

UNIVERSIDADE ESTADUAL PAULISTA JÚLIO DE MESQUITA FILHO – UNESP

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

COMPUTAÇÃO INSPIRADA PELA NATUREZA – TRABALHO 2

MAYARA EID ORLANDINI – CCO220388

São José do Rio Preto - SP

2023

Sumário

1. Questão 1	7
1.1 Enunciado	7
1.2 Linguagem e IDE	7
1.3 Descrição lógica do algoritmo	8
1.4 Análise do algoritmo	11
1.4.1 Influência da quantidade de neurônios da camada oculta sob o percentual de erros e acertos da classificação do <i>Multi Layer Perceptron</i> (MLP)	11
1.4.2 Influência da taxa de aprendizado sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do <i>Multi Layer Perceptron</i> (MLP)	13
1.4.3 Influência inicialização dos pesos sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do <i>Multi Layer Perceptron</i> (MLP)	19
1.4.3 Análise final do <i>Multi Layer Perceptron</i> (MLP)	25
2. Questão 2	29
2.1 Enunciado	29
2.2 Linguagem e IDE	29
2.3 Descrição lógica do algoritmo	29
2.4 Análise do algoritmo	30
2.4.1 Influência da quantidade de neurônios da camada oculta sob o percentual de erros e acertos da classificação do <i>Multi Layer Perceptron</i> (MLP)	30
2.4.2 Influência da taxa de aprendizado sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do <i>Multi Layer Perceptron</i> (MLP)	32
2.4.3 Influência inicialização dos pesos sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do <i>Multi Layer Perceptron</i> (MLP)	38
2.4.3 Análise final do <i>Multi Layer Perceptron</i> (MLP)	44
3. Questão 3	48
3.1 Enunciado	48
3.2 Linguagem e IDE	48
3.3. Banco de dados	48

3.4 Descrição lógica do algoritmo	48
3.5 Análise do algoritmo	49
3.5.1 Influência da quantidade de neurônios da camada oculta sob o percentual de erros e acertos da classificação do <i>Multi Layer Perceptron</i> (MLP).....	49
3.5.2 Influência da taxa de aprendizado sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do <i>Multi Layer Perceptron</i> (MLP)	51
3.5.3 Influência inicialização dos pesos sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do <i>Multi Layer Perceptron</i> (MLP)	56
3.5.3 Análise final do <i>Multi Layer Perceptron</i> (MLP).....	62
4. Link para acesso dos algoritmos.....	66
4.1 Link de acesso ao exercício 1	66
4.2 Link de acesso ao exercício 2	66
4.3 Link de acesso ao exercício 3	66
5. Referências bibliográficas	67

Lista de Figuras

Figura 1: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.	12
Figura 2: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.	14
Figura 3: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.	15
Figura 4: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.	16
Figura 5: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.	17
Figura 6: Influência da taxa de aprendizado do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.	18
Figura 7: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.	20
Figura 8: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.	21
Figura 9: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.	22
Figura 10: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.	23
Figura 11: Influência do processo de inicialização dos pesos do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.	24
Figura 12: Gráficos de acurácia e MSE e matrizes de confusão da primeira execução do algoritmo com as configurações finais dos parâmetros do MLP.	25
Figura 13: Gráficos de acurácia e MSE e matrizes de confusão da segunda execução do algoritmo com as configurações finais dos parâmetros do MLP.	26
Figura 14: Gráficos de acurácia e MSE e matrizes de confusão da terceira execução do algoritmo com as configurações finais dos parâmetros do MLP.	26
Figura 15: Gráficos de acurácia e MSE e matrizes de confusão da quarta execução do algoritmo com as configurações finais dos parâmetros do MLP.	27
Figura 16: Gráficos de acurácia e MSE e matrizes de confusão da quinta execução do algoritmo com as configurações finais dos parâmetros do MLP.	27

Figura 17: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.	31
Figura 18: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.	33
Figura 19: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.	34
Figura 20: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.	35
Figura 21: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.	36
Figura 22: Influência da taxa de aprendizado do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.	37
Figura 23: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.	39
Figura 24: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.....	40
Figura 25: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.	41
Figura 26: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.	42
Figura 27: Influência do processo de inicialização dos pesos do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.	43
Figura 28: Gráficos de acurácia e MSE e matrizes de confusão da primeira execução do algoritmo com as configurações finais dos parâmetros do MLP.	44
Figura 29: Gráficos de acurácia e MSE e matrizes de confusão da segunda execução do algoritmo com as configurações finais dos parâmetros do MLP.	45
Figura 30: : Gráficos de acurácia e MSE e matrizes de confusão da terceira execução do algoritmo com as configurações finais dos parâmetros do MLP.	45
Figura 31: Gráficos de acurácia e MSE e matrizes de confusão da quarta execução do algoritmo com as configurações finais dos parâmetros do MLP.	46
Figura 32: Gráficos de acurácia e MSE e matrizes de confusão da quinta execução do algoritmo com as configurações finais dos parâmetros do MLP.	46
Figura 33: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.	50

Figura 34: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.	51
Figura 35: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.	52
Figura 36: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.	53
Figura 37: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.	54
Figura 38: Influência da taxa de aprendizado do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.	55
Figura 39: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.	57
Figura 40: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.....	58
Figura 41: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.	59
Figura 42: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.	60
Figura 43: Influência do processo de inicialização dos pesos do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.	61
Figura 44: Gráficos de acurácia e MSE e matrizes de confusão da primeira execução do algoritmo com as configurações finais dos parâmetros do MLP.	62
Figura 45: Gráficos de acurácia e MSE e matrizes de confusão da segunda execução do algoritmo com as configurações finais dos parâmetros do MLP.	63
Figura 46: Gráficos de acurácia e MSE e matrizes de confusão da terceira execução do algoritmo com as configurações finais dos parâmetros do MLP.	63
Figura 47: Gráficos de acurácia e MSE e matrizes de confusão da quarta execução do algoritmo com as configurações finais dos parâmetros do MLP.	64
Figura 48: Gráficos de acurácia e MSE e matrizes de confusão da quinta execução do algoritmo com as configurações finais dos parâmetros do MLP.	64

Lista de Tabelas

Tabela 1: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.	11
Tabela 2: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.	13
Tabela 3: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.	19
Tabela 4: Configuração final dos parâmetros do MLP.	25
Tabela 5: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.	30
Tabela 6: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.	32
Tabela 7: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.	38
Tabela 8: Configuração final dos parâmetros do MLP.	44
Tabela 9: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.	49
Tabela 10: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.	51
Tabela 11: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.	56
Tabela 12: Configuração final dos parâmetros do MLP.	62

1. Questão 1

Nessa seção, serão apresentados os tópicos relacionados a resolução da questão 1 proposta no trabalho 2 da matéria “Computação Inspirada pela Natureza”, referente ao Programa de Pós-graduação em Ciência da Computação da UNESP.

1.1 Enunciado

Faça o download do conjunto de dados Iris do UCI Machine Learning Repository: <http://archive.ics.uci.edu/ml/datasets/Iris>

Este conjunto de dados tem 150 itens, sendo 50 de cada uma de três espécies da flor Iris, mostrada na foto ao lado.

As três espécies são I. setosa, I. virginica e I. versicolor

Para cada item foram obtidos quatro atributos: comprimento e largura da sépala e da pétala.

Você deve treinar uma rede Perceptron para reconhecer as três diferentes classes.

Divida aleatoriamente os exemplos em subconjuntos de treinamento, validação e testes (por exemplo: 70%/15%/15%).

Teste diferentes pesos de inicialização e taxas de aprendizado para encontrar os melhores valores. Lembre-se de repetir o experimento diversas vezes com cada configuração.

Entregue o relatório documentando seus passos, não se esqueça de mostrar os parâmetros iniciais e finais das redes ajustadas, a taxa de aprendizado, as taxas de acerto/erro. Inclua também um gráfico do erro médio quadrático mostrando a convergência do algoritmo, e outros gráficos e/ou tabelas que julgar relevantes.

Mostre uma tabela com as matrizes de confusão para cada subconjunto separadamente: treinamento, validação e teste. O que você pode concluir desses dados?

1.2 Linguagem e IDE

O código referente a questão 1 foi desenvolvido em *Python* [1], com suporte do *Google Colaboratory* [2].

1.3 Descrição lógica do algoritmo

Para o funcionamento do algoritmo, foi fundamental a implementação das seguintes bibliotecas: *numpy* [3], *pandas* [4], *matplotlib* [5], *sklearn* [6], *prettytable* [7], *seaborn* [8]. A biblioteca *numpy* [3], neste código, teve a função de permitir a utilização de recursos matemáticos, como o uso de funções exponenciais e a manipulação de matrizes. A *pandas* [4], por sua vez, foi utilizada para realizar a leitura do banco de dados. A biblioteca *matplotlib* [5] foi usada com o intuito de elaborar gráficos. A *sklearn* [6], ficou responsável por segmentar o banco de dados entre treinamento, validação e teste e, além disso, gerar as matrizes de confusão. A *prettytable* [7] foi a responsável por formatar a impressão dos resultados para análise. A *seaborn* [8], por fim, foi a responsável por gerar os gráficos das matrizes de confusão.

Seguinte a isso, iniciou-se a definição de funções responsáveis por comporem o algoritmo. A primeira função a ser definida, foi a sigmoide, responsável por realizar o processo de ativação da camada oculta da rede neural [9].

Partiu-se, então, para a inicialização dos pesos da camada oculta e da camada de saída. Nessa etapa foram testadas a inicialização de Xavier [10] e a inicialização aleatória simples. Os bias de ambas as camadas foram inicializados com zero.

Como sequência, definiu-se a função responsável por realizar a propagação direta do *perceptron* (*forward propagation*) [9]. Multiplicando as características dos indivíduos de entrada com os pesos da camada oculta, calculou-se o resultado da entrada desta. Foi acrescido a tal resultado o valor do bias referente a camada. Inserindo este resultado na função sigmoide, obteve-se a saída da camada oculta. A fim de calcular a entrada da camada de saída multiplicou-se a saída da camada oculta pelos pesos da camada de saída e, posterior a tal operação, somou-se o bias ao resultado. A saída da operação descrita anteriormente foi aplicada à função *softmax*.

Posteriormente, implementou-se a função de *backpropagation* [9]. A primeira etapa consistiu em calcular o delta da camada de saída da rede neural, encontrando a diferença entre a saída real e a saída esperada e, adicionalmente a isso, normalizando o resultado com base na quantidade de dados de entrada. Na sequência, calculou-se o delta da camada oculta, multiplicando o delta da camada de saída pela matriz transposta do peso desta mesma camada. Este resultado foi multiplicado pela saída da camada oculta e pela subtração da saída da camada de oculta em relação ao número 1. Tais valores de delta foram utilizados para atualizar os pesos e bias da rede neural. Em relação ao peso da camada de saída, subtraiu-se a ele, o resultado da multiplicação da taxa de aprendizado pelo produto da matriz transposta da saída da camada

oculta com o delta da camada de saída. O bias, por sua vez, sofreu uma subtração do produto da taxa de aprendizado pela soma do delta da camada de saída. O peso da camada oculta foi atualizado realizando, em seu próprio valor, uma subtração da multiplicação da taxa de aprendizado pelo produto da matriz transposta da camada de entrada com o delta da camada oculta. O bias passou pelo mesmo processo descrito anteriormente, porém, desta vez, considerando o delta da camada oculta.

A próxima função a ser definida foi a responsável por calcular o erro quadrático médio (MSE) do *perceptron* [9]. Para isso, foi necessário transformar a variável de categorização, ou seja, de saídas, em uma matriz binária. Na sequência, ativou-se a função *forward propagation* a fim de obter a saída calculada da rede neural. Tal saída foi comparada com a matriz binária e, então, o MSE foi calculado.

A última função é tida como a previsão [9]. Nela, é ativada, novamente, a função *forward propagation* à título de obter a saída da rede neural. Com auxílio da função *argmax*, encontra-se o índice do elemento com maior valor, correspondendo a escolha da classe com maior probabilidade de ocorrência para cada entrada.

Estruturadas as funções, parte-se para a implementação temporal do algoritmo. Primeiro, é realizada a leitura do banco de dados, armazenado em um arquivo *Microsoft Excel* [11]. Desse banco de dados, são extraídas as quatro características de cada indivíduo, bem como as saídas esperadas para cada um deles. Os dados são então normalizados utilizando a técnica *zscore*, conforme proposto neste presente trabalho, que consiste em calcular a média e o desvio padrão de cada uma dessas características e, na sequência, subtrair de cada característica o valor dessa média e dividir o resultado pelo desvio padrão anteriormente calculado. Com os dados normalizados, estes são segmentados em subgrupos de treinamento, validação e teste, utilizando o recurso *train_test_split* da biblioteca *sklearn* [6]. Para o grupo de treinamento, foram destinados 108 dados. Para o grupo de validação e teste, destinaram-se 27 e 15 dados, respectivamente.

O passo seguinte, consistiu em definir as configurações e parâmetros do *perceptron*: quantidade de neurônios da camada de entrada, quantidade de neurônios da camada oculta, quantidade de neurônios da camada de saída, taxa de aprendizado e quantidade de épocas.

Na sequência, inicializou-se os pesos da rede neural e, dentro de um loop iterativo, as funções *forward propagation* e *backpropagation* foram chamadas. Com base nesse

procedimento, as previsões, acurácias e MSEs dos três conjuntos de dados – treinamento, validação e teste – foram calculadas. Por fim, os gráficos e matrizes de confusão foram plotados.

Ao fim do processo, com a rede treinada e validada, espera-se que com os dados de teste, ocorra uma classificação correta e satisfatória, com elevada acurácia, baixo MSE e com uma matriz de confusão suficientemente adequada.

1.4 Análise do algoritmo

A fim de realizar uma análise relacionada ao algoritmo desenvolvido, o código foi executado algumas vezes e, durante esse processo, alguns parâmetros foram alterados e resultados foram coletados e interpretados.

1.4.1 Influência da quantidade de neurônios da camada oculta sob o percentual de erros e acertos da classificação do *Multi Layer Perceptron* (MLP)

A primeira análise consistiu em entender a influência da variação de neurônios presentes na camada oculta do *Multi Layer Perceptron* (MLP) em relação ao percentual de erros e acertos do algoritmo.

Para isso, executou-se o código 5 vezes para cada uma das configurações observadas na Tabela 1 e, para cada um desses arranjos, foi obtida a média percentual de erros dos três subconjuntos avaliados.

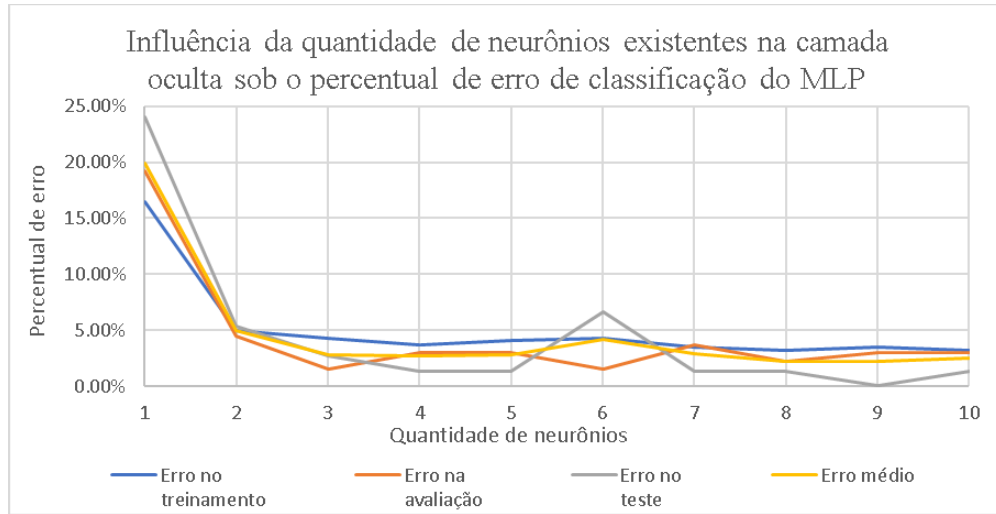
Tabela 1: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de inicialização dos pesos	Erro no treinamento	Erro na avaliação	Erro no teste	Erro médio
1000	0.1	1	Xavier	16.48%	19.26%	24.00%	19.91%
1000	0.1	2	Xavier	5.00%	4.44%	5.33%	4.93%
1000	0.1	3	Xavier	4.26%	1.48%	2.67%	2.80%
1000	0.1	4	Xavier	3.70%	2.96%	1.33%	2.67%
1000	0.1	5	Xavier	4.07%	2.96%	1.33%	2.79%
1000	0.1	6	Xavier	4.26%	1.48%	6.67%	4.14%
1000	0.1	7	Xavier	3.52%	3.70%	1.33%	2.85%
1000	0.1	8	Xavier	3.15%	2.22%	1.33%	2.23%
1000	0.1	9	Xavier	3.52%	2.96%	0.00%	2.16%
1000	0.1	10	Xavier	3.15%	2.96%	1.33%	2.48%

Fonte: Elaborado pelo autor (2023).

O gráfico obtido com os dados extraídos da Tabela 1 pode ser observado na Figura 1.

Figura 1: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.



Fonte: Elaborado pelo autor (2023).

É possível extrair da Tabela 1 e Figura 1, que a partir de 3 neurônios na camada oculta, ocorre uma tendência de linearização no que se relaciona ao percentual de erro do algoritmo em todos os grupos analisados. Nesse sentido, objetivando uma minimização da quantidade de classificações erradas no ao fim do processo de treinamento, validação e teste, optou-se por construir uma rede neural com 4 neurônios na camada oculta, visando equilibrar resultado com otimização computacional.

1.4.2 Influência da taxa de aprendizado sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do *Multi Layer Perceptron* (MLP)

A segunda análise teve como objetivo entender a influência da taxa de aprendizado do algoritmo em relação ao percentual de erros e acertos deste, bem como acurácia e MSE.

Para isso, executou-se o código 5 vezes para cada uma das configurações observadas na Tabela 2 e, para cada um desses arranjos, foi obtida a média percentual de erros dos três subconjuntos avaliados.

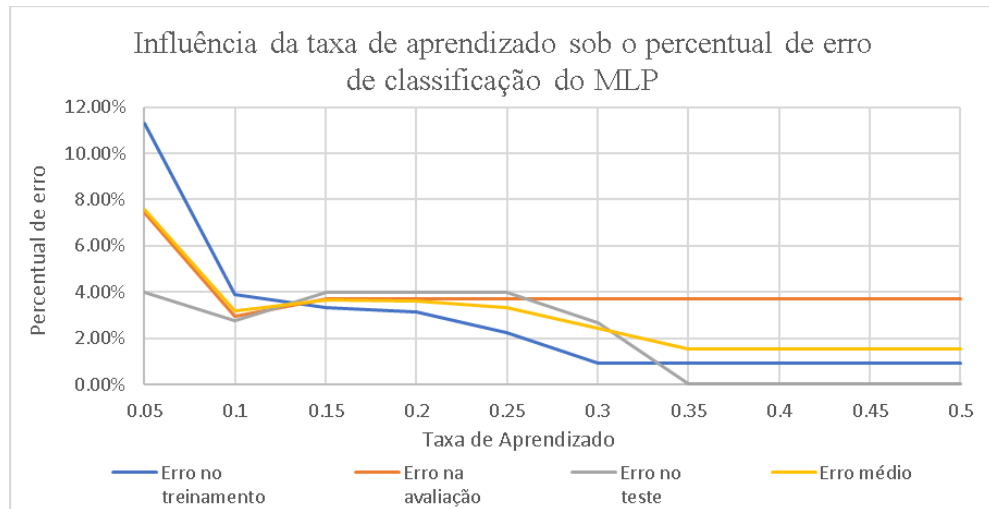
Tabela 2: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de inicialização dos pesos	Erro no treinamento	Erro na avaliação	Erro no teste	Erro médio
1000	0.05	4	Xavier	11.30%	7.41%	4.00%	7.57%
1000	0.1	4	Xavier	3.89%	2.96%	2.76%	3.20%
1000	0.15	4	Xavier	3.33%	3.70%	4.00%	3.68%
1000	0.2	4	Xavier	3.15%	3.70%	4.00%	3.62%
1000	0.25	4	Xavier	2.22%	3.70%	4.00%	3.31%
1000	0.3	4	Xavier	0.93%	3.70%	2.67%	2.43%
1000	0.35	4	Xavier	0.93%	3.70%	0.00%	1.54%
1000	0.4	4	Xavier	0.93%	3.70%	0.00%	1.54%
1000	0.45	4	Xavier	0.93%	3.70%	0.00%	1.54%
1000	0.5	4	Xavier	0.93%	3.70%	0.00%	1.54%

Fonte: Elaborado pelo autor (2023).

O gráfico obtido com os dados extraídos da Tabela 2 pode ser observado na Figura 2.

Figura 2: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.

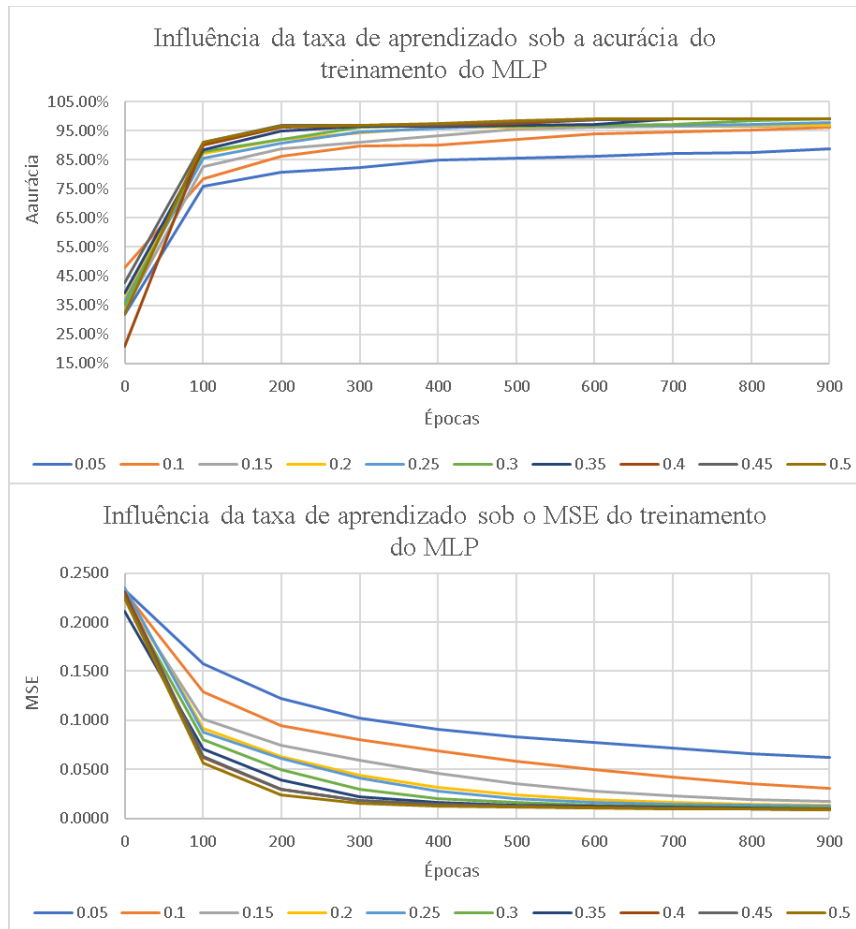


Fonte: Elaborado pelo autor (2023).

Observando a Figura 2, entende-se que quando a taxa de aprendizado do algoritmo atinge um valor igual ou maior do que 0.35, considerando as 1000 iterações propostas, o percentual de erro tende a se minimizar e a se estabilizar.

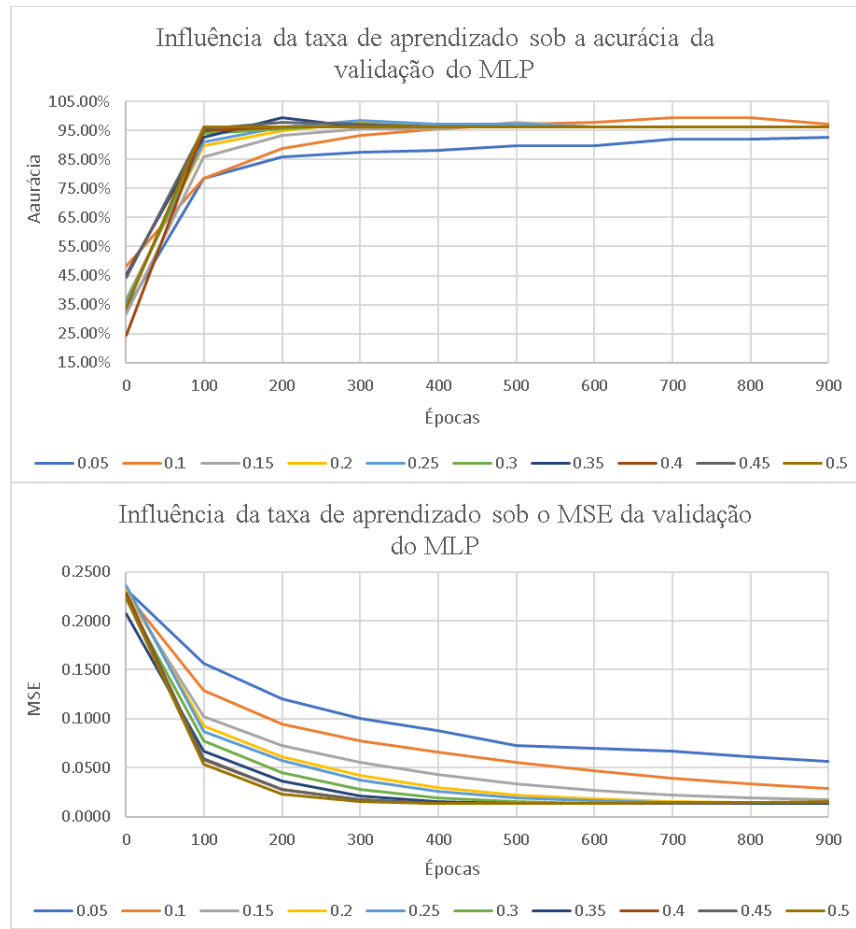
Nesse cenário, partiu-se para a avaliação da acurácia e MSE de cada uma das taxas de aprendizado mostradas na Tabela 2.

Figura 3: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.



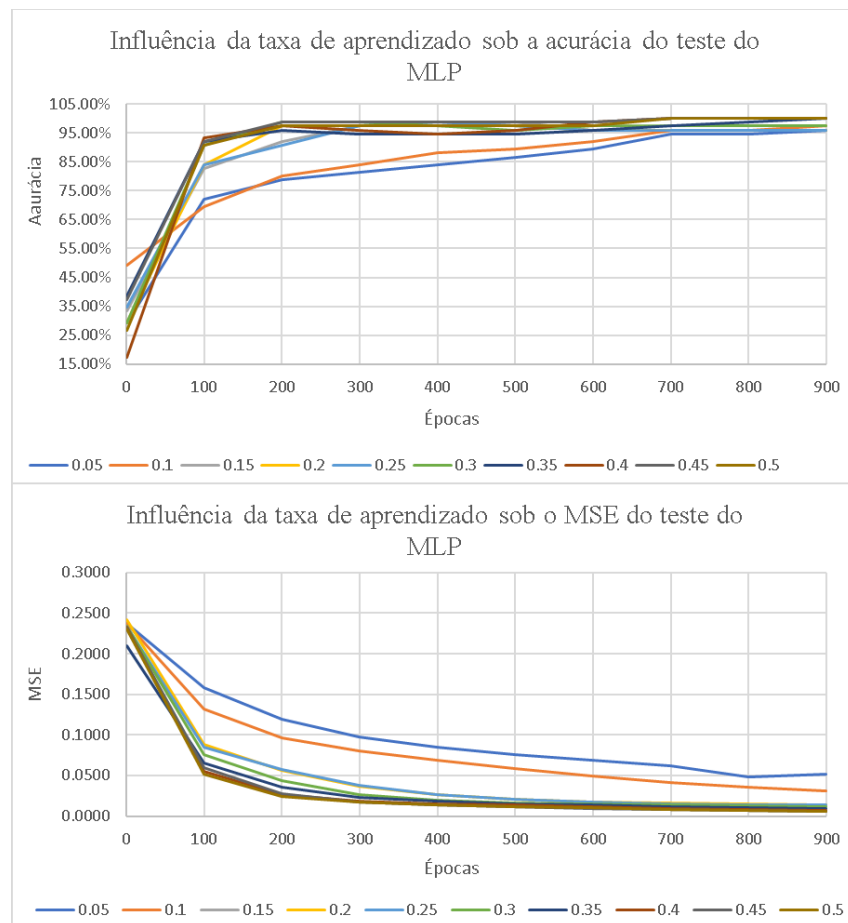
Fonte: Elaborado pelo autor (2023).

Figura 4: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.



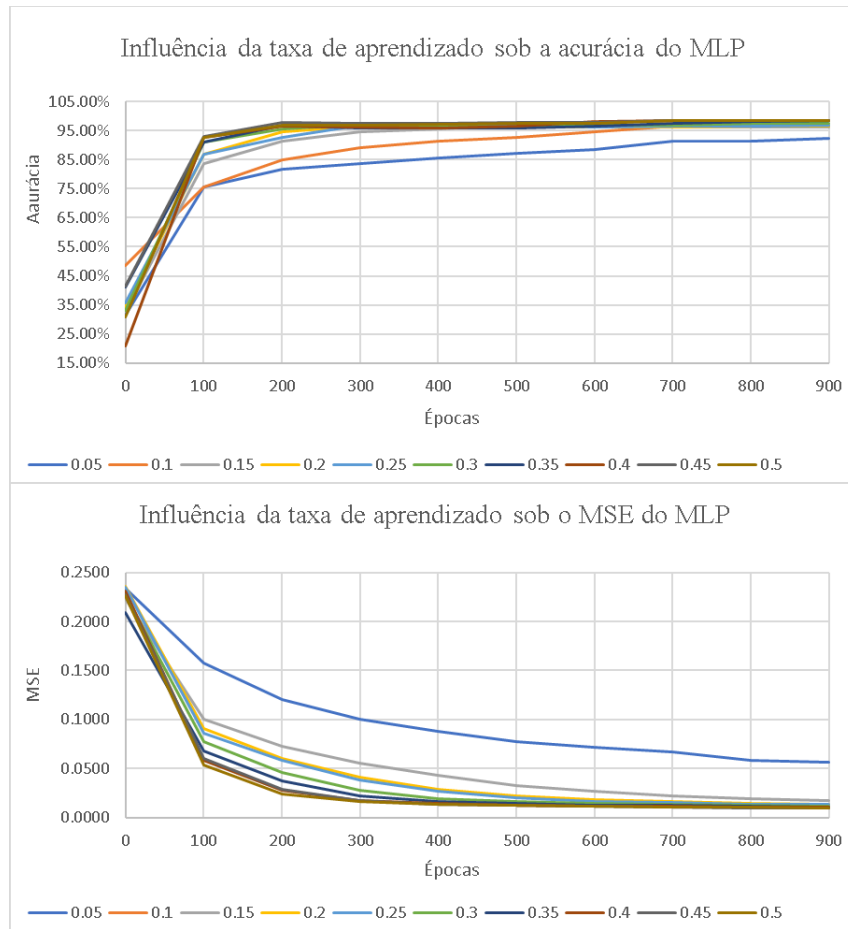
Fonte: Elaborado pelo autor (2023).

Figura 5: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

Figura 6: Influência da taxa de aprendizado do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

É possível interpretar que, conforme a taxa de aprendizado aumenta, dentro do espectro definido, existe uma tendência de minimização do MSE bem como aumento da acurácia. De modo a associar otimização com ótimos resultados, optou-se pela taxa de aprendizado de 0,35.

A escolha se valida uma vez em que, próximo a 500 iterações, uma certa estabilidade é garantida, atingindo resultados satisfatórios das métricas analisadas. Em paralelo a isso, o valor facilita o processo de convergência, não tornando o algoritmo excessivamente custoso computacionalmente, assim como reduz a probabilidade de que o resultado se prenda à ótimos locais.

1.4.3 Influência inicialização dos pesos sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do *Multi Layer Perceptron* (MLP)

A terceira análise buscou entender a influência da inicialização dos pesos em relação ao percentual de erros e acertos do código, bem como a acurácia e MSE.

Para isso, executou-se o código 5 vezes para cada uma das configurações observadas na Tabela 3 e, para cada um desses arranjos, foi obtida a média percentual de erros dos três subconjuntos avaliados.

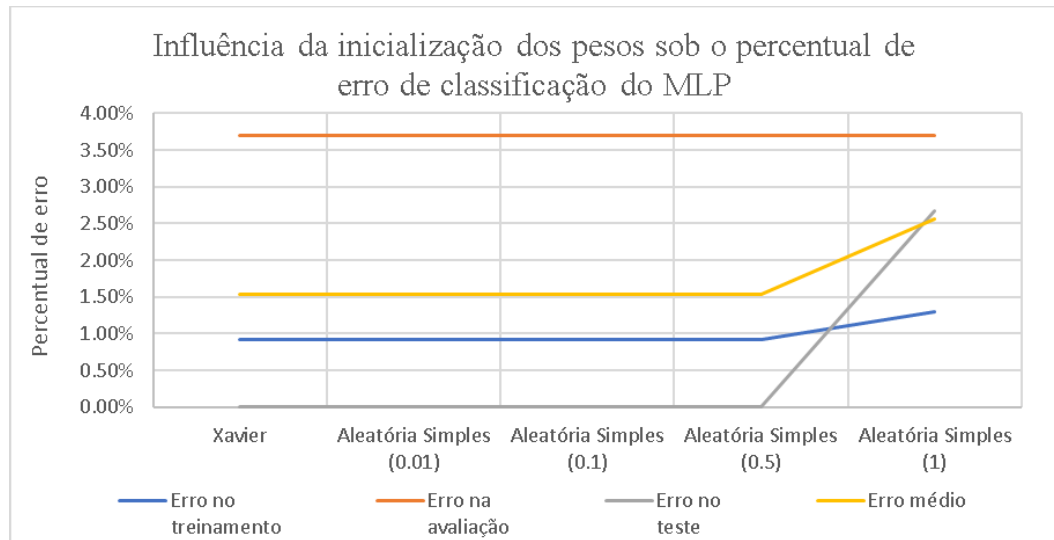
Tabela 3: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de inicialização dos pesos	Erro no treinamento	Erro na avaliação	Erro no teste	Erro médio
1000	0.35	4	Xavier	0.93%	3.70%	0.00%	1.54%
1000	0.35	4	Aleatória Simples (0.01)	0.93%	3.70%	0.00%	1.54%
1000	0.35	4	Aleatória Simples (0.1)	0.93%	3.70%	0.00%	1.54%
1000	0.35	4	Aleatória Simples (0.5)	0.93%	3.70%	0.00%	1.54%
1000	0.35	4	Aleatória Simples (1)	1.30%	3.70%	2.67%	2.56%

Fonte: Elaborado pelo autor (2023).

À título de explicação, as inicializações utilizadas foram inicialização de Xavier [10] e inicialização aleatória simples com fator multiplicativo. A primeira, consiste em ajustar a escala dos pesos com base no número de entradas e saídas das camadas da rede neural. A segunda, por sua vez, trabalha com aleatoriedade multiplicada por um fator constante e pré-definido. Tendo isso em vista, obteve-se a Figura 7.

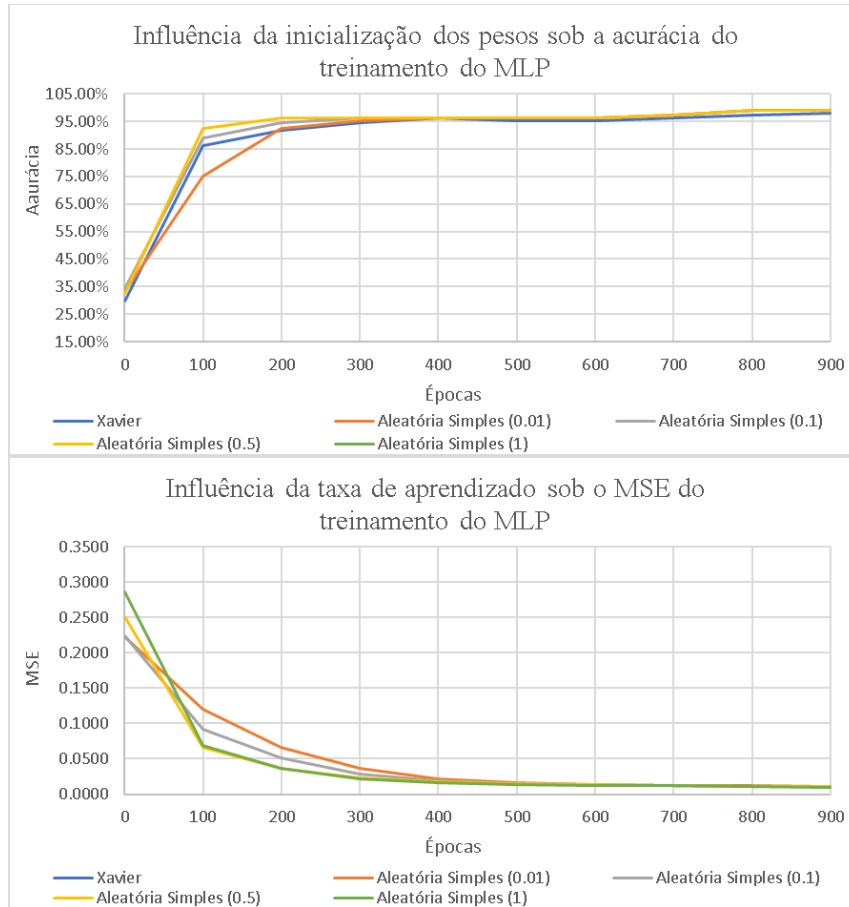
Figura 7: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.



Fonte: Elaborado pelo autor (2023).

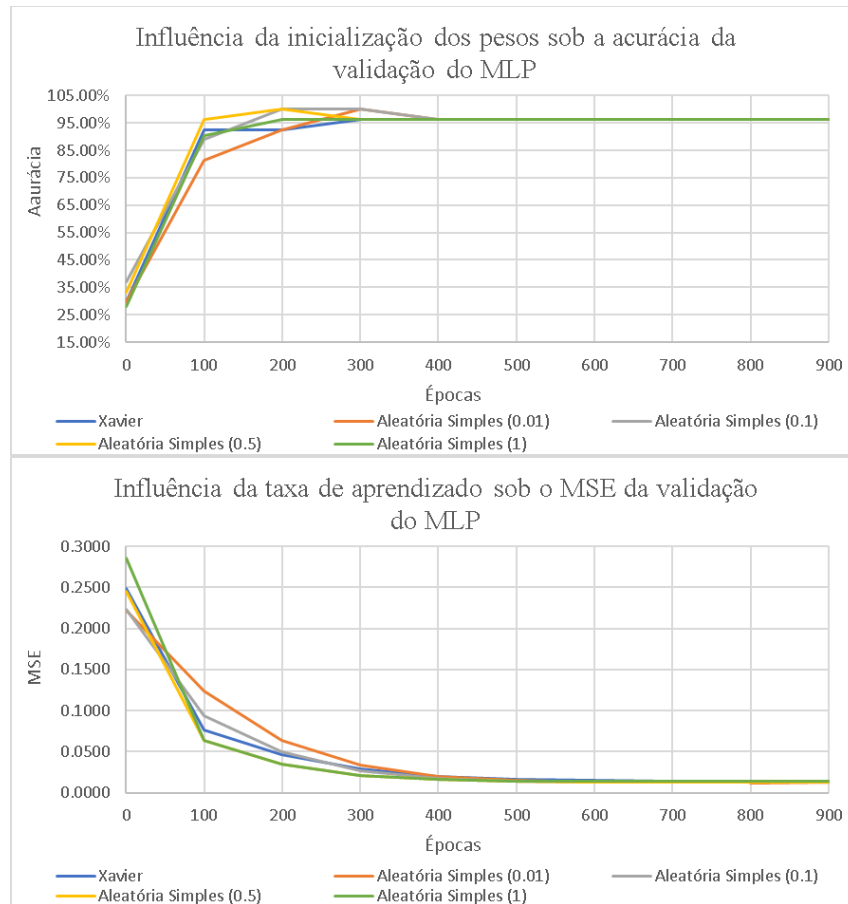
Analisando os dados, interpreta-se que, com exceção da inicialização aleatória simples com fator multiplicativo 1, todas as outras versões atingiram, sem variações, o mesmo percentual de erros. Portanto, analisou-se, seguinte a isso, a acurácia e o MSE proveniente de tais inicializações.

Figura 8: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.



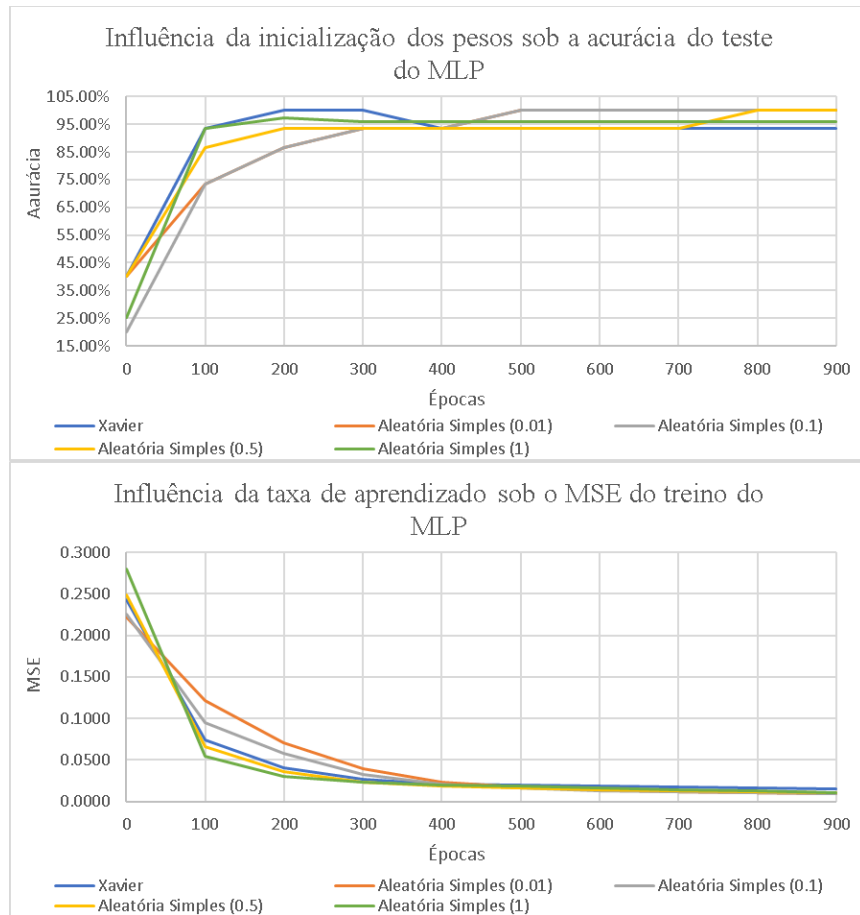
Fonte: Elaborado pelo autor (2023).

Figura 9: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.



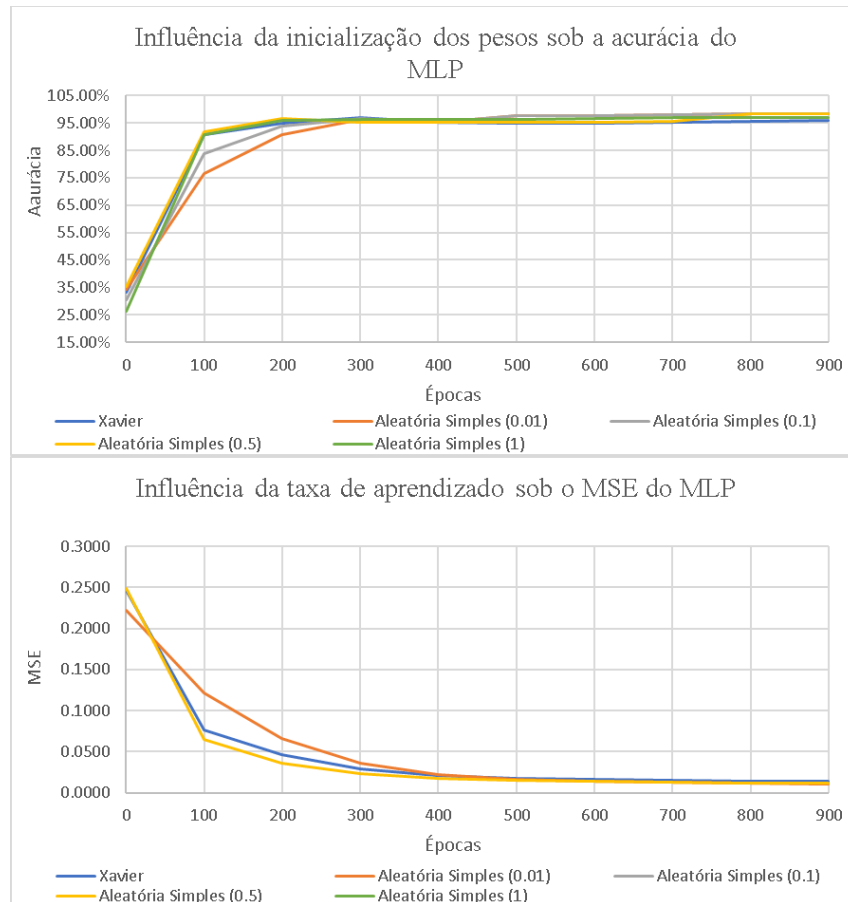
Fonte: Elaborado pelo autor (2023).

Figura 10: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

Figura 11: Influência do processo de inicialização dos pesos do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

Observando os gráficos que relatam o comportamento da acurácia e MSE do algoritmo com as diferentes inicializações propostas, optou-se por seguir com a inicialização aleatória simples com fator multiplicativo de 0,1. A opção é justificada uma vez que tende a alcançar uma acurácia satisfatória ao mesmo passo em que consegue minimizar significativamente o MSE em um número de épocas satisfatório.

1.4.3 Análise final do *Multi Layer Perceptron* (MLP)

À título de analisar o algoritmo com todos os parâmetros devidamente ajustados, com base nas explicações supracitadas, executou-se o código na configuração apresentada na Tabela 4 5 vezes, objetivando entender o comportamento da acurácia, MSE e matriz de confusão com os acertos e erros de cada subgrupo.

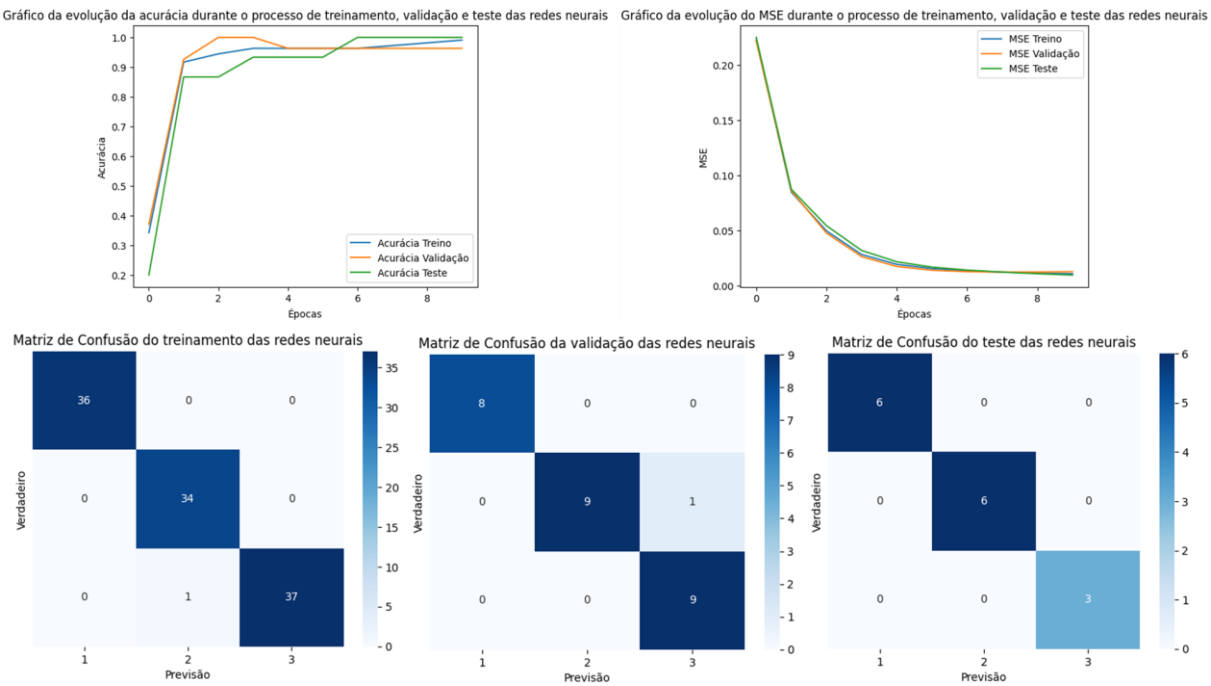
Tabela 4: Configuração final dos parâmetros do MLP.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de Inicialização dos pesos
1000	0.35	4	Aleatória Simples (0.1)

Fonte: Elaborado pelo autor (2023).

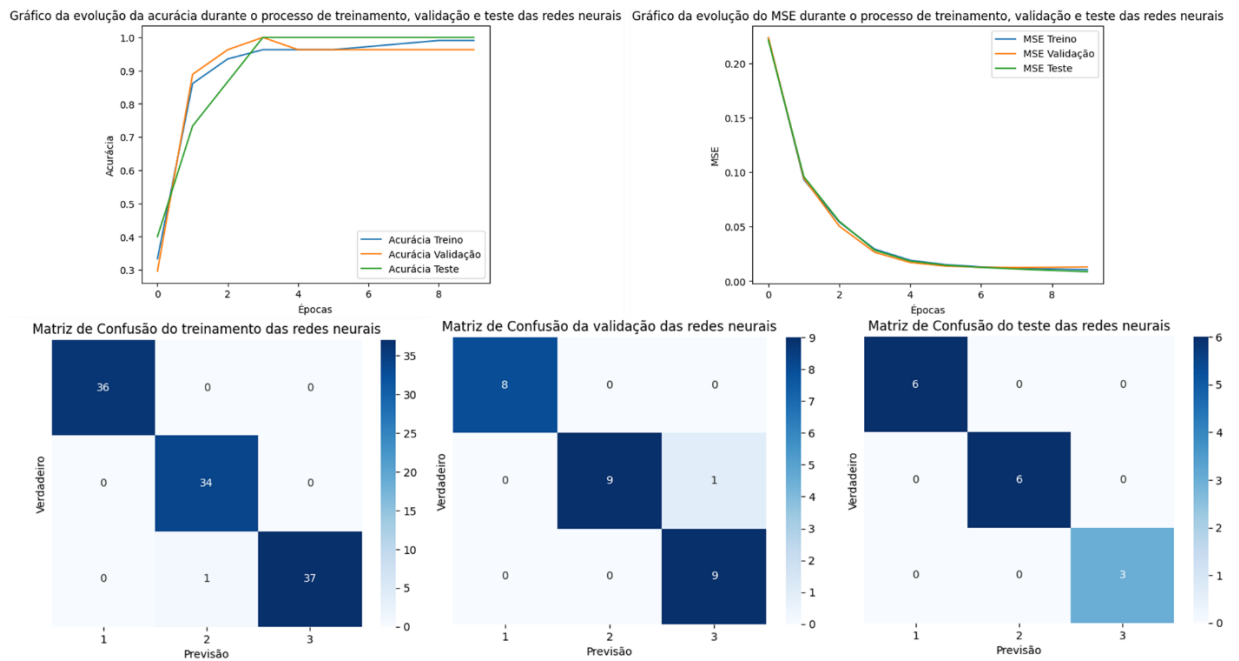
Os resultados de cada uma das simulações são vistos nas figuras abaixo.

Figura 12: Gráficos de acurácia e MSE e matrizes de confusão da primeira execução do algoritmo com as configurações finais dos parâmetros do MLP.



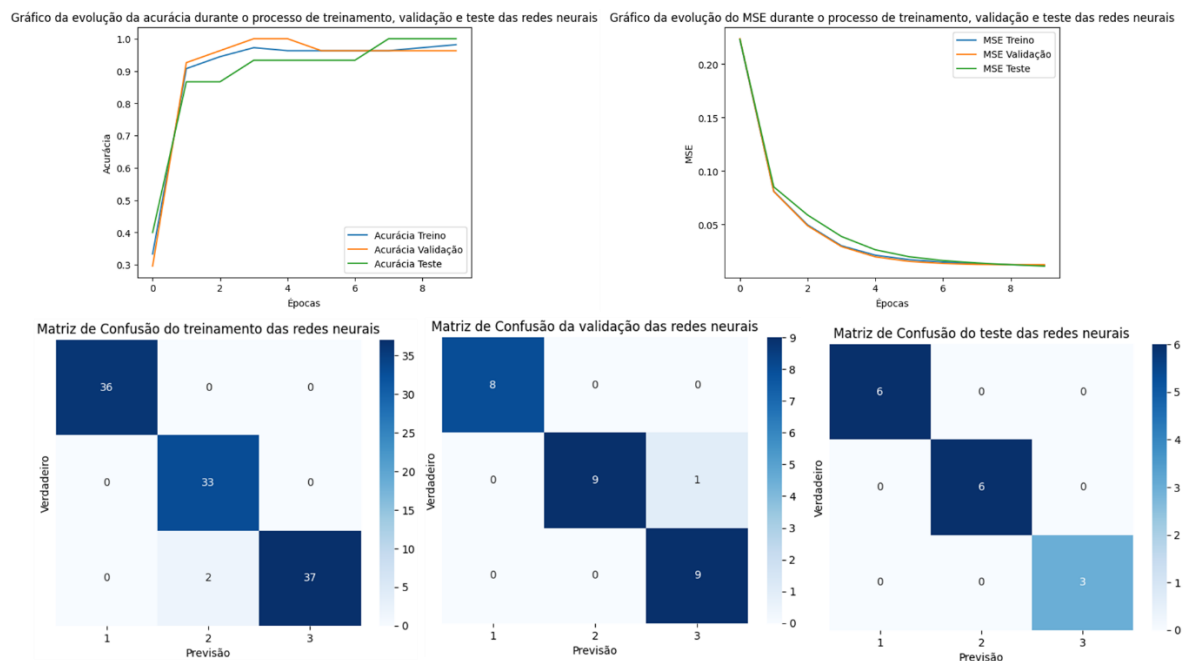
Fonte: Elaborado pelo autor (2023).

Figura 13: Gráficos de acurácia e MSE e matrizes de confusão da segunda execução do algoritmo com as configurações finais dos parâmetros do MLP.



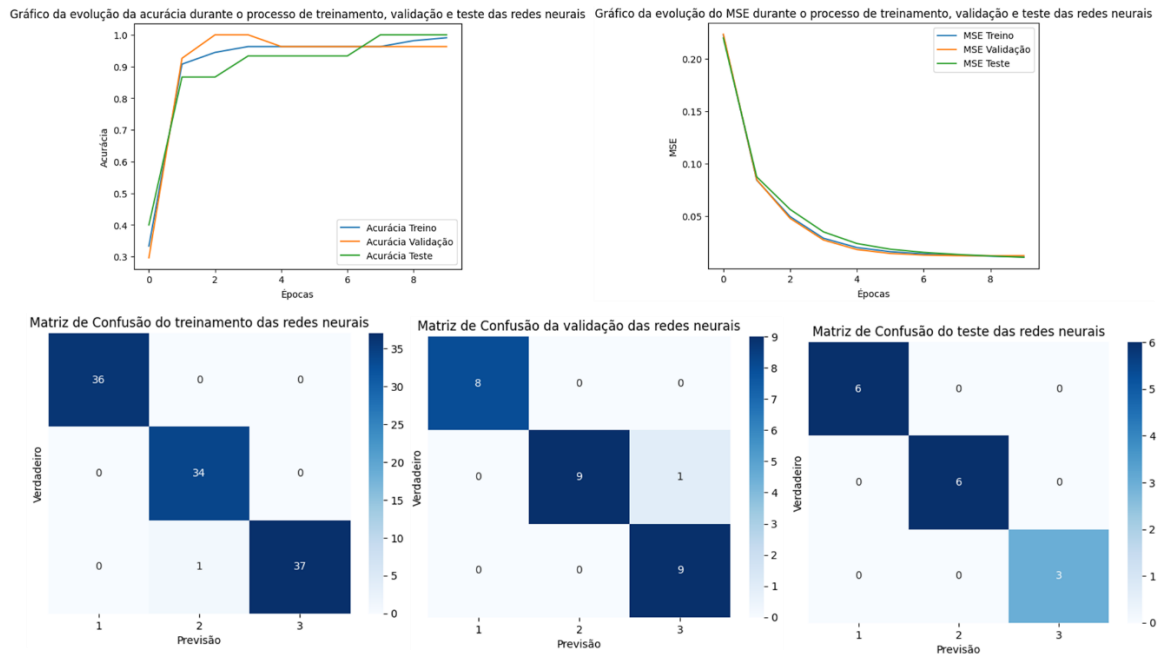
Fonte: Elaborado pelo autor (2023).

Figura 14: Gráficos de acurácia e MSE e matrizes de confusão da terceira execução do algoritmo com as configurações finais dos parâmetros do MLP.



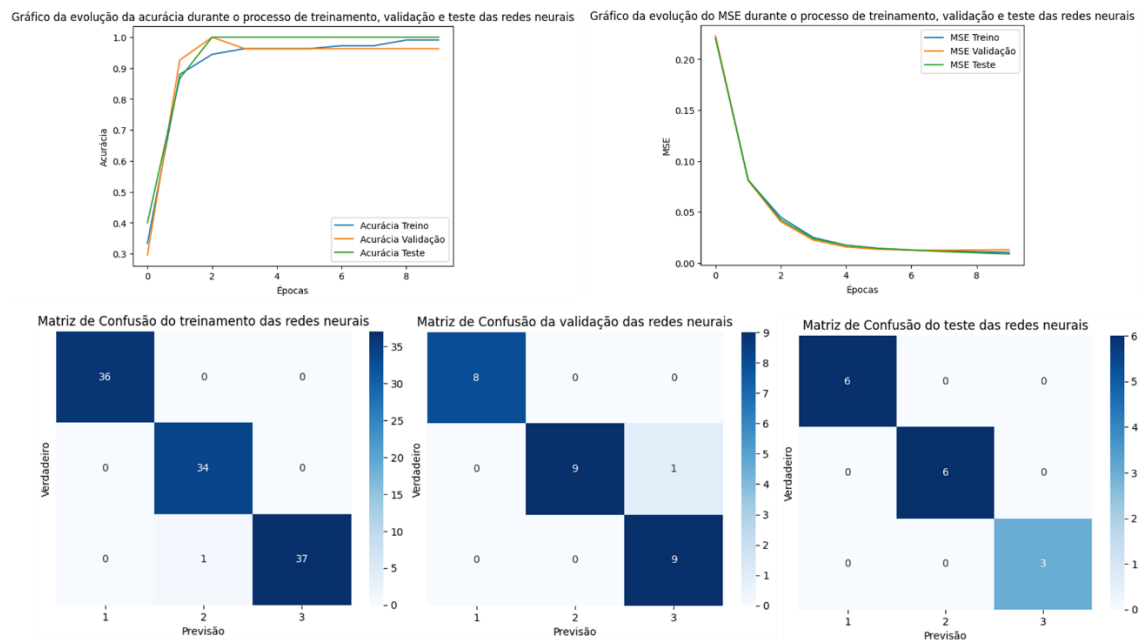
Fonte: Elaborado pelo autor (2023).

Figura 15: Gráficos de acurácia e MSE e matrizes de confusão da quarta execução do algoritmo com as configurações finais dos parâmetros do MLP.



Fonte: Elaborado pelo autor (2023).

Figura 16 Gráficos de acurácia e MSE e matrizes de confusão da quinta execução do algoritmo com as configurações finais dos parâmetros do MLP.



Fonte: Elaborado pelo autor (2023).

Analisando os resultados acima, tem-se que, no pior caso executado, a acurácia de treino atinge um valor próximo a 100% quanto aproxima-se de 1000 iterações, sofrendo, ainda, uma

pequena tendencia de crescimento. Quanto à acurácia da validação, tem-se que a pior situação ocorre atingindo um percentual de 95% em 450 iterações, se estabilizando. A acurácia do teste, por sua vez, no caso simulado mais crítico, estabiliza-se em 100% próximo à 700 iterações. O MSE, por sua vez, dentre as 1000 iterações, aproxima-se de 0. Por fim, quando se avalia as matrizes de confusão, tem-se que a pior situação traz como resultado um percentual de 1,85% de erro para o subconjunto de treino, 3,70% para o subconjunto de validação e 0% para os dados de teste, representados por 15 indivíduos.

Conclui-se, portanto, que os parâmetros estabelecidos são suficientes para possibilitar a distinção e classificação das classes das espécies presentes no banco de dados Iris.

2. Questão 2

Nessa seção, serão apresentados os tópicos relacionados a resolução da questão 2 proposta no trabalho 2 da matéria “Computação Inspirada pela Natureza”, referente ao Programa de Pós-graduação em Ciência da Computação da UNESP.

2.1 Enunciado

Faça o download do conjunto de dados Wine do UCI Machine Learning Repository:
<http://archive.ics.uci.edu/ml/datasets/Wine>.

Identifique quais são os atributos e classes do problema e coloque-os em seu relatório.

Realize experimentos similares aos realizados com o conjunto de dados Iris.

2.2 Linguagem e IDE

O código referente a questão 1 foi desenvolvido em *Python* [1], com suporte do *Google Colaboratory* [2].

2.3 Descrição lógica do algoritmo

A descrição lógica do algoritmo pode ser encontrada na seção 1.3.

É válido, adicionalmente, pontuar que a segmentação entre dados para treinamento, validação e teste consistiu em separar, respectivamente, 128, 32 e 18 dados.

2.4 Análise do algoritmo

A fim de realizar uma análise relacionada ao algoritmo desenvolvido, o código foi executado algumas vezes e, durante esse processo, alguns parâmetros foram alterados e resultados foram coletados e interpretados.

2.4.1 Influência da quantidade de neurônios da camada oculta sob o percentual de erros e acertos da classificação do *Multi Layer Perceptron* (MLP)

A primeira análise teve como intuito interpretar a influência da variação de neurônios presentes na camada oculta do *Multi Layer Perceptron* (MLP) em relação ao percentual de erros e acertos do algoritmo.

Para isso, executou-se o código 5 vezes para cada um dos arranjos observados na Tabela 5 e, para cada um desses, foi obtida a média percentual de erros dos três subconjuntos avaliados.

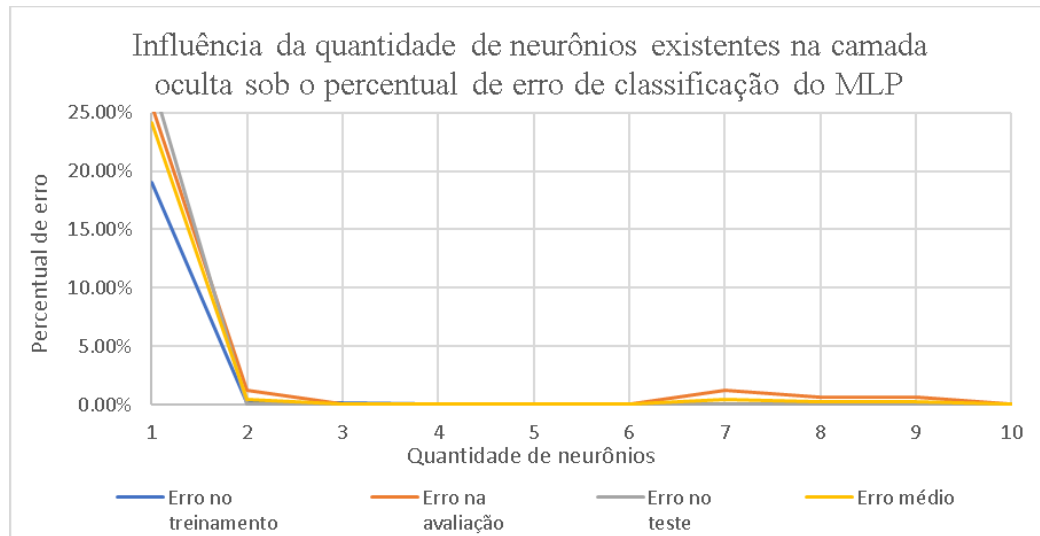
Tabela 5: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de inicialização dos pesos	Erro no treinamento	Erro na avaliação	Erro no teste	Erro médio
1000	0.1	1	Xavier	19.06%	25.63%	27.78%	24.16%
1000	0.1	2	Xavier	0.16%	1.25%	0.00%	0.47%
1000	0.1	3	Xavier	0.16%	0.00%	0.00%	0.05%
1000	0.1	4	Xavier	0.00%	0.00%	0.00%	0.00%
1000	0.1	5	Xavier	0.00%	0.00%	0.00%	0.00%
1000	0.1	6	Xavier	0.00%	0.00%	0.00%	0.00%
1000	0.1	7	Xavier	0.00%	1.25%	0.00%	0.42%
1000	0.1	8	Xavier	0.00%	0.63%	0.00%	0.21%
1000	0.1	9	Xavier	0.00%	0.63%	0.00%	0.21%
1000	0.1	10	Xavier	0.00%	0.00%	0.00%	0.00%

Fonte: Elaborado pelo autor (2023).

O gráfico obtido com os dados extraídos da Tabela 5 é implementado na Figura 17.

Figura 17: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.



Fonte: Elaborado pelo autor (2023).

Interpretando as informações apresentadas acima, fica claro que a partir da presença de 2 neurônios na camada oculta, com os parâmetros pré-definidos e apresentados na Tabela 5, o percentual de erro classificatório do MLP segue uma tendência de linearização. Além disso, percebe-se que os melhores cenários ocorrem com a existência de 4, 5, 6 e 10 neurônios.

Haja vista a situação explicitada acima, objetivando uma minimização da quantidade de classificações errôneas e combinando tal fator com a questão de complexidade da rede e processamento computacional, considerou-se um MLP de 1 camada oculta, sendo esta composta por 5 neurônios.

2.4.2 Influência da taxa de aprendizado sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do *Multi Layer Perceptron* (MLP)

A segunda análise buscou entender a influência da taxa de aprendizado do algoritmo em relação ao percentual de erros e acertos deste, bem como acurácia e MSE.

Para isso, executou-se o código 5 vezes para cada uma das configurações observadas na Tabela 6 e, para cada um desses arranjos, foi obtida a média percentual de erros dos três subconjuntos avaliados.

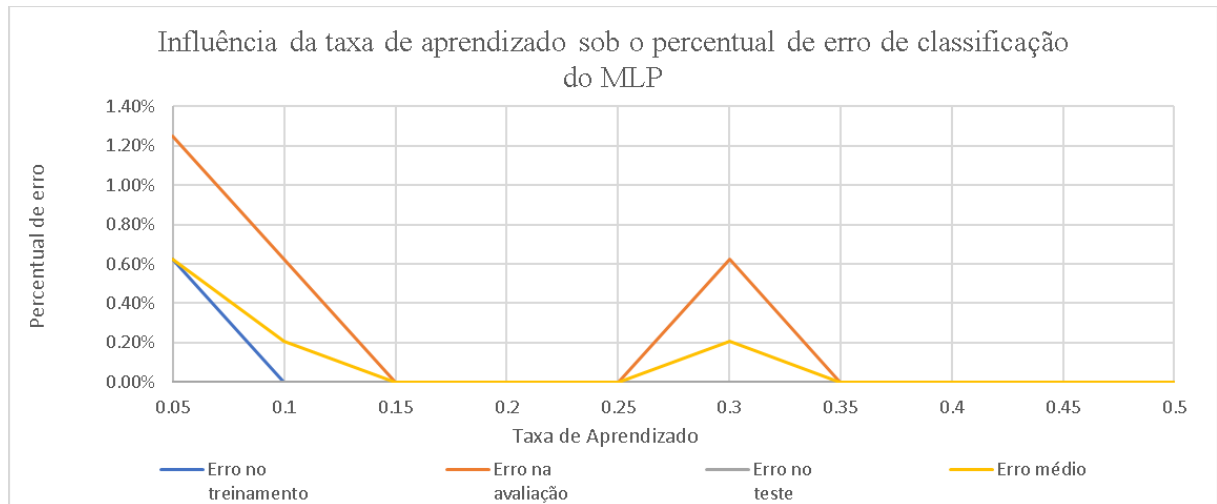
Tabela 6: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de inicialização dos pesos	Erro no treinamento	Erro na avaliação	Erro no teste	Erro médio
1000	0.05	5	Xavier	0.63%	1.25%	0.00%	0.63%
1000	0.1	5	Xavier	0.00%	0.63%	0.00%	0.21%
1000	0.15	5	Xavier	0.00%	0.00%	0.00%	0.00%
1000	0.2	5	Xavier	0.00%	0.00%	0.00%	0.00%
1000	0.25	5	Xavier	0.00%	0.00%	0.00%	0.00%
1000	0.3	5	Xavier	0.00%	0.63%	0.00%	0.21%
1000	0.35	5	Xavier	0.00%	0.00%	0.00%	0.00%
1000	0.4	5	Xavier	0.00%	0.00%	0.00%	0.00%
1000	0.45	5	Xavier	0.00%	0.00%	0.00%	0.00%
1000	0.5	5	Xavier	0.00%	0.00%	0.00%	0.00%

Fonte: Elaborado pelo autor (2023).

O gráfico que representa os dados da Tabela 6 é apresentado na Figura 18.

Figura 18: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.

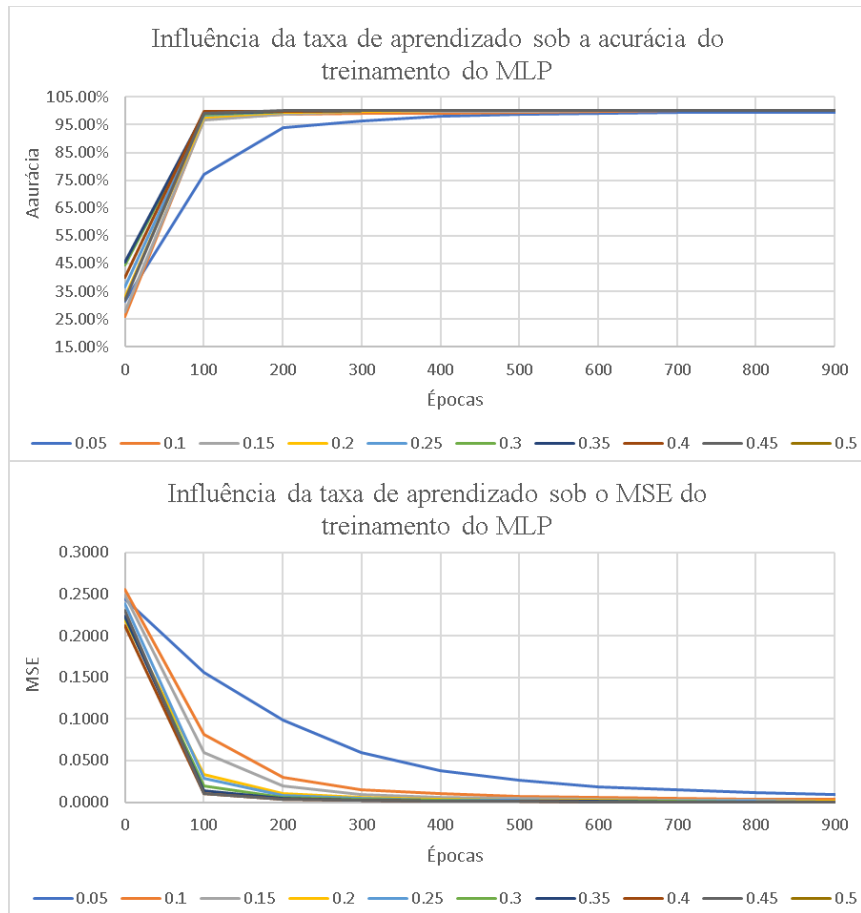


Fonte: Elaborado pelo autor (2023).

Observando os dados acima e considerando as 1000 iterações propostas, percebe-se que quanto a taxa de aprendizado é igual ou maior do que 0,15, com exceção da taxa 0,3, existe uma tendência de que a quantidade de acertos de classificação do MLP seja maximizada em todos os subgrupos existentes, ou seja, no processo de treinamento, validação e teste.

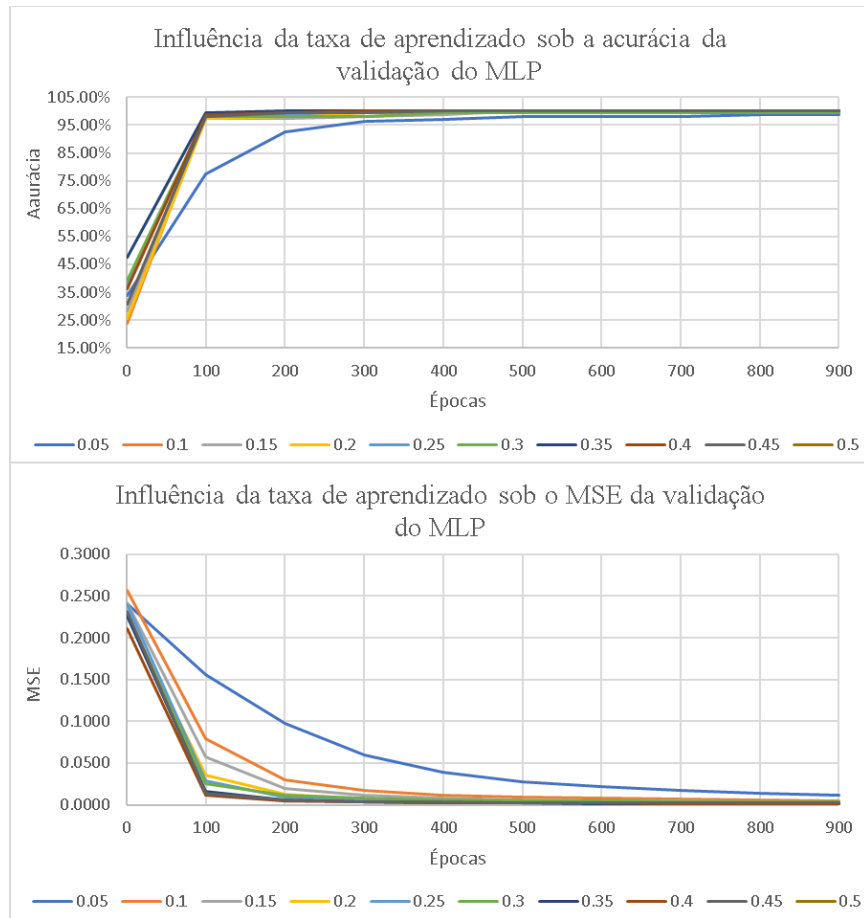
Com isso em mente, partiu-se para a análise da acurácia e MSE do algoritmo submetido à tais variações de taxa de aprendizado.

Figura 19: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.



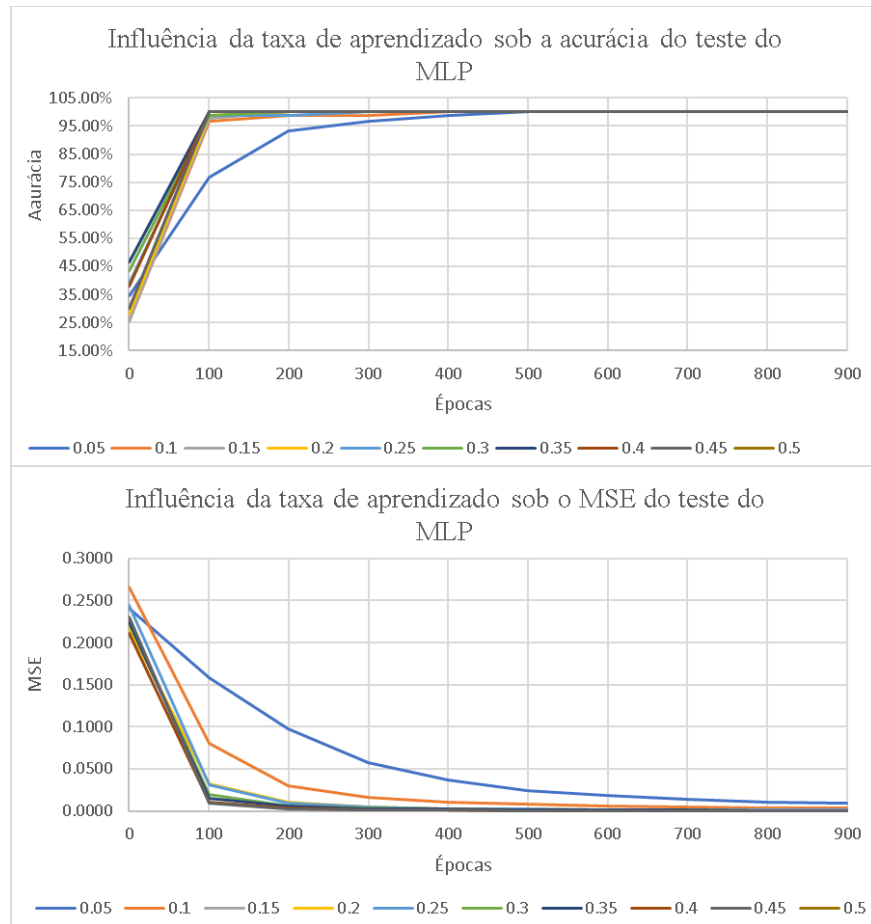
Fonte: Elaborado pelo autor (2023).

Figura 20: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.



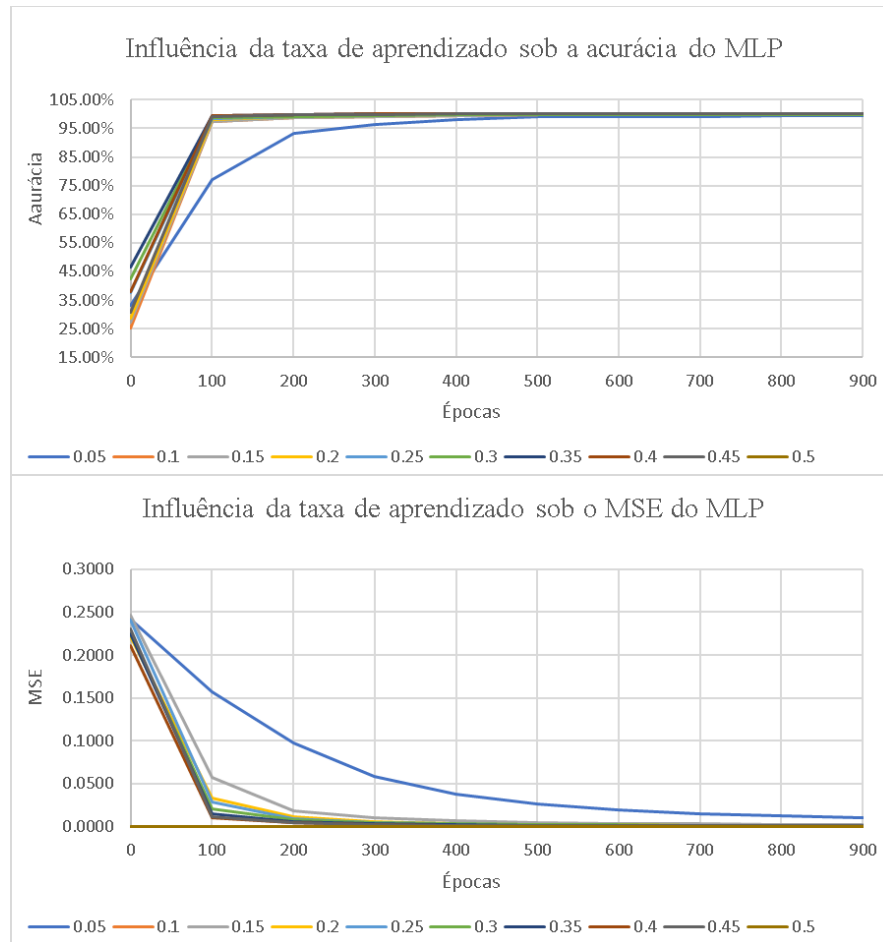
Fonte: Elaborado pelo autor (2023).

Figura 21: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

Figura 22: Influência da taxa de aprendizado do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

É pertinente exportar da análise gráfica que, com o aumento da taxa de aprendizado, considerando 1000 iterações, a tendência é de que ocorra, de maneira geral, uma minimização do MSE. A acurácia, por sua vez, tem seu atingimento máximo e convergência antecipado.

De modo a associar otimização com ótimos resultados, optou-se pela taxa de aprendizado de 0,2, de modo a garantir uma acurácia e convergência satisfatória em todos os subgrupos submetidos a análise, bem como um MSE suficientemente minimizado que não culmine em *overfitting* e erros de classificação, por exemplo. Em paralelo a isso, o valor contribui para que o algoritmo não se torne excessivamente custoso computacionalmente, assim como reduz a probabilidade de que o resultado se prenda à ótimos locais.

2.4.3 Influência inicialização dos pesos sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do *Multi Layer Perceptron* (MLP)

A terceira análise teve como objetivo entender a influência da inicialização dos pesos em relação ao percentual de erros e acertos do código, bem como a acurácia e MSE.

Para isso, executou-se o código 5 vezes para cada uma das configurações observadas na Tabela 7 e, para cada um desses arranjos, foi obtida a média percentual de erros dos três subconjuntos avaliados.

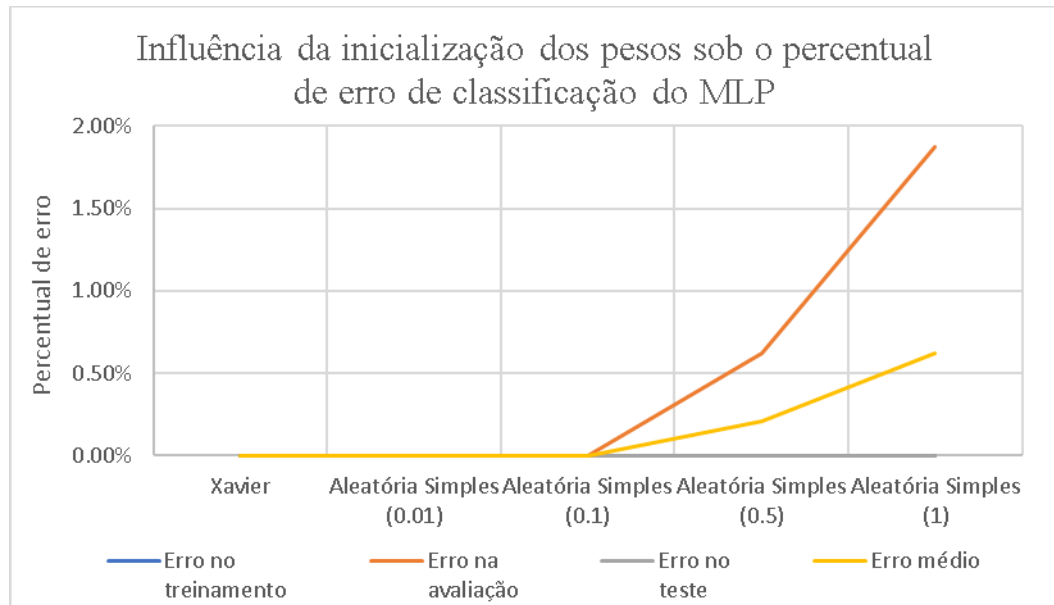
Tabela 7: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de inicialização dos pesos	Erro no treinamento	Erro na avaliação	Erro no teste	Erro médio
1000	0.2	5	Xavier	0.00%	0.00%	0.00%	0.00%
1000	0.2	5	Aleatória Simples (0.01)	0.00%	0.00%	0.00%	0.00%
1000	0.2	5	Aleatória Simples (0.1)	0.00%	0.00%	0.00%	0.00%
1000	0.2	5	Aleatória Simples (0.5)	0.00%	0.63%	0.00%	0.21%
1000	0.2	5	Aleatória Simples (1)	0.00%	1.88%	0.00%	0.63%

Fonte: Elaborado pelo autor (2023).

As inicializações de peso utilizadas foram inicialização de Xavier [10] e inicialização aleatória simples com fator multiplicativo. Com base nessa informação e nos dados extraídos da Tabela 7, obteve-se a Figura 23.

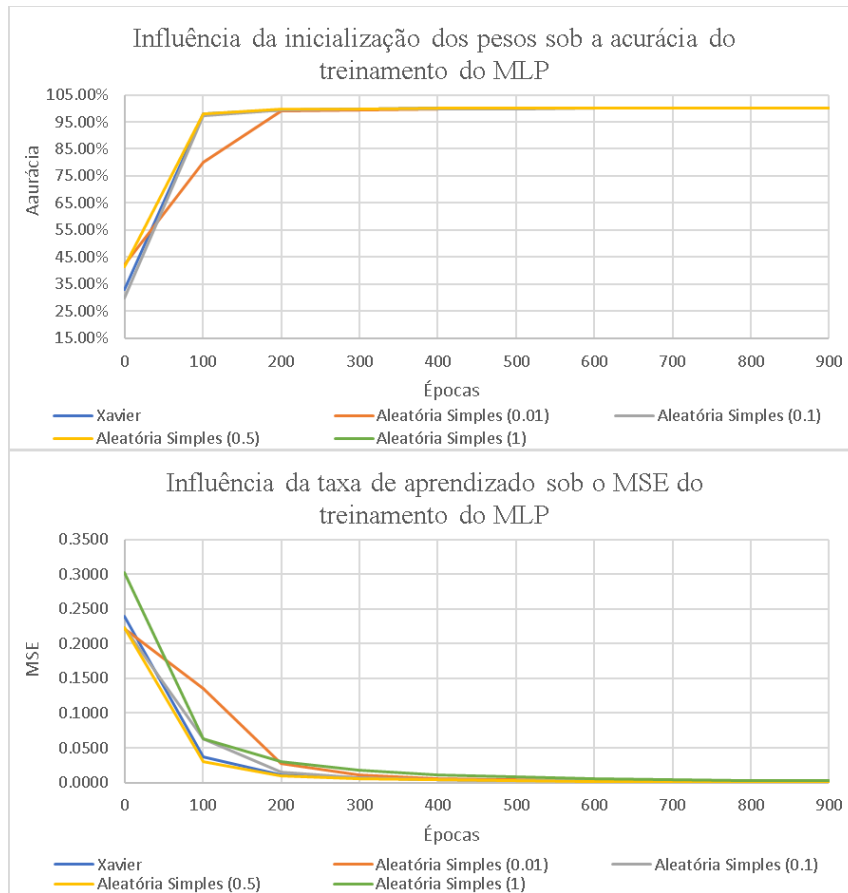
Figura 23: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.



Fonte: Elaborado pelo autor (2023).

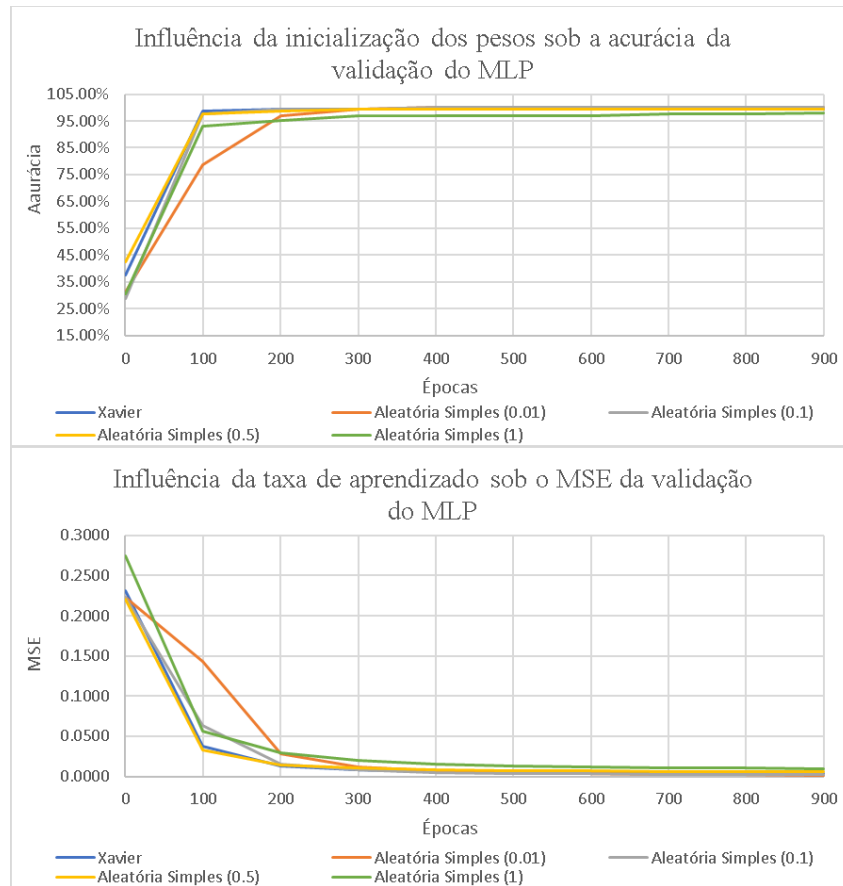
Analisando os dados, interpreta-se que os melhores resultados referentes à erro e acerto atrelam-se à inicialização de Xavier e a inicialização aleatória simples com fator multiplicativo de 0,01 e 0,1. À título de entender a variação da acurácia e MSE do MLP conforme altera-se o processo de inicialização dos pesos da rede neural, plotaram-se os gráficos abaixo.

Figura 24: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.



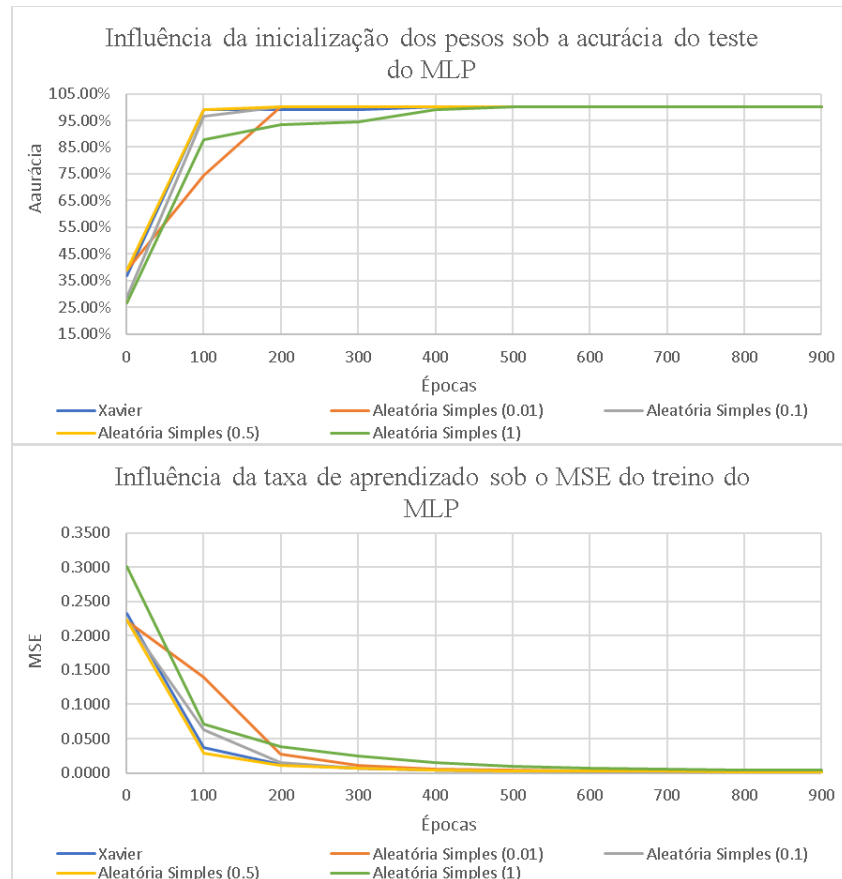
Fonte: Elaborado pelo autor (2023).

Figura 25: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.



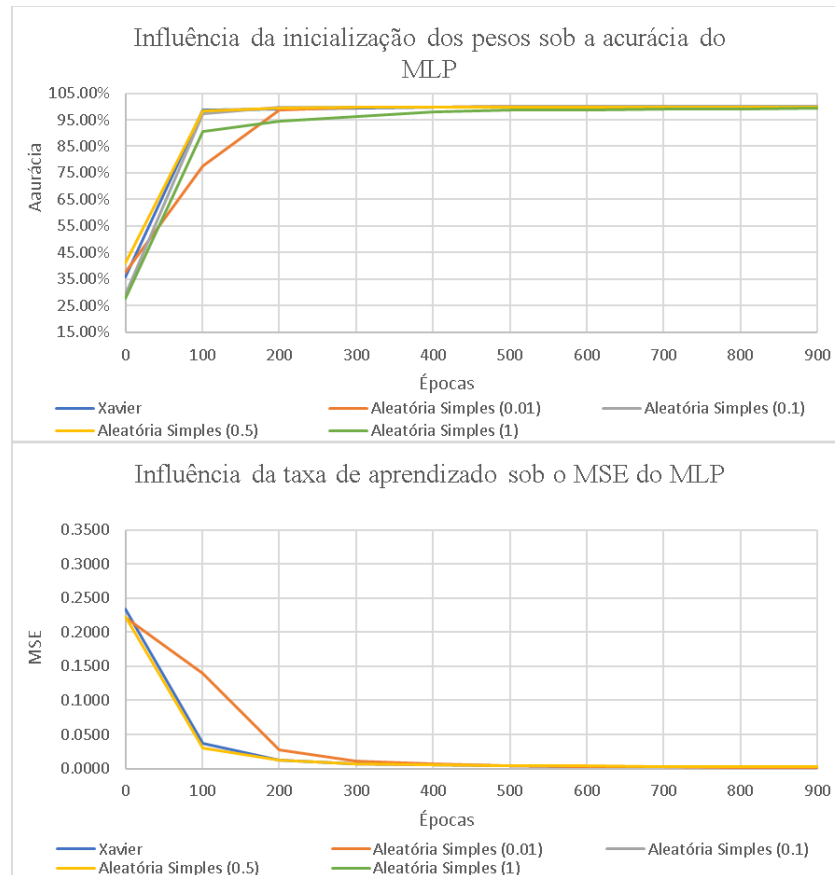
Fonte: Elaborado pelo autor (2023).

Figura 26: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

Figura 27: Influência do processo de inicialização dos pesos do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

Observando os gráficos que relatam o comportamento da acurácia e MSE do algoritmo com as diferentes inicializações propostas, e considerando a análise já realizada quanto ao percentual de erro e acerto no processo classificatório do MLP, optou-se por seguir com a inicialização de Xavier. A opção é justificada uma vez que tende a alcançar uma acurácia satisfatória ao mesmo tempo em que consegue minimizar significativamente o MSE em um número de épocas satisfatório. Adicionalmente, a inicialização escolhida apresenta uma classificação com elevado potencial de acerto.

2.4.3 Análise final do *Multi Layer Perceptron* (MLP)

A fim de analisar o resultado do algoritmo com todos os parâmetros devidamente ajustados (Tabela 8), executou-se o código 5 vezes e, com base em tais execuções, analisou-se a resposta do MLP quanto à acurácia, MSE e probabilidade de erro de classificação.

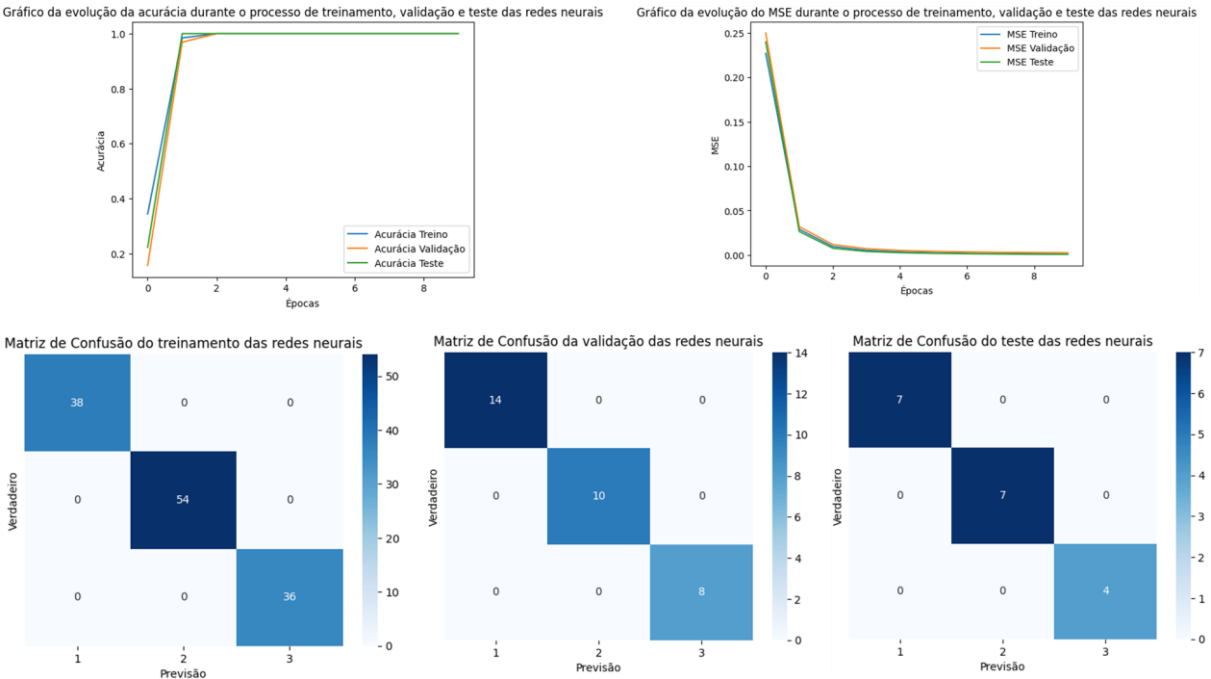
Tabela 8: Configuração final dos parâmetros do MLP.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de Inicialização dos pesos
1000	0.2	5	Xavier

Fonte: Elaborado pelo autor (2023).

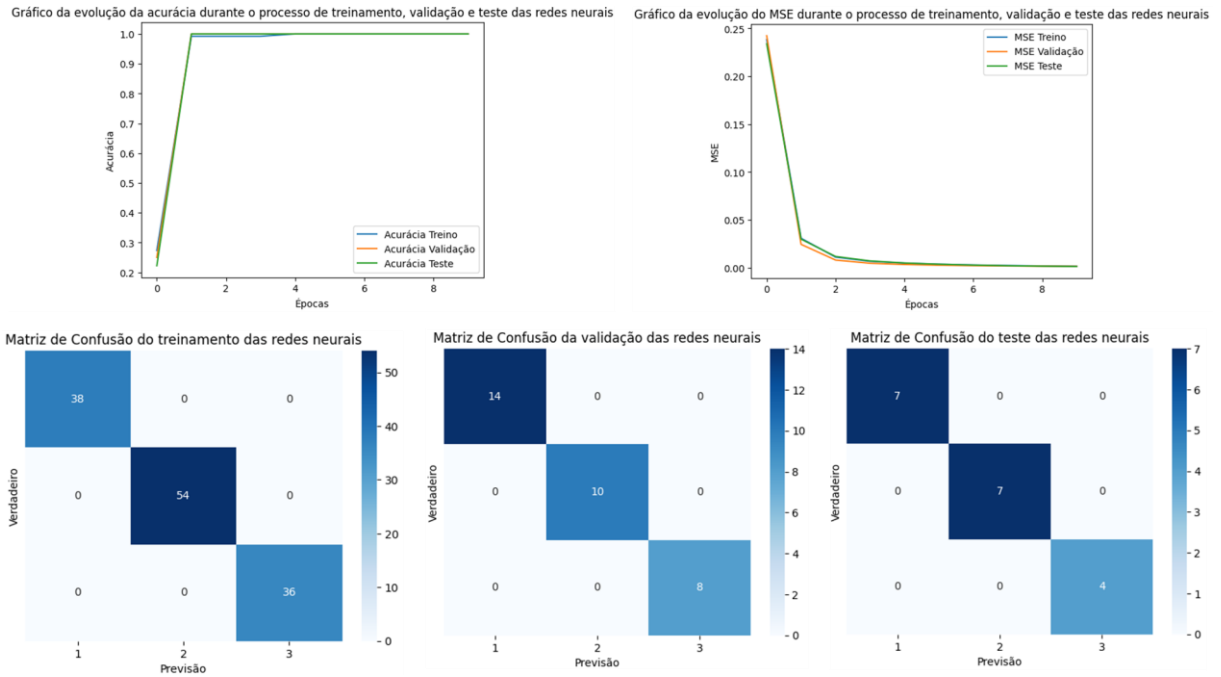
Os resultados de cada uma das simulações são vistos nas figuras abaixo.

Figura 28: Gráficos de acurácia e MSE e matrizes de confusão da primeira execução do algoritmo com as configurações finais dos parâmetros do MLP.



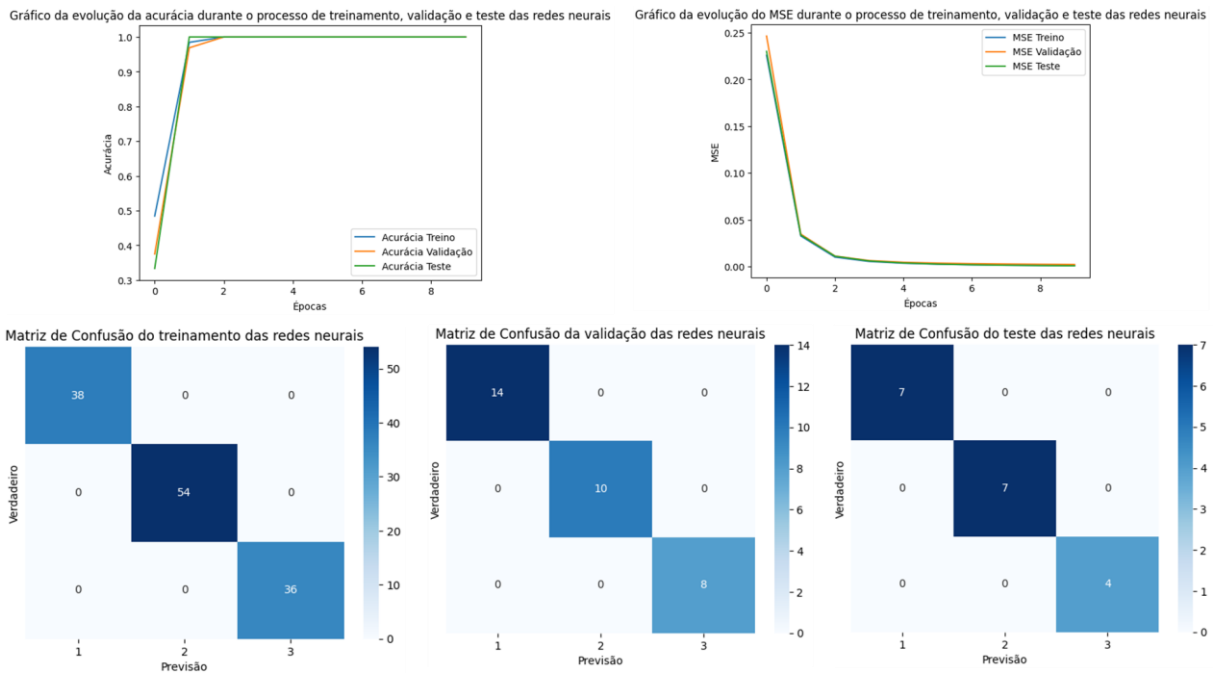
Fonte: Elaborado pelo autor (2023).

Figura 29: Gráficos de acurácia e MSE e matrizes de confusão da segunda execução do algoritmo com as configurações finais dos parâmetros do MLP.



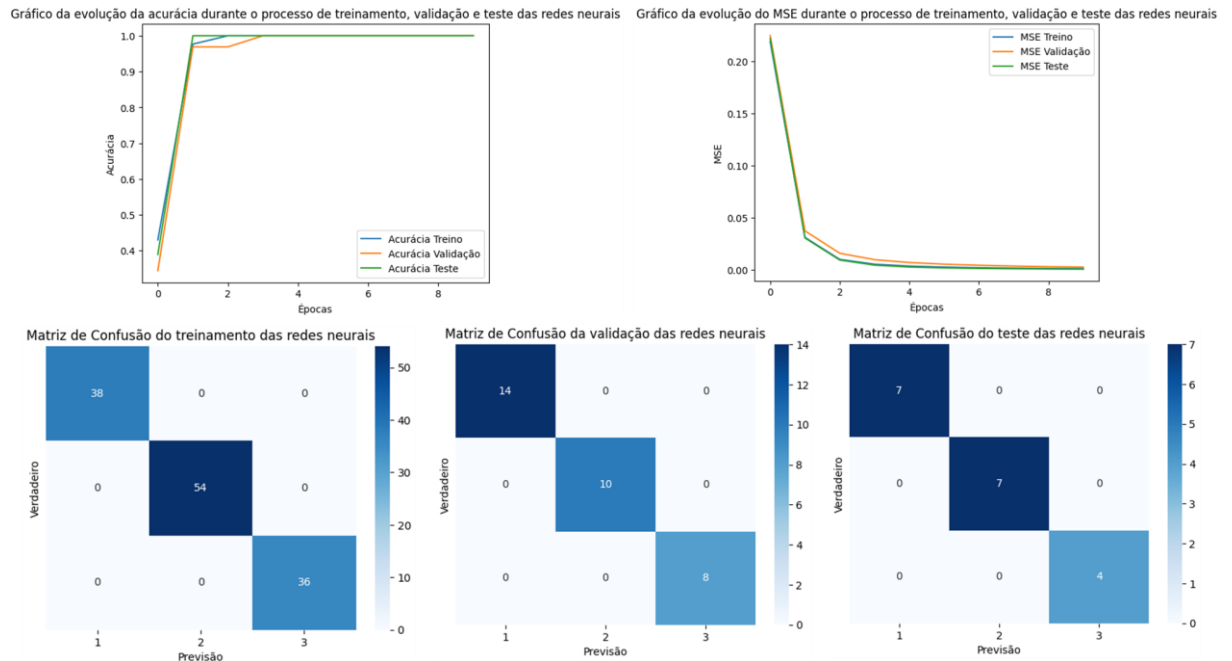
Fonte: Elaborado pelo autor (2023).

Figura 30: : Gráficos de acurácia e MSE e matrizes de confusão da terceira execução do algoritmo com as configurações finais dos parâmetros do MLP.



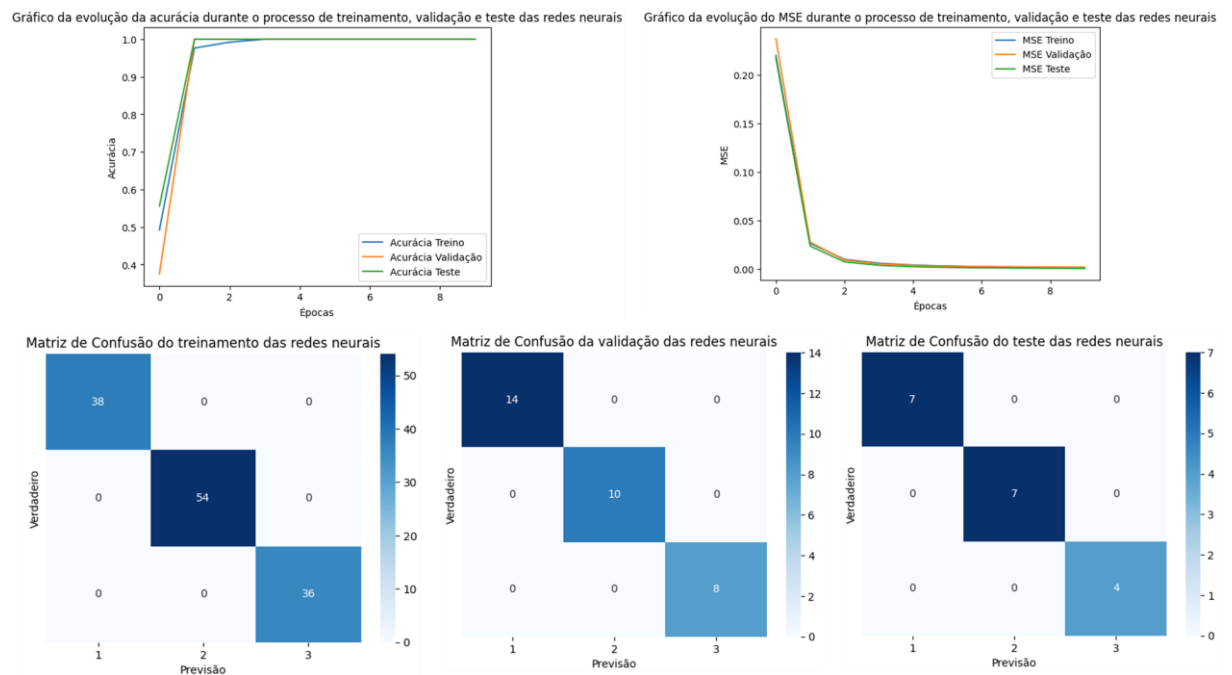
Fonte: Elaborado pelo autor (2023).

Figura 31: Gráficos de acurácia e MSE e matrizes de confusão da quarta execução do algoritmo com as configurações finais dos parâmetros do MLP.



Fonte: Elaborado pelo autor (2023).

Figura 32: Gráficos de acurácia e MSE e matrizes de confusão da quinta execução do algoritmo com as configurações finais dos parâmetros do MLP.



Fonte: Elaborado pelo autor (2023).

Analisando os resultados apresentados acima, tem-se que no pior caso executado dentre as 5 simulações, a acurácia de treinamento atinge um valor de 100% próximo à 100 iterações. O MSE, por sua vez, é minimizado, atingindo uma estabilidade considerável, próximo à 400 iterações. Quando se avalia a matriz de confusão, constituída pelos erros e acertos classificatórios, tem-se um percentual de 100% de acerto em todas as execuções do algoritmo.

Conclui-se, portanto, que os parâmetros estabelecidos são suficientes para possibilitar a distinção e classificação dos vinhos propostos.

3. Questão 3

Nessa seção, serão apresentados os tópicos relacionados a resolução da questão 3 proposta no trabalho 2 da matéria “Computação Inspirada pela Natureza”, referente ao Programa de Pós-graduação em Ciência da Computação da UNESP.

3.1 Enunciado

Realize experimentos de classificação ou qualquer outra tarefa de aprendizado de máquina com redes neurais em um conjunto de dados de sua escolha, que esteja disponível publicamente.

Qualquer rede neural pode ser utilizada e comparações entre diferentes redes neurais são encorajadas.

3.2 Linguagem e IDE

O código referente a questão 3 foi desenvolvido em *Python* [1], com suporte do *Google Colaboratory* [2].

3.3. Banco de dados

O banco de dados utilizado pode ser encontrado no link <https://www.kaggle.com/datasets/rwzhang/seeds-dataset>.

O banco de dados em questão é formado por 210 instâncias que compõem 3 classes balanceadas. As classes representam 3 espécies de semente de trigo: Kama, Rosa e Canadian.

As características, por sua vez, representam informações das sementes em análise: área, perímetro, compactude, comprimento, largura, coeficiente de assimetria e comprimento.

3.4 Descrição lógica do algoritmo

A descrição lógica do algoritmo pode ser encontrada na seção 1.3.

A segmentação dos três subgrupos – treinamento, validação e teste – baseiam-se em 151 instâncias, 17 instâncias e 42 instâncias, respectivamente.

3.5 Análise do algoritmo

A fim de realizar uma análise relacionada ao algoritmo desenvolvido, o código foi executado algumas vezes e, durante esse processo, alguns parâmetros foram alterados e resultados foram coletados e interpretados.

3.5.1 Influência da quantidade de neurônios da camada oculta sob o percentual de erros e acertos da classificação do *Multi Layer Perceptron* (MLP)

A primeira análise almejou identificar a influência da variação de neurônios presentes na camada oculta do *Multi Layer Perceptron* (MLP) em relação ao percentual de erros e acertos do algoritmo.

Para isso, executou-se o código 5 vezes para cada um dos arranjos observados na Tabela 9 e, para cada um desses, foi obtida a média percentual de erros dos três subconjuntos avaliados.

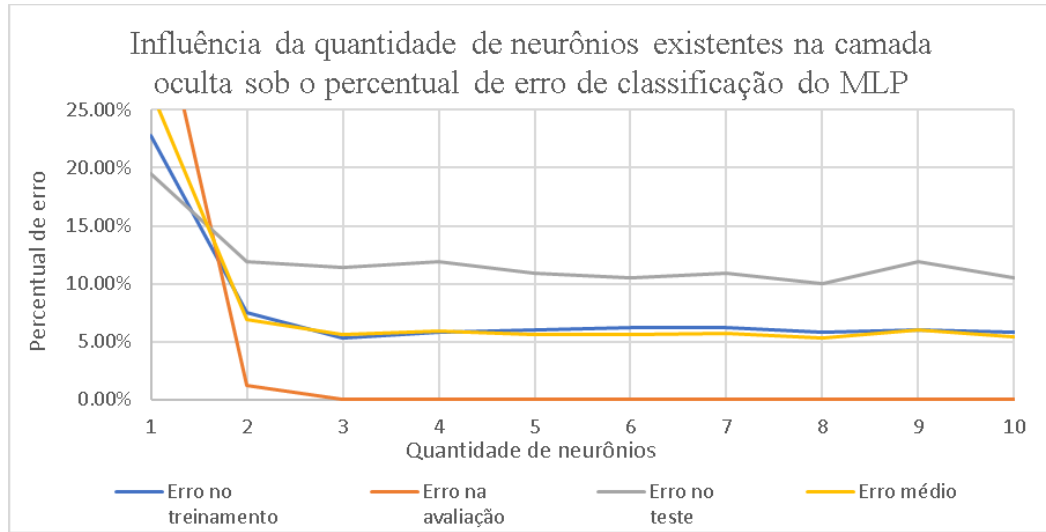
Tabela 9: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de inicialização dos pesos	Erro no treinamento	Erro na avaliação	Erro no teste	Erro médio
1000	0.1	1	Xavier	22.78%	37.65%	19.52%	26.65%
1000	0.1	2	Xavier	7.55%	1.18%	11.90%	6.88%
1000	0.1	3	Xavier	5.30%	0.00%	11.43%	5.58%
1000	0.1	4	Xavier	5.83%	0.00%	11.90%	5.91%
1000	0.1	5	Xavier	5.96%	0.00%	10.95%	5.64%
1000	0.1	6	Xavier	6.23%	0.00%	10.48%	5.57%
1000	0.1	7	Xavier	6.23%	0.00%	10.95%	5.73%
1000	0.1	8	Xavier	5.83%	0.00%	10.00%	5.28%
1000	0.1	9	Xavier	5.96%	0.00%	11.90%	5.96%
1000	0.1	10	Xavier	5.83%	0.00%	10.48%	5.43%

Fonte: Elaborado pelo autor (2023).

O gráfico obtido com os dados extraídos da Tabela 9 é implementado na Figura 33.

Figura 33: Influência da variação da quantidade de neurônios presentes na camada oculta do MLP em relação ao percentual de classificações errôneas do algoritmo.



Fonte: Elaborado pelo autor (2023).

É possível extrair da Tabela 9 e Figura 33 que, a partir da adição do terceiro neurônio na camada oculta, o algoritmo apresenta uma linearidade considerável quanto ao percentual de erros e acertos classificatórios. Objetivando balancear processamento computacional bem como a eficácia no processo classificatório do MLP, optou-se por construir a camada oculta da rede neural com 6 neurônios.

Os 6 neurônios foram escolhidos uma vez que apresentam um dos melhores resultados percentuais de acerto das instâncias de teste, bem como um baixo erro médio e, paralelo a isso, não tornam a arquitetura do MLP extremamente complexa, contribuindo com os requisitos de gastos computacionais.

3.5.2 Influência da taxa de aprendizado sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do *Multi Layer Perceptron* (MLP)

A segunda análise buscou entender a influência da taxa de aprendizado do algoritmo em relação ao percentual de erros e acertos deste, bem como acurácia e MSE.

Para isso, executou-se o código 5 vezes para cada uma das configurações observadas na Tabela 10 e, para cada um desses arranjos, foi obtida a média percentual de erros dos três subconjuntos avaliados.

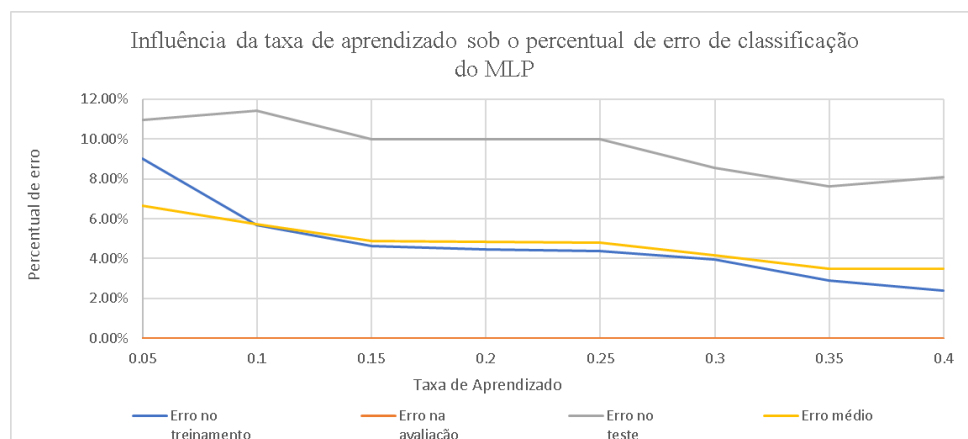
Tabela 10: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de inicialização dos pesos	Erro no treinamento	Erro na avaliação	Erro no teste	Erro médio
1000	0.05	6	Xavier	9.01%	0.00%	10.95%	6.65%
1000	0.1	6	Xavier	5.70%	0.00%	11.43%	5.71%
1000	0.15	6	Xavier	4.64%	0.00%	10.00%	4.88%
1000	0.2	6	Xavier	4.47%	0.00%	10.00%	4.82%
1000	0.25	6	Xavier	4.37%	0.00%	10.00%	4.79%
1000	0.3	6	Xavier	3.97%	0.00%	8.57%	4.18%
1000	0.35	6	Xavier	2.91%	0.00%	7.62%	3.51%
1000	0.4	6	Xavier	2.38%	0.00%	8.10%	3.49%

Fonte: Elaborado pelo autor (2023).

O gráfico que representa os dados da Tabela 10 é apresentado na Figura 34.

Figura 34: Influência da taxa de aprendizado do MLP em relação ao percentual de classificações errôneas do algoritmo.



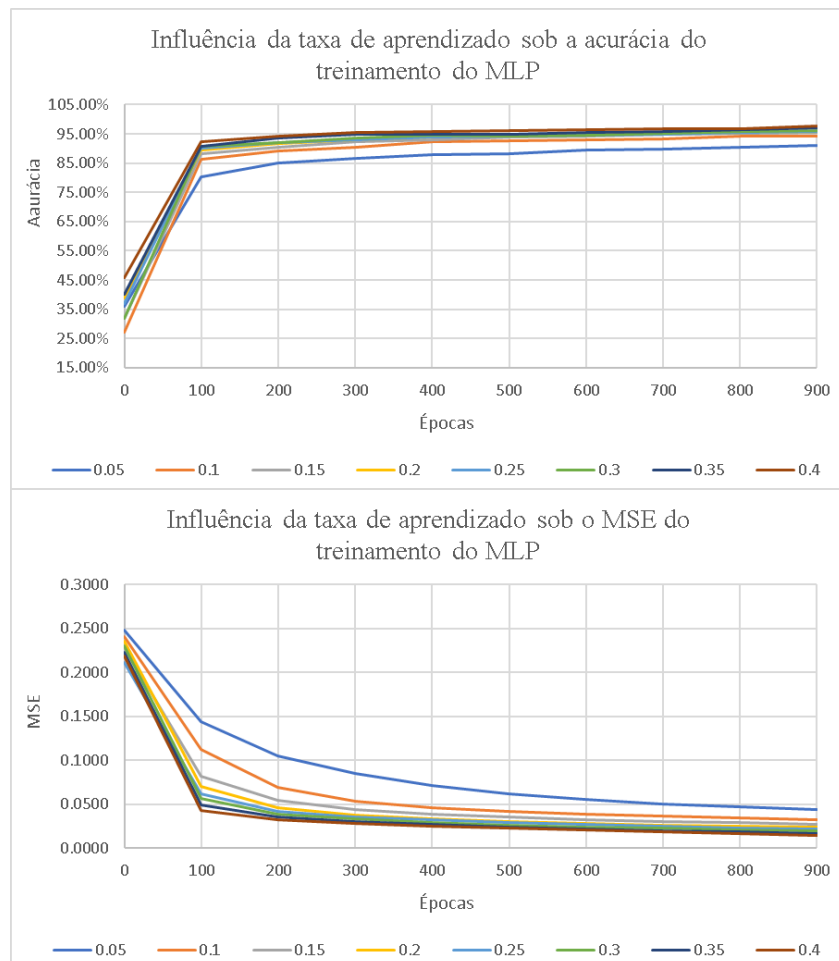
Fonte: Elaborado pelo autor (2023).

Com base nos dados apresentados anteriormente, percebe-se que a taxa de aprendizado de 0,35 é a responsável por equilibrar um menor percentil de erro do subconjunto de treinamento, validação, teste, resultando, consequente, um menor erro médio.

Analisando a taxa de 0,4, tem-se, comparativamente à taxa de aprendizado de 0,35, um menor percentil de erro do subconjunto de treinamento ao mesmo passo em que o conjunto de teste se apresenta pior. Esse é um indicativo significativo de *overffiting*, que se comprova posteriormente em conjunto à análise do MSE do conjunto. Por esse motivo, tal taxa de aprendizado, em conjunto com os valores maiores do que 0,4, não são consideradas como uma possível escolha.

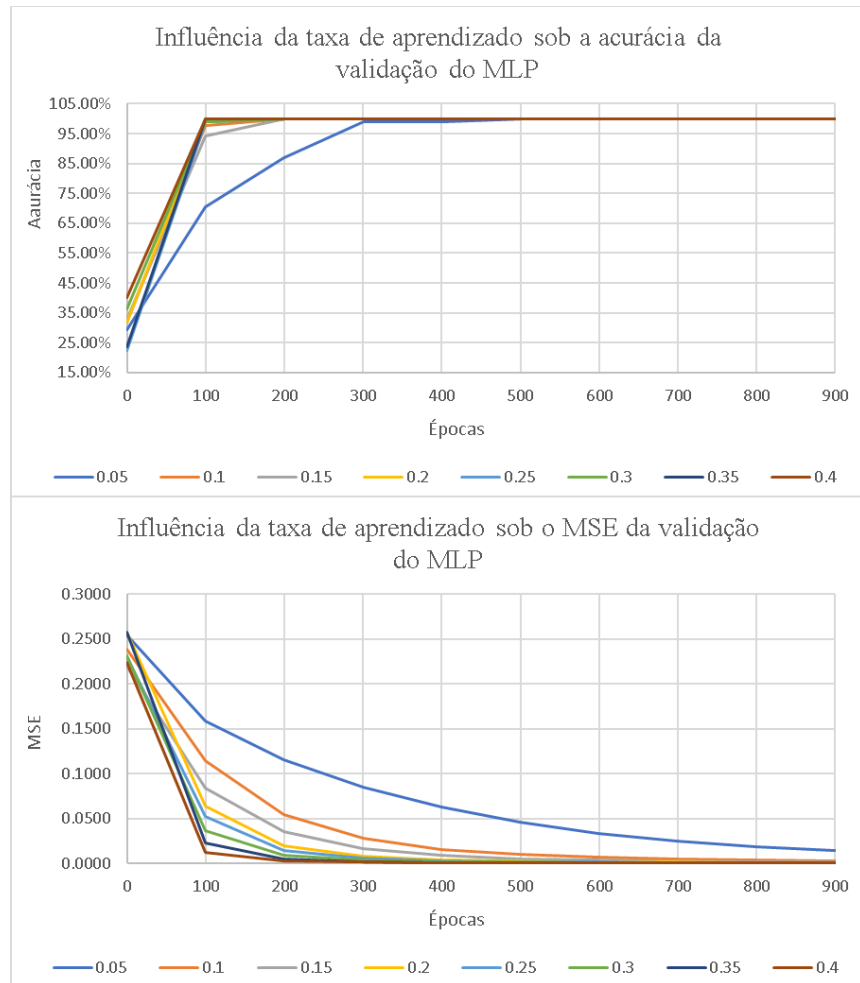
Com isso em mente, partiu-se para a análise da acurácia e MSE do algoritmo submetido à tais variações de taxa de aprendizado.

Figura 35: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.



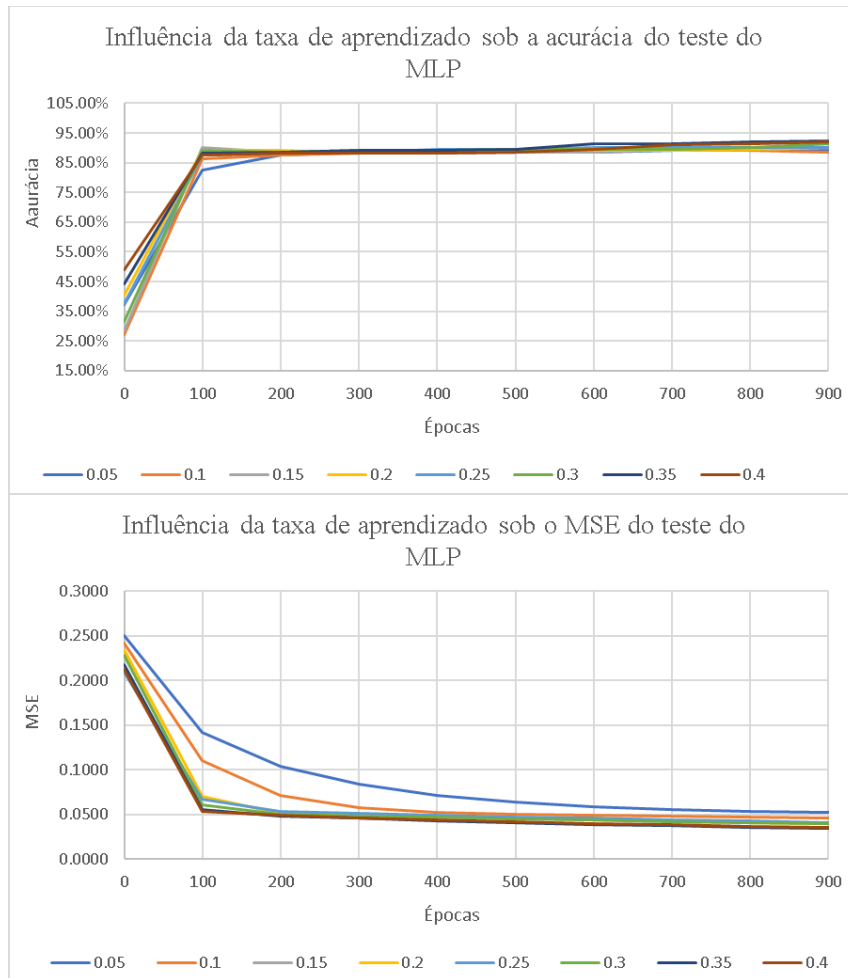
Fonte: Elaborado pelo autor (2023).

Figura 36: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.



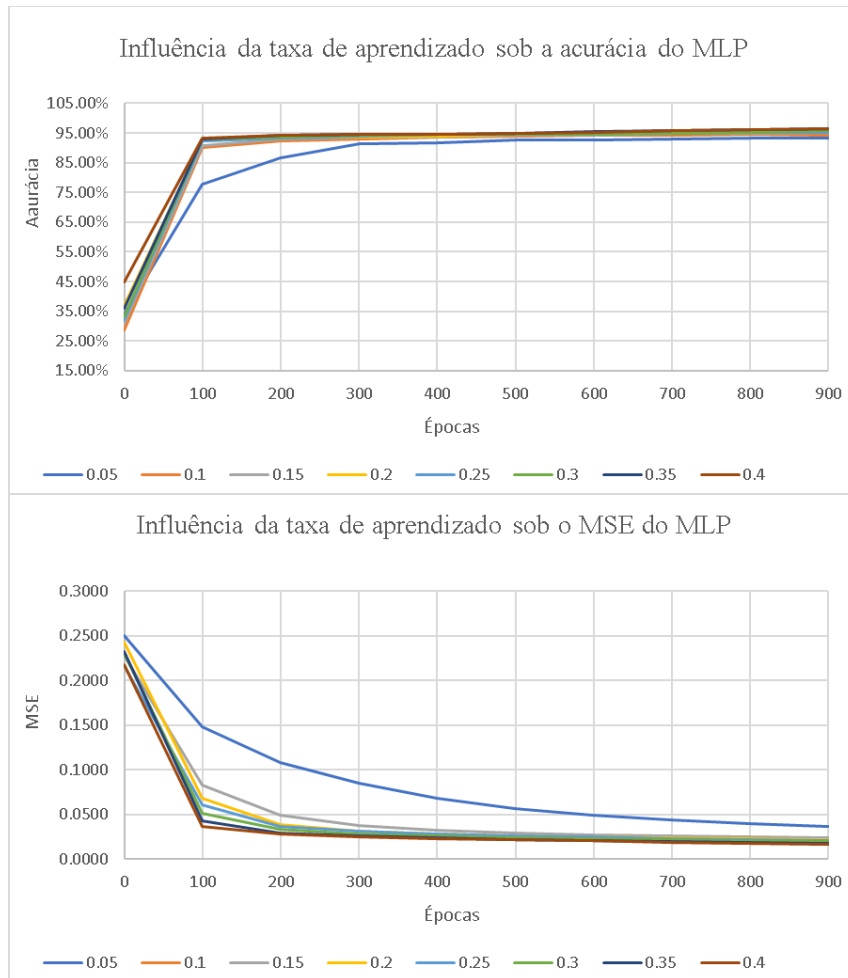
Fonte: Elaborado pelo autor (2023).

Figura 37: Influência da taxa de aprendizado do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

Figura 38: Influência da taxa de aprendizado do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

É pertinente entender da análise gráfica que, com o aumento da taxa de aprendizado, considerando as 1000 iterações propostas, a tendência é de que ocorra uma minimização do MSE e uma convergência antecipada do algoritmo, tendendo à uma estabilização mais precoce da acurácia.

À título de relacionar otimização com ótimos resultados, adotou-se uma taxa de aprendizado de 0,35, de modo a garantir uma máxima acurácia, rápida convergência e um mínimo MSE. Considerou-se, adicionalmente a todos esses fatores, um valor que não culminasse em um *overfitting* e aumentasse, por exemplo, a classificação errônea do subgrupo destinado a testagem do algoritmo.

3.5.3 Influência inicialização dos pesos sob o percentual de erros e acertos, acurácia e erro quadrático médio (MSE) do *Multi Layer Perceptron* (MLP)

A terceira análise teve como objetivo entender a influência da inicialização dos pesos em relação ao percentual de erros e acertos do código, bem como a acurácia e MSE.

Para isso, executou-se o código 5 vezes para cada uma das configurações observadas na Tabela 11 e, para cada um desses arranjos, foi obtida a média percentual de erros dos três subconjuntos avaliados.

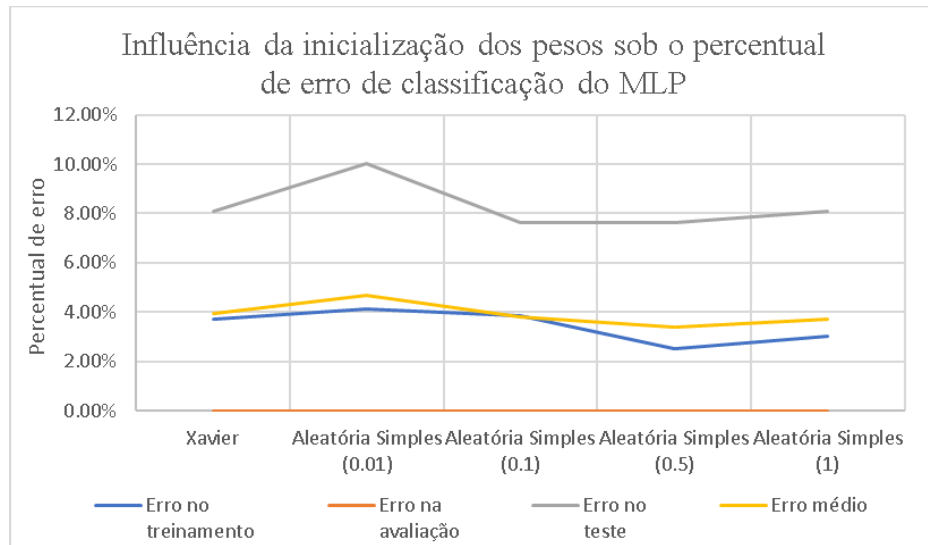
Tabela 11: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de inicialização dos pesos	Erro no treinamento	Erro na avaliação	Erro no teste	Erro médio
1000	0.35	6	Xavier	3.70%	0.00%	8.10%	3.93%
1000	0.35	6	Aleatória Simples (0.01)	4.11%	0.00%	10.00%	4.70%
1000	0.35	6	Aleatória Simples (0.1)	3.84%	0.00%	7.62%	3.82%
1000	0.35	6	Aleatória Simples (0.5)	2.52%	0.00%	7.62%	3.38%
1000	0.35	6	Aleatória Simples (1)	3.05%	0.00%	8.10%	3.71%

Fonte: Elaborado pelo autor (2023).

As inicializações de peso utilizadas foram inicialização de Xavier [10] e inicialização aleatória simples com fator multiplicativo. Com base nessa informação e nos dados extraídos da Tabela 11, obteve-se a Figura 23.

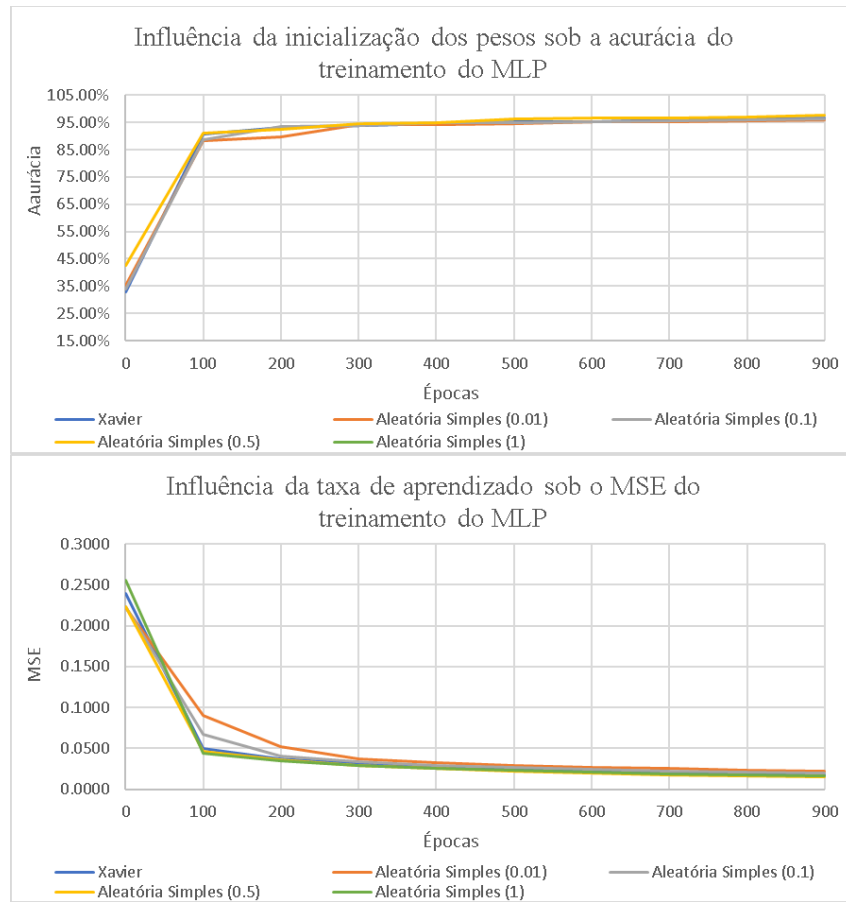
Figura 39: Influência do processo de inicialização dos pesos do MLP em relação ao percentual de classificações errôneas do algoritmo.



Fonte: Elaborado pelo autor (2023).

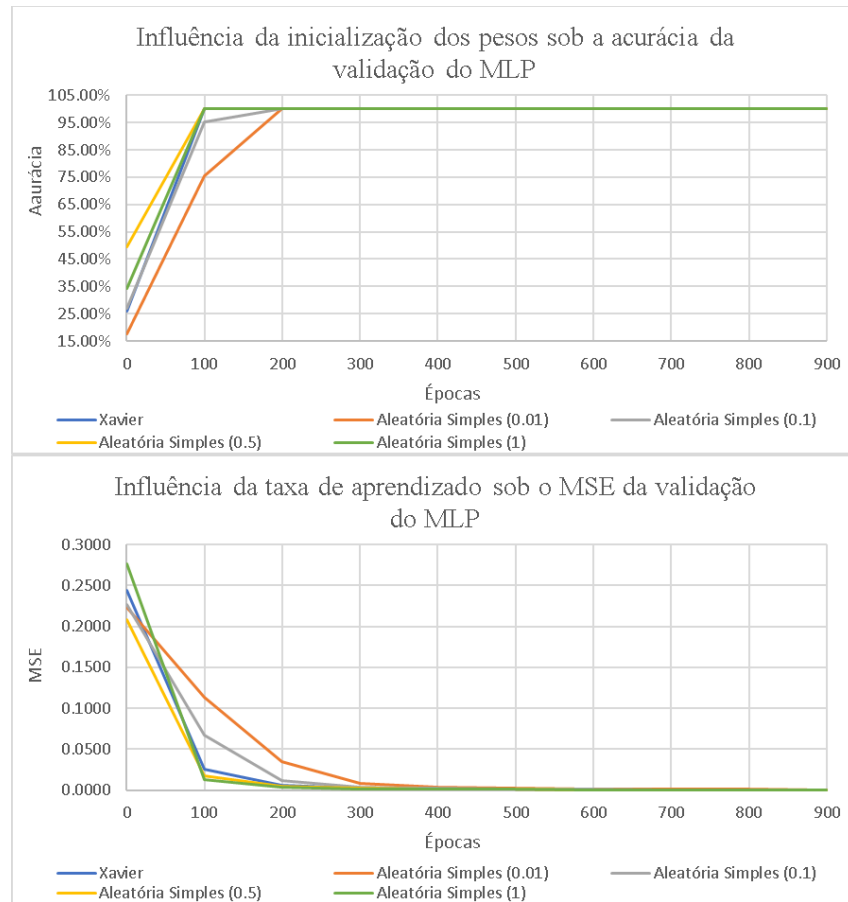
Analisando os dados, interpreta-se que os melhores resultados referentes à erro e acerto atrelam-se à inicialização aleatória simples com fator multiplicativo 0,5. À título de entender a variação da acurácia e MSE do MLP conforme altera-se o processo de inicialização dos pesos da rede neural, plotaram-se os gráficos abaixo.

Figura 40: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao treinamento do algoritmo.



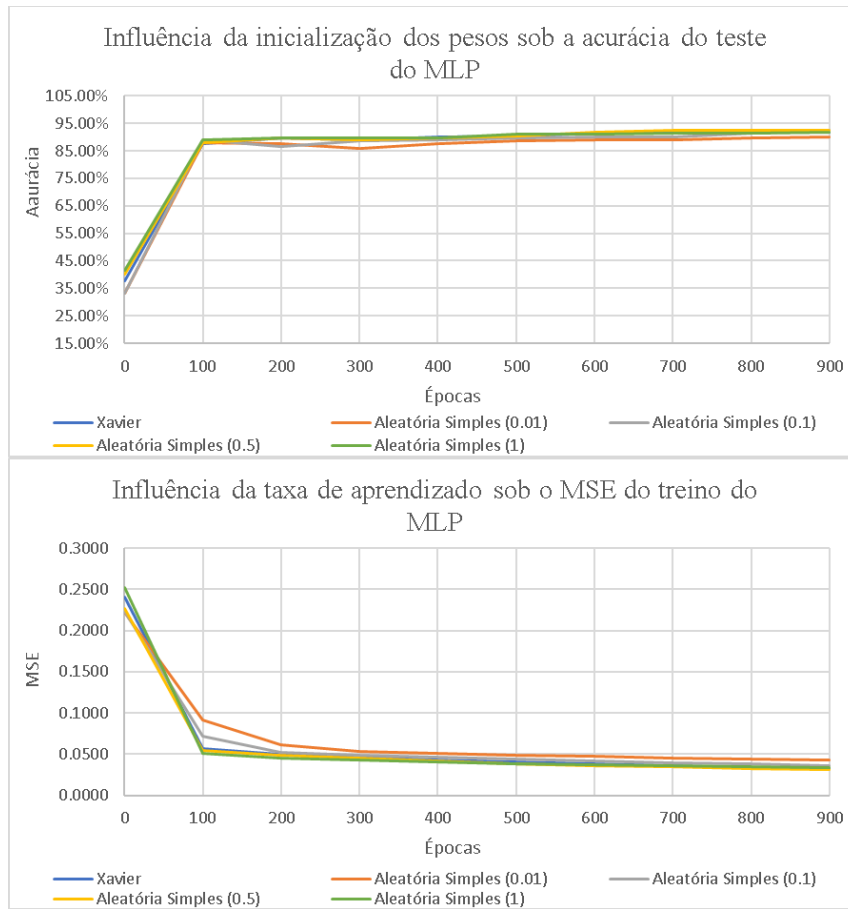
Fonte: Elaborado pelo autor (2023).

Figura 41: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado a validação do algoritmo.



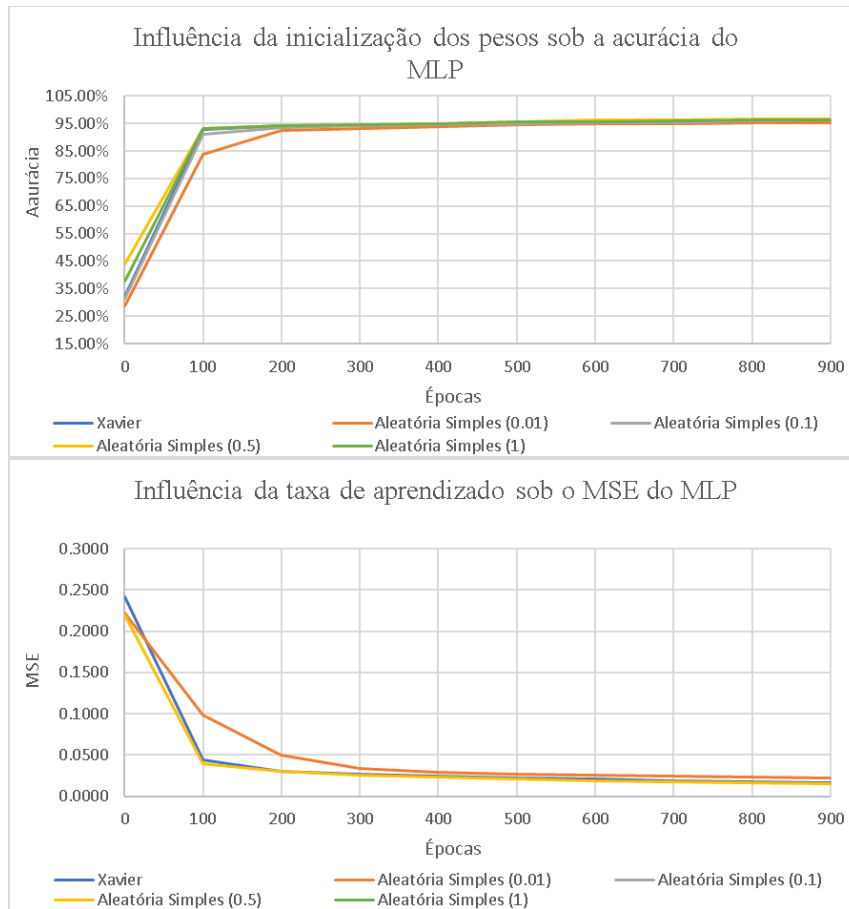
Fonte: Elaborado pelo autor (2023).

Figura 42: Influência do processo de inicialização dos pesos do MLP em relação à acurácia do subgrupo destinado ao teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

Figura 43: Influência do processo de inicialização dos pesos do MLP em relação à acurácia dos subgrupos de treinamento, validação e teste do algoritmo.



Fonte: Elaborado pelo autor (2023).

Observando os gráficos que relatam o comportamento da acurácia e MSE do algoritmo com as diferentes inicializações propostas, e considerando a análise já realizada quanto ao percentual de erro e acerto no processo classificatório do MLP, optou-se por seguir com a inicialização aleatória simples com fator multiplicativo de 0,5. A opção é justificada uma vez que tende a alcançar uma acurácia satisfatória ao mesmo tempo em que consegue minimizar significativamente o MSE no número de épocas definidos, combinando tais pontos com um fator de acerto classificatório relevante.

3.5.3 Análise final do *Multi Layer Perceptron* (MLP)

A fim de analisar o resultado do algoritmo com todos os parâmetros devidamente ajustados (Tabela 12), executou-se o código 5 vezes e, com base em tais execuções, analisou-se a resposta do MLP quanto à acurácia, MSE e probabilidade de erro de classificação.

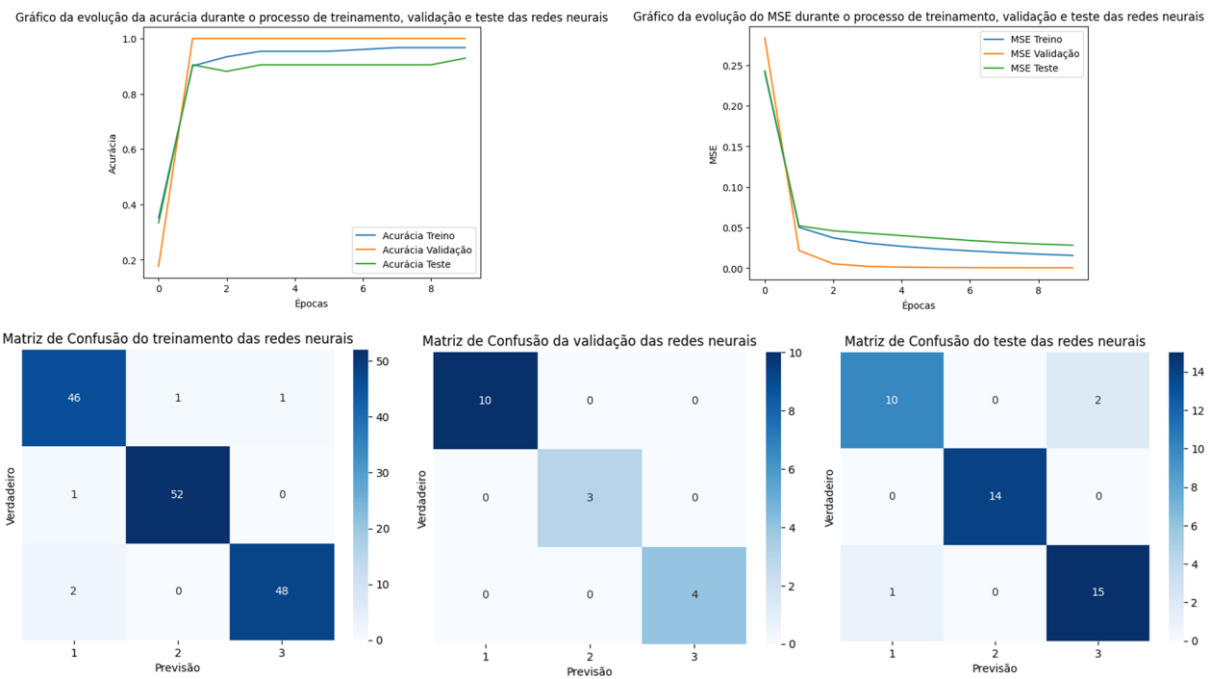
Tabela 12: Configuração final dos parâmetros do MLP.

Quantidade de Épocas	Taxa de Treinamento	Neurônios na camada oculta	Tipo de Inicialização dos pesos
1000	0.35	6	Aleatória Simples (0.5)

Fonte: Elaborado pelo autor (2023).

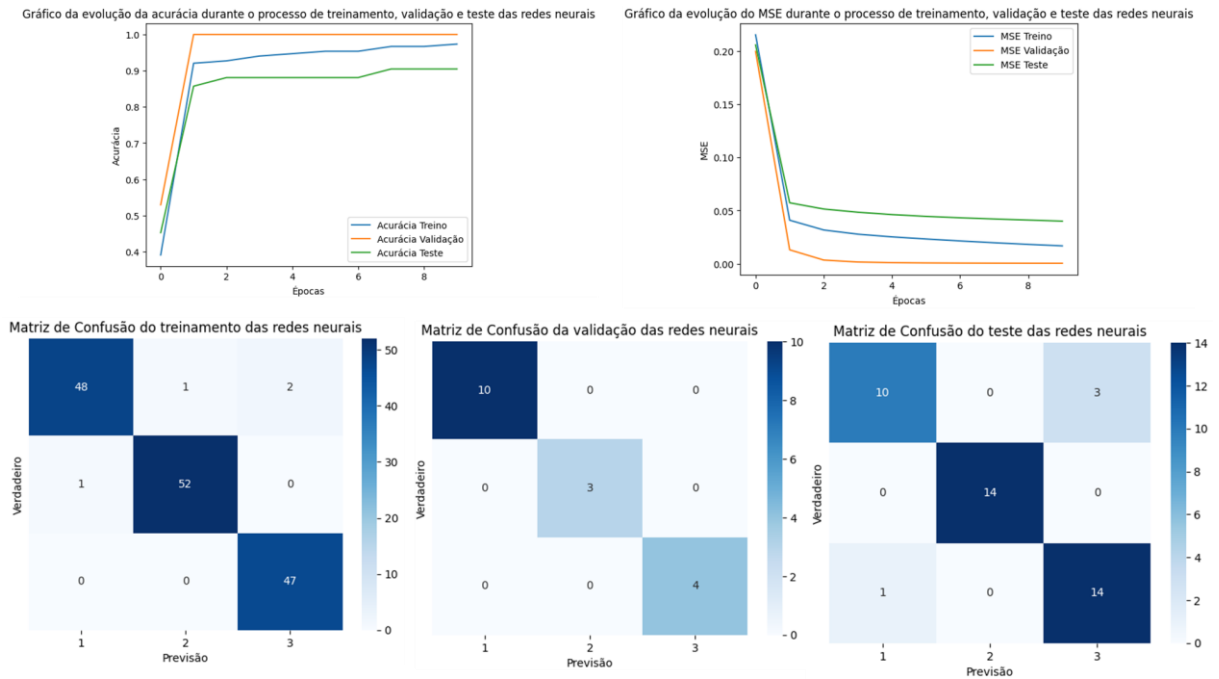
Os resultados de cada uma das simulações são vistos nas figuras abaixo.

Figura 44: Gráficos de acurácia e MSE e matrizes de confusão da primeira execução do algoritmo com as configurações finais dos parâmetros do MLP.



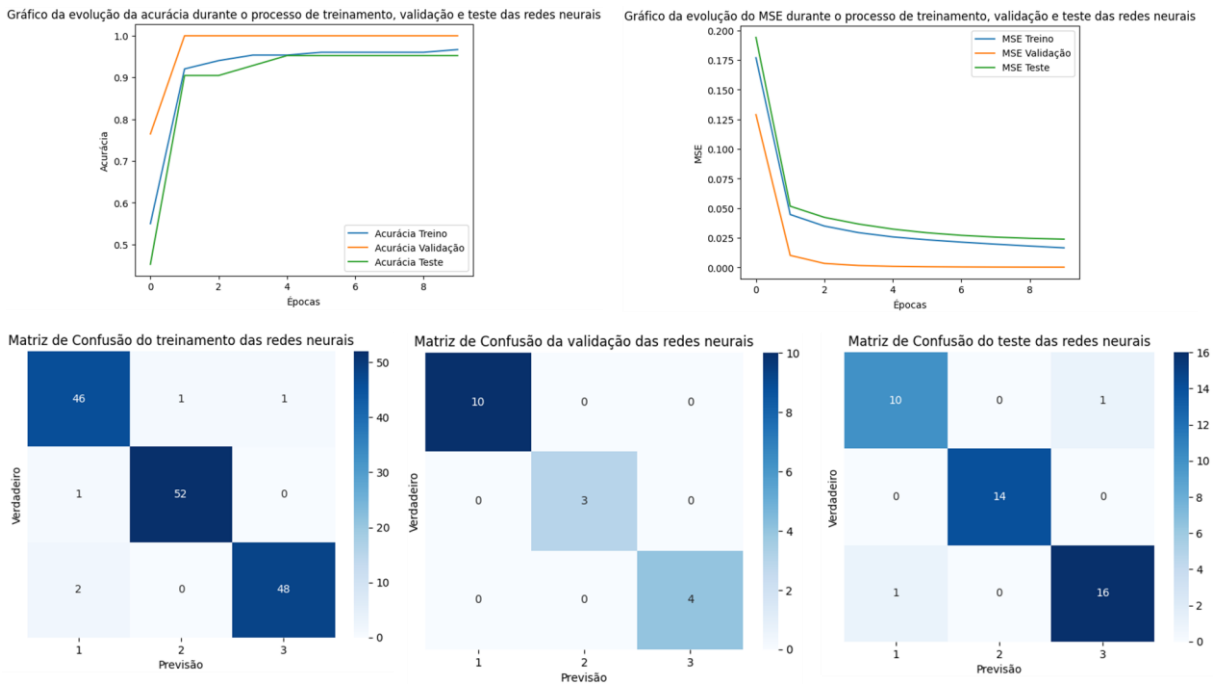
Fonte: Elaborado pelo autor (2023).

Figura 45: Gráficos de acurácia e MSE e matrizes de confusão da segunda execução do algoritmo com as configurações finais dos parâmetros do MLP.



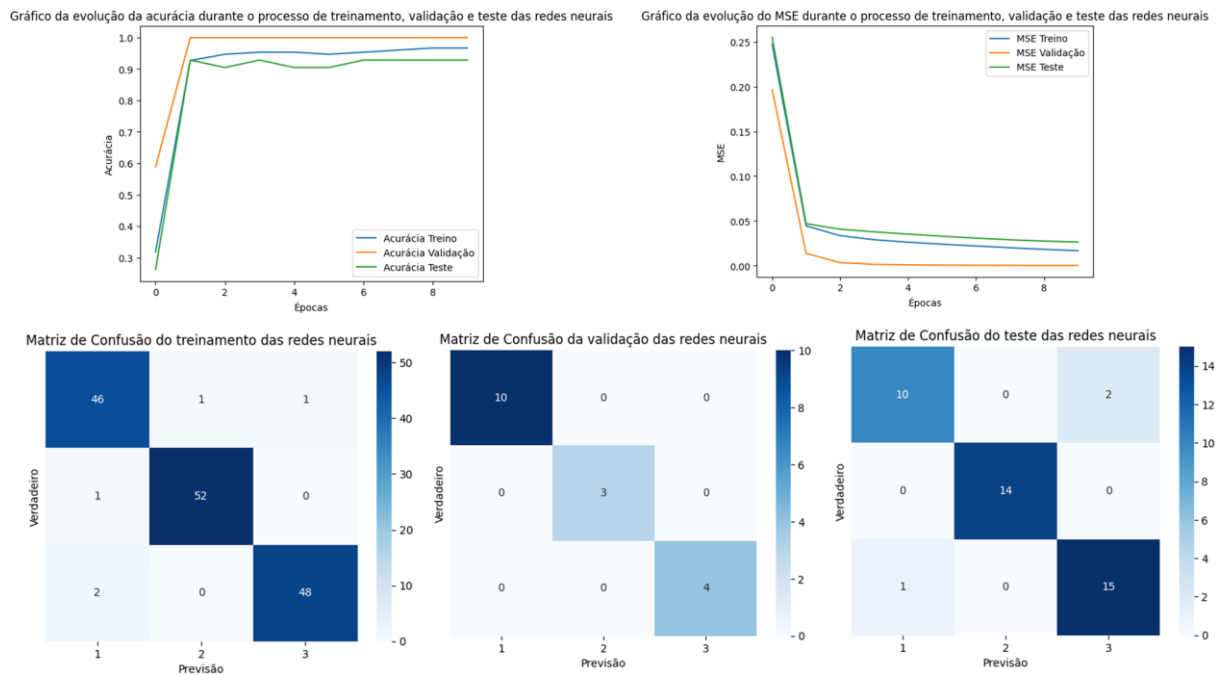
Fonte: Elaborado pelo autor (2023).

Figura 46: Gráficos de acurácia e MSE e matrizes de confusão da terceira execução do algoritmo com as configurações finais dos parâmetros do MLP.



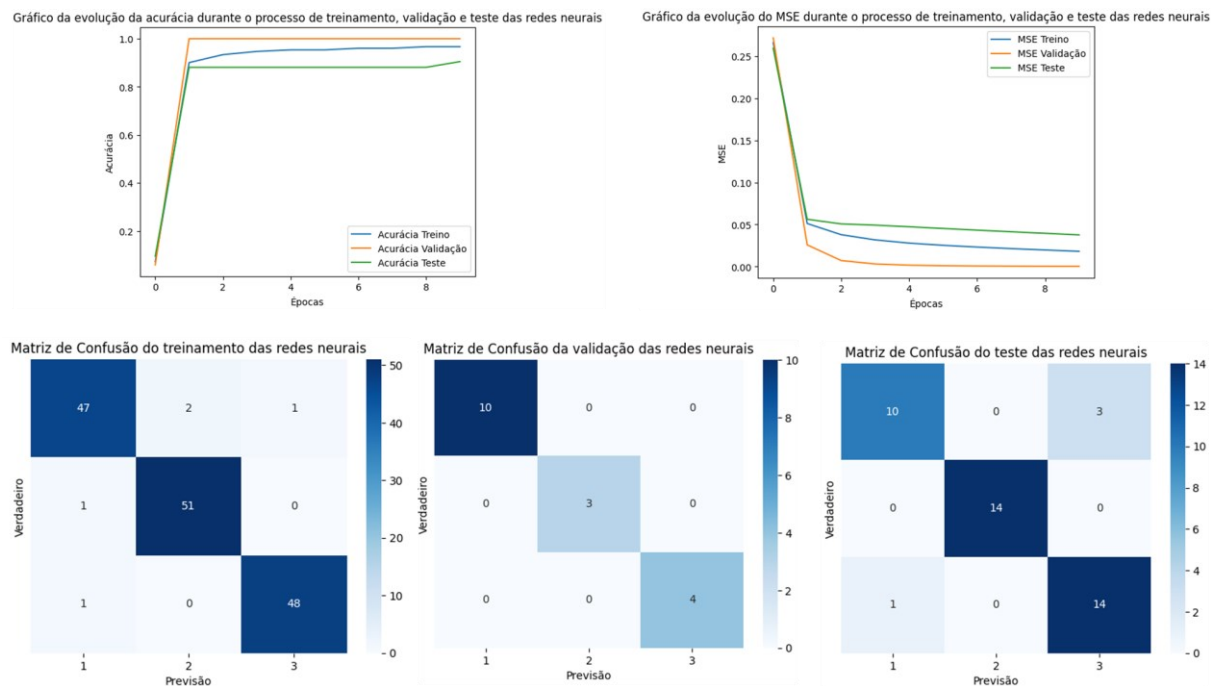
Fonte: Elaborado pelo autor (2023).

Figura 47: Gráficos de acurácia e MSE e matrizes de confusão da quarta execução do algoritmo com as configurações finais dos parâmetros do MLP.



Fonte: Elaborado pelo autor (2023).

Figura 48: Gráficos de acurácia e MSE e matrizes de confusão da quinta execução do algoritmo com as configurações finais dos parâmetros do MLP.



Fonte: Elaborado pelo autor (2023).

Analisando os resultados apresentados acima, tem-se que no pior caso executado dentre as 5 simulações, a acurácia de treinamento atinge um valor de próximo de 95% quando se aproxima de 300 iterações. A validação, por sua vez, é maximizada em 100% próximo às 100 iterações. Por fim, o subconjunto de teste tem sua pior performance quando atinge o patamar próximo à 90% próximo à 1000 iterações.

O MSE do conjunto de treinamento, na pior execução, atinge um valor próximo a 0,025 nas últimas iterações da MLP. O grupo de validação, por sua vez, é minimizado, atingindo um valor muito próximo à 0, quando é iterado cerca de 300 vezes. O teste, por fim, atinge um patamar que se aproxima de 0,4 na 1000^a iteração.

Quando se avalia a matriz de confusão, constituída pelos erros e acertos classificatórios, tem-se, nas piores execuções, os seguintes percentuais: 3,31% no conjunto de treinamento, 0% no processo de validação e 9,52% no teste do MLP.

Conclui-se, portanto, que os parâmetros estabelecidos são consideravelmente suficientes para possibilitar a distinção e classificação das sementes de trigo, conforme propostos.

4. Link para acesso dos algoritmos

Nesta seção, encontram-se os links de acesso aos algoritmos que solucionam os exercícios propostos.

4.1 Link de acesso ao exercício 1

<https://colab.research.google.com/drive/1A20kfkiBuzB5ZSsJb20kRxfAglCGm3Ps?usp=sharing>

4.2 Link de acesso ao exercício 2

<https://colab.research.google.com/drive/1jrH7SHOP5R9F7WUeJKZpwfFVxIUSjAEL?usp=sharing>

4.3 Link de acesso ao exercício 3

<https://colab.research.google.com/drive/19G5SzBh-B4dyk7A8q-fPxHoRaSfafGdW?usp=sharing>

5. Referências bibliográficas

- [1] PYTHON SOFTWARE FOUNDATION. Python. Disponível em: <https://www.python.org/>.
- [2] GOOGLE. Colaboratory. Disponível em: <https://colab.research.google.com/>.
- [3] NUMPY. NumPy. San Francisco: NumPy, 2023. Disponível em: <https://numpy.org/.7>.
- [4] PANDAS. Disponível em: <https://pandas.pydata.org/>.
- [5] MATPLOTLIB. Matplotlib: Visualization with Python: Disponível em: <https://matplotlib.org/>.
- [6] SCIKIT-LEARN. Scikit-learn: Machine Learning in Python. Disponível em: <https://scikit-learn.org/stable/>.
- [7] PRETTYTABLE. Prettytable 3.7.0. Disponível em: <https://pypi.org/project/prettytable/>.
- [8] SEABORN. Seaborn: statistical data visualization. Disponível em: <https://seaborn.pydata.org/>.
- [9] BREVE, Fabricio. Redes Neurais Artificiais - parte 2. 20 de abril de 2023. Apresentação de slides. UNESP.
- [10] GLOROT, Xavier; BENGIO, Yoshua. Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. PMLR, 2010. p. 249-256.
- [11] MICROSOFT EXCEL. Excel. Disponível em: https://www.microsoft.com/pt-br/microsoft-365/p/excel/CFQ7TTC0HR4R?activetab=pivot:overviewtab&ef_id=_k_Cj0KCQjwmN2iBhCrARIsAG_G2i7jC6AnQCMjC3Ibbx9arVkF_wfSFT0HtQh2xL5PersUxSarjkaw74IaApHZEALw_wcB_k_&OCID=AIDcmm409lj8ne_SEM__k_Cj0KCQjwmN2iBhCrARIsAG_G2i7jC6AnQCMjC3Ibbx9arVkF_wfSFT0HtQh2xL5PersUxSarjkaw74IaApHZEALw_wcB_k_&gclid=Cj0KCQjwmN2iBhCrARIsAG_G2i7jC6AnQCMjC3Ibbx9arVkF_wfSFT0HtQh2xL5PersUxSarjkaw74IaApHZEALw_wcB

*O trabalho foi embasado integralmente e desenvolvido com o suporte das notas de aula [9].