

# Relatório Trabalho Computacional 2

## Inteligência Computacional Aplicada

Emanuel Ávila C. Pires - 587130

Professor: Guilherme de Alencar Barreto

### A. Atividade 1

Abrir e executar o arquivo `face_preprocessing_column.m` sem aplicação do PCA. Ou seja, comentar as linhas 56-60. Escolha as dimensões para redução das imagens na linha 37. Note que quanto maiores os valores da redução, maior será a dimensão dos vetores de atributos após a vetorização das imagens e, obviamente, maior será o tempo de treinamento/teste dos classificadores. Exemplos: [20 20] →  $20 \times 20 = 400$ , [30 30] →  $30 \times 30 = 900$ .

**Resposta:** O arquivo foi executado da forma padrão; posteriormente, fiz testes com algumas variações, entre 15x15 e 30x30, nisso foi avaliado o tempo de execução e o desempenho. Após os testes, a dimensão de 20x20 foi escolhida por ser uma das melhores opções, visando tempo de processamento.

### B. Atividade 2

Abrir e executar o arquivo `compara_todos.m` usando `Ptrain = 80`; ou seja, 80% dos vetores de atributos serão usados para treinar os classificadores. Faça também `Nr = 50` (número de repetições independentes de treino/teste). Executar o código e preencher a tabela de estatísticas de desempenho abaixo. A figura de mérito é a taxa de acerto do classificador, determinando-se suas estatísticas descritivas ao final das 50 rodadas independentes, tais como valor médio, desvio padrão, valores mínimo/máximo e mediana. Classificadores a serem implementados: Classificador Linear de Mínimos Quadrados (MQ), Perceptron Logístico (PL) e Perceptron Multicamadas com uma (MLP-1H) e duas camadas (MLP-2H)<sup>1</sup>.

**Resposta:** Os resultados para os experimentos realizados conforme a Atividade 2, seguindo todas as configurações padrão. Os valores computados estão resumidos na Tabela I. Para efeito de comparação, foram realizados testes com variações de funções de ativação (`sigmoid`, `tanh`, `relu` e `leakyrelu`) e estratégias de gradiente descendente (`sgd`, `batch-gd` e `mini-batch gd`).

**Questão 1** - O que se pode concluir sobre os desempenhos dos classificadores avaliados?

**Resposta:** O classificador de Mínimos Quadrados (MQ) apresentou o melhor desempenho geral, com uma taxa média de acerto de 95,091%, conforme indicado na Tabela I. Este modelo destacou-se por sua alta acurácia, baixa variabilidade com desvio padrão de 3,347 e tempo de execução reduzido para 0,741 segundos. Os classificadores Linear Discriminant Analysis (LDA) e Perceptron Logistic (PSLog) também mostraram bom desempenho, com médias de 94,485% e 92,121%, respectivamente, mas com tempos de execução mais elevados. Por outro lado, os classificadores MLP-1H e MLP-2H apresentaram desempenhos inferiores, com médias de 90,485% e 86,667%, respectivamente, e maior variabilidade, especialmente no MLP-2H com desvio padrão de 6,537. Os resultados para os modelos de MLP com 1 camada oculta, a melhor combinação encontrada utilizando **Sigmoid** como função de ativação e **Adam** como otimizador. MLP com 2 camadas ocultas utilizou a função de ativação **Tanh** e também **Adam** como otimizador. O classificador Naive Bayes (NB) teve o pior desempenho, com média de

<sup>1</sup>Todos os códigos podem ser encontrados no repositório do GitHub: <https://github.com/meiazero/TC2>

56,545%, indicando que não é adequado para esta tarefa. Em geral, os modelos com normalização por mudança de escala [-1,+1] (MQ, Nearest Centroid (NC)) apresentaram resultados mais consistentes em comparação aos modelos com normalização z-score ou [0,+1].

Table I: Resumo de resultados dos classificadores sem a aplicação de PCA.

Classificador	Média	Mínimo	Máximo	Mediana	Desvio Padrão	Tempo de Execução (s)
MQ†	95.091	84.848	100.0	96.970	3.347	0.741
LDA	94.485	81.818	100.0	93.939	4.316	13.912
PSLog‡	92.121	81.818	100.0	93.939	4.974	9.100
MLP-1H*	90.485	81.818	100.0	90.909	4.499	12.810
MLP-2H*	86.667	69.697	96.970	87.879	6.537	20.356
NC†	77.818	69.697	90.909	78.788	4.974	0.218
KNN*	76.606	54.545	90.909	78.788	7.648	1.886
NB‡	56.545	39.394	75.758	57.576	8.969	3.580

\* Modelo utilizou normalização z-score.

† Modelo utilizou normalização por mudança de escala [-1,+1].

‡ Modelo utilizou normalização por mudança de escala [0,+1].

**Questão 2** - Qual deles teve o melhor desempenho em relação à taxa de acerto? E em relação ao tempo?

**Resposta:** O classificador MQ também apresentou uma mediana alta (96,970%) e baixa variabilidade com desvio padrão de 3,347, indicando consistência nos resultados. Em relação ao tempo de execução, o classificador NC foi o mais rápido, com um tempo médio de 0,218 segundos, significativamente inferior aos outros modelos. Por outro lado, o MQ, apesar de ter a melhor taxa de acerto, apresentou um tempo de execução competitivo com 0,741 segundos, tornando-o uma boa escolha entre desempenho e eficiência computacional.

### C. Atividade 3

Executar o arquivo `face_preprocessing_column.m` com aplicação do PCA. Ou seja, descomentar as linhas 56-60. Faça  $q=400$  ou  $q=900$  na linha 57, a depender do redimensionamento das imagens escolhido na Atividade 1. Note que para este valor de  $q$ , a aplicação de PCA não conduz a uma redução da dimensionalidade dos vetores de atributos, mas sim promove apenas a diagonalização da matriz de covariância dos dados transformados. Em outras palavras, os atributos para o novo conjunto de dados Z são descorrelacionados entre si.

Table II: Resumo de resultados dos classificadores com a aplicação de PCA, com valor de  $q=400$ .

Classificador	Média	Mínimo	Máximo	Mediana	Desvio Padrão	Tempo de Execução (s)
MQ†	70.182	39.394	87.879	69.697	9.191	0.773
LDA†	65.697	48.485	87.879	65.152	8.537	11.141
NC‡	64.727	45.455	81.818	63.636	7.810	0.177
NB*	56.909	33.333	69.697	60.606	8.401	3.997
KNN‡	43.394	18.182	72.727	42.424	10.576	1.767
PSLog†	9.212	3.030	18.182	9.091	3.568	7.447
MLP-1H*	8.606	0.000	18.182	9.091	3.889	12.046
MLP-2H*	8.606	0.000	18.182	9.091	4.168	21.951

\* Modelo utilizou normalização z-score.

† Modelo utilizou normalização por mudança de escala [-1,+1].

‡ Modelo utilizou normalização por mudança de escala [0,+1].

Table III: Resumo de resultados dos classificadores com a aplicação de PCA, com valor de  $q=39.400$ , equivalente a 98% da informação.

Classificador	Média	Mínimo	Máximo	Mediana	Desvio Padrão	Tempo de Execução (s)
LinearMQ*	96.788	90.909	100.0	96.970	2.554	0.133
LDA†	95.394	84.848	100.0	96.970	3.422	3.283
NC†	93.273	81.818	100.0	93.939	3.981	0.121
KNN‡	87.091	63.636	96.970	87.879	6.415	1.107
MLP-1H*	86.909	75.758	96.970	87.879	5.692	2.992
MLP-2H*	86.000	69.697	100.0	87.879	6.447	5.301
PSLog‡	82.545	69.697	93.939	81.818	5.922	1.691
NB‡	78.424	63.636	96.970	78.788	8.090	2.874

\* Modelo utilizou normalização z-score.

† Modelo utilizou normalização por mudança de escala [-1,+1].

‡ Modelo utilizou normalização por mudança de escala [0,+1].

#### D. Atividade 4

Executar novamente a Atividade 2, preenchendo a tabela de desempenho abaixo.

**Questão 4** - (i) O que se pode concluir sobre os desempenhos dos classificadores avaliados? Houve alguma mudança (melhora ou piora) nos desempenhos dos classificadores avaliados em relação à tabela anterior?

**Resposta:** Com a aplicação do PCA ( $q=400$ ), conforme apresentado na Tabela II, observa-se uma piora significativa no desempenho da maioria dos classificadores em comparação com a Tabela I (sem PCA). O classificador MQ, que anteriormente alcançava uma média de 95,091%, teve sua taxa de acerto reduzida para 70,182%, com maior variabilidade, com desvio padrão de 9,191. Similarmente, o LDA caiu de 94,485% para 65,697%, e o PSLog apresentou uma queda drástica de 92,121% para 9,212%. Os modelos MLP-1H e MLP-2H também sofreram reduções expressivas, com médias de 8,606% para ambos. Os MLPs estão configurados da seguinte forma, MLP-1H com função de ativação **Sigmoid** e otimizador **Adam**; o modelo MLP-2H utiliza como função de ativação **Tanh** e otimizador **Adam**. A aplicação do PCA com  $q=400$ , que não reduz a dimensionalidade, mas apenas descorrelaciona os atributos, parece introduzir maior complexidade nos dados, afetando negativamente a capacidade de generalização dos classificadores. Apenas o NC manteve um desempenho relativamente estável, com média de 64,727%, mas ainda assim inferior ao resultado sem PCA, média 77,818%.

#### E. Atividade 5

Com base na figura gerada durante a execução da atividade anterior, que mostra a variância explicada acumulada em função do número de componentes considerado, escolher um valor para  $q$  que preserve pelo menos 98% da informação (i.e., variância) dos dados originais. O valor de  $q$  adequado pode ser escolhido visualizando o conteúdo do vetor  $VEq$ , como sendo aquela componente cujo valor é maior que 98%. Executar o arquivo `face_preprocessing_column.m` com aplicação do PCA para o valor de  $q$  escolhido. Note que para este valor de  $q$ , a aplicação de PCA conduz a uma redução da dimensionalidade dos vetores de atributos, além de promover a descorrelação dos atributos dos dados transformados.

**Questão 5** - Qual foi a dimensão de redução  $q$  escolhida, de modo a preservar 98% da informação do conjunto de dados original?

**Resposta:** Com base na análise da variância explicada acumulada, visualizada na figura gerada durante a execução da Atividade 4, o valor de  $q=39,400$  foi escolhido para preservar pelo menos 98% da informação dos dados originais. Este valor foi determinado inspecionando o vetor  $VE_q$ , onde a componente correspondente atingiu um valor superior a 98%.

#### F. Atividade 6

Com base no valor escolhido para  $q$  na Atividade 5 e no conjunto de dados gerados correspondente, treine os modelos e preencha a tabela abaixo com os resultados de desempenho com os dados de teste.

**Questão 6** - O que se pode concluir sobre os desempenhos dos classificadores avaliados com a realização da redução de dimensionalidade via PCA? Houve alguma mudança (melhora ou piora) nos desempenhos dos classificadores avaliados em relação à tabela anterior? Quais classificadores pioraram/melhoraram de desempenho com a redução de dimensionalidade via PCA?

**Resposta:** Com a aplicação do PCA utilizando  $q=39,400$ , equivalente a 98% da informação, conforme mostrado na Tabela III, os desempenhos dos classificadores melhoraram significativamente em comparação com a Tabela II ( $q=400$ ). O classificador MQ alcançou a maior média de acertos em 96,788%, superando o desempenho sem PCA (95,091%). O LDA também apresentou melhora, passando de 65,697% para 95,394%, e o NC aumentou de 64,727% para 93,273%. Os modelos MLP-1H e MLP-2H, que tiveram desempenhos muito baixos com  $q=400$  (8,606% para ambos), recuperaram-se, atingindo 86,909% e 86,000%, respectivamente. O PSLog e o NB também mostraram melhorias, embora ainda com desempenhos inferiores (82,545% e 78,424%, respectivamente). A redução de dimensionalidade com  $q=39,400$ , além de descorrelacionar os atributos, permitiu uma representação mais compacta dos dados, beneficiando a generalização dos classificadores. Os maiores ganhos foram observados nos modelos MQ, LDA e NC, enquanto o NB apresentou a menor melhora.

#### G. Atividade 7

Repita a Atividade 6, porém aplicando a transformação de Box-Cox aos dados transformados após a aplicação de PCA. Em seguida, aplique a normalização z-score aos atributos dos dados transformados.

**Questão 7** - Houve alguma mudança (melhora ou piora) nos desempenhos dos classificadores avaliados em relação aos resultados da Atividade 6? Quais classificadores pioraram/melhoraram de desempenho com a aplicação da transformação Box-Cox juntamente com PCA?

Table IV: Resumo de resultados dos classificadores com a aplicação de PCA, com valor de  $q=39,400$ , equivalente a 98% da informação implementada com Box-Cox.

Classificador	Média	Mínimo	Máximo	Mediana	Desvio Padrão	Tempo de Execução (s)
LDA	96.788	90.909	100.000	96.970	2.697	3.137
LinearMQ	95.030	87.879	100.000	96.970	3.283	0.134
NC	94.545	84.848	100.000	93.939	3.872	0.103
KNN	89.152	75.758	100.000	87.879	5.198	1.140
MLP-1H	87.455	72.727	96.970	87.879	6.030	2.978
MLP-2H	86.606	72.727	96.970	87.879	4.977	5.429
PSLog	83.758	72.727	100.000	84.848	6.947	1.673
NB	77.091	57.576	93.939	78.788	7.325	2.884

\* Modelo utilizou normalização z-score.

† Modelo utilizou normalização por mudança de escala [-1,+1].

‡ Modelo utilizou normalização por mudança de escala [0,+1].

**Resposta:** A aplicação da transformação Box-Cox combinada com PCA ( $q=39,400$ ) e normalização z-score, conforme apresentado na Tabela IV, resultou em mudanças variadas nos desempenhos dos

classificadores em relação à Tabela III (Atividade 6). O classificador LDA apresentou a maior média de acertos (96,788%), igualando o MQ da Atividade 6, mas com menor variabilidade (desvio padrão de 2,697 contra 2,554). O MQ teve uma leve piora, caindo de 96,788% para 95,030%. O NC também melhorou ligeiramente, passando de 93,273% para 94,545%. O K-Nearest Neighbors (KNN) apresentou um aumento de 87,091% para 89,152%, enquanto o MLP-1H teve uma pequena melhora de 86,909% para 87,455%. Por outro lado, o MLP-2H e o PSLog apresentaram desempenhos ligeiramente inferiores, caindo de 86% para 86,606% e de 82,545% para 83,758%, respectivamente. O NB teve uma leve piora, de 78,424% para 77,091%. Em geral, a transformação Box-Cox parece ter estabilizado os desempenhos, reduzindo a variabilidade em alguns casos (como no LDA e NC), mas não resultou em melhorias consistentes para todos os classificadores. Os maiores ganhos foram observados no LDA e NC, enquanto o MQ e NB apresentaram ligeiras perdas.

#### *H. Atividade 8*

Usar os classificadores para aplicações de controle de acesso. Use a seguinte sequência de ações para projeto dos classificadores: Imagens vetorizadas + PCA + Box-Cox + normalização z-score + Classificador. Adicione 11 imagens próprias ao conjunto de dados para atuar como “intruso”; ou seja, indivíduo ao qual não deve ser dado acesso.

**Questão 8** - Calcule os seguintes índices de desempenho para os classificadores implementados: acurácia, taxa de falsos negativos (proporção de pessoas às quais acesso foi permitido incorretamente) e taxa de falsos positivos (pessoas às quais acesso não foi permitido incorretamente), sensibilidade e precisão. Os valores devem ser médios com inclusão de medida de dispersão (e.g., desvio padrão) para 50 rodadas.

**Resposta:** Sem Resposta.