

# Apresentação Projeto Prático - Grupo 3

## Classificadores e Sistemas Conexionistas (CSC)

Guilherme Martins, PG52214<sup>1</sup>  
Faustino Sachimuco, PG50944 <sup>1</sup>  
José Virgílio Loureiro, PG52252 <sup>1</sup>

<sup>1</sup>Escola de Ciências  
Mestrado em Matemática e Computação  
(Perfil MLFA: Machine Learning - Fundamentos e Aplicações)  
Universidade do Minho

Maio 2024



# Table of Contents

- 1 Introdução
- 2 Arquiteturas comuns para cada tipo de modelo
  - Tuning de hiperparâmetros - Benchmarking
- 3 Dataset Grupo Ímpar - MNIST
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 4 Dataset CIFAR-10
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 5 Dataset da Classificação de Baralhos de Cartas
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 6 Análise Comparativa dos Resultados
- 7 Conclusão



# Índice de conteúdos

- 1 Introdução
- 2 Arquiteturas comuns para cada tipo de modelo
  - Tuning de hiperparâmetros - Benchmarking
- 3 Dataset Grupo Ímpar - MNIST
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 4 Dataset CIFAR-10
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 5 Dataset da Classificação de Baralhos de Cartas
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 6 Análise Comparativa dos Resultados
- 7 Conclusão



O presente trabalho prático de grupo propõe o estudo da aplicação de variados modelos de *Deep Learning* em pelo menos dois problemas na área da Visão por Computador.

Neste trabalho, apresentaremos o estudo em três datasets: 2 clássicos e 1 real de maior complexidade.

## Objetivo de todos os datasets

Previsão correta da classe da imagem.

**Metodologia:** CRISP-DM.



# Índice de conteúdos

- 1 Introdução
- 2 Arquiteturas comuns para cada tipo de modelo
  - Tuning de hiperparâmetros - Benchmarking
- 3 Dataset Grupo Ímpar - MNIST
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 4 Dataset CIFAR-10
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 5 Dataset da Classificação de Baralhos de Cartas
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 6 Análise Comparativa dos Resultados
- 7 Conclusão



# Arquiteturas comuns para cada tipo de modelo (I)

```
model_MLP = tf.keras.Sequential([  
    tf.keras.layers.Dense(neurons, activation=activation, input_shape=(784,)),  
    tf.keras.layers.Dense(10, activation='softmax')  
])
```

Figure: Arquitetura comum MLP ('input\_shape' variável dependendo do dataset)

```
model_LSTM = Sequential()  
model_LSTM.add(LSTM(neurons, input_shape=X_train.shape[1:]), activation=activation, return_sequences=True))  
model_LSTM.add(LSTM(64))  
model_LSTM.add(Dense(64))  
model_LSTM.add(Dense(10, activation='softmax'))
```

Figure: Arquitetura comum LSTM



# Arquiteturas comuns para cada tipo de modelo (II)

```
model = models.Sequential(name="CNN_model")

model.add(layers.Conv2D(neurons, (3, 3), activation=activation, padding='same', input_shape=(28, 28, 1)))
model.add(layers.Conv2D(neurons, (3, 3), activation=activation, padding='same'))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Dropout(0.25))
model.add(layers.Conv2D(64, (3, 3), activation=activation, padding='same'))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Dropout(0.25))

model.add(layers.Flatten())

model.add(layers.Dense(128, activation=activation))
model.add(layers.BatchNormalization())
model.add(layers.Dropout(0.5))
model.add(layers.Dense(10, activation='softmax'))
```

Figure: Arquitetura comum CNN (*'input\_shape'* variável dependendo do dataset)



**Repetibilidade dos resultados:** *seed* estatística 2024

**Imutabilidade** entre os modelos do mesmo dataset

- **'learning\_rate'** do otimizador Adam: 0.001, 0.01
- **número de neurónios:** 128, 512
- **função de ativação:** 'relu', 'softmax'
- **número de épocas/epochs:** (variável dependendo do modelo a ser testado)
  - **MLP:** 15, 30
  - **LSTM:** 5, 10
  - **CNN:** 10, 20





# Índice de conteúdos

- 1 Introdução
- 2 Arquiteturas comuns para cada tipo de modelo
  - Tuning de hiperparâmetros - Benchmarking
- 3 Dataset Grupo Ímpar - MNIST
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 4 Dataset CIFAR-10
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 5 Dataset da Classificação de Baralhos de Cartas
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 6 Análise Comparativa dos Resultados
- 7 Conclusão



# MNIST - Contextualização, Exploração e Preparação dados

Treino: 60.000 ; Teste: 10.000

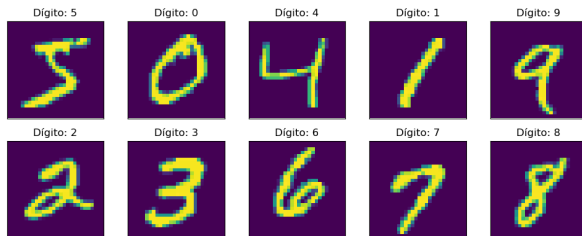


Figure: Exemplos de imagens no dataset MNIST

0	1	2	3	4	5	6	7	8	9
5923	6742	5958	6131	5842	5421	5918	6265	5851	5949

Table: Frequência dos números no conjunto de treino - MNIST



# MNIST - MLP (1/2)

Épocas	Neurónios	Learning Rate	Função Ativação	Accuracy	Tempo execução (s)
15	128	0.001	relu	0.977	68.576
30	128	0.001	relu	0.980	135.098
15	128	0.001	softmax	0.931	69.665
30	128	0.001	softmax	0.947	132.865
15	512	0.001	relu	0.982	95.748
30	512	0.001	relu	0.980	190.269
15	512	0.001	softmax	0.939	98.354
30	512	0.001	softmax	0.941	194.800
15	128	0.01	relu	0.963	68.778
30	128	0.01	relu	0.970	136.686
15	128	0.01	softmax	0.945	67.647
30	128	0.01	softmax	0.948	133.915
15	512	0.01	relu	0.965	93.847
30	512	0.01	relu	0.965	188.087
15	512	0.01	softmax	0.946	96.249
30	512	0.01	softmax	0.953	192.687

Table: Resultado do Tuning do modelo MLP e seus parâmetros



# MNIST - MLP (2/2)

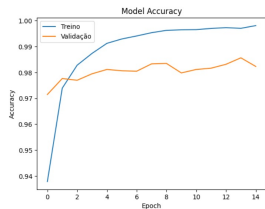


Gráfico de evolução da Accuracy ao longo das épocas - melhor modelo MLP

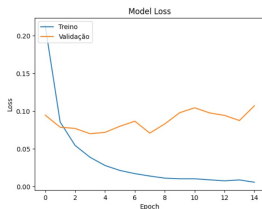
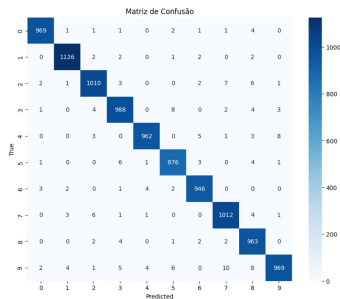


Gráfico da evolução da Loss ao longo das épocas - melhor modelo MLP



Matriz de confusão do melhor modelo MLP



# MNIST - LSTM (1/2)

Épocas	Learning Rate	Neurónios	Função Ativação	Accuracy	Tempo execução (s)
5	0.001	128	relu	0.9796	621.571
10	0.001	128	relu	0.9875	1246.888
5	0.001	128	softmax	0.9384	592.814
10	0.001	128	softmax	0.9646	891.819
5	0.001	512	relu	0.9863	3654.343
10	0.001	512	relu	0.9894	5554.391
5	0.001	512	softmax	0.9402	2550.720
10	0.001	512	softmax	0.9624	4718.027
5	0.01	128	relu	0.1028	236.172
10	0.01	128	relu	0.1711	567.382
5	0.01	128	softmax	0.9755	333.640
10	0.01	128	softmax	0.9818	668.423
5	0.01	512	relu	0.0974	1951.217
10	0.01	512	relu	0.1145	3854.036
5	0.01	512	softmax	0.9793	1886.023
10	0.01	512	softmax	0.9841	5131.517

Table: Resultado do Tuning do modelo LSTM e seus parâmetros



# MNIST - LSTM (2/2)

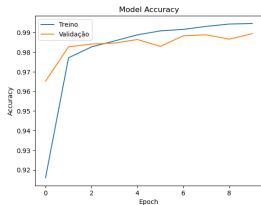


Gráfico de evolução da Accuracy ao longo das épocas - melhor modelo LSTM

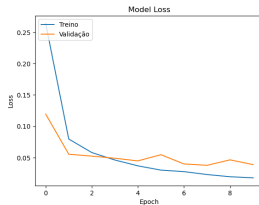
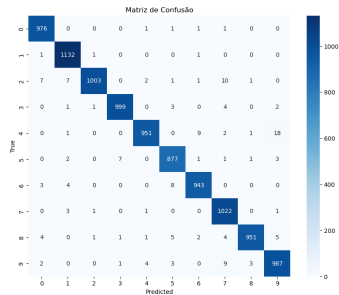


Gráfico da evolução da Loss ao longo das épocas - melhor modelo LSTM



Matriz de confusão do modelo LSTM



# MNIST - CNN (1/2)

Épocas	Neurónios	Learning Rate	Função Ativação	Accuracy	Tempo de Execução (s)
10	128	0.001	relu	0.9935	129.49
20	128	0.001	relu	0.9948	246.05
10	128	0.001	softmax	0.9901	165.03
20	128	0.001	softmax	0.9936	317.90
10	512	0.001	relu	0.9938	726.04
20	512	0.001	relu	0.9951	1396.60
10	512	0.001	softmax	0.9904	1115.99
20	512	0.001	softmax	0.9927	2208.66
10	128	0.01	relu	0.9942	127.40
20	128	0.01	relu	0.9935	247.38
10	128	0.01	softmax	0.9787	163.72
20	128	0.01	softmax	0.9880	324.57
10	512	0.01	relu	0.9924	679.33
20	512	0.01	relu	0.9936	1371.05
10	512	0.01	softmax	0.9804	1109.13
20	512	0.01	softmax	0.9891	2238.27

Table: Resultado do Tuning do modelo CNN - MNIST



# MNIST - CNN (2/2)

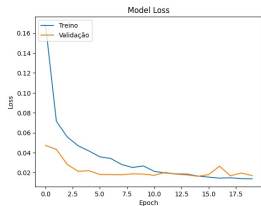


Gráfico de evolução da Accuracy ao longo das épocas - melhor modelo CNN

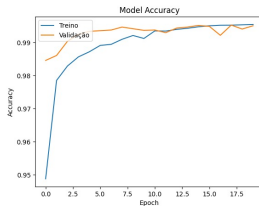
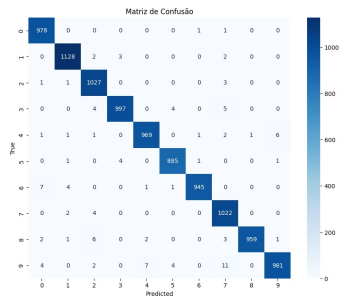


Gráfico da evolução da Loss ao longo das épocas - melhor modelo CNN



Matriz de confusão do melhor modelo CNN





# Índice de conteúdos

- 1 Introdução
- 2 Arquiteturas comuns para cada tipo de modelo
  - Tuning de hiperparâmetros - Benchmarking
- 3 Dataset Grupo Ímpar - MNIST
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 4 Dataset CIFAR-10
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 5 Dataset da Classificação de Baralhos de Cartas
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 6 Análise Comparativa dos Resultados
- 7 Conclusão



# CIFAR-10 - Contextualização Exploração Preparação dados

Treino: 50.000 ; Teste: 10.000

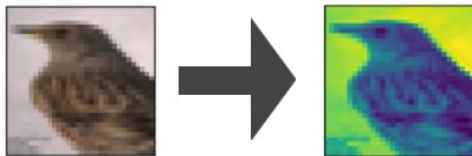


Figure: Exemplo da conversão para a escala de cinza no dataset CIFAR-10

0	1	2	3	4	5	6	7	8	9
5000	5000	5000	5000	5000	5000	5000	5000	5000	5000

Table: Frequência de cada classe no conjunto de treino - CIFAR-10



# CIFAR-10 - MLP (1/2)

Épocas	Neurónios	Learning Rate	Função Ativação	Accuracy	Tempo execução (s)
15	128	0.001	relu	0.3449	84.4642
30	128	0.001	relu	0.3744	202.6153
15	128	0.001	softmax	0.2825	79.9301
30	128	0.001	softmax	0.3186	153.6256
15	512	0.001	relu	0.3821	157.0000
30	512	0.001	relu	0.3947	322.7502
15	512	0.001	softmax	0.2888	139.5017
30	512	0.001	softmax	0.3274	325.7561
15	128	0.01	relu	0.1000	72.2893
30	128	0.01	relu	0.1594	202.6245
15	128	0.01	softmax	0.1805	82.5967
30	128	0.01	softmax	0.1791	202.5993
15	512	0.01	relu	0.1850	147.2418
30	512	0.01	relu	0.1587	322.7572
15	512	0.01	softmax	0.1842	202.7294
30	512	0.01	softmax	0.1000	301.5987

Table: Resultado do Tuning do modelo MLP - CIFAR-10



# CIFAR-10 - MLP (2/2)

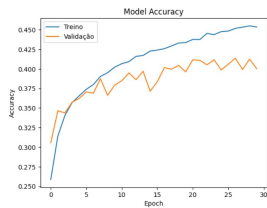


Gráfico de evolução da Accuracy por época - melhor modelo MLP

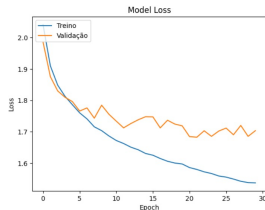
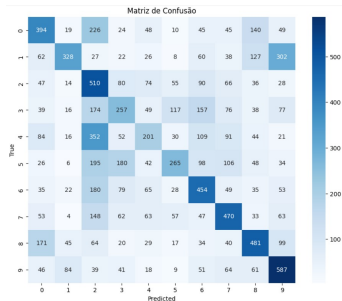


Gráfico da evolução da Loss ao longo das épocas - melhor modelo MLP



Matriz de confusão do melhor modelo MLP



# CIFAR-10 - LSTM (1/2)

Épocas	Neurónios	Learning Rate	Função Ativação	Accuracy	Tempo execução (s)
5	0.001	128	relu	0.4902	586.051
10	0.001	128	relu	0.5325	1098.715
5	0.001	128	softmax	0.3764	632.426
10	0.001	128	softmax	0.4123	741.526
5	0.001	512	relu	0.1000	2244.829
10	0.001	512	relu	0.1804	5623.663
5	0.001	512	softmax	0.3709	2208.641
10	0.001	512	softmax	0.4073	4434.035
5	0.01	128	relu	0.1372	549.627
10	0.01	128	relu	0.1000	927.422
5	0.01	128	softmax	0.4183	384.883
10	0.01	128	softmax	0.4546	1007.105
5	0.01	512	relu	0.1000	1838.224
10	0.01	512	relu	0.1000	4180.378
5	0.01	512	softmax	0.3899	2302.962
10	0.01	512	softmax	0.4758	5739.104

Table: Resultado do Tuning do modelo LSTM - CIFAR-10



# CIFAR-10 - LSTM (2/2)

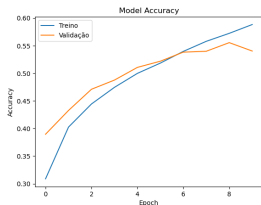


Gráfico de evolução da Accuracy ao longo das épocas - melhor modelo LSTM

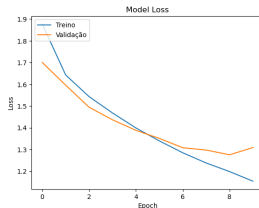
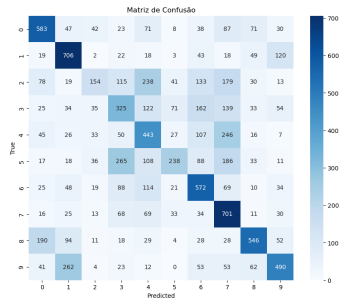


Gráfico da evolução da Loss ao longo das épocas - melhor modelo LSTM



Matriz de confusão do modelo LSTM



# CIFAR-10 - CNN (1/2)

Épocas	Neurónios	Learning Rate	Função Ativação	Accuracy	Tempo execução (s)
10	128	0.001	relu	0.7153	176.4
20	128	0.001	relu	0.7573	383.4
10	128	0.001	softmax	0.5707	263.7
20	128	0.001	softmax	0.6386	421.3
10	512	0.001	relu	0.7281	864.6
20	512	0.001	relu	0.7544	1613.4
10	512	0.001	softmax	0.5182	1047.2
20	512	0.001	softmax	0.6383	1276.4
10	128	0.01	relu	0.6840	203.6
20	128	0.01	relu	0.7085	383.4
10	128	0.01	softmax	0.5026	263.8
20	128	0.01	softmax	0.6204	443.5
10	512	0.01	relu	0.6340	932.2
20	512	0.01	relu	0.7287	1863.1
10	512	0.01	softmax	0.5476	876.8
20	512	0.01	softmax	0.7524	1823.5

Table: Resultado do Tuning do modelo CNN - CIFAR-10



# CIFAR-10 - CNN (2/2)

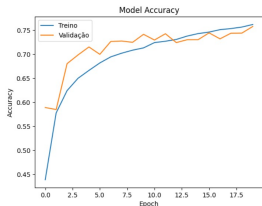


Gráfico de evolução da Accuracy por época - melhor modelo CNN

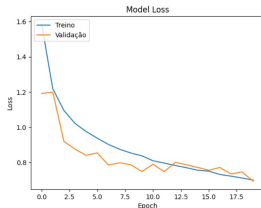
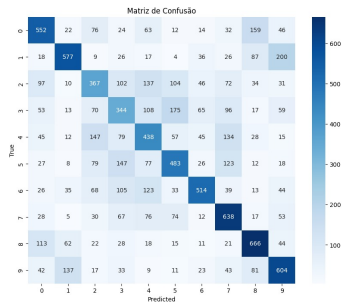


Gráfico da evolução da Loss ao longo das épocas - melhor modelo CNN



Matriz de confusão do melhor modelo CNN





# Índice de conteúdos

- 1 Introdução
- 2 Arquiteturas comuns para cada tipo de modelo
  - Tuning de hiperparâmetros - Benchmarking
- 3 Dataset Grupo Ímpar - MNIST
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 4 Dataset CIFAR-10
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 5 Dataset da Classificação de Baralhos de Cartas
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 6 Análise Comparativa dos Resultados
- 7 Conclusão



# CARDS - Contextualização Exploração Preparação dados

Treino: 7654; 53 classes = 53 tipos de cartas

Teste e Validação : 265 imagens (5 imagens para cada classe)

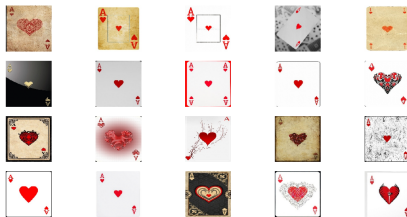


Figure: Exemplos de cartas originais a cores (Ás de Copas)



Figure: Exemplo de imagens neste Dataset depois de aplicada a conversão à escala de cinzento e redução de resolução

# CARDS - Contextualização Exploração Preparação dados

Carta	'ace of clubs'	'ace of diamonds'	'ace of hearts'	'ace of spades'	'eight of clubs'
Frequência	120	129	171	181	138
Carta	'eight of diamonds'	'eight of hearts'	'eight of spades'	'five of clubs'	'five of diamonds'
Frequência	159	152	135	150	138
Carta	'five of hearts'	'five of spades'	'four of clubs'	'four of diamonds'	'four of hearts'
Frequência	136	158	157	114	154
Carta	'four of spades'	'jack of clubs'	'jack of diamonds'	'jack of hearts'	'jack of spades'
Frequência	140	171	160	168	172
Carta	'king of spades'	'nine of clubs'	'nine of diamonds'	'nine of hearts'	'joker'
Frequência	115	128	135	125	151
Carta	'king of clubs'	king of diamonds'	'king of hearts'	'nine of spades'	'queen of clubs'
Frequência	124	129	133	154	161
Carta	'queen of diamonds'	'queen of hearts'	'queen of spades'	'seven of clubs'	'seven of diamonds'
Frequência	163	139	162	108	124
Carta	'seven of hearts'	'seven of spades'	'six of clubs'	'six of diamonds'	'six of hearts'
Frequência	143	165	152	139	127
Carta	'six of spades'	'ten of clubs'	'ten of diamonds'	'ten of hearts'	'two of spades'
Frequência	158	141	151	129	155
Carta	'ten of spades'	'three of clubs'	'three of diamonds'	'three of hearts'	
Frequência	158	126	153	113	
Carta	'three of spades'	'two of clubs'	'two of diamonds'	'two of hearts'	
Frequência	142	130	133	155	

Table: Frequência das 53 classes no dataset de treino das Cartas



# CARDS - MLP (1/2)

Épocas	Neurónios	Learning Rate	Função Ativação	Accuracy	Tempo execução (s)
15	128	0.001	relu	0.0189	12.7652
30	128	0.001	relu	0.0302	20.3248
15	128	0.001	softmax	0.0189	11.9511
30	128	0.001	softmax	0.0189	19.4607
15	512	0.001	relu	0.1509	15.4331
30	512	0.001	relu	0.1472	26.4616
15	512	0.001	softmax	0.0189	14.7316
30	512	0.001	softmax	0.0226	25.9281
15	128	0.01	relu	0.0189	11.2356
30	128	0.01	relu	0.0189	19.2754
15	128	0.01	softmax	0.0189	11.1407
30	128	0.01	softmax	0.0189	19.6244
15	512	0.01	relu	0.0189	14.7297
30	512	0.01	relu	0.0151	27.2976
15	512	0.01	softmax	0.0189	14.5116
30	512	0.01	softmax	0.0189	26.6482

Table: Resultado do Tuning do modelo MLP - Cartas



# CARDS - MLP (2/2)

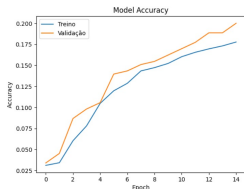


Gráfico de evolução da Accuracy por época - melhor modelo MLP

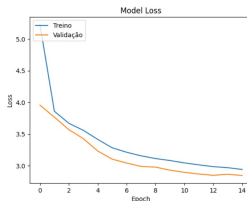
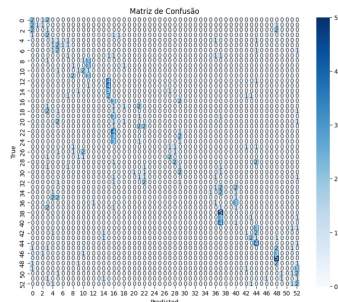


Gráfico da evolução da Loss ao longo das épocas - melhor modelo MLP



Matriz de confusão do melhor modelo MLP



# CARDS - LSTM (1/2)

Épocas	Neurónios	Learning Rate	Função Ativação	Accuracy	Tempo execução (s)
5	128	0.001	relu	0.1925	80.4472
10	128	0.001	relu	0.0189	154.7242
5	128	0.001	softmax	0.1132	77.2006
10	128	0.001	softmax	0.2226	162.4928
5	512	0.001	relu	0.0189	1652.5369
10	512	0.001	relu	0.0189	3118.3717
5	512	0.001	softmax	0.0377	1973.5203
10	512	0.001	softmax	0.2000	2388.8062
5	128	0.01	relu	NaN	588.5652
10	128	0.01	relu	NaN	1234.8789
5	128	0.01	softmax	0.1396	278.5909
10	128	0.01	softmax	0.0226	509.2516
5	512	0.01	relu	NaN	1453.0204
10	512	0.01	relu	NaN	2541.1126
5	512	0.01	softmax	0.0189	1203.2225
10	512	0.01	softmax	0.0868	2136.2038

Table: Resultado do Tuning do modelo LSTM - Cartas



# CARDS - LSTM (2/2)

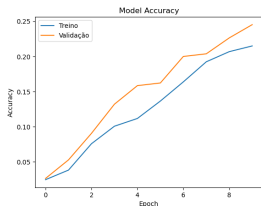


Gráfico de evolução da Accuracy ao longo das épocas - melhor modelo LSTM

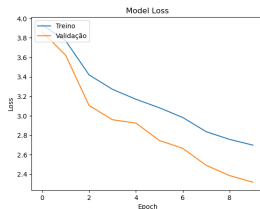
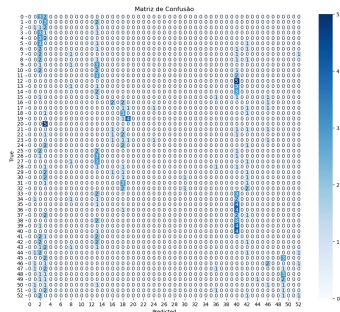


Gráfico da evolução da Loss ao longo das épocas - melhor modelo LSTM



Matriz de confusão do modelo LSTM



# CARDS - CNN (1/2)

Épocas	Neurónios	Learning Rate	Função Ativação	Accuracy	Tempo execução (s)
10	128	0.001	relu	0.842	119.575
20	128	0.001	relu	0.819	192.722
10	128	0.001	softmax	0.325	134.836
20	128	0.001	softmax	0.592	260.210
10	512	0.001	relu	0.755	727.131
20	512	0.001	relu	0.796	1239.635
10	512	0.001	softmax	0.019	961.231
20	512	0.001	softmax	0.019	1853.593
10	128	0.01	relu	0.664	101.420
20	128	0.01	relu	0.796	191.640
10	128	0.01	softmax	0.019	145.117
20	128	0.01	softmax	0.019	259.693
10	512	0.01	relu	0.755	601.010
20	512	0.01	relu	0.842	1177.517
10	512	0.01	softmax	0.019	962.368
20	512	0.01	softmax	0.019	1848.525

Table: Resultado do Tuning do modelo CNN - Cartas





# CARDS - CNN (2/2)

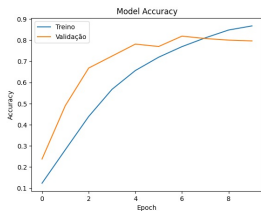


Gráfico de evolução da Accuracy ao longo das épocas - melhor modelo CNN

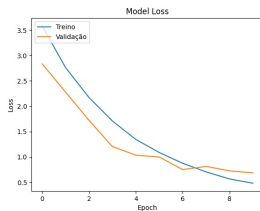
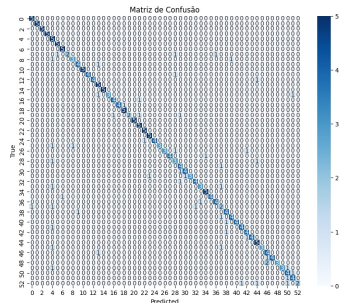


Gráfico da evolução da Loss ao longo das épocas - melhor modelo CNN



Matriz de confusão do melhor modelo CNN



# Índice de conteúdos

- 1 Introdução
- 2 Arquiteturas comuns para cada tipo de modelo
  - Tuning de hiperparâmetros - Benchmarking
- 3 Dataset Grupo Ímpar - MNIST
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 4 Dataset CIFAR-10
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 5 Dataset da Classificação de Baralhos de Cartas
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 6 Análise Comparativa dos Resultados
- 7 Conclusão



- MLP:
  - MNIST: construção de modelo com precisão elevada
  - CIFAR-10 e CARDS: accuracy relativamente baixa
- LSTM:
  - Tempo de execução muito elevado (muito pouco eficiente)
  - Inadequado para os datasets testados
- CNN:
  - Resultados bastante positivos em todos os datasets
  - Tempo de execução ligeiramente mais baixo que o das LSTMs
  - Modelo mais adequado para análise das imagens



# Índice de conteúdos

- 1 Introdução
- 2 Arquiteturas comuns para cada tipo de modelo
  - Tuning de hiperparâmetros - Benchmarking
- 3 Dataset Grupo Ímpar - MNIST
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 4 Dataset CIFAR-10
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 5 Dataset da Classificação de Baralhos de Cartas
  - Contextualização, Exploração e Preparação dos dados
  - MLP, LSTM e CNN
- 6 Análise Comparativa dos Resultados
- 7 Conclusão



- Limitações no uso de recursos computacionais (Google Colab)
  - Possibilidade melhor tuning de modelos
  - Uso das imagens na sua versão original
- **Benchmarking**: Importância da seleção adequada de modelos para os datasets propostos (Visão por Computador)
- CNN **vs** MLP e LSTM



Muito Obrigado

Muito Obrigado

**Figure:** Problemas da Inteligência Artificial (prompt: Agradecimento em várias línguas)

