19 atributos contínuos
- Carregamento dos dados - realizado em dois arquivos:
 - segmentation.data -----(210 amostras)
 - segmentation.test ----- (2100 amostras)
 - conjunto concatenado----- (2310 amostras)
- Dataset balanceado e sem missing values

Conjunto de Dados

- Separação dos Dados (requisitos do projeto)
 - O conjunto concatenado foi subdividido a nível de atributos
 - Foram criados 3 outros conjuntos de dados:
 - um considerando as colunas 4 a 9 (shape)
 - outro considerando as colunas 10 a 19 (rgb)
 - um terceiro considerando as colunas 4 a 19 (shape + rgb)

Questão 1

Clusterização

- Foi executado 50 iterações do algoritmo FCM
 - parâmetro de fuzzificação m foram tomados pelo conjunto $\{1.0, 1.6, 2.0\}$ - (difusividade no processo)
 - número máximo de iterações (T) definido para 100
 - nível de tolerância definido em 10^{-6} (convergência do algoritmo)

MELHORES RESULTADOS - FUNÇÃO OBJETIVO

Dataset	m	J
1	1.1	1464735.216323
	1.6	1145190.416235
	2.0	874369.638538
2	1.1	3083889.429403
	1.6	2744887.383073
	2.0	2106153.304136
3	1.1	13623060.345228
	1.6	10983111.670301
	2.0	6997221.004400

Critérios de Validação Interna

- Modified Partition Coefficient (MPC)
- Partition Entropy (PE)

CRITÉRIOS DE VALIDAÇÃO INTERNA

dataset	m	MPC	PE
shape	1.1	0.998537	0.002157
	1.6	0.985037	0.023425
	2.0	0.947474	0.097341
rgb	1.1	0.984820	0.021760
	1.6	0.854568	0.250481
	2.0	0.671102	0.599251
shape+rgb	1.1	0.979501	0.029664
	1.6	0.806279	0.336610
	2.0	0.617160	0.692518

CrITÉrios de ValidaÇão Externa

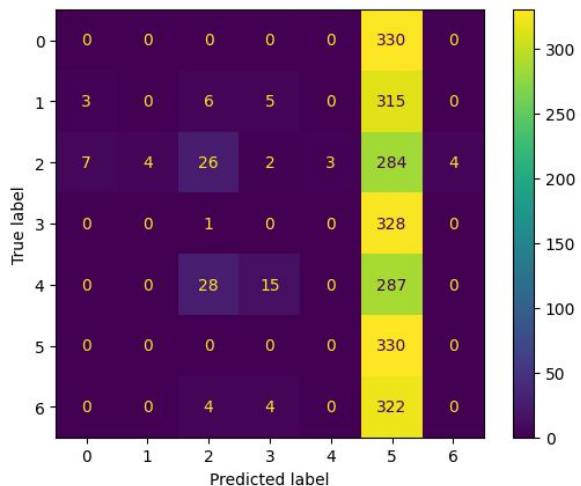
- Adjusted Rand Index (ARI)
- F-Measure (F-M)
- Comparativo com a partição *a priori*

dataset	m	ARI	F-M
shape	1.1	0.000563	0.126101
	1.6	0.000614	0.126206
	2.0	0.001604	0.126915
rgb	1.1	0.377698	0.287611
	1.6	0.390161	0.291068
	2.0	0.393355	0.291891
shape+rgb	1.1	0.365577	0.299482
	1.6	0.443022	0.317378
	2.0	0.432610	0.311162

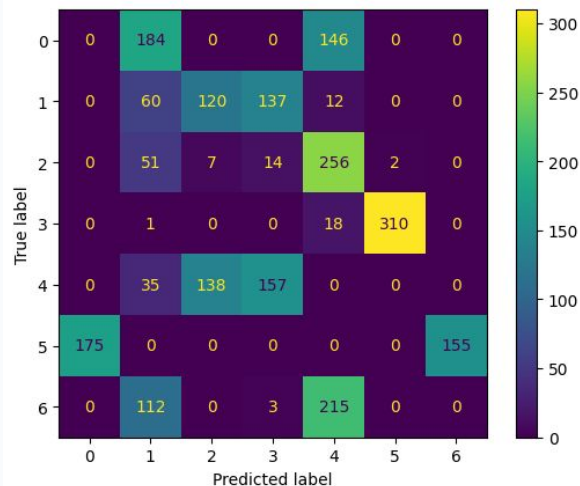
Matrizes de Confusão

- Comparativo com a partição *a priori*

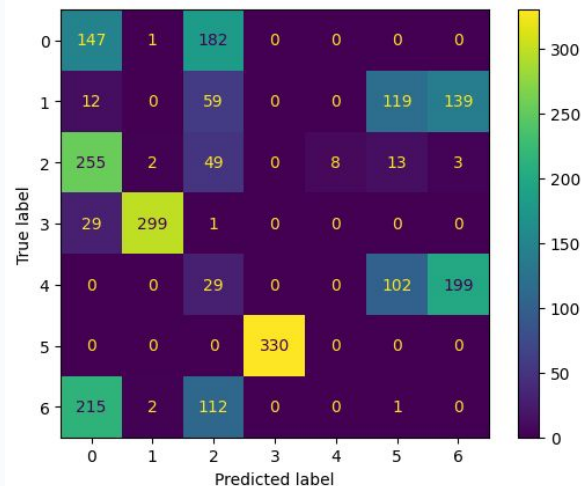
shape - 2.0



rgb - 2.0



shape+rgb - 1.6



Comparativo - Partições CRISP

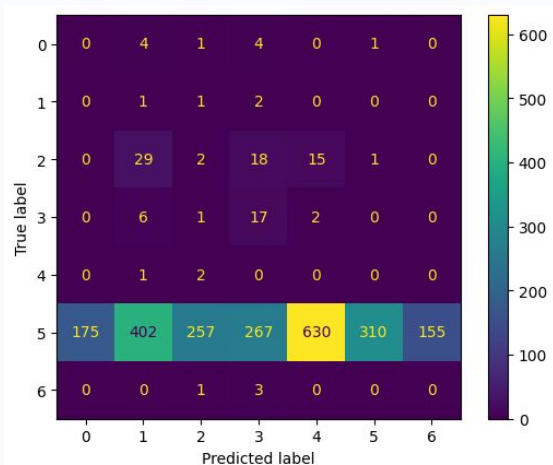
- Adjusted Rand Index (ARI)
- F-Measure (F-M)
- Maior correspondência entre rgb e shape+rgb

Partições CRISP		ARI	F-M
shape	rgb	0.000954	0.144433
shape	shape+rgb	0.006046	0.154688
rgb	shape+rgb	0.929994	0.468771

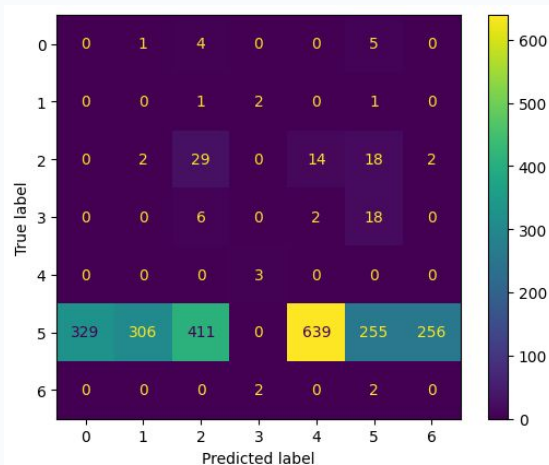
Comparativo - Partições CRISP

- Matrizes de Confusão

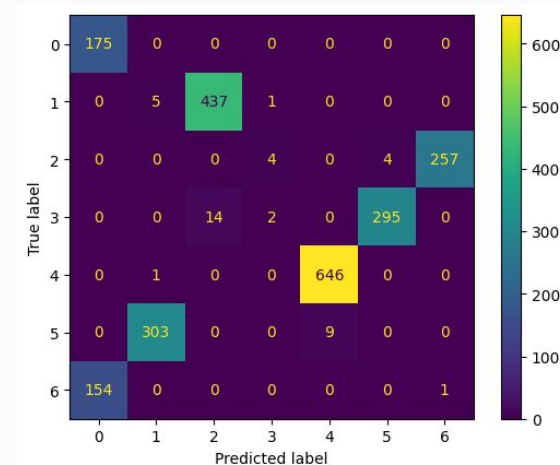
shape - rgb



shape - shape+rgb



rgb - shape+rgb



Questão 2

Tuning de Parâmetros

Foram utilizados os modelos:

1. Bayesiano Gaussiano
2. Bayesiano baseado em K-vizinhos
3. Bayesiano baseado em janela de Parzen
4. Regressão Logística

Para cada conjunto de dados, em cada modelo, foi realizado um ajuste de hiper-parâmetros através da ferramenta [Optuna](#) utilizando 20% dos dados. A cada iteração de modelo, a ferramenta realizou 10 treinos na busca.

Como heurística de seleção de melhores parâmetros, os parâmetros que resultaram no máximo F1-Score, independente do dataset, foram selecionados como parâmetros de treino completo em etapas posteriores.

Experimentação

Para realização do treinamento, cada modelo com o respectivo parâmetro ótimo foi submetido a 30 validações cruzadas estratificadas de 10 folds.

Foram coletadas as métricas: Tempo de treinamento, F1-Score, Precisão, Acurácia e Cobertura.

Experimentação

As tabelas a seguir apresentam os resultados de cada métrica nos conjuntos de teste de cada conjunto de dados. Para agregação dos folds, foi utilizada a média de cada métrica em todo lote gerado pela validação cruzada 30x10-folds.

Modelo	Tempo de Treinamento (segundos)	F1-Score	Precisão	Acurácia	Cobertura
BayesianNearestNeighbors	0,01	44,82%	46,48%	45,70%	45,70%
GaussianNB	0,00	14,97%	33,19%	21,39%	21,38%
LogisticRegression	0,06	30,25%	35,06%	33,30%	33,29%
ParzenWindowClassifier	0,01	49,72%	51,76%	50,64%	50,63%

Conjunto de dados 1

Experimentação

As tabelas a seguir apresentam os resultados de cada métrica nos conjuntos de teste de cada conjunto de dados.

Modelo	Tempo de Treinamento (segundos)	F1-Score	Precisão	Acurácia	Cobertura
BayesianNearestNeighbors	0,01	90,86%	91,22%	90,94%	90,94%
GaussianNB	0,00	61,46%	69,31%	64,82%	64,82%
LogisticRegression	0,06	78,83%	79,61%	80,03%	80,01%
ParzenWindowClassifier	0,01	91,27%	91,47%	91,29%	91,29%

Conjunto de dados 2

Experimentação

As tabelas a seguir apresentam os resultados de cada métrica nos conjuntos de teste de cada conjunto de dados.

Modelo	Tempo de Treinamento (segundos)	F1-Score	Precisão	Acurácia	Cobertura
BayesianNearestNeighbors	0,01	91,27%	91,65%	91,38%	91,38%
GaussianNB	0,00	61,61%	67,29%	66,38%	66,37%
LogisticRegression	0,07	86,41%	86,95%	86,60%	86,59%
ParzenWindowClassifier	0,01	88,90%	89,24%	88,87%	88,88%

Conjunto de dados 3

Testes de Hipótese

Teste de Friedman (3+ Grupos Pareados)

H0 (Hipótese Nula): Não há diferenças significativas entre os classificadores avaliados.

H1 (Hipótese Alternativa): Há diferenças significativas em pelo menos um classificador.

Teste de Nemenyi (*Post hoc*)

H0 (Hipótese Nula): Não há diferenças estatisticamente significativas entre o par avaliado.

H1 (Hipótese Alternativa): Existem diferenças estatisticamente significativas entre o par avaliado.

Interpretação dos Testes

D1 - Forma

Friedman Test

P-Value = 0.000 para todas as métricas avaliadas, indicando diferenças significativas entre o resultado dos modelos.

Nemenyi Post-Hoc

- Bayesian Nearest Neighbor e Janela de Parzen não apresentam diferenças estatisticamente significativas;
- Para a métrica de precisão, não há diferença estatística também entre o Gaussian NB e a Regressão Logística.

D2 - Cores

Friedman Test

P-Value = 0.000 para todas as métricas avaliadas, indicando diferenças significativas entre o resultado dos modelos.

Nemenyi Post-Hoc

- Bayesian Nearest Neighbor e Janela de Parzen não apresentam diferenças estatisticamente significativas;

D3 - Tudo

Friedman Test

P-Value = 0.000 para todas as métricas avaliadas, indicando diferenças significativas entre o resultado dos modelos.

Nemenyi Post-Hoc

- Janela de Parzen e Regressão Logística não apresentam diferenças estatisticamente significativas;

Modelos Sugeridos

D1

Formas da Imagem

- Bayesian Nearest Neighbors
- Janela de Parzen

D2

Cores da Imagem

- Bayesian Nearest Neighbors
- Janela de Parzen

D3

**Formas e Cores da
Imagem**

- Bayesian Nearest Neighbors

Ensembles

Ensemble 1

Bayesian Nearest Neighbors

C1 -> Formas
C2 -> Cores
C3 -> Ambos



Desempenho

	C1	C2	C3	E1
Acurácia (%)	43.1	90.7	<u>91.3</u>	90.7
F1 (%)	42.1	90.6	<u>91.2</u>	90.5

Ensemble 2

Janelas de Parzen

C1 -> Formas
C2 -> Cores
C3 -> Ambos



Desempenho

	C1	C2	C3	E2
Acurácia (%)	45.7	<u>86.1</u>	69.7	72.5
F1 (%)	44.9	<u>86.3</u>	70.6	72.1

Ensemble 3

Regressões Logísticas

C1 -> Formas
C2 -> Cores
C3 -> Ambos



Desempenho

	C1	C2	C3	E3
Acurácia (%)	34.4	84.2	<u>87.9</u>	86.4
F1 (%)	30.8	83.6	<u>87.1</u>	85.2

Obrigado!