



**MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO
TOCANTINS CAMPUS DIANÓPOLIS
CURSO SUPERIOR DE LICENCIATURA EM COMPUTAÇÃO**

Aparecida de Cássia Rodrigues Ávila

**COMPARATIVO DE DESEMPENHO DOS ALGORITMOS REGRESSÃO
LOGÍSTICA E REDES NEURAS ARTIFICIAIS NA CLASSIFICAÇÃO DE IMAGENS
UTILIZANDO A MATRIZ DE CONFUSÃO NA FERRAMENTA *ORANGE DATA
MINING***

**Dianópolis
2023**



**MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO
TOCANTINS CAMPUS DIANÓPOLIS
CURSO SUPERIOR DE LICENCIATURA EM COMPUTAÇÃO**

Aparecida de Cássia Rodrigues Ávila

COMPARATIVO DE DESEMPENHO DOS ALGORITMOS REGRESSÃO LOGÍSTICA E REDES NEURAIS ARTIFICIAIS NA CLASSIFICAÇÃO DE IMAGENS UTILIZANDO A MATRIZ DE CONFUSÃO NA FERRAMENTA *ORANGE DATA MINING*

Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do Título de Licenciado em Computação do Curso Superior de Licenciatura em Computação do Instituto Federal do Tocantins, *Campus* Dianópolis.

Orientador: Me. Robert Mady Nunes.

**Dianópolis
2023**

**Dados Internacionais de Catalogação na Publicação (CIP)
Bibliotecas do Instituto Federal do Tocantins**

Á958c Ávila, Aparecida de Cássia Rodrigues

Comparativo de desempenho dos algoritmos regressão logística e redes neurais artificiais na classificação de imagens utilizando a matriz de confusão na ferramenta Orange Data Mining / Aparecida de Cássia Rodrigues Ávila. – Dianópolis, TO, 2023.

74 p.

Trabalho de Conclusão de Curso (Licenciatura em Computação) – Instituto Federal de Educação, Ciência e Tecnologia do Tocantins, Campus Dianópolis, Dianópolis, TO, 2023.

Orientador: Robert Mady Nunes

1. Machine Learning. 2. Matriz de Confusão. 3. Orange Data Mining.
I. Nunes, Robert Mady. II. Título.

CDD 004

A reprodução total ou parcial, de qualquer forma ou por qualquer meio, deste documento é autorizada para fins de estudo e pesquisa, desde que citada a fonte.

Elaborado pelo sistema de geração automática de ficha catalográfica do IFTO com os dados fornecidos pelo(a) autor(a).

APARECIDA DE CÁSSIA RODRIGUES ÁVILA

COMPARATIVO DE DESEMPENHO DOS ALGORITMOS REGRESSÃO LOGÍSTICA E REDES NEURAS ARTIFICIAIS NA CLASSIFICAÇÃO DE IMAGENS UTILIZANDO A MATRIZ DE CONFUSÃO NA FERRAMENTA *ORANGE DATA MINING*

Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do Título de Licenciado em Computação do Curso Superior de Licenciatura em Computação do Instituto Federal do Tocantins, *campus* Dianópolis.

Orientador: Me. Robert Mady Nunes.

Aprovado em: ____/____/2023

Banca Avaliadora

Me. Robert Mady Nunes (Orientador)

IFTO – *campus* Dianópolis

Me. Válber Sardi Lopes

IFTO – *campus* Dianópolis

Esp. Cássia Cristina da Silva.

IFTO – *campus* Dianópolis

AGRADECIMENTOS

A Deus ao qual devo toda a minha vida e salvação, o seu consolo e presença por meio do Espírito Santo tem sido constante em minha vida, ao Senhor Jesus Cristo, o qual intercede por mim incansavelmente, sei que não dorme nem cochila o Guarda de Israel. A Ele toda honra e toda a glória. A todos os Professores que dedicam seu tempo a ensinar coisas que nos serão úteis na nossa vida, em especial ao Robert Mady. À Professora Cássia Cristina. À minha família. À Igreja Cristã Maranata, que oraram por mim em mais essa vitória.

“Bem-aventurado o homem que não anda segundo o conselho dos ímpios, nem se detém no caminho dos pecadores, nem se assenta na roda dos escarnecedores. Antes tem o seu prazer na lei do Senhor, e na sua lei medita de dia e de noite. Pois será como a árvore plantada junto a ribeiros de águas, a qual dá o seu fruto no seu tempo; as suas folhas não cairão, e tudo quanto fizer prosperará.”

Salmos 1:1-3

RESUMO

O presente trabalho treinou dois modelos de Aprendizado de Máquina, uma Regressão Logística e uma Rede Neural, para classificar imagens de três pintores. Foram utilizadas várias abordagens de Técnicas de Amostragens de Dados como a Validação Cruzada, Estratificação e Amostragem Aleatória. Foi utilizada uma pequena e diversificada quantidade de imagens e mudanças na arquitetura dos modelos durante os testes. O desempenho dos Algoritmos foi medido pela ferramenta Matriz de Confusão e toda a parte prática dos experimentos foram feitas na Ferramenta *Orange Data Mining*, uma caixa de ferramentas de mineração de dados gratuita própria para análise de dados. A Metodologia foi construída por meio de pesquisas bibliográficas e da leitura da documentação da ferramenta, ambas contribuíram para a geração de um Fluxo de trabalho, no qual os experimentos foram feitos baseados e executados. Foi utilizada uma quantidade mínima de imagens, devido a complexidade do problema e aos resultados alcançados nos testes iniciais. Os Experimentos demonstraram que as Técnicas de Amostragem de Dados, a arquitetura dos algoritmos, a quantidade de treinos e a diversificação das imagens são fatores relevantes para o desempenho dos modelos de aprendizagem de máquina, mesmo que o conjunto de dados seja limitado. A literatura já havia mencionado que o desempenho da Regressão Logística poderia ser superior ao de uma Rede Neural, tendo como base um conjunto de dados de pequena complexidade, fato que foi comprovado neste trabalho. A Matriz de Confusão se mostrou como uma ferramenta capaz de medir o desempenho dos algoritmos, dos modelos e das Técnicas de Amostragens de Dados, conforme também mencionado na literatura. O trabalho também concluiu que a Regressão Logística pode ser uma opção viável para a classificação de imagens em uma situação de pequena complexidade do problema envolvido e de recursos computacionais limitados.

Palavras-chave: *Machine learning*. Matriz de confusão. *Orange Data Mining*.

ABSTRACT

The present work trained two Machine Learning models, a Logistic Regression and a Neural Network, to classify images of three painters. Several approaches of Data Sampling Techniques were used, such as Cross Validation, Stratification and Random Sampling. A small and diverse amount of images and changes in the architecture of the models were used during the tests. The performance of the Algorithms was measured by the Confusion Matrix tool and all the practical part of the experiments were done in the Orange Data Mining Tool, a free data mining toolbox for data analysis. The Methodology was built through bibliographical research and reading the tool's documentation, both of which contributed to the generation of a Workflow, on which the experiments were based and executed. A minimal amount of images was used, due to the complexity of the problem and the results achieved in the initial tests. The Experiments demonstrated that the Data Sampling Techniques, the architecture of the algorithms, the amount of training and the diversification of the images are relevant factors for the performance of the machine learning models, even if the data set is limited. The literature had already mentioned that the performance of Logistic Regression could be superior to that of a Neural Network, based on a data set of low complexity, a fact that was proven in this work. The Confusion Matrix proved to be a tool capable of measuring the performance of algorithms, models and Data Sampling Techniques, as also mentioned in the literature. The work also concluded that Logistic Regression can be a viable option for image classification in a situation of low complexity of the problem involved and limited computational resources.

Keywords: Machine learning. Confusion matrix. Orange Data Mining.

LISTA DE ILUSTRAÇÕES

Figura 1 – Estrutura de uma Rede Neural simples	28
Figura 2 – Arquitetura de uma Rede Neural Convolutiva	29
Figura 3 – Matriz de Confusão Binária	31
Figura 4 – Matriz de Confusão Multiclasse.....	31
Figura 5 – Gráfico Curva ROC.....	34
Figura 6 – Interface da Orange Data Mining.....	35
Figura 7 – Workflow	38
Figura 8 – Ferramenta de captura de imagens Image Downloader.....	40
Figura 9 – Import Images	40
Figura 10 – Import Images	41
Figura 11 – Image Viewer	41
Figura 12 – Data Table	42
Figura 13 – Image Embedding	43
Figura 14 – <i>Test and Score</i>	47
Figura 15 – Matriz de Confusão	48
Figura 16 – Image Grid.....	49
Figura 17 – Desempenho da Classe Pintor1 do Experimento 2	53
Figura 18 – Desempenho da Classe Pintor2 do Experimento 2	53
Figura 19 – Desempenho da Classe Pintor3 do Experimento 2	54
Figura 20 – Desempenho da Classe Pintor1 do Experimento 5	58
Figura 21 – Desempenho da Classe Pintor2 do Experimento 5	58
Figura 22 – Desempenho da Classe Pintor3 Experimento 5	59
Figura 23 – Repetição do Experimento2 do <i>Logistic Regression</i>	61
Figura 24 – Repetição do Experimento 5 do Neural Network	62
Figura 25 - Workflow com as configurações finais	63
Figura 26 - Imagens separadas por semelhanças Rede Experimento 5- <i>Neural Network</i>	65
Figura 27 - Imagens separadas por semelhanças Rede Experimento 2- <i>Logistic Regression</i>	65

LISTA DE TABELAS

Tabela 1 – Configurações feitas no <i>Logistic Regression</i> e no <i>Test and Score</i>	50
Tabela 2 - Métricas extraídas do desempenho do Aprendiz <i>Logistic Regression</i> , de acordo com as técnicas de amostragem de dados.....	52
Tabela 3 - Matriz de Confusão do Experimento 2.....	52
Tabela 4 – Configurações feitas no <i>Neural Network</i> e no <i>Test and Score</i>	56
Tabela 5 – Métricas extraídas do desempenho do Aprendiz <i>Neural Network</i> , de acordo com as técnicas de amostragem de dados	57
Tabela 6 – Matriz de Confusão Experimento 5	57
Tabela 7- Experimentos 2 e 5	64

LISTA DE SIGLAS

ANN - Artificial Neural Networks

CI - Classificação De Imagens

CNN - Neural Network Convolutional

DP - Deep Learning

FN - False Negative

FP - False Positive

IA - Inteligência Artificial

JPG - *Joint Photographic Experts Group*

LASSO - Least Absolute Shrinkage And Selection Operator

LR - Logistic Regression

ML - Machine Learning

MLP - Multi-Layer Perceptron Backpropagation

RELU - Unidade Linear Retificada

RNA - Rede Neural Artificial

ROC - Receiver Operating Characteristic

SVM - Support Vector Machine

TP - True Positive

TN - True Negative

SUMÁRIO

1. INTRODUÇÃO	12
1.1. PROBLEMATIZAÇÃO	14
1.2. JUSTIFICATIVA	15
1.3. OBJETIVOS	19
1.3.1. Objetivo Geral	19
1.3.2. Objetivos Específicos	19
2. REFERENCIAL TEÓRICO	20
2.1. INTELIGÊNCIA ARTIFICIAL	20
2.2. CLASSIFICAÇÃO DE IMAGENS	21
2.2.1. Os Principais Métodos de Classificação de Imagens	21
2.2.2. Técnicas Utilizadas na Classificação	22
2.2.3. Alguns Classificadores Imagens	22
2.3. MACHINE LEARNING	24
2.3.1. Tipos de Aprendizado Utilizados em Machine Learning	24
2.4. REGRESSÃO LOGÍSTICA	26
2.5. REDES NEURAIS	27
2.6. MATRIZ DE CONFUSÃO.....	30
2.7. ORANGE DATA MINING	34
3. PROCEDIMENTOS METODOLÓGICOS	36
3.1. FONTES BIBLIOGRÁFICAS.....	36
3.2. AMBIENTE DE EXECUÇÃO E EQUIPAMENTOS UTILIZADOS.....	36
3.3. TIPO DA PESQUISA.....	37
3.4. CLASSIFICAÇÃO DE IMAGENS	37
3.5. WORKFLOW - FLUXO DE TRABALHO REALIZADO	37
4. RESULTADO E DISCUSSÕES	39
4.1. APRENDIZES.....	44
4.2. TESTE E PONTUAÇÃO	45
4.3. REFLEXÕES E EXPERIMENTOS REALIZADOS.....	49
5. CONCLUSÃO	66
REFERÊNCIAS	70

1. INTRODUÇÃO

“A vida de qualquer ser vivo é impossível se não tiver a capacidade de diferenciar entre várias coisas, objetos, cheiros, gostos, cores, etc” (CHANDRA; BEDI, 2018, p. 1). Várias decisões empresariais ou da vida cotidiana dependem de escolhas baseadas em informações já definidas ou de previsões feitas de acordo com acontecimentos passados ou ainda, de informações recorrentes. Pode-se comprar um imóvel baseado nos preços dos imóveis vizinhos ou pode-se deduzir que o valores desses imóveis tenderão a subir de acordo com determinada política econômica ou mais condições envolvidas. É possível, ao olhar um animal dizer se ele é um leão ou é um gato, baseado em conhecimentos armazenados e apreendidos ao longo da vida. Todas as situações acima dependem de aprender, posteriormente separar os resultados e então, tomar uma decisão.

Classificar ou separar baseado em pequenas quantidades de dados é algo que pode ser feito pelo ser humano, porém existem conhecimentos que só são descobertos por meio de grandes quantidades de informações e por extensos bancos de dados, para esses casos e para previsões complexas são usadas várias técnicas de Mineração de Dados, Estatística e *Machine Learning* - Aprendizado de Máquina com algoritmos capazes de classificar e aprender com os dados de forma relativamente rápida e com precisão.

O campo de classificação de imagens expandiu na última década (CHANDRA; BEDI, 2018). Essa expansão se deu devido ao aumento de dados disponíveis e a melhoria dos processos de computação. A medicina é um ramo que está se aperfeiçoando cada vez mais com o uso de classificação de imagens por meio de *Machine Learning* - ML e com o uso de Redes Neurais Profundas, essas modalidades de uso de Inteligência Artificial, tem sido muito utilizadas no diagnóstico de doenças como o câncer (JANOWCZYK, A.; MADABHUSHI, 2016).

Existem vários algoritmos de aprendizagem de máquina voltados para a classificação de imagens, como o *Random forest* (Florestas aleatórias), *Support Vector Machine* (Máquinas Vetoriais de Suporte), *K-NN* (K- vizinhos mais próximos), Redes Neurais (DHINGRA et al., 2019) e Regressão Logística (PATHAN et e al., 2018).

Nesse trabalho propôs-se analisar o desempenho de dois Algoritmos utilizados na classificação: Regressão Logística e Redes Neurais, e mensurar o desempenho

dessas ferramentas em identificar imagens dos pintores Jean Claude Monet, Vincet Van Gogh e Pablo Picasso, por meio da Matriz de Confusão na ferramenta de Mineração de Dados *Orange Data Mining*.

A Regressão Logística é um classificador linear, discriminativo que trabalha calculando a probabilidade de um elemento pertencer a determinada classe (TSANGARATOS; ILIA, 2016). É um dos algoritmos mais comuns usados e estatísticas e análise de dados discretos, problemas de classificação e regressão (TOMASEVIC et al., 2019).

Uma Rede Neural é uma ferramenta de aprendizado de máquina baseada na arquitetura e funcionamento do cérebro humano, cuja estrutura mais simples é composta por entrada, camada oculta e saída (KEIRON; NASH, 2015).

A Matriz de Confusão é uma maneira de observar de forma resumida o desempenho de um classificador (RASCHKA, 2014) e os métodos que têm como base a Matriz de Confusão, são os mais recomendados para se analisar o desempenho de classificadores (SMITS et e al., 1999), sendo esses, os motivos pelos quais foi adotada neste trabalho.

Por fim, a motivação da escolha da *Orange Data Mining* ocorreu por ser uma ferramenta gratuita, por possuir muitas ferramentas inclusas que são úteis para a Análise de Dados e de imagens, por possuir uma documentação bastante explicativa e Fluxos de Trabalho bastante intuitivos.

1.1. PROBLEMATIZAÇÃO

Existem ferramentas capazes de analisar grandes volumes de imagens em tempo relativamente curto, se comparado à capacidade humana e com detalhes que o olho humano não consegue captar. Cada algoritmo possui suas especificidades e seus ganhos e perdas de informação e necessidade maior ou menor de processamento. A *Orange Data Mining* possui vários algoritmos para análise e classificação de dados, dentre eles Regressão Logística e Redes Neurais Artificiais. Diante desse cenário surge a seguinte problemática: Entre o algoritmo de Regressão Logística e o algoritmo Redes Neurais Artificiais, qual se mostra mais eficiente na aprendizagem e classificação de imagens?

1.2. JUSTIFICATIVA

Classificação de imagens, está por trás de muitos serviços importantes para a sociedade, muitos destes serviços requerem confiabilidade para serem úteis ao seu fim. A Classificação de Imagens é um objeto importante do campo da Visão Computacional, que é uma ciência que procura visualizar os objetos tal qual a visão humana (MARENGONI; STRINGHINI, 2010).

Uma máquina que consegue ver e distinguir as imagens da forma mais assertiva, pode ser usada com mais confiança, e proporcionar melhor condição de aproveitamento para os usuários, é capaz ajudar a salvar vidas, como é o caso dos diagnósticos, colaborar com a educação, por meio de ferramentas mais interativas, ser usada na segurança, e muitas outras aplicações.

Para que uma máquina possa executar a tarefa de identificar e distinguir os objetos dentre os demais, são necessários recursos estatísticos, conjuntos de dados, abordagens probabilísticas, recursos computacionais sofisticados (OLIVEIRA et al., 2021) e algoritmos treinados e testados por ferramentas apropriadas.

Há vários algoritmos de aprendizagem de máquina que podem ser utilizados na classificação de imagens, para este trabalho foram escolhidas a Regressão Logística e a Rede Neural.

A Regressão Logística é amplamente usada em tarefas de classificação e regressão. O modelo de Regressão Logística tem a desvantagem de ajustes dos dados, ou seja, começa a decorar os dados de treino, em vez de aprender (TSANGARATOS; ILIA, 2016). A Regressão é um algoritmo de arquitetura simples e usado para problemas mais simples.

Uma outra condição verificada com o algoritmo é que quando há um aumento do conjunto de dados de treinamento a AUC (Área sob a Curva) aumenta, porém, com o aumento dos dados, aumenta a complexidade do modelo, e isso faz com que a Área sob a Curva diminua (TSANGARATOS; ILIA, 2016).

A Área sob a Curva é uma métrica que demonstra a relação entre a Especificidade e a Sensibilidade, de acordo com os limites de decisão de classificação do aprendiz e na visualização do gráfico, quanto mais a curva do modelo estiver do canto esquerdo do gráfico, melhor será esse modelo em contraposição do que estiver mais à direita (SILVA, 2019).

Especificidade é a capacidade do modelo em detectar resultados negativos enquanto que a Sensibilidade é a medida que mostra a proporção de amostras corretamente classificadas como positivas em relação ao total de amostras que são verdadeiramente positivas.

O algoritmo Regressão Logística, que é uma ferramenta probabilística e trabalha por meio de escolher os melhores coeficientes para poder classificar um objeto, possui desempenho similar a uma Rede Neural e a uma Máquina de Vetor de Suporte (Dreiseitl et al., 2001).

As redes neurais possuem um poder de aprendizado eficiente, uma vez que os neurônios trabalham semelhante a processadores paralelos, aprendendo e compartilhando informações e são ferramentas importantes para solução de problemas não lineares, problemas mais complexos (PATHAN et al, 2018).

As redes neurais propostas por (MCCULLOCH; PITTS, 1943), são um modelo computacional, projetados de forma matemática que imita processos biológicos no cérebro, e também é composto por inteligência artificial. Quanto mais os neurônios dispararam, mais aprendem (HEBB; WILEY, 1949). Esse fator mostra que a quantidade de treinos é importante para o aprendizado da rede, pois os neurônios disparam por meio da ação da função de ativação, demonstrando que aprenderam.

Uma das desvantagens em utilizar as redes neurais é que elas necessitam de quantidades significativas de dados para serem treinadas (LEK; GUÉGAN, 1999) e, conseqüentemente exigem maior poder de recursos computacionais (PAOLA; SCHOWENGERDT, 1995) a quantidade de tempo para treinar os dados também é maior, embora, após treinados, seja mais rápido o processamento (HEERMANN; KHAZENIE, 1992).

Muitas redes neurais usadas para classificação de imagens são treinadas com milhões de dados disponíveis, como o *ImageNet* (orangedatamining, 2023). As redes, na sua versão convolucional são amplamente usadas na classificação de imagens (JANOWCZYK; MADABHUSHI, 2016).

Em modelos de Machine Learning, para a escolha da quantidade de dados necessária ao treinamento e teste do algoritmos, é importante ponderar a dimensão do problema a ser resolvido e essa complexidade está relacionada ao número de variáveis envolvidas no processo (TSANGARATOS; ILIA, 2016).

Em termos gerais, um problema de alta dimensão refere-se a um problema em que o número de características (ou variáveis) envolvidas é muito grande em relação ao número de amostras disponíveis. Em classificação de imagens, a dimensão se refere ao número de atributos ou características que descrevem cada imagem. Quanto maior o número de características, maior a dimensão do problema (BISHOP, 2006).

Com base nas informações sobre a complexidade de um modelo de *Machine Learning*, foi observado que o problema envolvendo os três pintores para classificação de imagens, é um problema pequeno, pois envolve apenas 6 (seis) variáveis, segundo o que foi apurado pela ferramenta *Orange* por meio dos dados captados, conforme a Figura 10, e ainda, baseado em Jain et al., (2000), o qual afirma que é necessário haver uma quantidade de mínima de amostras 10 (dez) vezes maior que a dimensão do problema.

Por se tratar um problema de pequena complexidade, foram adotadas, para testes apenas 30 (trinta) imagens que seriam usadas inicialmente e seriam aumentadas de acordo com os testes, porém no decorrer da pesquisa observou-se que essa quantidade já era suficiente para a conclusão do experimento.

A motivação da comparação entre os algoritmos Regressão Logística e Rede Neural, foi uma necessidade de comprovar se realmente o desempenho da Regressão Logística poderia realmente ser similar ao desempenho de uma Rede Neural, como foi previsto na literatura, uma vez que o algoritmo de Regressão é usado para resolver problemas mais simples e é, por consequência um algoritmo de arquitetura mais simples, ao passo que uma rede neural é um algoritmos mais complexo usados para resolver problemas mais difíceis, embora os mesmos sejam modelos de *Machine Learning* usados na classificação.

E, diante da possibilidade da literatura estar correta em relação ao desempenho da Regressão Logística, um usuário, dependendo de suas limitações de quantidades de dados, com problema de baixa complexidade, e com poder computacional limitado poderia decidir qual dos dois algoritmos poderia utilizar com mais segurança e sem prejuízo de desempenho do algoritmo.

Diante do exposto acima e pelo fato da classificação de imagens ser uma tarefa de muita importância, fez-se necessário analisar o desempenho dos algoritmos Regressão Logística e Redes Neurais Artificiais utilizando a Matriz de Confusão para

extração de métricas para um comparativo entre os dois, na ferramenta de Análise de dados *Orange Data Mining*.

1.3.OBJETIVOS

1.3.1. Objetivo Geral

Realizar um comparativo de desempenho entre os algoritmos Regressão Logística e Redes Neurais Artificiais na classificação de imagens, utilizando a Matriz de Confusão para a extração de métricas de avaliação na ferramenta *Orange Data Mining*.

1.3.2. Objetivos Específicos

- Compreender a necessidade da aprendizagem e classificação de imagens;
- Compreender as potencialidades dos algoritmos Regressão logística e Redes Neurais;
- Identificar a extensão das possibilidades da ferramenta *Orange Data Mining*;
- Manipular os algoritmos para a classificação de imagens.

2. REFERENCIAL TEÓRICO

Esta seção trata da história, conceito e principais técnicas envolvendo Classificação de Imagens, Matriz de Confusão e temas afins, com o objetivo de situar e contextualizar o leitor no assunto.

2.1. Inteligência Artificial

Inteligência Artificial (IA) segundo (McCARTHY, 2007) é a ciência de se produzir máquinas inteligentes, máquinas capazes de imitar o conhecimento humano. “As primeiras conjecturas acerca do tema podem ter surgido no contexto da Segunda Guerra Mundial e assegura ser Alan Turing, matemático inglês, o pioneiro em IA” (McCarthy 2007 . p. 4). No final da década de 1950, havia muitos pesquisadores de IA, e a maioria deles baseava seu trabalho na programação de computadores.

No entanto, foi em 1956, no Dartmouth College, em Hannover, New Hampshire, nos EUA, que Claude E. Shannon, Marvin L. Minsky, Nathaniel Rochester e John McCarthy puseram em prática uma Proposta de Pesquisa em Inteligência Artificial, elaborada em 1955 (McCARTHY et al., 2006).

Nessa proposta, o estudo se baseava no princípio básico de que “todos os aspectos da aprendizagem ou outras características da inteligência humana poderiam ser descritos com muita precisão por meio de uma simulação em uma máquina” (McCARTHY et al., 2006. p. 1). O uso da linguagem e raciocínio lógico do ser humano a ser feito por um computador também foi considerado.

Desde que surgiu, a Inteligência Artificial tem sido objeto de estudo e alvo constante de definições sobre seu real significado e das ciências que a compõem. Embora a IA seja relativamente nova, tendo um pouco mais de sessenta anos, tem evoluído constantemente e passa por questionamentos sobre a coexistência de homem e máquina, ética e limites para uso de Inteligência Artificial.

Com o rápido avanço das tecnologias de Big Data, por exemplo, armazenamento de computação, tem aprimorado a capacidade de processamento e velocidade super-rápida das máquinas de processamento de dados (DUAN et al., 2019, p.1).

As aplicações da Inteligência Artificial estão presentes na medicina, a telemedicina e diagnósticos mais rápidos e sofisticados, agronegócio, estética, aplicativos como *facebook*, eletrodomésticos inteligentes, veículos autônomos, reconhecimento facial, assistentes personalizados para relacionamento com o cliente, aplicativos para geração de textos complexos, dentre outros.

2.2. Classificação de Imagens

A classificação de Imagens (CI), descreve o processo de identificar e rotular uma imagem em conformidade com o conteúdo visualizado. A CI faz parte do campo da Visão Computacional que diz respeito a como uma máquina enxerga as coisas ao seu redor, por meio de imagens capturadas por diversas formas e a partir dessas imagens retirar informações relevantes sobre as mesmas. (DE MILANO; HONORATO, 2010). O modo como a máquina enxerga é um indicativo para verificar o desempenho dos algoritmos de classificação.

Existem vários processos para se classificar uma imagem, um caminho bastante comum é a fase de Identificação de padrões, processo no qual uma regra de decisão é aplicada. Esse processo gera uma sequência de identificação de características que atribui a cada objeto uma categoria, com base nos dados ou observações. Essa identificação de padrão também é chamada de classificação de padrão ou simplesmente classificação. (SMITS et al.,1999). As Redes Neurais Artificiais por exemplo dividem uma imagem em vários pedaços dos quais são extraídas informações úteis.

2.2.1. Os Principais Métodos de Classificação de Imagens

Classificação Supervisionada- baseadas em função de identificação estatística, a qual por meio de métodos de treinamento e aprendizagem geram parâmetros para as classes e aplicam as regras de decisão geradas pela Identificação de Padrões. Nessa modalidade as classes são conhecidas previamente

Classificação Não-Supervisionada, é feita por meio de métodos *clusterizados* no qual o algoritmo geral de agrupamento seleciona o modelo número de pontos

como o a parte central do cluster, cada centro é uma categoria de acordo as semelhanças obtidas que tornam cada atributo padrão nos clusters representados pelo tipo. Nessa modalidade as classes não são conhecidas previamente.

Interpretação Visual - é feita pela distinção entre características geométricas e as características da superfície perceptíveis aos olhos humanos.

A Classificação baseada no conhecimento- São sistemas especialista para a classificação de dados, por meio de informações oferecidas por especialista (DHINGRA et al., 2019).

As Redes Neurais Artificiais- algoritmos com técnicas computacionais, representadas por um modelo matemático e inspiradas na biologia humana. Além dessas citadas, existe o Processamento técnico de imagem de sensoriamento remoto, que faz uso principalmente de imagens de satélites.

2.2.2. Técnicas Utilizadas na Classificação

Classificação por pixels; Medidas de distância ou probabilidades são usadas para determinar que um pixel pertença a determinadas áreas idênticas localizadas e determinadas pelos algoritmos. Há um conjunto de regras as quais o pixel é submetido antes de ser alocado em determinada classe.

Classificação por região- O Método Classificação Orientada a Objeto, que usa um agrupamento de pixel como unidade de classificação é utilizado, nessa modalidade, a textura e a forma dos objetos são levadas em consideração. As sub-técnicas que compõem a Classificação por região são a segmentação na qual há a conversão de uma imagem em múltiplos objetos e a classificação baseada em regras de decisão, que demonstram a composição dos objetos explicitados pelos seus atributos (DHINGRA et al., 2019).

2.2.3. Alguns Classificadores Imagens

Máquinas Vetoriais de Suporte (*Support Vector Machine*): Algoritmos muito utilizadas para classificar dados que podem ser linearmente separados, por isso sua aplicação em classificação binárias e de múltiplas classes. SVM é um conjunto de

métodos de aprendizado supervisionado, no qual os vetores de características são mapeados não linearmente para um espaço de elementos de alta dimensão. (PATHAN et al, 2018). Podem também ser usadas para problemas não lineares, se usadas com Kernels.

Florestas Aleatórias (*Random Forests*): Algoritmo de Aprendizado de máquina baseado em árvores de Decisão. Esse método combina várias árvores de decisão pra conseguir uma classificação melhor.

K - vizinhos mais próximos (*k-Nearest Neighbors*): Técnica que busca encontrar os pontos de dados mais próximos, a por meio dessa abordagem, ele faz previsões, usando os pontos encontrados. É um algoritmo simples, de baixo desempenho (DHINGRA et al., 2019).

Uma classificação de imagens confiável envolve muitos fatores a serem considerados, como as técnicas utilizadas, tempo de execução além de muitos outros requisitos. “O sucesso de uma classificação de imagem depende de muitos fatores, como conjunto de dados de imagem de alta qualidade, dados auxiliares, recurso extraído, seleção de recursos e projeto de um método de classificação apropriado” (CHANDRA; BEDI 2018 p. 9).

O aprendizado de um modelo e sua generalização em outras situações é o principal objetivo do uso dos algoritmos de aprendizagem na classificação de imagens, por meio dessa generalização é possível usar o aprendizado em Classificação de Imagens em aplicações como: mecanismos de pesquisa visual, detecção de logotipo: escuta visual e escuta social, Reconhecimento facial, aplicações na medicina por meio de diagnósticos e prognósticos confiáveis, agronegócio, previsões climáticas, além de influenciar nos processos de tomadas de decisões.

Existem vários algoritmos de classificação e várias são as técnicas utilizadas, com especificidades como precisão, baixo ou alto grau de processamento, tipos de dados que podem ser classificados e outras características que podem influenciar na escolha dos mesmos. Para esse trabalho optamos por Regressão Logística, Redes Neurais Artificiais ambos são classificadores e aprendem com um conjunto de dados.

2.3. Machine Learning

Aprender é um fator de muita importância para o humano, sua sobrevivência depende de como aprende a buscar o alimento, se higienizar, trabalhar, se proteger dos perigos que o ronda, são experiências que ele vivencia ou que são vividas por outras pessoas, mas que é apreendida por ele, nesse contexto, classificar as coisas para sua tomada de decisões é um resultado do aprendizado e primordial para todos os segmentos de sua vida.

No decorrer da história o ser humano criou máquinas com o objetivo de fazerem o seu trabalho, de tornar sua vida mais fácil e lucrativa e avançou para inventar ferramentas que o imitasse e até mesmo o superasse em algumas habilidades, surgiram assim as máquinas inteligentes.

2.3.1. Tipos de Aprendizado Utilizados em *Machine Learning*

O termo *Machine Learning* (ML) ou Aprendizado de máquina é definido por (MARSLAND, 2015), como a capacidade das máquinas em lidar com dados de forma eficiente, de interpretá-los e de aprenderem com eles mudando ou adaptando suas ações, o referido autor também classifica o Aprendizado de Máquina como:

Aprendizagem supervisionada - Um conjunto de exemplos de treinamento nos quais as respostas corretas (alvos) são fornecidas e, com base nesse conjunto de treinamento, o algoritmo generaliza para responder corretamente para todas as entradas possíveis. Isso também é chamado de aprendizado com exemplos. Um exemplo disponibilizamos um conjunto de fotos de gatos e as rotulamos com o nome desse animal, a máquina vai aprender o que é um gato, assim como o ser humano aprende quando é pequeno e o pai ou mãe lhe mostra um gato.

Aprendizado não supervisionado - As respostas corretas não são fornecidas, mas sim o algoritmo tenta identificar semelhanças entre as entradas para que as entradas que tenham algo em comuns são categorizados juntos. A abordagem estatística para o aprendizado não supervisionado é conhecida como estimativa de densidade. A máquina observa o que aqueles dados têm em comum e assim os agrupa.

Aprendizado por reforço - Isso está em algum lugar entre o aprendizado supervisionado e não supervisionado. O algoritmo é informado quando a resposta está errada, mas não é informado como corrija isto. Tem que explorar e experimentar diferentes possibilidades até descobrir como acertar a resposta. O aprendizado por reforço às vezes é chamado de aprendizado com um crítico por causa desse monitor que pontua a resposta, mas não sugere melhorias. O aprendizado por reforço é utilizado também com animais, sempre que eles acertam, recebem uma recompensa.

Aprendizagem evolutiva - A evolução biológica pode ser vista como um processo de aprendizagem: organismos adaptam-se para melhorar suas taxas de sobrevivência e chance de ter descendentes em seu meio Ambiente. Veremos como podemos modelar isso em um computador, usando uma ideia de fitness, que corresponde a uma pontuação de quão boa é a solução atual.

A arquitetura e funcionamento dos algoritmos de aprendizagem estão modelados de acordo com um ou mais modalidades de aprendizagem (Bishop, 2006), e também de acordo com a natureza e disponibilidade dos dados (LeCun et al., 2015).

O *Machine Learning* conta com diferentes algoritmos para resolver problemas envolvendo dados. Os cientistas de dados gostam de salientar que não existe um único tipo de algoritmo único para todos que seja melhor para resolver um problema. “O tipo de algoritmo empregado depende do tipo de problema que você deseja resolver, do número de variáveis, do tipo de modelo que melhor se adequa a ele e assim por diante” (MAHESH, 2018, p. 2)

No assunto sobre ML, também é necessário mencionar sobre *Deep Learning*, que é o aprendizado profundo. *Deep Learning* é um termo usado para designar a técnica que utiliza redes neurais profundas no aprendizado. *Deep Learning* é uma sub-área do *Machine Learning*.

2.4. Regressão Logística

Logistic Regression ou Regressão Logística (LR) é a aplicação de uma função *Sigmoid* a um modelo linear. O modelo generalizado que pode utilizar dados contínuos ou discretos, é muito utilizada em estatísticas probabilísticas, em classificação e no aprendizado de máquinas. A *Logistic Regression* divide os espaços de forma linear, e mesmo que algumas variáveis sejam correlacionadas, ela funciona. O seu mecanismo de aprendizado é um modelo discriminativo que estima a probabilidade de um determinado recurso (x) e o rótulo (y) diretamente dos dados de treinamento, minimizando o erro (TSANGARATOS; ILIA, 2016).

A LR também recebe o nome de Função *Logit*, é uma técnica de análise multivariada, apropriada e muito utilizada em situações nas quais a variável dependente estudada é categórica e binária (RIOS, 2016). O algoritmo provê probabilidades utilizando novos dados com medidas contínuas e discretas baseadas na Máxima Verossimilhança. A função *Logit* é mais apropriada para conjuntos de dados menores, pois apresentam maiores taxas de convergência. A classificação vai de 0 a 1 e o gráfico é de uma função *Sigmoid*, na qual quanto mais perto de 1 estiver o objeto, maiores são as chances de ele pertencer a classe 1 e quanto mais perto de 0, menor a chance desse objeto pertencer a essa classe, no caso da regressão retornar o valor 0, considera-se que o objeto pertence a classe 1, mas com baixa probabilidade.

Duas dificuldades importantes na Regressão Logística são o *Overfitting* e o *Underfitting*, no primeiro caso, “O LR produz resultados que superajustam os dados, uma condição quando um modelo começa a memorizar o treinamento dados em vez de aprender a generalizar a partir da tendência” (RIOS, 2016, p.2). Nessa situação o modelo oferece uma acurácia de 100% no momento em que é treinado, mas quando é testado em um conjunto diferente de dados, o desempenho é bem menor. Quando há um conjunto grande de dados, com muitas variáveis e atributos, pode ocorrer o *Underfitting*, situação que deixa o programa lento, pouco produtivo, necessitando de maiores recursos de processamento para ser executado (Idem, 2016)

Para resolver ou minimizar esses problemas a Regularização é utilizada, tanto para casos de diminuir a quantidade de variáveis, como para casos de *Overfitting* ou

de *Underfitting*. As Regularizações presentes na Regressão Logística são a *Lasso1* (*Least Absolute Shrinkage and Selection Operator*), que traduzido é Menor Encolhimento Absoluto e Operador de Seleção e *Ridge1*, que significa *Regressão de cumeeira*. A *Lasso1*, ou L1, aplica restrições aos coeficientes, igualando-os a zero, isso faz com que atributos sejam eliminados, tornando o modelo mais leve e eficiente, ideal para *Datasets* grandes.

A Regularização *Ridge1* Ou L2, age de forma parecida, porém, aplicando uma penalização equivalente ao valor ao quadrado da magnitude dos coeficientes. O método Ridge, com o parâmetro lambda, penaliza os coeficientes que assumem valores muito grandes, levando-os a tender a zero.

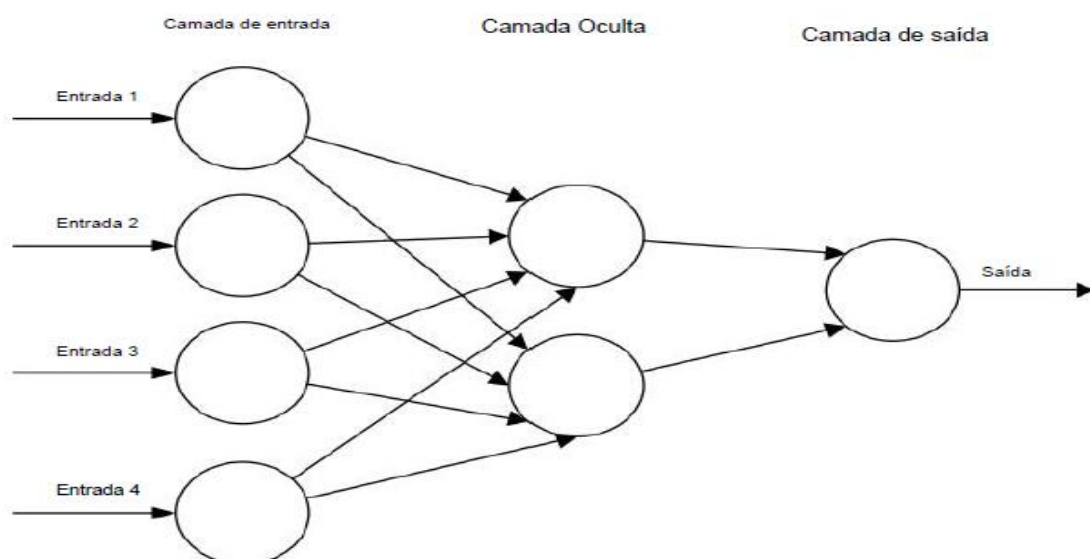
No *Ridge* os valores dos coeficientes são minimizados, não ficando, portanto, fora do modelo, com os coeficientes relacionados diminuídos, o problema da multicolinearidade é resolvido, no L2, os atributos não são eliminados, útil em conjunto de dados menores (HASTIE et al., 2009).

2.5. Redes Neurais

Redes Neurais Artificiais ou *Artificial Neural Networks* - ANN, ou ainda, RNA, são algoritmos inspirados na arquitetura do cérebro humano e está classificada como algoritmo biológico (FIGURA 1).

Dos algoritmos inspirados na Neurociência, as redes neurais são as que obtiveram maior sucesso (RAWAT; WANG, 2017). ANN é um recurso bastante avançado e amplamente utilizado no *Machine* e no *Deep Learning*.

Figura 1 – Estrutura de uma Rede Neural simples



Fonte: Rede Neural (O'SHEA ; NASH, 2015).

As Redes Neurais Artificiais são um sistema de processamento de informações que semelhante a um cérebro humano. É composta por uma quantidade enorme de neurônios, comparado a um grande processador distribuído paralelo, que resolve de forma eficaz problemas categóricos. No caso, é um sistema integrado entre Inteligência Artificial e processamento de informações (PATHAN et al., 2018).

Na representação matemática de uma RNA, estão presentes o conceito de entradas, pesos, função de soma, função de ativação e saídas. A rede auxilia na decisão de ajustes dos pesos, mudanças de parâmetros e no tipo de aprendizado, no final, após a análise do sistema, é feito o treinamento da Rede, então é feita a implementação da ANN (A.D.DONGARE et al., 2012).

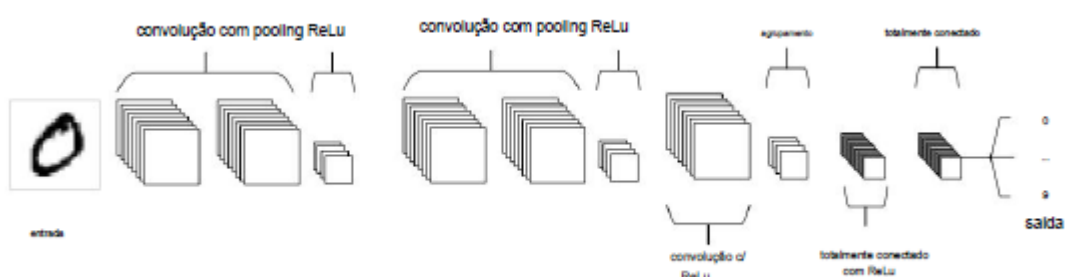
Em uma Rede, os neurônios são as unidades de processamento, cada entrada associada a ele possui um peso associado, a entrada é multiplicada pelo peso, em seguida, uma função não linear é aplicada a esse resultado e segue para a saída. Os Neurônios são interligados uns aos outros, formando uma rede e podendo alimentar entradas de um ou mais neurônios da camada seguinte.

Os pesos têm papel fundamental no desempenho de uma Rede Neural. “Tais redes se mostraram dispositivos computacionais muito capazes. “Se os pesos forem definidos corretamente, uma rede neural com neurônios suficientes e uma função de ativação não linear pode aproximar uma gama muito ampla de funções matemáticas”

(GOLDBERG YOAV, 2016, p.10). O ajuste dos pesos é repetido até que a condição de aprendizado desejada seja atendida.

As redes neurais convolucionais são amplamente usadas em visão computacional, na realidade, são recursos de vanguarda para análise e classificação de imagens (FIGURA 2). Por meio dessas redes é feita a extração de recursos das imagens ou vídeos, com os quais são feitos mapeamentos e posterior classificação (O'SHEA ; NASH, 2015).

Figura 2 – Arquitetura de uma Rede Neural Convolucional



Fonte: Rede Neural Convolucional (O'SHEA ; NASH, 2015).

Uma *Convolutional Neural Network* (CNN) - rede neural convolucional simples é formada por três tipos de camadas: Camada Convolucional, Camada de Agrupamento e Camada totalmente conectada, além das camadas de entrada e saída, quando empilhadas, essas camadas formam a arquitetura de uma CNN. Na camada de entrada, os valores de pixel das imagens são mantidos, na Camada Convolucional, as saídas dos neurônios são determinadas por meio do cálculo do produto escalar entre seus pesos e a região conectada ao volume de entrada.

A função de Ativação é aplicada pela ReLu-Unidade Linear Retificada para a saída da camada anterior. Na camada de Agrupamento as amostragens são reduzidas e os parâmetros diminuídos, trazendo eficiência para o desempenho da Rede, já as camadas totalmente conectadas efetuam o mesmo papel da Rede Neural Artificial e produzem pontuações de classe por meio das ativações para serem usadas na classificação (KEIRON; NASH, 2015).

A única diferença notável entre as CNNs e as ANNs tradicionais é que as CNNs são usadas principalmente no campo de reconhecimento de padrões em imagens. Isso nos permite codificar recursos específicos da imagem na arquitetura, tornando a rede mais adequada para tarefas focadas na imagem

- reduzindo ainda mais os parâmetros necessários para configurar o modelo (KEIRON; NASH, 2015, p. 3).

Tanto as redes neurais artificiais e as recorrentes, são usadas como classificadores, a escolha de uma ou outra depende da complexidade do modelo na qual serão aprendizes. Problemas de *Overfitting* também ocorre nas redes, a diminuição de parâmetros tem sido eficaz para solucionar esse problema, essa prática também minimiza tempo e robustez de processamento.

Rede neural recorrente é um modelo de rede usada para modelar sequência de dados, bastante eficaz quando se trata de temporalidade. A arquitetura desse modelo de rede é diferente das duas citadas anteriormente, aqui, a saída de um neurônio é usada como entrada para o neurônio seguinte, e cada etapa é memorizada e lembrada na etapa seguinte.

As Redes Neurais Recorrentes são muito usadas no Processamento da Linguagem Natural, na previsão da próxima palavra a ser digitada, em análise de sentimentos, traduções simultâneas e outras.

2.6. Matriz de Confusão

A Matriz de confusão é uma ferramenta utilizada para medir o desempenho de algoritmos de classificação, as técnicas de aprendizado de máquina e ainda, avaliar o desempenho individual de cada classe estudada por meio de métricas que permitem verificar onde o algoritmo precisa ser aperfeiçoado e corrigir as possíveis falhas, também pode se medir a regularidade com que cada classe é classificada erradamente ou acertadamente seus dados (FIGURA 3).

A Matriz de Confusão é composta por linhas nas quais estão as classes reais e por colunas, nas quais estão as classes previstas, ou vice-versa, na forma de uma matriz quadrada (RASCHKA, 2014).

A Matriz pode ser binária quando possui apenas duas classes, como mostra a Figura 3, ou Multiclasse, quando possui mais de uma classe, conforme a Figura 4.

Na Matriz Binária, assim como na Multiclasse existe o conceito da classificação das classes em Verdadeiro Positivo, e Verdadeiro Negativo, Falso Positivo e Falso Negativo. Os valores Verdadeiro Positivo e Verdadeiro Negativo são os dados que foram classificados corretamente, os Falso Positivo e Falso Negativo são os que foram

classificados erroneamente pelos algoritmos (RASCHKA, 2014). Na Matriz de Confusão há dois tipos de acertos que são Verdadeiros Positivos e Verdadeiros Negativos e dois tipos de erros, a saber: Falsos Positivo e os falsos Negativo.

Figura 3 – Matriz de Confusão Binária

		classe prevista	
		P	N
Real Aula	P	Verdadeiro Positivos (PT)	Falso Negativos (NF)
	N	Falso Positivos (FP)	Verdadeiro Negativos (TN)

Fonte: (RASCHKA , 2014).

Em Uma matriz, elementos localizados na Diagonal Principal os quais são os Verdadeiro Positivo e os Verdadeiro Negativo indicam os acertos, já os erros estão nas outras células, conforme mostra a Figura 3. Essa regra vale tanto para a Matriz Binária, quanto para a Multiclasse, conforme mostra a Figura 4, porém, na Multiclasse o cálculo é feito célula a célula, exceto para a diagonal Principal (GRANDINI et al., 2020).

Figura 4– Matriz de Confusão Multiclasse

		Predicted			Σ
		pintor1	pintor2	pintor3	
Actual	pintor1	9	0	1	10
	pintor2	1	6	3	10
	pintor3	0	5	5	10
Σ		10	11	9	30

Fonte: Autora (2023).

A Figura 4 demonstra a estrutura de uma Matriz Multiclasse contendo 3 (três classe) com 30 (trinta) elementos. Na Matriz Binária as métricas que podem ser extraídas estão descritas abaixo, juntamente com suas respectivas fórmulas, as mesmas métricas e outras mais avançada podem ser extraídas de uma Matriz Multiclasse.

Acurácia: é a métrica obtida pela soma de acertos da classificação, dividida pelo total de instâncias classificadas. Na prática somam-se os elementos da diagonal principal e os divididos pela soma das outras células. Essa métrica nos diz, de forma geral o quanto o modelo acertou. A acurácia funciona bem em modelos com dados equilibrados com classes com mesma quantidade de exemplares.

A Acurácia é chamada de Precisão para classificações binárias e de Precisão Balanceada na Matriz Multiclasse, isso quer dizer que na Acurácia comum, a soma dos valores Verdadeiro Positivo e dos Verdadeiro Negativo são somente dois e são divididos pela soma dos Falso Negativo e dos Falso positivo, gerando um valor global que pode apresentar muitos erros, principalmente se os Verdadeiro Negativo forem mais do que os Verdadeiro Positivo.

Em uma Matriz Multiclasse essa métrica é obtida calculando as classes de forma individual e depois fazendo uma média (GRANDINI et al., 2020). A Equação da Acurácia na matriz Binária é: $Acurácia = \frac{\text{previsões corretas}}{\text{total de previsões}}$

Recall: chamado de Revocação ou ainda, Sensibilidade e Taxa de Verdadeiros Positivos), essa métrica mostra a proporção de amostras corretamente classificadas como positivas em relação ao total de amostras que são de fato positivas. Para chegar a esse valor, divide-se o número de verdadeiros positivos (TP) pela soma dos verdadeiros positivos com os falsos negativos (FN).

Essa métrica questiona se de todos os classificados como positivos, quantos poderiam ser identificados (CHAUHAN, 2020). Então temos que a equação do Recall seria: $Recall = \frac{TP}{TP + FN}$. Em uma Matriz Multiclasse o Recall, que é uma média dos acertos de cada classe, dividido pelo número de classes totais, é usado para conseguir a Precisão Balanceada.

Precisão: se refere à proporção de instâncias corretamente classificadas como positivas em relação ao total de instâncias classificadas como positivas. Esse cálculo é feito dividindo o número de verdadeiros positivos (TP) pelo total de verdadeiros positivos e falsos positivos (FP). Por meio da Precisão, podemos verificar se das

instâncias classificadas como positivas, quantas verdadeiramente são positivas (RASCHKA, 2014).

Em uma Matriz Multiclasse essa métrica é a chamada Precisão Balanceada. Isso porque ela representa uma média dos valores das classes, na Acurácia da Matriz Binária os valores individuais de cada célula têm muito peso, e isso pode levar a enganos quando se trata de classes desbalanceadas, aquelas nas quais a quantidade de elementos é muito desigual.

Já na Multiclasse os valores são balanceados de forma geral, o conjunto de dados fica nivelado. Precisão é a probabilidade de que a previsão do modelo esteja correta (GRANDINI et al., 2020). “A fórmula da Precisão Balanceada é essencialmente uma média de recalls. Primeiro avaliamos o Recall para cada classe, depois calculamos a média dos valores para obter a pontuação de Precisão Balanceada” (Idem,2020, p.4).

Medida F1-Score, ou F-measure: É uma média harmônica entre precisão e recall. Ela leva em consideração tanto falsos positivos quanto falsos negativos. É calculada usando a seguinte fórmula: $F1\ Score = 2 * (Precisão * Recall) / (Precisão + Recall)$.

Esta métrica é influenciada diretamente pela Precisão e pelo Recall, quando os mesmos dão como resultado zero ou número muito baixo, também o F1 será baixo, e isso quer dizer que a qualidade do modelo não é boa, resumindo, o F1 é a medida, dentre as demais acima como a mais importante para se avaliar um modelo (kdnuggets, 2020).

Especificidade: avalia a capacidade do modelo em detectar resultados negativos. Usa-se a seguinte fórmula para esse cálculo: dividem-se os verdadeiros Negativos pela soma dos Verdadeiros Negativos com os Falsos Positivos) (CHAUHAN, 2020). A Equação para cálculo da Especificidade é: $Especificidade = TN/(TN+FP)$.

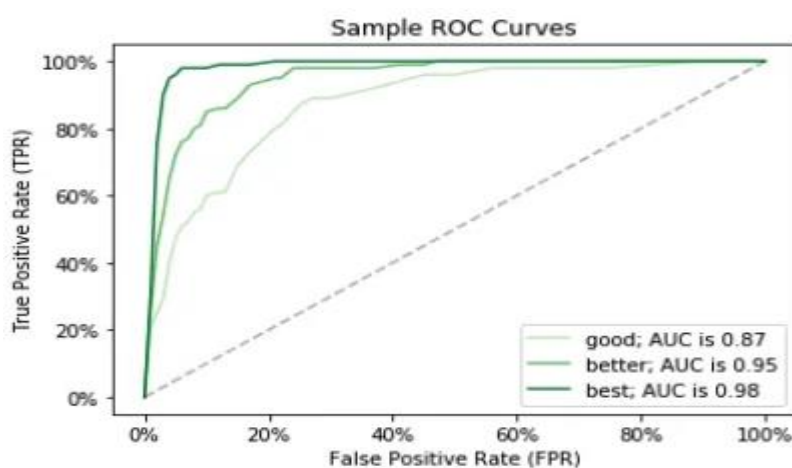
Curva ROC (Receiver Operating Characteristic) Receptor Operacional de Características): é uma ferramenta que apresenta os dados em forma de um gráfico e é utilizada na avaliação de aprendizes de aprendizado de máquina. A Curva ROC demonstra a relação entre a Especificidade e a Sensibilidade de acordo com os limites de decisão de classificação do aprendiz e na visualização do gráfico, quanto mais a

curva do modelo estiver do canto esquerdo do gráfico, melhor será esse modelo em contraposição do que estiver mais à direita (SILVA, 2019).

Área sob a Curva- AUC: ferramenta que quantifica o quanto o modelo é capaz de separar as classificações verdadeiras das falsas. AUC calcula a área da forma bidimensional formada abaixo da curva (FIGURA 5).

Essa métrica indica a probabilidade de duas previsões serem corretamente ranqueadas. A AUC será um valor entre 0 e 1. Quanto maior esse valor, melhor a capacidade do modelo em separar classes e quanto maior for a área dentro da curva, melhor será o desempenho do modelo (SILVA, 2019).

Figura 5 – Gráfico Curva ROC



Fonte: kdnuggets (2020).

2.7. Orange Data Mining

A Orange Data Mining - Laranja Mineração de Dados, é uma ferramenta de análise de dados e está localizada no sítio: <https://orangedatamining.com/>, é definida na própria *Home Page* como uma “caixa de ferramentas gratuitas”, nela estão inclusos diversos algoritmos utilizados em análise estatística descritiva, classificação de imagens, Análise de Sentimentos e muitas outras aplicações baseadas em *Machine* e em *Deep Learning*. A Plataforma é desenvolvida na Linguagem *Python*. O *Orange* é desenvolvido pelo Laboratório de Bioinformática da Universidade de Ljubljana,

3. PROCEDIMENTOS METODOLÓGICOS

Neste capítulo serão apresentadas as ferramentas com suas aplicações as quais foram utilizadas para o desenvolvimento deste trabalho, constam também o local e o período de realização e ainda, a modalidade da pesquisa, bem como suas fontes bibliográficas.

3.1. Fontes Bibliográficas

Para a construção do referencial teórico do trabalho foram pesquisados:

- Artigos de Revisão;
- Livros;
- Publicações Científicas;
- Documentação da Ferramenta Orange;
- Sites técnicos sobre o assunto.

3.2. Ambiente de Execução e Equipamentos Utilizados

Foi utilizado um computador com um processador *AMD Ryzen 7 3800x 8-Core Processor, de 3,90 Ghz, com 16GB* de memória e Sistema Operacional de x64 bits, com placa de vídeo GTX 1060 5GB.

Na coleta de dados (Imagens) foi usada a ferramenta *Image Downloader*, uma extensão do Google disponível no sítio: <https://chrome.google.com/webstore/detail/imagdownloader/cnpniohnpfhjihaiiggeabnkjhpaldj>. Por ela foi possível baixar imagens de páginas na internet, escolher tamanho além de também ter a opção de escolher somente imagens com *links*, para o caso de não preferir imagens conectadas a determinadas páginas por meio de *script*.

3.3. Tipo da Pesquisa

Quanto ao aspecto da condução e dos procedimentos adotados para esse trabalho, foi adotada a Análise Descritiva, para descrever as partes concernentes à ferramenta Orange. De acordo com Gil (2008), as pesquisas descritivas objetivam a descrição das características de uma população, fenômeno ou de uma experiência, nesse aspecto, ao descrever os experimentos realizados, esse trabalho assume características desse tipo de pesquisa.

O uso da Análise Descritiva foi o mais apropriado para o tipo de pesquisa nesse trabalho.

A Análise Descritiva é a fase inicial deste processo de estudo dos dados coletados. Utilizamos métodos de Estatística Descritiva para organizar, resumir e descrever os aspectos importantes de um conjunto de características observadas ou comparar tais características entre dois ou mais conjuntos (REIS e REIS, 2002, p.5).

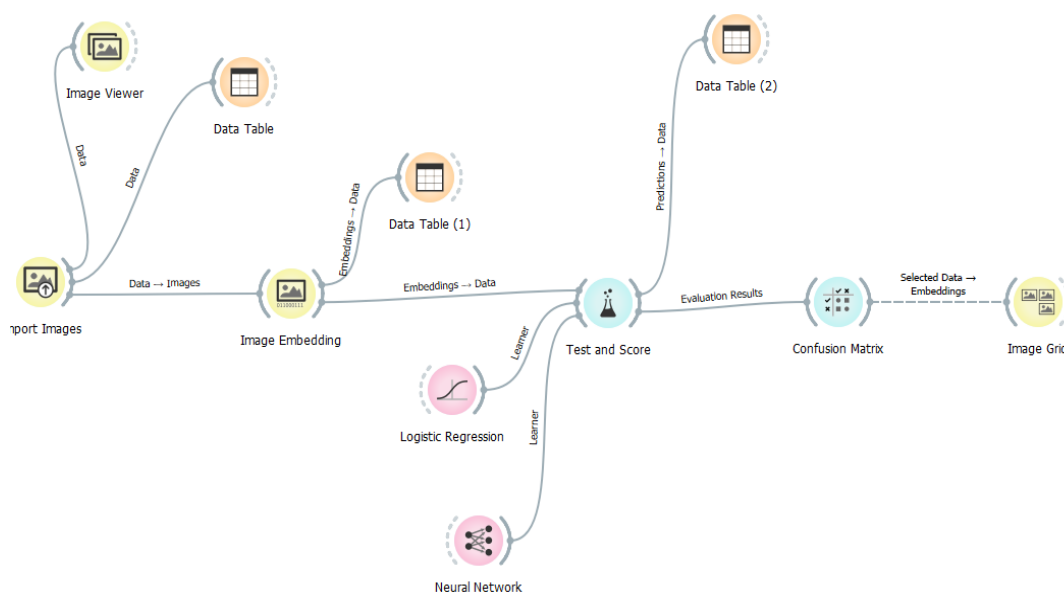
3.4. Classificação de Imagens

A classificação de imagens é um requisito muito importante no campo da Visão Computacional, um sistema necessita traduzir corretamente ou com as mínimas falhas o que lhe exerga para que seja confiável. Para o desenvolvimento do Fluxo de trabalho apresentado a seguir foi utilizada a ferramenta *Orange Data Mining* e os seguintes *widgets*, ou ferramentas.

3.5. Workflow - Fluxo de trabalho realizado

Na Figura 7, temos o fluxo de trabalho e uma visualização das ferramentas para uma melhor compreensão.

Figura 7 – Workflow



Fonte: Autora (2023).

O Fluxo de Trabalho apresentado na Figura 7 é um demonstrativo do que a ferramenta executou: Em um primeiro momento houve o processo de importação das imagens de uma pasta do computador local para a *Orange Data Mining* pelo *widget Import Image*. Após essa etapa, as mesmas passaram pelo processo de Incorporação de imagens, no *Image Embedding* para que fossem mapeadas por redes neurais, após esse processo, etapa seguinte foi configurar os parâmetros da ferramenta *Test and Score* e, em sequência as imagens já mapeadas foram disponibilizadas para os algoritmos de classificação Redes Neurais e Regressão Logística, os chamados Aprendizes para treinamento.

A etapa posterior de verificação dos desempenhos dos Aprendizes por meio da ferramenta *Test and Score* e a Matriz de confusão. Por último, o *widget Image Grid* comparou as semelhanças das imagens classificadas. As Tabelas de dados foram usadas para mostrar as mudanças ocorridas após as imagens passarem pelas ferramentas mostradas no fluxo.

4. RESULTADO E DISCUSSÕES

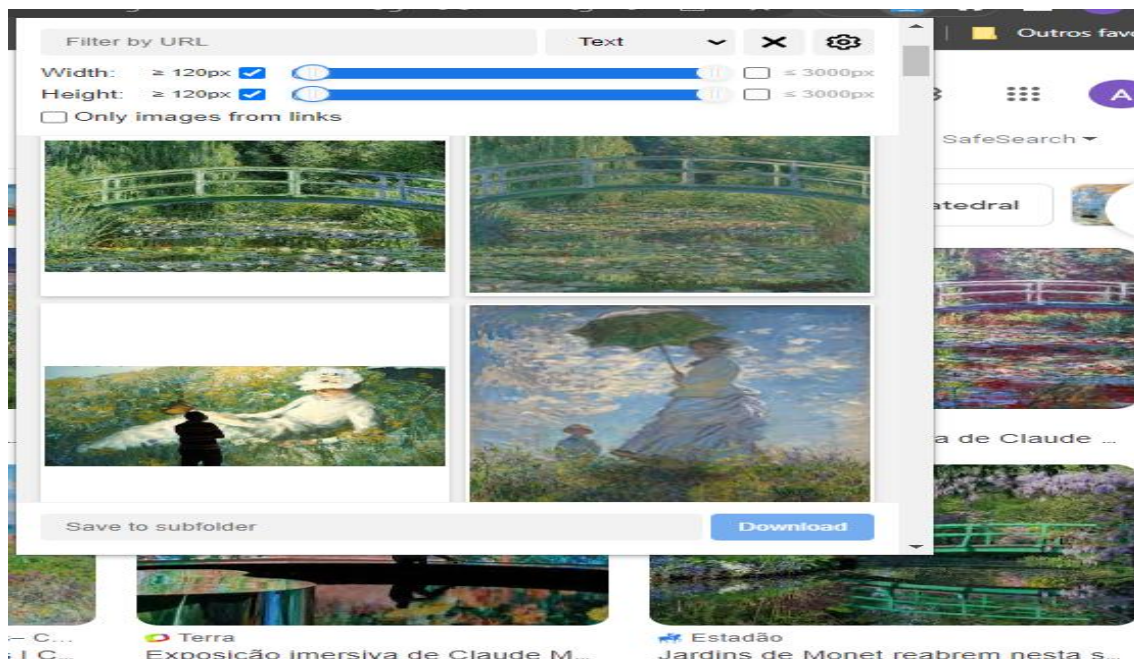
A primeira etapa desse projeto foi escolher a ferramenta adequada, uma que fosse interessante do ponto de vista voltado para a Classificação de Imagens, com funcionalidades que permitissem avaliar o desempenho dos algoritmos. Tendo em vista o objetivo principal mencionado anteriormente, a Orange Data Mining, por ser gratuita, e ser voltada para análise de dados e ser intuitiva, apresentou-se como adequada ao projeto.

Diante desse cenário, e após a delimitação do fluxo de trabalho, foi descrito de forma sequencial, de maneira a acomodar uma melhor compreensão das etapas realizadas, aos algoritmos destinados a classificação, por uma boa prática existente no ramo do *Machine e Deep Learning*, são denominados aprendizes.

Diante do exposto, o *Workflow*, conforme mostra a Figura 7, foi sequenciado da seguinte forma: Importação e tratamento de dados, Incorporação da Imagens, Aprendizes, Teste e Validação e Visualização, foram utilizados os seguintes *Widgets* - ferramentas, organizadas de acordo com a sequência de uso.

O Trabalho de classificação de imagens requer como principal matéria prima vídeos e imagens. Para tanto, de forma geral, é sabido que para um aprendizado eficiente das máquinas é requerida uma quantidade significativa de dados para treinamento e testes, sendo, portanto, oportuno o uso de ferramenta de captação de imagem. Na coleta de dados (Imagens) foi usada a ferramenta *Image Downloader*, conforme mostra a Figura 8. A ferramenta é uma extensão do *Google*, com a qual foram coletadas 30 (trinta) imagens em formato jpg (*joint photographic experts group*) de tamanhos variados para o treino.

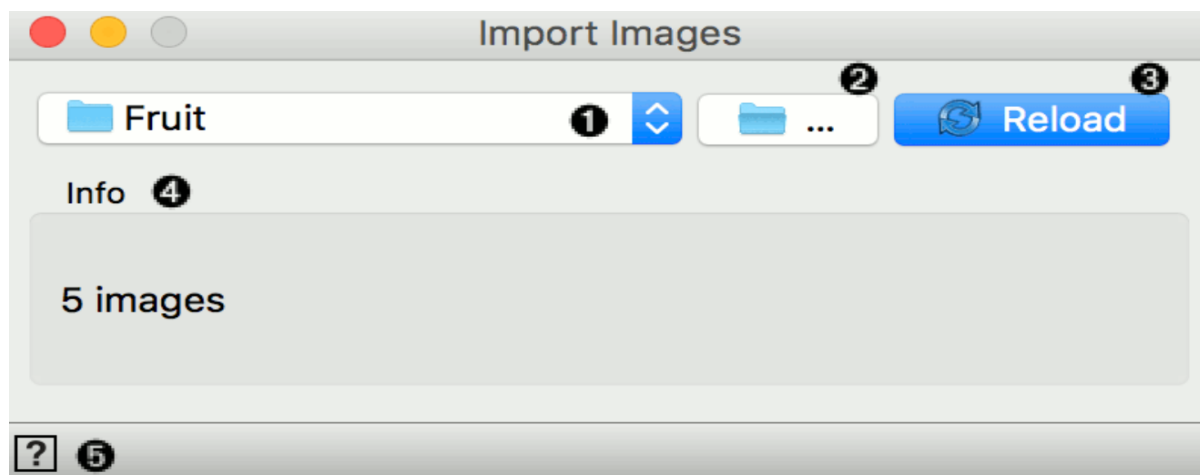
Figura 8 – Ferramenta de captação de imagens Image Downloader



Fonte: Autora (2023).

A importação das imagens foi feita pela ferramenta seguinte.

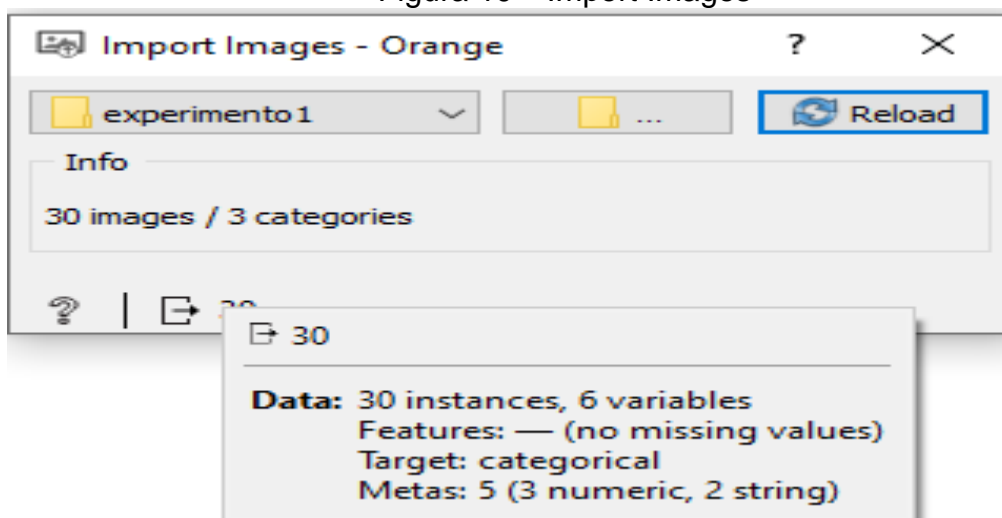
Figura 9 – Import Images



Fonte: Autora (2023).

Import images: Importar imagens - é um *widget* que percorre um diretório e retorna uma linha por imagem localizada, conforme mostra a Figura 9. Essa ferramenta carrega as imagens e cria categorias que são usadas como valores de classe. No *Import Image*, também é possível verificar a quantidade de variáveis, de instâncias, o tipo de dados e outras informações, conforme a Figura

Figura 10 – Import Images

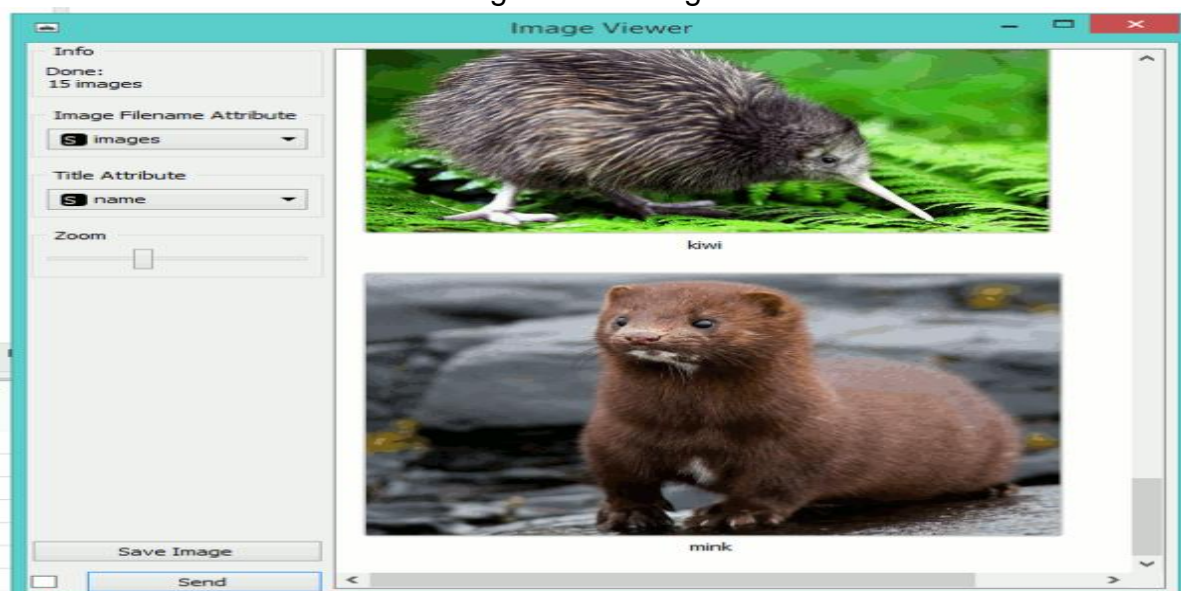


Fonte: Autora (2023).

Na Figura 10, estão as características dos dados usados nos experimentos, sendo 30 (trinta amostras), 6 (seis) variáveis categóricas e 5 (cinco) Metadados.

Os resultados podem ser visualizados em uma tabela de dados, *Data Table*.

Figura 11 – Image Viewer



Fonte: Autora (2023).

Image Viewer, conforme mostra a Figura 11, é utilizada para visualizar as imagens que são acompanhadas de um conjunto de dados, ficam armazenados

localmente ou na internet. Ele pode ser usado para comparação de imagens, enquanto procura semelhanças ou discrepâncias entre instâncias de dados selecionadas (por exemplo, crescimento bacteriano ou representações de *bitmap* de caligrafia). Essa ferramenta permite visualizar todo o conjunto de dados ou somente as instâncias selecionadas como tamanho e nome.

Figura 12 – Data Table

	y	RI	Na	Mg	Al	
172	5	1.513	13.020	0.000	3.040	70.486
173	5	1.513	13.000	0.000	3.020	70.700
174	5	1.520	13.380	0.000	1.400	72.250
175	5	1.521	12.850	1.610	2.170	72.180
176	5	1.521	12.970	0.330	1.510	73.390
177	6	1.519	14.000	2.390	1.560	72.370
178	6	1.519	13.790	2.410	1.190	72.760
179	6	1.518	14.460	2.240	1.620	72.380
180	6	1.519	14.090	2.190	1.660	72.670
181	6	1.513	14.400	1.740	1.540	74.550
182	6	1.519	14.990	0.780	1.740	72.500
183	6	1.519	14.150	0.000	2.090	72.740
184	6	1.520	14.560	0.000	0.560	73.480
185	6	1.511	17.380	0.000	0.340	75.410
186	7	1.511	13.690	3.200	1.810	72.810
187	7	1.518	14.320	3.260	2.220	71.250
188	7	1.523	13.440	3.340	1.230	72.380
189	7	1.522	14.860	2.200	2.060	70.260

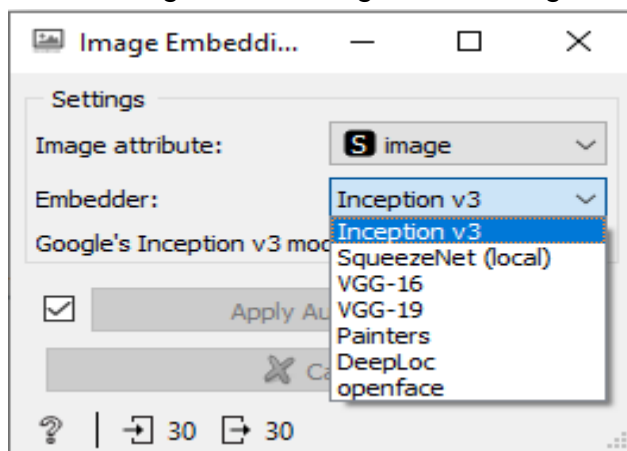
Fonte: Autora (2023).

Data Table - Tabela de dados, conforme mostra a Figura 12, recebe um ou mais conjuntos de dados e os apresenta em forma de planilha no qual são demonstradas a quantidade de instâncias e as *features* ou características. As instâncias estão em linhas e os valores de atributos em colunas. Cores pode ser usada para visualizar as classes. E cada linha pode ser selecionada e enviada para o canal de saída da ferramenta.

Incorporação é o processo de inserção das imagens na ferramenta. Para a análise de imagens, o Orange possui o algoritmo *Imagem Embedding* composto por redes neurais artificiais, convolucionais e recursivas, e que opera da seguinte forma: A incorporação de imagens lê imagens e as carrega em um servidor remoto ou as avalia localmente.

Modelos de aprendizado profundo são usados para calcular um vetor de recursos para cada imagem, ou seja, cria um grupo de recursos gráficos que podem ser úteis no manuseio de imagens, como redimensionamento e edição, sem perda de qualidade Excel. A interface do algoritmo *Imagem Embedding* pode ser vista na Figura 13.

Figura 13 – Image Embedding



Fonte: Autora (2023).

Além disso, esse *widget* pode receber dados do *Import Image* e também do *Widget File*, nos formatos de planilhas. Ele retorna uma tabela de dados aprimorada com colunas adicionais (descritores de imagem). Essa ferramenta utiliza várias Redes Neurais Convolucionais de acordo com a Figura 13, tais como:

- *SqueezeNet*: Modelo pequeno e rápido para reconhecimento de imagem treinado no *ImageNet* que é um banco de dados gratuito, cujo objetivo é aprimorar a visão computacional, atualmente o *ImageNet* possui mais de 20.000 (vinte e mil) categorias de imagens treinadas.
- *Inception v3*: modelo *Inception v3* do Google treinado no *ImageNet*. O *Inception v3* possui cerca de 25mi de parâmetros.
- *VGG-16*: modelo de reconhecimento de imagem de 16 camadas treinado no *ImageNet*.
- *VGG-19*: modelo de reconhecimento de imagem de 19 camadas treinado no *ImageNet*. Assim como o modelo anterior é uma rede neural profunda.
- *Painters*: um modelo treinado para prever pintores a partir de imagens de obras de arte.
- *DeepLoc*: Um modelo treinado para analisar imagens de células de levedura. Esse modelo é amplamente usado em biotecnologia.
- *OpenFace*: Um software de visão computacional utilizada para realizar reconhecimento facial e análise de expressão facial por meio de

extração de recursos de imagens, vídeos e etc. É composto por redes neurais convolucionais e recorrentes.

4.1. Aprendizes

Logistic Regression (Regressão Logística): é um algoritmo de análise estatística, amplamente utilizado para classificação. Possui as regularizações *LASSO (L1)* ou *RIDGE (L2)* que tem como entrada um conjunto de dados ou pré-processador de dados e entrega como saída um aprendiz, um modelo treinado com coeficientes de regressão logística.

Esse modelo funciona apenas para tarefas de classificação. Esse aprendiz pode receber um nome e é possível escolher a regularização apropriada para melhor aprendizado. Existe a opção de definir a intensidade, porém, a recomendada na documentação é a $C=1$. A LR executa a tarefa de pré-processamento padrão, no caso da inexistência de um pré-processador de dados, ela remove colunas vazias, imputa valores ausentes com valores médios, remove instâncias com valores alvo desconhecidos e transforma variáveis categóricas em contínuas (com codificação one-hot). Previsões podem ser feitas a partir do uso desse algoritmo.

Neural Network (Redes Neurais): é um algoritmo Multi-layer Perceptron (MLP) da *Scikit Learn* com Backpropagation (Retropropagação). O *Multi-layer Perceptron* é um algoritmo de aprendizado supervisionado que aprende uma função treinando um conjunto de dados com as vantagens de aprender modelos não lineares e aprender modelos em tempo real. e aceita como entrada um conjunto de dados ou um pré-processador.

Assim como o *Logistic Regression*, retorna um aprendiz e um modelo treinado. Essa ferramenta utiliza o Backpropagation, que é um algoritmo usado pra treinar redes neurais por meio do cálculo do gradiente da função de erro em relação aos pesos. Nesse *Widget* é permitido dar um nome ao Aprendiz, e ainda definir os seguintes parâmetros:

- Quantidade de neurônios por camada: Define a profundidade da rede . Quando a camada de neurônios não está definida, essa ferramenta se comporta como a regressão logística;

- Função de Ativação para a camada oculta: estão disponíveis as funções: Identidade, Logística, *Tanh* e a *Relu*, que a mais utilizada;
- Otimizadores de peso:
 - L-BFGS-B: um otimizador da família dos métodos quasi-Newton
 - SGD: descida de gradiente estocástico
 - *Adam*: otimizador baseado em gradiente estocástico, minimiza uma função de perda escolhida com uma função linear;
 - Alfa: Parâmetro de penalidade L2 (prazo de regularização) flutuação alfa, padrão = 0,0001;
 - Força do termo de regularização L2. O termo de regularização L2 é dividido pelo tamanho da amostra quando adicionado à perda.
 - *Max iterations*: número máximo de iterações.

Além disso, a ferramenta pode produzir um relatório e também aplicar as opções escolhidas de forma automática. Assim como o *Logistic Regression*, a Rede Neural usa um pré-processador padrão executa todas as etapas de pré-processamento de dados da LR, além normalizar os dados centralizando para a média e dimensionando para o desvio padrão de 1.

4.2. Teste e pontuação

"O objetivo de um modelo de aprendizado de máquina é generalizar para novos dados "(JAMES et al., 2013, p. 25). Nesse sentido, avaliação do desempenho do modelo é fundamental para garantir que o modelo possa generalizar bem. Para tanto, as ferramentas *Test na Score* e *Matrix Confusion* foram selecionadas para pontuarem o desempenho dos Aprendizes.

Test and Score: Teste e pontuação: testa algoritmos de aprendizagem em dados, tem como entrada dados, dados em teste ainda pode receber um ou mais aprendizes. O *widget* apresenta como saída os resultados da avaliação. Essa ferramenta mostra uma tabela com medidas do classificador tais como Precisão, precisão de classificação e área sob a curva e em seguida gera resultados da avaliação.

Métodos de amostragem suportado pelo *Test and Score*.

Os métodos de amostragem são úteis para avaliar o quanto um modelo é capaz de generalizar por meio de um determinado conjunto de dados e como será seu desempenho em um conjunto de dados diferente, ou seja, verifica se o modelo realmente aprendeu, testando-o em um novo conjunto de dados (KOHAVI, 1995).

Validação cruzada (*Cross-validation*): Essa técnica divide os dados em um determinado número de dobras (partes), geralmente 5 ou 10. O algoritmo é testado mantendo exemplos de uma dobra de cada vez; o modelo é induzido a partir de outras dobras e exemplos da dobra retida são classificados. Isso é repetido para todas as dobras. Essa técnica de validação cruzada é denominada K-Fold.

Validação cruzada por recurso (*Cross validation by feature*): A validação cruzada por recurso executa a validação cruzada, mas as dobras são definidas pelo recurso categórico selecionado a partir dos meta-recursos.

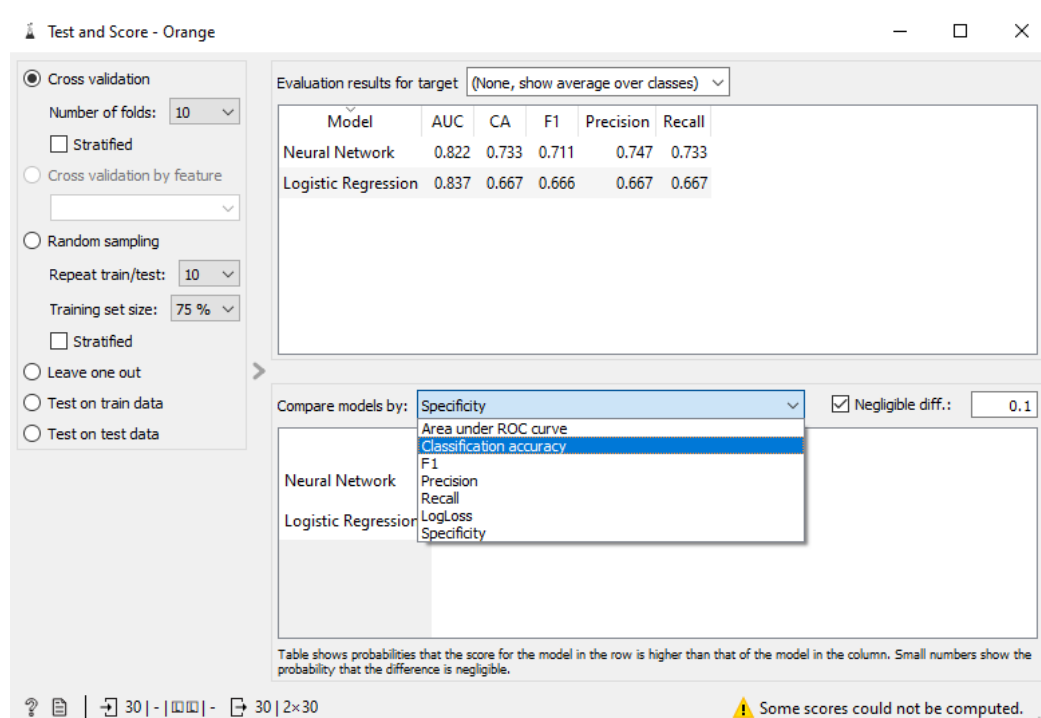
Amostragem aleatória (*Random sampling*): amostragem aleatória- A amostragem aleatória divide aleatoriamente os dados no conjunto de treinamento e teste na proporção dada (por exemplo, 70:30); todo o procedimento é repetido por um número especificado de vezes. Essa técnica de validação cruzada é denominada Holdout.

“Deixar algo de fora” (*Leave-one-out*): assim como os métodos acima, também é uma técnica de validação cruzada, é semelhante, às demais, mas mantém uma instância de cada vez, induzindo o modelo de todas as outras e, em seguida, classificando as instâncias mantidas. Este método estável, confiável, porém, muito lento.

Teste em dados de treinamento (Test on train data): o teste em dados de treinamento usa todo o conjunto de dados para treinamento e, em seguida, para teste. Este método praticamente sempre dá resultados errados.

Teste em dados de teste (Test on test data): os métodos acima usam apenas os dados do sinal de dados. Para inserir outro conjunto de dados com exemplos de teste (por exemplo, de outro arquivo ou alguns dados).

Além dos métodos acima citados há também a Estratificação, mostrada na ferramenta como Estratificação (*Stratify*): esse recurso estratifica amostras dentro o conjunto de dados a servir como treino e essa(s) amostra(s) se comporta como os dados, representando assim todo o conjunto de dados.

Figura 14 – *Test and Score*

Fonte: Autora (2023).

A ferramenta, de acordo com a Figura 14, permite comparar os dois aprendizes de acordo com as métricas abaixo, para verificar qual dos algoritmos é definitivamente superior ao outro segundo as métricas usadas por ele. As seguintes métricas são usadas pela ferramenta para avaliar os algoritmos de aprendizagem na tarefa de classificação:

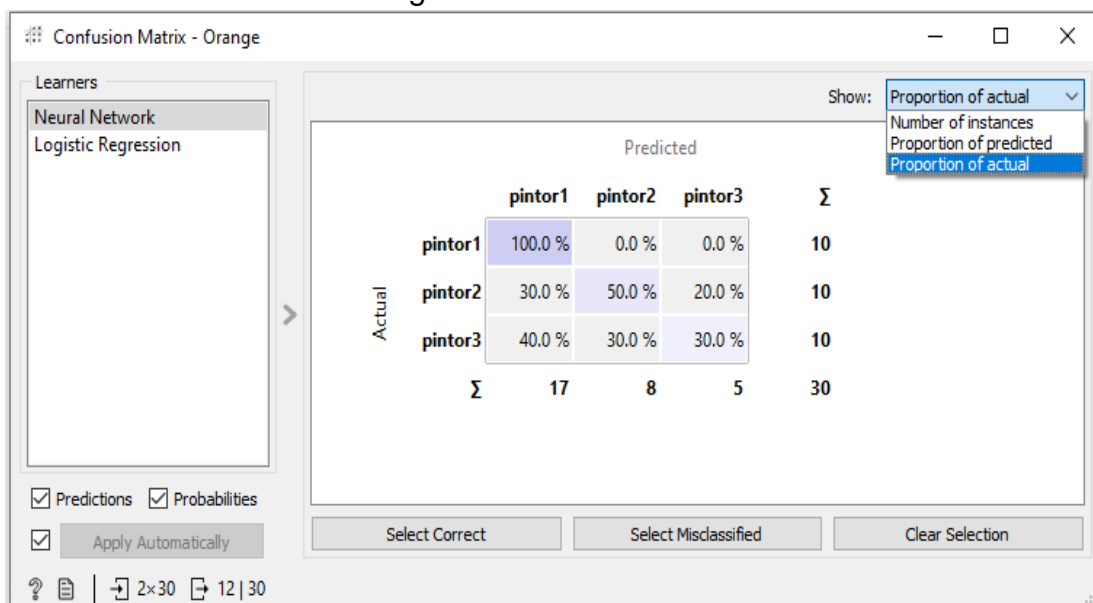
- Curva ROC: área sob a curva de operação do receptor;
- Precisão da classificação- proporção de exemplos corretamente classificados;
- F-1: é uma média harmônica ponderada de precisão e recall;
- A precisão é a proporção de verdadeiros positivos entre as instâncias classificadas como positivas;
- Recall: proporção de verdadeiros positivos entre todas as instâncias positivas nos dados;
- Especificidade: é a proporção de verdadeiros negativos entre todas as instâncias negativas;
- LogLoss: ou perda de entropia cruzada leva em consideração a incerteza de sua previsão com base enquanto ela varia do rótulo real;

- Tempo de treinamento: tempo cumulativo em segundos usado para modelos de treinamento;
- Tempo de teste - tempo cumulativo em segundos usado para testar modelos.

Matriz de Confusão: Ferramenta de avaliação dos aprendizes, ela demonstra as proporções de acertos e erros entre a classe real e a prevista e por ela pode-se extrair métricas importantes para a avaliação de um modelo de classificação como Acurácia, Recall, precisão, dentre outras. Os resultados são em forma de uma matriz quadrada, onde as colunas representam as classes previstas e as linhas, as classes reais, ou seja, corretas.

A Matriz de confusão admite como entrada resultados de testes de avaliação e retorna subconjuntos de dados selecionados da matriz e dados com informações adicionais sobre uma instância de dados selecionada. Quando há mais de uma aprendiz, é necessário selecionar o que se deseja avaliar. Na opção Show pode-se selecionar quais dados se deseja visualizar, números de instâncias classificadas correta ou erradamente, proporção da previsão e proporção real.

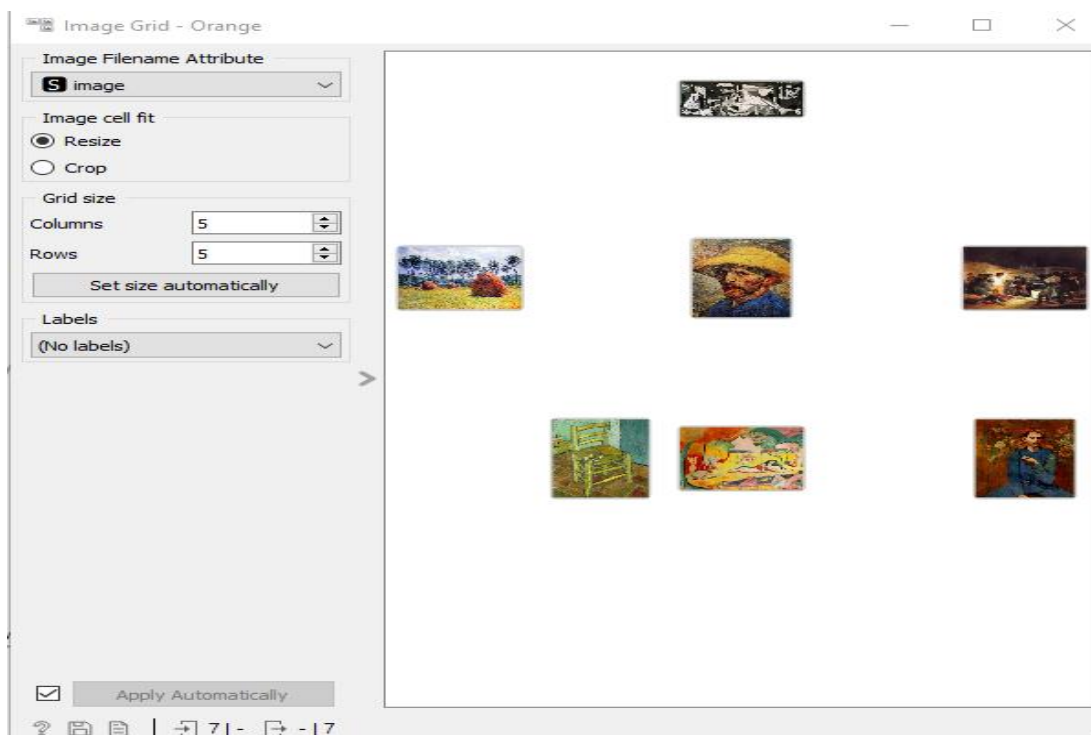
Figura 15 – Matriz de Confusão



Fonte: Autora (2023).

Conforme mostra a Figura 15, pode-se fazer a opção de demonstração dos resultados da Matriz sob os seguintes critérios: O número de instâncias, proporção da classe prevista, ou Proporção da classe real.

Figura 16 – Image Grid.



Fonte: Autora (2023).

Grade de imagens(*Image Grid*): é uma grade de imagens que mostra as imagens classificadas e agrupadas por semelhança, conforme mostra a Figura 16.

4.3. Reflexões e experimentos realizados

Neste trabalho, foram coletadas 30 imagens dos pintores Monet, Picasso e Van Gogh, ambas sem qualquer rótulo que as pudesse identificar. O que identifica esse tipo de aprendizagem como Aprendizagem não - supervisionada e a técnica de

classificação usada é também chamada de Classificação supervisionada, visto que o nome das classes era previamente conhecida pintor1, pintor2 e pintor3.

A Rede Neural do *Image Embedding* configurada para esses testes foi a *Inception V3*.

Para análise detalhada foram usados o melhor e o pior desempenho de cada algoritmo como exemplares e no final os dois melhores resultados foram testados novamente com novos parâmetros.

Como ferramentas de avaliação e desempenho foram usadas o *Test and Score*, o qual utiliza métricas da Matriz de Confusão, juntamente com as técnicas de amostragens de dados, Validação cruzada (*Cross validation*), Validação cruzada por recurso (*Cross validation by feature*), Amostragem Aleatório (*Random sampling*), “Deixar algo de fora” (*Leave-one-ou*), Teste em dados de teste (*Test on train data*), Teste em dados de teste (*Test on test data*) e Estratificação (*Stratified*), para avaliar os Aprendizes *Logistic Regression* e *Neural Network*, os quais foram escolhidos como Aprendizes.

Os seguintes parâmetros foram configurados nos Aprendizes: *Logistic Regression* e *Neural Network*, e na Ferramenta *Test and Score*, conforme a Tabela 1.

Tabela 1 – Configurações feitas no *Logistic Regression* e no *Test and Score*

Algoritmo	Experimento	Captador de imagem	Incorporador	Regularização	Tamanho	repeat trainites	Random Sampling	Training set size	Number of folds	Evaluation results for target	Test and Score					Abordagem
											AuC	CA	F1	Precision	Recall	
Logistic Regression	1	Import Image	Image Embedding	Lasso (L1)	C=1	10	75%	5	x	0.830	0.700	0.700	0.700	0.700	Cross validation	
	2			Lasso (L1)	C=1	10	75%	5	x	0.862	0.767	0.767	0.767	0.767	Cross validation, Stratified1	
	3			Lasso (L1)	C=1	10	75%	5	x	0.835	0.700	0.698	0.698	0.700	Stratified1, Random Sampling	
	4			Lasso (L1)	C=1	10	75%	5	x	0.803	0.675	0.671	0.670	0.675	Stratified1, Stratified2, Rando Sampling	
	5			Lasso (L1)	C=1	10	75%	5	x	0.847	0.700	0.689	0.752	0.667	Stratified1, Stratified2, Leave-one-out	
	6			Lasso (L1)	C=1	10	75%	5	x	1.000	1.000	1.000	1.000	1.000	Stratified1, Stratified2, Test on train data	
	7			Lasso (L1)	C=1	10	75%	5	x	Não retornou nenhum valor					Stratified1, Stratified2, Test on test data	
	8			Lasso (L1)	C=1	10	75%	5	x	0.862	0.767	0.767	0.767	0.767	Cross validation	

Fonte: Autora (2023).

Em um primeiro momento foram feitos 16 (dezesseis), testes, sendo 8 (oito) testes para cada ferramenta. Após essa etapa de testes, foi foram feitos mais 2(dois) testes para cada algoritmo.

A primeira etapa de testes, conforme a Tabela 1, está com as seguintes configurações: os itens Algoritmo, Captador de imagem, Incorporador, Regularização Lasso1, Tamanho, Quantidade de repetição treino/teste, Tamanho do conjunto de treino e número de dobras foram deixados fixos e os demais, foram sendo configurados à medida que cada teste era feito e os resultados mensurados.

O objetivo dessa escolha dos itens fixos, foi para poder observar como o Aprendiz se comporta de acordo com os métodos de amostragem de dados presentes na ferramenta. Quando uma configuração de método de amostragem não é selecionada, a ferramenta Teste and Score, por padrão deixa marcada a opção *Random sampling* (Amostragem aleatória), para que haja separação de dados de treino de dados de teste.

A Regularização Lasso1, que separa os recursos relevantes e torna o modelo mais leve, zerando os coeficientes, embora seja mais apropriada para conjuntos de dados grandes, foi deixada como fixa para poder ser comparada com a Ridge2 que outra regularização mas, que não zera os coeficientes.

A quantidade de dados usadas para treinar é de 75% (setenta e cinco por cento), definida pela opção *Training set size*, da ferramenta *Test and Score*, isso quer dizer que da quantidade de dados iniciais, no caso 30(trinta imagens), 75% ficará para treino e 25% servirá para teste.

Dos parâmetros escolhidos acima, resultaram as seguintes métricas descritas na tabela 2:

Tabela 2 - Métricas extraídas do desempenho do Aprendiz *Logistic Regression*, de acordo com as técnicas de amostragem de dados

Algoritmo	Experimento	Test and Score					Abordagem
		AuC	CA	F1	Precision	Recall	
Logistic Regression	1	0.830	0.700	0.700	0.700	0.700	Cross validation
	2	0.862	0.767	0.767	0.767	0.767	Cross validation, Stratified1
	3	0.835	0.700	0.698	0.698	0.700	Stratified1, Random Sampling
	4	0.803	0.675	0.671	0.670	0.675	Stratified1, Stratified2, Rando Sampling
	5	0.847	0.700	0.689	0.752	0.667	Stratified1, Stratified2, Leave-one-out
	6	1.000	1.000	1.000	1.000	1.000	Stratified1, Stratified2, Test on train data
	7	Não retornou nenhum valor					Stratified1, Stratified2, Test on test data
	8	0.862	0.767	0.767	0.767	0.767	Cross validation

Fonte: Autora (2023).

As métricas, conforme mostra a Tabela 2 são extraídas da Matriz de Confusão, as quais são: AuC - área sob a Curva Roc, ca - Acurácia, F1-Score, Precisão e Sensibilidade, que é o Recall. Observa-se que os resultados variaram de acordo com as técnicas de Amostragem de dados empregada. Nos resultados acima, verificou-se que o Experimento 2 obteve as melhores métricas, e gerou a seguinte Matriz de Confusão:

Tabela 3 - Matriz de Confusão do Experimento 2

Confusion Matrix					
Confusion matrix for Logistic Regression (showing number of instances)					
		Predicted			
		pintor1	pintor2	pintor3	Σ
Actual	pintor1	9	0	1	10
	pintor2	1	7	2	10
	pintor3	0	3	7	10
Σ		10	10	10	30

Fonte: Autora (2023).

Na Tabela 3, na Diagonal Principal, verificamos a quantidade de classificações acertadas (Verdadeiro Positivo – TP e os Verdadeiros Negativos-TN), nas outras células, estão os classificados incorretamente, que são os Falso Negativo – FN, Falso Positivo – FP. Em uma Matriz de confusão multiclasse, como é o caso da Tabela 3, as demais classes, exceto as que compõem a Diagonal Principal são analisadas como Falso Negativo – FN ou Falso Positivo – FP de forma individualizada. Dessa forma, o desempenho das classes individuais do Experimento 2 são demonstrados nas Figuras 17, 18 e 19.

Figura 17 – Desempenho da Classe Pintor1 do Experimento 2

Test and Score					
Settings					
Sampling type: Stratified 5-fold Cross validation					
Target class: pintor1					
Scores					
Model	AUC	CA	F1	Precision	Recall
Neural Network	0.975	0.8	0.7692307692307693	0.625	1.0
Logistic Regression (1)	1.0	0.9333333333333333	0.9	0.9	0.9

Fonte: Autora (2023).

Figura 18 – Desempenho da Classe Pintor2 do Experimento 2

Test and Score					
Settings					
Sampling type: Stratified 5-fold Cross validation					
Target class: pintor2					
Scores					
Model	AUC	CA	F1	Precision	Recall
Neural Network	0.925	0.7333333333333333	0.5555555555555556	0.625	0.5
Logistic Regression (1)	0.825	0.8	0.7	0.7	0.7

Fonte: Autora (2023).

Figura 19 – Desempenho da Classe Pintor3 do Experimento 2

Test and Score						
Settings						
Sampling type: Stratified 5-fold Cross validation						
Target class: pintor3						
Scores						
Model	AUC	CA	F1	Precision	Recall	
Neural Network	0.85	0.7333333333333333	0.5	0.6666666666666666	0.4	
Logistic Regression (1)	0.8		0.8	0.7	0.7	

Fonte: Autora (2023).

Nas Figuras 17, 18 e 19, estão as métricas das classes Pintor1, Pintor2 e Pintor3, referentes ao do Experimento 2, as quais são: Área sob a curva Roc, Acurácia, Precisão e Recall. Esse conjunto de medidas são indicadores que mostram o desempenho do modelo de aprendizado Regressão Logística .

O Modelo Pintor1, conforme a Figura 17, Obteve a melhor Acurácia: 0,93, ou seja, dentre os números de amostras, ele conseguiu classificar acertadamente (Verdadeiro Positivo TP), 93% dos casos. O modelo Pintor1 Também obteve a melhor medida da Área sob a curva Roc: 1. A Melhor Precisão, F1_Score e Recall, todas com valor 0.9, também foi da classe Pintor1. Isso indica que o modelo Pintor1 tem uma capacidade superior em generalizar que os demais modelos, caso fosse para escolher um dos modelos, esse seria o melhor.

De maneira global, conforme mostram as a Tabelas 2 e 3, verifica-se que o modelo que melhor performou foi o Pintor1, pois, de forma geral, conseguiu os melhores índices. O algoritmo gerou as seguintes métricas: na área sob a Curva ROC a classe gerou a pontuação AuC = 0.862. A acurácia, que é medida total de acertos, dividida pelo conjunto de treinamento foi de: CA= 0.767, indicando que o modelo classificou como correto 76% das imagens. O F1-Score, que é a relação harmônica entre a Precisão e Recall ficou em: F1= 0.767.

A Precisão Balanceada, medida pela qual é possível verificar a média do desempenho das classes, teve o valor de Precisão =0.767, no modelo Pintor1. O valor do Recall foi de 0.767. O recall mede a precisão preditiva do modelo. O

Experimento 2 usou com técnica de amostragem de dados, a Validação cruzada, técnica na qual os há uma separação de dados para treinamento dos dados usados para teste.

O Experimento 4, conforme mostram as Tabelas 2 e 3, apresentou as métricas mais baixas as quais são: AuC = 0.803, CA = 0.675, F1= 0.671, Precision =0.670, e Recall = 0.675. As abordagens de amostragem de dados utilizadas foram: Estratificação, técnica na qual é garantido que todas as classes tenham representatividade (Stratified1) com uma segunda estratificação, juntamente com a Amostragem aleatória (Random Sampling), técnica na qual os dados no conjunto de treinamento e teste são divididos aleatoriamente na proporção dada (por exemplo, 70:30)

Analisando o Experimento 6, conforme mostram as Tabelas 2 e 3, foi de maior pontuação verifica-se que a opção Teste nos dados de treino (*Test on train*), do *Test and Score* estava marcada, ou seja, os testes são feitos nos dados de treino, sem serem confrontados com dados de teste, nesse caso, o modelo decora o resultado e acerta todas as opções, a própria ferramenta Orange, assim como a literatura afirmam não ser uma técnica de amostragem de dados confiável que superestima os resultados.

Por questão de analogia, todas as medidas do Experimento 6 serão deixadas de fora das análises futuras deste trabalho.

De uma forma geral, conclui-se que no Experimento 2 o modelo aprendendo melhor os demais, pois o F1-Score, indica uma pontuação superior às demais e isso é confirmado pela Área sob a Curva Roc, que indica que quanto mais a métrica se aproxima de 1, mais capaz é o modelo de separar as classificações verdadeiras das falsas. A Acurácia, foi bem elevada, porém não é indicativo totalmente confiável para avaliar um modelo, visto que ela considera as classes individualmente.

A Avaliação do Aprendiz Neural Network também seguiu as configurações gerais do *Test and Score* mencionadas acima, e também a mesma quantidade de imagens, os testes dos dois aprendizes foram feitos simultaneamente, uma condição que o *Widget Test and Score* aceita normalmente. A função de Ativação ReLu, o

otimizador de pesos Adam e a Regularização, são padrões para início de uso da ferramenta. A tabela 4 apresenta as configurações do Algoritmo *Neural Network*.

Tabela 4 – Configurações feitas no *Neural Network* e no *Test and Score*

Algoritmo	Experimento	Captador de imagem	Incorporador	Neurônios por camada	Função de Ativação	Solucionador de otimização de pesos	Regularização	Max iterations	repeat train/test	Random Sampling	Training set size	Number of folds	Evaluation results for target	Test and Score					Abordagem	
														None, show average over classes	AuC	CA	F1	Precision		Recall
Neural Network	1			20	ReLU	Adam	Alfa = 0,0001	30	10	75%	5	x	0.830	0.567	0.518	0.606	0.567	Cross validation		
	2								10	75%	5	x	0.858	0.633	0.608	0.639	0.633	Cross validation, Stratified1		
	3								10	75%	5	x	0.813	0.625	0.599	0.591	0.625	Stratified1, Random Sampling		
	4	Import Image							10	75%	5	x	0.817	0.588	0.564	0.579	0.588	Stratified1, Stratified2, Rando Sampling		
	5	Image Embedding							10	75%	5	x	0.842	0.667	0.648	0.752	0.667	Stratified1, Stratified2, Leave-one-out		
	6								10	75%	5	x	1.000	1.000	1.000	1.000	1.000	Stratified1, Stratified2, Test on train data		
	7								10	75%	5	x	Não retornou nenhum valor					Stratified1, Stratified2, Test on test data		
	8								10	75%	5	x	0.858	0.633	0.608	0.639	0.633	Cross validation		

Fonte: Autora (2023).

A tabela 4 mostra as configurações feitas no algoritmo *Neural Network* e na ferramenta *Test and Score*. A rede Neural fez os testes com 20 Neurônios por camada, tendo como otimizador de pesos, o Adam, como função de Ativação a ReLU e com 30 (trinta) iterações, ou seja, ela vai repetir o treino 30 vezes. Esses valores, como já mencionados são valores padrão de início de operação da máquina.

Tabela 5 – Métricas extraídas do desempenho do Aprendiz Neural Network, de acordo com as técnicas de amostragem de dados

Algoritmo	Experimento	Test and Score					Abordagem
		AuC	CA	F1	Precision	Recall	
Neural Network	1	0.830	0.567	0.518	0.606	0.567	Cross validation
	2	0.858	0.633	0.608	0.639	0.633	Cross validation, Stratified1
	3	0.813	0.625	0.599	0.591	0.625	Stratified1, Random Sampling
	4	0.817	0.588	0.564	0.579	0.588	Stratified1, Stratified2, Random Sampling
	5	0.842	0.667	0.648	0.752	0.667	Stratified1, Stratified2, Leave-one-out
	6	1.000	1.000	1.000	1.000	1.000	Stratified1, Stratified2, Test on train data
	7	Não retornou nenhum valor					Stratified1, Stratified2, Test on test data
	8	0.858	0.633	0.608	0.639	0.633	Cross validation

Fonte: Autora (2023).

As métricas da Tabela 5 são extraídas da Matriz de Confusão, as quais são: AuC - área sob a Curva Roc, ca - Acurácia, F1-Score, Precisão e Sensibilidade, que é o Recall.

Observa-se que os resultados variaram de acordo com as técnicas de amostragem de dados empregada, assim como o algoritmo testado anteriormente. Na Avaliação das métricas geradas observou-se que o Experimento 5 obteve as melhores métricas, gerando a seguinte Matriz de Confusão:

Tabela 6 – Matriz de Confusão Experimento 5

Confusion matrix for Neural Network (showing number of instances)						
		Predicted			Σ	
		pintor1	pintor2	pintor3		
Actual	pintor1	10	0	0	10	
	pintor2	4	6	0	10	
	pintor3	3	3	4	10	
Σ		17	9	4	30	

Fonte: Autora (2023).

A tabela 6 é uma Matriz de Confusão do Experimento 5 mostrando os resultados de acertos na Diagonal Principal e os erros nas outras células. O desempenho das classes individuais do Experimento 5 são demonstrados nas Figuras 20, 21 e 22.

Figura 20 – Desempenho da Classe Pintor1 do Experimento 5

Test and Score					
Settings					
Sampling type: Leave one out					
Target class: pintor1					
Scores					
Model	AUC	CA	F1	Precision	Recall
Neural Network		0.7666666666666667	0.7407407407407407	0.5882352941176471	1.0
Logistic Regression (1)		0.9333333333333333		0.9	0.9

Fonte: Autora (2023).

A Figura 20 mostra o desempenho da Classe Pintor1 do Experimento 5 do Aprendiz *Neural Network*.

Figura 21 – Desempenho da Classe Pintor2 do Experimento 5

Test and Score					
Settings					
Sampling type: Leave one out					
Target class: pintor2					
Scores					
Model	AUC	CA	F1	Precision	Recall
Neural Network		0.7666666666666667	0.631578947368421	0.6666666666666666	0.6
Logistic Regression (1)		0.7333333333333333	0.6666666666666666	0.5714285714285714	0.8

Fonte: Autora (2023).

A Figura 21 mostra o desempenho da Classe Pintor2 do Experimento 5 do Aprendiz *Neural Network*.

Figura 22 – Desempenho da Classe Pintor3 Experimento 5

Test and Score						
Settings						
Sampling type: Leave one out						
Target class: pintor3						
Scores						
Model	AUC	CA	F1		Precision	Recall
Neural Network			0.8	0.5714285714285715		1.0 0.4
Logistic Regression (1)		0.7333333333333333			0.5 0.6666666666666666	0.4

Fonte: Autora (2023).

Nas Figuras 20, 21 e 22, estão as métricas das classes Pintor1, Pintor2 e Pintor3, concernentes ao Experimento 5, as quais são: Área sob a curva Roc, Acurácia, Precisão e Recall. Esse conjunto de medidas são indicadores que mostram o desempenho do modelo de aprendizado Rede Neural de forma individualizada por cada classe.

A partir dessas métricas observou-se que: o modelo Pintor1, conforme a figura 20, obteve o melhor desempenho. A Acurácia foi de 67%, ou seja, o modelo classificou como Verdadeiro Positivo ou Verdadeiro Negativo 67% de um total de 30 imagens. O *F1-Score* foi de 64%, A Precisão foi de 75%, já o *Recall* foi de 66%.

As Técnicas de Amostragens de dados adotadas foram *Stratified1*, *Stratified2*, *Leave-one-out*, que são Estratificação no qual o conjunto de dados é dividido de forma que cada classe possa ter representatividade e o *Leave-one-out*, que quer dizer um fora, é uma técnica de validação cruzada no qual uma amostra é retirada para teste enquanto do conjunto de dados é usado para treino .

Para esse conjunto de dados e as configurações usadas não foi feito o cálculo da área sob a curva Roc. Esse conjunto de valores reflete o desempenho do algoritmo.

A Classe Pintor2, conforme mostra a Figura 21, obteve 76% de Acurácia, 63% de *F1-Score*, 66% de Precisão e 60% de Recall. São valores medianos e bem

homogêneos, o que pode facilitar na escolha ou no descarte do modelo, uma vez que os ganhos ou perdas estão balanceados, porém, cada caso tem suas especificidades para adoção de um modelo de classificação.

A classe Pintor3, de acordo com a Figura 22, obteve a Precisão no valor 1, esse é o valor máximo para essas métricas, considerado um ótimo valor. A Precisão é uma métrica muito importante em casos em que a perda em se avaliar um dado como falso positivo é grande, em casos de saúde em especial. Também essa classe teve o menor valor de Recall: 40%. O Recall é a proporção de verdadeiros positivos para tudo que é positivo. O modelo Pintor3 a Acurácia foi de 80% e o F1-Score de 57%. Os resultados estão bastante heterogêneos, porém com um bom índice de precisão.

Como métricas globais, consoante com a Tabela 5, os indicadores do Experimento 5 foram: Auc=0.842, Acurácia=0.667, F1=Score=0.648, *Precision*=0.752 e Recall=0.667. Área abaixo da Curva Roc foi a melhor métrica, as medidas são satisfatórias, esse índice nos diz que o modelo possui uma capacidade mediana de abstrair, ou seja de ser usado em uma classe real de dados.

O Experimento 1 obteve as piores métricas globais, que são: Área sob a curva Roc 83% de pontuação, Acurácia 56%, F1-Score 51%, Precisão 60% e Recall de 56%. Valores considerados medianos para uma Rede Neural.

A abordagem para amostragem dos dados foi a Validação cruzada, técnica na qual os há uma separação de dados para treinamento dos dados usados para teste. As medidas obtidas são razoáveis e são passíveis de serem melhoradas.

O melhor resultado do Aprendiz Regressão Logística foi o Experimento 2, com medidas Área sob a curva ROC, 0.862, Acurácia, 0.767, F1-Score, 0.767 Precisão Balanceada, 0.767 e Recall 0.767, tendo como Abordagem de amostragem de dados a Validação cruzada e a estratificação.

O Aprendiz *Neural Network* foi o Experimento 5, com as seguintes métricas: Área sob a curva ROC, 0.842, Acurácia, 0.667, *F1-Score*, 0.648, Precisão Balanceada, 0.752 e Recall, 0.667, tendo com abordagem de amostragem de dados *Stratified1*, *Stratified2*, *Leave-one-out*.

Diante dos dados acima verificou-se que o desempenho do Aprendiz Regressão Logística foi superior ao do Aprendiz Neural Network. Nos experimentos feitos observa-se que o conjunto de dados era pequeno, 30 (trinta amostras) divididas em três classes, com dados equilibrados, as classes continham 10 (dez) amostras

cada. As técnicas de Amostragem influenciaram nos resultados, no Experimento 2 com o LR, a abordagem que gerou o melhor resultado foi a validação cruzada, já no *Neural Network* as abordagens *Stratified1*, *Stratified2*, *Leave-one-out*, geraram melhores métricas.

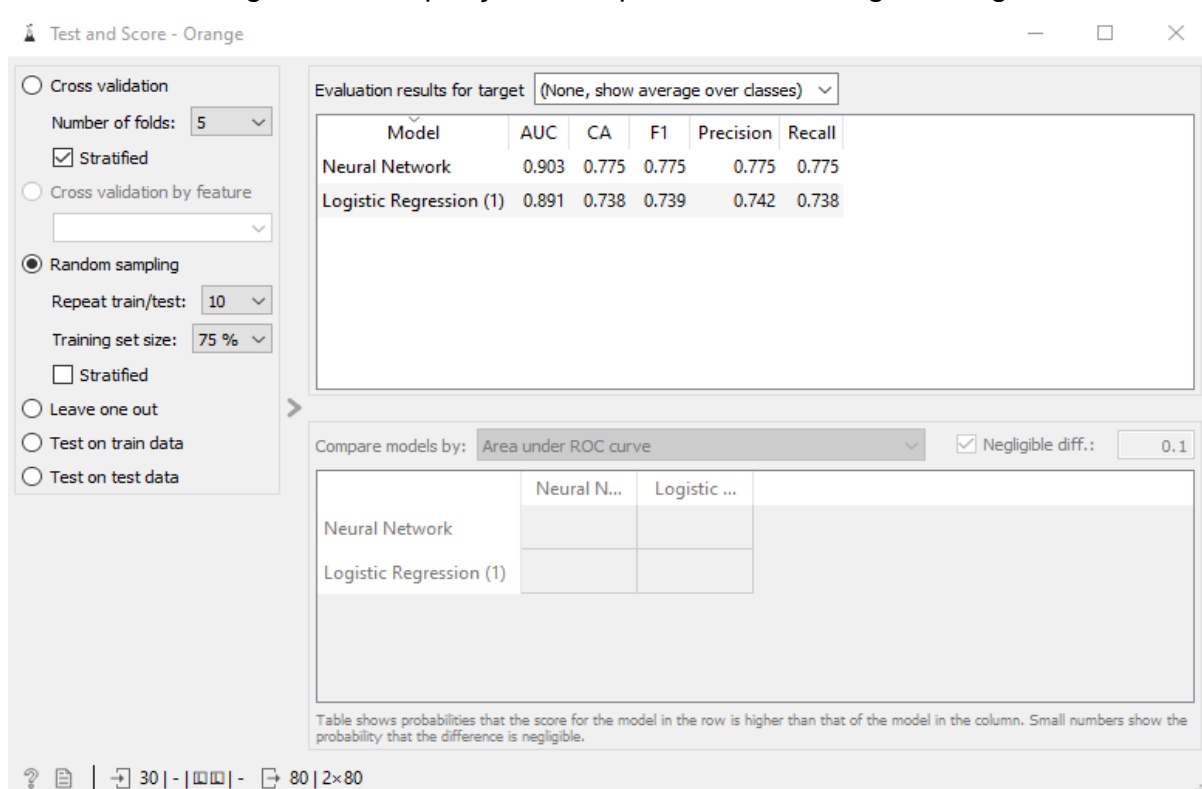
Em todo o *WorkFlow* foram usadas duas redes neurais ao todo, a Inception V3, conforme a Figura 24, que colaborou com os dois Aprendizes e a Neural Network do na evolução do *Logistic Regression*.

A Matriz de Confusão foi determinante para analisar a performance dos dois algoritmos e ela indicou que a Regressão Logística possui o melhor desempenho para esse caso concreto em comparação com a Rede Neural.

Diante dos fatos e de que um princípio do Machine Learning é a repetição do treino para aperfeiçoar o aprendizado, foi feita a segunda etapa de testes de repetição dos Experimentos 2 e 5, dos Aprendizes, mas com parâmetros diferentes.

As seguintes configurações foram usadas para o Experimento2 do *Logistic Regression*: Regularização Ridge1, validação cruzada e a opção da Rede Neural Painters (pintores) no *Embedding Image*, conforme a Figura 25.

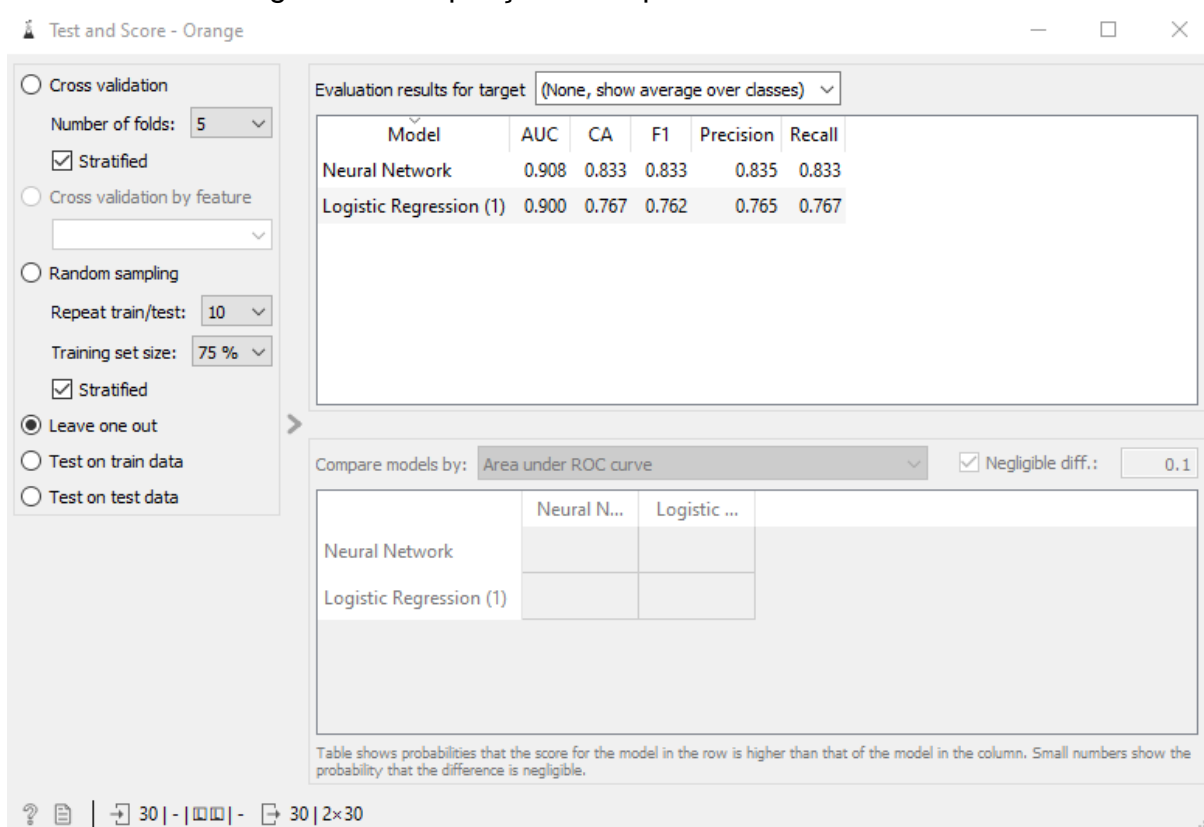
Figura 23 – Repetição do Experimento2 do *Logistic Regression*



Fonte: Aurora (2023).

A Figura 23 mostra os parâmetros usados no Experimento 2 do *Logistic Regression*, cujas métricas são mostradas na Tabela 6

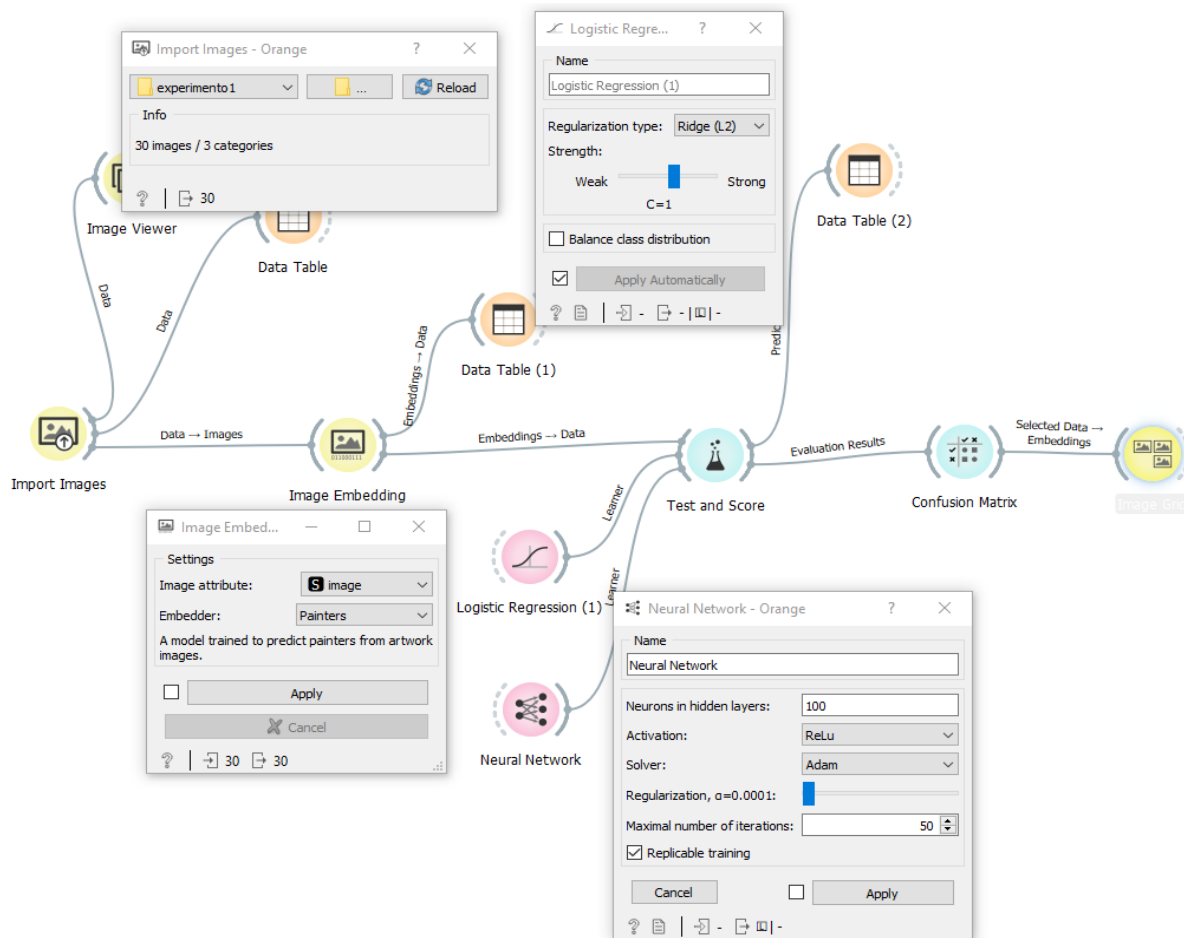
Figura 24 – Repetição do Experimento 5 do Neural Network



Fonte: Autora (2023).

A Figura 24 mostra o desempenho e as configurações da ferramenta *Test and Score* para a repetição do Experimento 5. A configuração *Neural Network* foi feita com 100 (cem) neurônios por camada, 50 (cinquenta) iterações Regularização Adam e com as abordagens de amostragem de dados *Stratified1*, *Stratified2* e *Leave-one-not* e com a opção da Rede Neural *Painters* (pintores) no *Embedding Image*, mostrada na Figura 25.

Figura 25 - Workflow com as configurações finais



Fonte: Autora (2023).

Na Figura 25 está o fluxo das ferramentas configuradas para a repetição dos Experimentos 2 e 5 como Os Aprendizes *Logistic Regression* e *Neural Network*. A Rede *Painters*, especializada em classificação de pintores foi usada para essa fase dos testes. Os resultados finais dos testes estão na Tabela 7.

Tabela 7- Experimentos 2 e 5

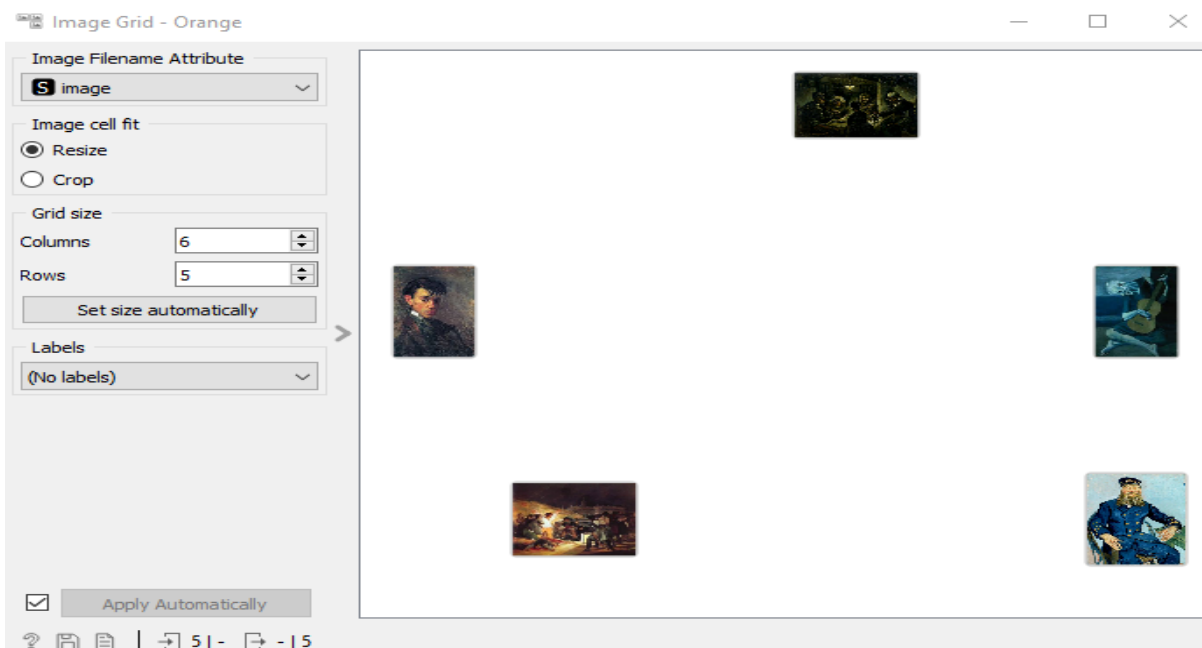
		AuC	CA	F1	precisio	Recall
Experimento 2	Logistic Regression	0.891	0.738	0.739	0.742	0.738
Experimento 2	Logistic Regression	0.900	0.767	0.762	0.765	0.767
Experimento 5	Neural Network	0.908	0.833	0.833	0.835	0.833
Experimento 5	Neural Network	0.908	0.833	0.833	0.835	0.833

Fonte: Autora (2023).

A Tabela 7 mostra a evolução dos modelos treinados nos algoritmos *Logistic Regression* e *Neural Network*, onde constam suas respectivas métricas. Foi verificado que com as mudanças na arquitetura dos algoritmos, conforme mostra Figura 25, e no aumento de quantidade de treino houve uma melhora significativa para os dois Aprendizes.

Comparando resultados finais dos Aprendizes *Logistic Regression* e *Neural Network*, Foi constatado que ambos melhoraram seus índices com o aumento da quantidade de treino e mudanças nos parâmetros da ferramenta Test and Score, juntamente com as mudanças na arquitetura dos Aprendizes. Os testes finais geraram uma grade de semelhanças das imagens exibidas pelo Image Grid, nas Figuras 26 e 27.

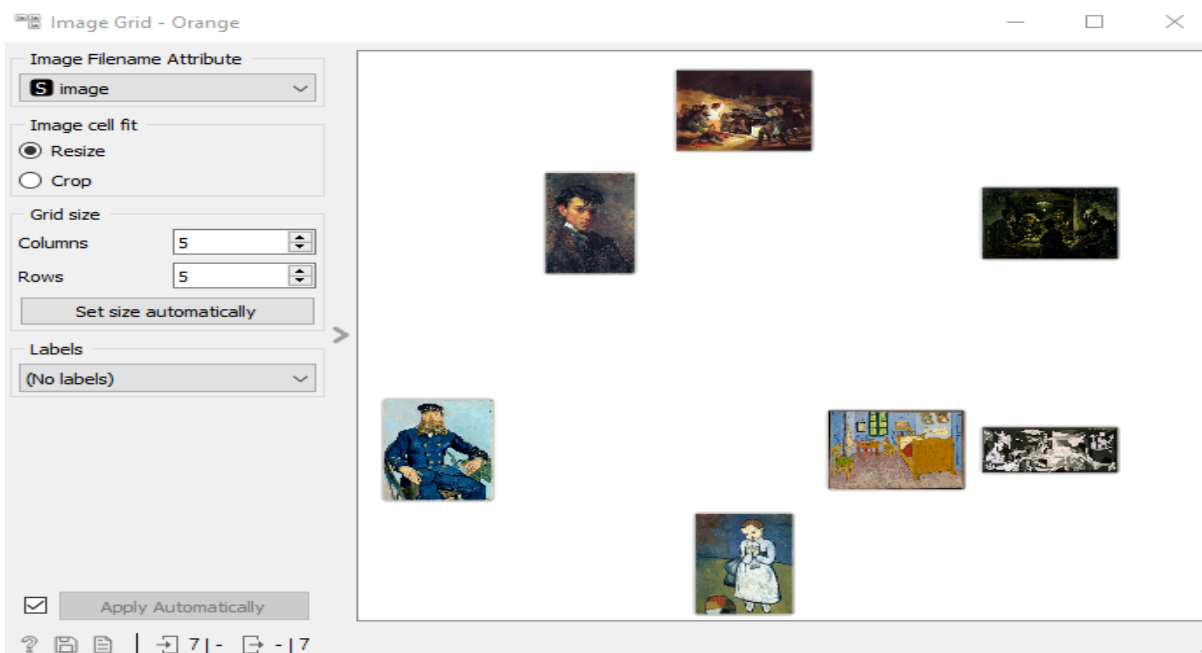
Figura 26 - Imagens separadas por semelhanças Rede Experimento 5-
Neural Network



Fonte: Autora (2023).

Na Figura 26 a disposição das imagens e a organização dos conjuntos, que são agrupados por semelhanças fica latente a clareza da classificação.

Figura 27 - Imagens separadas por semelhanças Rede Experimento 2-
Logistic Regression



Fonte: Autora (2023).

Na Figura 27 a disposição das imagens mostra um conjunto de imagens um pouco menos fácil de verificar o agrupamento, que é feito por semelhanças entre as imagens.

5. CONCLUSÃO

O trabalho procurou medir o desempenho dos algoritmos Redes Neurais e Regressão Logística por meio da Matriz de Confusão, na ferramenta Orange Data Mining usando uma base de dados de 30 (trinta amostras), sendo 10 (dez) para cada classe, ou seja, foi usado um valor mínimo de amostras, baseado na recomendação da literatura, porém não o valor exato recomendado, foi um valor abaixo, mas dentro das orientações exigidas de acordo com a complexidade do problema. Embora a quantidade de dados tenha sido restrita, as técnicas de validação cruzada, e a diversificação dos dados compensaram essa falta.

Para alcançar o objetivo proposto, foram realizadas a princípio pesquisas bibliográficas que vieram a gerar a Revisão da Literatura com assuntos envolvendo Inteligência Artificial, Classificação de Imagens, *Machine Learning* e outros assuntos correlatos ao tema. O Fluxo de Trabalho foi definido após as pesquisas bibliográficas e da escolha ferramenta. Os procedimentos metodológicos foram pesquisas realizadas na documentação da ferramenta além das mencionadas.

Os testes iniciais, no caso, 16 (dezesseis), foram feitos com os dois Aprendizes simultaneamente, como os mesmos parâmetros, na ferramenta *Test and Score*, porém com as especificidades da arquitetura de cada algoritmo.

A Regressão Logística tinha a Regularização Lasso¹ e a Rede Neural possuía 20 neurônios por camada. Os treinos foram repetidos por trinta 30 (trinta vezes). A Rede Neural da ferramenta da Incorporadora *Image embedding* era a *Inception V3*.

Para a etapa final, foram feitos mais dois testes com as mudanças na arquitetura da Regressão Logística, que passou a ter a Regularização Ridge¹ e a Rede Neural, passou a ter 100 neurônios por camada. Os testes foram repetidos por 50 (Cinquenta vezes). A Rede Neural da ferramenta da Incorporadora *Image embedding* era a *Painters*.

Os algoritmos tiveram um bom desempenho, geraram bons resultados na primeira etapa de testes, considerando que o conjunto de dados era limitado, os

resultados foram muitos bons, indicando que a arquitetura dos Aprendizes Regressão Logística e Redes Neurais são muito bons de uma forma geral.

A rede neural *Painters* pode ter influenciado o Aprendiz Regressão Logística, o que sugere que, na classificação de imagens em casos específicos pode necessitar de uso de mais de um classificador para gerar o desempenho desejado.

A Regressão Logística teve um desempenho inicial superior ao da rede neural na primeira etapa, corroborando para referendar o afirmado na literatura. Quando a quantidade de treino da rede e a quantidade de neurônios aumentaram, seu desempenho foi superado.

Uma das explicações para o desempenho superior da Regressão Logística e também mencionado na literatura é que com grandes quantidades de dados e modelos muito complexos sua Área sob a Curva Roc tende a diminuir. Em algoritmos de *Machine Learning*, quanto maior for a AuC, mais capaz de aprender com os dados e de poder ter um desempenho melhor em um conjunto de dados de teste, é o algoritmo.

Outro fator que favoreceu a Regressão Logística foi o fato de que esse algoritmo é indicado para resolução de problemas menos complexos, embora seja um algoritmo amplamente utilizado na Aprendizagem de Máquina como classificador de imagem.

O Modelo Regressão Logística teve o *Overfitting* (situação na qual o modelo não apreende, apenas memoriza), corrigido pela Regularização Lasso¹ e não sofreu *Underfitting* (Situação que deixa o modelo lento, devido a grandes conjuntos de dados), pois o conjunto de dados era pequeno.

O Aprendiz Rede Neural, na Etapa inicial de testes teve um bom resultado, comparado a Regressão Logística, não se mostrou muito bom a aprender a classificar os dados. Alguns fatores podem ter colaborado para esse desempenho da Rede.

O conjunto de dados era muito pequeno, a literatura recomenda conjuntos de dados grandes para treino e teste das Redes, e relatam também que, embora o treino de uma Rede Neural com grandes quantidades de dados demande tempo e recursos computacionais, após treinadas o seu desempenho é muito superior a muitos algoritmos de *Machine Learning*.

Devido a pequena quantidade de dados disponíveis para treino e teste, provavelmente, na primeira etapa dos experimentos, é possível que a Rede Neural

possa ter sofrido *Overfitting*, ou seja, tenha se ajustado demais aos exemplos de treinamento específicos e não tenha sido capaz de aprender bem, ou seja, de generalizar bem para novos exemplos.

O Otimizador usado para os experimentos da Rede foi o *Adam*, um otimizador baseado em gradiente estocástico, minimiza uma função de perda escolhida com uma função linear é o otimizador que vem como padrão na ferramenta Orange Data Mining quando não se escolhe outro otimizador. Não houve a exploração de outros otimizadores, embora eles sejam importantes para os ajustes de pesos das redes, foi percebido que a aplicação e teste dos mesmos merece um estudo a parte e devido também ao projeto ter alcançado o resultado convincente, e escassez de tempo, não houve essa variação de otimizadores.

A maneira de aprendizagem da Rede Neural, favorece ao uso de uma quantidade mais expressiva de dados, porém, ficou observado também que, a qualidade e diversidade dos dados, a quantidade de treinos e as Técnicas de Amostragem de Dados são fatores relevantes para um bom desempenho e confiabilidade de um modelo de Rede Neural, embora com dados escassos.

Na segunda etapa de treinos houve mudanças na arquitetura dos Aprendizes, na quantidade de treinos, e as Técnicas de Amostragem de Dados foram repetidas para os melhores modelos que tiveram melhores desempenhos na primeira etapa de testes.

A quantidade de repetição foi aumentada de 30 (trinta) para 50 (cinquenta) em ambos os Aprendizes, na Regressão Logística a regularização escolhida foi a *Ridge 1* e a Rede Neural passou a ter 100 (cem) neurônios por camada. A Rede Neural se mostrou superior ao algoritmo Regressão Logística, porém ambos foram muito bons em todas as métricas da Matriz de Confusão.

Os Experimentos demonstraram que a quantidade de treinamentos feita pelos Aprendizes é fundamental para uma melhor performance dos mesmos e aliados às Técnicas de Amostragem de Dados, podem melhorar profundamente um modelo de *Machine Learning*, além da arquitetura dos algoritmos que influenciou bastante para uma melhor performance dos algoritmos.

O experimento demonstra o que alguns teóricos afirmam sobre os algoritmos de *Machine Learning* de que não existe uma ou outra técnica melhor, o cientista de dados

usa um algoritmo de acordo com a complexidade do problema, a natureza do mesmo, os tipos de dados e outras especificidades.

A Matriz de Confusão se mostrou uma excelente ferramenta de análise, na verdade, até insubstituível talvez. A ferramenta *Orange Data Mining* se mostrou uma excelente ferramenta para análises estatísticas e descritivas de um modo geral, porém o *widget Image grid* não se mostrou muito bom visualizador de imagem, pelo menos para aquele conjunto de dados.

A experiência foi enriquecedora e como sugestão de trabalhos futuros, outros algoritmos de classificação de imagens poderiam ser testados na ferramenta *Orange Data Mining* como as Florestas aleatórias e a Máquinas Vetoriais de Suporte e com o uso de bancos de dados maiores.

REFERÊNCIAS

- A.D.Dongare; R.R. kharde; AMIT D.Kachare. **Introduction to Artificial Neural Network**. International Journal of Engineering and Innovative Technology (IJEIT). Volume 2, Issue 1, July 2012.
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1082.1323&rep=rep1&type=pdf>. Acesso em 16 de janeiro 2023.
- BISHOP, C. M. **Pattern recognition and machine learning**. Springer. (2006). Disponível em: <https://link.springer.com/book/9780387310732>. Acesso em: 17 de fevereiro de 2023.
- CHANDRA, M. A., & BEDI, S. S. **Survey on SVM and their application in image classification**. International Journal of Information Technology. (2018). doi:10.1007/s41870-017-0080-1. Disponível em: <https://link.springer.com/article/10.1007/s41870-017-0080-1>. Acesso em: 20 de março de 2023.
- CHAUHAN, Nagesh Singh. **Model Evaluation Metrics in Machine Learning**. Disponível em: <https://www.kdnuggets.com/2020/05/model-evaluation-metrics-machine-learning.html>. Acesso em 15 de abril de 2023.
- DE MILANO, Danilo; HONORATO, Luciano Barrozo. **Visão computacional**. Faculdade de Tecnologia, Universidade Estadual de Campinas, 2010. https://www.academia.edu/download/35825905/2010_IA_FT_UNICAMP_visaoComputacional.pdf Acesso em: 18 de janeiro 2023.
- DHINGRA, Sakshi; KUMAR, Dharminder. **A review of remotely sensed satellite image classification**. International Journal of Electrical and Computer Engineering. 2019. v. 9, n. 3, p. 1720. Disponível em: https://www.researchgate.net/publication/333563411_A_review_of_remotely_sensed_satellite_image_classification. Acesso em 11 de maio de 2023.
- DREISEITL, S., OHNO-MACHADO, L., KITTLER, H., VINTERBO, S., BILLHARDT, H., & BINDER, M. **A Comparison of Machine Learning Methods for the Diagnosis of Pigmented Skin Lesions**. Journal of Biomedical Informatics. 2001. 34(1), 28–36. doi:10.1006/jbin.2001.1004. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1532046401910044>. Acesso em 21 de abril de 2023.
- DUAN, Y; EDWARDS, J. S., & DWIVEDI, Y. K. **Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda**. International Journal of Information Management. 2019. 48, 63–71. doi:10.1016/j.ijinfomgt.2019.01.0. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0268401219300581>. Acesso em 13 de janeiro de 2023.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2008. Acesso em: 21 de fevereiro 2023.

GOLDBERG, Yoav. **A Primer on Neural Network Models for Natural Language Processing**. 2016. *Journal of Artificial Intelligence Research* 57. 2016. 345–420. Disponível em: <https://www.jair.org/index.php/jair/article/view/11030/26198> Acesso em: 21 de fevereiro 2023.

GRANDINI, Margherita; BAGLI, Enrico; VISANI, Giorgio. **Metrics for Multi-Class Classification: an Overview**. 2020. Disponível em: <https://doi.org/10.48550/arXiv.2008.05756>. Acesso em 24 de maio de 2023.

HASTIE, T.; TIBSHIRANI, R. & FRIEDMAN, J. **The elements of statistical learning: Data mining, inference, and prediction**. Springer. (2009). Disponível em: <https://link.springer.com/book/10.1007/978-0-387-84858-7>. Acesso em 26 de fevereiro de 2023.

HEBB, Donald Olding. **The organization of behavior: A neuropsychological theory**. Psychology press, 2005. Acesso em 12 de maio de 2023.

HEERMANN, P. D., & KHAZENIE, N. **Classification of multispectral remote sensing data using a back-propagation neural network**. *Geoscience and Remote Sensing*. IEEE Transactions. 1992. 30(1), 81-88. Disponível em: <https://ieeexplore.ieee.org/document/124218>. Acesso em 14 de junho de 2023.

JAIN, AK., DUIN, PW. & JIANCHANG Mao. **Reconhecimento de padrões estatísticos: uma revisão**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. IEEE. 2000. 22(1), 4–37. doi:10.1109/34.824819. Disponível em: <https://ieeexplore.ieee.org/abstract/document/824819>. Acesso em 20 de janeiro de 2023.

JANOWCZYK, A. & MADABHUSHI, A. **Deep Learning for Digital Pathology Image Analysis: A Comprehensive Tutorial with Selected Use Cases**. *J. Pathol. Informar*. 2016. 7, 29 Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4977982/>. Acesso em 26 de fevereiro de 2023.

JAMES, G.; WITTEN, D.; HASTIE, T., & TIBSHIRANI, R. **Introduction to Statistical Learning: with Applications in R**. Springer. 2013. (p. 25). Disponível em: <https://link.springer.com/book/10.1007/978-1-0716-1418-1> Acesso em 24 de maio de 2023.

KOHAVI, R. **A study of cross-validation and bootstrap for accuracy estimation and model selection**. In: *International joint Conference on artificial intelligence*. [S.l.: s.n.], 1995. v. 14, p. 1137–1145. Acesso em: 21 de fevereiro 2023.

LECUN, Y.; BENGIO, Y., & HINTON, G. **Deep learning**. *Nature*. 2015. 521(7553), 436-444. Disponível em: <https://www.nature.com/articles/nature14539>. Acesso em: 14 de maio de 2023.

LEK, S., & GUÉGAN, J. F. **Artificial neural networks as a tool in ecological modelling, an introduction.** Ecological modelling. 1999. 120(2), 65-73. Disponível em: https://horizon.documentation.ird.fr/exl-doc/pleins_textes/2021-06/010020707.pdf. Acesso em 14 de junho de 2023.

OLIVEIRA, Dario Augusto Borges; PEREIRA, Luiz Gustavo Ribeiro; BRESOLIN, Tiago; FERREIRA, Rafael Ehrich Pontes; DOREA, Joao Ricardo Rebouças. **A review of deep learning algorithms for computer vision systems in livestock.** Livestock Science. 2021. 253, p. 104700. Disponível em: <https://www.sciencedirect.com/science/article/am/pii/S1871141321003085>. Acesso em 14 de maio de 2023.

McCARTHY, J. **What is artificial intelligence?** Department of Computer Science Universidade de Stanford, Stanford, CA. 2007. 94305jmc@cs.stanford.edu <http://www-formal.stanford.edu/jmc/> 12 de novembro de 2007. Disponível em: http://35.238.111.86:8080/jspui/bitstream/123456789/274/1/McCarthy_John_What%20is%20artificial%20intelligence.pdf. Acesso em 24 de maio de 2023.

McCARTHY, J., MINSKY, M. L., ROCHESTER, N., & SHANNON, C. E. **A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955.** AI Magazine. (2006). 27(4), 12. Disponível em: <https://doi.org/10.1609/aimag.v27i4.1904>. Acesso em 24 de maio de 2023.

MAHESH, Batta. **Machine Learning Algorithms - A Review.** Revista Internacional de Ciência e Pesquisa (IJSR). 2018. Disponível em: <https://www.researchgate.net/publication/344717762>. Acesso em: 14 de janeiro 2023.

MARSLAND, S. **MACHINE LEARNING: An Algorithmic Perspective.** CRC.2015. Second Edition. International Standard Book Number-13: 978-1-4665-8333-7 (eBook - PDF).

MARENGONI, M., & STRINGHINI, S. **Tutorial: Introduction to Computer Vision using OpenCV.** Revista De Informática Teórica E Aplicada. 2010). 16(1), 125–160. Disponível em: <https://doi.org/10.22456/2175-2745.11477>. Acesso em 16 de março de 2023.

O'SHEA, Keiron; NASH, Ryan. **An Introduction to Convolutional Neural Networks.** UK. 2015. Disponível em: <https://dblp.org/rec/journals/corr/OSheaN15.bib>. Acesso em 16 de maio de 2023.

Documentação. <https://orangedatamining.com/docs/>. Acesso em 13 de janeiro de 2023.

PAOLA, J. D., & SCHOWENGERDT, R. **A detailed comparison of backpropagation neural network and maximum-likelihood classifiers for urban land use classification.** Geoscience and Remote Sensing. IEEE Transactions. 1995. 33(4), 981-996. Disponível em: <https://www.semanticscholar.org/paper/A-detailed-comparison-of-backpropagation-neural-and-Paola>

Schowengerdt/0bb826c719c316ecaf087c50b9e81ea1f36116dd. Acesso em 14 de junho de 2023.

PATHAN, S.; PRABHU, KG, & SIDDALINGASWAMY, PC. **Techniques and algorithms for computer-aided diagnosis of pigmented skin lesions—A review.** Processamento e controle de sinais biomédicos. 2018. 39, 237-262. doi:10.1016/j.bspc.2017.07.010 Acesso em 24 de março de 2023.

RASCHKA, Sebastian. **An overview of general performance metrics of binary classifier systems.** arXiv preprint arXiv:1410.5330, 2014. Disponível em: <https://arxiv.org/pdf/1410.5330>. Acesso em 24 de março de 2023.

RAWAT, W., & WANG, Z. **Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review.** Neural Computation. 2017. 29(9), 2352–2449. doi:10.1162/neco_a_00990 10.1162/neco_a_00990. Disponível em: <https://ieeexplore.ieee.org/abstract/document/8016501>. Acesso em 25 de fevereiro de 2023.

REIS, E.A., REIS I.A. **Análise Descritiva de Dados. Relatório Técnico do Departamento de Estatística da UFMG.** 2002 . Disponível em: www.est.ufmg.br <http://www.est.ufmg.br/portal/arquivos/rts/rte0202.pdf>. Acesso em 24 de março de 2023.

RIOS, Cleisson Da Silva. **Análise sobre a relevância da informação contábil para prever a captação de doações das organizações da sociedade civil de interesse público no período de 2006 a 2014**[manuscrito] / Cleisson da Silva Rios. – 2016. Disponível em: <https://1library.org/document/zpd6re4z-analise-relevancia-informacao-contabil-captacao-organizacoes-sociedade-interesse.html> Acesso em 24 de maio de 2023.

SILVA, Marcos. **Matriz de Confusão e AUC ROC.** Data Hackers – Medium. 2019. Disponível em: <https://medium.com/data-hackers/matriz-de-confus%C3%A3o-e-auc-roc-f7e446dca107>. Acesso em 24 de março de 2023.

SMITS, P. C., DELLEPIANE, S. G., & SCHOWENGERDT, R. A. **Quality assessment of image classification algorithms for land-cover mapping: A review and a proposal for a cost-based approach.** International Journal of Remote Sensing. 1999. 20(8), 1461–1486. doi:10.1080/014311699212560 Acesso em 24 de março de 2023.

TSANGARATOS, P., & ILIA, I. **Comparison of a logistic regression and Naïve Bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size.** CATENA. 2016. 145, 164–179. doi:10.1016/j.catena.2016.06.004 Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0341816216302090>. Acesso em 24 de março de 2023.

TOMASEVIC, N., GVOZDENOVIC, N., & VRANES, S. **An overview and comparison of supervised data mining techniques for student exam performance prediction.** Computers & Education. 2019. 103676.

Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S0360131519302295>

doi:10.1016/j.compedu.2019.103676

Acesso em 20 de abril de 2023.

MCCULLOCH, W.S., PITTS, W. **A logical calculus of the ideas immanent in nervous activity.** The bulletin of mathematical biophysics. 1943. 5(4):115–133.

Disponível em: <https://link.springer.com/article/10.1007/bf02478259>.