

UNIVERSIDADE FEDERAL DE SANTA CATARINA - CENTRO TECNOLÓGICO
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA

**Desenvolvimento de um Curso On-line para o
Ensino de *Machine Learning* voltado a Classificação de Imagens de
Macrofungos no Ensino Médio usando Python**

Stefano Bergamini Poletto

Florianópolis
2022
Universidade Federal de Santa Catarina
Departamento de Informática e Estatística

**Desenvolvimento de um Curso On-line para o
Ensino de *Machine Learning* voltado a Classificação de Imagens de
Macrofungos no Ensino Médio Usando Python**

Trabalho de Conclusão de Curso de Graduação em Ciências da Computação, do Departamento de Informática e Estatística, do Centro Tecnológico da Universidade Federal de Santa Catarina, requisito parcial à obtenção do título de Bacharel em Ciências da Computação.

Autor: Stefano Bergamini Poletto

Orientadora: Prof.^a Dr.^a rer. nat. Christiane Gresse von Wangenheim, PMP

Florianópolis
2022

Stefano Bergamini Poletto

Cogufy - Desenvolvimento de um Curso On-line para o
Ensino de Machine Learning voltado a Classificação de Imagens de Macrofungos no
Ensino Médio Usando Python

Trabalho de Conclusão de Curso de
Graduação em Sistemas de Informações,
do Departamento de Informática e
Estatística, do Centro Tecnológico da
Universidade Federal de Santa Catarina,
requisito parcial à obtenção do título de
Bacharel em Ciências da Computação.

Florianópolis, 16 de Dezembro de 2022

Prof.^a Dr.^a rer. nat. Christiane Gresse von Wangenheim, PMP
Orientadora
Universidade Federal de Santa Catarina

Prof. Dr. Mauro Roisenberg
Avaliador
Universidade Federal de Santa Catarina

Prof. Me. Ramon Mayor Martins
Avaliador
Universidade Federal de Santa Catarina

RESUMO

Machine Learning (ML) é atualmente uma das tecnologias com maior crescimento do mundo. Porém atualmente o conhecimento sobre esta tecnologia é pouco popularizado pelo fato da mesma não ser abordada na Educação Básica. Esse trabalho tem como proposta criar um curso *on-line* interdisciplinar para ensinar conceitos de *ML* no ensino médio utilizando da linguagem de programação Python. O curso visa ensinar a aplicação de conceitos de *ML* de forma prática na biologia voltada na classificação de espécies de Macrofungos. O curso será desenvolvido baseado na análise do estado da arte, adotando o design instrucional resultando na criação de um plano de ensino, o desenvolvimento de todo o material didático em forma de um curso *on-line*, infraestrutura técnica e a definição da avaliação da aprendizagem do aluno. Espera-se que os resultados deste trabalho contribuam para a popularização do conhecimento de *Machine Learning* e a geração de interesse nos jovens nas áreas de Computação e Biologia.

Palavras Chaves: *Machine Learning*, *Deep Learning*, Biologia, Micologia, Computação.

SUMÁRIO

1. INTRODUÇÃO	6
1.1. Contextualização	6
1.2. Objetivos	8
Objetivo Geral	8
Objetivos específicos	8
1.3. Metodologia de pesquisa	9
1.4. Estrutura do documento	10
2. FUNDAMENTAÇÃO TEÓRICA	11
2.1. Conceitos, processo e ferramentas de Machine Learning	11
2.1.1. Redes Neurais	13
2.1.2. Deep Neural Network (DNN)	14
2.1.3. Convolutional Neural Networks (CNN)	15
2.2. Ensino de Machine Learning no Ensino Médio	26
2.3. Ensino de Micologia no Ensino Médio	34
2.3.1 Conceitos de Micologia	34
2.3.2 Ensino de Micologia	42
3. ESTADO DA ARTE	44
3.1. Definição do protocolo de revisão	44
3.2. Execução da busca	46
3.3. Análise dos Resultados	47
3.3.1 Quais unidades instrucionais existem?	47
3.3.2 Quais conceitos de ML são ensinados nessas Unidades Instrucionais?	48
3.3.3 Quais os conteúdos de ML abordados?	51
3.3.4 Quais são as características instrucionais dessas unidades instrucionais?	55
3.3.5 Como foram desenvolvidas as unidades instrucionais e qual é a avaliação de qualidade dessas unidades instrucionais?	59
3.4. Discussão	63
4. Curso Cogufy (Nível Modifique)	66
4.1 Análise de Contexto	66
4.2 Objetivos de Aprendizagem	69
4.3 Design do Curso	72
4.4 Plano de Ensino	73
4.5 Desenvolvimento do material didático	77
5. CONCLUSÃO	81

1.INTRODUÇÃO

1.1. Contextualização

O uso da Inteligência Artificial (IA) e especificamente soluções utilizando *Machine Learning (ML)* no dia-a-dia já é realidade, muitas vezes de maneira despercebida ela auxilia em tarefas cotidianas nos diferentes aparelhos e serviços utilizados pela população. Muitos já sabem sobre a existência da IA e seu grande crescimento nas últimas décadas, com estimativas alcançando um valor de \$62,5B em 2022, um crescimento de 21,3% de 2021 (Gartner, 2021). Porém a ideia sobre IA nessa geração atual é muito focada no que os jovens consomem, como meios de entretenimento ou assistentes de pesquisas, por exemplo, *videogames*, assistentes pessoais (Siri, Google, etc.) e nas notícias, como, vitórias de computadores sobre oponentes humanos (Royal Society, 2017). Mesmo com todas essas informações ao seu redor poucos sabem como funciona essa IA e a existência de *ML* (Royal Society, 2017). Com isso se percebe a necessidade de abordar o tema de *ML* e computação já na educação básica para que as pessoas tenham um melhor conhecimento sobre as tecnologias que estão ao seu redor (Burgsteiner et al., 2016). Desta maneira desmistificar a ideia que *IA/ML* é complexa e inacessível para quem não está no ramo de tecnologia a muito tempo (Wangenheim et al., 2020). E, evitar que ocorra escassez de mão-de-obra qualificada para assumir tais postos de trabalho, já que a evolução tecnológica de *ML/IA* é cada vez mais rápida (Scaico et al., 2012). Apresentar *ML* e computação desde a Educação Básica tem também o potencial de atrair grupos mais diversificados de pessoas para a área de computação (Kelleher et al., 2005).

Visando a formar criadores de soluções de TI (Resnick, 2012), existem propostas de diretrizes curriculares e unidades instrucionais para o ensino de *IA/ML* na Educação Básica (Marques et al., 2020a). No Brasil a Sociedade Brasileira de Computação (SBC) apresenta diretrizes para a Educação Básica brasileira. Estas diretrizes têm um foco maior na aprendizagem de pensamento computacional e conceitos básicos necessárias para entender as questões éticas envolvendo IA (SBC, 2017). O Ministério da Educação (MEC) ressalta que grande parte das futuras profissões utilizarão, direta ou indiretamente, computação e tecnologias digitais, com objetivo de formar jovens e adultos com habilidades para utilizar das

tecnologias digitais futuras se percebe a necessidade para abordar esse tema já na educação básica (Brasil, 2017).

Percebendo a importância de *ML*, alguns países como Austrália e Reino Unido começaram a inserir IA nos currículos de Educação Básica (Marques et al., 2020a). No Brasil e em muitos outros países o estudo de IA e principalmente *ML* na Educação Básica é escassa e basicamente inexistente (Marques et al., 2020a).

Para realizar a inserção desses conteúdos sobre IA/*ML* na educação básica são criadas unidades instrucionais. Essas unidades instrucionais tipicamente visam ensinar conceitos básicos de *ML*, redes neurais, questões éticas e implicações de *ML* na sociedade. Geralmente se adota estratégias ativas visando levar o aluno a aprender e aplicar estes conceitos, por exemplo desenvolvendo modelos de deep learning (DL) e realizando a implantação dentro de jogos ou aplicativos. Frequentemente estas unidades instrucionais são realizadas em forma de cursos extracurriculares ou inserido de forma interdisciplinar. São utilizados diversos ambientes de desenvolvimento de modelos de *ML*, desde ferramentas visuais até ferramentas textuais como *Jupyter notebooks* (Gresse von Wangenheim, 2020). *Jupyter Notebooks* são artefatos que podem conter código, texto, imagens e/ou equações (O'Hara et al., 2015). Os *Jupyter Notebooks* podem ser lidos normalmente porém possuem a possibilidade de executar pedaços de código, como um programa de computador, tudo diretamente do web browser (O'Hara et al., 2015). Esse tipo de ambiente vem mostrando bons resultados para ensinar diversos tópicos para estudantes.

Porém observa-se atualmente uma falta de cursos sobre *ML* no ensino médio brasileiro, utilizando linguagens de programação textual de maneira mais completa, abrangendo o processo de *ML* completo, desde a análise de requisitos até a sua implantação.

Assim, este projeto tem o objetivo de criar um curso on-line utilizando *Jupyter Notebook* e Google Colab usando uma linguagem textual (Python) no estágio Modify para ensinar conceitos e o processo de *ML* aos alunos do ensino médio. Visa-se a criação de um curso interdisciplinar com conteúdo da área de Biologia, especificamente macrofungos/micologia, possibilitando uma integração deste conteúdo de *ML* no contexto de cursos do ensino médio.

1.2. Objetivos

Objetivo Geral

O objetivo geral deste projeto é o desenvolvimento de um curso on-line para ensinar *Machine Learning* de forma interdisciplinar junto com o ensino de tópicos da Micologia para ser aplicado em escolas brasileiras em nível de Ensino Médio. O curso será alinhado aos currículos de referência do ensino de computação da ACM/CSTA (Seehorn e Clayborn, 2017) e SBC (SBC, 2017) e especificamente Inteligência Artificial utilizando como base as diretrizes de currículo da AI4ALL (Touretzky, 2019) e da Long e Magerko (2020).

Este projeto utilizará uma linguagem de programação textual (Python) dentro do *Notebook Jupyter* para realizar a computação das imagens de macrofungos e realizar a classificação dos mesmos, ensinando as etapas para criação dos modelos de *Machine Learning*.

Este projeto visa todo o design instrucional relativo ao curso, incluindo a análise do contexto (caracterizando o público alvo, ambiente, necessidades, etc.), definição dos objetivos de aprendizagem, definição da estratégia instrucional e a preparação de todos os recursos didáticos e infraestrutura técnica a serem utilizados durante a sua execução. Como resultado, será criado o plano de ensino, todo o material didático em forma de um curso interativo on-line, como também a definição da avaliação da aprendizagem do aluno.

Objetivos específicos

Os objetivos específicos são:

- O1. Elaborar a fundamentação teórica em relação ao ensino de *Machine Learning* no ensino médio e os processos e ferramentas de desenvolvimento de seus modelos para classificação utilizando *Machine Learning*
- O2. Levantar o estado da arte em relação a unidades instrucionais semelhantes com foco em estudantes do ensino médio e em relação ao modelos de *Machine Learning* de classificação de macrofungos.

- O3. Definir o *design* instrucional do curso (análise do contexto e definição do plano de ensino);
- O4. Desenvolver o material didático e infraestrutura técnica para o curso on-line;

1.3. Metodologia de pesquisa

A fim de alcançar os resultados esperados com este trabalho, é adotada uma combinação de metodologias de pesquisa de acordo com o respectivo objetivo a ser buscado. Então, de acordo com os objetivos específicos do projeto, são adotadas etapas da seguinte forma:

Etapa 1 – Fundamentação teórica: análise e síntese de conceitos básicos envolvidos no tema. São abordados conceitos de Micologia e ensino de *Machine Learning* no ensino médio e sobre o processo e ferramentas de desenvolvimento de *Machine Learning*, principalmente o *Jupyter Notebook*. Este estudo é feito por meio de uma análise e síntese da literatura.

Atividade 1.1: Sintetizar conceitos de ensino de *ML* no ensino médio;

Atividade 1.2: Sintetizar conceitos, processo e ferramentas de *ML*.

Atividade 1.3 Sintetizar conceitos de micologia.

Etapa 2 - Levantamento do estado da arte: levantamento sobre trabalhos existentes relacionados à área do projeto. É realizado um estudo de mapeamento seguindo um processo proposto por Petersen, Vakkalanka e Kuzniarz (2008; 2015) para identificar e analisar unidades instrucionais/estratégias de ensino, atualmente sendo utilizadas e voltadas ao ensino de computação em escolas.

Atividade 2.1: Realizar a revisão das unidades instrucionais atualizando a revisão da Marques et al. (2020b);

Atividade 2.1.1: Definir o protocolo de revisão;

Atividade 2.1.2: Executar a busca atualizada;

Atividade 2.1.3: Extrair e analisar as informações.

Etapa 3 - Design da unidade instrucional: engloba toda a parte de planejamento e *design* da unidade instrucional a ser realizada. A definição do *design* da unidade segue a metodologia de design instrucional ADDIE (BRANCH, 2010).

Atividade 3.1: Analisar o contexto (público alvo, infraestrutura, etc.)

Atividade 3.2: Definir os objetivos de aprendizagem e sequenciar o conteúdo da unidade instrucional e definir uma estratégia instrucional, criando o plano de ensino.

Atividade 3.3: Definir a avaliação da aprendizagem do aluno;

Atividade 3.4 Criar o modelo de ML de classificação de macrofungos no Jupyter;

Atividade 3.5: Desenvolver o material de atividades on-line;

1.4. Estrutura do documento

No Capítulo 2 é apresentada a fundamentação teórica dos conceitos necessários que sustentam a proposta deste trabalho. No capítulo 3 é apresentado o estado da arte, a situação atual em que se encontram os trabalhos e propostas existentes nas áreas de categorização de macrofungos utilizando de *deep learning* e de cursos de ensino sobre IA/ML para jovens do ensino médio brasileiro. O capítulo 4 mostra a análise do contexto dos estudantes e ambiente escolar e apresenta a proposta do plano de ensino e o desenvolvimento do curso, além das ferramentas utilizadas. No capítulo 5 é a conclusão do trabalho e sugestões de trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados conceitos com relação ao trabalho desenvolvido. O primeiro tópico a ser apresentado é IA e suas diversas técnicas. Dentre elas, redes neurais e analisar suas subdivisões *Machine Learning* e *Deep Learning*. Seguido por uma análise do ensino sobre *ML* no ensino médio. Finalizando com uma análise teórica sobre Micologia, com foco maior em Macrofungos.

2.1. Conceitos, processo e ferramentas de *Machine Learning*

Desde os primeiros computadores se perguntava se essas máquinas poderiam se tornar inteligentes, atualmente nos vemos rodeados de *softwares* inteligentes realizando diversos tipos de aplicações práticas de maneira automatizada e diversos tópicos de estudando sobre IA e suas áreas (Goodfellow et al., 2016).

Para McCarthy (1956), IA se define como “a ciência e engenharia para criar máquinas inteligentes que tenham a habilidade de alcançar resultados como humanos fazem”. Ou seja, Inteligência Artificial é a inteligência humana demonstrada por máquinas (Steels, 1993).

Machine Learning é uma subdivisão de IA que lida com o campo de estudos que dá a habilidade aos computadores de aprenderem sem serem programados explicitamente (Steels, 1993), ou seja, se fornecidos dados suficientes a máquina pode prever resultados de maneira precisa. *ML* é inerentemente um campo multidisciplinar. IA se expandiu em outras subdivisões e diversas técnicas para implementação (Mitchell, 1997), sendo uma delas a *Deep Learning (DL)*. Dentre as diversas técnicas e subdivisões presentes para *ML*, a *DL* vem se destacando atualmente (Fan et al., 2020), sendo considerada a próxima evolução de *ML*. Assim como IA tenta aproximar o computador ao humano, *DL* utiliza algoritmos inspirados no sistema de aprendizado humano, gerando padrões e classificando os diferentes tipos de informação recebidas, como, p.ex., imagens, som, texto entre outros (Tang et al., 2016). A Figura 1 ilustra o relacionamento entre as diferentes áreas de IA, com o objetivo de facilitar o entendimento sobre suas divisões, mostrando que *DL* é um tipo de *ML* que por sua vez é uma parte de IA.

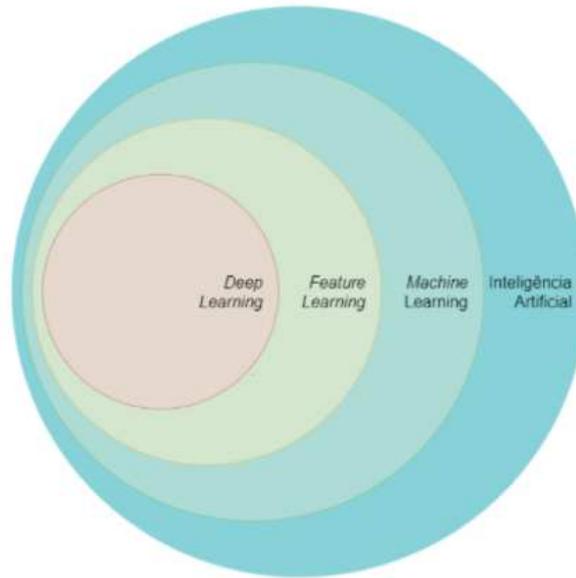


Figura 1 - Relação entre IA, *ML* e *DL* (Fonte: Adaptado a partir de Goodfellow et al., 2016)

Deep Learning vem sendo utilizada para resolver problemas nas mais diversas áreas (LeCun, 2021). Suas aplicações abrangem diversas áreas de conhecimento, como por exemplo, setor automotivo, financeiro, biologia, saúde e até a área de entretenimento. *DL* pode realizar diversos tipos de tarefas, sendo as principais:

- Classificação de imagens
- Detecção de objetos, faces e movimentos
- Reconhecimento de linguagem natural
- Prevenção de possíveis erros

A utilização de *DL* na área de classificação de imagens mostrou uma melhoria significativa na taxa de precisão, se comparado a outros algoritmos (Zhu et al., 2018).

Diferente de outras técnicas de *ML*, *DL* pode descobrir automaticamente recursos a serem usados para a classificação e conseqüentemente necessita de um banco de dados maior e computadores melhores, se comparado a *ML* (Goodfellow, 2016). A técnica de *DL* realiza a análise de suas informações dividindo em diversas camadas de softwares, de forma não linear, tendo seu resultado encontrado após combinar cada camada e identificar a resposta com maior probabilidade de estar correta (Soares, 2017). Isso é possível pois *DL* utiliza de redes neurais (*Neural Networks*) com múltiplas camadas de rede.

2.1.1. Redes Neurais

A primeira proposta de rede neural foi realizada em 1944 por McCulloch e Pitts que ficaram conhecidos também por serem fundadores do primeiro departamento de ciência cognitiva (Hardesty, 2017). A pesquisa sobre redes neurais sofreu diversos altos e baixos, Poggio até comparou o estudo sobre rede neural a um tipo de vírus que acontece em períodos alternados. Atualmente o “renascimento” das pesquisas sobre redes neurais se dá pelo fato da técnica de *Deep Learning* (Hardesty, 2017).

Redes neurais possuem seu modelo baseado no cérebro humano, consistindo de milhares ou até mesmo milhões de nós que são densamente interconectados (Hardesty, 2017). Nós são “lugares” onde se realiza computação. Os nós são vagamente baseados nos neurônios que são ativados quando recebem estímulo suficiente (Russel, 2009). Cada nó combina o dado recebido com um peso numérico, que determina a força e o sinal de conexão (Russel, 2009). A Figura 2 mostra um exemplo da estrutura de um nó.

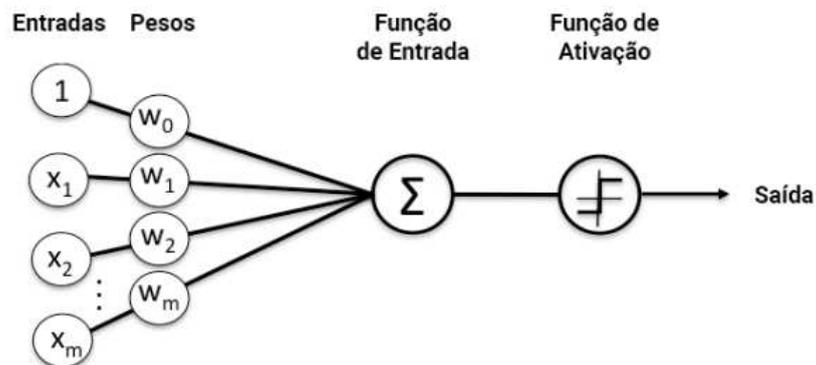


Figura 2 - Diagrama exemplo de um nó (Fonte: Adaptado de Pathmind, 2019)

Atualmente a maioria das redes neurais são organizadas em camadas que possuem diversos nós que são “alimentados” pelos valores processados pelos nós anteriores movendo em uma só direção para alcançar a saída (Hardesty, 2017). Cada saída de camada é simultaneamente a entrada da camada subsequente, começando por uma camada inicial que recebe os dados iniciais (Russel, 2009). A Figura 3 mostra um exemplo de rede neural que apresenta diversos nós ocultos que realizam o processamento até alcançar a camada de saída.

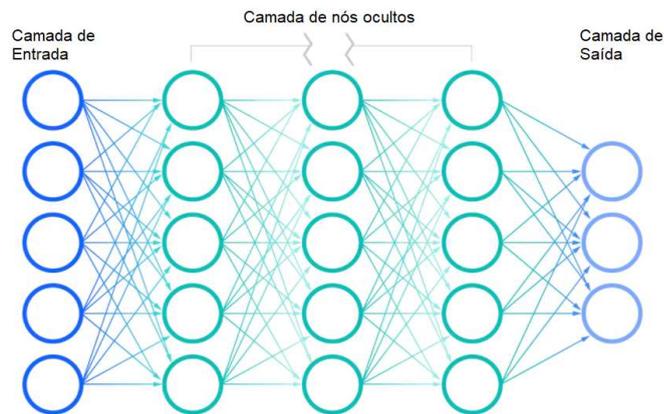


Figura 3 - Diagrama exemplo de uma rede neural
 (Fonte: adaptado a partir de IBM Cloud Education, 2020)

Redes neurais nos ajudam a agrupar e classificar dados. As redes neurais agrupam os dados não rotulados de acordo com as semelhanças entre eles e as entradas de exemplo, e realiza as classificações dos dados quando se tem um *dataset* rotulado para que seja realizado o treinamento (Pathmind, 2020a).

Redes neurais também possuem diferentes tipos de implementação e de construção.

2.1.2. *Deep Neural Network (DNN)*

A DNN se diferencia das demais redes neurais pela sua profundidade, ou seja, a quantidade de camadas de nós ocultos até alcançar o nó de saída, mais do que 3 camadas, incluindo a camada de entrada e saída, qualifica-se como uma DNN (Albawi, 2017). A Figura 3 apresenta uma representação de uma DNN, com mais de 3 camadas de nós.

Na DNN, cada camada de nó treina em conjuntos distintos de recursos baseado na camada anterior. Quanto mais se avança nas camadas, os nós conseguem reconhecer dados de entrada cada vez mais complexos, já que se realiza a junção e combinação dos recursos da camada anterior (Szegedy, 2013). A Figura 4 exemplifica esse processo da DNN.

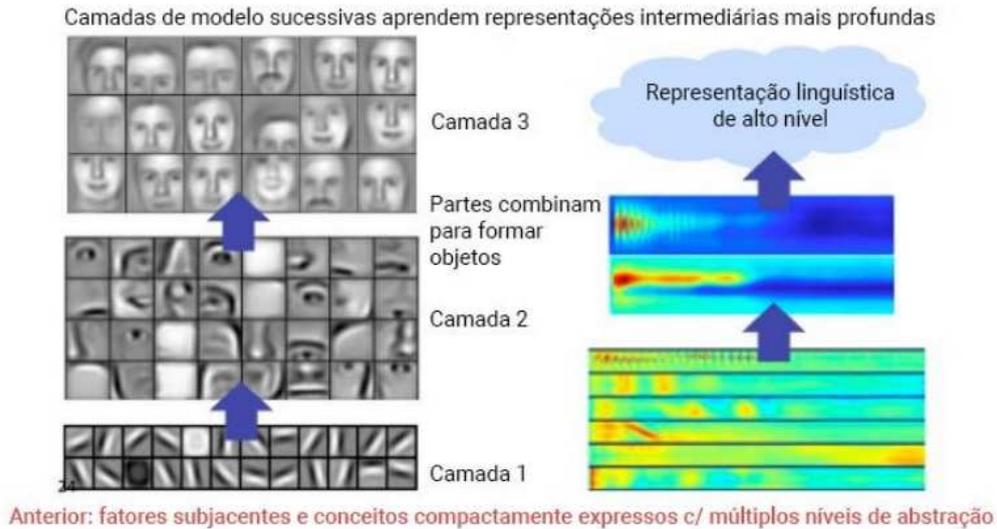


Figura 4 - Exemplo do funcionamento de uma DNN (Fonte: Adaptado de Pathmind, 2019)

Essa característica da DNN é chamada de hierarquia de recursos, com ela é possível processar dados grandes com bilhões de parâmetros (Albawi, 2017). Com isso também é possível descobrir estruturas latentes em dados não-estruturados e não categorizados, que representam a grande maioria dos dados presentes no mundo real (Pathmind, 2020a).

2.1.3. Convolutional Neural Networks (CNN)

Uma das principais DNNs é a *Convolutional Neural Network* (CNN) que está sendo também amplamente utilizada para tarefas de classificação de imagens (Albawi, 2017). O principal ponto positivo da CNN é o fato de que as características não precisam ser espacialmente dependentes. Ou seja, por exemplo, em aplicações de detecção de faces, não é preciso prestar atenção onde as faces estão localizadas na imagem (Albawi, 2017). Outro aspecto importante da CNN é a obtenção de características abstratas quando a entrada se propaga para as camadas mais internas. Por exemplo, em classificação de imagens, uma simples beirada pode ser detectada na primeira camada, já na segunda se detecta formatos mais complexos, e na terceira em diante já apresenta características de uma face (Albawi, 2017) (Figura 5).

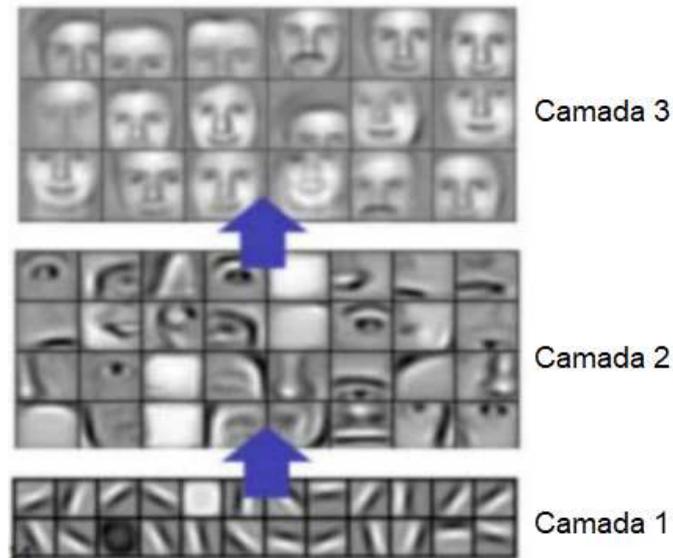


Figura 5 - Aprendizado entre camadas de uma CNN
(Fonte: Adaptado a partir de Albawi , 2017)

CNN, como o próprio nome diz “redes neurais convolucionais”, utiliza da operação matemática de convolução. Em definição CNNs são redes neurais que usam convolução no lugar de multiplicação de matrizes em ao menos uma de suas camadas (Goodfellow et al., 2016). Convolução é uma integral que expressa a sobreposição de uma função g ao ser deslocada sobre uma outra função f , resultando na “combinação” das duas funções (Weisstein, 2019). De maneira mais simplificada pode se dizer que uma convolução é o produto de duas funções: f e g (Weisstein, 2019).

A utilização de convolução vem para diminuir o peso em cima das camadas ocultas de processamento, pois, por exemplo, caso seja enviado uma informação de $32 \times 32 \times 3$ de pesos diretamente para um nó e seja adicionado um novo nó, são no total $32 \times 32 \times 3 \times 2$ parâmetros sendo passados para processar, isso com somente 2 nós. Para evitar esse problema e achar um método mais eficiente, se percebeu que ao invés de enviar todos os dados, pode ser enviado uma região da imagem (Albawi, 2017). Com isso se tem uma drástica redução de conexões entre camadas (Albawi, 2017). Nessa técnica também provê a oportunidade de detectar e conhecer características em qualquer posição da imagem (Albawi, 2017). A Figura 6 ilustra um exemplo.

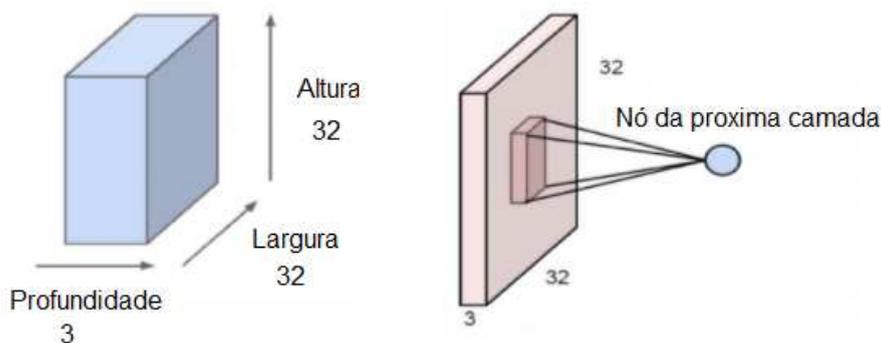


Figura 6 - Exemplo de entrada e a utilização da convolução ao invés de conexão direta
(Fonte: Adaptado a partir de Albawi , 2017)

Como é possível observar pelo exemplo, essa rede neural identifica as imagens como volumes, ou seja, tridimensionais. Isto ocorre pois as imagens digitais utilizam a codificação *Red-Green-Blue (RGB)*, com seus canais de cores sendo armazenados separadamente e misturados para produzir cores perceptíveis pelo ser humano. Esses canais são representados por uma matriz de valores, que representam a intensidade do canal em um determinado pixel (Pathmind, 2020b). A Figura 7 mostra uma aproximação dessa matriz.



Figura 7 - Comparação entre uma imagem e sua representação em matriz
(Fonte: Adaptado de Karkare, 2019)

Para realizar a detecção das características e detalhes, em geral são arestas, presentes nas imagens, são utilizados filtros que percorrem a imagem um pixel por vez. O filtro é uma pequena matriz de valores que multiplica seus valores com os valores da imagem, somando-os em um único valor a cada passagem sobre o pixel (Karkare, 2019). A Figura 8 demonstra um exemplo de convolução com o filtro.

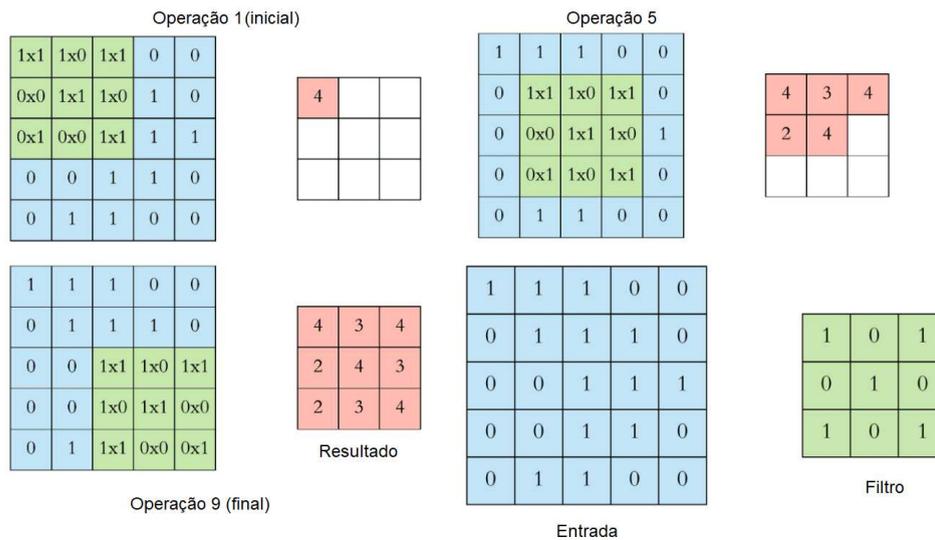


Figura 8 - Exemplo de convolução (Fonte: Adaptado a partir de Karkare , 2019)

Assim como filtros clássicos para processamento de imagens, os filtros da CNN também agem de maneira parecida. A diferença maior é que os filtros da CNN são “treinados” para serem mais adequados para a tarefa dada (Albawi, 2017). A Figura 9 mostra exemplos de filtros com “treinamento” diferentes, sendo um para melhor detecção de linhas horizontais e outro para linhas verticais.

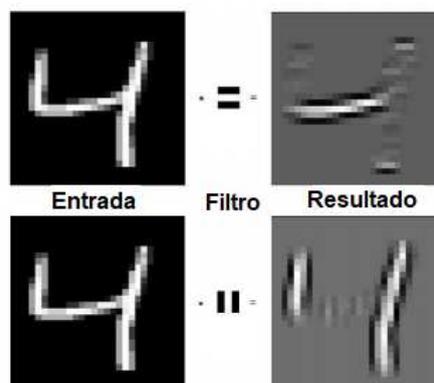


Figura 9 - Exemplo de diferentes filtros (Fonte: Adaptado de Karkare, 2019)

Para que esse método seja ainda mais eficiente são adicionadas camadas e cada camada possui um filtro diferente para alcançar o resultado mais detalhado (Albawi, 2017). A Figura 10 demonstra essa operação de maneira ilustrativa. A Figura 11 mostra os resultados possíveis ao utilizar diferentes matrizes na convolução.

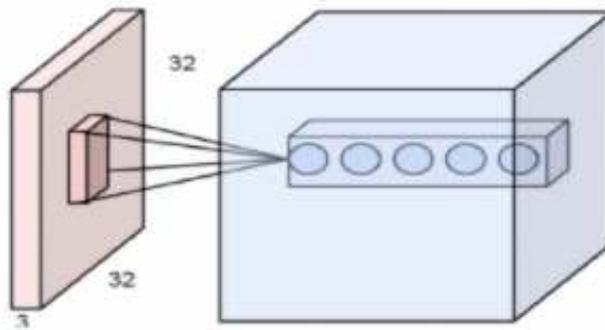


Figura 10 - Múltiplas camadas com filtros diferentes processando a mesma região de entrada
(Fonte: Albawi , 2017)

Operação	Filtro	Resultado
Identity	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Edge detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	

Figura 11 - Efeito de diferente filtros e filtros sequenciais, na mesma imagem de entrada
(Fonte: Adaptado a partir de Albawi, 2017)

Para reduzir ainda mais a complexidade para outras camadas, se utiliza de uma técnica chamada de *pooling*, no domínio de imagens pode se considerar como se reduzisse a resolução da imagem (Albawi, 2017). A técnica de *pooling* mais comumente utilizada é a *max-pooling* (Albawi, 2017). Essa técnica particiona a imagem em sub regiões retangulares, e nela só se retorna o valor máximo dentro dessa sub-região. Em geral o tamanho mais utilizado em max-pooling são matrizes 2x2 (Albawi, 2017). A Figura 12 demonstra um max-pooling 2x2.

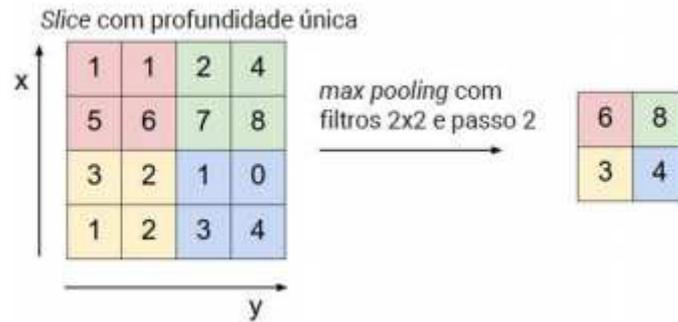


Figura 12 - Max-pooling (Fonte: Adptado a partir de Albawi, 2017)

Por fim é necessário realizar todas as conexões das camadas anteriores para gerar o resultado final, essa etapa tem sua camada/camadas específicas, que recebem o nome de camadas totalmente conectadas (Albawi, 2017). A Figura 13 mostra a estrutura total padrão de uma CNN, mostrando suas camadas de convoluções, *pooling* e as camadas totalmente conectadas.

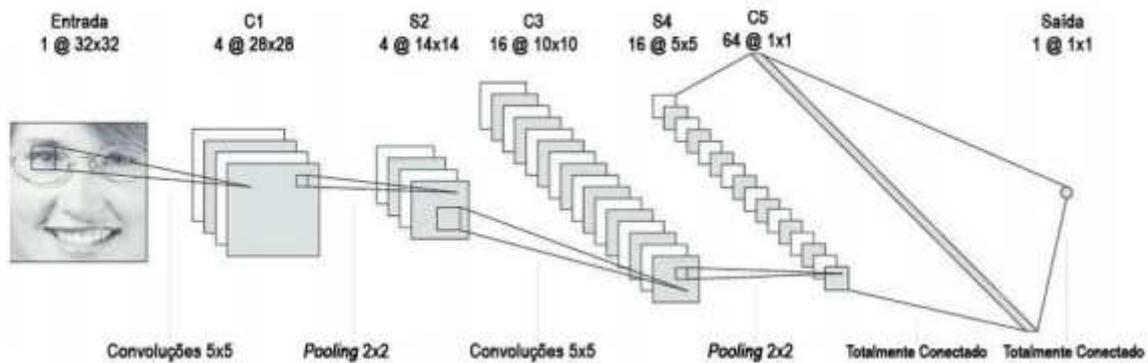


Figura 13 - Arquitetura padrão de uma CNN (Fonte: Adaptado de Karkare, 2019)

Com todas as combinações e técnicas presentes na CNN, atualmente é considerada uma forte ferramenta para detecção de faces e imagens, reconhecimento de vídeo e de voz, tudo por um custo computacional reduzido (Albawi, 2017)

Assim como os outros tópicos comentados, a CNN também apresenta diversas arquiteturas (ModelZoo, 2019), variando suas funções de ativação dos nós, tipos de camadas etc. A Figura 14 apresenta um esquema dos diversos modelos existentes para visão computacional que utiliza a CNN como base.

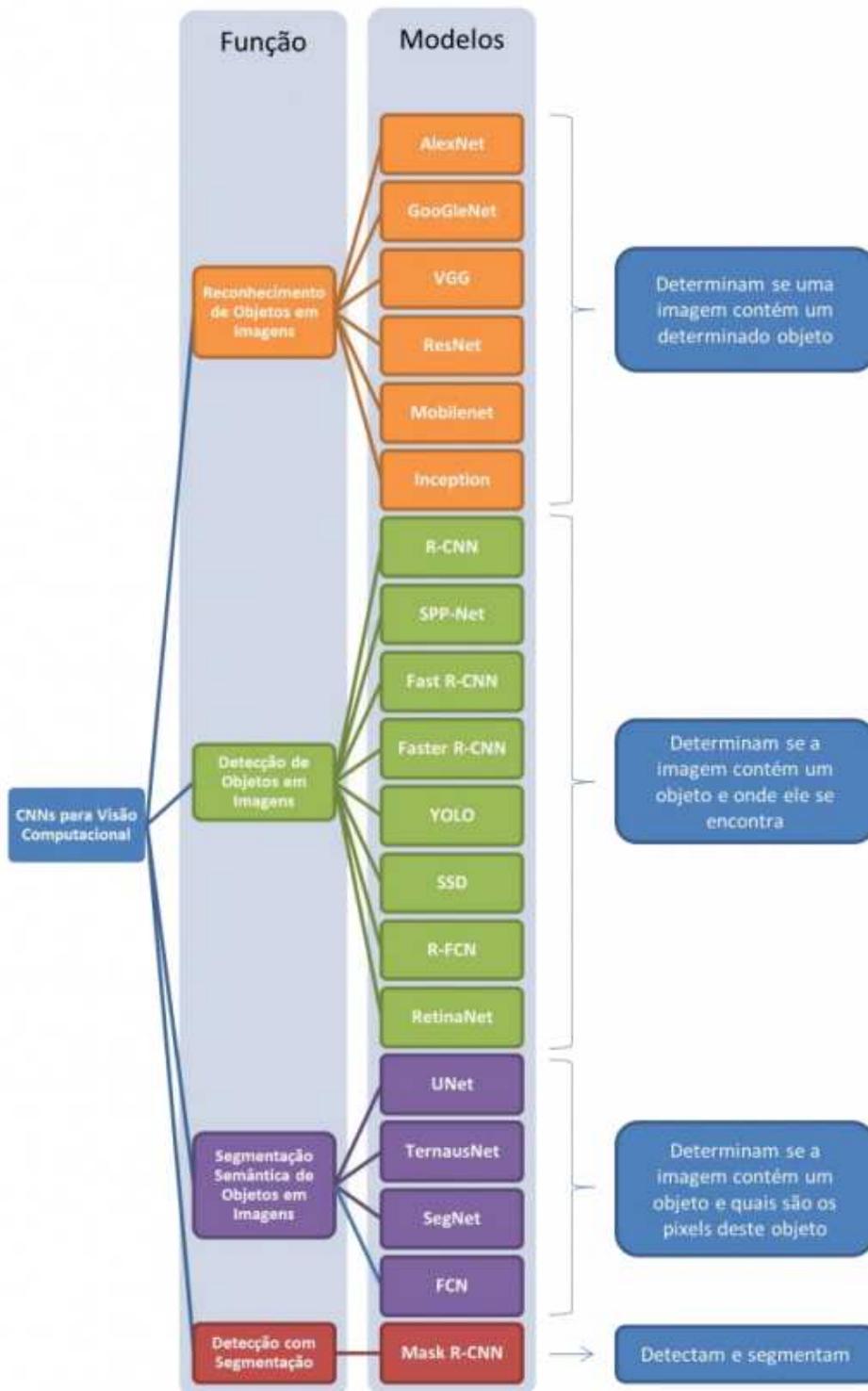


Figura 14 - Diferentes tipos de arquiteturas baseadas em CNNs para Visão Computacional
(Fonte: von Wangenheim , 2018)

O uso de CNNs vem crescendo cada vez mais na comunidade de visão computacional, desde que Krizhevsky propôs AlexNet em 2012 (Wu, 2019). A ResNet (Residual Network) é a mais profunda das redes existentes, apresentando

1202 camadas treináveis e vem sendo utilizada cada vez mais para classificação de imagens (Wu, 2019). A ResNet criou uma solução para que seja possível ter as vantagens de aumento de camadas sem perder precisão. Essa solução apresenta o aprendizado residual e seu bloco de construção (He et al., 2015). A Figura 15 mostra o esquemático deste método.

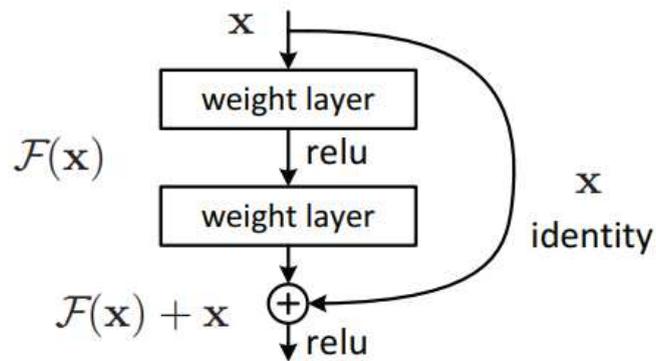


Figura 15 - Aprendizado residual: um bloco de construção (Fonte: He, 2015)

Com isso a ResNet se mostrou facilmente treinada e com os excelentes resultados, alcançando erros de treinamento e de teste menores que outras redes. Por consequência a ResNet vem sendo uma das principais redes para detecção de imagens (Li et al., 2017).

2.1.4. Frameworks e Bibliotecas

Frameworks e bibliotecas são códigos disponibilizados para que possam ser utilizados por outras pessoas resolvendo problemas comuns. Com isso se elimina a necessidade de repetir códigos que já existem. Bibliotecas são invocadas pelos programadores em seu código, já o *framework* possui um fluxo de controle, deixando para o programador a responsabilidade de preenchimento das lacunas faltantes.

Cada *framework* e biblioteca tem seus pontos positivos e negativos, fazendo que a escolha para uso dependa da aplicação em desenvolvimento. Em geral, *frameworks* são mais complexos de utilizar inicialmente, pois o usuário precisa entender seu funcionamento e como fazer uso de suas funções, se comparado com bibliotecas. Para solucionar isso, foi desenvolvido bibliotecas que se baseiam nos

frameworks, com o intuito de facilitar o uso para o programador, tornando o *deep learning* ainda mais acessível e disponível para qualquer pessoa (Fastai, 2019).

Atualmente uma das bibliotecas predominantes é a *fast.ai*. Ela utiliza de um framework *PyTorch* e uma *Application Programming Interface* (API) para aplicações de DL (Howard, 2018).

Em muitos casos, além do uso de frameworks e bibliotecas, se usa em conjunto *notebooks* computacionais. Esses notebooks têm a capacidade de executar as aplicações desenvolvidas pelos programadores diretamente de seu browser, com seus resultados sendo também apresentados diretamente em browser (O'Hara et al., 2015). Atualmente um dos notebooks predominantes é o Jupyter e Google Colab vem sendo um dos ambientes de execução predominantes, ambos tendo Python como a linguagem predominante para esse tipo de aplicação (O'Hara et al., 2015).

2.1.5. Treinamento de Redes Neurais

Todas as arquiteturas necessitam de treinamento para que os resultados sejam o mais próximos da realidade possível. O treinamento de um modelo é realizado a partir de dados que generalizam uma decisão contra uma medida de desempenho (Mitchell, 1997).

Para assegurar que os resultados apresentados pela rede neural sejam precisos é necessário garantir que a base de dados esteja correta. Para garantir a corretude é necessário corrigir os possíveis erros encontrados, por exemplo, dados ausentes, dados com ruído e dados inconsistentes (Pant, 2019).. Algumas técnicas são utilizadas para realizar a limpeza dos dados, conversão de dados, ignorar dados ausentes, completar dados ausentes, detecção de pontos fora da curva e até mesmo a própria *Machine Learning* (Pant, 2019). Com o conjunto de dados limpos disponível se dá início ao treinamento da rede neural.

O treinamento de redes neurais podem ser categorizadas de acordo com seu método de aprendizado (Goodfellow et al., 2016). Os dois principais tipos de aprendizados são: Aprendizado Supervisionado e Aprendizado Não-Supervisionado.

No aprendizado supervisionado, o algoritmo constrói um modelo matemático com o banco de dados, que possui os dados de entrada e os dados desejados de

saída. Algoritmos de classificação e de regressão utilizam do aprendizado supervisionado.

Já o aprendizado não-supervisionado, o algoritmo constrói um modelo matemático de um banco de dados que apresenta somente os dados de entrada. Esse aprendizado é utilizado para encontrar estruturas/padrões presentes no dado de entrada, agrupando os dados em categorias.

A Figura 16 ilustra os modelos e aprendizados comentados.

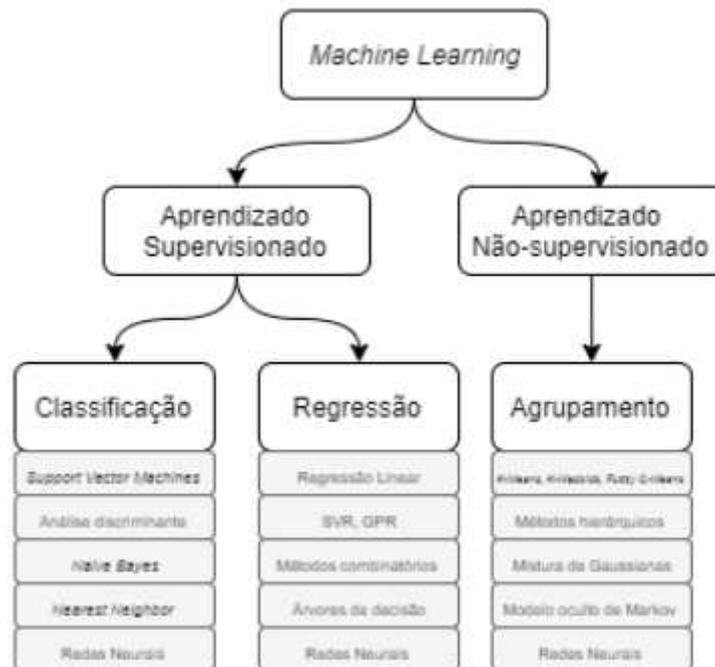


Figura 16 - Modelos de treinamento (Fonte: Adaptado a partir de Pant, 2019)

O treinamento da rede neural para ML é um processo iterativo que apresenta uma sequência de passos, em geral os passos são (Amazon, 2019):

1. Análise de Requisitos. Nessa fase é identificada e especificada as funcionalidades e modelos adequados para solucionar o problema em questão.
2. Gerenciamento de dados. Nessa fase é realizada a coleta de dados e as técnicas apresentadas anteriormente para limpar e revisar a qualidade dos dados do banco de dados que será alimentado a rede neural.

Essa revisão e limpeza pode ser feita com transformações nas imagens, como por exemplo, realizando a mudança de tamanho da figura de maneira que a mesma se encontre no tamanho aceito pela rede neural.

Outra técnica utilizada nessa etapa é a de *data augmentation*, que se trata de fazer pequenas alterações na imagem como por exemplo, translações, rotações e/ou espelhar a imagem. Para a rede neural em treinamento essas alterações fazem com que a rede considere essas imagens como imagens diferentes.

Por fim, essa etapa pode realizar a rotulação dos dados para a aprendizagem supervisionada. A rotulação é realizada de maneira manual atribuindo *tags* importantes nos dados que serão utilizados pela rede neural.

3. Engenharia de *Features* (Características/Recursos). Nessa etapa são selecionados os dados que são relevantes para o aprendizado, ou seja, indicar para a máquina quais características focar na hora de identificar e classificar o dado.
4. Treinamento de Modelo. Nessa etapa pode chegar a durar semanas para realizar o treinamento com *conjunto de dados* grandes, para encurtar esse tempo se utiliza de uma técnica chamada de *transfer learning*. Essa técnica aproveita os recursos de modelos pré-treinados, para que não seja necessário criar um novo modelo do zero. A escolha entre utilizar o *transfer learning* ou criar o modelo depende do programador, em geral, se escolhe utilizar de modelos pré-treinados, testados e com eficiência comprovada em diversos problemas, por exemplo ModelZoo (ModelZoo, 2019). O modelo então é treinado e ajustado recebendo as *features* designadas na fase anterior.

Nessa etapa também existe uma técnica optativa chamada de *fine tuning*, ela se baseia em utilizar pesos (*weights*) de algoritmos anteriores de deep learning similares ao modelo atual.

Uma variável importante no treinamento de redes neurais é introduzido nessa etapa, o *learning rate*. Essa variável indica o quanto os pesos mudam seu valor durante a atualização do modelo, variando entre 0 e 1,0.

5. Avaliação do Modelo. Nessa fase é realizada a avaliação da saída do modelo após seu treinamento. Essa avaliação é realizada analisando seu desempenho, em que suas medidas dependem do tipo de tarefa, incluindo tipicamente acurácia, taxa de erro e matriz de confusão.
6. Implantação do modelo. Ao final o modelo é aplicado em apps, sistemas web, etc., variando de acordo da finalidade que o modelo foi criado.

2.2. Ensino de Machine Learning no Ensino Médio

Para preparar cidadãos, incluindo os jovens, para os diversos dispositivos e serviços que utilizam de *Machine Learning (ML)* no nosso dia-a-dia, é importante que eles se popularizem com um entendimento básico de *ML* (Touretzky et al., 2019).

A ideia de ensinar *ML* para jovens é comumente categorizado como complexo para essa faixa etária, porém estudos mostram que os conceitos sobre IA/*ML* podem ser compreendidos por pessoas mais jovens (Hitron et al., 2019). A exposição do jovem para esse tipo de conhecimento pode ajudar a aumentar o potencial das habilidades importantes no século XXI. Assim, equipando os jovens para lidar com mudanças constantes, econômicas e éticas que estão presentes ao utilizar de *ML* (Kahn et al., 2021).

Apesar de se ter tentativas de ensinar IA para os jovens em educação básica nos anos 70 (Kahn, 1977) e *ML* nos anos 90 (Bemley, 1999), somente recentemente houve um retorno de foco para o ensino de IA/*ML* para os jovens na educação básica (Marques et al., 2020b). Mesmo com um foco maior atualmente, ainda se tem tradicionalmente o ensino de IA e *ML* somente no ensino superior (Torrey, 2012).

Para o ensino de *ML* diretrizes de currículo da AI4K12 (2020), apontam as “grandes ideias” do que cada aluno deve aprender sobre o assunto de IA. Essas “grandes ideias” apresentam ao aluno a percepção, representação e raciocínio, aprendizado, interações naturais e impactos sociais que IA pode causar (Touretzky et al., 2019) (Figura 17).

Cinco Grandes Ideias em Inteligência Artificial

5. Impacto Social

A IA tem um impacto na sociedade que assume formas positivas e negativas. As tecnologias de IA estão a mudar a forma como trabalhamos, viajamos, comunicamos e cuidamos uns dos outros. Mas devemos estar atentos aos danos que podem potencialmente ocorrer. Por exemplo, enviesamentos nos dados usados para treinar um sistema de AI podem levar a situações em que algumas pessoas serão prejudicadas em detrimento de outras. Por essa razão, é importante discutir os impactos que IA está a causar na nossa sociedade e desenvolver critérios para uma conceção e aplicação ética de sistemas baseados em IA.

4. Interação Natural

Os agentes inteligentes requerem vários tipos de conhecimento para interagir naturalmente com os seres humanos. Os agentes têm de ser capazes de conversar em idiomas humanos, reconhecer expressões e emoções faciais, e aproveitar o conhecimento da cultura e das convenções sociais para inferir intenções do comportamento observado. Tudo isto conduz a problemas difíceis. Os sistemas de IA atuais são capazes de utilizar a linguagem até determinados limites, mas não possuem sequer capacidades de conversação e de raciocínio geral semelhantes às de uma criança.

A AI for X-12 Initiative é um projeto conjunto da Associação para o Avanço da Inteligência Artificial (AAAI) e da Associação de Professores de Ciência da Computação (CSTA), financiado pelo prémio da Fundação Nacional da Ciência dos Estados Unidos DRL-1846073

1. Percepção

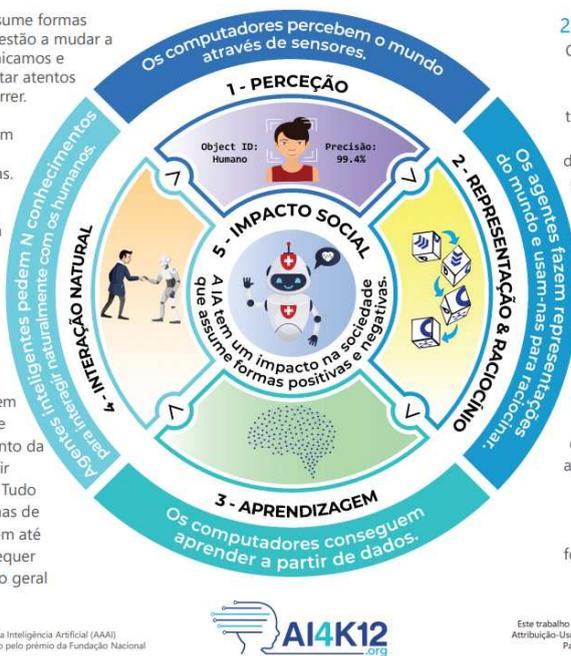
Os computadores percebem o mundo através de sensores. A percepção é o processo de extrair significado a partir de sinais sensoriais. Fazer os computadores “ver” e “ouvir” suficientemente bem para o seu uso prático é uma das mais importantes conquistas da IA.

2. Representação & Raciocínio

Os agentes fazem representações do mundo e usam-nas para raciocinar. A representação é um dos principais problemas da inteligência, tanto natural como artificial. Os computadores constroem representações usando estruturas de dados e estas representações estão na base de algoritmos de raciocínio que derivam nova informação a partir do que já é conhecido. Embora os agentes de IA sejam capazes de raciocinar sobre problemas complexos, não pensam da mesma forma que os seres humanos.

3. Aprendizagem

Os computadores conseguem aprender a partir de dados. A aprendizagem das máquinas é uma espécie de inferência estatística que identifica padrões nos dados. Nos últimos anos, houve várias áreas da IA que evoluíram consideravelmente graças aos algoritmos de aprendizagem que criam novas representações. Para que esta abordagem funcione, são necessárias enormes quantidades de dados. Estes dados de formação são habitualmente fornecidos pelos humanos, mas por vezes podem ser adquiridos pela própria máquina.



Este trabalho está licenciado com uma licença Creative Commons Atribuição-Use Não Comercial-Compartilhável 4.0 Internacional. Para consultar uma cópia desta licença, visite a página: <http://creativecommons.org/licenses/by-nc-sa/4.0/>.



Figura 17 - Grandes Ideias (Fonte: AI4K12, 2020)

Considerando as “grandes ideias”, a ideia do “aprendizado” se refere à aprendizagem de ML. Neste contexto o ensino de ML para ensino médio tem como objetivo principal ensinar aos jovens como ML funciona e suas limitações, preocupações éticas e impactos sociais. Com isso os principais conceitos a serem mostrados para os jovens são (Touretzky, 2019):

- Ensinar o que é o “learning” (aprendizado) em Machine Learning e as abordagens para ML (como por exemplo, algoritmos de regressão, algoritmos baseados em instanciação, máquinas de vetores de suporte, algoritmos de árvores de decisão, algoritmos Bayesianos, algoritmos de agrupamento, algoritmos de rede neural artificial), assim como os tipos de algoritmos de aprendizado, ou seja, supervisionado e não supervisionado.
- Fundamentos sobre redes neurais e seus tipos, inclusive Deep Learning, que com ela é possível realizar computação de redes neurais de multicamadas, incluindo CNNs, e por fim tipos de arquiteturas de redes neurais e como influenciar no seu aprendizado.
- Limitações, preocupações e o impacto de *Machine Learning*.

Para alcançar as competências de ensino de ML em um nível de aplicação, é necessário que os alunos aprendam a desenvolver aplicações de ML para que eles se tornem criadores de soluções inteligentes (Kahn et al., 2018; Touretzky et al., 2019). Construir um aplicativo de ML personalizado de maneira centrada no ser humano é um processo iterativo que exige que os alunos executem uma sequência de etapas conforme apresentado no tópico 2.1.5.

No Brasil a Base Nacional Comum Curricular (BNCC) (Brasil, 2017), propõe tópicos de aprendizagem essenciais para que o jovem tenha um desenvolvimento satisfatório com relação às 10 competências gerais na Educação Básica (Ministério da Educação, 2018). Dentre essas competências, algumas merecem destaque por ter relação às tecnologias e podendo viabilizar a inclusão de ML na Educação brasileira:

- *Valorizar e utilizar os conhecimentos historicamente construídos sobre o mundo físico, social, cultural e **digital** para entender e explicar a realidade, continuar aprendendo e colaborar para a construção de uma sociedade justa, democrática e inclusiva.*
- *Exercitar a curiosidade intelectual e recorrer à abordagem própria das ciências, incluindo a investigação, a reflexão, a análise crítica, a imaginação e a criatividade, para investigar causas, elaborar e testar hipóteses, formular e resolver problemas e criar soluções (**inclusive tecnológicas**) com base nos conhecimentos das diferentes áreas.*
- *Utilizar diferentes linguagens – verbal (oral ou visual-motora, como Libras, e escrita), corporal, visual, sonora e **digital** –, bem como conhecimentos das linguagens artística, matemática e científica, para se expressar e partilhar informações, experiências, ideias e sentimentos em diferentes contextos e produzir sentidos que levem ao entendimento mútuo.*
- ***Compreender, utilizar e criar tecnologias digitais de informação e comunicação de forma crítica, significativa, reflexiva e ética nas diversas práticas sociais (incluindo as escolares) para se comunicar, acessar e disseminar informações, produzir conhecimentos, resolver problemas e exercer protagonismo e autoria na vida pessoal e coletiva.***” (Ministério da Educação, 2018).

A última competência destacada, comenta sobre utilizar tecnologias digitais, em detalhe de compreender, utilizar e criar essas tecnologias com senso crítico de forma a gerar soluções para os problemas na vida social e pessoal. Mesmo não sendo declarado explicitamente, todos os aspectos comentados podem ser inseridos com relação ao ensino de ML.

A BNCC possui um posicionamento sobre ensino de pensamento computacional no ensino fundamental e médio, porém apresenta poucas abordagens práticas para o ensino tecnológico. A Sociedade Brasileira de Computação (SBC) criou diretrizes que apresentam o pensamento computacional na Educação Básica, realizando agrupamento de competências de computação em três eixos, sendo eles, Pensamento Computacional, Cultura Digital e Mundo Digital (Figura 18).

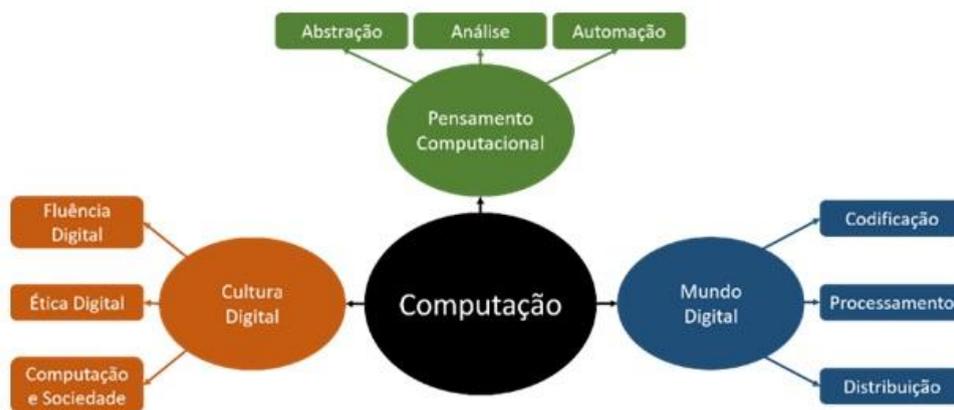


Figura 18 - Três eixos da computação de acordo com SBC (Fonte: SBC, 2019)

De acordo com o documento disponibilizado pela SBC sobre suas diretrizes de computação, o tópico sobre IA é pouco comentado e ML não chega a ser mencionado. A IA de acordo com SBC se enquadra no eixo de Pensamento Computacional, nesse eixo são trabalhadas a técnica de transformação de problemas e o paradigma de metaprogramação (algoritmos que recebem outros algoritmos como entrada), que são conceitos necessários para a compreensão dos limites da computação, ou seja, dos limites da formalização/racionalização. Com esse entendimento então se junta com fundamentos sobre IA para discussões sobre o que é Homem e o que é Máquina.

O plano da SBC não apresenta um ensino teórico e principalmente prático sobre o tema de IA, e o fato de não ser inserido o tópico de ML também não é algo esperado considerando o dia-a-dia cada vez mais competitivo e atualizado com relação às tecnologias de ML utilizadas.

Considerando os planos propostos pela BNCC e SBC se tem uma baixa consideração com relação ao ensino de IA e principalmente ML, se tem especificado na Tabela 1 os objetivos de aprendizado com relação à terceira “grande ideia” do IA4K12, para a Educação Básica.

Tabela 1: Objetivos de aprendizado da terceira “grande ideia”. Traduzido de (AI4K12, 2020)

Conceito		3-5 ano escolar (EUA)	6-8 ano escolar (EUA)	9-12 ano escolar (EUA)
		Anos iniciais do Ensino Fundamental	Anos finais do Ensino Fundamental	Ensino Médio
Natureza de Aprendizagem	Humanos vs. máquinas	Diferenciar entre como pessoas e computadores aprendem	Contrastar as características únicas do aprendizado humano e das maneiras que máquinas operam	Definir aprendizagem supervisionada, não supervisionada e por reforço e dar exemplos de aprendizagem humana similares a cada algoritmo
	Achando padrões em dados	Modelar como aprendizagem supervisionada identifica padrões em dados rotulados	Modelar como aprendizagem supervisionada identifica padrões em dados rotulados	Modelar como ML constrói um raciocinador para classificação ou predição ajustando os parâmetros do raciocinador (suas interpretações internas)

	Treinando modelo	Treinar um modelo de classificação usando machine learning, e então examinar a precisão do mesmo com entradas novas	Treinar e avaliar um modelo de classificação ou predição usando machine learning em um conjunto de dados tabular	Usar um algoritmo de aprendizagem supervisionada ou não supervisionada para treinar um modelo com dados reais e avaliar os resultados
	Construir vs. usar um raciocinador	Demonstrar como dados de treino são rotulados quando se usa uma ferramenta de ML	Explicar a diferença entre treinar e usar um modelo de raciocínio	Ilustrar o que acontece em cada etapa necessária ao usar ML para construir um classificador ou preditor
	Ajustar representações internas	Analisar um jogo onde se constrói a árvore de decisão, descrevendo a organização da árvore e do algoritmo usado para adicionar nós	Comparar como um algoritmo de aprendizado de árvore de decisão funciona vs. como um algoritmo de aprendizado de redes neurais funciona	Descrever como vários tipos de algoritmos de ML aprendem ao ajustar suas interpretações internas
	Aprendendo por experiência	Explicar como aprendizado de reforço permite que um computador aprenda por experiência (tentativa)	Explicar a diferença entre aprendizado supervisionado e não supervisionado	Selecionar o tipo de algoritmo de ML apropriado (aprendizado supervisionado, não supervisionado ou por reforço) para resolver um problema de raciocínio
Redes Neurais	Estrutura de uma rede neural	Ilustrar como uma rede neural de 1 a 3 neurônios é uma função que computa uma saída	Ilustrar a estrutura de uma rede neural e descrever como suas partes formam um conjunto de funções que computam uma saída	Descrever as arquiteturas e usos das redes neurais feedforward, redes convolucionais 2D, recorrentes e adversariais generativas
	Ajuste de peso	Demonstrar como pesos são designados em uma rede neural para produzir o comportamento desejado de entrada e saída	Demonstrar como uma regra de aprendizado pode ser usada para ajustar os pesos em uma rede neural de um nível	Treinar uma rede neural multicamadas usando o algoritmo de aprendizagem de retropropagação e descrever como os pesos dos neurônios e as saídas das unidades ocultas

				mudam como resultado da aprendizagem
Conjunto de dados	Conjuntos de características	Criar um conjunto de dados rotulado com features explícitas de diferentes tipos e usa uma ferramenta de machine learning para treinar um classificador nestes dados	Criar um conjunto de dados para treinar um classificador de árvore de decisão ou preditor e explorar o impacto que diferentes features têm na árvore de decisão	Comparar dois conjuntos de dados com dados reais em termos de features inclusas e como estas features estão codificadas
	Conjunto de dados grande	Ilustrar como treinar um classificador para um conceito amplo como “cachorro” requer uma grande quantidade de dados para capturar a diversidade do domínio	Ilustrar como objetos em uma imagem podem ser segmentados e rotulados para construir um conjunto de treino para reconhecimento de objetos	Avaliar um conjunto de dados usado para treinar um sistema de IA real considerando o tamanho do conjunto de dados, a forma com que os dados foram adquiridos e rotulados, o armazenamento necessário, e o tempo estimado para produzi-lo
	Viés	Examinar características e rótulos de dados de treino para detectar possíveis fontes de viés	Explicar como a escolha dos dados de treino molda o comportamento do classificador, e como esse viés pode ser introduzido se o conjunto de treino não for balanceado apropriadamente	Investigar desequilíbrios nos dados de treinamento em termos de gênero, idade, etnia e outras variáveis demográficas que podem resultar em um modelo enviesado, utilizando alguma ferramenta de visualização

Os objetivos e habilidades que se ganham para o ensino de IA e ML podem ser ainda melhor inseridos em cursos de cunho técnico, juntamente com ensino médio, oferecidos por escolas técnicas brasileiras, em específico o Instituto Federal de cada estado. Nos diversos cursos disponibilizados alguns merecem destaque por ter seu enfoque na área de computação e suas tecnologias.

Atualmente já existem diversas unidades instrucionais que adotam, principalmente, as metodologias ativas, utilizando de ferramentas visuais para facilitar no ensinamento. Com elas o aluno alcança o nível de aplicação mostrando o ciclo de uso-modifique-crie (Lee, 2011), muito utilizado nas áreas de tecnologia (Gresse von Wangenheim et al., 2020).

Com as ferramentas visuais faz com que o aluno possa interagir e executar processos de ML centrado no ser humano de maneira interativa (Gresse von Wangenheim et al., 2020). Muitas dessas ferramentas possuem um foco para iniciantes em nível de educação básica (Gresse von Wangenheim et al., 2020).

Além de ferramentas visuais que correspondem a um nível de ensino mais básico focado para iniciantes, existem as ferramentas que utilizam de linguagens textuais de programação, como por exemplo, Python. Essas ferramentas que utilizam de linguagens de programação tem um nível de dificuldade maior pelo fato de ser necessário ter um conhecimento prévio da linguagem que está sendo utilizada. Esse conhecimento envolve não somente o conhecimento da sintaxe da linguagem mas também a parte lógica para realizar a programação de um programa. Com isso o nível dessa ferramenta é considerado de intermediário a até nível avançado considerando o conhecimento de programação do usuário.

Atualmente a ferramenta que vem sendo utilizada para realizar essa programação de modelos com linguagem de programação é o Jupyter notebook, onde nele é possível realizar a compilação e a execução de pedaços de códigos além de poder adicionar blocos de texto para facilitar a explicação. Juntamente com o Jupyter vem sendo muito utilizado o ambiente de execução Google Colab, onde nele é possível ter acesso a nuvem e realizar a execução do código nela, acelerando o processo de treinamento dos modelos.

2.3. Ensino de Micologia no Ensino Médio

2.3.1 Conceitos de Micologia

Os estudos com fungos são feitos em diferentes áreas da ciência, mas a principal inclui trabalhos investigativos, cujo objetivo é encontrar e divulgar a grande diversidade de espécies de fungos para o mundo (Silva, 2018).

Atualmente são reconhecidas aproximadamente 145 mil espécies de fungos (Catalogue of Life, 2021).

O Reino Fungi possui uma grande importância ecológica e no ciclo de material biológico, isso porque esses organismos são essencialmente decompositores, realizando assim a decomposição de detritos e matéria orgânica devolvendo para o ecossistema moléculas menores, que então são utilizadas por diversas plantas (Neves et al., 2013). Além do papel essencial de decompositor, os fungos participam de diversas associações com outros fungos, bactérias, plantas e animais, onde em geral se tem simbioses benéficas entre eles (Neves et al., 2013).

Os fungos podem ser unicelulares ou filamentosos. Os filamentosos são mais conhecidos por terem a possibilidade de se tornar macroscópico, ou seja, possível de se ver a olho nu. Para isso, as células tubulares e alongadas dos fungos filamentosos, chamadas de hifas, crescem e formam um micélio, que por sua vez, pode se organizar em estruturas chamadas esporóforos, que são visíveis a olho nu (Neves et al., 2013). Um exemplo simples de esporóforo são os cogumelos que são consumidos em alimentos. Neste trabalho são tratados de fungos filamentosos e visíveis a olho nu, para que seja possível realizar a captura de imagens do mesmo. A Figura 19 mostra exemplos de cogumelos macroscópicos.



Figura 19 - Cogumelos macroscópicos (Fonte: iStock, 2020)

Além da separação de unicelulares e filamentosos, os fungos são agrupados em filós dependendo de suas características. Atualmente são reconhecidos sete filós (Blackwell et al., 2012): Neocallimastigomycota, Chytridiomycota, Blastocladiomycota, Zygomycota, Glomeromycota, Ascomycota e Basidiomycota.

Cada um dos sete filós apresentam características próprias e ainda existem discussões sobre as classificações dos fungos (Neves et al., 2013). Para este trabalho será focado em dois filós do sub reino Dikarya, são eles o Filo Ascomycota e o Filo Basidiomycota:

- Filo Ascomycota: É o maior grupo de fungos existentes. Estima-se que cerca de 75% dos fungos descritos pertençam a esse grupo, tendo um registro de cerca de 32 mil espécies. Na natureza são encontrados como parasitas, saprófitos e formando líquens. Nesse filó, se observa uma estrutura de propagação conhecida como asco, a qual possui ascósporos (esporos sexuais).
- Filo Basidiomycota: Nesse filó existem os fungos tradicionalmente conhecidos como cogumelos. São considerados os fungos mais evoluídos do reino em virtude de sua complexidade. Em maioria são fungos terrestres. Para sua reprodução, que é sexuada, forma uma estrutura conhecida como basídio, o qual contém basidiósporos

(esporos sexuados). Esses fungos são macroscópicos e se diferenciam pela forma, coloração e tamanho.

A maioria dos macrofungos são dos filios Ascomycota e Basidiomycota, com algumas exceções no filo Zygomycota (Mueller et al., 2006). Por esse motivo se decidiu utilizar fungos dos Filios Ascomycota e Basidiomycota, para que seja possível realizar a captura de imagens dos mesmos, utilizando de aparelhos fotográficos como câmeras ou celulares, e realizar a sua categorização com a rede neural.

Os fungos do filo Basidiomycota, comumente conhecidos como cogumelos, são os mais facilmente reconhecidos por serem mais comuns no cotidiano das pessoas. São utilizados como símbolo de cogumelos, em jogos ou até mesmo em representações simples, a maioria dos cogumelos utilizados para alimentação são desse filo. A Figura 20 mostra em resumo a fisiologia desse tipo de fungo.

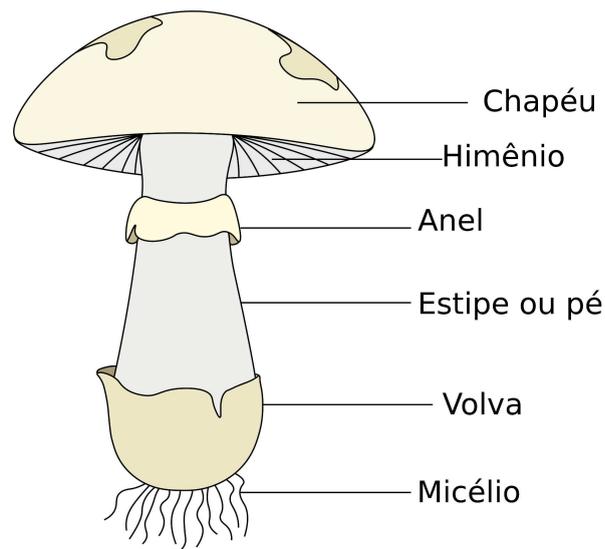


Figura 20 - Fisiologia de um fungo do filo Basidiomycota, comumente chamado de Cogumelos
(Fonte: Silva, 2018)

O Chapéu é a parte mais importante do cogumelo, pois em sua parte interior, o Himênio, se encontram os esporos reprodutivos. O Anel, caso ele exista, é uma sobra de um véu que cobre a parte inferior do Chapéu no estado jovem do cogumelo, com o crescimento do cogumelo esse véu então é rompido para o surgimento do chapéu podendo assim manter essa sobra aparente formando o anel. O Estipe ou Pé pode adquirir diferentes formas. A Volva é o resto do Véu comentado anteriormente porém podendo ter essa sobra na base do Estipe. O

Micélio é a parte vegetativa de um fungo, que consiste de uma massa de ramificação formada por um conjunto de hifas emaranhadas e possui como responsabilidade carregar nutrientes até onde o fungo necessita.

Em Santa Catarina assim como no restante do Brasil existe uma variedade enorme de fungos, porém para esse trabalho se pretende ter um foco maior em dez tipos de fungos encontrados em Santa Catarina para sua utilização nas redes neurais. São eles:

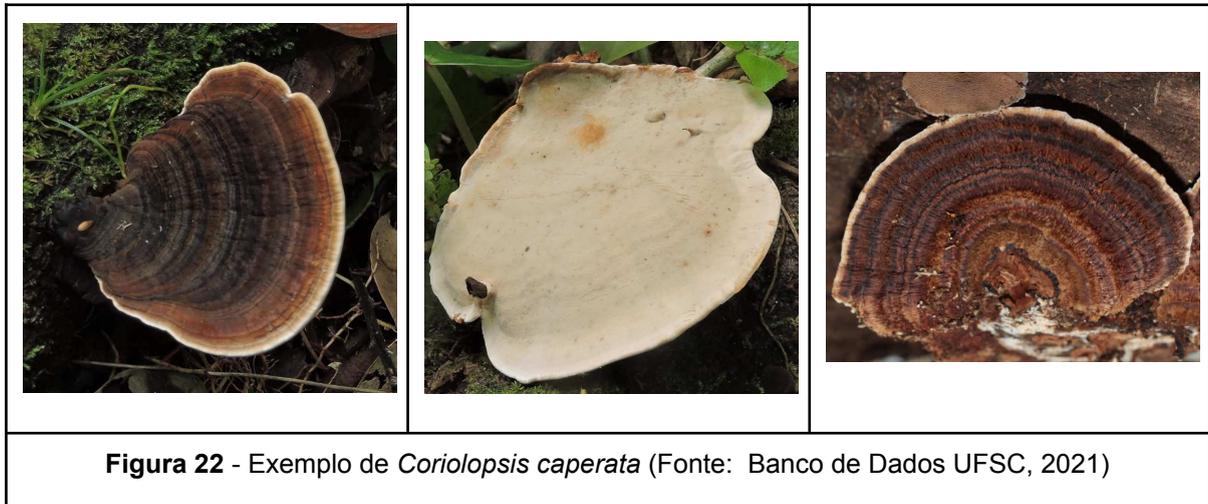
- *Auricularia fuscusuccinea*
- *Coriolopsis caperata*
- *Cyptotrama asprata*
- *Ganoderma australe*
- *Oudemansiella cubensis*
- *Phallus indusiatus*
- *Phillipsia domingensis*
- *Pycnoporus sanguineus*
- *Schizophyllum commune*
- *Trametes villosa*

Auricularia fuscusuccinea novamente é um fungo do filo Basidiomycota, é uma espécie comestível descrita cientificamente em 1842. A Figura 21 mostra exemplos de *Auricularia fuscusuccinea*.



Figura 21 - Exemplo de *Auricularia fuscusuccinea* (Fonte: Banco de Dados UFSC, 2021)

Coriolopsis caperata é um fungo do filo Basidiomycota que é encontrado em troncos em decomposição em quase todo Brasil. A Figura 22 mostra exemplos de *Coriolopsis caperata*.



Cyptotrama asprata é um fungo do filo Basidiomycota, comumente encontrado em regiões tropicais pelo mundo, tem como características a sua cor alaranjada e amarela e espécies novas possui uma camada de fibras que parecem espinhos, como mostra a Figura 23.



Ganoderma australe é um fungo do filo Basidiomycota que também é encontrado em árvores, nos estágios iniciais de sua colonização se acredita que o fungo é parasita, porém à medida que a árvore morre o fungo se torna saprófita, como ocorre com a maioria dos fungos parasitas. A Figura 24 mostra exemplos de *Ganoderma australe*.



Figura 24 - Exemplo de *Ganoderma australe* (Fonte: Banco de Dados UFSC, 2021)

Oudemansiella cubensis, com nome popular de Fungo-de-Porcelana, nascem em troncos e raízes de árvores, é comestível e tem textura delicada e notas aromáticas remetem ao coco. Seu filo é o Basidiomycota. A Figura 25 mostra exemplos do *Oudemansiella cubensis*.



Figura 25 - Exemplo de *Oudemansiella cubensis* (Fonte: Banco de Dados UFSC, 2021)

O *Phallus indusiatus* é um fungo do filo Basidiomycota, chamado popularmente de véu-de-noiva por sua estrutura peculiar, como mostra a Figura 26. Possui grande importância nos ecossistemas por ter a capacidade de desintoxicar vários poluentes ambientais.



Figura 26 - Exemplo de *Phallus indusiatus* (Fonte: Banco de Dados UFSC, 2021)

O *Phillipsia domingensis*, o único da lista do filo Ascomycota, encontrado em maior quantidade nas regiões subtropicais e tropicais. O fungo foi circunscrito por Miles Joseph Berkeley em 1881.



Figura 27 - Exemplo de *Phillipsia domingensis* (Fonte: Banco de Dados UFSC, 2021)

O *Pycnoporus sanguineus*, conhecido como Urupê ou orelha de pau, é um fungo do filo Basidiomycota que costuma crescer sobre troncos de árvores. Esse fungo possui o trabalho de decompositor na cadeia alimentar, se alimentando de matéria morta, podendo assim até indicar o estado da árvore. A Figura 28 mostra exemplos de *Pycnoporus sanguineus*.



Figura 28 - Exemplo de *Pycnoporus sanguineus* (Fonte: Banco de Dados UFSC, 2021)

O *Schizophyllum commune* também é encontrado em árvores em decomposição e são colhidos por ter valor medicinal e ter um perfil de sabor aromático, possui uma característica interessante que o fungo se torna endófito antes da árvore morrer. O *Schizophyllum commune* também é um fungo do filo Basidiomycota. A Figura 29 mostra exemplos de *Schizophyllum commune*.



Figura 29 - Exemplo de *Schizophyllum commune* (Fonte: Banco de Dados UFSC, 2021)

Trametes villosa do filo Basidiomycota apresenta uma cor branca e apresenta interesse comercial, pois possui potencial biotecnológico para de degradação de compostos organoclorados e descoloração de corantes têxteis. A Figura 30 mostra exemplos de *Trametes villosa*.



Figura 30 - Exemplo de *Trametes villosa* (Fonte: Banco de Dados UFSC, 2021)

Analisando as figuras exemplos dos fungos que serão utilizados, é possível identificar a grande diferença entre eles, além de não seguirem por completo a base do filo Basidiomycota apresentada na Figura 20.

2.3.2 Ensino de Micologia

A matéria de micologia é aplicada juntamente a área de conhecimento de biologia referente ao tópico de biodiversidade, políticas ambientais e ecossistemas. Tendo o foco maior com os macrofungos, é contemplado na habilidade “*EF04C106: Relacionar a participação de fungos e bactérias no processo de decomposição, reconhecendo a importância ambiental deste processo*” (MEC, 2017).

Com a existência de diversos tipos de fungos e seu importante papel no ecossistema e até mesmo na economia, a Micologia é um tema que vem tendo seu crescimento em interesse, com seus tópicos e assuntos sendo comentados em diferentes mídias abrangendo desde os jogos até a culinária. Assim como na Biologia em que sempre se tem adição de novas descobertas, a Micologia também sempre está em constante atualização com novas descobertas de espécies e organismos.

A micologia é uma área da Biologia que tem como campo de estudo, os fungos. Os fungos são organismos cosmopolitas, eucariontes, unicelulares (leveduriformes) ou multicelulares (filamentosos), que possuem reprodução sexuada ou assexuada e com parede celular predominantemente de quitina. Sua importância pode estar ligada a três aspectos principais: aspectos ecológicos, por serem seres decompositores; aspectos econômicos, envolvendo principalmente alimentação e

meio ambiente; e aspectos clínicos, por causarem patologias ao homem (SOUZA, 2021).

Com o crescimento da área cada vez maior, se tem a noção que é necessário realizar mudanças no modo em que se é ensinado aos jovens a matéria de Micologia em sua educação básica.

Considerando a falta de materiais extras para complementar a utilização de livros didáticos, esse projeto pretende utilizar de tecnologias de Machine Learning para complementar e até mesmo atualizar o método de ensino sobre Micologia para jovens do Ensino Médio, sendo esse método até uma maneira mais prática para ensinar os estudantes.

3. ESTADO DA ARTE

A fim de levantar o estado da arte sobre o ensino de *Machine Learning* para jovens do ensino médio utilizando de macrofungos como base para inserção ao currículo, são apresentados neste capítulo os resultados de um mapeamento sistemático da literatura com base no procedimento proposto por Petersen et al. (2008). Este mapeamento complementa/atualiza uma pesquisa anterior (MARQUES et al., 2020) fazendo uma busca por publicações no período de 2019-2021.

3.1. Definição do protocolo de revisão

A pergunta de pesquisa é: Quais unidades instrucionais existem para ensinar conceitos de ML no contexto do ensino médio?

Essa pergunta de pesquisa é refinada nas seguintes questões de análise:

- AQ1. Quais unidades instrucionais existem?
- AQ2. Quais conceitos de ML são ensinados nessas unidades instrucionais?
- AQ3. Quais são as características instrucionais dessas unidades instrucionais em relação ao conteúdo de ML?
- AQ4. Quais são as características instrucionais dessas unidades instrucionais?
- AQ5. Como foram desenvolvidas as unidades instrucionais e qual é a avaliação de qualidade dessas unidades instrucionais

Critérios de inclusão/exclusão: Foram consideradas quaisquer unidades instrucionais (curso, atividade, tutorial, etc.) que tivessem foco no ensino de computação, incluindo conceitos de ML no ensino Médio, publicadas em língua inglesa entre 2019 e 2021. Unidades instrucionais que focam no ensino de ML no ensino superior, ou infantil/fundamental e/ou que abordassem o ensino de computação sem aprofundar-se em conceitos de ML foram excluídas. Além dessas, publicações que não especificam o estágio escolar na educação básica foram excluídas. Também foram excluídas publicações de blogs, vídeos ou ferramentas que não possuem uma unidade instrucional.

Critérios de qualidade: Foram considerados apenas artigos ou materiais com informações suficientes relacionados a conceitos de ML, indicando, por exemplo, conteúdo de aulas, materiais de apoio, etc.

Fontes dos dados: Foram examinados todos os materiais e artigos publicados em inglês e disponíveis nas mais importantes bibliotecas digitais (Incluindo ACM Digital Library, IEEEExplore, Scopus) acessíveis por meio do portal CAPES. O Google também foi incluído, por indexar um grande conjunto de dados de diferentes fontes (Haddaway et al., 2015), já que nessa área emergentes muitas unidades instrucionais não foram publicadas. Dado o foco de pesquisa do MIT Media Lab nessa área, publicações desse grupo também foram procuradas. Literatura secundária foi também incluída.

Definição da *string* de busca: A *string* de busca foi composta de conceitos relacionados à questão de pesquisa, incluindo sinônimos. Dessas palavras chave, a search string foi adaptada para cada fonte de dados apresentada na Tabela 2.

Tabela 2: Search string para cada fonte

Fonte	Search String
ACM	[[Abstract: teach*] OR [Abstract: education] OR [Abstract: course] OR [Abstract: mooc] OR [Abstract: learn*]] AND [[Abstract: "machine learning"] OR [Abstract: "data science"] OR [Abstract: "artificial intelligence"] OR [Abstract: "deep learning"]] AND [[Abstract: "k-12"] OR [Abstract: school*] OR [Abstract: kids] OR [Abstract: children] OR [Abstract: teen*]] AND [Publication Date: (01/01/2019 TO *)]
IEEE	((("Abstract":teach*) OR ("Abstract":education) OR ("Abstract":course) OR ("Abstract":MOOC) OR ("Abstract":learn*)) AND ((("Abstract": "machine learning") OR ("Abstract": "data science") OR ("Abstract": "artificial intelligence") OR ("Abstract": "deep learning"))) AND ((("Abstract": "k-12") OR ("Abstract": "school*") OR ("Abstract": "kids") OR ("Abstract": "children") OR ("Abstract": "teen*")))) 2019 - 2021
Scopus	TITLE-ABS-KEY ((teach* OR education OR course OR mooc OR learn*) AND ("machine learning" OR "data science" OR "artificial intelligence" OR "deep learning") AND ("k-12" OR school* OR kids OR children OR teen*)) AND (LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR ... OR LIMIT-TO (PUBYEAR , 2010)) AND (LIMIT-TO (SUBJAREA , "COMP"))
Google Scholar	"machine learning" teach "K-12" OR school
MIT media lab ¹	--

Com o intuito de realizar o complemento/atualização de trabalhos realizados anteriormente se teve o foco de realizar a busca nos anos de 2019 a 2021.

¹ <https://appinventor.mit.edu/explore/research>
<https://www.media.mit.edu/groups/lifelong-kindergarten/publications/>

3.2. Execução da busca

A busca foi realizada entre outubro e novembro de 2021 pelo autor do presente trabalho e revisada pela orientadora. A busca inicial resultou em 18.957 artigos, dos quais foram realizados filtros para selecionar artigos com maior relevância de acordo com os critérios de inclusão, exclusão e qualidade, chegando em um resultado de 14 artigos relevantes.

Tabela 3: Resultados de pesquisa

Base de Dados	Nº de artigos resultantes da busca	Nº de resultados analisados	Nº de resultados potencialmente relevantes	Nº de resultados relevantes
ACM	210	210	12	4
IEEE	632	250	41	1
SCOPUS	1.815	300	7	1
Google Scholar	16.100	300	27	3
MIT media lab	200	200	12	5
Total (sem duplicados)				14 (11)

O número de resultados analisados variou de acordo com o número de artigos resultantes da busca. As pesquisas retornaram um grande número de resultados mesmo com a calibração da *string* de busca. Isso ocorre pelo fato de muitos dos resultados descreverem sobre como utilizar da IA para auxiliar na educação de jovens ou na identificação de possíveis problemas com as crianças, como queda de rendimento escolar ou até mesmo doenças, que corresponde a alguns termos utilizados na pesquisa. Então se manteve a *string* de busca e se limitou a análise para aqueles que apresentam maior relevância. No primeiro estágio, foi realizada uma rápida revisão dos títulos e resumos que possivelmente coincidia com o critério de inclusão. Já na segunda etapa, os materiais foram lidos para verificar sua relevância em relação aos critérios de inclusão e exclusão. Muitos artigos foram excluídos por terem um foco somente no uso de IA para a educação.

3.3. Análise dos Resultados

3.3.1 Quais unidades instrucionais existem?

Um total de 14 unidades instrucionais abrangendo o ensino de *ML* no ensino médio foram identificadas nos artigos pesquisados (Tabela 3). Algumas unidades encontradas são do mesmo autor(es) e sobre o mesmo tema as quais foram realizadas uma junção, resultando em 11 artigos relevantes descritos na tabela abaixo (Tabela 4).

Tabela 04 - Unidades Instrucionais e suas descrições

Referência	Nome da UI	Breve Descrição
(Biehler e Fleischer, 2021)	Introducing Students to Machine Learning with Decision Trees Using CODAP and Jupyter Notebooks	Utiliza de um programa que utiliza de ML, se recebe como dados a média de utilização de jogos eletrônicos para prever quem joga jogos online mais frequentemente.
(Bhatia, 2020) (Bhatia e Lao, 2020)	Personal Audio Classifier (PAC)	Ensino de ML utilizando de <i>Áudio Bites</i> para alunos do ensino médio.
(Mariescu-Istodor e Jormanainen, 2019)	Machine Learning for High School Students	Realiza o ensino de ML para jovens de 13 a 19 anos utilizando de categorização de imagens. Apresenta o processo de construção do APP e os resultados.
(Reyes et al., 2020)	A Preliminary Work on Visualization-based Education Tool for High School Machine Learning Education	Para realizar o ensino para os estudantes, é desenvolvido um jogo 2D utilizando da Unity3D como engine.
(Rodríguez-García et al., 2021)	Evaluation of an Online Intervention to Teach Artificial Intelligence with LearningML to 10-16-Year-Old Students	Apresenta um curso para ensinar IA e ML para jovens de 10 a 16 anos utilizando da plataforma LearningML.
(Tamashiro et al., 2021)	Introducing Teenagers to Machine Learning through Design Fiction: An Exploratory Case Study	Realiza o ensino de ML tendo o foco em análise facial com ML supervisionada.
(Tang et al., 2019) (Tang, 2019)	PIC: A Personal Image Classification Webtool for High School Students	Realiza a classificação de imagens utilizando o App Inventor como base.
(Voulgari et al., 2021)	Learn to Machine Learn: Designing a Game Based Approach for Teaching Machine Learning to Primary and Secondary Education Students	Utiliza de jogos eletrônicos e do framework LearnML para o ensino de IA e ML.
(Wan et al., 2020)	SmileyCluster: Supporting Accessible Machine Learning in K-12 Scientific Discovery	Programa em web para ensinar ML realizando a identificação das emoções de faces simplificadas desenhadas.
(Zimmermann-Niefeld et al., 2019)	Youth Learning Machine Learning through Building Models of Athletic Moves	O ensino de ML é feito a partir da captura e a análise dos movimentos feitos pelos estudantes em exercícios físicos e esportes, para que a

		máquina aponte os movimentos que podem ser aprimorados.
(Zhu, 2019)	An Educational Approach to Machine Learning with Mobile Applications	Uma aplicação web para criar um app mobile que pode utilizar de uma ML escolhida pelo estudante.

Comparando com relação ao resultados encontrados por Marques et al. (2020b) é possível identificar uma queda no número de unidades instrucionais publicadas em 2021 se comparado a 2019. Outro ponto observado foi o aumento da utilização de cursos realizados de maneira remota, já que até 2019 se tinha majoritariamente a aplicação de cursos via presencial. Por conta da pandemia do COVID-19 algumas unidades instrucionais tiveram que se adaptar para realizar os estudos de maneira remota, visando a segurança de todos os envolvidos.

3.3.2 Quais conceitos de ML são ensinados nessas Unidades Instrucionais?

Se observou que a grande maioria das unidades instrucionais ensina o que é o aprendizado de *ML* antes de mostrar com mais detalhes o funcionamento de *IA/ML*. Com relação ao ensino da *ML* se observou que as abordagens de árvore de decisão e redes neurais são os tópicos mais mostrados nas unidades instrucionais. Com relação aos algoritmos utilizados se teve uma pequena maioria a utilização do método de aprendizagem supervisionado, com a aprendizagem não supervisionada logo atrás. Somente duas unidades instrucionais apresentaram os fundamentos de redes neurais enquanto a maioria apresenta aos alunos as limitações de *ML* e como os dados de treinamento podem influenciar no resultado final da *ML*. Também é abordado, na minoria das unidades instrucionais, os tópicos sobre aspectos éticos e sociais envolvendo *ML* e as oportunidades de carreira com relação a *ML*.

Algumas exceções são , Reyes et al.(2020) e Wan et al.(2020) que apresentaram para os estudantes a abordagem de *clustering*, pouco apresentada nas outras unidades instrucionais por ser um tópico considerado mais avançado. A unidades instrucionais apresentada por Voulgari et al.(2021) apresenta o tipo de aprendizagem de *Reinforcement*, sendo a única unidade instrucional das observadas a apresentar esse método

Tabela 04 - Quais conceitos de ML ensinados nas Unidades Instrucionais?

Referência	O que é "aprendizado"	Abordagens a machine learning							Tipos de algoritmos de aprendizagem por estilo de aprendizagem			Fundamentos de redes neurais	Como o treinamento de dados influencia no aprendizado	Limitações do Machine Learning	Preocupações éticas / implicações sociais	Oportunidades de carreira
		Algoritmo de regressão	Baseado em Instância	Árvore de decisão	Algoritmo de Bayesian	Clustering	Rede Neural	Outros	Supervisionado	Não Supervisionado	Reinforcemento					
(Biehler e Fleischer, 2021)				x					x				x		x	
(Bhatia, 2020) (Bhatia e Lao, 2020)	x						x		x			x	x			
(Mariescu-Istodor e Jormanainen, 2019)				x			x			x						
(Reyes et al., 2020)	x	x				x			x	x				x		x
(Rodríguez-García et al., 2021)	x						x		x				x	x		
(Tamashiro et al., 2021)	x						x			x			x	x		
(Tang et al., 2019) (Tang, 2019)		x							x				x	x	x	
(Voulgari et al., 2021)	x			x					x		x		x	x	x	
(Wan et al., 2020)	x					x				x			x	x		x

(Zimmerman n-Niefeld et al., 2019)	x			x						x						
(Zhu, 2019)							x		x			x	x			

3.3.3 Quais os conteúdos de ML abordados?

Dentre as unidades instrucionais estudadas se observou uma abrangência do tipo de tarefa realizada pela ML, tendo em geral a classificação como a tarefa mais utilizada. As unidades instrucionais que propõem atividades e conteúdos voltados à classificação, a grande maioria utiliza seu conjunto de dados imagens ou são convertidos para imagem, como é o caso de Bhatia e Lao (2020), que utilizam de representações visuais das gravações de áudios, chamados de espectrograma. Outras unidades instrucionais não utilizaram de imagens e sim de vídeos, muitas vezes “ao vivo”, utilizando de webcams e câmeras de celular para realizar a gravação dos dados. A utilização de algo mais interativo como a utilização de vídeos aumentou se comparado a estudos anteriores, como por exemplo Marques et al. (2020b). Em específico a UI de Zimmermann-Niefield et al. (2019) utiliza de movimentos realizados em esportes para que a *ML* mostre aos estudantes onde existem melhorias a serem feitas em seus movimentos. Com esse estudo se mostra a evolução com relação ao modo de ensinar os estudantes, não sendo obrigatório o ensino direto de salas de aulas, podendo ser algo mais interativo e divertido para o estudante participante.

O trabalho de Voulgari et al (2021), Biehr e Fleischer (2021) e leva em consideração que atualmente crianças com 4 anos de idade já começam a ter contato com jogos, com isso se utiliza do meio de comunicação sendo um jogo virtual para realizar o ensino do tópico de IA de maneira mais fácil para estudantes. O jogo ArtBot ensina a importância do *labeling* e o funcionamento de classificação de imagens, tendo como objetivo classificar imagens se são esculturas ou pinturas.

A grande maioria das unidades instrucionais apresenta aos participantes as etapas básicas para o treinamento de ML, como Gerenciamento de dados, Treinamento de modelo e Implantação do modelo. Por serem unidades instrucionais com duração pequena não é realizado as etapas de avaliação e ajustes, sendo comentado para os participantes e deixando como atividades extras para os mesmos. As que realizaram a avaliação utilizaram da acurácia como medição principal.

Em geral as unidades instrucionais utilizaram de ferramentas próprias para facilitar o ensino aos participantes. Para a linguagem de programação se utilizou de

JavaScript, porém os participantes poucas vezes tiveram que escrever código do zero, somente realizando pequenas alterações para seus casos individuais ou em atividades extras onde o projeto inicial é alterado para atender algum objetivo do aluno.

A maioria das unidades instrucionais utiliza ferramentas de programação baseadas em blocos, como por exemplo o Scratch. Poucos mostraram para os estudantes o código por trás da *ML* optando por um ensino controlado orientando somente as partes que devem ser alteradas para alcançar o objetivo final, sem dar a liberdade para os estudantes de escrever do zero seus programas.

O curso feito por Mariescu-Istodor e Jormanainen (2019) é o único que apresenta o ensino utilizando de programação textual, utilizando da linguagem de preferência do estudante transferindo para JavaScript para realizar a implementação do App. Para observar o aprendizado dos estudantes foi criado um ambiente de programação onde se observa as alterações em tempo real, comportamento parecido ao Google Docs, nesse ambiente os professores conseguem observar o que cada estudante/grupo de estudantes está desenvolvendo e realizar as alterações e conselhos necessários.

Para a implementação, as unidades instrucionais em geral utilizam de sistemas web ou aplicativos. Os aplicativos em geral utilizam do App Inventor desenvolvido pelo MIT, isso devido a facilidade para realizar a implantação.

O curso desenvolvido por Biehler e Fleischer (2021) se destaca por utilizar do Jupyter Notebook, o qual foi utilizado em um desenvolvimento controlado, dando ao estudante pequenos trechos de código para realizar as alterações. Como descrito em seu artigo: “Um benefício especial (do Jupyter) é que o código Python pode ser ocultado por trás de widgets interativos para que os alunos possam trabalhar com poderosas ferramentas Python em um ambiente baseado em menu sem sequer ver o código.”.

Tabela 05 - Conteúdos abordados nas Unidades Instrucionais?

Referência	Objetivo de aprendizado	Domínio de Aplicação	Tipo de tarefa de ML	Processos de ML	Conjunto de dados	Framework/Ferramentas para ML	Linguagens de programação	Medição de Desempenho do modelo de ML?	Implantação
(Biehler e Fleischer, 2021)	Ensino de ML utilizando de árvores de decisão	Jogos eletrônicos	Previsão de comportamento de acordo com base de dados	Gerenciamento de dados, Treinamento de modelo	Dados sobre utilização de mídia pelos adolescentes	NI	CODAP, Jupyter Notebook	NI	NI
(Bhatia, 2020) (Bhatia e Lao, 2020)	Introduzir conceitos básicos relacionados a IA e ML	Falas comuns gravados pelos alunos	Classificação de áudio de trechos curtos de áudio	Gerenciamento de dados, Treinamento de modelo, Implantação do modelo	Imagens de trechos curtos de áudio	TensorFlow.js	Ambiente visual PAC, App Inventor com extensão (PAC)	Acurácia e Recall	Implantado um aplicativo móvel com App Inventor
(Mariescu-Istodor e Jormanainen, 2019)	Introduzir ML para estudantes com o conhecimento adquirido em aulas de programação presentes no currículo.	Biologia, identificando espécies.	Categorização de imagens	Implantação do modelo	Imagens tiradas na hora para interpretação com a máquina	NI	HTML, JavaScript (sempre utilizados) C, C++ e Python (dependendo da escolha do participante, sendo convertido para JavaScript depois)	NI	Sistema web
(Reyes et al., 2020)	Ensino de ML utilizando de uma educação baseada em jogos eletrônicos	Jogos eletrônicos	NI	Gerenciamento de dados, Engenharia de Features (Características/Recursos), Treinamento de modelo, Avaliação do modelo	NI	NI	NI	NI	Sistema web
(Rodríguez-García et al., 2021)	Ensino do ML para alunos utilizando da plataforma web criada pelos autores, LearningML	Pode variar dependendo da escolha do estudante	Reconhecimento de texto ou imagens	Gerenciamento de dados, Treinamento do modelo e avaliação do modelo	Depende da escolha do estudante, variando entre imagens e textos	NI	LearningML (ambiente visual) Scratch	NI	Aplicação
(Tamashiro et al., 2021)	Utilizar da ficção de design para introduzir estudantes sobre ML	Criação de personagens, design de personagens.	Análise facial	Gerenciamento de dados e Treinamento do modelo	Imagens tiradas na hora para geração de personagens	Teachable Machine e Adacraft	NI	NI	NI
(Tang et al., 2019) (Tang, 2019)	Utilizar do PIC para que qualquer pessoa consiga entender o funcionamento de classificação de imagens	Reconhecimento facial	Análise facial	Gerenciamento de dados e Treinamento do modelo	Imagens tiradas da webcam do usuário	Webcam Pacman, MobileNet e SqueezeNet. TensorFlow.js	PIC, App inventor	Acurácia, tabela de correção e gráfico de confiança	Implantado um aplicativo móvel com App Inventor ou diretamente ao PIC

(Voulgari et al., 2021)	Utilizar de um jogo desenvolvido em Unity para ensinar os estudantes sobre ML	Objetos de arte (pinturas e/ou esculturas)	Reconhecimento de Imagens	Gerenciamento de dados e Treinamento do modelo	Imagens tiradas pelos usuários de obras de arte e um dataset de artes do Museu de Arte Metropolitana.	NI	Unity	Learning rate	Sistema web no Learn ML utilizando de WebGL
(Wan et al., 2020)	Como criar ambientes de aprendizagem que suporta conceitos de ML para jovens estudantes e como esse ambiente pode melhorar mais a descoberta científica no contexto STEM	Reconhecimento facial	Análise facial	Gerenciamento de dados e Treinamento do modelo	Ilustrações de rostos apresentando diferentes emoções	NI	Scratch	NI	Sistema web
(Zimmermann-Niefeld et al., 2019)	Utilizar do esporte para fazer a ligação com o assunto de ML ensinando seus conceitos de maneira mais prática e descontraída.	Melhoria de movimentos em esportes	Movimentos de esporte	Gerenciamento de dados, Treinamento do modelo e avaliação do modelo	Imagens e Vídeos dos movimentos no esporte	AlpacaML	NI	NI	Aplicação iOS
(Zhu, 2019)	Utilizar das extensões criadas para o App Inventor para criar aplicações divertidas e educacionais	Identificação de objeto e controle de robôs	Categorização de imagens e áudio	Implantado um aplicativo móvel com App Inventor	Imagens, vídeos e áudios	TensorFlow.js, MobileNet, Teachable Machine e OCR	Extensões feitas em Java e JavaScript. App Inventor	NI	Implantado um aplicativo móvel com App Inventor

3.3.4 Quais são as características instrucionais dessas unidades instrucionais?

Algumas unidades instrucionais se identificam como cursos, oferecendo aulas com mais tópicos podendo chegar a 80 horas, ou cursos mais curtos de até 8 horas. Estes cursos em geral disponibilizam o plano de ensino utilizado. As demais unidades instrucionais se consideram oficinas tendo em sua maioria somente um encontro durando algumas horas.

Em geral, as unidades instrucionais analisadas apresentam o nível de dificuldade considerado para novatos, por não possuírem pré-requisitos e não ter programação direta. A única exceção é Mariescu-Istodor e Jormanainen (2019), isto porque além de apresentar mais informações específicas, a escola onde foi aplicado a oficina já apresentava aula de programação no currículo.

Como comentado anteriormente houve um aumento da utilização do modo instrucional online, mas ainda assim várias unidades instrucionais têm como plano inicial o modo presencial, aplicando eles antes no início da pandemia do COVID-19. Alguns apresentaram dificuldades para encontrar participantes justamente por causa dos problemas enfrentados pelo mundo inteiro.

Para expor o assunto aos participantes, em grande maioria se aplicou aulas expositivas ou o contato direto com a aplicação. As aulas expositivas utilizavam slides e vídeos para realizar o ensino aos participantes e somente após a exposição do conteúdo aos participantes que se teve início a parte prática do curso ou oficina. Todas as unidades instrucionais chegam ao ponto de prática com os participantes, seja por atividades práticas ou utilizando da ferramenta. Para a parte prática das unidades instrucionais muitas utilizaram de seus projetos próprios, com poucas exceções utilizando somente de sistemas terceiros, como por exemplo, Jupyter Notebook ou Google *Teachable Machine*.

Em grande maioria as unidades instrucionais utilizam do ciclo “use” pelo fato de ser mais fácil de apresentar para participantes novatos na área, onde se tem os programas já feitos e só é realizada a explicação do que já foi construído anteriormente. Nos casos que não ficam somente em “use” se chega até o estágio de “create” propondo para o participante criar seu próprio projeto, visando o entendimento do participante na etapa de “modify”.

Para auxiliar nas explicações muitos utilizam de slides como material didático junto com videos explicativos. Em alguns casos se utiliza da própria ferramenta como o material para ensinar os participantes.

Todas as unidades instrucionais apresentam material em inglês com exceção a do Biehler e Fleischer (2021) Mariescu-Istodor e Jormanainen (2019) que apresentam também material na língua em que foi aplicada a unidade, alemão e romeno, respectivamente. Se observa com isso a falta de unidades instrucionais na língua portuguesa, com a maioria focando na língua inglesa a qual é a mais utilizada mundialmente.

Poucos artigos comentaram sobre o treinamento do instrutor, considerando que em muitos casos quem realizava a apresentação da unidade instrucional foi o próprio criador dos projetos. Mariescu-Istodor e Jormanainen (2019) e Voulgari et al. (2021) foram os únicos a comentarem sobre o conhecimento que o instrutor deve ter, em que ambos comentam e experiência básica sobre HTML e conhecimento sobre JavaScript no caso do Mariescu-Istodor e Jormanainen (2019).

Tabela 06 - Características das Unidades Instrucionais?

Referência	Tipo de UI	Nível de dificuldade	Estágio(s) do ciclo use-modify-create	Duração	Modo instrucional	Método instrucional	Material Instrucional	Método de avaliação da aprendizagem do aluno	Idioma	Licença	Treinamento do Instrutor
(Biehler e Fleischer, 2021)	Curso (projeto interdisciplinar entre matemática e ciências da computação)	Novato	use	6 fases totalizando 8 aulas	NI	Aulas expositivas, demonstrações, discussões e atividades práticas	CODAP, Jupyter Notebook e slides	NI	Inglês e Alemão	Acesso gratuito	NI
(Bhatia, 2020) (Bhatia e Lao, 2020)	Oficina	Novato	use, modify, create	Duas aulas de 55 minutos cada	Presencial	Aulas expositivas, demonstrações, discussões e atividades práticas	Sistema web e extensão PAC, App inventor, aplicações de ML disponível na internet e slides	NI	Inglês	Creative Commons Attribution-Share Alike 4.0	NI
(Mariescu-Istodor e Jormanainen, 2019)	Oficina	Intermediário	use, modify, create	Aula de 2 horas	Presencial	Atividades práticas e vídeos demonstrativos	Aplicação web, slides e ambiente de escrita sincronizada.	NI	Romeno e Inglês	Gratuito desde que não seja distribuído para ganhos monetários	Conhecimento com HTML e JavaScript
(Reyes et al., 2020)	Curso	Novato	use	4 módulos educacionais	Online	Explicação, Animação, aplicação no mundo real e tarefas	Aplicação web em formato de jogo.	NI	Inglês	NI	NI
(Rodríguez-García et al., 2021)	Oficina	Novato	use, modify, create	1 seminário	Online	Explicação via seminário online e implementação prática	Sistema web LearningML	NI	Inglês	GNU Affero GPL free, liberados a todos.	NI
(Tamashiro et al., 2021)	Curso	Novato	use	3 sessões	Online	Aula expositiva, atividades práticas e discussões	Teachable Machine e Adacraft	NI	Inglês	NI	NI
(Tang et al., 2019) (Tang, 2019)	Oficina	Novato	use, modify, create	2 aulas de 50 minutos	Presencial	Aula expositiva e a criação de um jogo mobile que depende das expressões faciais do jogador	Sistema web e extensão PIC, App inventor e Google Slides	NI	Inglês	Creative Commons Attribution-Share Alike 4.0	NI
(Voulgari et al., 2021)	Curso	Novato	use	NI	NI	Exposição direta ao jogo e explicações utilizando do jogo como base	Jogo utilizado (ArtBot)	NI	NI	Gratuito desde que não seja distribuído para ganhos monetários	Conhecimento dos conceitos básicos de ML

(Wan et al., 2020)	Oficina	Novato	use	2.5 horas	Presencial	Aula expositiva, atividades práticas e discussões	Slides e o sistema próprio SmileyCluster	NI	Inglês	Direitos autorais para componentes deste trabalho pertencente a terceiros que não a ACM deve ser homenageado	NI
(Zimmermann-Niefield et al., 2019)	Oficina	Novato	use	3 horas	Presencial	Atividades práticas e tutorial introdutório a ferramenta	Sistema próprio AlpacaML	NI	Inglês	Gratuito desde que não seja distribuído para ganhos monetários	NI
(Zhu, 2019)	Curso	Novato	use, modify, create	6 classes de 1 hora e 30 minutos	Presencial	Aulas expositivas e atividades práticas	Slides e o Sistema Web com as extensões desenvolvidas no artigo	NI	Inglês	Creative Commons Attribution-Share Alike 4.0	NI

3.3.5 Como foram desenvolvidas as unidades instrucionais e qual é a avaliação de qualidade dessas unidades instrucionais?

Na maioria das unidades instrucionais o método de desenvolvimento não é descrito, focando mais na sua aplicação aos participantes de seus respectivos cursos ou oficinas. Em geral os estudos realizados nas unidades instrucionais são estudos de casos, em que existe um planejamento do que será avaliado, qual grupo de alunos vão participar do curso e como é possível realizar a replicação dos estudos com diversas outras turmas de participantes, tendo casos com mais de 100 participantes. Em casos específicos como Tamashiro et al. (2021) e Zimmermann-Niefeld et al. (2019) se utilizou do método de estudo *ad-hoc*, um método mais informal com contato direto aos participantes, isso foi decidido pelos autores pelo fato do número de participantes ser pequeno, 2 e 5 participantes respectivamente.

Em geral as unidades instrucionais focaram mais nos quesitos de interesse e a qualidade dos cursos, se encontrou menos avaliações abordando a avaliação da aprendizagem do aluno.

Para realizar a coleta dos foram adotados diferentes tipos, focando mais em questionários via texto ou entrevistas de forma oral, sendo semi-estruturada/estruturada.

Em todas as unidades instrucionais foi apresentado resultados positivos, de acordo com os autores, indicando um aumento do conhecimento sobre *ML* e *IA* entre os participantes e uma boa recepção com relação aos cursos e oficinas aplicados.

Tabela 07 -Desenvolvimento das Unidades Instrucionais?

Referência	Método de desenvolvimento	Tipo de estudo	Fatores de qualidade avaliados	Tipo de coleta de dados utilizado para a avaliação das unidades instrucionais	Métodos de análise de dados	Tamanho da amostra	Principais resultados
(Biehler e Fleischer, 2021)	NI	Estudos de casos	Compreensão dos conceitos principais sobre ML e árvores de decisão	NI	NI	NI	Apresentou as diferentes maneiras que os alunos pensam com relação ao mesmo problema utilizando de ML. E uma recepção positiva com relação a utilização de CODAP e Jupyter Notebook.
(Bhatia, 2020) (Bhatia e Lao, 2020)	NI	Série de 3 estudos de casos	Compreensão básica dos conceitos fundamentais de aprendizado de máquina. Conscientização em torno dos aplicativos de aprendizado de máquina no mundo real. Eficácia da ferramenta de aprendizado de máquina PAC no navegador> Eficácia da extensão PAC e aplicativos MIT App Inventor	Questionário pré e pós aplicação do curso	Estatística descritiva (média, mediana e alcance) e análise qualitativa	80 estudantes com 28 acima de 18 anos	Houve um crescimento geral dos estudantes que participaram da oficina, aumentando sua confiança sobre o tópico mas também seu conhecimento geral sobre a área.
(Mariescu-Istodor e Jormanainen, 2019)	NI	Estudos de casos	Interesse e prazer Competência percebida Esforço e importância Valor e utilidade	Questionário pré e pós aplicação da aula	NI	Três conjuntos diferentes de alunos, o primeiro S1 com 22, S2 com 11 e S3 com 18 totalizando 51	Apresentou os ensinamentos sobre ML para os estudantes, treinando primeiro com imagens reais de peixe, elefante e girafa. Ao se introduzir os desenhos das crianças muitos desenhos não foram reconhecidos, se explicando para a turma os motivos do porque não foi identificado
(Reyes et al., 2020)	Desenvolvido o jogo na engine Unity3D.	NI	NI	NI	NI	NI	A utilização de um método educacional utilizando de jogos se espera ter um maior engajamento com novos usuários interessados em entrar na área de ML. O projeto ainda é um protótipo e ainda vai ser testado.

(Rodríguez-García et al., 2021)	NI	Estudos de casos	Se a criança consegue melhorar seu conhecimento de ML utilizando de LearningML. Como a criança pensa sobre LearningML depois de desenvolver projetos com ele, Se as crianças mudam sua percepção sobre IA depois de realizar atividades no LearningML.	Questionário com 14 perguntas aplicadas antes do seminário, o mesmo questionário é aplicado após o seminário com questões extras sobre a plataforma utilizada.	Estatística descritiva (média, mediana e alcance) e análise qualitativa	494 Participantes de 10 a 16 anos, onde 469 responderam o pré-questionário, 184 o pós e 162 fizeram ambos, retirando os que não responderam todas as respostas se alcançou 135 participantes válidos	Se concluiu que é possível sim utilizar da ferramenta desenvolvida, já que obteve sucesso ensinando as crianças de 10-16 anos sobre ML e IA. Muitos objetivos ainda são planejados a serem alcançados, como por exemplo, uma versão offline do LearningML.
(Tamashiro et al., 2021)	NI	Ad-Hoc	Interesse e desempenho dos estudantes	Dados coletados via notas consistindo de observações, reflexões e pequenas entrevistas com os estudantes	NI	2 estudantes	Apresentou um resultado de grande interesse dos estudantes envolvidos onde a ideia sobre a tecnologia e o tópico mostrado no curso foram alterados para uma visão melhor perante os estudantes. Se tem o objetivo de realizar esse curso com um número maior de estudantes.
(Tang et al., 2019) (Tang, 2019)	NI	Estudos de casos	Os alunos conseguiram entender conceitos básicos sobre ML, se os alunos entenderam como ML é utilizada no mundo real e se as ferramentas criadas para essa oficina foram efetivos em ajudar a ensinar ML para os alunos	Questionário pré e pós aplicação do curso	Análise qualitativa	23 estudantes foram oficialmente avaliados com os questionários	Se concluiu a partir dos dados coletados durante as oficinas que se obteve um sucesso em facilitar o aprendizado de ML utilizando do PIC para pessoas não especialistas
(Voulgari et al., 2021)	O algoritmo realiza a classificação das imagens decompondo a paleta de cores em número de cores primárias, com esses valores se treina uma árvore de decisão utilizando um algoritmo C4.5. Na segunda etapa do jogo se tem um foco maior no reinforcement learning o qual utiliza de um algoritmo Q-learning.	Estudos de casos	Facilidade em utilizar o jogo, diversão com o jogo e caso gostaria de ver o jogo em uma sala de aula	Questionário online, 19 perguntas para estudantes e 20 perguntas caso seja professor	NI	130 Estudantes e 17 professores	O jogo ArtBot foi bem recebido por estudantes e professores, apresentando facilidades para ensinar conceitos de ML.
(Wan et al., 2020)	Utilizou 16 expressões faciais em estilo de Emoji, com relação ao cluster, é realizada uma sobreposição do rosto que suporta comparações de pares entre pontos de dados em pares e em grupos arrastando e soltando o rosto um por cima do outro.	1 caso de estudo	Ganho de conhecimento, comportamento de aprendizado.	Questionário pré e pós o estudo	Estatística descritiva (média, mediana e alcance) e análise qualitativa	8 estudantes	Os resultados encontrados neste estudo mostram que a metáfora com o overlay de faces foi efetiva em ajudar os estudantes a aprender ML.

	Quando se faz o empilhamento cada face é semitransparente, com a soma de opacidades iguais a 100% se tem o resultado.						
(Zimmermann-Niefeld et al., 2019)	O modelo recebe as capturas de imagens ou vídeos, o <i>labeling</i> é realizado em 3 categorias, 'good', 'bad' e 'no action' (bom, ruim e sem ação)	Ad-Hoc	NI	Video e audio dos participantes durante a oficina e fichas de trabalho que os participantes preencheram durante a oficina	NI	5 participantes	O trabalho foi feito de maneira informal e ainda com a aplicação em protótipo. Mesmo sendo uma pesquisa inicial, se alcançou bons resultados com relação ao interesse dos participantes e a utilização de ML com os esportes.
(Zhu, 2019)	O sistema de identificação de objetos utiliza de uma detecção de uma única imagem multi caixas (Single Shot MultiBox Detector (SSD)) para desenhar as caixas.	Estudos de casos	Vontade de aprender sobre ML, seu conhecimento prévio, o nível de entendimento ganho após as aulas e o quanto as extensões ajudaram em aprender ML	Perguntas antes e depois das aulas em uma escala de 1 a 5	Análise quantitativa	14 participantes	O projeto realizou um total de 7 extensões, essas extensões foram utilizadas para diversas aplicações para auxiliar no ensino de ML, alcançando esse objetivo de ensino com sucesso de acordo com o curso aplicado.

3.4. Discussão

Os resultados analisados na busca mostram uma estabilidade no número de unidades instrucionais sendo publicadas com o alvo sendo os estudantes do ensino médio, tendo como comparativo os dados coletados por Marques, et al. (2020b). Nessa comparação, se observou um crescimento significativo de unidades instrucionais sendo aplicadas de maneira online ou remoto. Um dos possíveis motivos desse crescimento de utilização do ensino online é a pandemia de COVID-19 que afetou o mundo inteiro e o modo de ensino como um todo. Um dos exemplos de UI que foi desenvolvida visando essa mudança de ensino foi a desenvolvida por Rodríguez-García, et al. (2021). A mudança para modo online não foi propriamente planejada, podendo afetar os resultados encontrados pela pesquisa. Mesmo com as mudanças se observou um aprendizado dos alunos com relação ao tema.

Com a mudança para o modo de ensino remoto surgiram dúvidas sobre o nível de aprendizagem dos participantes, porém todas as unidades instrucionais estudadas apresentaram resultados positivos, alegando que o modo online não interferiu na aprendizagem de seus participantes, tendo um impacto positivo no entendimento dos alunos sobre o tema. Com a pesquisa realizada por Gresse von Wangenheim et al. (2020) se tem uma possível solução para o desafio sendo enfrentado, com a UI utilizando não só uma versão remota/online mas também tendo todo o material e plano de ensino para o modo presencial. Na versão online se tem se tem todas as apresentações, questionários e outras atividades interativas que podem auxiliar os alunos no ritmo próprio. Se percebe que tendo um curso online deixa o curso ainda mais acessível e flexível a todos os participantes, com isso os educadores podem optar pela maneira mais viável para todos os participantes e que os mesmo consigam realizar as atividades no tempo individual.

Analisando as unidades instrucionais pesquisadas, se identifica que a tendência para ensinar estudantes do ensino médio é utilizar da aprendizagem supervisionada e introduzir os participantes aos seus conceitos básicos e processos presentes em *ML*. Classificação e identificação são as tarefas mais utilizadas e visão computacional o mais utilizado para o domínio de aplicação. Em grande

maioria as unidades instrucionais analisadas utilizam do ciclo Use-Modifique-Crie em sua grande maioria no *Use*, com isso é possível perceber que é o tipo de tarefa que vem apresentando melhores resultados, sendo mais acessível ao estágio escolar aplicado.

Se identificou também, em diferentes unidades instrucionais, participantes que apresentavam um conhecimento prévio de programação tiveram um desempenho ainda melhor se comparado a participantes novatos na área. Se observou também um interesse do gênero feminino nas atividades, ainda que bem menor se comparado ao gênero masculino, em específico a pesquisa realizada por Tamashiro et al. (2021) em que se teve somente 2 participantes e ambas eram mulheres. O resultado encontrado nesta pesquisa foi de grande interesse e de curiosidade, por ser um estudo com um número reduzido de participantes, as mesmas puderam expor todas as dúvidas e curiosidades sobre o tema surpreendendo os pesquisadores.

É possível concluir que os cursos e oficinas voltados ao ensino de ML para estudantes do ensino médio com pouco ou nenhum conhecimento de programação e conhecimento anterior sobre o tema se beneficiaram ao serem introduzidos aos conceitos e aplicar as atividades presentes na UI em seu currículo. Unidades instrucionais que utilizaram do recurso online também tiveram uma vantagem ao serem mais facilmente aplicadas como cursos extras para participantes interessantes.

Um ponto observado é a quantidade de unidades instrucionais que utilizam do nível “use” para o ensino de suas atividades, onde estudantes em geral não possuem grande liberdade para alterações e construções de código de maneira independente.

Se observa, com os resultados da pesquisa, que ainda há falta de cursos em língua portuguesa. Também se observa a falta de cursos no nível mais avançado, sendo a maioria focado em um ensino mais básico e de fácil entendimento para os participantes. Outra observação é o fato de não ocorrer a interdisciplinaridade, então se tem a introdução do assunto utilizando de outras disciplinas, como a matemática e no caso deste trabalho a biologia, por exemplo.

Ameaças à validade da revisão da literatura. Algumas ameaças podem afetar a validade do estudo de mapeamento. Com a finalidade de minimizar essas possíveis ameaças foram identificadas estas potenciais ameaças e estratégias

foram organizadas para minimizar seus impactos. Mapeamentos sistemáticos podem se tornar tendenciosos, e resultados positivos podem ter mais chances de serem publicados do que os negativos. Porém, como o presente estudo não leva em consideração os resultados, sendo eles positivos ou negativos, pois possui o objetivo de levantar e analisar as características das unidades instrucionais, não é considerada uma ameaça neste caso. A possível omissão de estudos relevantes é outra ameaça que foi reduzida com a construção cuidadosa de *strings* de busca que seja a mais inclusiva possível, com uma variedade de termos e seus sinônimos pensados para serem o mais concisos possível. A predominância de artigos encontrados em língua inglesa pode ser associada ao fato da *search string* ser feita em inglês. Ameaças relacionadas a seleção de unidades instrucionais relevantes assim como a extração de dados foi mitigada com o desenvolvimento de regras de definição de inclusão e exclusão e critérios de qualidade. Todos os estudos foram analisados de maneira meticulosa para evitar essa possível ameaça.

4. Curso Cogufy (Nível Modifique)

Neste capítulo é apresentado o curso de Cogufy voltado ao público do Ensino Médio. O curso visa ensinar a classificação de imagens no nível de modifique seguindo o ciclo Use-Modifique-Crie (Lytle et al., 2019) supondo que o aluno já tem um conhecimento inicial de ML, considerando esse curso uma continuação a ser realizado depois de um curso inicial no nível *use* como por exemplo o curso ML4TEENS (Cardozo, 2022). Visa-se ensinar o desenvolvimento do modelo de ML com uma linguagem textual (Python) trazendo o aluno de um ambiente visual para um ambiente mais convencional de programação. O curso é projetado de forma interdisciplinar aplicado na área de biologia, especificamente voltado ao desenvolvimento de um modelo de classificação de 10 espécies de macrofungos no contexto da micologia presente na área de biologia.

4.1 Análise de Contexto

Primeiramente se analisa o cenário em que se tem o intuito de aplicar o curso, em busca de assegurar a coerência com a realidade.

Público alvo. O ensino médio brasileiro corresponde ao período de no mínimo três anos finalizando a educação básica. Em média, a faixa etária dos estudantes desse período é de 15 a 18 anos. Por ser a consolidação do que foi estudado durante o ensino fundamental se tem o objetivo de preparar o estudante para o mercado de trabalho, mesmo que em trabalhos não especializados. Observa que os estudantes desta faixa estudantil são fluentes em língua portuguesa e buscam conhecimento de outras línguas em principal inglês e espanhol (MEC, 2017).

Mesmo não estando explicitamente incluído no currículo de ensino médio atualmente, os alunos dessa faixa etária vem aumentando sua procura para cursos extracurriculares presenciais ou online buscando ter um conhecimento inicial de programação. Podendo ter adquirido essa competência de programação com linguagens baseadas em blocos e/ou também baseado em texto, como por exemplo, Python.

Assuma-se também que o estudante já tem um conhecimento básico de *ML* para classificação de imagens adquirido realizando cursos para iniciantes, como por exemplo, curso *Machine Learning* para Todos (Gresse von Wangenheim, 2020) ou

ML4TEENS (Cardozo, 2022). Estes cursos introduzem conceitos básicos de *Machine Learning*, como algoritmos de aprendizagem e fundamentos de redes neurais, bem como discutem as limitações e conceitos éticos relacionados a *Machine Learning*. Os cursos também ensinam a aplicação desses conceitos, levando o estudante a criar o seu primeiro modelo de *Machine Learning* de reconhecimento de imagens para separação de lixo (Gresse von Wangenheim, 2020) usando o ambiente visual Google Teachable Machine. O curso ML4TEENS também apresenta um ensino da aplicação de modelos de ML criando um projeto que realiza a classificação de imagens de árvores nativas selecionadas (Cardozo, 2022).

Assuma-se também que uma grande parte dos estudantes neste nível educacional tem acesso a computadores em casa ou realizam acesso a internet de alguma maneira, seja utilizando computadores ou celulares. Estudos mostram que o equipamento mais utilizado para acessar a Internet foi o celular chegando a 98,6% (IBGE, 2019). Com relação a utilização da Internet com o fim educacional se tem que 88,1% dos estudantes possuem acesso a Internet, porém dos estudantes sem acesso 95,9% são estudantes da rede pública de educação (IBGE, 2021). Considerando os números de acesso a Internet pode se inferir que uma grande parte do público alvo terá condições de realizar cursos online de maneira remota utilizando da Internet como meio de comunicação entre estudante e professor. Se evidenciou com o estudo do Instituto Federal de Santa Catarina em 2021 em que além do celular muitos dos estudantes possuem um *notebook* ou computador de mesa em casa, sendo 75% possuindo pelo menos um *notebook* ou mais e 30% possuindo ao menos um computador de mesa (Figura 30).

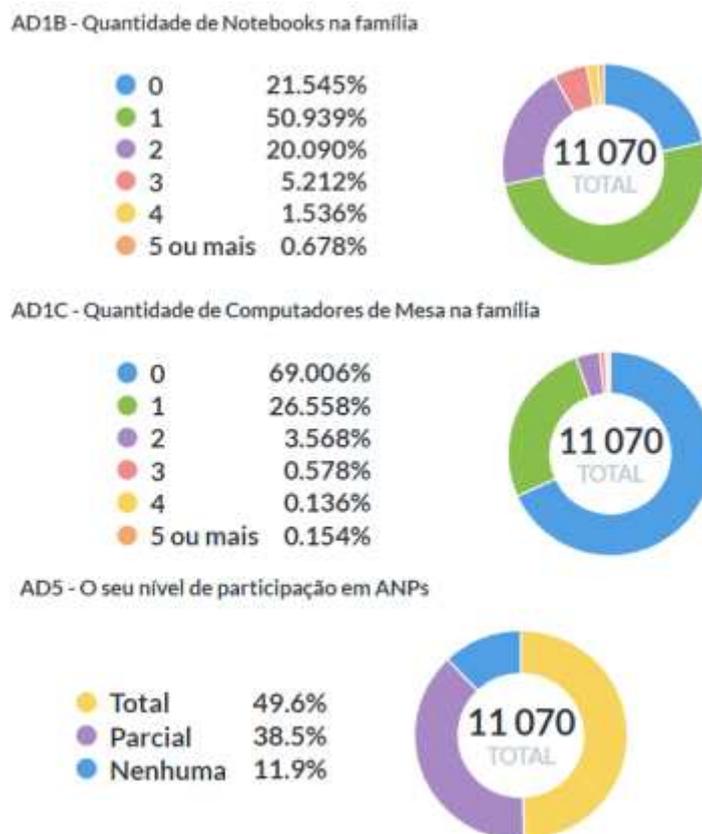


Figura 30 - Gráficos resultantes da pesquisa realizada pelo IFSC. (Fonte: IFSC, 2021)

O último gráfico da Figura 30 apresenta o índice de participação dos estudantes com atividades de ensino a distância, chegando a quase 90% de participação total ou parcial. Chegando a conclusão, nesse contexto identificado pelo IFSC, que a maioria dos alunos possuem a infraestrutura para participar de cursos on-line.

Contexto educacional. O Ensino Médio brasileiro tem como objetivo garantir aos estudantes a compreensão dos fundamentos científicos e tecnológicos de maneira a relacionar a teoria com a prática. Isso é alcançado por meio da junção de diferentes áreas do conhecimento: Linguagens, Matemática, Ciência da natureza e Ciências humanas (MEC, 2017). No currículo atual ainda não tem disciplinas de computação, porém existe a possibilidade de abordar conteúdos de pensamento computacional e de IA no ensino médio de forma interdisciplinar.

Interdisciplinaridade. Considerando que atualmente ainda não existem disciplinas específicas de computação no ensino médio, uma alternativa é o ensino de forma interdisciplinar. Nesse contexto, uma alternativa é a integração com a disciplina de biologia. No ensino médio nessa área de conhecimento tipicamente

são abordados temas com relação a microfungos e em alguns casos explicações de macrofungos e suas estruturas biológicas (MEC, 2017), conforme mais detalhado na seção 2.3.1.

Em termos de infraestrutura nas escolas brasileiras, estudos do Cetic TIC domicílios (2021) observa-se um crescimento de 40% das plataformas para atividades de ensino e aprendizagem nas escolas e laboratórios no Brasil, chegando a 62% em 2020. Também neste estudo se constatou que 82% das escolas brasileiras possuem acesso à internet, variando de acordo com a região, sendo a norte a pior com 51% e a centro-oeste a melhor com 98% (Cetic, 2021).

Os professores que promovem o uso pedagógico das Tecnologias Digitais de Interação e Comunicação (TDIC) (Bastos, 2020) devem ter macrocompetências:

- Compreender o contexto e a promoção do uso ético, seguro e responsável das TDIC;
- Pesquisar, utilizar e manipular ferramentas e Conteúdos Digitais;
- Desenvolver as habilidades dos estudantes em TDIC;
- Compreender a colaboração das TDIC na comunicação, processos avaliativos e na gestão da sala de aula;
- Comprometer-se com o aprendizado individual do aluno com TDIC;
- Garantir o alinhamento das TDIC com o Currículo e implementação dessas tecnologias em sala de aula;
- Comprometer-se com seu desenvolvimento pessoal e profissional com as TDIC.

Porém, atualmente no cenário brasileiro de educação observa-se que tem um número pequeno de professores da educação básica formado na área de computação que tem o nível desejado para realizar o ensino dessa área, pois atualmente depende somente do esforço individual do professor para prestar a essas competências listadas, considerando o fato de não ocorrer alterações nos currículos e Licenciaturas para a adição do ensino do TDIC (Bastos, 2020).

4.2 Objetivos de Aprendizagem

Os objetivos de aprendizagem do curso de Cogufy foram desenvolvidos com base nas competências de alfabetização de IA de Long e Magerko (2020) e as diretrizes

estabelecidas pelo AI4K12 (Touretzky et al., 2019). Também foram acrescentados objetivos com base em aprendizagem de algoritmos e programação conforme o guia de currículo CSTA (2016) o qual abrange tópicos de *user story* e casos de teste. Os objetivos foram selecionados considerando o contexto educacional e a base para se ter o conhecimento dos fundamentos sobre IA/ML, além de introduzir conceitos para a realização da tarefa de classificação de imagens desenvolvida no final do curso. Em relação a interdisciplinaridade foi incluído também o objetivo de ensinar Micologia considerando ser inserido na matéria de Biologia presente no currículo atual do ensino médio.

Tabela 08 - Objetivos de Aprendizagem

Categoria	Objetivo de Aprendizagem	Nível de aprendizagem com base na Taxonomia Bloom	Referências(s)
OA1. Conceitos básicos de ML e redes neurais	Utilizar de exemplos de aprendizagem em ML e comparar com o método de aprendizagem humana. Compreender a estrutura de uma rede neural e explicar como ela é composta, o caminho necessário para alcançar uma saída que realize a identificação de padrões.	Compreender	3-A-i K-2, 3-A-i 3-5, 3-A-i 6-8 (AIK12, 2020); 1, 2, 3, 5 (Long e Magerko, 2020) 3-A-ii 3-5, 3-B-i 6-8, 3-B-ii 3-5 (AIK12, 2020)
OA2. Gerenciamento de dados	Toda a preparação para alcançar um conjunto de dados usado para treinar um modelo de ML, levando em consideração o tamanho do conjunto de dados e a forma com que é realizado a rotulação desses dados coletados. Além disso, comentar sobre a qualidade do conjunto de dados.	Aplicar	3-C-ii 9-12 (AI4K12, 2020); 11, 12 (Long e Magerko, 2020); 1A-DA-05 (CSTA, 2017)
OA3. Treinamento de Modelo de ML	Realizar o treinamento de um modelo de ML para classificação/categorização utilizando de um algoritmo de aprendizagem supervisionada com dados reais e realizando os ajustes necessários para se obter o resultado desejado.	Aplicar	3-A-ii 9-12, 3-A-iii 9-12 (AI4K12, 2020) CSTA/Engenharia de SW
OA4. Avaliação e interpretação do desempenho do modelo ML	Realizar a análise e a interpretação do desempenho do modelo ML para classificação de imagens.	Aplicar	(Amershi et al., 2019)
OA5. Processo de ML	Compreender sobre Machine Learning e conseguir aplicar esse conhecimento superando os desafios envolvidos na área	Aplicar	3-A-iv 9-12 (AI4K12, 2020); 9 (Long e Magerko, 2020)
OA6. Ética e impactos de ML	Apresentar e discutir sobre as questões éticas envolvidas com ML e Identificar os prós e contras que existem ao utilizar de IA/ML e sobre as possíveis carreiras que essas tecnologias podem gerar ou extinguir	Compreender	3-C-iii 6-8 (AI4K12); 3A-AP-24 (CSTA, 2017); 13, 16 (Long e Magerko, 2020) 2-IC-21 (CSTA, 2017); 6 (Long e Magerko, 2020)
OA7. Algoritmos e programação	Criar programas de computador modificando de maneira inicial projetos	Aplicar	1B-AP-10, 1B-AP-12, 1B-AP-15, 2-DA-09,

	já existentes, realizando assim um aperfeiçoamento na área		3A-IC-25 (CSTA, 2017)
OA8. Biologia	Discutir sobre a importância da preservação e conservação da biodiversidade presente em nosso meio ambiente, realizando o reconhecimento de espécies de cogumelos presentes em Santa Catarina e região sul do país	Lembrar	(MEC, 2017)

Neste curso visa-se revisar conceitos básicos de ML que os estudantes tenham aprendido anteriormente por cursos extracurriculares como, por exemplo, Machine Learning para todos (Gresse von Wangenheim, 2020) ou ML4TEENS (Cardozo, 2022).

Com foco no nível “modifique” no ciclo use-modifique-crie (Lee, 2011) esse curso modifica o domínio de aplicação da classificação de lixo reciclável apresentado no curso Machine Learning para todos (Gresse von Wangenheim, 2020) ou de árvores nativas do curso ML4TEENS (Cardozo, 2022), para a classificação de macrofungos da região sul do Brasil.

Outra modificação é com relação a área de desenvolvimento de ao invés usar um ambiente visual introduzir uma linguagem textual de programação. Neste curso é utilizado programação textual Python em Jupyter Notebook no Google Colab. O desenvolvimento do modelo é realizado no ambiente Jupyter/Colab por ser um ambiente já estabelecido para criação de modelos educacionais, pela facilidade de armazenar os dados de maneira gratuita utilizando do Google Drive e do Colab utilizando do recurso da nuvem da Google para treinar os modelos de maneira gratuita.

Visa-se a exportação do modelo treinado no ambiente criado do Jupyter/Colab para posteriormente ser implantado dentro de um aplicativo móvel ou web site. Porém, a implantação do modelo está fora do escopo do curso atualmente. Assim, o curso vai cobrir todos os processos na criação de um modelo *ML* da análise de requisitos até a exportação para implantação futura.

Visa se a possibilidade da aplicação do curso nas aulas de biologia no ensino médio inclusive podendo oferecer uma forma educacional alternativa para fixar esse conteúdo sobre macrofungos e suas propriedades e características.

4.3 Design do Curso

Considerando o contexto e os objetivos de aprendizagem, foi projetado um curso intermediário para ensinar a tarefa de classificação de macrofungos utilizando a tecnologia de *Deep Learning*. A classificação de imagens é considerada uma das tarefas mais simples e com teor introdutório, por isso se escolheu dessa tarefa. A utilização de *Deep Learning* vem pelo fato de ser uma das tecnologias mais adotadas atualmente.

Antes da aplicação do curso, por ser um curso de nível intermediário, assume-se que os participantes já tenham um conhecimento básico prévio sobre o assunto de IA/ML.

Outro ponto a ser considerado é a utilização da linguagem Python, onde o curso ensina aos participantes as partes básicas em um nível suficiente para que o participante consiga realizar as tarefas projetadas no curso. Para isso é realizado aulas com o objetivo único de apresentar de maneira concisa os principais conceitos da linguagem de programação Python.

O curso será aplicado de maneira interdisciplinar, considerando a falta de uma área explícita sobre IA/ML atualmente na BNCC. Assim, o curso será aplicado juntamente a área de conhecimento de biologia referente ao tópico de biodiversidade, políticas ambientais e ecossistemas, tendo o foco maior com os macrofungos, contemplando a habilidade EF04CI06: Relacionar a participação de fungos e bactérias no processo de decomposição, reconhecendo a importância ambiental deste processo (MEC, 2017). Os alunos são estimulados a exercitarem o entendimento sobre as diferentes espécies de macrofungos presentes em Santa Catarina a fim de realizar um projeto de classificação de imagens de 10 espécies de cogumelos presentes na região sul do país. Espera-se que os estudantes consigam identificar os diferentes tipos de espécies nas imagens para que possam catalogá-las corretamente. Com isso se tem o objetivo de classificar imagens de 10 espécies de cogumelos da região sul, sendo os 10 fungos selecionados com a cooperação de professores do departamento de Biologia da UFSC. Esses fungos são apresentados e explicados no tópico 2.3 deste trabalho.

Levando em consideração a aplicação do curso para estudantes do ensino médio, se projetou um curso de curta ou média duração. O curso foi projetado como

curso on-line de maneira que seja possível ser aplicado de modo remoto em turma ou individualmente ou até de forma presencial utilizando o material interativo.

4.4 Plano de Ensino

O plano de ensino foi desenvolvido levando em consideração principalmente as diretrizes apresentadas pelo AI4K12 (Touretzky et al., 2019) e as competências de Long e Magerko (2020) que foram utilizadas como base para levantar os objetivos do curso. Todas as informações relevantes coletadas durante o mapeamento sistemático durante a pesquisa do estado da arte foram consideradas, como os itens mais comumente citados e utilizados alcançando os melhores resultados, moldando o plano de ensino do curso onde se tem o foco em classificação de imagens e suas etapas, desafios e processos.

Tabela 09 - Plano de ensino

Aula	Conteúdo	Objetivos de Aprendizagem	Método Instrucional	Material Instrucional	Avaliação
Aula 1: Motivação e revisão de conceitos básicos					
1 h	Revisar a motivação, importância e aplicações de IA no dia a dia	OA1, OA6	Material interativo, Demonstrações	Slides aula 1, Conjunto de dados	Quiz
	Revisar conceitos básicos de IA e ML.	OA1, OA6			
	Recapitular de maneira concisa os conceitos de ML, Tipos de abordagens e Etapas de ML	OA1, OA2, OA3, OA5, OA7			
	Resumo sobre o funcionamento de classificação de imagens	OA2, OA3			
	Fundamentos básicos sobre Redes neurais	OA1			
	Resumo sobre os Processo de ML	OA1, OA2, OA3, OA4, OA5, OA6			
Aula 1: Revisão de Micologia					
20min	Introdução do assunto de Macrofungos, mostrando as características básicas de cada filo que será	OA8	Material interativo, Demonstrações	Slide aula 1	Quiz

	utilizado para realizar a classificação				
Aula 2: Conceitos básicos de programação com Jupyter Notebook usando python no Google Colab					
1h30min	Apresentar o funcionamento e ensinar como utilizar do Jupyter e Google Colab	OA7	Jupyter Notebook, Google Colab	Slides aula 2, Jupyter Notebook e Colab	Quiz
	Conceitos básicos de python, como por exemplo, "hello world", já utilizando do ambiente Jupyter no Colab.	O7	Jupyter Notebook, Google Colab		
	Ensino dos conceitos de programação utilizados durante o curso, desde variáveis, condicionais a funções.	OA7, OA3	Material interativo, Demonstrações no notebook		
Aula 3: Desenvolvimento do modelo de ML, etapas: Análise de requisitos e Preparação de dados					
1h15min	Análise de requisitos	OA3	Material interativo, Vídeos, Demonstrações	Slides aula 3 , Conjunto de dados, Jupyter e Colab	Quiz
	Preparação dos dados. Utilizando dos conceitos aprendidos de ML e utilizando de Python	OA2, OA5, OA7	Apresentação, Vídeos, Atividades		
Aula 4: Classificação de Imagens de Fungos da Região Sul focado no estado de Santa Catarina					
2h	Realizar a apresentação do que será feito durante o projeto	OA7	Demonstrações e Slides	Slides aula 4, Jupyter Notebook e Google Colab	Quiz
	Construção do modelo utilizando da linguagem de programação Python.	OA1, OA2, OA3, OA5, OA8	Atividades no Jupyter Notebook com python		
	Treinamento, avaliação do loss Utilizando dos conceitos aprendidos de ML e utilizando de Python	OA3, OA4, OA8	Jupyter Notebook com python		
	Avaliação do desempenho	OA4	Slides, Jupyter Notebook		
Aula 5: Testes e Exportação do Modelo					
45min	Predição	OA6	Slides, Atividade Jupyter	Slides aula 5, Jupyter	Quiz

	Exportação do modelo	OA5, OA7	Slides e Jupyter	Notebook, Google Colab e Conjunto de dados	
--	----------------------	----------	------------------	--	--

Considerando um curso mais prático e com o nível quase intermediário, se prevê que os participantes já tenham uma noção sobre ML, podendo assim ter mais tempo programando e realizando as alterações no modelo para alcançar o objetivo de ter um aplicativo que realiza a classificação de macrofungos.

O desenvolvimento do modelo “Cogufy” tem como objetivo classificar imagens de macrofungos em 10 espécies de macrofungos e 1 categoria de não macrofungo para caso de teste com imagens sem cogumelos. Os objetivos do modelo de ML são documentados conforme a Tabela 10.

Tabela 10:Objetivos do modelo de ML para classificação de macrofungo

Objetivo do modelo de ML	
Tarefa	Classificar/predizer a espécie de macrofungo por meio de imagens de macrofungos da região sul brasileira.
Tipo da tarefa	<i>Single-label</i> classificação de imagens
Categorias	10 categorias de macrofungos e 1 de não-macrofungo Auricularia fuscosuccinea, Corioloopsis caperata,Cyptotrampa asprata, Ganoderma australe, Oudemansiella cubensis, Phallus indusiatus,Phillipsia domingensis, Pycnoporus sanguineus, Schyzophyllum commune,Trametes villosa e não-macrofungo
Contexto de uso	O modelo é utilizado como exemplo no contexto de ensino de na Educação Básica.
Público alvo	Cidadãos (15+ anos) Foco em alunos do Ensino Médio

Para o curso é utilizado um conjunto de dados já preparado pela UFSC para realizar o treinamento do modelo de *ML*. O conjunto é composto de 540 imagens divididas entre as categorias (Tabela 11) .

Tabela 11: Template conjunto de dados.

Conjunto de dados	
Fonte de dados	Conjunto de total de 540 imagens das 10 espécies de macrofungos da região sul brasileira

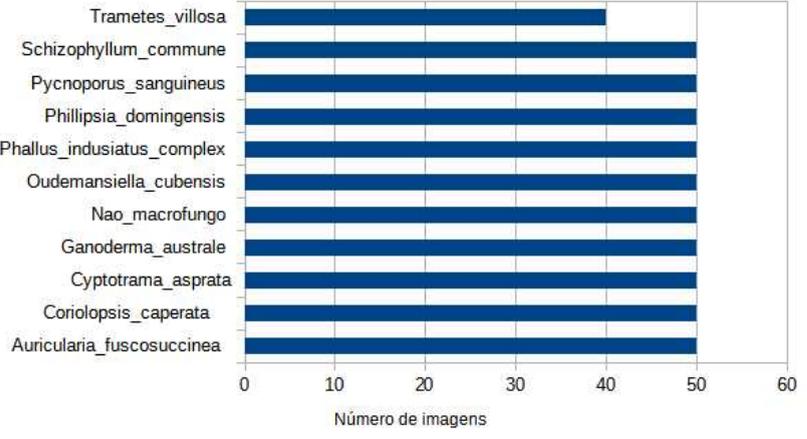
Conjunto de dados																										
gráfico com a distribuição de de imagens por categoria para treinamento	 <table border="1" data-bbox="571 286 1378 725"> <thead> <tr> <th>Categoria</th> <th>Número de imagens</th> </tr> </thead> <tbody> <tr><td>Trametes_villosa</td><td>40</td></tr> <tr><td>Schizophyllum_commune</td><td>50</td></tr> <tr><td>Pycnoporus_sanguineus</td><td>50</td></tr> <tr><td>Phillipsia_domingensis</td><td>50</td></tr> <tr><td>Phallus_indusiatus_complex</td><td>50</td></tr> <tr><td>Oudemansiella_cubensis</td><td>50</td></tr> <tr><td>Nao_macrofungo</td><td>50</td></tr> <tr><td>Ganoderma_australe</td><td>50</td></tr> <tr><td>Cyptotrama_asprata</td><td>50</td></tr> <tr><td>Coriopsis_caperata</td><td>50</td></tr> <tr><td>Auricularia_fuscosuccinea</td><td>50</td></tr> </tbody> </table>		Categoria	Número de imagens	Trametes_villosa	40	Schizophyllum_commune	50	Pycnoporus_sanguineus	50	Phillipsia_domingensis	50	Phallus_indusiatus_complex	50	Oudemansiella_cubensis	50	Nao_macrofungo	50	Ganoderma_australe	50	Cyptotrama_asprata	50	Coriopsis_caperata	50	Auricularia_fuscosuccinea	50
Categoria	Número de imagens																									
Trametes_villosa	40																									
Schizophyllum_commune	50																									
Pycnoporus_sanguineus	50																									
Phillipsia_domingensis	50																									
Phallus_indusiatus_complex	50																									
Oudemansiella_cubensis	50																									
Nao_macrofungo	50																									
Ganoderma_australe	50																									
Cyptotrama_asprata	50																									
Coriopsis_caperata	50																									
Auricularia_fuscosuccinea	50																									
Padronização das imagens	Formato: .jpg	Tamanho: 224x224 pixels																								
Rotulação de dados	Feito pelos pesquisadores do projeto MIND.Funga/UFSC																									



Figura 35: Exemplo de imagem coletada para a categoria de macrofungo, Trametes Villosa



Figura 36: Exemplo de imagem coletada para a categoria não-macrofungo

O treinamento do modelo foi realizado no ambiente Google Colab utilizando do Jupyter Notebook para realizar a programação, realizada em Python, o qual foi disponibilizado para o aluno.

Tabela 12: Exemplo de treinamento do modelo.

Treinamento																																																																																																																																																	
Quantidade de épocas	10																																																																																																																																																
Taxa de aprendizagem	0,0001																																																																																																																																																
Curva de loss	<p>The graph displays the loss curves for the training and validation sets over 60 epochs. The y-axis represents the loss value, ranging from 0.0 to 3.5. The x-axis represents the number of epochs, ranging from 0 to 60. The training loss (blue line) starts at 3.5 at epoch 0 and decreases steadily, reaching approximately 0.3 by epoch 60. The validation loss (orange line) starts at 0.8 at epoch 5 and decreases to approximately 0.05 by epoch 20, remaining stable thereafter.</p>																																																																																																																																																
Modelo	Resnet18																																																																																																																																																
Biblioteca utilizada	Fast.ai																																																																																																																																																
Avaliação do desempenho																																																																																																																																																	
Acurácia total	0.935185																																																																																																																																																
Matriz de confusão	<p>Confusion matrix</p> <table border="1"> <thead> <tr> <th>Actual \ Predicted</th> <th>Auricularia fuscusuccinea</th> <th>Coriolopsis caperata</th> <th>Cyptotrama asprata</th> <th>Ganoderma australe</th> <th>Nao macrofungo</th> <th>Oudemansiella cubensis</th> <th>Phallus indusiatus complex</th> <th>Phillipsia domingensis</th> <th>Pycnoporus sanguineus</th> <th>Schizophyllum commune</th> <th>Trametes villosa</th> </tr> </thead> <tbody> <tr> <th>Auricularia fuscusuccinea</th> <td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td> </tr> <tr> <th>Coriolopsis caperata</th> <td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td> </tr> <tr> <th>Cyptotrama asprata</th> <td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td> </tr> <tr> <th>Ganoderma australe</th> <td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td> </tr> <tr> <th>Nao macrofungo</th> <td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td> </tr> <tr> <th>Oudemansiella cubensis</th> <td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td> </tr> <tr> <th>Phallus indusiatus complex</th> <td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td> </tr> <tr> <th>Phillipsia domingensis</th> <td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td> </tr> <tr> <th>Pycnoporus sanguineus</th> <td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td> </tr> <tr> <th>Schizophyllum commune</th> <td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td> </tr> <tr> <th>Trametes villosa</th> <td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td> </tr> </tbody> </table>	Actual \ Predicted	Auricularia fuscusuccinea	Coriolopsis caperata	Cyptotrama asprata	Ganoderma australe	Nao macrofungo	Oudemansiella cubensis	Phallus indusiatus complex	Phillipsia domingensis	Pycnoporus sanguineus	Schizophyllum commune	Trametes villosa	Auricularia fuscusuccinea	0	0	0	0	0	0	0	0	0	0	0	Coriolopsis caperata	0	1	0	0	0	0	0	0	0	0	0	Cyptotrama asprata	0	0	1	0	0	0	0	0	0	0	0	Ganoderma australe	0	0	0	1	0	0	0	0	0	0	0	Nao macrofungo	0	0	0	0	1	0	0	0	0	0	0	Oudemansiella cubensis	0	0	0	0	0	1	0	0	0	0	0	Phallus indusiatus complex	0	0	0	0	0	0	1	0	0	0	0	Phillipsia domingensis	0	0	0	0	0	0	0	1	0	0	0	Pycnoporus sanguineus	0	0	0	0	0	0	0	0	1	0	0	Schizophyllum commune	0	0	0	0	0	0	0	0	0	1	0	Trametes villosa	0	0	0	0	0	0	0	0	0	0	1
Actual \ Predicted	Auricularia fuscusuccinea	Coriolopsis caperata	Cyptotrama asprata	Ganoderma australe	Nao macrofungo	Oudemansiella cubensis	Phallus indusiatus complex	Phillipsia domingensis	Pycnoporus sanguineus	Schizophyllum commune	Trametes villosa																																																																																																																																						
Auricularia fuscusuccinea	0	0	0	0	0	0	0	0	0	0	0																																																																																																																																						
Coriolopsis caperata	0	1	0	0	0	0	0	0	0	0	0																																																																																																																																						
Cyptotrama asprata	0	0	1	0	0	0	0	0	0	0	0																																																																																																																																						
Ganoderma australe	0	0	0	1	0	0	0	0	0	0	0																																																																																																																																						
Nao macrofungo	0	0	0	0	1	0	0	0	0	0	0																																																																																																																																						
Oudemansiella cubensis	0	0	0	0	0	1	0	0	0	0	0																																																																																																																																						
Phallus indusiatus complex	0	0	0	0	0	0	1	0	0	0	0																																																																																																																																						
Phillipsia domingensis	0	0	0	0	0	0	0	1	0	0	0																																																																																																																																						
Pycnoporus sanguineus	0	0	0	0	0	0	0	0	1	0	0																																																																																																																																						
Schizophyllum commune	0	0	0	0	0	0	0	0	0	1	0																																																																																																																																						
Trametes villosa	0	0	0	0	0	0	0	0	0	0	1																																																																																																																																						

4.5 Desenvolvimento do material didático

Com base no material didático de cursos já existentes da iniciativa Computação na Escola (FILHO, 2019) (FERREIRA, 2019) (GRESSE VON WANGENHEIM et al., 2020a) (OLIVEIRA, 2022) foi preparado e completado o material didático para o presente curso. De acordo com os métodos de ensino definidos foram desenvolvidos diversos materiais. O material didático é composto por slides interativos em formato h5p.

Tabela 13: Exemplo de slides do curso.

Slides	
<p>Aula 1</p> <p>FASE DE TREINAMENTO</p> <p>Fase de treinamento: A rede aprende ajustando-se para encontrar a resposta certa por conta própria, aumentando a acurácia de suas previsões</p> <p>Para fazer isso, a rede compara os resultados iniciais com uma resposta correta fornecida</p>	<p>Aula 2</p> <p>VARIÁVEIS</p> <p>Quando o programador for utilizar da variável o valor armazenado nela será utilizado</p> <p>No caso do armário ao acessar a gaveta de sapatos vamos ter acesso aos sapatos e suas propriedades. Em programação essas propriedades podem ser variadas porém seguem regras</p>
<p>Aula 3</p> <p>IMPORTAÇÃO DAS IMAGENS</p> <p>Após realizar seu login clique no seguinte botão para criar uma nova pasta chamada de dataset</p> <p>Clique aqui para abrir as opções de novos documentos</p> <p>Clique aqui para criar uma nova pasta e chame de Dataset</p>	<p>Aula 4</p> <p>TREINAMENTO DA REDE</p> <p>Após realizado o treinamento vai ser obtido um resultado parecido com o mostrado ao lado</p> <p>É possível dividir em três partes:</p> <ol style="list-style-type: none"> 1. Os ciclos de aprendizado da rede 2. O gráfico representando os ciclos e seu treinamento 3. Mensagem indicando qual ciclo teve a melhor perda válida (<i>valid loss</i>) <p>Cada parte apresenta dados importantes sobre o aprendizado realizado</p>
<p>Aula 5</p> <p>CRIAÇÃO DA IMAGEM</p> <p>Com a imagem que será analisada criada dentro do ambiente do Colab é realizado uma célula para garantir que o carregamento da imagem foi correta</p> <p>É criado a variável <code>img_externa</code> que vai apresentar um resultado</p> <pre>img_externa = widgets.Output() img_externa.clear_output() with img_externa: display(img_to_thumb(128,128)) img_externa</pre> <p>Por fim é mostrada a imagem que foi realizado o carregamento</p> <p>É limpo o resultado caso a célula já tenha sido executada antes</p>	

Foi criado também o *Notebook* no ambiente Google Colab onde já possui todas as células para que o aluno consiga acompanhar os slides disponibilizados do curso finalizando com um modelo de ML que realiza a classificação de 10 espécies de macrofungos.

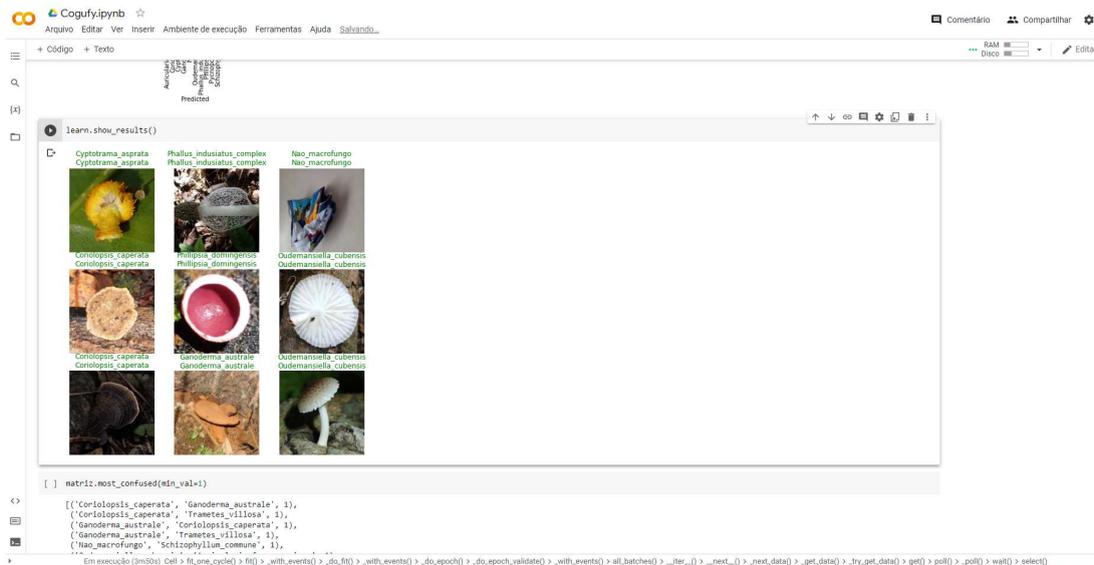


Figura 34 - Exemplos de células presentes no *Jupyter Notebook*.

Para realizar o treinamento do modelo é necessário além do Jupyter Notebook é preciso que sejam incluídas as imagens que são utilizadas, esse conjunto de imagens também já é disponibilizado para os alunos via Google Drive, onde nesse conjunto temos 540 imagens ao total onde somente a classe de Trametes Villosa possui 40 imagens e o restante das classes possuem 50 imagens. As imagens disponibilizadas para os alunos vem de um conjunto de dados da UFSC compilado pelo projeto MIND.Funga que encarecidamente liberou a utilização das mesmas nesse curso. Além das imagens de treinamento existe outra pasta com outras imagens para serem utilizadas nos testes de validação ao final da criação do modelo.

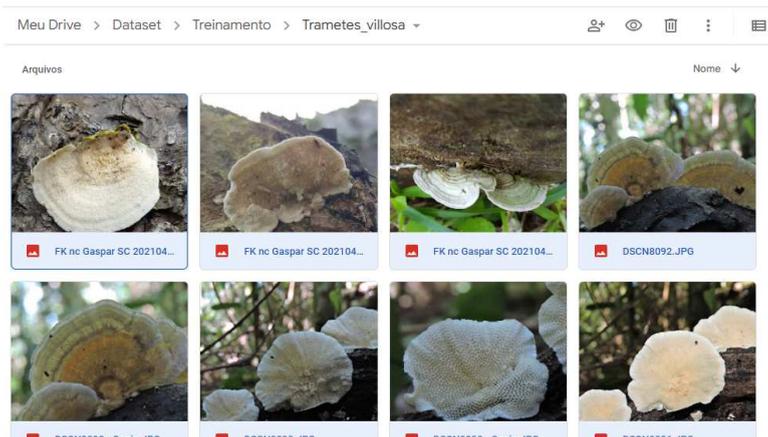
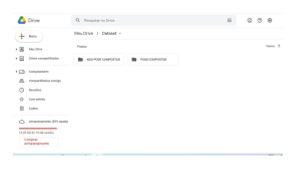


Figura 35 - Exemplos de imagens presente no conjunto de dados.

4.6 Ferramentas utilizadas no curso

O curso faz uso de ferramentas online, que são gratuitas e/ou de código aberto. A escolha dessas ferramentas tem como objetivo facilitar o processo de aprendizagem dos alunos.

Tabela 14: Ferramentas utilizadas no curso.

Ferramenta	Função	Exemplo	Link
Google Drive	Armazenamento de conjunto de dados		https://www.google.com/intl/pt-BR/drive/
Google Colab	Ambiente para realizar a programação e o treinamento do modelo de ML para o curso "Cogufy"		https://colab.research.google.com/

4.7 Avaliação da aprendizagem do aluno

A avaliação da aprendizagem é realizada em forma de quizzes interativos por meio do H5P (<https://h5p.org/>), e são distribuídos no material interativo (slides) ao longo de cada módulo do curso. A avaliação de desempenho dos artefatos de ML criado é realizada por meio do modelo de avaliação de desempenho proposto por Gresse von Wangenheim et al. (2021b).

As perguntas foram desenvolvidas de acordo com os objetivos de aprendizagem e são disponibilizadas após a exposição do conteúdo nos slides

anteriores. O *framework HP5* possui diversos tipos de avaliação, como Questões de Múltipla Escolha, em que o aluno pode selecionar mais de uma resposta correta dentre algumas opções, Escolha Única, Arrastar e Soltar para completar frases etc.

Aula 02 - Conceitos básicos

The slide features a green header with the title "VARIÁVEIS" and a cartoon character with a lightbulb. The main content includes a text prompt: "Utilizando do conhecimento ensinado até aqui responda Qual o resultado da operação mostrada na imagem?". Below this are four radio button options: 15, 6, 12, and 5. To the right is a code block with a play button icon, containing the following text: $a = 3$, $b = 2$, $a = 5$, $c = 13$, $a = 1$, and $a - b + c$. At the bottom, there is a progress bar showing "37 / 105" and navigation icons.

Figura 36 - Exemplo de quiz aplicado em h5p

Aula 3 - Análise de requisitos e preparação de dados

The slide has a green header with the title "ANÁLISE DE REQUISITOS" and a cartoon character. The main content asks: "Como deve ser a especificação do nosso projeto Cogufy?". Below is a form with two rows: "Tarefa" with the text "Classificar/predizer a espécie de <xxx> por meio de imagens de <xxx>" and "Tipo de tarefa" with the text "Single-label classificação de imagens". Below the form is the instruction "Preencha as palavras que faltam" and a sentence with two empty input boxes: "Classificar/predizer a espécie de [] por meio de imagens de []". A blue button labeled "Verificar resposta" is at the bottom. The footer shows a progress bar with "6 / 29" and navigation icons.

Figura 37 - Exemplo de quiz aplicado em h5p.

5. CONCLUSÃO

O presente estudo apresenta o desenvolvimento do curso Cogufy para o ensino de *ML* com foco em classificação de imagens de Macrofungos para alunos do Ensino Médio. Como parte do presente TCC foi sintetizada a fundamentação teórica centralizada nas estratégias instrucionais de *ML* no contexto do Ensino Médio. Analisando o estado da arte das unidades instrucionais semelhantes observou-se que muitas têm o foco no nível básico de ensino com o intuito de fazer uma apresentação inicial ao mundo de IA e *ML*.

Assim com base nas diretrizes de currículo foi desenvolvido o curso Cogufy buscando o objetivo de facilitar o ensino dos processos de *ML* para alunos do Ensino Médio de forma interativa e on-line, de maneira a apresentar a construção do modelo de *ML* com foco na linguagem textual Python e a biblioteca Fast.ai.

Utilizando de um ensino que os alunos são desafiados a desenvolver e acompanhar a construção de um modelo que realiza a classificação de imagens, com isso os alunos aprendem o passo a passo para a construção de um modelo de *ML* para que os alunos realizem seus próprios projetos futuramente.

Destaca-se também o caráter inédito do presente curso, avançando o estado da arte em que atualmente existem poucos cursos especificamente voltados a esse estágio educacional e menos ainda em português voltados a escolas brasileiras. A possível junção com a matéria de biologia com o intuito de ensinar sobre macrofungos, componente muito importante para a fauna e flora, também é de caráter inédito.

Como trabalhos futuros, sugere-se a aplicação do curso com usuários do público alvo com objetivo de refinar e identificar potenciais lacunas no material instrucional e aumentar a validade da avaliação, além da possibilidade para que sejam realizados novos cursos onde a categorização é de outro tipo ou de outras categorias. Visa-se também a aplicação do curso com o intuito de avaliar se o mesmo pode contribuir efetivamente para o ensino de *ML* no ensino médio e com base no *feedback* dessas aplicações também revisar e melhorar o curso.

REFERÊNCIAS

Amazonas, M. A. L. d. A. Biodiversidade de Macrofungos e Potencial de Uso Para O Desenvolvimento Sustentável. In: Anais da Semana do Estudante Universitário, 1. 2003, Colombo, PR, Brasil, 2003.

Banco de dados UFSC, MIND.Funga, disponível em: <https://mindfunga.ufsc.br/>. Acesso em: junho de 2020.

Basili, V., Caldiera, G., Rombach, H. D. The Goal Question Metric Approach. John Wiley & Sons, Encyclopedia of Software Engineering, 1994.

Bastos, T. B. M. C. e Boscaroli, C. 2020. Os Professores do Ensino Básico e as Tecnologias Digitais: Uma reflexão emergente e necessária em tempos de pandemia. ISSN: 2175-9235. Disponível em: <http://horizontes.sbc.org.br/index.php/2020/04/22/professores-do-ensino-basico-e-as-tecnologias-digitais/> Acesso em: fev. 2022.

Bemley J. L., Neural networks for precollege students, Proc. of the International Joint Conference on Neural Networks. IEEE, 4422–4427 vol.6, 1999.

Blackwell, M., The fungi: 1, 2, 3 ... 5.1 million species? American Journal of Botany, 98, 426–38, 2011.

Branch, R. Instructional design: The ADDIE approach. Springer-Verlag USA 2009, 2009.

Brasil. Base Nacional Comum Curricular. Brasília: MEC, 2017. Disponível em: http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518_versaofinal_sit e.pdf. Acesso em: outubro de 2020.

Burgsteiner, H., Kandlhofer, M., & Steinbauer, G. iRobot: Teaching the Basics of Artificial Intelligence in High Schools. In: Proc. of the AAAI Conference on Artificial Intelligence, Symposium on Educational Advances in Artificial Intelligence, Phoenix, AZ, USA, 2016.

Cardozo, J., Martins, R. M., Gresse Von Wangenheim, C. ML4Teens – Introduzindo Machine Learning no Ensino Médio. In: Anais do 30º WEI – Workshop sobre Educação em Computação, Niterói, Brasil, 2022.

Catalogue of Life, disponível em: <https://www.catalogueoflife.org/>. Acesso em: nov de 2021.

Columbus, L. Roundup Of Machine Learning Forecasts And Market Estimates, Forbes. 2020. Disponível em: <https://www.forbes.com/sites/louiscolombus/2020/01/19/roundup-of-machine-learning-forecasts-and-market-estimates-2020/?sh=3e0083d95c02>. Acesso em: fev. 2021.

Gardner-McCune, C., Touretzky, D., C., Martin, F., Seehorn, D. AI for K-12: Making Room for AI in K-12 CS Curricula. In: Proc. of the 50th ACM Technical Symposium on Computer Science Education . ACM, New York, NY, USA, 2019.

Gardner, Inc. Gartner Forecasts Worldwide Artificial Intelligence Software Market to Reach \$62 Billion in 2022. Disponível em: <https://www.gartner.com/en/newsroom/press-releases/2021-11-22-gartner-forecasts-worldwide-artificial-intelligence-software-market-to-reach-62-billion-in-2022>. Acesso em: outubro de 2022.

Gresse von Wangenheim, C., Marques, L. S., Hauck, J. C. R. Machine Learning for All - Introducing Machine Learning in K-12. SocArXiv Paper, version 3. 2020.

Haddaway, N.R. et al. The role of Google Scholar in evidence reviews and its applicability to grey literature searching. PloS one, 10(9). 2015.

Hitron, T., Orlev, Y., Wald I., Can Children Understand Machine Learning Concepts? The Effect of Uncovering Black Boxes, 2019 Conference on Human Factors in Computing Systems (CHI), 2019.

Howard, J. fastai v1 for PyTorch: Fast and accurate neural nets using modern best practices, 2018. Disponível em: <https://www.fast.ai/2018/10/02/fastai-ai/>. Acesso em: fevereiro de 2020.

IFSC. IFSC Analytics: Censo Escolar Covid-19. 2021. Disponível em: Acesso em: Jul de 2022.

Kahn, K., Three interactions between AI and education. In Elcock, E., & Michie, D. (Eds.), *Machine intelligence 8: Machine representations of knowledge*. Ellis Horwood Ltd. and John Wylie & Sons, 1977.

Kahn, K., & Winters, N.. Constructionism and AI: A history and possible futures. *British Journal of Educational Technology*, 52(3), 1130–1142. <https://doi.org/10.1111/bjet.13088>, 2021.

Kelleher, C., Pausch, R. Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers. *ACM Computing Surveys*, vol. 37, no 2, 2005.

Lecun, Y., Bengio, Y., Hinton, G. Deep learning for AI. *Commun. ACM* 64, 7 (July 2021), 58–65. <https://doi.org/10.1145/3448250>.

Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., Malyn-Smith, J., Werner, L. Computational thinking for youth in practice. *ACM Inroads*, 2(1), 2011.

Long, D., Magerko, B. What is AI Literacy? Competencies and Design Considerations. In: *Proc. of the CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, USA, 2020.

Lytle, N. Catete, V. Boulden, D. Dong, Y. Houchins, J. Milliken, A. Isvik, A. Bounajim, D. Wiebe, E. Use, Modify, Create: Comparing Computational Thinking Lesson Progressions for STEM Classes. Association for Computing Machinery, New York, NY, USA, 2019.

Marques, L., Gresse von Wangenheim, C., Hauck, J. C. R. Introdução à Machine Learning Desenvolvimento de uma unidade instrucional para ensino de ML na educação básica. Relatório Técnico INCoD/GQS.03.2020.P., INCoD/UFSC, Florianópolis/Brasil, 2020a.

Marques, L. Gresse von Wangenheim, C. Hauck, J. C. R. Teaching Machine Learning in School: A Systematic Mapping of the State of the Art. *Informatics in Education*, vol 19, no. 2. 2020b.

Marques, L., Gresse von Wangenheim, C., Hauck, J. C. R. Ensino de Machine Learning na Educação Básica: um Mapeamento Sistemático do Estado da Arte. In: Anais do Simpósio Brasileiro de Informática na Educação, Natal, Brasil, 2020c.

MEC. Base Nacional Comum Curricular, 2018. Disponível em <http://basenacionalcomum.mec.gov.br/abase/>. Acesso em: abril 2021.

ModelZoo. Disponível em: <https://modelzoo.co/about>, 2019. Acesso em: fev. 2020.

Neves, M. A., Baseia, I. G., Drechsler-Santos, E. R., Góes-Neto A. Guide to the Common Fungi of the Semiarid Region of Brazil, 2013.

O'Hara, K., Blank, D., and Marshall, J. Computational Notebooks for AI Education. Twenty-Eighth International Florida Artificial Intelligence Research Society Conference , Hollywood, Florida, USA, 2015.

Pesquisa TIC Domicílios.
https://cetic.br/media/docs/publicacoes/2/20211124201505/resumo_executivo_tic_do_micilios_2020.pdf. 2021.

Petersen, K., Vakkalanka, S. , Kuzniarz, L. Guidelines for conducting systematic mapping studies in software engineering: An update, Information and Software Technology, vol. 64, 2015.

Rosa, M. D'A., Mohr, A. Os Fungos na Escola: Análise dos Conteúdos de Micologia em Livros Didáticos do Ensino Fundamental de Florianópolis. Experiências em Ensino de Ciências, vol. 5, 2010.

Resnick, M. Mother's Day, Warrior Cats, and Digital Fluency: Stories from the Scratch Online Community. In: Proc. of the Constructionism Conference: Theory, Practice and Impact, Athens, Greece, 2012.

Royal Society. Machine learning: the power and promise of computers that learn by example.
<https://royalsociety.org/~media/policy/projects/machine-learning/publications/machine-learning-report.pdf>, 2017.

SBC. Diretrizes para ensino de Computação na Educação Básica, 2017. Disponível em:

<https://www.sbc.org.br/documentos-da-sbc/send/203-educacao-basica/1220-bncc-e-m-itinerario-informativo-computacao-2>. Acesso em: fev. 2021.

Scaico, P. D., Lima, A. A., Silva, J. B. B., Azevedo, S., Paiva, L. F., Raposo, E. H. S., Alencar, Y., Mendes, J. P. Programação no Ensino Médio: Uma Abordagem de Ensino Orientado ao Design com Scratch. In: Anais do XVII Workshop de Informática na Escola, Rio de Janeiro, Brasil, 2012.

Seehorn, D., Clayborn, L. CSTA K-12 CS Standards for All. In: Proc. of the ACM SIGCSE Technical Symposium on Computer Science Education, ACM, New York, NY, USA, 2017.

Silva, V. dos A., Santos, G. da S., Rodrigues, L. F. de S., Guimarães, S. S. M. O “Misterioso” Universo Dos Fungos e o Ensino de Ciências: Um Relato de Experiência. Experiências em Ensino de Ciências, vol 14, no. 1, 2019.

Torrey, L. Teaching Problem-Solving in Algorithms and AI. Proc. of the 3rd Symposium on Educational Advances in Artificial Intelligence, Toronto, Ontario, Canada, 2012.

Touretzky, D. S., Gardner-McCune, C., Martin, F., Seehorn, D. Envisioning AI for K-12: What Should Every Child Know about AI?. In: Proc. of the 33rd AAAI Conference on Artificial Intelligence, Honolulu, Hawaii, USA. 2019.

APÊNDICE A

Cogufy - Desenvolvimento de um Curso On-line para o Ensino de Machine Learning voltado a Classificação de Imagens de Macrofungos no Ensino Médio usando Python

Stefano Bergamini Poletto

Universidade Federal de Santa Catarina (UFSC)

Departamento de Informática e Estatística

Florianópolis SC Brasil

stefano.bp@grad.ufsc.br

Resumo. *Machine Learning* (ML) é atualmente uma das tecnologias com maior crescimento do mundo. Porém atualmente o conhecimento sobre esta tecnologia é pouco popularizado pelo fato da mesma não ser abordada na Educação Básica. Esse trabalho tem como proposta criar um curso on-line interdisciplinar para ensinar conceitos de ML no ensino médio utilizando da linguagem de programação Python. O curso visa ensinar a aplicação de conceitos de ML de forma prática na biologia voltada na classificação de espécies de Macrofungos. O curso será desenvolvido baseado na análise do estado da arte, adotando o design instrucional resultando na criação de um plano de ensino, o desenvolvimento de todo o material didático em forma de um curso on-line, infraestrutura técnica e a definição da avaliação da aprendizagem do aluno. Espera-se que os resultados deste trabalho contribuam para a popularização do conhecimento de Machine Learning e a geração de interesse nos jovens nas áreas de Computação e Biologia.

Abstract. Machine Learning (ML) is currently one of the fastest growing technologies in the world. However, currently knowledge about this technology is not well popularized because it is not addressed in Basic Education. This work proposes to create an interdisciplinary online course to teach ML concepts in high school using the Python programming language. The course aims to teach the application of ML concepts in a practical way in biology considering the classification of Macrofungi species. The course will be developed based on the analysis of the state of the art, adopting the instructional design generated in the creation of a teaching plan, the development of all didactic material in the form of an online course, technical infrastructure and the definition of the evaluation of the student learning. It is expected that the results of this work will contribute to the popularization of Machine Learning knowledge and the generation of interest in young people in the areas of Computing and Biology.

PALAVRAS CHAVES

Machine Learning, Deep Learning, Biologia, Micologia, Classificação de Imagens, Ensino médio, Jupyter Notebook, Google Colab

1. Introdução

O uso da Inteligência Artificial (IA) e especificamente soluções utilizando Machine Learning (ML) no dia-a-dia já é realidade, muitas vezes de maneira despercebida ela auxilia em tarefas cotidianas nos diferentes aparelhos e serviços utilizados pela população. Muitos já sabem sobre a existência da IA e seu grande crescimento nas últimas décadas, com estimativas alcançando um valor de \$62,5B em 2022, um crescimento de 21,3% de 2021 [Gartner, 2021]. Porém a ideia sobre IA nessa geração atual é muito focada no que os jovens consomem, como meios de entretenimento ou assistentes de pesquisas, por exemplo, videogames, assistentes pessoais [Siri, Google, etc.] e nas notícias, como, vitórias de computadores sobre oponentes humanos [Royal Society, 2017]. Mesmo com todas essas informações ao seu redor poucos sabem como funciona essa IA e a existência de ML [Royal Society, 2017]. Com isso se percebe a necessidade de abordar o tema de ML e computação já na educação básica para que as pessoas tenham um melhor conhecimento sobre as tecnologias que estão ao seu redor [Burgsteiner et al., 2016]. Desta maneira desmistificar a ideia que IA/ML é complexa e inacessível para quem não está no ramo de tecnologia a muito tempo [Wangenheim et al., 2020]. E, evitar que ocorra escassez de mão-de-obra qualificada para assumir tais postos de trabalho, já que a evolução tecnológica de ML/IA é cada vez mais rápida [Scaico et al., 2012]. Apresentar ML e computação desde a Educação Básica tem também o potencial de atrair grupos mais diversificados de pessoas para a área de computação [Kelleher et al., 2005].

Visando a formar criadores de soluções de TI [Resnick, 2012], existem propostas de diretrizes curriculares e unidades instrucionais para o ensino de IA/ML na Educação Básica [Marques et al., 2020a]. No Brasil a Sociedade Brasileira de Computação (SBC) apresenta diretrizes para a Educação Básica brasileira. Estas diretrizes têm um foco maior na aprendizagem de pensamento computacional e conceitos básicos necessários para entender as questões éticas envolvendo IA [SBC, 2017]. O Ministério da Educação (MEC) ressalta que grande parte das futuras profissões utilizarão, direta ou indiretamente, computação e tecnologias digitais, com objetivo de formar jovens e adultos com habilidades para utilizar das tecnologias digitais futuras se percebe a necessidade para abordar esse tema já na educação básica [Brasil, 2017].

Percebendo a importância de ML, alguns países como Austrália e Reino Unido começaram a inserir IA nos currículos de Educação Básica [Marques et al., 2020a]. No Brasil e em muitos outros países o estudo de IA e principalmente ML na Educação Básica é escassa e basicamente inexistente [Marques et al., 2020a].

Para realizar a inserção desses conteúdos sobre IA/ML na educação básica são criadas unidades instrucionais. Essas unidades instrucionais tipicamente visam ensinar conceitos básicos de ML, redes neurais, questões éticas e implicações de ML na sociedade. Geralmente se adota estratégias ativas visando levar o aluno a aprender e aplicar estes conceitos, por exemplo desenvolvendo modelos de deep learning (DL) e realizando a implantação dentro de jogos ou aplicativos. Frequentemente estas unidades instrucionais são realizadas em forma de cursos extracurriculares ou inserido de forma interdisciplinar. São utilizados diversos ambientes de desenvolvimento de modelos de ML, desde ferramentas visuais até ferramentas textuais como Jupyter notebooks [Gresse von Wangenheim, 2020]. Jupyter Notebooks são artefatos que podem conter código, texto, imagens e/ou equações [O'Hara et al., 2015]. Os Jupyter Notebooks podem ser lidos normalmente porém possuem a possibilidade de executar pedaços de código, como um programa de computador, tudo diretamente do web browser [O'Hara et al., 2015]. Esse tipo de ambiente vem mostrando bons resultados para ensinar diversos tópicos para estudantes.

Porém observa-se atualmente uma falta de cursos sobre ML no ensino médio brasileiro, utilizando linguagens de programação textual de maneira mais completa, abrangendo o processo de ML completo, desde a análise de requisitos até a sua implantação. Assim, este projeto tem o objetivo de criar um curso on-line utilizando Jupyter Notebook e Google Colab usando uma linguagem textual (Python) no estágio Modify para ensinar conceitos e o processo de ML aos alunos do ensino médio. Visa-se a criação de um curso interdisciplinar com conteúdo da área de Biologia, especificamente macrofungos/micologia, possibilitando uma integração deste conteúdo de ML no contexto de cursos do ensino médio.

3. Cogufy

O curso de Cogufy foi desenvolvido com foco no público do Ensino Médio, assim como a análise do contexto do público alvo e a definição do plano de ensino baseado na revisão bibliográfica. A partir desses resultados, o contexto dos estudantes foi analisado e o material do curso foi desenvolvido.

3.1. Análise de contexto

O ensino médio brasileiro corresponde ao período de no mínimo três anos finalizando a educação básica. Em média, a faixa etária dos estudantes desse período é de 15 a 18 anos. Por ser a consolidação do que foi estudado durante o ensino fundamental se tem o objetivo de preparar o estudante para o mercado de trabalho, mesmo que em trabalhos não especializados. Observa-se que os estudantes desta faixa etária são fluentes em língua portuguesa e buscam conhecimento de outras línguas em principal inglês e espanhol [MEC, 2017].

Mesmo não estando explicitamente incluído no currículo de ensino médio atualmente, os alunos dessa faixa etária vem aumentando sua procura para cursos extracurriculares presenciais ou online buscando ter um conhecimento inicial de programação. Podendo ter adquirido essa competência de programação com linguagens baseadas em blocos e/ou também baseado em texto, como por exemplo, Python.

Considerando os números de acesso a Internet pode se inferir que uma grande parte do público alvo terá condições de realizar cursos online de maneira remota utilizando da Internet como meio de comunicação entre estudante e professor. Se evidenciou com o estudo do Instituto Federal de Santa Catarina em 2021 em que além do celular muitos dos estudantes possuem um notebook ou computador de mesa em casa, sendo 75% possuindo pelo menos um notebook ou mais e 30% possuindo ao menos um computador de mesa.

3.2. Objetivos de Aprendizagem

Os objetivos de aprendizagem do curso de Cogufy foram desenvolvidos com base nas competências de alfabetização de IA de Long e Magerko [2020] e as diretrizes estabelecidas pelo AI4K12 [Touretzky et al., 2019]. Também foram acrescentados objetivos com base em aprendizagem de algoritmos e programação conforme o guia de currículo CSTA [2016] o qual abrange tópicos de user story e casos de teste. Os objetivos foram selecionados considerando o contexto educacional e a base para se ter o conhecimento dos fundamentos sobre IA/ML, além de introduzir conceitos para a realização da tarefa de classificação de imagens desenvolvida no final do curso. Em relação a interdisciplinaridade foi incluído também o objetivo de ensinar Micologia considerando ser inserido na matéria de Biologia presente no currículo atual do ensino médio.

Tabela 1. Objetivos de aprendizagem do curso Cogufy.

Categoria	Objetivo de Aprendizagem	Referências(s)
OA1. Conceitos básicos de ML e redes neurais	Utilizar de exemplos de aprendizagem em ML e comparar com o método de aprendizagem humana. Compreender a estrutura de uma rede neural e explicar como ela é composta, o caminho necessário para alcançar uma saída que realize a identificação de padrões.	3-A-i K-2, 3-A-i 3-5, 3-A-i 6-8 [AIK12, 2020]; 1, 2, 3, 5 [Long e Magerko, 2020] 3-A-ii 3-5, 3-B-i 6-8, 3-B-ii 3-5 [AIK12, 2020]
OA2. Gerenciamento de dados	Toda a preparação para alcançar um conjunto de dados usado para treinar um modelo de ML, levando em consideração o tamanho do conjunto de dados e a forma com que é realizado a rotulação desses dados coletados. Além disso, comentar sobre a qualidade do conjunto de dados.	3-C-ii 9-12 [AI4K12, 2020]; 11, 12 [Long e Magerko, 2020]; 1A-DA-05 [CSTA, 2017]
OA3. Treinamento de Modelo de ML	Realizar o treinamento de um modelo de ML para classificação/categorização utilizando de um algoritmo de aprendizagem supervisionada com dados reais e realizando os ajustes necessários para se obter o resultado desejado.	3-A-ii 9-12, 3-A-iii 9-12 [AI4K12, 2020] CSTA/Engenharia de SW
OA4. Avaliação e interpretação do desempenho do modelo ML	Realizar a análise e a interpretação do desempenho do modelo ML para classificação de imagens.	[Amershi et al., 2019]
OA5. Processo de ML	Compreender sobre Machine Learning e conseguir aplicar esse conhecimento superando os desafios envolvidos na área	3-A-iv 9-12 [AI4K12, 2020]; 9 [Long e Magerko, 2020]
OA6. Ética e impactos de ML	Apresentar e discutir sobre as questões éticas envolvidas com ML e Identificar os prós e contras que existem ao utilizar de IA/ML e sobre as possíveis carreiras que essas tecnologias podem gerar ou extinguir	3-C-iii 6-8 [AI4K12]; 3A-AP-24 [CSTA, 2017]; 13, 16 [Long e Magerko, 2020] 2-IC-21 [CSTA, 2017]; 6 [Long e Magerko, 2020]
OA7. Algoritmos e programação	Criar programas de computador modificando de maneira inicial projetos já existentes, realizando assim um aperfeiçoamento na área	1B-AP-10, 1B-AP-12, 1B-AP-15, 2-DA-09, 3A-IC-25 [CSTA, 2017]
OA8. Biologia	Discutir sobre a importância da preservação e conservação da biodiversidade presente em nosso meio ambiente, realizando o reconhecimento de espécies de cogumelos presentes em Santa Catarina e região sul do país	[MEC, 2017]

3.3. Design do Curso

Considerando o contexto e os objetivos de aprendizagem, foi projetado um curso intermediário para ensinar a tarefa de classificação de macrofungos utilizando a tecnologia de Deep Learning. A classificação de imagens é considerada uma das tarefas mais simples e com teor introdutório, por isso se escolheu dessa tarefa. A utilização de Deep Learning vem pelo fato de ser uma das tecnologias mais adotadas atualmente.

Antes da aplicação do curso, por ser um curso de nível intermediário, assume-se que os participantes já tenham um conhecimento básico prévio sobre o assunto de IA/ML.

Os alunos são estimulados a exercitarem o entendimento sobre as diferentes espécies de macrofungos presentes em Santa Catarina a fim de realizar um projeto de classificação de imagens de 10 espécies de cogumelos presentes na região sul do país. Espera-se que os estudantes consigam identificar os diferentes tipos de espécies nas imagens para que possam catalogá-las corretamente. Com isso se tem o objetivo de classificar imagens de 10 espécies de cogumelos da região sul, sendo os 10 fungos selecionados com a cooperação de professores do departamento de Biologia da UFSC.

3.4. Plano de Ensino

O plano de ensino foi desenvolvido levando em consideração principalmente as diretrizes apresentadas pelo AI4K12 (Touretzky et al., 2019) e as competências de Long e Magerko (2020) que foram utilizadas como base para levantar os objetivos do curso. Todas as informações relevantes coletadas durante o mapeamento sistemático durante a pesquisa do estado da arte foram consideradas, como os itens mais comumente citados e utilizados alcançando os melhores resultados, moldando o plano de ensino do curso onde se tem o foco em classificação de imagens e suas etapas, desafios e processos.

Tabela 02 - Plano de ensino do curso Cogufy

Aula	Conteúdo	Objetivos de Aprendizagem	Método Instrucional	Material Instrucional	Avaliação
Aula 1: Motivação e revisão de conceitos básicos					
1 h	Revisar a motivação, importância e aplicações de IA no dia a dia	OA1, OA6	Material interativo, Demonstrações	Slides aula 1, Conjunto de dados	Quiz
	Revisar conceitos básicos de IA e ML.	OA1, OA6			
	Recapitular de maneira concisa os conceitos de ML, Tipos de abordagens e Etapas de ML	OA1, OA2, OA3, OA5, OA7			
	Resumo sobre o funcionamento de classificação de imagens	OA2, OA3			
	Fundamentos básicos sobre Redes neurais	OA1			
	Resumo sobre os Processo de ML	OA1, OA2, OA3, OA4, OA5, OA6			
Aula 1: Revisão de Micologia					
20min	Introdução do assunto de Macrofungos, mostrando as características básicas de cada filo que será utilizado para realizar a classificação	OA8	Material interativo, Demonstrações	Slide aula 1	Quiz
Aula 2: Conceitos básicos de programação com Jupyter Notebook usando python no Google Colab					
1h30min	Apresentar o funcionamento e ensinar como utilizar do Jupyter e Google Colab	OA7	Jupyter Notebook, Google Colab	Slides aula 2, Jupyter Notebook e Colab	Quiz
	Conceitos básicos de python, como por exemplo, "hello world", já utilizando do ambiente Jupyter no Colab.	O7	Jupyter Notebook, Google Colab		
	Ensino dos conceitos de programação utilizados durante o curso, desde variáveis, condicionais a funções.	OA7, OA3	Material interativo, Demonstrações no notebook		
Aula 3: Desenvolvimento do modelo de ML, etapas: Análise de requisitos e Preparação de dados					
1h15min	Análise de requisitos	OA3	Material interativo, Vídeos, Demonstrações	Slides aula 3, Conjunto de dados, Jupyter e	Quiz

			es	Colab	
	Preparação dos dados. Utilizando dos conceitos aprendidos de ML e utilizando de Python	OA2, OA5, OA7	Apresentação, Vídeos, Atividades		
Aula 4: Classificação de Imagens de Fungos da Região Sul focado no estado de Santa Catarina					
2h	Realizar a apresentação do que será feito durante o projeto	OA7	Demonstrações e Slides	Slides aula 4, Jupyter Notebook e Google Colab	Quiz
	Construção do modelo utilizando da linguagem de programação Python.	OA1, OA2, OA3, OA5, OA8	Atividades no Jupyter Notebook com python		
	Treinamento, avaliação do loss Utilizando dos conceitos aprendidos de ML e utilizando de Python	OA3, OA4, OA8	Jupyter Notebook com python		
	Avaliação do desempenho	OA4	Slides, Jupyter Notebook		
Aula 5: Testes e Exportação do Modelo					
45min	Predição	OA6	Slides, Atividade Jupyter	Slides aula 5, Jupyter Notebook, Google Colab e Conjunto de dados	Quiz

3.5. Desenvolvimento do Material Didático

Com base no material didático de cursos já existentes da iniciativa Computação na Escola [FILHO, 2019] [FERREIRA, 2019] [GRESSE VON WANGENHEIM et al., 2020a] [OLIVEIRA, 2022] foi preparado e completado o material didático para o presente curso. De acordo com os métodos de ensino definidos foram desenvolvidos diversos materiais. O material didático é composto por slides interativos em formato h5p.

Tabela 03 - Slides do curso Cogufy

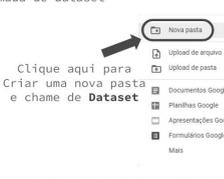
Slides	
Aula 1	Aula 2
<p>FASE DE TREINAMENTO</p> <p>Fase de treinamento: A rede aprende ajustando-se para encontrar a resposta certa por conta própria, aumentando a acurácia de suas previsões</p> <p>Para fazer isso, a rede compara os resultados iniciais com uma resposta correta fornecida</p> <p>Amarelo, Branco, Preto, Branco, Branco, Branco, Branco, Branco, Branco, Branco</p>	<p>VARIÁVEIS</p> <p>Quando o programador for utilizar da variável o valor armazenado nela será utilizado</p> <p>No caso do armário ao acessar a gaveta de sapatos vamos ter acesso aos sapatos e suas propriedades. Em programação essas propriedades podem ser variadas por meio de regras</p> <p><small>Copyright © Computação na Escola/Cogufy. Todos os Direitos Reservados. Proibida a distribuição e reprodução sem autorização prévia. Imagem: https://www.shutterstock.com/usuario/monocolor-illustration</small></p>
Aula 3	Aula 4

IMPORTAÇÃO DAS IMAGENS

Após realizar seu login clique no seguinte botão para criar uma nova pasta chamada de dataset



Clique aqui para abrir as opções de novos documentos



Clique aqui para Criar uma nova pasta e chame de Dataset

TREINAMENTO DA REDE

Após realizado o treinamento vai ser obtido um resultado parecido com o mostrado ao lado

É possível dividir em três partes:

1. Os ciclos de aprendizado da rede e seu treinamento
2. O gráfico representando os ciclos e
3. Mensagem indicando qual ciclo teve a melhor perda válida (*valid loss*)

Cada parte apresenta dados importantes sobre o aprendizado realizado

epoch	train_loss	train_acc	valid_loss	valid_acc	valid_f1
0	0.80172	0.18663	0.7014	0.17766	0.09101
1	1.00716	0.18721	0.63203	0.17021	0.09747
2	1.00716	0.24271	0.60723	0.17821	0.09747
3	0.70851	0.28624	0.58624	0.18751	0.10275
4	0.50381	0.32024	0.56746	0.19751	0.11171
5	0.40754	0.35023	0.54975	0.19751	0.11171
6	0.40427	0.38723	0.53248	0.19751	0.12026
7	0.31023	0.42723	0.51618	0.19666	0.12147
8	0.26023	0.45723	0.50123	0.19523	0.12147
9	0.22023	0.48723	0.49723	0.19523	0.12147

Aula 5

CRIAÇÃO DA IMAGEM

Com a imagem que será analisada criada dentro do ambiente do Colab é realizado uma célula para garantir que o carregamento da imagem foi correta

É criado a variável `img_externa` que vai apresentar um resultado

```
img_externa = widgets.Output()
with img_externa: display(img_to_thumb(120,120))
img_externa
```

É limpo o resultado caso a célula já tenha sido executada antes

Por fim é mostrada a imagem que foi realizado o carregamento

Foi criado também o Notebook no ambiente Google Colab onde já possui todas as células para que o aluno consiga acompanhar os slides disponibilizados do curso finalizando com um modelo de ML que realiza a classificação de 10 espécies de macrofungos.

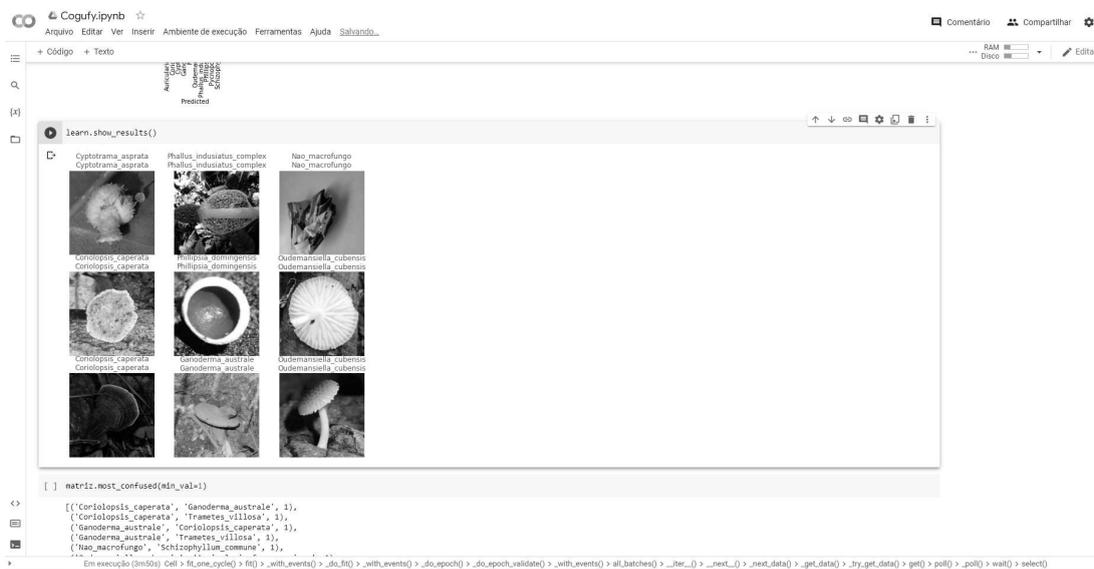


Figura 1: Exemplos de células presentes no *Jupyter Notebook*.

Para realizar o treinamento do modelo é necessário além do Jupyter Notebook é preciso que sejam incluídas as imagens que são utilizadas, esse conjunto de imagens também já é disponibilizado para os alunos via Google Drive, onde nesse conjunto temos 540 imagens ao total onde somente a classe de *Trametes Villosa* possui 40 imagens e o restante das classes possuem 50 imagens. As imagens disponibilizadas para os alunos vem de um conjunto de dados da UFSC compilado pelo projeto MIND.Funga que encarecidamente liberou a

utilização das mesmas nesse curso. Além das imagens de treinamento existe outra pasta com outras imagens para serem utilizadas nos testes de validação ao final da criação do modelo.

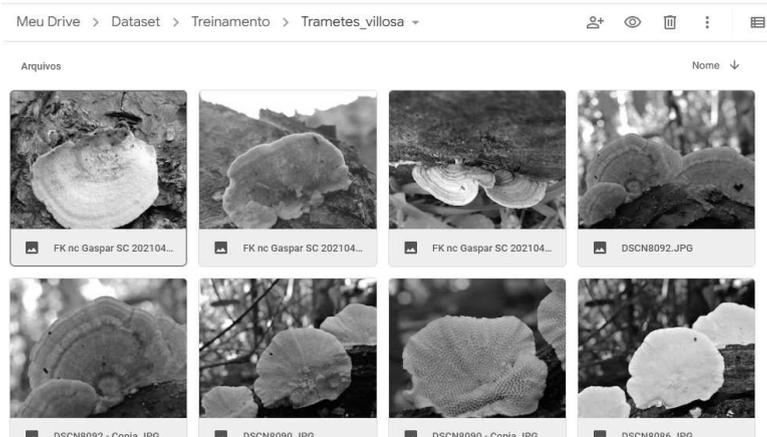


Figura 2: Exemplos de imagens presente no conjunto de dados.

4. Conclusão

O presente estudo apresenta o desenvolvimento do curso Cogufy para o ensino de ML com foco em classificação de imagens de Macrofungos para alunos do Ensino Médio. Como parte do presente TCC foi sintetizada a fundamentação teórica centralizada nas estratégias instrucionais de ML no contexto do Ensino Médio. Analisando o estado da arte das unidades instrucionais semelhantes observou-se que muitas têm o foco no nível básico de ensino com o intuito de fazer uma apresentação inicial ao mundo de IA e ML.

Assim com base nas diretrizes de currículo foi desenvolvido o curso Cogufy buscando o objetivo de facilitar o ensino dos processos de ML para alunos do Ensino Médio de forma interativa e on-line, de maneira a apresentar a construção do modelo de ML com foco na linguagem textual Python e a biblioteca Fast.ai.

Utilizando de um ensino que os alunos são desafiados a desenvolver e acompanhar a construção de um modelo que realiza a classificação de imagens, com isso os alunos aprendem o passo a passo para a construção de um modelo de ML para que os alunos realizem seus próprios projetos futuramente.

Destaca-se também o caráter inédito do presente curso, avançando o estado da arte em que atualmente existem poucos cursos especificamente voltados a esse estágio educacional e menos ainda em português voltados a escolas brasileiras. A possível junção com a matéria de biologia com o intuito de ensinar sobre macrofungos, componente muito importante para a fauna e flora, também é de caráter inédito.

Como trabalhos futuros, sugere-se a aplicação do curso com usuários do público alvo com objetivo de refinar e identificar potenciais lacunas no material instrucional e aumentar a validade da avaliação, além da possibilidade para que sejam realizados novos cursos onde a categorização é de outro tipo ou de outras categorias. Visa-se também a aplicação do curso com o intuito de avaliar se o mesmo pode contribuir efetivamente para o ensino de ML no ensino médio e com base no feedback dessas aplicações também revisar e melhorar o curso.

REFERÊNCIAS

Amazonas, M. A. L. d. A. Biodiversidade de Macrofungos e Potencial de Uso Para O Desenvolvimento Sustentável. In: Anais da Semana do Estudante Universitário, 1. 2003, Colombo, PR, Brasil, 2003.

Banco de dados UFSC, MIND.Funga, disponível em: <<https://mindfunga.ufsc.br/>>. Acesso em: junho de 2020.

Basili, V., Caldera, G., Rombach, H. D. The Goal Question Metric Approach. John Wiley & Sons, Encyclopedia of Software Engineering, 1994.

Bastos, T. B. M. C. e Boscarioli, C. 2020. Os Professores do Ensino Básico e as Tecnologias Digitais: Uma reflexão emergente e necessária em tempos de pandemia. ISSN: 2175-9235. Disponível em: <<http://horizontes.sbc.org.br/index.php/2020/04/22/professores-do-ensino-basico-e-as-tecnologias-digitais/>> Acesso em: fev. 2022.

Bemley J. L., Neural networks for precollege students, Proc. of the International Joint Conference on Neural Networks. IEEE, 4422–4427 vol.6, 1999.

Blackwell, M., The fungi: 1, 2, 3 ... 5.1 million species? American Journal of Botany, 98, 426–38, 2011.

Branch, R. Instructional design: The ADDIE approach. Springer-Verlag USA 2009, 2009.

Brasil. Base Nacional Comum Curricular. Brasília: MEC, 2017. Disponível em: http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518_-versaofinal_site.pdf. Acesso em: outubro de 2020.

Burgsteiner, H., Kandhofer, M., & Steinbauer, G. iRobot: Teaching the Basics of Artificial Intelligence in High Schools. In: Proc. of the AAAI Conference on Artificial Intelligence, Symposium on Educational Advances in Artificial Intelligence, Phoenix, AZ, USA, 2016.

Cardozo, J., Martins, R. M., Gresse Von Wangenheim, C. ML4Teens – Introduzindo Machine Learning no Ensino Médio. In: Anais do 30º WEI – Workshop sobre Educação em Computação, Niterói, Brasil, 2022.

Catalogue of Life, disponível em: <<https://www.catalogueoflife.org/>>. Acesso em: nov de 2021.

Columbus, L. Roundup Of Machine Learning Forecasts And Market Estimates, Forbes. 2020. Disponível em: <<https://www.forbes.com/sites/louiscolumbus/2020/01/19/roundup-of-machine-learning-forecasts-and-market-estimates-2020/?sh=3e0083d95c02>>. Acesso em: fev. 2021.

Gardner-McCune, C., Touretzky, D., C., Martin, F., Seehorn, D. AI for K-12: Making Room for AI in K-12 CS Curricula. In: Proc. of the 50th ACM Technical Symposium on Computer Science Education . ACM, New York, NY, USA, 2019.

Gardner, Inc. Gartner Forecasts Worldwide Artificial Intelligence Software Market to Reach \$62 Billion in 2022. Disponível em: <<https://www.gartner.com/en/newsroom/press-releases/2021-11-22-gartner-forecasts-worldwide-artificial-intelligence-software-market-to-reach-62-billion-in-2022>>. Acesso em: outubro de 2022.

Gresse von Wangenheim, C., Marques, L. S., Hauck, J. C. R. Machine Learning for All - Introducing Machine Learning in K-12. SocArXiv Paper, version 3. 2020.

Haddaway, N.R. et al. The role of Google Scholar in evidence reviews and its applicability to grey literature searching. PloS one, 10(9). 2015.

Hitron, T., Orlev, Y., Wald I., Can Children Understand Machine Learning Concepts? The Effect of Uncovering Black Boxes, 2019 Conference on Human Factors in Computing Systems (CHI), 2019.

Howard, J. fastai v1 for PyTorch: Fast and accurate neural nets using modern best practices, 2018. Disponível em: <<https://www.fast.ai/2018/10/02/fastai-ai/>>. Acesso em: fevereiro de 2020.

IFSC. IFSC Analytics: Censo Escolar Covid-19. 2021. Disponível em: Acesso em: Jul de 2022.

Kahn, K., Three interactions between AI and education. In Elcock, E., & Michie, D. (Eds.), Machine intelligence 8: Machine representations of knowledge. Ellis Horwood Ltd. and John Wylie & Sons, 1977.

Kahn, K., & Winters, N.. Constructionism and AI: A history and possible futures. British Journal of Educational Technology, 52(3), 1130–1142. <https://doi.org/10.1111/bjet.13088>, 2021.

Kelleher, C., Pausch, R. Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers. *ACM Computing Surveys*, vol. 37, no 2, 2005.

Lecun, Y., Bengio, Y., Hinton, G. Deep learning for AI. *Commun. ACM* 64, 7 (July 2021), 58–65.
<https://doi.org/10.1145/3448250>.

Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., Malyn-Smith, J., Werner, L. Computational thinking for youth in practice. *ACM Inroads*, 2(1), 2011.

Long, D., Magerko, B. What is AI Literacy? Competencies and Design Considerations. In: *Proc. of the CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, USA, 2020.

Lytle, N. Catete, V. Boulden, D. Dong, Y. Houchins, J. Milliken, A. Isvik, A. Bounajim, D. Wiebe, E. Use, Modify, Create: Comparing Computational Thinking Lesson Progressions for STEM Classes. *Association for Computing Machinery*, New York, NY, USA, 2019.

Marques, L., Gresse von Wangenheim, C., Hauck, J. C. R. Introdução à Machine Learning Desenvolvimento de uma unidade instrucional para ensino de ML na educação básica. Relatório Técnico INCoD/GQS.03.2020.P., INCoD/UFSC, Florianópolis/Brasil, 2020a.

Marques, L. Gresse von Wangenheim, C. Hauck, J. C. R. Teaching Machine Learning in School: A Systematic Mapping of the State of the Art. *Informatics in Education*, vol 19, no. 2. 2020b.

Marques, L., Gresse von Wangenheim, C., Hauck, J. C. R. Ensino de Machine Learning na Educação Básica: um Mapeamento Sistemático do Estado da Arte. In: *Anais do Simpósio Brasileiro de Informática na Educação*, Natal, Brasil, 2020c.

MEC. Base Nacional Comum Curricular, 2018. Disponível em <http://basenacionalcomum.mec.gov.br/abase/>. Acesso em: abril 2021.

ModelZoo. Disponível em: <https://modelzoo.co/about>, 2019. Acesso em: fev. 2020.

Neves, M. A., Baseia, I. G., Drechsler-Santos, E. R., Góes-Neto A. *Guide to the Common Fungi of the Semiarid Region of Brazil*, 2013.

O'Hara, K., Blank, D., and Marshall, J. *Computational Notebooks for AI Education*. Twenty-Eighth International Florida Artificial Intelligence Research Society Conference, Hollywood, Florida, USA, 2015.

Pesquisa TIC Domicílios, 2021. disponível em:
<https://cetic.br/media/docs/publicacoes/2/20211124201505/resumo_executivo_tic_domicilios_2020.pdf>.
Acesso em: jan 2022.

Petersen, K., Vakkalanka, S., Kuzniarz, L. Guidelines for conducting systematic mapping studies in software engineering: An update, *Information and Software Technology*, vol. 64, 2015.

Rosa, M. D'A., Mohr, A. Os Fungos na Escola: Análise dos Conteúdos de Micologia em Livros Didáticos do Ensino Fundamental de Florianópolis. *Experiências em Ensino de Ciências*, vol. 5, 2010.

Resnick, M. Mother's Day, Warrior Cats, and Digital Fluency: Stories from the Scratch Online Community. In: *Proc. of the Constructionism Conference: Theory, Practice and Impact*, Athens, Greece, 2012.

Royal Society. Machine learning: the power and promise of computers that learn by example. <<https://royalsociety.org/~media/policy/projects/machine-learning/publications/machine-learning-report.pdf>>., 2017.

SBC. Diretrizes para ensino de Computação na Educação Básica, 2017. Disponível em:
<<https://www.sbc.org.br/documentos-da-sbc/send/203-educacao-basica/1220-bncc-em-itinerario-informativo-computacao-2>>. Acesso em: fev. 2021.

Scaico, P. D., Lima, A. A., Silva, J. B. B., Azevedo, S., Paiva, L. F., Raposo, E. H. S., Alencar, Y., Mendes, J. P. Programação no Ensino Médio: Uma Abordagem de Ensino Orientado ao Design com Scratch. In: *Anais do XVII Workshop de Informática na Escola*, Rio de Janeiro, Brasil, 2012.

Seehorn, D., Clayborn, L. CSTA K-12 CS Standards for All. In: Proc. of the ACM SIGCSE Technical Symposium on Computer Science Education, ACM, New York, NY, USA, 2017.

Silva, V. dos A., Santos, G. da S., Rodrigues, L. F. de S., Guimarães, S. S. M. O “Misterioso” Universo Dos Fungos e o Ensino de Ciências: Um Relato de Experiência. Experiências em Ensino de Ciências, vol 14, no. 1, 2019.

Torrey, L. Teaching Problem-Solving in Algorithms and AI. Proc. of the 3rd Symposium on Educational Advances in Artificial Intelligence, Toronto, Ontario, Canada, 2012.

Touretzky, D. S., Gardner-McCune, C., Martin, F., Seehorn, D. Envisioning AI for K-12: What Should Every Child Know about AI?. In: Proc. of the 33rd AAAI Conference on Artificial Intelligence, Honolulu, Hawaii, USA. 2019.