

# PROJETO 2: Reconhecimento de Comandos de Voz

Prof. Dr. Guilherme de Alencar Barreto

06/Setembro/2024

Departamento de Engenharia de Teleinformática (DETI)

Disciplina: TI0097 - Introdução ao Reconhecimento de Padrões

Universidade Federal do Ceará (UFC), Campus do Pici, Fortaleza-CE

**Objetivo** - Desenvolver as habilidades de síntese de classificadores de padrões estudados na disciplina em problemas de processamento e reconhecimento de comandos de voz. Linguagem de programação adotada: Octave/Matlab. Material fornecido: kit de projeto com arquivos de voz (em formato .WAV), código para processamento dos sinais de voz e geração do banco de amostras de treinamento e teste, biblioteca com implementações do classificador quadrático gaussiano e 4 variantes estudadas em sala de aula, além do classificador linear de mínimos quadrados (linearMQ) e o de máxima correlação (MaxCorr).

## 1 Sequência de Atividades

- **Atividade 1:** Abrir e executar o arquivo `audio_preprocessing.m` sem aplicação do PCA. Ou seja, comentar as linhas 74-78.
- **Atividade 2:** Abrir e executar o arquivo `compara_todos.m` usando  $P_{train} = 80$ ; ou seja, 80% dos vetores de atributos serão usados para treinar os classificadores. Faça também  $N_r = 100$  (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 100 rodadas independentes, tais como valor médio, desvio padrão, valores mínimo/máximo e mediana.

Classificador	Média	Mínimo	Máximo	Mediana	Desvio Padrão	Tempo de execução
Quadrático						
Variante 1						
Variante 2						
Variante 3						
Variante 4						
LinearMQ						
MaxCorr						
DMC						
1-NN						

Tabela 1: Tabela de resultados sem a aplicação de PCA.

**OBS-1:** Os classificadores MaxCorr, DMC e 1-NN não estão implementados na toolbox fornecida, devendo o aluno viabilizar de alguma forma sua implementação, podendo ser usado o código Octave/Matlab disponibilizado no SIGAA.

**OBS-2:** Antes de preencher a tabela acima, testes os classificadores MaxCorr, DMC e 1-NN, para diferentes tipos de normalização dos atributos (sem normalização, com normalização z-score e normalização por mudança de escala  $[0,+1]$  ou  $[-1,+1]$ ). Inclua na tabela apenas o resultado da versão que deu melhor resultado para as várias normalizações testadas.

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

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

**Questão 3 :** Houve problemas de inversão das matrizes de covariância? Se sim, para quais classificadores? Este problema foi contornado por alguma das variantes avaliadas? Se sim, descreva sucintamente o mecanismo usado para resolvê-lo.

- **Atividade 3:** Executar o arquivo `audio_preprocessing.m` com aplicação do PCA. Ou seja, descomentar as linhas 74-78. Faça  $q = 13$  na linha 75. 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  $\mathbf{Z}$  são descorrelacionados entre si.
- **Ação 4:** Executar novamente a **Ação 2**, preenchendo a tabela de desempenho abaixo.

Classificador	Média	Mínimo	Máximo	Mediana	Desvio Padrão	Tempo de execução
Quadrático						
Variante 1						
Variante 2						
Variante 3						
Variante 4						
LinearMQ						
MaxCorr						
DMC						
1-NN						

Tabela 2: Tabela de resultados com a aplicação de PCA sem redução de dimensionalidade.

**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? **(ii)** Note que, com a aplicação de PCA aos dados originais, a matriz de covariância dos dados transformados é diagonal. Isso faz com que o classificador quadrático e a Variante 4 sejam *teoricamente* equivalentes. Estes classificadores tiveram de fato desempenho equivalente nos experimentos relacionados?

- **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 `audio_preprocessing.m` com aplicação do PCA para o valor de  $q$  escolhido. Note

que para estes valor de  $q$ , a aplicação de PCA conduz a uma redução da dimensionalidade dos vetores de atributos, além de promover a desconexão 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?

- **Atividade 6:** Com base no valor escolhido para  $q$  na **Atividade 5** e no conjunto de dados gerados correspondente, preencha a tabela de desempenho abaixo.

Classificador	Média	Mínimo	Máximo	Mediana	Desvio Padrão	Tempo de execução
Quadrático						
Variante 1						
Variante 2						
Variante 3						
Variante 4						
LinearMQ						
MaxCorr						
DMC						
1-NN						

Tabela 3: Tabela de resultados com a aplicação de PCA com redução de dimensionalidade.

**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?

- **Atividade 7** - Repita a **Atividade 6**, porém aplicando a transformação de BOX-COX ao conjunto de dados original antes de aplicar PCA.

**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?

**BOA SORTE!!**