

UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA

JACQUELINE CARDOZO

**Desenvolvimento de um Curso On-line
para o Ensino de *Machine Learning*
voltado à Classificação de Imagens no Ensino Médio**

FLORIANÓPOLIS

2022

JACQUELINE CARDOZO

**Desenvolvimento de um Curso On-line
para o Ensino de *Machine Learning*
voltado à Classificação de Imagens no Ensino Médio**

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

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

FLORIANÓPOLIS

2022

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Cardozo, Jacqueline

Desenvolvimento de um Curso On-line para o Ensino de
Machine Learning voltado à Classificação de Imagens no
Ensino Médio / Jacqueline Cardozo ; orientador, Christiane
Gresse von Wangenheim, 2022.

116 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Centro Tecnológico,
Graduação em Sistema de Informação, Florianópolis, 2022.

Inclui referências.

1. Sistema de Informação. 2. Inteligência Artificial. 3.
Machine Learning. 4. Ensino Médio. 5. Classificação de
Imagens. I. Gresse von Wangenheim, Christiane . II.
Universidade Federal de Santa Catarina. Graduação em
Sistema de Informação. III. Título.

Jacqueline Cardozo

**Desenvolvimento de um Curso On-line
para o Ensino de *Machine Learning*
voltado à Classificação de Imagens no Ensino Médio**

Trabalho de conclusão de curso submetido ao Departamento de Informática e Estatística da Universidade Federal de Santa Catarina para a obtenção do Grau de Bacharelado em Sistemas de Informação.

Florianópolis, 20 de março de 2022.

Orientadora:

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

Banca Examinadora:

Prof. Dr. Jean Carlo R. Hauck
Avaliador Universidade Federal de Santa Catarina

Prof. Dr. Fernando Pacheco
Avaliador Instituto Federal de Santa Catarina

AGRADECIMENTOS

Deixo aqui meus agradecimentos:

A minha família e amigos, por proporcionar o melhor ambiente para meus estudos e fornecerem apoio em todos os momentos necessários.

A Professora Christiane pela orientação e gentileza durante todo o período de desenvolvimento deste trabalho.

Ao doutorando Ramon Mayor Martins por toda a ajuda na elaboração do conteúdo do curso.

Aos membros da banca que se mostraram bastante acessíveis durante o desenvolvimento do trabalho.

Aos participantes do painel de especialistas, Ramon Mayor Martins, Nathalia da Cruz Alves, Jean Hauck, Lúcia Helena Martins Pacheco, Ana Maria Bender Seidenfuss das Neves, Marcelo Fernando Rauber, Gustavo Salvador, Mirian França, Bianca Cristina da Silva Almeida, Leonardo Martins Gonçalves, Pedro Fiaschi e Christiane Gresse von Wangenheim.

A minha terapeuta Joana, por ter me ajudado a conseguir finalizar este trabalho.

Ao meu professor Rodrigo de Souza Mota que não está mais entre nós, mas que as suas ideias ainda vivem nos alunos que inspirou.

RESUMO

O entendimento sobre *Machine Learning* (ML) tem se tornado cada vez mais importante para as sociedades que almejam formar estudantes na Educação Básica com as competências necessárias para enfrentar os impactos das transformações digitais. Porém, existem poucas iniciativas de ensino de *ML* voltadas para esse público. Assim, o presente trabalho visa ajudar os alunos da Educação Básica, mais especificamente estudantes do Ensino Médio, a entenderem fundamentos de *Machine Learning* por meio de um curso on-line. A partir da análise do estado da arte, foi criado sistematicamente um curso on-line adotando a abordagem de design instrucional. A unidade instrucional desenvolvida poderá ser utilizada por educadores e professores do Ensino Médio e desta forma contribuir na popularização dessa área de conhecimento neste estágio educacional.

Palavras-chave: Inteligência Artificial, *Machine Learning*, *Deep Learning*, Unidade Instrucional, Ensino Médio, Classificação de Imagens.

SUMÁRIO

1 INTRODUÇÃO	9
1.1 Contextualização	9
1.2 Objetivos	11
Objetivo geral	11
Objetivos específicos	11
1.3 Metodologia de Pesquisa e Trabalho	11
1.4 Estrutura do Documento	14
2 REFERENCIAL TEÓRICO	15
2.1 Ensino de Machine Learning	15
2.1.1 Contexto Nacional	16
2.1.2 Iniciativas Internacionais	19
2.1.2.1 Natureza da Aprendizagem	20
2.1.2.2 Redes Neurais	21
2.1.2.3 Conjuntos de Dados	22
2.1.3 Estratégias Instrucionais	24
2.2 Processos e Ferramentas de ML	26
2.2.1 Machine Learning	26
2.2.2 Classificação de Imagens com Deep Learning	29
2.2.3 Arquiteturas, Frameworks e Ferramentas de ML	30
3 ESTADO DA ARTE	32
3.1 Definição do Protocolo de Revisão	32
3.2 Execução da Busca	34
3.3 Análise dos Resultados	35
3.3.1 Quais UIs existem?	35
3.3.2 Quais conceitos de Machine Learning são ensinados nessas UIs?	37
3.3.3 Quais são as características instrucionais dessas UIs?	42
3.4 Discussão	45

4 CURSO ML4TEENS	48
4.1 Análise de Contexto	48
4.2 Objetivos de Aprendizagem	51
4.3 Design do Curso	52
4.4 Plano de Ensino	55
4.5 Material Didático	56
4.6 Avaliação da Aprendizagem	60
5 AVALIAÇÃO DO CURSO	64
5.1 Definição da Avaliação	64
5.2 Execução da Avaliação	66
5.3 Análise da Avaliação	66
5.3.1 Utilidade do curso ML4Teens	67
5.3.2 Adequação funcional	67
5.3.3 Usabilidade	68
5.3.4 Experiência de aprendizagem	69
5.3.4 Pontos fortes e fracos	70
5.4 Discussão	72
6 CONCLUSÃO	74
REFERÊNCIAS	75
APÊNDICES	82

LISTA DE FIGURAS

- Figura 1:** Os 3 eixos de conhecimentos da área de Computação (SBC, 2017).
- Figura 2:** Conceitos de Computação no Ensino Médio (SBC, 2017).
- Figura 3:** Etapas do processo de ML.
- Figura 4:** Número de UIs publicadas por ano com foco no Ensino Médio.
- Figura 5:** Número de UIs que introduzem cada abordagem.
- Figura 6:** Porcentagem dos tipos de tarefas ML abordadas nas UIs.
- Figura 7:** Número de UIs que abordam determinado domínio da aplicação.
- Figura 8:** Porcentagem de tipos de UIs que utilizam as etapas Use-Modify-Create.
- Figura 9:** Porcentagem de utilização dos diferentes métodos instrucionais.
- Figura 10:** Nuvem de palavras e termos mais associados ao objetivo das UIs.
- Figura 11:** quantidade de notebooks na família de acordo com o Censo Escolar Covid-19 (IFSC, 2021).
- Figura 12:** quantidade de computadores de mesa na família (IFSC, 2021).
- Figura 13:** participação dos alunos em atividades não presenciais (IFSC, 2021).
- Figura 14:** Ilustração com exemplos de uso de IA no cotidiano.
- Figura 15:** Conceito ilustrado de uma rede neural convolucional.
- Figura 16:** Vídeo sobre como usar o Visual Jupyter para classificar imagens.
- Figura 17:** Vídeo introduzindo o conceito de *ML*.
- Figura 18:** Vídeo introduzindo o conceito de redes neurais.
- Figura 19:** Zêd personifica o popular meme baseado no anime *Brave of the Sun Fighbird* (Romano, 2018).
- Figura 20:** Exemplo de quiz do tipo “arraste e solte”.
- Figura 21:** Exemplo de quiz do tipo “escolha única”.
- Figura 22:** Exemplo do recurso externo “Rápido, desenhe!”.
- Figura 23:** Contagem de participantes em relação a sua atuação.
- Figura 24:** Resultados sobre a utilidade do curso.
- Figura 25:** Resultados em relação a adequação funcional e usabilidade do curso.
- Figura 26:** Resultados em relação a usabilidade do curso.
- Figura 27:** Resposta sobre o conteúdo ser suficiente para realizar as atividades.

LISTA DE TABELAS

- Tabela 1:** Objetivos sobre a Natureza da Aprendizagem da *Big Idea #3* (CSTA, 2017).
- Tabela 2:** Objetivos sobre Redes Neurais da *Big Idea #3* (CSTA, 2017).
- Tabela 3:** Objetivos sobre *conjuntos de dados* da *Big Idea #3* (CSTA, 2017).
- Tabela 4:** Competências relacionadas ao ensino de ML (Long & Magerko, 2020).
- Tabela 5:** Search string para cada fonte.
- Tabela 6:** Número de artigos identificados por repositório e por fase de seleção.
- Tabela 7:** Unidades Instrucionais de ensino de ML.
- Tabela 8:** Análise do conteúdo das Unidades Instrucionais.
- Tabela 9:** Características do conteúdo relacionadas a *ML*.
- Tabela 10:** Características das Unidades Instrucionais.
- Tabela 11:** Objetivos de aprendizagem do curso ML4Teens.
- Tabela 12:** Espécies de árvores nativas selecionadas.
- Tabela 13:** Plano de ensino do curso ML4Teens.
- Tabela 14:** Plano de avaliação referente aos objetivos e competências.
- Tabela 15:** Quizzes inseridos como atividades interativas utilizando a ferramenta H5P (<https://h5p.org/>).
- Tabela 16:** Rubrica de Salvador (2021) alinhada aos objetivos do curso ML4Teens.
- Tabela 17:** Visão geral da decomposição das características de qualidade e operacionalização do curso.
- Tabela 18:** Respostas abertas sobre elementos desnecessários no curso.
- Tabela 19:** Respostas sobre os pontos positivos do curso.
- Tabela 20:** Respostas sobre os pontos negativos do curso.

LISTA DE ABREVIATURAS

ACM - Association for Computing Machinery

BNCC - Base Nacional Comum Curricular

CNN - Convolutional Neural Network

CSTA - Computer Science Teachers Association, Associação de Professores de Ciências da Computação

GQM - Goal Question Metric

IA - Inteligência Artificial

MEC - Ministério da Educação

ML - Machine Learning, “Aprendizado de Máquina” na língua inglesa

SBC - Sociedade Brasileira de Computação

TI - Tecnologia da Informação

UI - Unidade Instrucional

ANP - Atividades Não-Presenciais

1 INTRODUÇÃO

1.1 Contextualização

A Inteligência Artificial (IA) tem cada vez mais se tornado parte do cotidiano das pessoas. Assim como a Educação Básica abrange as áreas clássicas do conhecimento, como linguagens e matemática, a alfabetização em IA e computação tem se tornado uma grande questão a ser indagada (Burgsteiner et al., 2016). Essa discussão do que os alunos da Educação Básica devem aprender sobre computação tem impulsionado o desenvolvimento de diversas diretrizes curriculares (Marques et al., 2020a). A Sociedade Brasileira de Computação (SBC) enfatiza a importância do ensino de pensamento computacional e IA para alunos do Ensino Médio, provendo a base necessária para entender grandes questões éticas envolvidas em IA (SBC, 2017). O Ministério da Educação (MEC) enfatiza que grande parte das futuras profissões envolverá, direta ou indiretamente, computação e tecnologias digitais, fazendo-se necessário abordar o tema na Educação Básica a fim de formar cidadãos com os valores e habilidades necessárias para enfrentar os impactos das transformações digitais (MEC, 2017).

No mundo todo, a população jovem se desenvolve em meio à tecnologia digital computadores, celulares e outros dispositivos - que permitem que eles usem estes dispositivos para se comunicar, se divertir, realizar pesquisas on-line, etc. (Resnick, 2012). Embora tenham muito contato com o mundo digital, essa geração de jovens, em grande parte, não possui real fluência no uso de tais tecnologias para desenvolver seus próprios artefatos (Resnick, 2012), assim como usar essas ferramentas virtuais para entender e transformar o mundo (SBC, 2017).

Para os jovens se inspirarem na importância do desenvolvimento dessas tecnologias, abrindo caminho para que se tornem criadores de soluções de TI, é necessário ter conhecimento na área de Inteligência Artificial, que é o campo em que se estuda o desenvolvimento de máquinas inteligentes (Royal Society, 2017). Uma das principais áreas de IA atualmente é o *Machine Learning (ML)*, tecnologia que permite a um sistema aprender por meio de exemplos, dados e experiência,

desempenhando tarefas específicas, dada uma coleção específica de entradas de dados (Royal Society, 2017).

O entendimento de IA/ML pode encorajar estudantes a considerar carreiras nas áreas de computação e promover uma preparação dos alunos para o ensino superior (Marques et al., 2020a). Assim, é importante que estudantes encontrem esses elementos nos seus currículos escolares o mais cedo possível e dentro da Educação Básica. O Ensino Médio é um ótimo momento para que os estudantes entrem em contato com IA, pois estão maduros o suficiente para encarar problemas mais complexos (Estevez et al., 2019).

Enquanto iniciativas de ensino de IA estão sendo propagados no Ensino Médio dos Estados Unidos, China e Grã Bretanha (Marques et al., 2020a), no Brasil há ainda poucos trabalhos sobre a aplicação de cursos de programação com o objetivo de ensinar programação fora do ensino superior (Scaico et al., 2010)(Marques et al., 2020b). O ensino de IA é sugerido apenas como itinerário formativo - parte flexível do currículo ou atividade optativa - no novo currículo de Educação Básica brasileiro definido pela Base Nacional Comum Curricular (MEC, 2017). A SBC defende que o ensino de IA provê a base necessária para uma discussão sobre a interação Homem-Máquina e suas implicações éticas (SBC, 2017), sendo condizente com a previsão da BNCC quanto à necessidade de preparar os alunos para enfrentar os impactos das transformações digitais.

Assim, observa-se atualmente uma falta de unidades instrucionais para ensinar *Machine Learning* no Ensino Médio em escolas brasileiras. Portanto, este estudo visa conduzir a análise e síntese da literatura sobre as características instrucionais necessárias para o ensino de *ML* para alunos do Ensino Médio. O estudo levará em consideração os processos e ferramentas de desenvolvimento e como resultado, será criado um curso on-line visando o ensino de conceitos básicos de *ML* no Ensino Médio.

1.2 Objetivos

Objetivo geral

O objetivo geral deste projeto é o desenvolvimento de um curso on-line para ensinar *Machine Learning* em escolas brasileiras de Ensino Médio. O curso é alinhado aos currículos de referência do ensino de computação da ACM/CSTA (Seehorn e Clayborn, 2017) e SBC (SBC, 2017) e especificamente de Inteligência Artificial (Gardner-McCune et al., 2019)(Judd, 2020). O curso será projetado de forma que permita a sua adoção no formato interdisciplinar, inserido em conteúdo programático do Ensino Médio.

Assim, este projeto visa 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 a serem utilizados durante a sua execução. Como resultado, será criado um plano de ensino, todo o material didático em forma de um curso 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 a ensino de *ML* no Ensino Médio e os processos e ferramentas de desenvolvimento de seus modelos;
- O2: Analisar o contexto referente aos alunos, ambiente e currículos de referência e definir os objetivos de aprendizagem;
- O3: Levantar o estado da arte em relação a unidades instrucionais semelhantes e atualizar o conteúdo com foco em estudantes do Ensino Médio.
- O4: Definir o *design* instrucional do curso (seleção e sequenciamento do conteúdo, estratégias de ensino, etc.);
- O5: Desenvolver o material didático para o curso on-line;
- O6: Avaliar o curso desenvolvido por meio de um painel de especialistas.

1.3 Metodologia de Pesquisa e Trabalho

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 ensino de *ML* no Ensino Médio e sobre o processo e ferramentas de desenvolvimento de *ML*, 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 de processo e ferramentas de *ML*.

Etapa 2 - Análise do contexto: análise do contexto da unidade instrucional a ser desenvolvida, identificando características e restrições em relação ao público alvo, à infraestrutura e ao contexto educacional. É seguido o modelo de design instrucional ADDIE (Branch, 2009), representando uma orientação para a construção de ferramentas eficazes de treinamento.

Atividade 2.1: Analisar o contexto em termos de necessidades e objetivos de aprendizagem alinhado às diretrizes dos currículos;

Atividade 2.2: Analisar o contexto em termos de perfil dos aprendizes e instrutores;

Atividade 2.3: Analisar o contexto em termos de ambiente em escolas brasileiras.

Etapa 3 - Levantamento do estado da arte: levantamento sobre trabalhos existentes relacionados a área do projeto. É realizado um estudo de mapeamento seguindo um processo proposto por Petersen et al. (2015) para identificar e analisar unidades instrucionais/estratégias de ensino, atualmente sendo utilizadas e voltadas ao ensino de computação em escolas.

Atividade 3.1: Definir o protocolo de revisão;

Atividade 3.2: Executar a busca;

Atividade 3.3: Extrair e analisar as informações.

Etapa 4 - 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, 2009).

Atividade 4.1: Definir os objetivos de aprendizagem e sequenciar o conteúdo da unidade instrucional e definir uma estratégia instrucional, criando o plano de ensino.

Atividade 4.2: Definir a avaliação da aprendizagem do aluno;

Etapa 5 - Desenvolvimento da unidade instrucional: nesta etapa será realizado o desenvolvimento de todo o material didático para a aplicação da unidade instrucional.

Atividade 5.1: Desenvolver o material de atividades *on-line*;

Atividade 5.2: Desenvolver vídeos interativos;

Etapa 6 - Avaliação do curso: a avaliação será realizada em forma de um painel de especialistas, definindo os objetivos de avaliação usando o Goal-Question-Metric (GQM) (Basili et al., 1994), a execução com representantes do perfil alvo, incluindo alunos e professores do Ensino Médio/Superior e analisando os resultados.

Atividade 6.1: Definição da avaliação;

Atividade 6.2: Execução do curso e coleta de dados;

Atividade 6.3: Analisar e interpretar os dados.

1.4 Estrutura do Documento

No capítulo 2, os conceitos que fundamentam o presente estudo serão apresentados: as abordagens nacionais e internacionais do ensino de ML e IA para alunos do Ensino Médio, estratégias instrucionais e metodologias de ensino, tais quais o Use-Modifique-Crie, processos e ferramentas de ML - conceituando os tipos de aprendizagem e etapas de ML. No capítulo 3, são apresentados os resultados da análise e síntese do mapeamento sistemático do estado da arte em relação ao ensino de ML para o Ensino Médio. O capítulo 4 introduz a análise do contexto dos perfil dos professores, estudantes e ambiente escolar e traz a proposta de plano de ensino e modelos de avaliação. O capítulo 5 sintetiza os resultados da avaliação do curso realizada por um painel de especialistas. O capítulo 6 apresenta as conclusões e sugestões para trabalhos futuros.

2 REFERENCIAL TEÓRICO

Neste capítulo, serão discutidas iniciativas de ensino de *ML* para alunos do Ensino Médio e de que forma elas se organizam.

2.1 Ensino de *Machine Learning*

Embora existam muitos recursos online para os jovens começarem a aprender sobre *Machine Learning*, a maioria desses jovens necessitam de conhecimentos sólidos de matemática ou programação, o que pode desencorajar este público. Também há escassez de literatura que indique as formas de ensinar *ML* (Evangelista et al., 2018) e de uma metodologia unificada no ensino de inteligência artificial para alunos do Ensino Médio (Estevez et al., 2020). Isso acontece devido à natureza da multidisciplinaridade que envolve *ML*.

Sem o entendimento básicos dos conceitos de *ML*, a inteligência artificial ao qual temos contato todos os dias pode parecer ameaçadora ou como uma caixa preta mágica (Evangelista et al., 2018). No entanto, não é necessário ter formação acadêmica para aprender sobre *ML* (Howard et al., 2020) e há evidências que comprovam que é possível despertar o interesse no tópico sem que haja a necessidade de se pautar a jargões matemáticos (Evangelista et al., 2018).

Existem iniciativas que defendem o ensino de *ML* para alunos do Ensino Médio, como é o exemplo da Inglaterra. O parlamento britânico reconheceu a necessidade de se ensinar os tópicos de Inteligência Artificial começando do Ensino Médio. Os alunos dessa faixa etária já possuem a base de conhecimentos como estatística e probabilidade para ter entendimento básico do funcionamento de um modelo de *ML* (Estevez et al., 2020).

No Brasil, não existem iniciativas concretas em relação ao ensino de *ML* para jovens. No entanto, há alguns indicativos de sugestão de ensino de IA para alunos da escola básica, como é o caso da Base Nacional Comum Curricular (BNCC) e da Sociedade Brasileira de Computação (SBC). Existem também diretrizes mais completas para o ensino de *ML* de iniciativas internacionais como a AI4K12.

2.1.1 Contexto Nacional

A Base Nacional Comum Curricular (BNCC) é um documento de caráter normativo elaborado e distribuído pelo Ministério da Educação (MEC), cujo objetivo é nortear os currículos das redes de ensino de Educação Básica do Brasil. Definido de acordo com a Lei de Diretrizes e Bases da Educação Nacional (LDB, Lei nº 9.394/1996), o registro traz propostas pedagógicas que tange, entre outros temas, o ensino de princípios da computação para alunos do Ensino Médio.

O documento cita a necessidade de ensinar os jovens aprendizes sobre o uso de tecnologias digitais de informação e comunicação de maneira que tenham os meios necessários para atuar em uma sociedade em constante mudança e que estejam preparados para atuar em profissões que ainda não existem. O documento também salienta que grande parte das futuras profissões envolverá conhecimento de computação, direta ou indiretamente.

Em relação ao ensino de tecnologias digitais, é sugerido o ensino de pensamento computacional em todos os estágios de ensino na Educação Básica. Para o Ensino Médio, a BNCC sugere currículos que estimulem os jovens a implementar soluções para identificar, analisar, modelar e solucionar problemas complexos de diversas áreas do cotidiano, utilizando-se de conhecimentos do pensamento computacional e raciocínio lógico. No entanto, o documento não cita especificamente o ensino de *ML*, mas sugere que Inteligência Artificial faça parte de um itinerário formativo. Estes itinerários são parte flexível da organização curricular, que podem ser implementados em cada unidade de ensino considerando o contexto local e de acordo com a disponibilidade.

O documento da BNCC, ao sugerir o ensino de Inteligência Artificial como parte flexível do currículo do Ensino Médio por meio de implementação e análise de soluções computacionais.

A Sociedade Brasileira de Computação (SBC) tem como uma de suas funções incentivar o ensino de computação no Brasil. Para isso, elaborou um documento com diretrizes curriculares referentes à formação em Computação que sugere que os fundamentos da área sejam ensinados na profundidade compatível com a Educação Básica.

O ensino de Computação segue um modelo de três eixos que englobam áreas de conhecimento: Pensamento Computacional, Cultura Digital e Mundo Digital. Cada eixo, por sua vez, possui uma série de habilidades e competências referentes a cada área que constituem o processo curricular sugerido pela SBC para a Educação Básica.

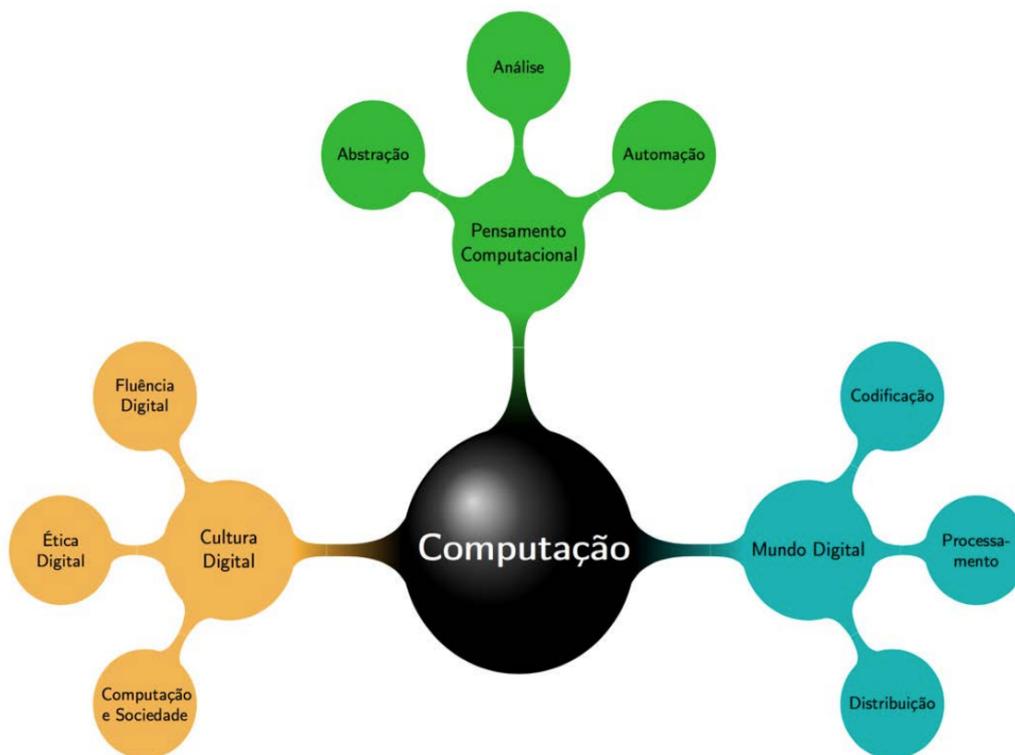


Figura 1: Os 3 eixos de conhecimentos da área de Computação (SBC, 2017).

É de muita relevância, segundo o documento, que os eixos de Pensamento Computacional e Mundo Digital sejam trabalhados entre os alunos do Ensino Médio. A experiência de elaboração de projetos de modelagem computacional fornece meios aos alunos de buscarem soluções para problemas de diversas áreas e entendimento de como a Computação pode ser usada não só no ambiente de trabalho, mas nas tarefas do cotidiano.

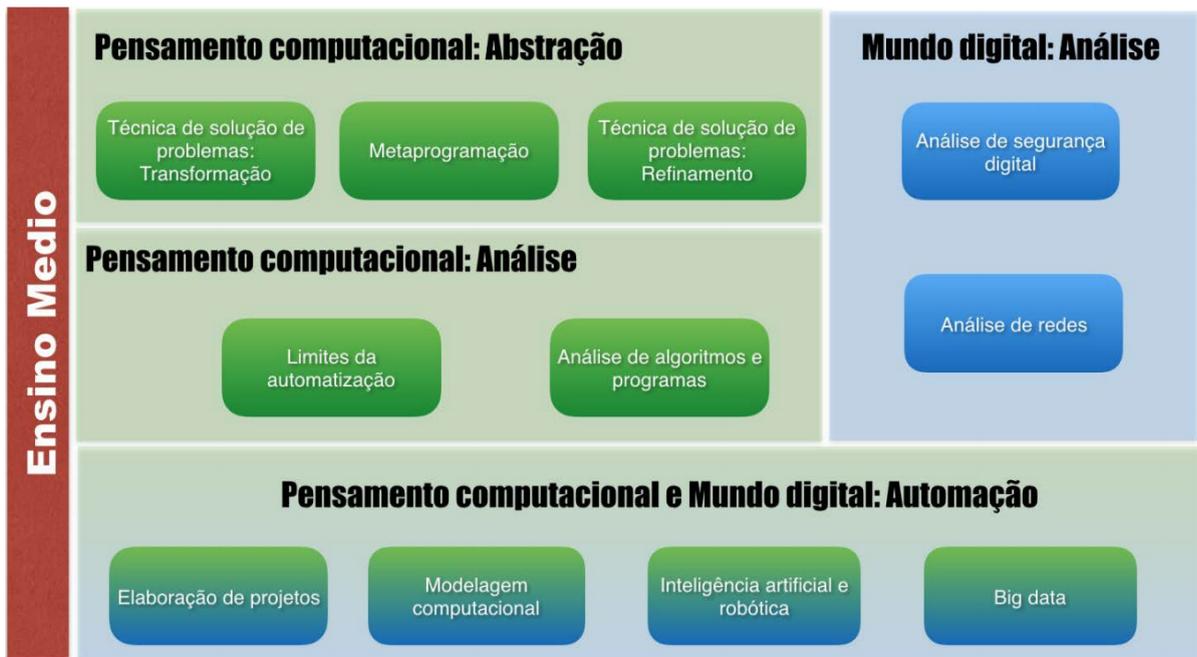


Figura 2: Conceitos de Computação no Ensino Médio (SBC, 2017).

No nível do Ensino Médio, as diretrizes da SBC também incluem o ensino de fundamentos de Inteligência Artificial e robótica com o objetivo de que os alunos entendam as limitações do uso de computação e questões éticas envolvidas no uso de Inteligência Artificial. No entanto, o documento não cita especificamente o ensino de *ML* para alunos da Educação Básica.

2.1.2 Iniciativas Internacionais

O AI4K12 (Touretzky et al. 2019) é uma iniciativa educacional que tem como objetivo propor diretrizes nacionais para o ensino de IA, disponibilização de recursos que facilitam a instrução de IA e fomentar uma comunidade de praticantes, pesquisadores, recursos e ferramentas focados no ensino de IA para alunos da Educação Básica (CSTA, 2017). AI4K12 é amparada pela Associação para o Avanço da Inteligência Artificial (AAAI) e Associação de Professores de Ciência da Computação (CSTA), financiada pela Fundação Nacional da Ciência dos Estados Unidos.

As diretrizes do AI4K12 (Touretzky et al. 2019) são divididas em cinco grande ideais: Percepção, Representação e Raciocínio, Aprendizagem, Interação Natural e Impacto Social. Essas diretrizes têm como objetivo servir de estrutura para o desenvolvimento de currículos que abranjam os conceitos de IA, tendo em vista o conhecimento essencial que deve ser abordado em cada tema, assim como as habilidades necessárias levando em consideração a faixa etária dos alunos.

A iniciativa disponibiliza por meio da idéia relacionada à Aprendizagem (*Big Idea #3*), diretrizes concretas sobre a progressão pedagógica do ensino de *Machine Learning* para alunos do Ensino Médio, que incluem os objetivos de aprendizagem e compreensão a longo prazo para cada tópico abordado no ensino de *ML*. É importante salientar que será considerado que os alunos do Ensino Médio que participarão do curso desenvolvido no presente estudo não terão nenhum conhecimento prévio de *ML*. A partir do uso das diretrizes da AI4K12 para o desenvolvimento curricular deste curso, é necessário que alguns conceitos sugeridos para faixas etárias mais jovens também sejam abordados para alunos do Ensino Médio.

Dentro do foco do presente trabalho, a *Big Idea #3* é constituída por três partes: Natureza de Aprendizagem, Redes Neurais e Conjuntos de Dados. Cada área é dividida em conceitos a serem abordados no currículo do curso. Para cada conceito, são descritos os objetivos de aprendizagem e compreensão a longo prazo que apresentam, respectivamente, o que os estudantes podem fazer e o que aprenderam tendo concluído o conteúdo de determinado conceito.

2.1.2.1 Natureza da Aprendizagem

Conceito	Faixa	Objetivo de Aprendizagem
Humano Vs. Máquina	K-2	Demonstrar como treinar um computador para reconhecer algo.
	3-5	Treinar um modelo de classificação usando aprendizado de máquina e, em seguida, examinar a precisão do modelo em novas entradas.
	6-8	Treinar e avaliar um modelo de classificação ou previsão usando aprendizado de máquina em um conjunto de dados tabular.
	9-12	Definição de algoritmos de aprendizado supervisionado, não supervisionado e reforçado e exemplos de aprendizado humano similar a cada algoritmo.
Observando padrões em dados	K-2	Identificar padrões em dados rotulados e determinar os recursos que prevêm rótulos.
	3-5	Entender como o aprendizado supervisionado identifica padrões em dados rotulados.
	6-8	Entender como a aprendizagem não supervisionada identifica padrões em dados não rotulados.
	9-12	Compreender como ML constrói um raciocínio para classificação ou previsão através do ajuste dos parâmetros do raciocínio da máquina (sua representação interna).
Treinando um modelo	K-2	Demonstrar como treinar um computador para reconhecer algo.
	3-5	Treinar um modelo de classificação usando aprendizado de máquina e, em seguida, examinar a precisão do modelo em novas entradas.
	6-8	Treinar e avaliar um modelo de classificação ou previsão usando aprendizado de máquina em um conjunto de dados tabular.
	9-12	Usar um algoritmo de aprendizado supervisionado ou não supervisionado para treinar um modelo em dados do mundo real e, em seguida, avaliar os resultados.
Construindo vs. usando um raciocinador	K-2	N/A
	3-5	Demonstrar como os dados de treinamento são rotulados ao usar uma ferramenta de aprendizado de máquina.
	6-8	Explicar a diferença entre treinar e usar um modelo de raciocínio.
	9-12	Ilustrar o que acontece durante cada uma das etapas necessárias ao usar o aprendizado de máquina para construir um classificador ou preditor.
Ajustando representações internas	K-2	N/A
	3-5	Analisar um jogo onde se constrói uma árvore de decisão, descrevendo a organização da árvore e o algoritmo de aprendizagem usado para adicionar nós.
	6-8	Comparar como funciona um algoritmo de aprendizagem de árvore de decisão e um de aprendizagem de rede neural.

	9-12	Descrever como vários tipos de algoritmos de aprendizado de máquina aprendem ajustando suas representações internas.
Aprendendo através da experiência	K-2	N/A
	3-5	Explicar como a aprendizagem por reforço permite que um computador aprenda com a experiência (ou seja, tentativa e erro).
	6-8	Explicar as diferenças entre aprendizagem supervisionada e aprendizagem por reforço.
	9-12	Selecionar o tipo apropriado de algoritmo de aprendizado de máquina (aprendizado supervisionado, não supervisionado ou de reforço) para resolver um problema de raciocínio.

Tabela 1: Objetivos sobre a Natureza da Aprendizagem da *Big Idea #3* (CSTA, 2017).

2.1.2.2 Redes Neurais

Conceito	Faixa	Objetivo de Aprendizagem
Estrutura de uma Rede Neural	K-2	N/A
	3-5	Ilustrar como uma rede neural de 1 a 3 neurônios é uma função que calcula uma saída.
	6-8	Ilustrar a estrutura de uma rede neural e descrever como suas partes formam um conjunto de funções que calculam uma saída.
	9-12	Descrever as seguintes arquiteturas de rede neural e seus usos: rede feed-forward, rede convolucional 2D, rede recorrente, rede adversarial geradora.
Atualização de Parâmetros	K-2	N/A
	3-5	Demonstrar como os pesos são atribuídos em uma rede neural para produzir um comportamento de entrada / saída desejado.
	6-8	Demonstrar como uma regra de aprendizado pode ser usada para ajustar os pesos em uma rede neural de uma camada.
	9-12	Treinar uma rede neural multicamadas usando o algoritmo de aprendizado de retropropagação e descrever como os parâmetros dos neurônios e as saídas das unidades ocultas mudam como resultado do aprendizado.

Tabela 2: Objetivos sobre Redes Neurais da *Big Idea #3* (CSTA, 2017).

2.1.2.3 Conjuntos de Dados

Conceito	Faixa	Objetivo de Aprendizagem
<i>Feature sets</i>	K-2	Criar um conjunto de dados rotulados com recursos explícitos para ilustrar como os computadores podem aprender a classificar coisas como alimentos, filmes ou brinquedos.
	3-5	Criar um conjunto de dados rotulado com recursos explícitos de vários tipos e usar uma ferramenta de aprendizado de máquina para treinar um classificador nesses dados.
	6-8	Criar um conjunto de dados para treinar um classificador ou preditor de árvore de decisão e explorar o impacto que diferentes codificações de recursos têm na árvore de decisão.
	9-12	Comparar dois conjuntos de dados do mundo real em termos dos recursos que eles abrangem e como esses recursos são codificados.
Grandes Conjuntos de Dados	K-2	N/A
	3-5	Ilustrar como treinar um classificador para um conceito amplo, como "cachorro", requer uma grande quantidade de dados para capturar a diversidade do domínio.
	6-8	Ilustrar como os objetos em uma imagem podem ser segmentados e rotulados para construir um conjunto de treinamento para reconhecimento de objetos.
	9-12	Avaliar um conjunto de dados usado para treinar um sistema de IA real, considerando o tamanho do conjunto de dados, a maneira como os dados foram adquiridos e rotulados, o armazenamento necessário e o tempo estimado para produzir o conjunto de dados.
<i>Bias</i>	K-2	Examinar um conjunto de dados rotulados e identificar problemas nos dados que podem levar um computador a fazer previsões incorretas.
	3-5	Examinar os recursos e rótulos dos dados de treinamento para detectar fontes potenciais de distorção.
	6-8	Explicar como a escolha dos dados de treinamento molda o comportamento do classificador e como o viés pode ser introduzido se o conjunto de treinamento não estiver adequadamente equilibrado.
	9-12	Investigar desequilíbrios nos dados de treinamento em termos de gênero, idade, etnia ou outras variáveis demográficas que possam resultar em um modelo enviesado, usando uma ferramenta de visualização de dados.

Tabela 3: Objetivos sobre *conjuntos de dados* da *Big Idea #3* (CSTA, 2017).

Complementando, Long e Magerko (2020) trazem uma investigação concreta sobre as necessidades curriculares em relação ao ensino de IA. A respeito do ensino de *ML*, a pesquisa sugere que é de grande importância que os aprendizes entendam como computadores aprendem, além de trazer exemplos de competências relacionadas a *ML*. Assim, desenvolveram uma definição concreta para *AI Literacy* (Alfabetização de IA), sugerindo competências chave para *AI Literacy*, conceitos que o aluno deve entender, divididos em 5 temas diferentes (Long e Magerko, 2020):

- O que é IA?;
- O que IA pode fazer?;
- Como IA funciona?;
- Como IA deve ser usada?; e
- Como as pessoas enxergam IA? .

Competência	Descrição
Competência 1 (Reconhecendo IA)	Distinguir entre artefatos tecnológicos que usam ou não IA
Competência 2 (Entendendo Inteligência)	Analisar e discutir criticamente características que fazem uma entidade “inteligente”, incluindo discutir diferenças entre humanos, animais e inteligência de máquina
Competência 3 (Interdisciplinaridade)	Reconhecer que existem muitos jeitos diferentes de pensar e desenvolver máquinas “inteligentes”. Identificar uma variedade de tecnologias que usam IA, incluindo sistemas cognitivos, robótica e ML
Competência 4 (Generalista vs. Específico)	Diferenciar entre IA generalista e IA específica
Competência 5 (Pontos Fortes e Fracos de IA)	Identificar tipos de problemas que IA se sobressai e problemas que são mais difíceis para a mesma. Usar essa informação para determinar quando é apropriado usar IA e quando usar habilidades humanas.
Competência 6 (Imaginar IA do Futuro)	Imaginar possíveis aplicações futuras de IA e considerar os efeitos de tal aplicação no mundo
Competência 7 (Representação)	Entender o que representação de conhecimento é e descrever alguns exemplos
Competência 8 (Tomada de Decisão)	Reconhecer e descrever exemplos de como computadores raciocinam e fazem decisões
Competência 9 (Passos de ML)	Entender os passos envolvidos em <i>machine learning</i> , as práticas e desafios de cada passo.
Competência 10 (Papel Humano na IA)	Reconhecer que humanos fazem um papel importante na programação, escolha de modelos e no refinamento de sistemas de IA
Competência 11 (<i>Data Literacy</i>)	Entender conceitos básicos de <i>data literacy</i>

Competência 12 (Aprender pelos Dados)	Reconhecer que computadores geralmente aprendem a partir dos dados
Competência 13 (Interpretar Criticamente os Dados)	Entender que dados não podem ser tomados por si só e requerem interpretação. Descrever como os exemplos dos dados de treinamento de um <i>dataset</i> inicial podem afetar o resultado de um algoritmo
Competência 14 (Ação e Reação)	Entender que alguns sistemas de IA têm a habilidade de fisicamente agir no mundo. Essa ação pode ser direcionada por raciocínio de alto nível (como andar em um caminho planejado) ou pode ser reativo (pular para trás para evitar um obstáculo percebido)
Competência 15 (Sensores)	Entender o que são sensores, reconhecer que computadores percebem o mundo usando eles e como identificá-los em uma variedade de dispositivos. Reconhecer que diferentes sensores suportam diferentes tipos de representação e raciocínio sobre o mundo.
Competência 16 (Ética)	Identificar e descrever diferentes perspectivas nos problemas éticos chave acerca de IA (privacidade, singularidade, tomadas de decisões éticas, etc.)

Tabela 4: Competências relacionadas ao ensino de ML (Long e Magerko, 2020).

2.1.3 Estratégias Instrucionais

A maioria dos cursos de *ML* usa estratégias de ensino ativo (Marques et al., 2020a). Essas pesquisas mostram que existem estratégias que têm se mostrado bem sucedidas no ensino de computação dentro das salas de aula e que tem o potencial de diminuir o estresse cognitivo por meio da disponibilização de materiais didáticos que servem como esqueleto para atividades interativas. Essa metodologia também tem o objetivo de limitar a ansiedade dos estudantes em relação a atividades percebidas como "muito difíceis" como é a programação para iniciantes (Lytle et al., 2019). Lee et al. (2011) sugere a implementação do ciclo de aprendizagem Use-Modifique-Crie. Este método consiste em três fases. **Use** refere-se a fase de disponibilização de um modelo já criado para que os estudantes possam interagir. Na fase **Modifique**, os estudantes são convidados a alterar o código existente para se adequar aos seus desejos. **Crie** é a fase final, onde se espera que os alunos já tenham conhecimento suficiente para desenvolver um modelo completamente novo (Lytle et al., 2019). Isso demonstra que existe um reconhecimento por parte da comunidade científica que é importante investir em aulas mais interativas e experimentais nos ambientes escolares. Alunos do Ensino Médio são usuários nativos de dispositivos digitais, portanto é importante que os professores usem os mesmos para cativar a atenção dos alunos, por meio de

recursos atrativos, como exposição de fatos interessantes, jogos e interfaces interativas (Estevez et al., 2020).

Voltado a essa metodologia ativa, no contexto de Educação Básica, é comum encontrar iniciativas que utilizam linguagens e mecanismos baseados em blocos (Weintrop, 2019), tornando mais acessível o ensino de computação para essa estudantes que não tem conhecimento mais aprofundado dos fundamentos de programação (Marques et al., 2020a). No entanto, já existem ferramentas que permitem o ensino de forma simples dos fundamentos de *ML* com uso de mecanismos baseado em texto, como é o caso da biblioteca Fastai (<https://www.fast.ai/>) e da plataforma Jupyter. Para o público do Ensino Médio, é esperado que essa faixa etária já seja capaz de treinar uma rede neural usando ferramentas interativas (Touretzky et al., 2019). Alunos com conhecimento mais avançado tem capacidade de codificar aplicações simples usando linguagens de programação como o Python e seus *frameworks*. Porém, para auxiliar iniciantes também existem ferramentas visuais de *ML* (Gresse von Wangenheim et al., 2020) como Google Teachable Machine ou Vulcan (Franz, 2021), que usando um ambiente mais convencional (Jupyter Notebook) ajuda a remover algumas barreiras de aprendizado e a necessidade de adicionar pré-requisitos de conhecimentos de programação.

É também importante durante o processo de participação do curso que o desempenho dos alunos seja avaliado tanto para guiá-lo em seu processo de aprendizagem quanto para usar esse *feedback* para melhorias na unidade instrucional. Essa avaliação pode ser feita de diversas formas, por exemplo, tanto por meio de *quizzes* interativos quanto por automatização da avaliação de artefatos criados pelos alunos como resultados da aprendizagem.

2.2 Processos e Ferramentas de *ML*

2.2.1 *Machine Learning*

Machine Learning usa as habilidades de computadores para complementar a inteligência humana, muitas vezes executando tarefas que seriam impossíveis de serem realizadas dentro das habilidades humanas. A habilidade de examinar e processar grandes quantidades de dados permite que programas de *ML* possam detectar padrões fora do escopo de percepção humana (Shalev-Shwartz e Ben-David, 2014). Assim como a programação, *ML* é um meio de fazer com que os computadores completem um determinado trabalho. A diferença é que a primeira necessita que o programador forneça passos específicos para que o computador complete a tarefa, enquanto *ML* usa mecanismos para aprender o que precisa para completar uma tarefa. Isso ocorre por meio do uso de variáveis chamadas de pesos (ou parâmetros) que são atualizados automaticamente a fim de atingir determinada performance (Howard et al., 2020). Algoritmos de *ML* são capazes de criar suas próprias representações por meio do processo de treinamento usando dados que podem tanto ser fornecidos por pessoas quanto adquiridos pela própria máquina (Touretzky et al., 2019).

Alguns exemplos de uso de *ML* no cotidiano são os carros autônomos que identificam partes do sistema de trânsito, como outros carros, estradas, sinalização e pedestres. O uso de *ML* pode variar desde reconhecimento de fala humana, linguagem natural e até classificação de imagens através do uso de múltiplas camadas de redes neurais. Cada uma dessas camadas recebe entradas das camadas anteriores, refinando-as de forma progressiva. Existem algoritmos envolvidos que treinam as camadas de maneira que os erros sejam diminuídos e a acurácia seja melhorada.

Existem diferentes tipos de abordagens de aprendizagem, sendo os três mais conhecidos:

- **Aprendizagem supervisionada:** O sistema é treinado com dados rotulados, indicando para cada dado a qual categoria pertence. Esses são conhecidos como dados de treinamento e o sistema usa esse conhecimento para prever as categorias de novos dados (Royal Society, 2017).

- **Aprendizagem não supervisionada:** Aprendizado ocorre sem a rotulação prévio dos dados e tem como objetivo detectar características em comum do grupo de dados, usando, por exemplo, *clusters* e atribuindo dados a esses *clusters* (Royal Society, 2017).
- **Aprendizagem de reforço:** O aprendizado por reforço foca no aprendizado com a experiência, sendo uma intersecção entre supervisionado e não supervisionado. Usa mecanismos de recompensa para otimizar os resultados (Royal Society, 2017). Após completar uma sequência, a rede neural recebe valores lógicos e reais que definem se o resultado foi considerado certo ou errado (Kriesel, 2007).

A construção de aplicações de ML ocorre em um processo iterativo de etapas sequenciadas (Figura 3), normalmente incluindo:



Figura 3: Etapas do processo de ML (Gresse von Wangenheim, 2021).

1. **Análise de Requisitos** - Identificar e especificar o objetivo do modelo e que problemas ele resolve, assim como a caracterização dos *input* e *outputs* esperados (Amazon, 2019).
2. **Gerenciamento de Dados** - etapa em que os dados que serão consumidos pelo modelo de ML passam por processo de coleta, limpeza e preparação. Para a coleta de dados, existe a identificação de conjuntos de dados disponíveis, tanto genéricos quanto conjuntos de dados especializados em aprendizado por transferência. O tipo de dado pode variar dependendo do problema. Em aprendizagem supervisionada, os dados são rotulados. Os *conjuntos de dados* são divididos em conjunto de treinamento, usados para treinar o modelo, conjunto de validação que seleciona o modelo que melhor se encaixa na resolução do problema e conjunto de teste que avalia o desempenho do modelo escolhido (Marques et al., 2020a).
3. **Treinamento do modelo** - o algoritmo de ML aprende com o conjunto de dados de treinamento, os padrões que mapeiam as variáveis para a saída

esperada, produzindo um modelo que captura esses relacionamentos. A partir desse momento, o modelo de *ML* passa a estar disponível para gerar previsões para novas entradas de dados. Hiperparâmetros como taxa de aprendizado, tamanho do modelo, regularização são usados para otimização do modelo (Amazon, 2019).

4. **Avaliações de modelo** - O objetivo do modelo de *ML* é aprender padrões de dados genéricos não conhecidos ao invés de aprender a memorizar os dados de treinamento (Amazon, 2019). Para garantir que essa memorização não ocorra, os dados de treinamento são divididos em dois: o primeiro agrupamento é usado para o treinamento de fato e o segundo é chamado de modelo de validação, que é usado para testar o progresso. O treinamento é finalizado a partir do momento que os resultados se mostram satisfatórios (Kriesel, 2007) e iterações de atualização de parâmetros e treinamento podem ser realizadas até que o objetivo de predição seja atingido. Os resultados podem ser avaliados com diversas medidas. Para tarefas de classificação de imagem, por exemplo, a tabela de acurácia é mais tipicamente usada e tem como objetivo disponibilizar valores que mostram a qualidade do modelo treinado. Alguns exemplos desses valores são *loss function* (pontua quão bom ou ruim é o modelo) e *error rate* (proporção de imagens que foram identificadas incorretamente). Há também a matriz de confusão que permite visualizar facilmente quais categorias de imagens o modelo errou ou acertou (Howard, 2020). Há ainda outras métricas como a acurácia, que informa o quanto o modelo está acertando, precisão que identifica quantos resultados classificados como certos estavam realmente certos, *recall* que mostra com que frequência o classificador está encontrando exemplos de uma classe e o *F1 score* que combina o recall com a precisão.
5. **Implantação do modelo** - Nesta fase, o modelo pode ser exportado para ambientes de produção para ser utilizado em artefatos computacionais como aplicações web ou móveis.

2.2.2 Classificação de Imagens com Deep Learning

O uso de classificação de imagens é um dos tópicos mais discutidos em *ML*, devido a sua notoriedade em atuar em processos complexos, como por exemplo o reconhecimento de tumores malignos por meio de imagens de exames. Um modelo de classificação tenta prever a qual categoria o dado de entrada pertence. O modelo prediz baseado em um número de possibilidades discretas. Existem também outros tipos de modelos, como o de regressão que tenta prever uma ou mais quantidades numéricas, como temperatura ou localidade (Howard et al., 2020).

Deep Learning é a técnica de *ML* tipicamente usada para extrair e transformar dados através do uso de múltiplas camadas de redes neurais e seu uso varia desde reconhecimento de voz humana a classificação de imagens de animais (Howard, 2020). A rede neural é um tipo de função tão flexível que pode ser usada para solucionar uma grande gama de soluções através da descoberta dos parâmetros corretos. É uma abordagem de *ML* em que camadas de unidades computacionais se conectam umas às outras, de maneira similar a forma como os neurônios do nosso sistema nervoso se conectam (Royal Society, 2017). Na prática, são funções matemáticas tão flexíveis que, a partir da variação de seus parâmetros, podem resolver qualquer problema (Howard et al., 2020).

Redes neurais artificiais são baseadas na ideia de conectividade local e pesos compartilhados, reduzindo o número de parâmetros de aprendizagem, resultando em um treinamento mais eficiente (Sadouk, 2019). Esses parâmetros podem ser obtidos automaticamente por meio do uso de um mecanismo chamado retropropagação, que permite que a máquina aprenda por experiência (Howard, 2020).

Além das etapas envolvidas em desenvolvimento de *features* de *ML*, existem uma série de conceitos associados ao desenvolvimento de classificadores de imagens que são importantes de entender a fim de assimilar alguns processos de *ML* (Howard, 2020):

- **Transformações:** aplicadas em cada imagem ou lote de imagens, como redimensionamento.

- **Overfitting:** quando o modelo começa a memorizar o conjunto de treinamento ao invés de generalizar padrões nos dados.
- **Métrica:** função que mede a qualidade da predição do modelo usando o conjunto de validação.
- **Error rate:** métrica cuja porcentagem de imagens do conjunto de validação que foi classificado incorretamente.
- **Acurácia:** métrica cuja porcentagem de imagens do conjunto de validação que foi classificado corretamente.
- **Modelo pré-treinado:** modelo cujos parâmetros já foram gerados através do treinamento com outro *dataset*.
- **Aprendizagem por transferência:** o uso de um modelo pré-treinado para uma tarefa diferente.
- **Epoch:** processo em que cada lote do conjunto de dados teve a oportunidade de atualizar seus parâmetros internos do modelo.
- **Fine-tuning:** uma técnica de aprendizagem por transferência que atualiza os parâmetros do modelo pré-treinado através do treinamento de *epochs* adicionais, usando uma tarefa diferente da tarefa em que o modelo foi originalmente treinado.
- **Loss:** define a “medida de performance” que o sistema de treinamento pode usar para atualizar os parâmetros do modelo automaticamente.
- **CNN:** *Convolutional Neural Network*, um tipo de rede neural que trabalha particularmente bem com tarefas de visão computacional.

2.2.3 Arquiteturas, *Frameworks* e Ferramentas de *ML*

Existe uma grande diversidade de bibliotecas e *frameworks* que facilitam a criação de projetos *ML*. O Pytorch é um *framework* de *ML* de código aberto amplamente utilizado em pesquisas, que tem como objetivo acelerar o caminho desde a prototipagem até a implantação. Ele permite experimentação rápida e flexível, fornecendo a sua aplicação pronta para produção, por meio de uma interface amigável, treinamento distribuído e ecossistema de ferramentas e bibliotecas. Para simplificar o desenvolvimento de modelos de *ML*, existem *frameworks* de nível mais alto de abstração, como o Fastai (Howard et al., 2020). Fastai é uma biblioteca de código aberto que se propõe em trazer funcionalidades

de alto nível ao PyTorch, um *framework* de baixo nível. O foco da Fastai é fornecer uma camada de abstração sobre o PyTorch, fazendo com que o código seja escrito da maneira mais clara possível, focando no ensino eficiente dos fundamentos de *ML* e suas técnicas.

Algumas ferramentas como o Fastai, fazem uso de uma plataforma de programação experimental chamada *Jupyter Notebook* (jupyter.org). Esta é a ferramenta mais popular para desenvolver *data science* usando a linguagem Python (Howard et al., 2020). Ela tem como ponto forte uma interface que se assemelha a editores de texto, mas que ao mesmo tempo fornece mecanismos poderosos e flexíveis para desenvolver e rodar código, sem que seja necessário conhecimento de linha de comando ou a instalação adicional de programas.

O Google Colab (<https://colab.research.google.com>) é um serviço gratuito baseado em *Jupyter Notebook*, onde qualquer pessoa pode rodar código Python no seu navegador. As vantagens para o ensino de *ML* é que o Colab oferece acesso gratuito a GPUs, não necessita de configuração e permite o compartilhamento de notebooks. O serviço também permite salvar os notebooks usando o Google Drive, necessitando apenas que o usuário tenha uma conta do Google. A natureza interativa do Colab permite que os estudantes aprendam no seu ritmo próprio e possam brincar com a aplicação à medida que o código é modificado (Norouzi, 2020), fazendo com que esta ferramenta seja uma boa escolha para a realização de cursos online, usando as etapas do *Use-Modify-Create*.

3 ESTADO DA ARTE

Visando a análise do estado de arte em relação a unidades instrucionais voltadas ao ensino de *ML* no Ensino Médio, este trabalho tem como objetivo a atualização e complementação do mapeamento sistemático feito por Marques et al. (2020a), que levantou as unidades instrucionais publicadas entre 2009 e 2019, para o ensino de *ML* incluindo o ensino fundamental e Ensino Médio. Assim, de acordo com o foco do presente trabalho, nesta análise serão consideradas somente unidades instrucionais voltadas ao Ensino Médio que foram publicadas no período entre 2020-2021. Para levantar o estado da arte foi conduzido um mapeamento sistemático seguindo os procedimentos propostos por Petersen et al. (2015).

3.1 Definição do Protocolo de Revisão

A pergunta de pesquisa é: Quais unidades instrucionais existem para ensinar conceitos de *Machine Learning* no contexto do Ensino Médio (publicados em 2020-2021)? Essa pergunta de pesquisa é refinada nas seguintes questões de análise:

AQ1. Quais UIs existem?

AQ2. Quais conceitos de Machine Learning são ensinados nessas UIs?

AQ3. Quais são as características instrucionais dessas UIs?

AQ4. Como cada UI foi desenvolvida e como a sua qualidade foi avaliada?

Critérios de inclusão/exclusão. Foram consideradas quaisquer unidades instrucionais (curso, atividade, tutorial, etc.) que tivessem foco no ensino de computação, incluindo conceitos de *ML* do Ensino Médio, publicadas entre 2020 e 2021. Unidades instrucionais que focam no ensino de *ML* para alunos do ensino superior, ensino infantil ou fundamental, e/ou que abordassem o ensino de computação sem aprofundar-se em conceitos de *ML* foram excluídas. Também foram excluídas publicações em forma de resumos, blogs, vídeos ou ferramentas que não apresentam uma unidade instrucional.

Critérios de qualidade. Foram considerados apenas artigos ou materiais com informações suficientes relacionadas 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 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 emergente muitas UIs 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 Search String. A *search string* 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 1.

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: (01/01/2020 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*))) 2020 - 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 , 2021) OR LIMIT-TO (PUBYEAR , 2020)) AND (LIMIT-TO (SUBJAREA , "COMP"))
Google	"machine learning" teach "K-12" OR school [Período personalizado 1 de Janeiro de 2020 - 29 de abril de 2021]
MIT media lab ¹	--

Tabela 5. Search string para cada fonte.

¹ <https://appinventor.mit.edu/explore/research>

3.2 Execução da Busca

A pesquisa foi realizada em abril de 2021 pela autora e revisada pela orientadora (Tabela 2). Algumas pesquisas retornaram muitos resultados mesmo depois da adaptação das *search strings*, devido ao fato de que artigos que descrevem como utilizar IA na educação, como no aprendizado de análise de dados, correspondem aos mesmos termos de busca.

Fonte	No. de resultados da busca	No. de resultados analisados	No. de resultados potencialmente relevantes	No. de resultados relevantes
ACM	107	107	13	4
IEEE	320	320	17	0
SCOPUS	1,196	400	18	1
Google	204	204	13	3
MIT media lab	12	12	12	2
Total	1839	1043	73	10

Tabela 6: Número de artigos identificados por repositório e por fase de seleção.

Na fase inicial da análise, títulos e resumos foram verificados. Nas plataformas de pesquisa científica, diversos artigos foram descartados por focar na aplicação de *ML* no contexto da educação e não focar no ensino de *ML* para K-12. Também foram excluídos, a partir da pesquisa no Google, artigos superficiais e com poucas referências, assim como cursos não disponíveis gratuitamente para o público geral.

Na análise inicial, foram identificados 73 artefatos potencialmente relevantes. Posteriormente, todos os documentos foram lidos integralmente de maneira que fossem assegurados os critérios de inclusão e exclusão. Unidades Instrucionais claramente limitadas a estágios educacionais que não incluem o Ensino Médio foram excluídas, como, p.ex., (Toivonen et al., 2020)(Vartiainen et al., 2020)(Lin et al., 2020). Algumas UIs foram descartadas por não conter detalhes da estratégia instrucional. Este é o caso do estudo brasileiro de Santana et al. (2020), que foca mais aprofundadamente na análise do engajamento dos alunos em relação ao ensino de *ML* e Reyes et al. (2020) por não prover dados suficientes sobre o currículo do curso ou avaliação da aprendizagem. Apesar de prover a estratégia instrucional, outra pesquisa foi excluída por não focar em IA e trazer poucas conclusões sobre a efetividade da abordagem (Julie et al., 2020). A unidade

instrucional do Markham (2021) também foi excluída dos resultados relevantes por falta de informação disponível em relação ao estágio educacional, assim como esclarecimentos sobre seus pré-requisitos. Por outro lado, UIs sem essas limitações explícitas foram mantidas. Materiais duplicados e documentos que descrevem o mesmo projeto de UI foram unificados.

3.3 Análise dos Resultados

3.3.1 Quais UIs existem?

Foram identificadas 10 UIs consideradas relevantes para o ensino de *ML* para alunos do Ensino Médio publicados no período de 2020-2021 (Tabela 6). Relacionando com os achados de Marques et al. (2020a) referentes ao estágio educacional que compreende o Ensino Médio (Anexo 1), houve uma sutil queda no número de UIs publicadas em 2020 em relação a 2019. Em relação a 2021, podemos supor que apenas duas UIs foram classificadas em decorrência do período em que a busca e análise do presente estudo foi executada (abril de 2021). Existe uma mudança visível no modo de aplicação das UIs (Figura 4), que até 2019 eram majoritariamente aplicadas de forma presencial. A partir de 2020, nota-se um crescimento de UIs investindo em cursos on-line e atividades remotas. Pode-se inferir que a pandemia de COVID-19 pode ter afetado a maneira como algumas UIs foram conduzidas, visando manter a segurança dos participantes e professores. Isso é confirmado em pelo menos um estudo (Rodríguez-García et al. 2021), que menciona que essa mudança repentina pode ter afetado os resultados da pesquisa.

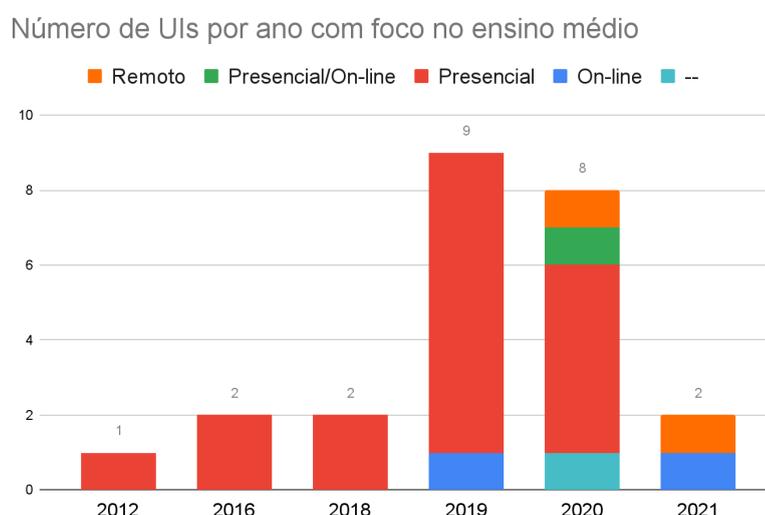


Figura 4: Número de UIs publicadas por ano com foco no Ensino Médio.

Referência	Nome da UI	Descrição
(Bhatia e Lao, 2020) (Bhatia, 2020)	Using Transfer Learning, Spectrogram Audio Classification, and MIT App Inventor to Facilitate Machine Learning Understanding	Projeto de pesquisa com foco em aplicar aprendizado por transferência e métodos de classificação de modelos de espectrograma de áudio para estimular o aprendizado de conceitos de ML.
(Bilstrup et al., 2020)	Staging Reflections on Ethical Dilemmas in Machine Learning: A Card-Based Design Workshop for High School Students	Oficina com foco nas implicações éticas por traz no processo de desenvolvimento de ML.
(Chittora e Baynes, 2020)	Interactive Visualizations to Introduce Data Science for High School Students	Ferramenta educacional interativa e visual para ensinar <i>data science</i> e conceitos básicos de ML.
(Google, 2021)	Introduction to Machine Learning	Curso on-line que introduz os conceitos básicos com tarefas práticas como análise de dados e recomendações além de trazer discussão sobre implicações éticas.
(Gresse von Wangenheim et al., 2020)	Machine Learning for All – Introducing Machine Learning in K-12	Curso introdutório sobre conceitos básicos de ML, utilizando um modelo de reconhecimento de imagem para reciclagem de lixo através da plataforma Google Teachable Machine.
(Mike et al., 2020)	Equalizing Data Science Curriculum for Computer Science Pupils	Oficina cujo objetivo é adaptar um curso que incorpora tanto uma visão global de <i>data science</i> como o entendimento mais profundo de algoritmos de processamento de dados com foco em ML.
(Norouzi et al., 2020)	Lessons Learned from Teaching Machine Learning and Natural Language Processing to High School Students	Curso usa processamento de linguagem natural para ensinar alunos do Ensino Médio conceitos de ML, assim como avalia a efetividade do curso através da análise do feedback dos estudantes e propõe sugestões para os impedimentos identificados.
(Rodríguez-García et al. 2021) (Rodríguez-García et al. 2020)	Introducing Artificial Intelligence Fundamentals with LearningML & Evaluation of an Online Intervention to Teach Artificial Intelligence With LearningML to 10-16-Year-Old Students	Oficina que incentiva criação de projetos de programação usando a plataforma <i>LearningML</i> para identificar o impacto na aprendizagem dos alunos sobre IA. <i>LearningML</i> é uma aplicação que permite criar modelos de ML usando textos e imagens através de aprendizagem supervisionada.
(Van Brummelen et al., 2020) (Van Brummelen, 2019)	Teaching Tech to Talk: K-12 Conversational Artificial Intelligence Literacy Curriculum and Development Tools	Oficina com o intuito de estimular o entendimento dos estudantes em relação à inteligência artificial conversacional, respeitando as 8 Competências de IA identificadas por Long e Magerko (2020).
(Wan et al., 2020)	SmileyCluster: supporting accessible machine learning in K-12 scientific discovery	Ambiente que explora visualização de dados, exploração prática e aprendizagem colaborativa para maximizar as oportunidades de aprendizado de conceitos abstratos e complexos de ML para jovens estudantes.

Tabela 7: Unidades Instrucionais de ensino de ML.

3.3.2 Quais conceitos de Machine Learning são ensinados nessas UIs?

A grande maioria das UIs trabalha com aprendizagem supervisionada e suas etapas de treinamento, aprendizado e avaliação do modelo (Figura 5). A única exceção é a UI de Wan et al. (2020), que introduz conceitos e métodos de *k-means clustering* e algoritmos de aprendizagem não-supervisionada. Os resultados deste estudo mostram que essa UI proporcionou pouco entendimento dos alunos em relação ao processo de *clustering*. A maioria das UIs conceituam como os dados de treinamento influenciam o aprendizado e abordam as limitações de *ML*. Metade das UIs encontradas discutem os aspectos éticos e sociais. Bilstrup et al. (2020) trabalham especificamente com esta abordagem, estimulando os estudantes a explorarem as limitações e implicações éticas envolvidas em *ML*. Quatro das UIs promovem o fundamento de redes neurais em seus planos de ensino e três delas promovem alguma atividade prática relacionada ao tema, enquanto Mike et al. (2020) introduz algoritmos de rede neural artificial.

Abordagens introduzidas pelas UIs

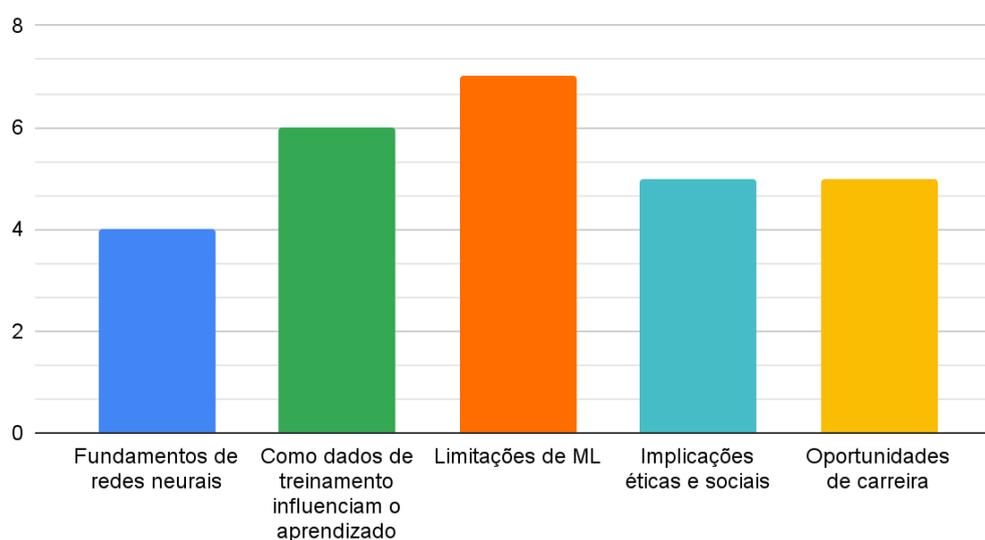


Figura 5: Número de UIs que introduzem cada abordagem.

Conceitos de regressão são introduzidos em algumas UIs, mas apenas Chittora e Baynes (2020) promovem o uso de um algoritmo para prever dados de condições climáticas. Esta mesma UI cita o classificador Naive Bayes e conceitos de Máquina de Vetores de Suporte (MVS). Mike et al. (2020) introduz os algoritmos de *k-nearest neighbors*, *perceptron* e MVS.

Referência	Abordagens de <i>Machine Learning</i>						Tipos de abordagens de aprendizagem		Outras abordagens				
	O que é aprendido?	Algoritmos de regressão	Classificador Naive Bayes	Clustering	Rede Neural	Outro	Super- visionada	Não Super- visionada	Fundamentos de Redes Neurais	Como dados de treinamento influenciam o aprendizado	Limitações de ML	Implicações éticas e sociais	Oportunidades de carreira
(Bhatia e Lao, 2020) (Bhatia, 2020)					x		x		x	x	x		
(Bilstrup et al., 2020)							x				x	x	
(Chittora e Baynes, 2020)		x	x			x	x			x	x		x
(Google, 2021)	x						x			x	x	x	x
(Gresse von Wangenheim et al., 2020)	x						x		x	x	x	x	x
(Mike et al., 2020)		x		x	x	x	x		x				
(Norouzi et al., 2020)	x	x			x		x		x	x		x	
(Rodríguez-García et al. 2021) (Rodríguez-García et al. 2020)							x						
(Van Brummelen et al., 2020) (Van Brummelen, 2019)	x						x			x	x	x	x
(Wan et al., 2020)				x				x			x		x

Tabela 8: Análise do conteúdo das Unidades Instrucionais.

A grande maioria das UIs propõe atividades e conteúdos voltados à tarefa de classificação (Figura 6). Os conjuntos de dados são tipicamente voltados ao uso de imagens ou são de alguma maneira convertidos em imagem, caso do estudo de Bhatia e Lao (2020), que usam espectrogramas, representações visuais das gravações de áudios. O artigo de Chittora e Baynes (2020) difere por usar séries temporais de dados de condições meteorológicas para introduzir conceitos tanto de regressão quanto de classificação e o curso do Google (2021), por usar dados tabulares para o desenvolvimento de um sistema de recomendação (Figura 7). Apenas uma das UIs fundamenta uma tarefa *ML* diferente, introduzindo *clustering* por meio da análise e interpretação de dados na ferramenta *SmileyCluster*. Wan et al. (2020) usam glifos para criar uma abstração que tem como objetivo ajudar os estudantes a entenderem como funciona o processo de *k-means clustering*.

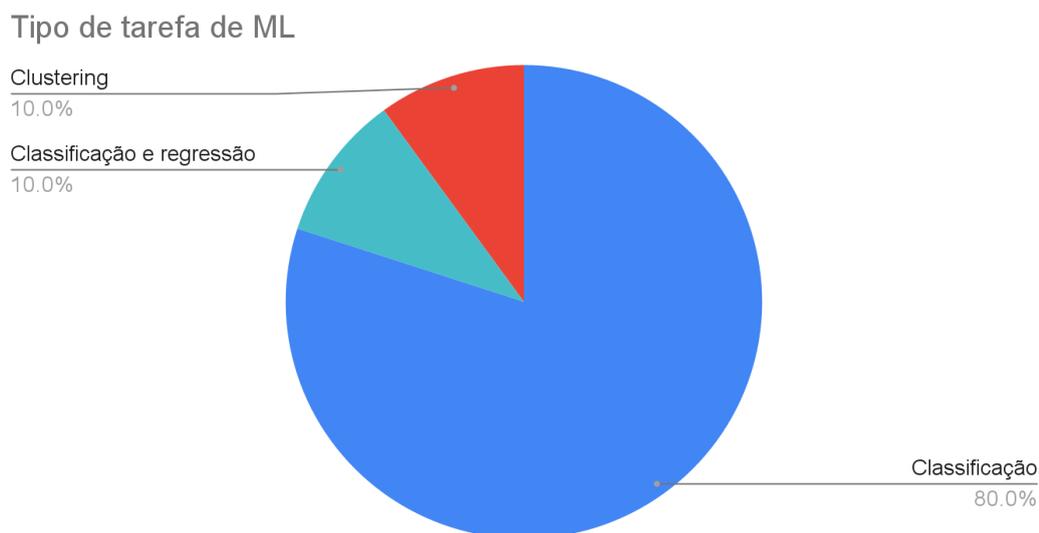


Figura 6: Porcentagem dos tipos de tarefas ML abordadas nas UIs.

Quatro UIs utilizam visão computacional e duas Processamento de Linguagem Natural, sendo que a UI de Rodríguez-García et al. (2020) permite treinar modelos de classificação tanto de texto quanto de imagens. Apenas Bilstrup et al. (2020) não especifica o domínio da aplicação, visto que esta UI é mais focada no ensino e discussão dos aspectos éticos envolvidos no desenvolvimento de soluções de *ML*.

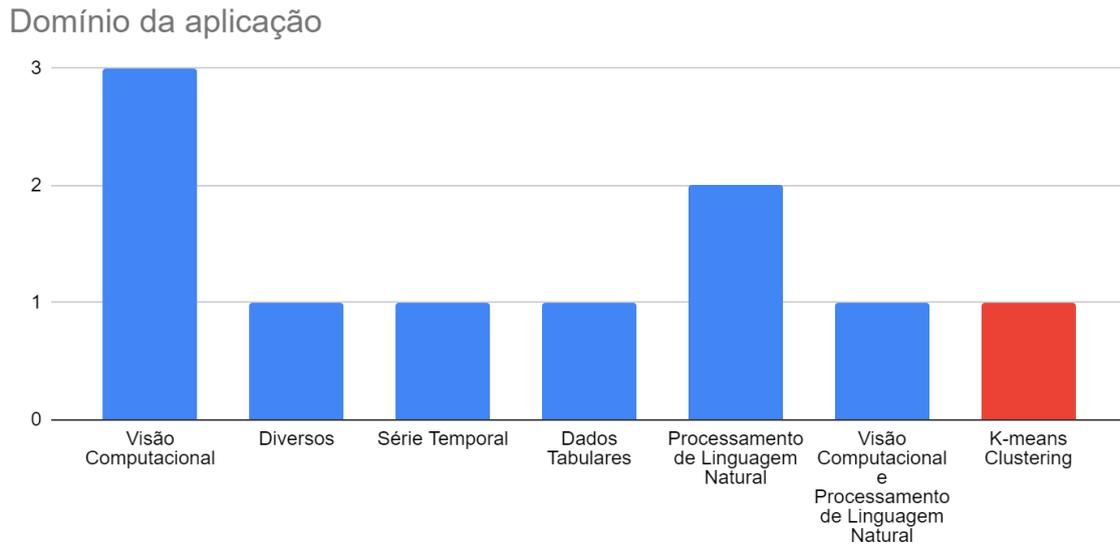


Figura 7: Número de UIs que abordam determinado domínio da aplicação.

As UIs escolheram utilizar *dashboards* e aplicações interativas para o entendimento dos processos de *ML*, entretanto existem duas UIs que utilizam a linguagem Python, sendo que Mike et al. (2020) descreve o uso da biblioteca *scikit-learn* para o desenvolvimento de atividades de Máquina de Vetores de Suporte e Redes Neurais Artificiais. É importante notar que estas duas UIs são cursos têm uma duração mais longa, ensinando conhecimento básico de programação no decorrer do curso (Norouzi et al., 2020) ou faz parte da grade curricular dos estudantes do Ensino Médio que fazem parte do estudo (Mike et al., 2020). Estas também são as UIs que utilizam o Google Colab como ferramenta de desenvolvimento e que guiam as tarefas práticas. Outras ferramentas conhecidas utilizadas são Google Teachable Machine e o MIT App Inventor, sendo que Bhatia e Lao (2020) ainda disponibilizam um extensão para facilitar o desenvolvimento de projetos que usam os dados produzidos pelo *Personal Audio Classifier*. As demais UIs utilizam suas próprias ferramentas Web para conduzir as lições. Bilstrup et al. (2020) não usa nenhum tipo de framework ou ferramenta. As UIs que fornecem instrução e atividade prática sobre o processo de *implantação* utilizam o MIT App Inventor e Scratch como ferramentas para implantação. As UIs que analisam a performance dos modelos criados, utilizam predominantemente tabelas de acurácia. Apenas a UI de Wan et al. (2020) promove uma interpretação informal do *cluster*.

Referência	Objetivo de aprendizagem	Domínio da aplicação	Tipo de tarefa de ML	Processo de ML	Dataset	Frameworks/ Ferramentas de ML	Linguagens de programação	Análise de performance	Deployment
(Bhatia e Lao, 2020) (Bhatia, 2020)	Incitar a curiosidade dos estudantes sobre ML através do uso de uma ferramenta de classificação de áudio pessoal	Visão Computacional (Classificação de espectrogramas)	Classificação	Gerenciamento de dados, treinamento do modelo, avaliação, deployment	Áudios de 1 segundo	Personal Audio Classifier, MIT App Inventor	NI	Teste de acurácia	MIT App Inventor
(Bilstrup et al., 2020)	Ajudar os estudantes a refletir sobre as implicações éticas por trás do desenvolvimento de soluções de ML	Diversos	Classificação	NI	NI	NI	NI	NI	NI
(Chittora e Baynes, 2020)	Criar uma ferramenta educacional e interativa de ML sem pré-requisitos de matemática e programação	Série Temporal (Análise de Dados Ambientais)	Classificação e regressão	Gerenciamento de dados, treinamento do modelo, avaliação	Dados de condições meteorológicas	Ferramenta Web (Chittora e Baynes, 2020)	Ferramenta de Fluxo de Trabalho Visual (sem código)	Acurácia e Precisão	NI
(Google, 2021)	Explorar ML por meio de atividades práticas relacionadas aos processos de ML e refletir nas implicações éticas.	Dados Tabulares (Sistema de Recomendação)	Classificação	Gerenciamento de dados, Coleta de dados, Treinamento do modelo, Avaliação do modelo	Dados de navegação	Google Spreadsheets, Google My Maps	NI	Acurácia	NI
(Gresse von Wangenheim et al., 2020)	Introduzir os conceitos básicos de Inteligência Artificial e <i>Machine Learning</i>	Visão Computacional	Classificação	Gerenciamento de dados, treinamento do modelo, avaliação do modelo, deployment	Imagens de objetos para reciclagem de lixo	Google Teachable Machine	Ferramenta de Fluxo de Trabalho Visual (sem código)	Tabela de acurácia e matriz de confusão.	MIT App Inventor
(Mike et al., 2020)	Prover uma visão geral sobre <i>data science</i> e ML sem pré-requisitos sólidos de matemática e algoritmos	Visão Computacional	Classificação	Gerenciamento de dados, treinamento do modelo, avaliação do modelo	Imagens	Google Colab	Python, scikit-learn	Matriz de confusão, outras métricas de erro.	NI
(Norouzi et al., 2020)	Engajar alunos do Ensino Médio a resolver problemas reais usando IA enquanto retém o conhecimento intelectual sobre o tema	Processamento de Linguagem Natural	Classificação	Gerenciamento de dados, treinamento do modelo, avaliação do modelo	Texto	Google Colab	Python	Tabela de acurácia	NI
(Rodríguez-García et al. 2021) (Rodríguez-García et al. 2020)	Ajudar jovens a entender os fundamentos de AI através de uma ferramenta interativa	Visão Computacional (Classificação de Imagens) e Processamento de Linguagem Natural	Classificação	Gerenciamento de dados, treinamento do modelo, avaliação do modelo, deployment	Imagens, Textos escolhidos pelo usuário	LearningML (Rodríguez-García et al. 2020)	Ferramenta de Fluxo de Trabalho Visual (sem código)	Tabela de acurácia	Scratch
(Van Brummelen et al., 2020)(Van Brummelen, 2019)	Testar e avaliar a interface e o conteúdo da oficina que segue as competências propostas por Long e Magerko (2020)	Processamento de Linguagem Natural (Agentes Conversacionais)	Classificação	Treinamento do modelo, Deployment	Áudios	MIT App Inventor, Amazon Alexa	NI	NI	MIT App Inventor
(Wan et al., 2020)	Ajudar jovens estudantes a entender os conceitos básicos e métodos de ML	<i>K-means Clustering</i>	Clustering	Visualização de dados baseada em glifos	NI	Ferramenta Web (Wan et al. 2020)	NI	Interpretação informal do <i>cluster</i>	NI

Tabela 9: Características do conteúdo relacionadas a ML.

3.3.3 Quais são as características instrucionais dessas UIs?

Algumas UIs se identificam como cursos, com carga horária que varia entre 4 e 80 horas. Estes cursos disponibilizam algum tipo de plano de ensino. Há também oficinas, que geralmente são lecionados em apenas um encontro, não possuem nenhum pré-requisito e duram poucas horas. Algumas UIs apresentam ferramentas próprias para facilitar o ensino de ML e conduzem oficinas para analisar sua efetividade. A única exceção é a pesquisa de Chittora e Baynes (2020), que apesar de fornecer a descrição detalhada da UI, não possui ainda resultados relacionados a experiências com os estudantes. O nível de dificuldade dos cursos considera que os alunos são novatos, tendo pouco ou nenhum conhecimento de algoritmos e nenhum conhecimento dos processos e métodos de ML. Mike et al. (2020) é a única exceção, utilizando-se da vantagem dos alunos da escola onde o estudo foi conduzido já terem aulas de programação em sua grade curricular.

Metade das UIs encontradas usam o ciclo completo Use-Modifique-Crie e a maioria destas propõe um projeto final com tema livre (Figura 8). Quatro UIs permitem que os estudantes apenas usem a aplicação, providenciando tutoriais ou aulas expositivas onde os alunos seguem passo a passo para realizar a tarefa de ML. A única UI que não usa nenhuma etapa do ciclo Use-Modifique-Crie é a de Bilstrup et al. (2020), por não usar ferramenta ou aplicação, visto que foca mais nos aspectos éticos de ML.

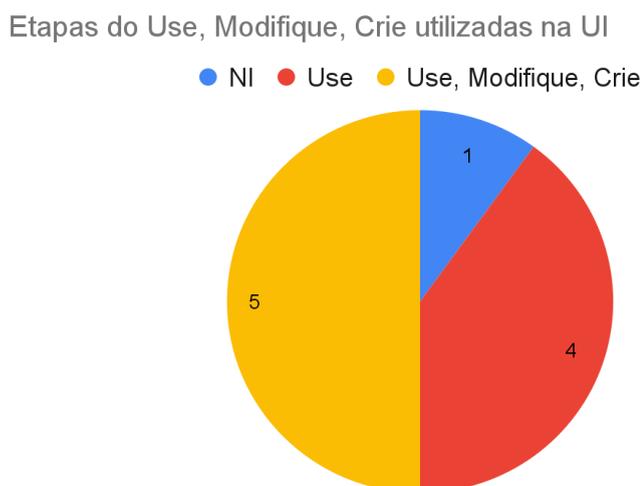


Figura 8: Porcentagem de tipos de UIs que utilizam as etapas Use-Modifique-Crie (Total n=10).

Grande parte das UIs usam algum tipo de atividade prática e aulas para introduzir os conceitos de *ML* (Figura 9). Discussões são estimuladas em certas UIs e a pedagogia construtivista é citada em algumas delas (Norouzi et al., 2020) (Mike et al., 2020) (Wan et al., 2020). Uma UI utiliza tutoriais para que os alunos consigam entender a parte de desenvolvimento de apps com o MIT App Inventor (Van Brummelen et al., 2020) e duas UIs que disponibilizam ferramentas próprias usam atividades de exploração nas mesmas (Bhatia e Lao, 2020)(Chittora e Baynes, 2020). A maioria das UIs usa uma combinação desses métodos instrucionais.

Métodos instrucionais

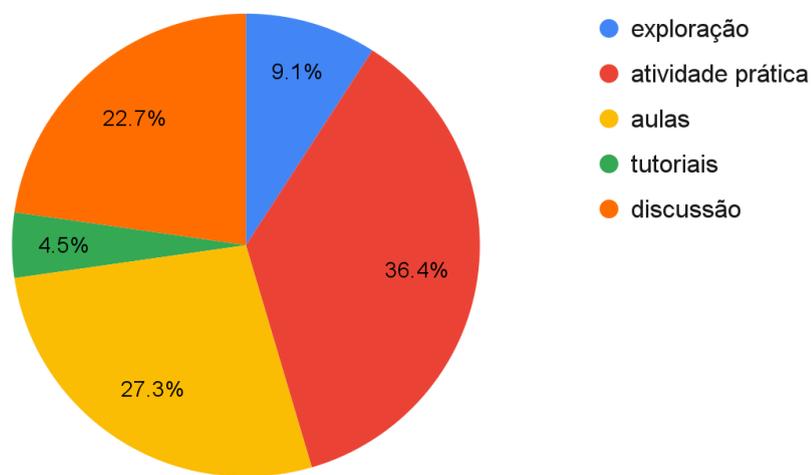


Figura 9: Porcentagem de utilização dos diferentes métodos instrucionais (Total n= 10).

Alguns termos relacionados aos objetivos de aprendizagem das UIs podem indicar o tipo de características que as UIs buscam para o ensino de *ML*. Através da coleta e síntese da nuvem de palavras encontradas nos artigos e materiais disponíveis sobre as UIs (Figura 10), podemos notar que estas têm investido em experiências interativas, que estimulem a reflexão e a curiosidade dos estudantes.



Figura 10: Nuvem de palavras e termos mais associados ao objetivo das UIs.

Referência	Tipo de UI	Nível de dificuldade	Etapas do use-modifique-crie	Modo de aplicação	Método instrucional	Características do método instrucional	Material instrucional	Método de avaliação da aprendizagem	Idioma	Licença	Treinamento do instrutor
(Bhatia e Lao, 2020) (Bhatia, 2020)	Ferramenta de Classificação de Áudio	Novato	Use, Modifique, Crie	Presencial	exploração da ferramenta	curiosidade	Google Quick Draw, plugin para MIT App Inventor	Questionários	Inglês	NI	Currículo, descrição das atividades, tutorial
(Bilstrup et al., 2020)	Oficina	Novato	NI	Presencial	atividade prática, discussão	exploração, reflexão	Apresentação, cartões e quadros	Observação, análise de gravações, fotos, vídeos e artefatos	NI	NI	Descrição das atividades
(Chittora e Baynes, 2020)	Ferramenta visual interativa	Novato	Use	NI	atividade de exploração	interativo, familiarizar, desmistificar	Ferramenta educacional de ML / Dashboard Interativo	NI	Inglês	NI	NI
(Google, 2021)	Curso	Novato	Use	On-line	aula, atividade prática	inferência, reflexão	Vídeos, tutoriais, spreadsheets, journal	Quiz no fim de cada lição, Rúbrica de Avaliação	Inglês	CC BY-SA 4.0	Currículo, descrição das atividades, vídeos
(Gresse von Wangenheim et al., 2020)	Curso	Novato	Use, Modifique, Crie	Presencial/ On-line	aula, atividade prática, discussão	interativo, disseminação	Slides, quiz	Rubrica de Avaliação	Português Brasileiro	NI	Currículo, descrição das atividades, slides
(Mike et al., 2020)	Curso	Intermediário	Use, Modifique, Crie	Presencial	aula, atividade prática, <i>active learning</i>	visão geral, entendimento	Apresentação, projeto final, trabalhos de classe	Entendimento das etapas de processamento de dados e qualidade de código	NI	NI	Oficina voltada aos professores
(Norouzi et al., 2020)	Curso	Novato	Use, Modifique, Crie	Presencial	aula, atividade prática, discussão	conscientizar	apresentação, tutoriais, surveys	Questionários	Inglês	NI	Currículo
(Rodríguez-García et al. 2021) (Rodríguez-García et al. 2020)	Plataforma educacional	Novato	Use, Modifique, Crie	Remoto	atividade prática	útil, atrativo, fácil	Tutoriais, vídeos	Questionários	Inglês, Espanhol, Galego, Catalão, Italiano, Holandês, Grego	GNU Affero General Public License	Instrução remota (e-mail e encontros)
(Van Brummelen et al., 2020)(Van Brummelen, 2019)	Oficina	Novato	Use	Remoto	tutoriais, aulas, discussão, atividade prática	interativo, engajar	Apresentação, questionários	Questionários	NI	NI	Currículo, descrição das atividades, tutorial
(Wan et al., 2020)	Plataforma educacional	Novato	Use	Presencial	atividade prática, discussão	acessível	Questionários	Questionários, entrevista semiestruturada com um grupo de foco	NI	NI	Descrição das atividades

Tabela 10: Características das Unidades Instrucionais.

3.4 Discussão

Esses resultados mostram uma estabilidade no número de UIs publicadas para o público de Ensino Médio, comparando com os dados coletados de Marques et al. (2020a). Neste comparativo, considerando todos os anos anteriores a 2020, é interessante notar uma mudança no modo de aplicação das UIs que até 2019 mostrava uma tendência de serem majoritariamente presenciais. A partir de 2020, percebemos uma porcentagem maior de UIs optando pelo modo remoto ou on-line.

É possível que a pandemia de COVID-19 tenha afetado a forma como algumas UIs foram conduzidas, caso da UI desenvolvida por Rodríguez-García et al. (2021). No artigo, os autores citam que a princípio, a oficina não foi idealizada para o ensino remoto e que a mudança abrupta pode ter afetado os resultados da pesquisa. Estes, no entanto, mostram que a UI obteve um impacto positivo no aprendizado dos alunos sobre o tema.

Pode-se inferir sobre a eficiência do ensino que foi idealizado como remoto de Van Brummelen et al. (2020), em que a maioria das competências de IA foram absorvidas e houve alto nível de engajamento por parte dos alunos. Como a oficina foi aplicada em um ritmo que todos os participantes pudessem finalizar as tarefas, é possível que alguns alunos sintam-se desmotivados esperando que todos os colegas completem as atividades. Para isso, são sugeridas atividades extras para encorajar alunos mais avançados.

A pesquisa de Gresse von Wangenheim et al. (2020) traz uma possível solução para este desafio, uma vez que esta UI também fornece uma versão on-line do curso, além de todo o material e plano de ensino para o modo presencial. A versão on-line do curso contém apresentações, questionários e outros artefatos interativos que podem auxiliar os alunos a concluir as atividades no seu próprio ritmo. Google (2021) usa uma estratégia similar, disponibilizando o curso integralmente on-line, vídeo-aulas e todos os documentos necessários para que o estudante conclua as atividades de forma virtual, ao mesmo tempo que a UI fornece o material necessário para que os professores possam aplicá-lo de forma presencial.

É visível que a presença de um curso integralmente on-line torna o ensino mais flexível e acessível, permitindo que os educadores possam optar pelo modo

de ensino que achar mais viável para seus alunos e estes possam assistir às aulas e realizar as atividades num ritmo próprio.

A análise também mostrou que as UIs para o público do Ensino Médio seguem uma tendência de utilizar aprendizagem supervisionada para introduzir os conceitos e processos de *ML*. Classificação é a tarefa mais abordada e visão computacional o domínio de aplicação mais utilizado. Estas também são as UIs que propõem atividades práticas que usam o ciclo Use-Modifique-Crie. Portanto, pode-se inferir que este tipo de tarefa se encaixa melhor e é mais acessível para este estágio escolar. Outro dado que fomenta essa inferência é o fato dos resultados da única UI que introduz um tipo de tarefa *ML* diferente (Wan et al., 2020) mostrarem que houve uma aprendizagem fraca sobre os procedimentos de *k-means clustering* por parte dos alunos.

Alguns outros achados pontuais são importantes de notar, como a necessidade de ensinar conceitos básicos de programação que podem ajudar a remover bloqueios dos alunos enquanto praticam as atividades propostas (Van Brummelen et al., 2020), assim como torna o ensino de *ML* mais acessível para estudantes do gênero feminino que tem maior tendência a evitar tarefas mais técnicas (Norouzi et al., 2020).

É possível concluir que cursos e oficinas sobre *ML* para alunos com pouco ou nenhum conhecimento de programação e conhecimento prévio sobre *ML* se beneficiaram em contemplar aprendizagem supervisionada e tarefas de classificação em seu currículo. UIs que ofertam uma versão on-line também permitiriam que educadores pudessem introduzir IA em concordância com o itinerário formativo sugerido pelo MEC (2017).

Observa-se ainda a falta de cursos em língua portuguesa. Também se observa uma falta de cursos no nível de Ensino Médio que já usem ambientes mais alinhados ao desenvolvimento “profissional” de modelos de *ML* e de forma interdisciplinar considerando a atual não inclusão de conceitos de IA na BCNN, necessitando então uma introdução por meio de outras disciplinas, como a biologia, por exemplo.

Ameaças à validade da revisão da literatura. Algumas ameaças podem afetar a validade do estudo de mapeamento. 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, já que tem como foco levantar e analisar as características das UIs, não é considerada uma ameaça neste caso. A possível omissão de estudos relevantes é outra ameaça que foi mitigada com a construção de *search strings* amplos, 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 *search string* com termos em inglês. Ameaças relacionadas à seleção de UIs 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, por meio de um protocolo rígido, pela autora e revisados pela orientadora.

4 CURSO ML4TEENS

Neste capítulo é apresentado o curso de ML4Teens voltado para o público do Ensino Médio, assim como a análise do contexto do público alvo e a definição do plano de ensino.

4.1 Análise de Contexto

Inicialmente, analisa-se o domínio do cenário em que pretende-se aplicar o curso, buscando assegurar que será coerente com a realidade.

Público alvo. O Ensino Médio brasileiro corresponde ao período de três anos que finaliza a Educação Básica. A faixa etária dos estudantes do Ensino Médio tipicamente é de 15 a 18 anos. Os estudantes desta faixa são fluentes em língua portuguesa e possuem conhecimento básico em pelo menos uma língua estrangeira, como o inglês ou espanhol. Espera-se que os estudantes do Ensino Médio consolidem os conhecimentos adquiridos no ensino fundamental, obtendo preparo básico para o mercado de trabalho em termos das disciplinas “tradicionais” (MEC, 2017).

Os jovens brasileiros (15-29 anos) são o grupo etário com maior acesso a tecnologias de informação e comunicação, computadores e acesso a energia elétrica em relação a outros grupos etários. Eles também possuem acesso ao celular em um percentual 3,5 vezes maior do que os outros grupos (Atlas das Juventudes, 2021). De acordo com o Censo Escolar Covid-19 (IFSC, 2021) realizado entre alunos de cursos técnicos integrados ao Ensino Médio e graduação do Instituto Federal de Santa Catarina, é possível inferir que os alunos, de modo geral, possuem acesso a computadores para desenvolver as atividades de ensino a distância (Figura 11), sendo que 75% dos entrevistados indicam ter ao menos um *notebook* disponível na família e 30% possuem ao menos um computador de mesa (Figura 12). Quase 90% participaram das atividades a distância de forma total ou parcial (Figura 13). Dos alunos que não participaram das aulas a distância, poucos justificaram a ausência por não ter computador disponível ou por possuírem internet fraca (2,7% e 2,4% respectivamente). Diante deste recorte no contexto do

IFSC, é possível supor que a maior parte dos alunos do Ensino Médio possam ter infraestrutura necessária para realizar cursos on-line de curta duração.

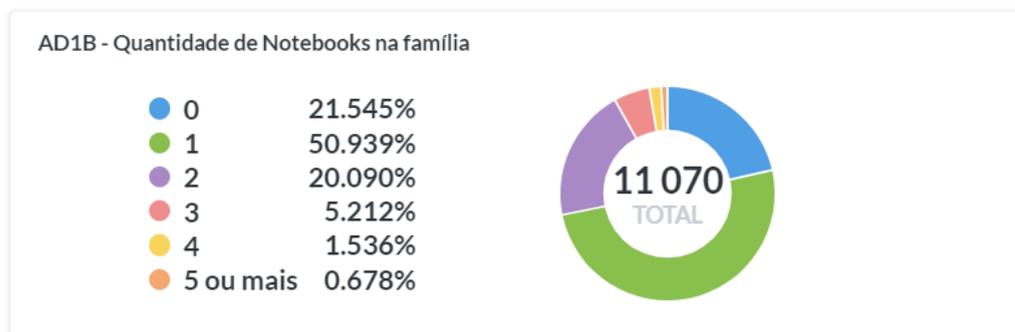


Figura 11: Quantidade de notebooks na família de acordo com o Censo Escolar Covid-19 (IFSC, 2021)(Total n= 11070).

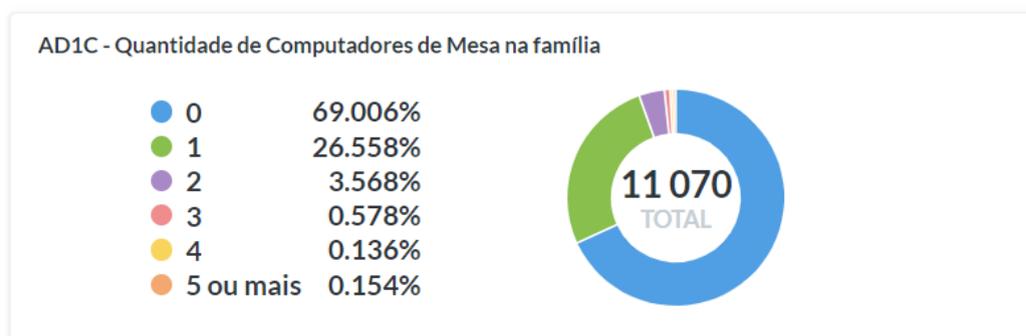


Figura 12: Quantidade de computadores de mesa na família (IFSC, 2021)(Total n= 11070).

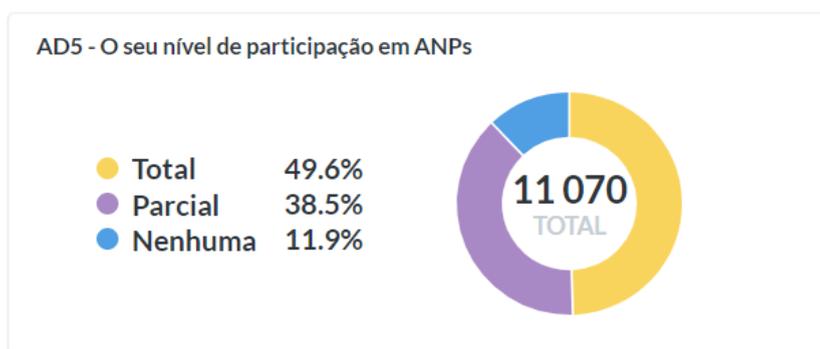


Figura 13: Participação dos alunos em atividades não presenciais (ANPs) (IFSC, 2021)(Total n= 11070).

O Ensino Médio é a etapa dos estudos em que os jovens sentem mais “peso”, exigindo determinação e esforço para estudar. O percentual de jovens

estudantes que buscam um futuro em áreas de ciência, tecnologia, engenharia e matemática (STEM) costuma ser baixo. A promoção de orientação, informação e estímulo por parte dos professores nas escolas pode ser uma estratégia importante para apoiar a decisão de mais alunos buscarem aperfeiçoamento nessas áreas (Atlas das Juventudes, 2021). A grande maioria dos alunos não têm conhecimento de computação e/ou inteligência artificial. Alguns talvez aprendam em cursos extracurriculares, já que neste momento este conteúdo ainda não faz parte da BCNN.

Existem evidências de que os jovens desta faixa etária passam a maior parte do seu tempo dedicados ao lazer no meio *on-line*. Nodari et al. (2014) mostra que “Navegar na internet” já era considerado o principal hobby dos jovens de 15 a 19 anos que compõem a pesquisa, ganhando de outras atividades como “Assistir TV” e “Ouvir ou tocar música”. A pesquisa ainda mostra que o uso de dispositivos móveis já está inserido no cotidiano dos jovens, sendo esta a principal ferramenta usada para o acesso a internet. Entre as atividades envolvidas no uso da rede estão “Fazer trabalhos escolares”, “Comunicação com amigos” e “Jogar” (Nodari et al., 2014).

Contexto. O Ensino Médio brasileiro tem como objetivo garantir aos estudantes a compreensão dos fundamentos científicos e tecnológicos dos processos produtivos, de modo a relacionar a teoria com a prática (MEC, 2017). Isto é feito por meio da articulação entre diferentes áreas do conhecimento: Linguagens e suas Tecnologias, Matemática e suas Tecnologias, Ciências da Natureza e suas Tecnologias, Ciências Humanas e Sociais Aplicadas. O limite do total da carga horária deste estágio educacional é de 1800 horas. Os currículos do Ensino Médio são compostos pela formação geral básica em conjunto com os itinerários formativos (MEC, 2017).

Em 2018, o Brasil contava com 28.673 escolas que ofertavam o Ensino Médio. Aproximadamente 60% dessas escolas possuem disponibilidade de infraestrutura e recursos tecnológicos como laboratório de informática e internet (INEP, 2019). O acesso à internet está disponível em 95,1% das escolas que ofertam Ensino Médio.

A Pandemia do Coronavírus causou a transição de aulas presenciais para o ensino on-line, afetando o aprendizado dos jovens. Metade deles acreditavam que os estudos seriam adiados e 9% pensavam que fracassariam nos estudos. Sobre as perspectivas futuras de carreira, 38% dos jovens ao redor do mundo relatam incerteza e esperam que a crise sanitária cause maiores obstáculos para transição da escola para o trabalho (Cetic.br. et al., 2020).

Interdisciplinaridade. O BNCC descreve no itinerário EM13CNT206, habilidades e competências que propõe a introdução da vida vegetal e conhecimentos relacionados à biodiversidade, políticas ambientais e ecossistemas (MEC, 2017). Essas competências vão ao encontro da importância do ensino de botânica nas escolas do Ensino Médio (Anjos et al., 2021). A disciplina de botânica é uma das mais tradicionais de todas as ciências, faz parte do currículo da Educação Básica e seus conceitos são revisados no terceiro ano do Ensino Médio.

Existe um consenso de que é uma disciplina desafiadora de ensinar visto que as aulas são majoritariamente expositivas e conceituais, em que o uso de nomenclatura técnica latinizada pode causar perda da atenção dos alunos. Também é de conhecimento que o uso de metodologias ativas e de mídias como vídeos e gravações parecem contribuir para o aprendizado (Atlas das Juventudes, 2021) e motivar os estudantes a se interessarem pela aprendizagem de botânica (Lindenmaier, 2017) (Dos Anjos et al., 2021).

4.2 Objetivos de Aprendizagem

Os objetivos de aprendizagem do curso ML4Teens foram desenvolvidos com base nas competências de alfabetização de IA de Long e Magerko (2020) e as diretrizes do AI4K12 (Touretzky et al. 2019). Os objetivos foram selecionados considerando o contexto educacional e a base necessária para conhecimento dos fundamentos de IA/ML e introduzindo conceitos para a realização da tarefa de classificação de imagens da atividade final do curso em concordância com a multidisciplinaridade junto ao tema de Botânica (MEC, 2017).

Categoria	Objetivo de aprendizagem	Referência(s)
OA1. Conceitos básicos de ML	Identificar exemplos de Machine Learning e diferenciá-lo da aprendizagem humana	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)
OA2 .Redes neurais	Compreender a estrutura de uma rede neural e descrever como suas partes formam um conjunto de funções que computam uma saída capaz de identificar padrões em dados	3-A-ii 3-5, 3-B-i 6-8, 3-B-ii 3-5 (AIK12, 2020)
OA3. Gerenciamento de dados	Preparar um conjunto de dados usado para treinar um modelo de ML considerando o tamanho deste conjunto, a forma com que os dados foram coletados e rotulados, além de sua qualidade (equilíbrio, balanceamento, viés)	3-C-ii 9-12 (AI4K12, 2020); 11, 12 (Long e Magerko, 2020); 1A-DA-05 (CSTA, 2017)
OA4. Treinamento de modelo de ML	Treinar um modelo de ML para classificação/predição usando um algoritmo de aprendizagem supervisionada com dados reais e ajustando os parâmetros de treinamento	3-A-ii 9-12, 3-A-iii 9-12 (AI4K12, 2020)
OA5. Processo de ML	Compreender e aplicar as etapas envolvidas no Machine Learning e suas práticas e desafios	3-A-iv 9-12 (AI4K12, 2020); 9 (Long e Magerko, 2020)
OA6. Ética de ML	Identificar e descrever diferentes questões éticas acerca de ML (privacidade, viés introduzido por características dos dados de treinamento, tomadas de decisões éticas, etc.)	3-C-iii 6-8 (AI4K12); 3A-AP-24 (CSTA, 2017); 13, 16 (Long e Magerko, 2020)
OA7. Impactos do IA/ML	Identificar prós e contras de IA e ML para atividades cotidianas e opções de carreira atuais e futuras	2-IC-21 (CSTA, 2017); 6 (Long e Magerko, 2020)
OA8. Biologia/ Botânica	EM13CNT206 / Discutir a importância da preservação e conservação da biodiversidade (MEC, 2017) por meio do reconhecimento de espécies de árvores nativas de Santa Catarina	(MEC, 2017)

Tabela 11: Objetivos de aprendizagem do curso ML4Teens.

4.3 Design do Curso

Alinhado ao contexto e aos objetivos de aprendizagem, o curso foi projetado para iniciantes para ensinar a tarefa de classificação de imagens utilizando *Deep Learning*. Escolheu-se a tarefa de classificação de imagens por ser uma das mais simples e introdutória e *Deep Learning* por ser uma das técnicas mais adotadas atualmente.

Prevê-se uma aplicação interdisciplinar deste curso, pela falta de inclusão explícita da área de conhecimento de IA/ML atualmente na BNCC. Assim, o curso será aplicado à área de conhecimento de biologia referente ao tópico de biodiversidade, políticas ambientais e ecossistemas (MEC, 2017) reforçando a aprendizagem de conceitos de botânica. Os alunos serão estimulados a exercitarem o entendimento sobre as diferentes espécies de árvores endêmicas e nativas de Santa Catarina a fim de realizar um projeto de classificação de imagens

de seis espécies de árvores presentes no campus da Universidade Federal de Santa Catarina. Espera-se que os estudantes consigam identificar os diferentes tipos de espécies nas imagens para que possam catalogá-las corretamente.

Assim, visa-se um projeto de classificação de imagens de seis espécies de árvores nativas/endêmicas de Santa Catarina (Tabela 12). Foram selecionadas 6 espécies comuns - limitada a seis para manter o esforço de desenvolvimento de modelo de ML razoável para o escopo das aulas do Ensino Médio. A seleção dessas espécies foi realizada em cooperação com o professor Pedro Fiaschi do departamento de Biologia da UFSC.

Nome popular	Imagem	Espécie	Família	Origem
Aroeira-vermelha		<i>Schinus terebinthifolia</i>	Anacardiaceae	Nativa
Jerivá		<i>Syagrus romanzoffiana</i>	Arecaceae	Nativa
Mulungu, Eritrina		<i>Erythrina speciosa</i>	Fabaceae	Endêmica

Capororoca		Myrsine guianensis	Primulaceae	Nativa
Embaúba		Cecropia glaziovii	Urticaceae	Endêmica
Pitangueira		Eugenia uniflora	Myrtaceae	Nativa

Tabela 12: Espécies de árvores nativas selecionadas.

A partir da aplicação do curso no Ensino Médio, levando em consideração que ainda não faz parte da BNCC, projetou-se um curso de curta duração. Planeja-se também a criação de um curso online que pode ser aplicado a distância (considerando a situação da pandemia Covid-19). Ele também poderá ser aplicado na forma presencial, em sala de aula, sem a necessidade da assistência de professores graduados na área de computação, uma vez que o curso fornecerá todo o conteúdo necessário para o entendimento do tema, assim como avaliação de aprendizagem dos alunos.

4.4 Plano de Ensino

O plano de ensino foi desenvolvido levando em consideração principalmente as diretrizes de AI4K12 (Touretzky et al. 2019) e as competências de Long e Magerko (2020) para levantar os objetivos do curso, levando em consideração o estágio escolar dos alunos do Ensino Médio. Também foram adicionadas atividades extras com o objetivo de estimular os estudantes a modificarem os códigos por trás do projeto de classificação de imagens, assim como a introdução do conteúdo necessário para entender o básico de algoritmos e uso de bibliotecas como o Fastai.

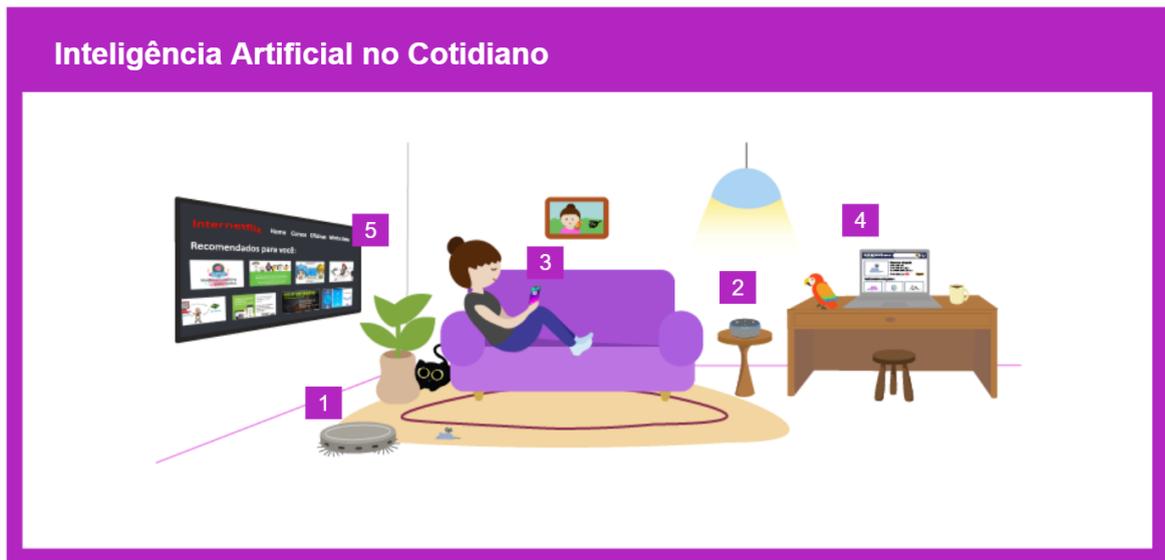
Todas as informações relevantes coletadas no mapeamento sistemático foram consideradas, como tipo de tarefa e processos de ML mais comumente usados que moldaram o plano de ensino do curso para focar em classificação de imagens, suas etapas e processos (Tabela 13). Para as atividades práticas, foi utilizada a ferramenta VisualJupyter, também conhecido como Vulcan (Franz, 2021).

Aula	Conteúdo	Objetivos de Aprendizagem	Método Instrucional	Material Instrucional	Avaliação
Introdução					
1. (15min)	Motivação, Aplicações de IA no cotidiano e Conceito de ML	OA1, OA7, OA8	Apresentação interativa, Vídeos, Demonstrações	Slides 1, "Rápido, desenhe!"	Quizzes
Conceitos Básicos					
2. (30min)	Aprendizagem Supervisionada e Classificação de imagens	OA1, OA3, OA4, OA5	Apresentação, Demonstrações, Vídeos, Atividade no Vulcan	Slides 2, Vulcan (Franz, 2021)	Quizzes
3. (30min)	Redes neurais	OA1, OA2	Apresentação, Vídeos	Slides 3, "Digit Recognizer"	Quizzes
Processo de ML					
4.1. (15min)	Preparação dos Dados (Limpeza, Classificação e Pré-processamento)	OA3, OA5, O8	Apresentação, Atividades	Slides 4, Vulcan (Franz, 2021)	Quizzes
4.2. (30min)	Treinamento, Avaliação do Modelo (Aprendizado por Transferência e Ajuste Fino)	OA4, OA5, O8	Apresentação, Atividade no Vulcan	Slides 4, Vulcan (Franz, 2021)	Quizzes, rubrica (Salvador, 2021)
4.3. (15min)	Predição	OA5, O8	Apresentação, Atividade no Vulcan	Slides 4, Vulcan (Franz, 2021)	Quizzes, rubrica (Salvador, 2021)
Implicações Éticas e Impactos					
5. (15min)	Responsabilidade, Defeitos, Preconceito e Desinformação	OA6	Apresentação, Vídeos	Slides 5, "Máquina Moral"	Quizzes
6. (15min)	Impactos e Oportunidades	OA7	Apresentação, Vídeos	Slides 6	Quizzes

Tabela 13: Plano de ensino do curso ML4Teens.

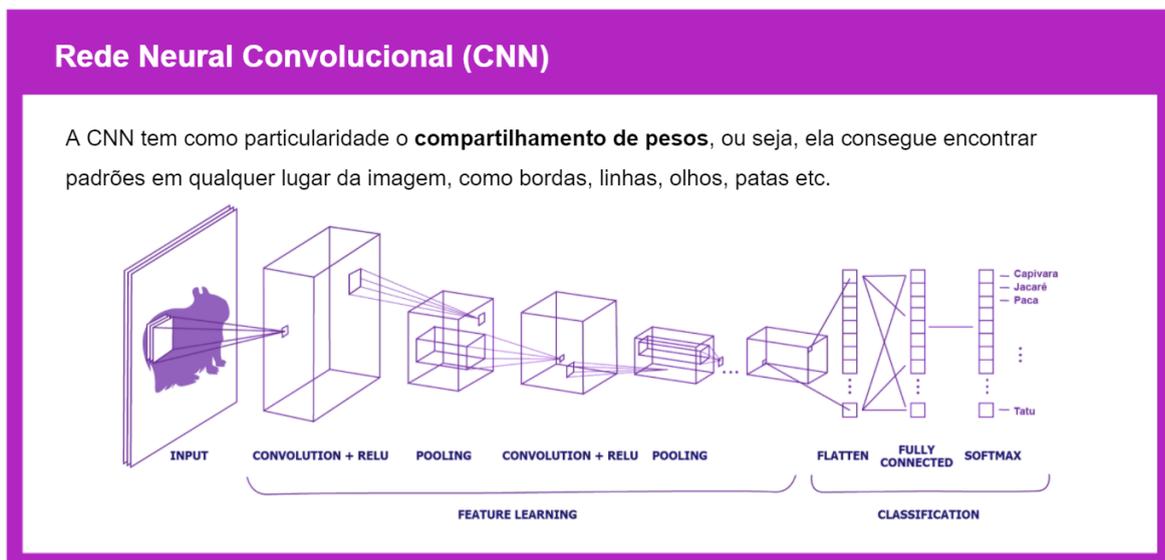
4.5 Material Didático

O material didático desenvolvido é composto por apresentações interativas, vídeos com conteúdo imersivo, como a demonstração do Vulcan (Figura 16) e introduzindo conceitos mais desafiadores ou que demandam mais exemplos.



Copyright © Computação na Escola/INCoD/INE/UFSC. Todos os Direitos Reservados. Proibida a distribuição e reprodução sem autorização prévia.

Figura 14: Ilustração com exemplos de uso de IA no cotidiano.



Copyright © Computação na Escola/INCoD/INE/UFSC. Todos os Direitos Reservados. Proibida a distribuição e reprodução sem autorização prévia.

Figura 15: Conceito ilustrado de uma rede neural convolutiva.

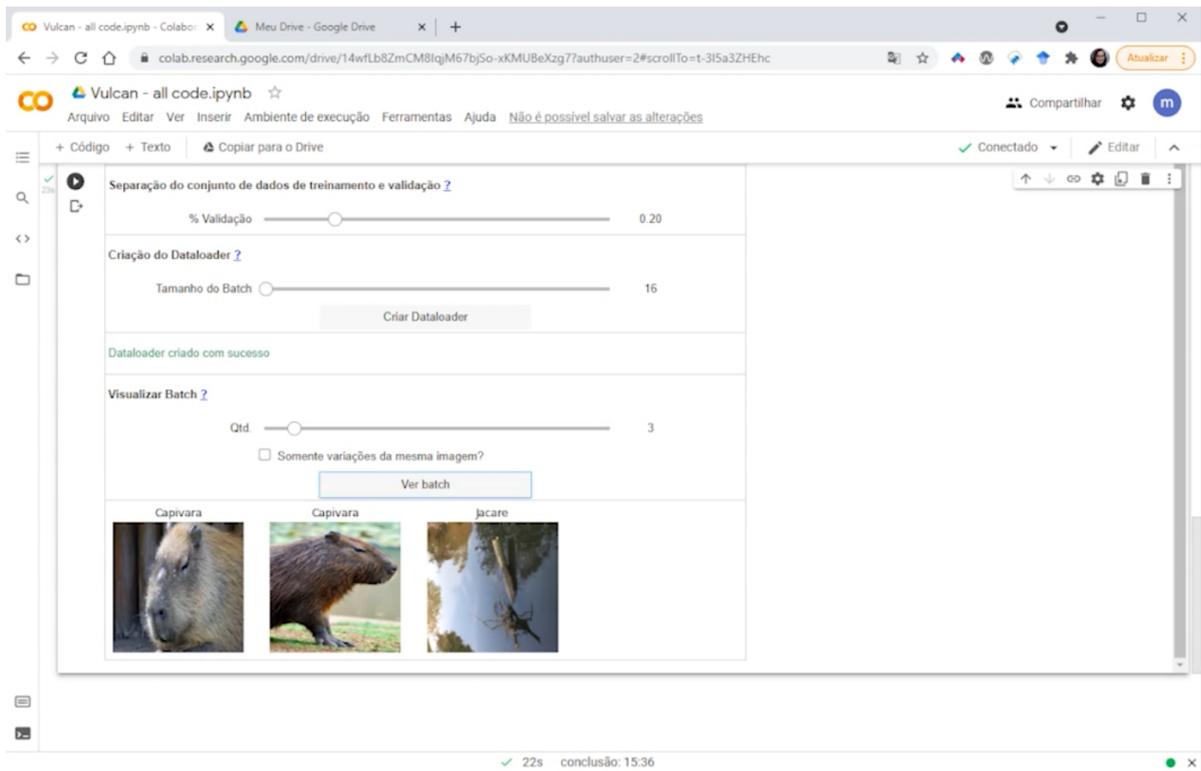


Figura 16: Vídeo sobre como usar o Vulcan para classificar imagens.

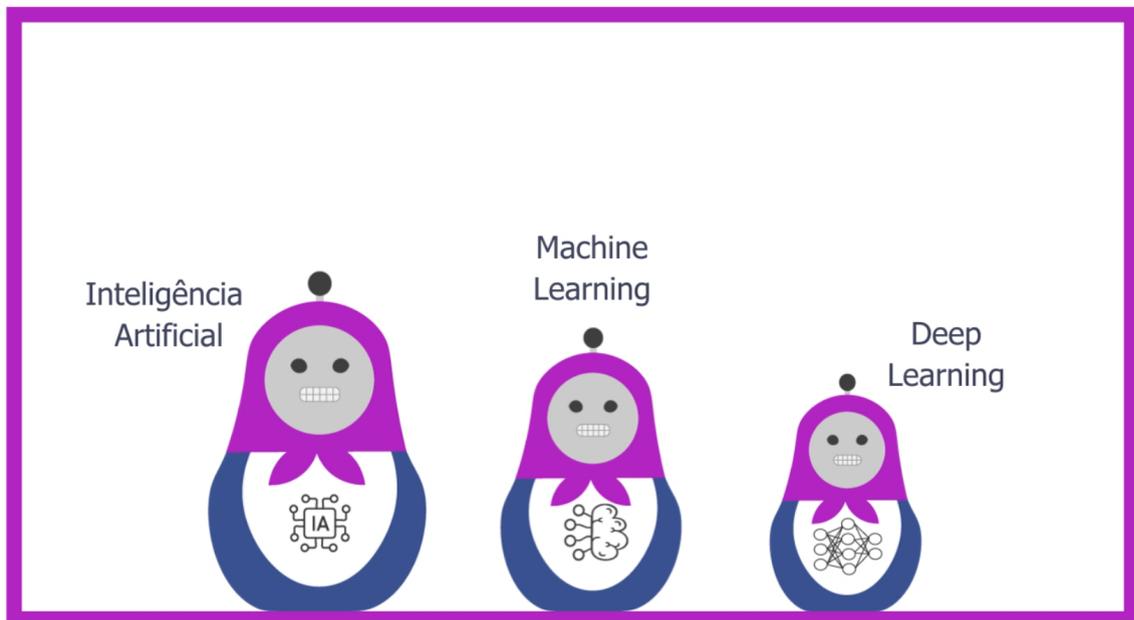


Figura 17: Vídeo introduzindo o conceito de *ML*.

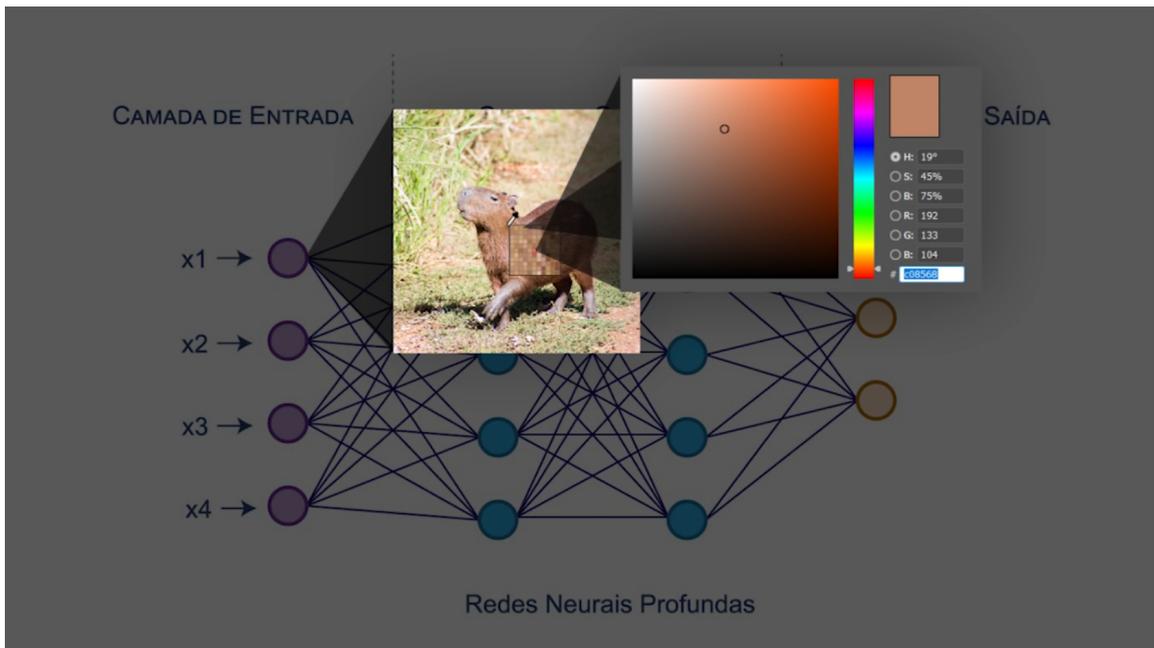


Figura 18: Vídeo introduzindo o conceito de redes neurais.

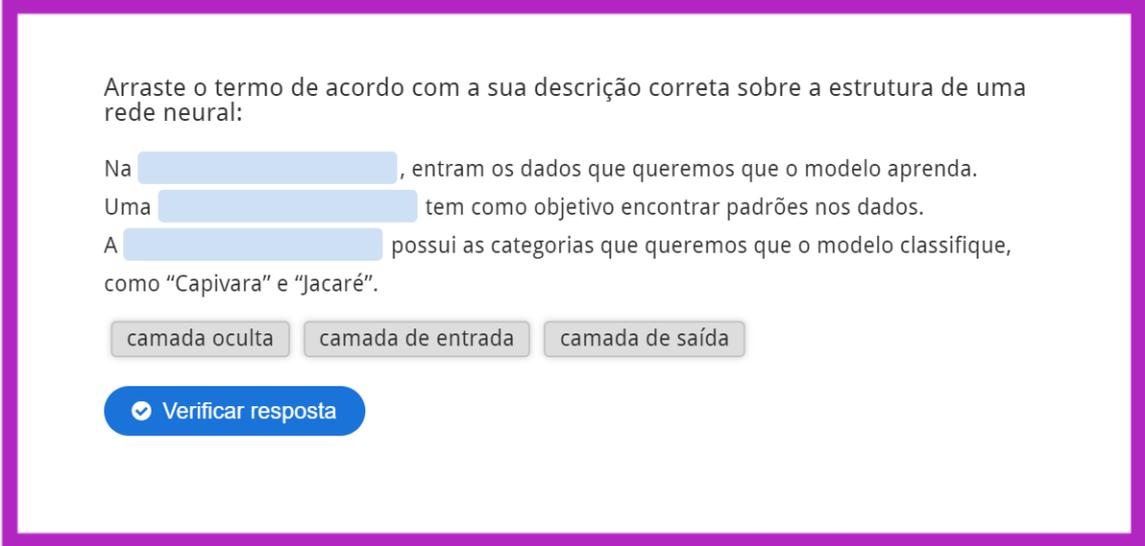
O design do curso foi desenvolvido com foco no público jovem, dispondo de ilustrações vetoriais, personagens exclusivos, memes e *emoticons*. O curso também possui narrativas como a de Zéd (Figura 19), o alienígena que tem como objetivo mapear a vida terrestre, que convida os estudantes a desenvolverem os modelos de classificação de imagens para ajudá-lo. Essa personagem foi definido com base em uma enquete on-line, com representantes do público alvo que puderam dar sugestões de mascote para o curso.



Figura 19: Zéd personifica o popular meme baseado no anime *Brave of the Sun Fighbird* (Romano, 2018).

As aulas também possuem quizzes interativos intercalados aos slides, utilizando essas funcionalidades disponíveis na plataforma HP5 (<https://h5p.org/>). Os quizzes possuem tipos variados, como:

- “arraste e solte” onde o estudante é convidado a completar termos em frases relacionados ao conteúdo da aula (Figura 20),
- perguntas com resposta única (Figura 21),
- múltipla escolha e respostas abertas (Tabela 15).

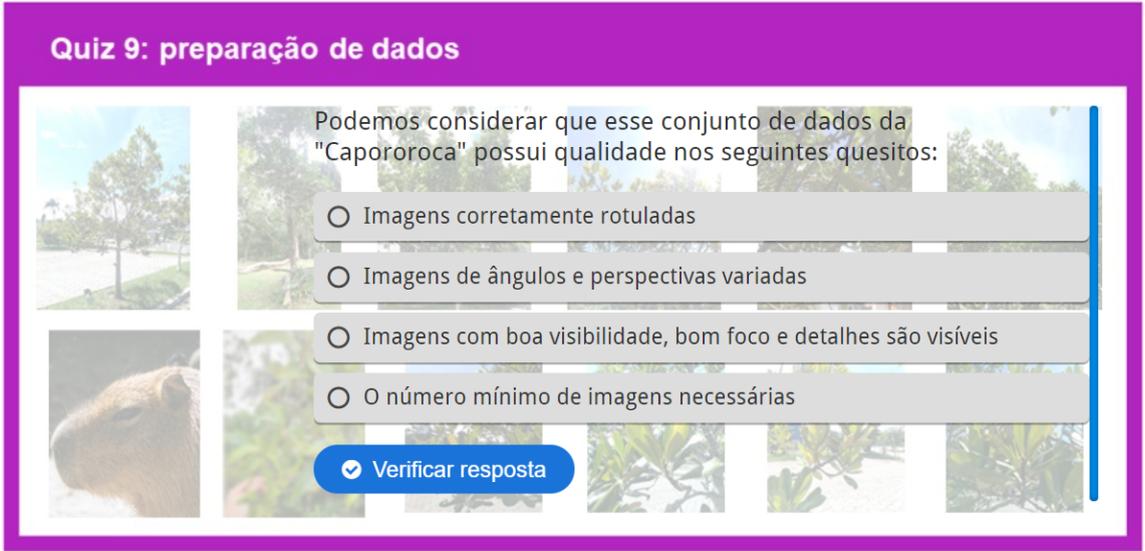


Arraste o termo de acordo com a sua descrição correta sobre a estrutura de uma rede neural:

Na , entram os dados que queremos que o modelo aprenda.
Uma tem como objetivo encontrar padrões nos dados.
A possui as categorias que queremos que o modelo classifique, como “Capivara” e “Jacaré”.

Copyright © Computação na Escola/INCoDiNE/UFSC. Todos os Direitos Reservados. Proibida a distribuição e reprodução sem autorização prévia.

Figura 20: Exemplo de quiz do tipo “arraste e solte”.



Quiz 9: preparação de dados

Podemos considerar que esse conjunto de dados da "Capororoca" possui qualidade nos seguintes quesitos:

Imagens corretamente rotuladas

Imagens de ângulos e perspectivas variadas

Imagens com boa visibilidade, bom foco e detalhes são visíveis

O número mínimo de imagens necessárias

Copyright © Computação na Escola/INCoDiNE/UFSC. Todos os Direitos Reservados. Proibida a distribuição e reprodução sem autorização prévia.

Figura 21: Exemplo de quiz do tipo “escolha única”.

Alguns links externos com atividades interativas são disponibilizados para que os alunos explorem recursos fora do curso. Um exemplo é a aplicação “Rápido, desene!” (Figura 22), onde os estudantes podem brincar com um modelo treinado para reconhecer desenhos simples, dando a oportunidade de entrarem em contato com dados rotulados.



Figura 22: Exemplo do recurso externo “Rápido, desene!”.

O curso on-line está disponível via moodle UFSC. Ele é dividido em aulas formatadas na plataforma HP5, que permitem a apresentação de todo material do curso, assim como a parte interativa (vídeos e quizzes) e será disponibilizado em 2022 no site de cursos da iniciativa Computação na Escola/INCoD/INE/UFSC.

4.6 Avaliação da Aprendizagem

A avaliação da aprendizagem dos alunos do curso é realizada por meio de quizzes interativos em cada módulo do curso. Para a avaliação de desempenho a partir do trabalho final criado como resultado da aprendizagem no Jupyter Notebook (Tabela 14), é utilizada a rubrica do modelo de avaliação de Salvador (2021).

ID	Critério de avaliação	Avaliação	Fonte
Introdução			
A1	Diferenciar entre como pessoas e computadores aprendem	Quizzes	3-A-i 3-5 (AI4K12, 2020)
A2	Reconhecer que existem muitos jeitos diferentes de pensar e desenvolver máquinas "inteligentes". Identificar uma variedade de tecnologias que usam IA, incluindo sistemas cognitivos, robótica e ML	Quizzes	3 (Long & Magerko, 2020)
A3	Distinguir entre artefatos tecnológicos que usam ou não IA	Quizzes	1 (Long & Magerko, 2020)
A4	Reconhecer e descrever exemplos de como computadores raciocinam e fazem decisões	Quizzes	8 (Long & Magerko, 2020)
Conceitos Básicos			
A5	Entender como aprendizagem supervisionada identifica padrões em dados rotulados	Quizzes	3-A-ii 3-5 (A4IK12, 2020)
A6	Ilustrar a estrutura de uma rede neural e descrever como suas partes formam um conjunto de funções que computam uma saída	Quizzes	3-B-i 6-8 (AI4K12, 2020)
A7	Demonstrar como pesos são designados em uma rede neural para produzir o comportamento desejado de entrada e saída	Quizzes	3-B-ii 3-5 (AI4K12, 2020)
Processo de ML			
A8	Ilustrar o que acontece em cada etapa necessária ao usar ML para construir um classificador ou preditor	Quizzes, rubrica (Salvador, 2021)	3-A-iv 9-12 (AI4K12, 2020)
A9	Entender as etapas envolvidas em Machine Learning e suas práticas e desafios	Quizzes, rubrica (Salvador, 2021)	9 (Long & Magerko, 2020)
Implicações Éticas e Impactos			
A10	Entender que os dados não devem ser considerados pelo seu valor bruto e requerem interpretação; Entender como os conjuntos de treinamento podem afetar os resultados de um algoritmo	Quizzes	13 (Long & Magerko, 2020)
A11	Comparar os prós e contras associados às tecnologias de computação que afetam as atividades cotidianas das pessoas e as opções de carreira	Quizzes	2-IC-21 (CSTA, 2017)
A12	Identificar tipos de problemas que IA se sobressai e problemas que são mais difíceis para a mesma. Usar essa informação para determinar quando é apropriado usar IA e quando usar habilidades humanas	Quizzes	5 (Long & Magerko, 2020)

Tabela 14: Plano de avaliação referente aos objetivos e competências.

Os quizzes foram formulados de acordo com os critérios de avaliação e objetivos de aprendizagem e são disponibilizados imediatamente após a exposição do conteúdo. A plataforma HP5, utilizada para a administração do curso, permite a inserção de questionários de diversos tipos, como Múltipla Escolha, onde o aluno pode selecionar mais de uma resposta correta dentre algumas opções, Escolha Única, Arrastar e Soltar para completar frases etc.

Id	Pergunta	Tipo	Critério de Avaliação
Introdução			
Q1	Usa inteligência artificial?	Múltipla Escolha	A3
Q2	Ligue a tarefa ao método ML em que ela se encaixa:	Arrastar e soltar	A2
Q3	Selecione exemplos de como os computadores raciocinam para tomam decisões:	Múltipla Escolha	A4
Conceitos Básicos			
Q4	Arraste o termo ao espaço correto sobre os processos de aprendizagem supervisionada:	Arrastar e soltar	A5
Q5	Como os humanos aprendem melhor do que os computadores?	Escolha Única	A1
Q6	Qual deve ser o valor dos pesos w_1 e w_2 para que o resultado indique para o banco se ele deve conceder o empréstimo ou não, considerando a tabela abaixo?	Respostas Abertas	A7
Q7	Arraste o termo de acordo com a sua descrição correta sobre a estrutura de uma rede neural:	Arrastar e soltar	A6
Q8	Arraste o termo de acordo com a sua descrição correta sobre as funções usadas em redes neurais:	Arrastar e soltar	A6
Processo de ML			
Q9	Podemos considerar que esse conjunto de dados possui qualidade nos seguintes quesitos:	Escolha Única	A9
Q10	Sobre a etapa de Transfer Learning, arraste os termos para os espaços corretos:	Arrastar e soltar	A9
Q11	Calcule o desempenho do modelo de acordo com as seguintes informações:	Respostas Abertas	A9
Q12	Junte o termo a sua respectiva descrição sobre a etapa de ajuste fino:	Arrastar e soltar	A9
Q13	Combine a etapa do processo de ML com a sua respectiva descrição:	Arrastar e soltar	A9
Implicações Éticas e Impactos			
Q14	Ligue a categoria do problema a sua respectiva descrição sobre os principais problemas éticos envolvidos no desenvolvimento em ML:	Arrastar e soltar	A9
Q15	Baseado em tudo que você aprendeu sobre preconceito, feedback loops e desinformação, você considera que os modelos de machine learning:	Escolha Única	A10
Q16	Complete com o termo relacionado a sua descrição correta sobre os tipos de problemas resolvidos por humanos e máquinas:	Arrastar e soltar	A12

Tabela 15: Quizzes inseridos como atividades interativas utilizando a ferramenta H5P.

Para a avaliação do desempenho em relação às principais atividades práticas do curso do processo de ML é adotada a rubrica de Salvador (2021), alinhadas aos objetivos de aprendizagem do curso ML4Teens e a ferramenta Vulcan de Franz (2021). Essa rubrica (Tabela 16) pode ser utilizada manualmente pelo instrutor em aplicações em turmas do curso. A sua aplicação de forma

automatizada depende ainda de uma integração do modelo de avaliação do Salvador (2021) dentro do ambiente visual Vulcan.

ID	Critério	Níveis de desempenho		
		Baixo - 0 pt	Aceitável - 1 pt.	Aceitável - 1 pt.
Preparação de dados (OA3)				
C1	Quantidade de imagens	Menos de 5 imagens por categoria	6 de 10 imagens por categoria	Mais de 10 imagens por categoria
C2	Distribuição do conjunto de dados	Quantidade de imagens por categoria varia muito	Quantidade de imagens por categoria varia pouco	Todas as categorias possuem a mesma quantidade de imagens
Preparação de dados/Botânica (OA3/OA8)				
C3	Rotulagem das imagens	Menos de 20% das imagens rotuladas corretamente	De 20% a 99% das imagens rotuladas corretamente	Todas as imagens rotuladas corretamente
Treinamento de modelo de ML/Transfer Learning e Fine-Tuning (OA4)				
C4	Treinamento - <i>Transfer Learning</i>	O modelo não foi treinado (<i>transfer learned</i>)	O modelo foi treinado com os parâmetros padrão	O modelo foi treinado com parâmetros ajustados (arquitetura, época e taxa de aprendizagem)
C5	Treinamento - Ajuste Fino	O modelo não sofreu o ajuste fino	Foi feito <i>unfreeze</i> das camadas e melhor taxa de aprendizagem não encontrada ou modelo não treinado	Foi feito <i>unfreeze</i> das camadas, a melhor taxa de aprendizagem foi encontrada e o modelo sofreu o ajuste fino
Avaliação e interpretação do desempenho de um modelo de ML (Transfer Learning e Fine-Tuning) (OA5)				
C6	Interpretação de acurácia	Categorias com baixa acurácia não identificadas	Categorias com baixa acurácia identificadas e interpretação incorreta em relação ao modelo	Categorias com baixa acurácia identificadas corretamente e interpretação correta em relação ao modelo
C7	Interpretação da matriz de confusão	Classificações incorretas não identificadas e interpretação incorreta em relação ao modelo	Classificações incorretas identificadas e interpretação incorreta em relação ao modelo	Classificações incorretas identificadas e interpretação correta em relação ao modelo
C8	Ajustes/melhorias feitas	Sem novas iterações de desenvolvimento	Uma nova iteração com alterações no conjunto de dados e/ou parâmetros de treinamento	Diversas novas iterações com alterações no conjunto de dados e/ou parâmetros de treinamento
Avaliação e interpretação do desempenho de um modelo de ML/Testar e aperfeiçoar programas (OA5)				
C9	Testes com novos objetos	Nenhum novo objeto testado	1-2 novos objetos testados	Mais de dois novos objetos testados
C10	Interpretação dos testes	Interpretação errada	---	Interpretação correta

Tabela 16: Rubrica de Salvador (2021) alinhada aos objetivos do curso ML4Teens.

5 AVALIAÇÃO DO CURSO

Neste capítulo é apresentada a avaliação do curso ML4Teens realizada por um painel de especialistas.

5.1 Definição da Avaliação

Uma avaliação preliminar foi realizada com o objetivo de analisar a qualidade do curso desenvolvido em termos de utilidade, adequação funcional, experiência de aprendizagem e usabilidade do ponto de vista de especialistas no contexto do ensino de *Machine Learning*.

Com base nas normas ISO/IEC 25010 (2011), ISO/IEC 9241 (1998), em TAM (Davis, 1989) e SUS (BROOKE, 1996), os fatores de qualidade a serem avaliados foram decompostos conforme apresentado na Tabela 17.

A avaliação do curso foi feita por meio de um painel de especialistas. Esta tem como objetivo avaliar a percepção dos participantes do painel em relação aos critérios de qualidade descritos na Tabela 15. O teste consiste em analisar o curso, assim como seus materiais, vídeos e links externos, por meio da plataforma Moodle, disponibilizada pela Universidade Federal de Santa Catarina.

A partir dos fatores de qualidade foi projetado um questionário de avaliação publicado no Google Forms. Dados demográficos são coletados para a análise das diferentes perspectivas dos membros do painel em relação a adequação do curso. O formulário é dividido em cinco sessões, com perguntas que utilizam a escala Likert de 5 pontos e respostas abertas para permitir sugestões de melhorias e apontamento de erros. O questionário é finalizado com perguntas de respostas abertas sobre os pontos fortes e fracos do curso.

Característica	Subcaracterística	Item no questionário	Escala de resposta
Utilidade	--	Acho o curso útil para ensinar Machine Learning no Ensino Médio.	Escala Likert de 5 pontos (Concordo totalmente, Concordo, Não concordo nem discordo; Discordo; Discordo totalmente)
		Acho que em sua forma atual, o curso atende objetivos de aprendizagem de ML no Ensino Médio.	
Adequação funcional	Completude	Você acha que está faltando algum conteúdo no curso?	Sim, não (se não, o que está faltando)
		Você acha que tem algo desnecessário no curso?	Sim, não (se sim, qual)
	Corretude	Você nota algum erro em relação ao conteúdo do curso?	Sim, não (se sim, qual)
Usabilidade	Ambiguidade	Você acha que o curso possui elementos ambíguos ou difíceis de entender?	Sim, não (se sim, qual)
	Estética	O design visual do curso é atraente.	Escala Likert de 5 pontos (Concordo totalmente, Concordo, Não concordo nem discordo; Discordo; Discordo totalmente)
		Acho que o design visual (fonte, cores, imagens) são compreensíveis.	
	Aprendizibilidade	Acho que alunos iniciantes no Ensino Médio conseguem alcançar os objetivos de aprendizagem.	
		O curso adota uma estratégia instrucional (tipos de atividades etc.) adequada para um curso online.	
Achei o curso desnecessariamente complexo.			
	Acho que é fácil alcançar os objetivos de aprendizagem com esse curso.		
Experiência de aprendizagem	Eficiência	Acho que o curso ensina o conteúdo de forma eficiente (permite aos alunos alcançar os objetivos com o esforço, tempo e recursos adequados).	Escala Likert de 5 pontos (Concordo totalmente, Concordo, Não concordo nem discordo; Discordo; Discordo totalmente)
	Eficácia	Acho que o conteúdo fornecido é suficiente para concluir todas as atividades propostas.	
		Acho que vídeos e animações ajudaram na compreensão do conteúdo.	
	Desafio	Acho que o curso é adequadamente desafiador aos alunos.	
	Satisfação	O curso não se torna monótono.	
		Eu recomendaria o uso deste curso para outros professores/instrutores.	
		Achei o curso fácil de entender.	
	Acho que a organização do conteúdo ajudará os alunos a aprender com este curso.		
Atenção focada	Há algo interessante no curso que captura a atenção dos alunos no início do curso.		
Relevância	O conteúdo é relevante.		
Pontos fortes e fracos	--	O que você mais gostou no curso?	Aberto (texto longo)
		O que você menos gostou no curso?	Aberto (texto longo)
		Mais alguma sugestão?	Aberto (texto longo)

Tabela 17: Visão geral da decomposição das características de qualidade e operacionalização da coleta de dados da avaliação.

5.2 Execução da Avaliação

A avaliação do curso foi realizada por um painel de especialistas, representados por professores do Ensino Médio, do ensino superior, de ambos e alunos da graduação (Figura 23). Devido a falta de competências em *ML* de alunos do Ensino Médio atualmente, foram convidados alunos da graduação e pós-graduação para integrarem o painel. Ao todo, foram convidados 12 participantes por email, dos quais todos participaram na pesquisa (100% taxa de resposta). Foi definido um período de coleta de avaliações de duas semanas para dar tempo necessário para os membros do painel conseguirem realizar as atividades do curso, que tem uma duração média de duas horas e meia. A avaliação ocorreu durante as últimas semanas de novembro de 2021.

Atuação dos participantes

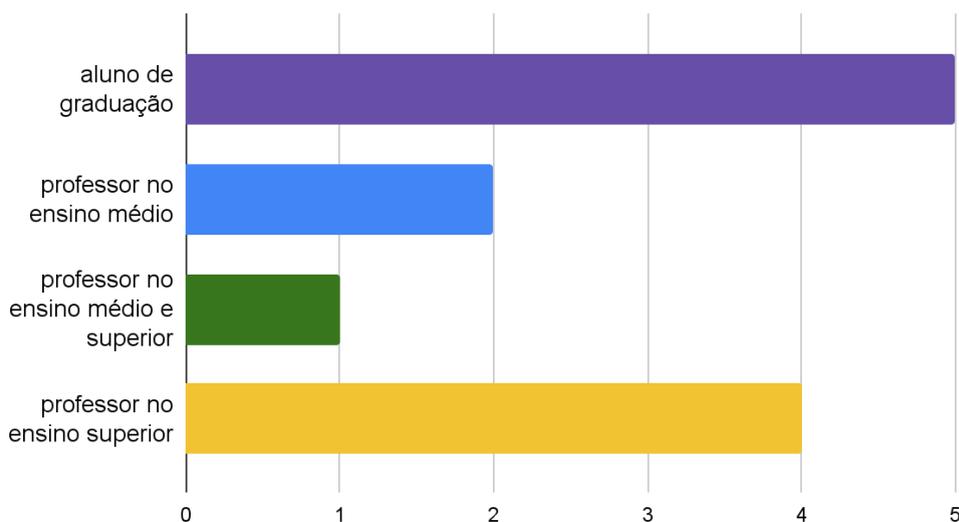


Figura 23: Quantidade de participantes em relação a sua atuação.

5.3 Análise da Avaliação

Os dados foram analisados de acordo com os fatores de qualidade definidos (Tabela 17). A análise e inferência sobre as respostas da avaliação estão sintetizadas a seguir.

5.3.1 Utilidade do curso ML4Teens

Em relação à utilidade do curso, todos os participantes concordam que ele é útil para o ensino de *ML* para estudantes do Ensino Médio e atende aos objetivos de aprendizagem propostos (Figura 24).

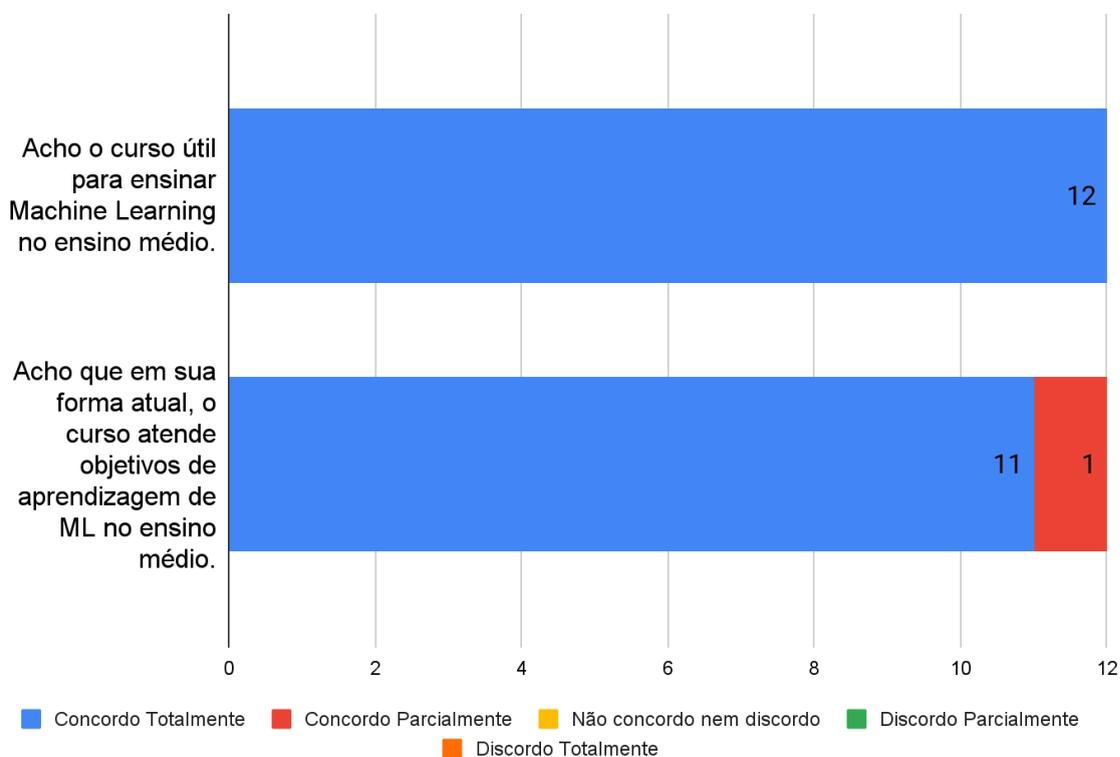


Figura 24: Resultados sobre a utilidade do curso.

5.3.2 Adequação funcional

Em relação à funcionalidade do curso, de forma geral os participantes o acharam adequado. Poucos participantes relataram que há elementos faltantes no curso e alguns erros foram reportados e corrigidos (Figura 25). Sobre tópicos desnecessários e difíceis de entender, apenas a aula sobre Redes Neurais foi citada, sendo que a maioria dos participantes demonstraram preocupação em relação à complexidade do conteúdo ministrado (Tabela 18), particularmente sobre os conceitos de Perceptron e Redes Neurais Convolucionais, o que pode desmotivar os alunos mais iniciantes. Como solução a esses pontos, a aula sobre Redes Neurais foi atualizada para possuir um conteúdo mais simplificado e abstraído e ainda permitindo que estudantes interessados possam acessar os conteúdos mais complexos na forma de aulas extras.

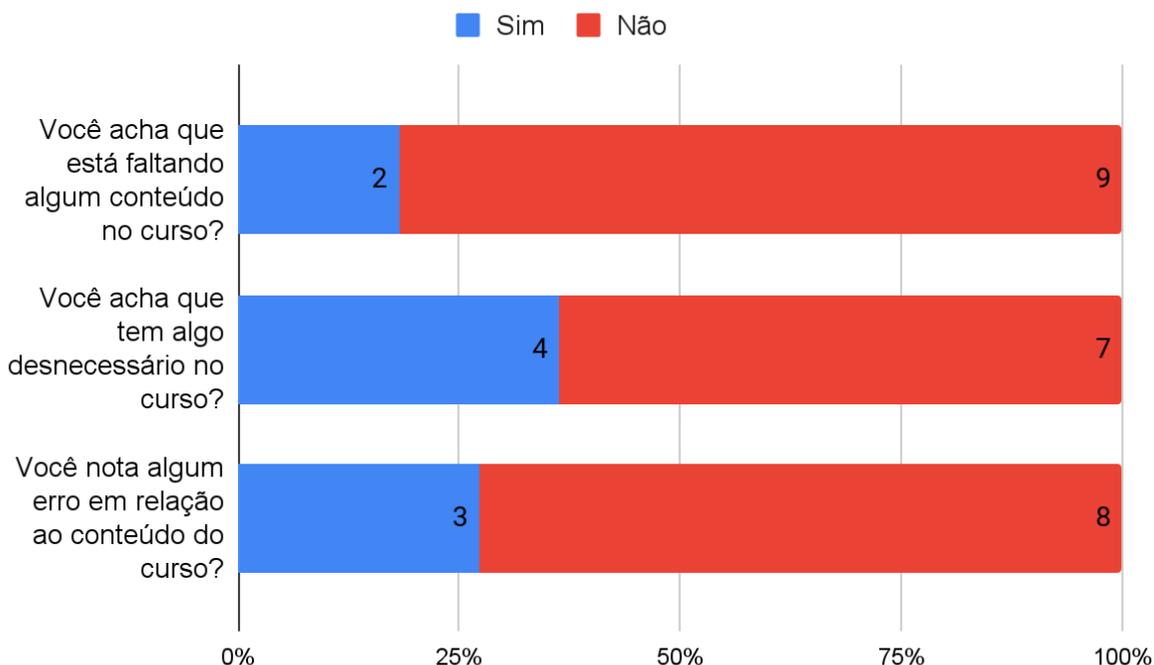


Figura 25: Resultados em relação a adequação funcional e usabilidade do curso.

Se você acha que o curso possui elementos ambíguos ou difíceis de entender, o que seria?
Talvez o cálculo dos pesos de uma rede perceptron (Unidade 3, slide 21) seja um pouco pesado para quando o curso for aplicado sem instrutor.
Muitos conceitos complexos poderiam ser abstraídos ou simplificados de forma a tornar menos formal e quantitativo para algo mais motivacional e qualitativo. Em alguns pontos, me parece, criam uma ideia de muita complexidade e dificuldade (p.ex. redes CNN em detalhes). Poderia ser falado de uma maneira mais geral e funcional sem entrar em detalhes.
Aula 3 - A repetição do quadro sobre a estrutura de um perceptron (slide 12 até slide 17) , eu não sei se não atrapalha do foco que é prestar atenção em cada elemento separado nos próximos 5 slides, eu senti um pouco poluído visualmente. No slide 11, por exemplo, fala dos sinais de entrada e depois disso não aparece mais... Por que os outros parâmetros seriam mais importantes? Aula 3 - Repetição da imagem nos slides 43 e 44, sendo que só vai falar dela a partir do slide 44.
Na aula 3 não precisaria detalhar tanto um perceptron, pois introduz uma complexidade desnecessária.
Não exatamente desnecessário, mas senti que alguns conteúdos poderiam ser mais simplificados e/ou até omissos, principalmente na Aula 3, pensando em não desmotivar alguns estudantes.

Tabela 18: Respostas abertas sobre elementos desnecessários no curso.

5.3.3 Usabilidade

De modo geral, houve boa aceitação em relação ao design e estrutura do curso, havendo duas respostas em discordância, focadas na percepção de

complexidade do conteúdo ministrado na aula de Redes Neurais citada anteriormente (Figura 26).

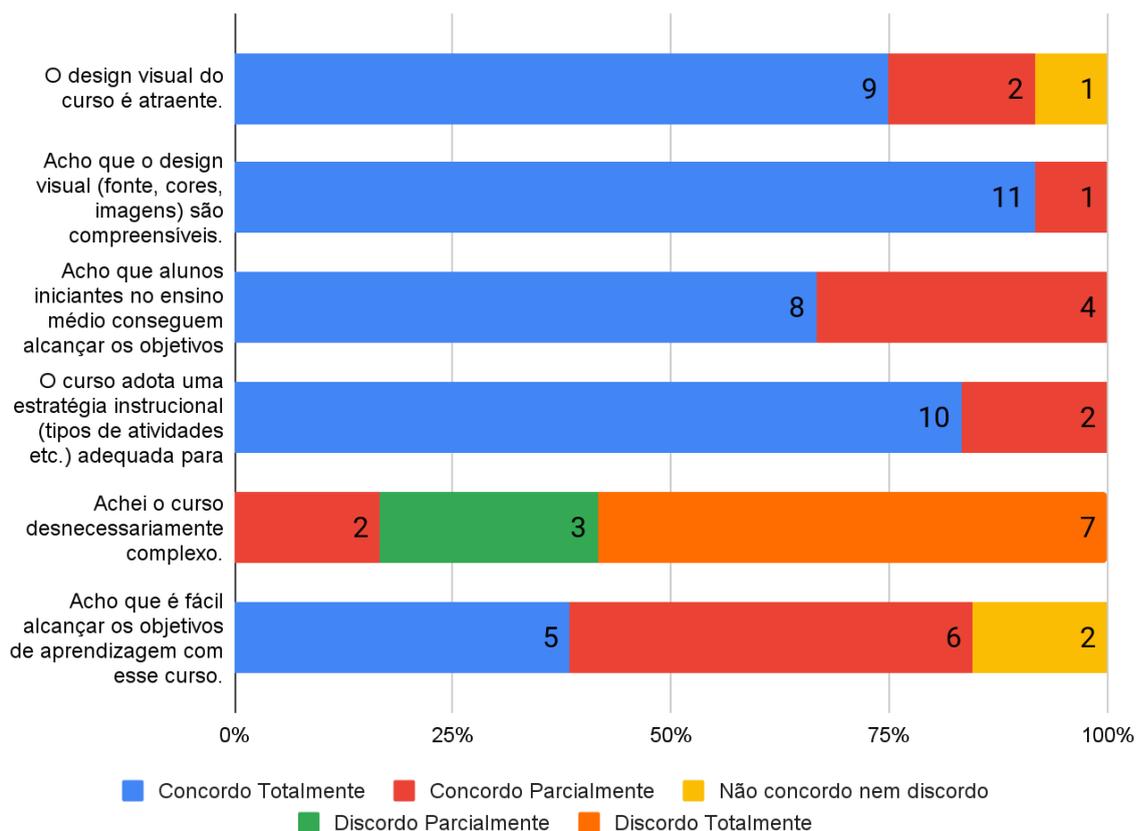


Figura 26: Resultados em relação a usabilidade do curso.

5.3.4 Experiência de aprendizagem

Os resultados em relação à experiência de aprendizagem mostram que os participantes reconhecem que o curso ensina de forma eficiente, se mantém adequadamente desafiador para o público alvo, além de ser bem organizado, relevante e fácil de entender (Figura 27). Todos participantes relatam que os vídeos e animações foram úteis para a compreensão do conteúdo, o que demonstra que esse tipo de material didático pode ter grande valor em cursos on-line. É também possível inferir que o curso possui potencial para ser aplicado em escolas uma vez que todos os respondentes recomendariam o uso do curso para outros professores.

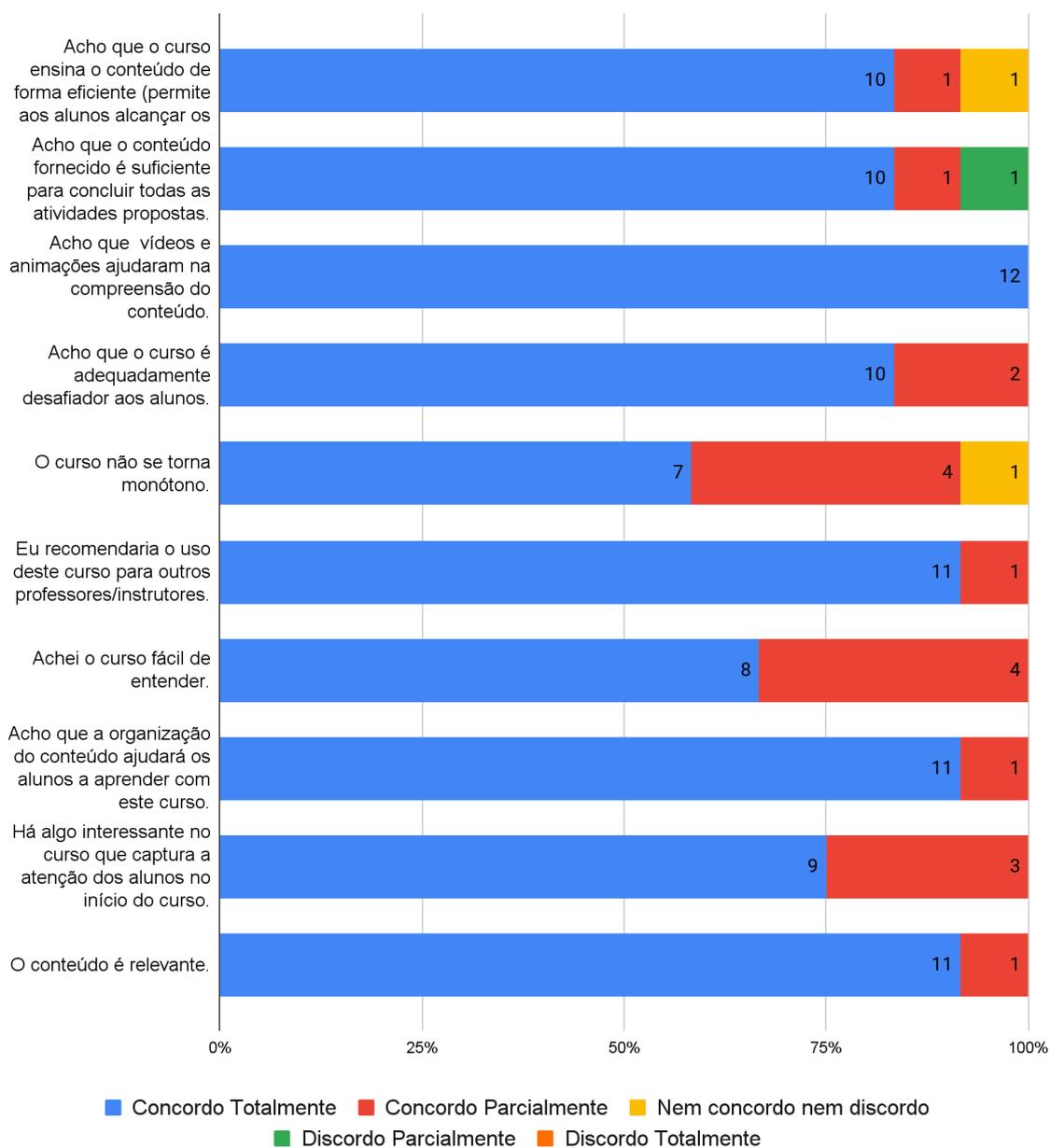


Figura 27: Frequência das respostas referente a experiência da aprendizagem.

5.3.5 Pontos fortes e fracos

Os pontos fortes mais citados são o *design* atrativo e vídeos do curso, além do uso do Jupyter como ferramenta de desenvolvimento de atividades práticas. A explicação matemática do funcionamento de um Perceptron foi citado como ponto positivo por alguns participantes (Tabela 19), embora outros tenham relatado o mesmo ponto como negativo por aumentar a complexidade (Tabela 20),

evidenciando a variedade de opiniões distintas dos participantes do painel que torna a avaliação deste curso bastante rica.

O que você mais gostou no curso?
A forma simples e objetiva como aborda questões complexas como a matemática por trás das redes neurais. É sintetizado e se atém àquilo que importa. Também há elementos gráficos muito agradáveis. Gostei muito dos exemplos e a forma com que leva o aluno em forma de degrau de conhecimento.
Material didático muito bem organizado e bonitinho!
Excelente abrangência do conteúdo. Vídeos com animações atrativas.
Do design dos slides, dos vídeos e roteiros mostrando aplicações práticas, do uso do Vulcan, do banco de imagens, da utilidade para a formação profissional, dos cuidados com a ética.
Gostei da linguagem ter sido adaptada para o público alvo, e do sentimento de acolhimento que o material passa, junto aos vídeos bem explicados; Gostei que o curso não passou correndo por temas como a perpetuação de preconceitos que pode ocorrer se não tivermos cuidado ao criar um modelo de ML.
A apresentação/utilização de forma mais visual dentro do Jupyter facilita muito e agiliza a utilização.
Os vídeos estão muito bons, auxiliando muito na absorção do conteúdo. A interatividade dos recursos do H5P foi bem aproveitada. O plano de ensino e objetivos de aprendizagem estão condizentes com o público alvo. Gostei bastante também do Zêd :) O curso todo no geral está ótimo, parabéns pelo trabalho!
Gostei dos vídeos, ficou bem editado e áudio muito bom. O curso foi direto e claro no conteúdo proposto.
Vídeos para explicação de alguns conceitos.
A forma que cativou o aluno por meio de animações, vídeos, interações, quizzes e também a criação de personagens e histórias que facilitam o entendimento do conteúdo. A forma que assuntos complexos foram simplificados e ensinados foi ótima.
Gostei do empenho em tornar algo complexo em uma atividade (mais ou menos) fácil de entender e não muito "pesada" em matemática.
As atividades práticas e a diversidade de tipos de material (slides, quizzes, vídeos, atividades práticas, etc.).

Tabela 19: Respostas sobre os pontos positivos do curso.

Ainda como pontos negativos, além do uso de matemática no exemplo do Perceptron, foi novamente citada a complexidade da aula 3 sobre Redes Neurais (Tabela 19), além de pequenas sugestões que foram levadas em consideração durante o processo de atualização do material didático.

O que você menos gostou no curso?
O que me preocupa é o tempo para a aplicação do curso, afinal, ele tem muitos conceitos novos, que precisam ser explicados além do exposto nos slides.
Nada, o curso está excelente, parabéns pelo trabalho!
No plano de ensino, como documento isolado, seria melhor que os quizzes utilizados para a avaliação dos tópicos de ensino fossem de alguma forma identificados unicamente ou numerados. No tópico 1 Motivação, e na primeira atividade do tópico 1 (Quiz 1), gostaria de sugerir uma alteração. As câmeras IP do sistema de circuito fechado que tenho instalado em casa já possuem ferramentas de IA embarcadas na própria câmera, como a funcionalidade que identifica a presença de pessoas no ambiente ;-). Com essas câmeras inteligentes, fica mais difícil de separar. Talvez trocar o exemplo.
A Aula 3 achei excessivamente técnica e com detalhes que não me pareceram necessários, pois as tarefas no Vulcan não exigiam todos aqueles conhecimentos.
O exemplo com números ficou um pouco difícil de entender; aula 3.
Para os alunos falta uma parte onde explica melhor o que é, como usar e o que editar no Jupyter, talvez um vídeo ou uma aula extra.
Somente os pontos que comentei nas respostas anteriores.
Não teve algo que não gostei.
Não sei até que ponto ele é adequado aos alunos do ensino médio. Acredito que seja, já que os jovens hoje em dia estão bastante habituados com a linguagem computacional e com a dinâmica das redes sociais e da internet como um todo. Estou bastante curioso quanto à aceitação do conteúdo pelos estudantes.

Tabela 20: Respostas sobre os pontos negativos do curso.

5.4 Discussão

Os resultados da avaliação indicam inicialmente que o curso ML4Teens pode fornecer uma experiência de aprendizagem útil e interessante, tem boa usabilidade e se adequa funcionalmente aos objetivos de aprendizagem, indicando ser uma opção instrucional de valor para o ensino de *ML* e projetos de classificação de imagens para alunos do Ensino Médio. O *design* atrativo, exemplos e os vídeos que fornecem uma forma mais interativa de expor o conteúdo foram citados como os pontos fortes do curso, além do uso do Vulcan dentro do Jupyter Notebook como ferramenta de desenvolvimento de atividades práticas.

Por meio da análise da avaliação do curso, é possível inferir também que abrir a caixa preta para ensinar o funcionamento de redes neurais ainda é um desafio que precisa ser explorado, especialmente em cursos introdutórios onde os

estudantes têm pouco ou nenhum conhecimento prévio dos processos de *ML*. No *ML4Teens*, foi utilizado o recurso de caminhos alternativos como materiais extras para que estudantes que se sintam confortáveis possam explorar mais a fundo temas que foram considerados complexos pelo painel de especialistas.

É importante notar também que não houve coleta de dados sobre a avaliação do desempenho dos aplicantes, visto que o modelo de avaliação de desempenho dos alunos, utilizada para desenvolvimento das atividades dentro do *VisualJupyter*, ainda não foi integrada. Trabalhos futuros com avaliação do desempenho são importantes para verificar se o conteúdo do curso é suficiente para a aprendizagem das atividades de desenvolvimento de modelos de *ML*.

Ameaças à Validade. Os resultados obtidos nesta avaliação devem ser interpretados com cautela, levando-se em consideração potenciais ameaças à sua validade. Devido a falta de uma aplicação no contexto real com estudantes do Ensino Médio (fora do escopo do presente trabalho), os resultados são limitados a fornecer apenas uma primeira indicação sobre a qualidade e adequação do curso *ML4Teens*. Embora os participantes do painel de especialistas foram escolhidos com base no perfil que correspondesse aos usuários alvo, a ausência de feedback de estudantes do Ensino Médio pode afetar os resultados obtidos. Sem participantes do público alvo, há possíveis lacunas de mais conteúdo e explicações que podem ter sido relevadas pelos participantes do painel devido ao seu conhecimento prévio no assunto. A partir disso, é necessário realizar estudos futuros com maior número de participantes com o perfil específico do público alvo.

Devido a natureza do teste do projeto e limitações práticas, o curso foi realizado em ambiente controlado para fins de avaliação e a análise do desempenho dos participantes nas atividades práticas não foram coletadas. Mais estudos de avaliação em contextos educacionais são, portanto, necessários para confirmar os resultados. Visando neutralizar a ameaça de possíveis problemas na definição da medição em si, os questionários foram desenvolvidos decompondo sistematicamente o objetivo da avaliação em itens de questionário que adotam a abordagem GQM (BASILI et al., 1994). Mais estudos sobre a avaliação de desempenho dos alunos em relação aos quizzes e atividades no *Jupyter* podem ser realizados para identificar lacunas, adicionar e corrigir conteúdo e adequar a complexidade dos quizzes.

6 CONCLUSÃO

O presente estudo apresenta o desenvolvimento do curso ML4Teens para o ensino de *ML* com foco em classificação de imagens para alunos do Ensino Médio. Como parte do presente TCC foi sintetizada a fundamentação teórica centralizada nas estratégias instrucionais de *ML* no contexto de Ensino Médio. O estado da arte das unidades instrucionais semelhantes e com o mesmo público alvo foi levantado e suas características foram analisadas. Com base nessas unidades e estratégias, o material instrucional do curso ML4Teens foi desenvolvido, levando-se em conta todos os fatores coletados na revisão bibliográfica do estado da arte. Apresentações, vídeos e quizzes interativos foram criados, assim como todo o conteúdo necessário para que os usuários possam realizar as atividades propostas utilizando a ferramenta Vulcan no Jupyter Notebook. Uma avaliação preliminar do curso foi realizada por meio de um painel de especialistas composto por professores do Ensino Médio e graduação e alunos de graduação. Os resultados obtidos nesta avaliação indicam que o curso ML4Teens pode ter valor no ensino de *ML* com foco em classificação de imagens para alunos do Ensino Médio.

Desta forma, espera-se criar uma contribuição importante para o ensino de *ML* nas escolas brasileiras, buscando facilitar a introdução dos processos de *ML* para alunos do Ensino Médio de forma interativa e on-line. 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.

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. Será possível verificar também, por meio da análise do *feedback* deste público, quais são os pontos fortes do curso para que este possa inspirar futuras iniciativas semelhantes de ensino de *ML* para jovens estudantes e inspirar futuras gerações a ingressar na área de tecnologia. Para dar continuidade, também sugere-se a criação de novos cursos voltados aos estágios de *modify* e *create*, levando o aluno a criação dos seus próprios modelos de classificação de imagens com *ML*.

REFERÊNCIAS

AI4K12. 2020a. Disponível em: <https://github.com/touretzkyds/ai4k12/wiki>. Acesso em: 17 maio. 2021.

AI4K12. Draft Big Idea 3 - Progression Chart. 2020b. Disponível em: https://drive.google.com/file/d/1QL6I_I5cdNTVnYBIZ3_Lxur2DgFjmG_d/view

Amazon. Amazon Machine Learning, AWS Documentation. 2019. Disponível em: <https://docs.aws.amazon.com/machine-learning/latest/dg/building-machine-learning.html>. Acesso em: 24 de out de 2020.

Atlas das Juventudes. Evidências Para A Transformação Das Juventudes (2021).

Disponível em:

<https://atlasdasjuventudes.com.br/wp-content/uploads/2021/06/ATLAS-DAS-JUVENTUDES-COMPLETO.pdf> Acesso em: 21 de Jun de 2021.

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

Bilstrup, K., Kaspersen, M., and Petersen, M. Staging Reflections on Ethical Dilemmas in Machine Learning: A Card-Based Design Workshop for High School Students. In Proc. of the ACM Designing Interactive Systems Conference . Association for Computing Machinery, New York, NY, USA, 2020.

Bhatia, N., Lao, N., Using Transfer Learning, Spectrogram Audio Classification, and MIT App Inventor to Facilitate Machine Learning Understanding. Massachusetts Institute of Technology, Cambridge, MA, USA 2020.

Bhatia, N. Using Transfer Learning, Spectrogram Audio Classification, and MIT App Inventor to Facilitate Machine Learning Understanding. Master thesis, Massachusetts Institute of Technology, Cambridge, MA, USA 2020.

Branch, R. Instructional design: The ADDIE approach. Springer US, 2009.

Burgsteiner, H., Kandlhofer, M., & Steinbauer, G. iRobot: Teaching the Basics of Artificial Intelligence in High Schools. In: Proceedings of the AAI Conference on

Artificial Intelligence, Symposium on Educational Advances in Artificial Intelligence , Phoenix, AZ, United States, 2016.

Cetic.br., NIC.br, CGI.br. Painel TIC COVID-19: Pesquisa sobre o uso da internet no Brasil durante a Pandemia do Coronavírus. 3 Edição. 2020. Disponível em: <https://cetic.br/media/docs/publicacoes/2/20201104182616/painel_tic_covid19_3edicao_livro%20eetr%C3%B4nico.pdf> Acesso em: 21 de Jun de 2021.

Chittora, S. and Baynes, A. Interactive Visualizations to Introduce Data Science for High School Students. In Proc. of the 21st Annual Conference on Information Technology Education. ACM, New York, NY, USA, 2020.

Dos Anjos, C. B., De Moura, O. S., Bigio, N. C. A Percepção Do Ensino De Botânica No Ensino Médio. 2021.

Evangelista, I. Blesio, G. Benatti, E. Why Are We Not Teaching Machine Learning at High School? A Proposal. In Proc. of the World Engineering Education Forum - Global Engineering Deans Council, Albuquerque, NM, USA, 2018.

Estevez, J. G., Garate, G., Graña, M. Gentle Introduction to Artificial Intelligence for High-School Students Using Scratch. IEEE Access, vol. 7, 2019.

Gardner-McCune, C., Touretzky, D., Martin, F., and Seehorn, D. AI for K-12: Making Room for AI in K-12 CS Curricula. In: Proceedings of the 50th ACM Technical Symposium on Computer Science Education. ACM, New York, NY, USA, 1244, 2019.

Google, Introduction to Machine Learning. 2021. Disponível em: <<https://applieddigitalskills.withgoogle.com/c/middle-and-high-school/en/introduction-to-machine-learning/materials.html>>. Acesso em: 24 de mar de 2021.

Gresse von Wangenheim, C. Marques, L.& Hauck, J., Machine Learning for All – Introducing Machine Learning in K-12, socarxivxxx, 2020.

Gresse von Wangenheim, C. Hauck, J. Pacheco, F. S.; Bertoncely Bueno, M. F. Visual Tools for Teaching Machine Learning in K-12: A Ten-Year Systematic Mapping, Education and Information Technologies, 2021.

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.

Howard, J., Guggen, S. Deep Learning for Coders with Fastai and Pytorch. Sebastopol : O'Reilly Media, Incorporated, c2020.

INEP. Dados Do Censo Escolar. 2019. Disponível em:

<http://portal.inep.gov.br/artigo/-/asset_publisher/B4AQV9zFY7Bv/content/dados-do-censo-escolar-noventa-e-cinco-por-cento-das-escolas-de-ensino-medio-tem-acesso-a-internet-mas-apenas-44-tem-laboratorio-de-ciencias/21206> Acesso em: 22 de Jun de 2021.

IFSC. IFSC Analytics: Censo Escolar Covid-19. 2021. Disponível em:

<https://censo.ifsc.edu.br/home_alunos.php> Acesso em: 1 de Jul de 2021.

Judd, S. All Means All: Bringing Project-based, Approachable AI Curriculum to More High School Students through AI4ALL Open Learning. In Proceedings of the 51st ACM Technical Symposium on Computer Science Education . ACM, New York, NY, USA, 1409. 2020.

Julie, H., Alyson, H., & Anne-Sophie, C. Designing digital literacy activities: an interdisciplinary and collaborative approach. In 2020 IEEE Frontiers in Education Conference. IEEE, 2020, 1-5.

Franz, A. C. M. Desenvolvimento de uma Ferramenta Visual de Classificação de Imagens para o Ensino de Machine Learning no Ensino Médio. 2021. Trabalho de Conclusão de Curso. (Graduação em Sistemas de Informação) – Universidade Federal de Santa Catarina. 2021.

Kelleher, C. and Pausch, R. Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers. ACM Computing, 2005.

Kriesel, D., A Brief Introduction to Neural Networks. 2007. Disponível em: <<http://www.dkriesel.com>>. Acesso em: 24 de mar de 2021.

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.

Lin, P., Van Brummelen, J., Lukin, G., Williams, R., & Breazeal, C. Zhorai: Designing a Conversational Agent for Children to Explore Machine Learning Concepts. In Proc. of the AAAI Conference on Artificial Intelligence, AAAI, 2020.

Lindenmaier, D. Herbários Digitais: Uma Experiência No Ensino De Botânica Através Da Arborização E Tecnologias Da Informação E Comunicação (TIC). 2017.

Long, D. Magerko, B. What is AI Literacy? Competencies and Design Considerations. In Proceedings of the CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, 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. Teaching Machine Learning in School: A Systematic Mapping of the State of the Art. Informatics in Education, 19(2), 283-321. 2020a.

Marques, L. & Gresse von Wangenheim, C. & Hauck, J. 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, 2020b.

Markham K., Introduction to Machine Learning with scikit-learn. 2021. Disponível em:

<<https://courses.dataschool.io/introduction-to-machine-learning-with-scikit-learn>>.

Acesso em: 21 de Jun de 2021.

MEC. Base Nacional Comum Curricular. Brasília: MEC. 2017. Disponível em:

<http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518_versaofinal_site.pdf>. Acesso em: 24 de out de 2020.

Mike, K., Hazan, T., and Hazzan, O. Equalizing Data Science Curriculum for Computer Science Pupils. In Proc. of the 20th Koli Calling International Conference on Computing Education Research . ACM, New York, NY, USA, 2020.

Norouzi, N., Chaturvedi, S., & Rutledge, M. Lessons Learned from Teaching Machine Learning and Natural Language Processing to High School Students. In Proc. of the AAAI Conference on Artificial Intelligence, AAAI, 2020.

Nodari, M. P. M., Rosa, E. M., Nascimento, C. R. R., & Guerra, V. M. (2017). Os usos do tempo livre entre jovens de classes populares. *Psicologia: Teoria e Pesquisa*, 32.

O'Hara, K., Blank, D., and Marshall, J. Computational Notebooks for AI Education. In Proc. of the 28th International Florida Artificial Intelligence Research Society Conference. Hollywood, Florida, 2015.

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

Reyes, A. A., Elkin, C., Niyaz, Q., Yang, X., Paheding, S., & Devabhaktuni, V. K. . A Preliminary Work on Visualization-based Education Tool for High School Machine Learning Education. In Proc. of the IEEE Integrated STEM Education Conference. IEEE, 2020, 1-5.

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

Royal Society. 2017. Machine learning: the power and promise of computers that learn by example. Disponível em:

<<https://royalsociety.org/~media/policy/projects/machine-learning/publications/machine-learning-report.pdf>>. Acesso em: 24 de out de 2020.

Rodríguez-García, J. D., Moreno-León, J., Román-González, M., Robles, G. Introducing Artificial Intelligence Fundamentals with LearningML: Artificial Intelligence made easy. In Proc. of the 8th International Conference on Technological Ecosystems for Enhancing Multiculturality. ACM, New York, NY, USA. 2020.

Rodríguez-García, J. D., Moreno-León, J., Román-González, M. and Robles, G. Evaluation of an Online Intervention to Teach Artificial Intelligence with LearningML

to 10-16-Year-Old Students. In Proc. of the 52nd ACM Technical Symposium on Computer Science Education. ACM, New York, NY, USA. 2021.

Romano, A. Vox. 2018. Disponível em:

<<https://www.vox.com/2018/5/15/17351806/is-this-a-pigeon-anime-butterfly-meme-explained>> Acesso em: 3 de Nov de 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. Revista Brasileira de Informática na Educação, 21, 2010.

Santana, O., Braga, G., Braga, J., Carvalho, H. Interactive Model Tool about Center of Mass during COVID-19 Pandemic: A New Learning Path in STEM for K-12 Education. In Proc. of the IEEE International Conference on Teaching Assessment and Learning for Engineering , 2020, 503-508.

Seehorn, D. & Clayborn, L. CSTA K-12 CS Standards for All. In Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education. ACM, New York, NY, USA, 730. 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: 24 de out de 2020.

Salvador, G. C. Desenvolvimento de um Modelo de Avaliação de Aprendizagem de Machine Learning voltado a Classificação de Imagens no Ensino Médio. 2021. Trabalho de Conclusão de Curso. (Graduação em Sistemas de Informação) – Universidade Federal de Santa Catarina. 2021.

Sadouk, L. CNN Approaches for Time Series classification. IntechOpen, 2019.

Shalev-Shwartz, S., Ben-David, S. 2014. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press, USA. 2014.

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

Toivonen, T., Jormanainen, I., Kahila, J., Tedre, M., Valtonen, T., & Vartiainen, H. Co-Designing Machine Learning Apps in K–12 With Primary School Children. In Proc. of the IEEE 20th International Conference on Advanced Learning Technologies , IEEE, 2020, 308-310.

Vartiainen, H., Toivonen, T., Jormanainen, I., Kahila, J., Tedre, M., & Valtonen, T. . Machine learning for middle-schoolers: Children as designers of machine-learning apps. In Proc. of the IEEE Frontiers in Education Conference, IEEE, 2020, 1-9.

Van Brummelen, J., Heng, T., Tabunshchyk, V. Teaching Tech to Talk: K-12 Conversational Artificial Intelligence Literacy Curriculum and Development Tools. Massachusetts Institute of Technology, Cambridge, MA, US. 2020a.

Van Brummelen, J. Tools to Create and Democratize Conversational Artificial Intelligence, M.S. thesis, Elect. Eng. Comput. Sci., Massachusetts Inst. of Technol., Cambridge, MA, USA, 2019.

Wan, X., Zhou, X., Ye. Z., Mortensen, C. and Bai, Z. SmileyCluster: supporting accessible machine learning in K-12 scientific discovery. In Proc. of the Interaction Design and Children Conference . Association for Computing Machinery, New York, NY, USA, 2020.

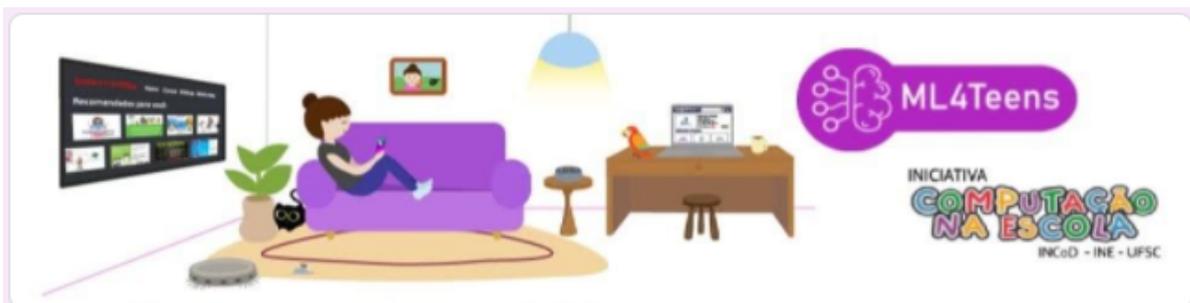
Weintrop, D. Block-based Programming in Computer Science Education. Communications of the ACM, 62(8), 2019.

APÊNDICES

Apêndice 1: Tabela relacionando as UIs com foco no Ensino Médio encontradas por Marques et al. (2020a).

Estudo	Referência	Modo de Aplicação	Ano
(Marques et al., 2020a)	(Apps for Good, 2019a)	Presencial	2019
	(Burgsteiner et al., 2016)(Burgsteiner, 2016)	Presencial	2016
	(Essinger & Rosen, 2019)	Presencial	2019
	(Evangelista et al., 2018)	Presencial	2018
	(MIT App Inventor, 2019)	Presencial	2019
	(Mobasher et al., 2019)	Presencial	2019
	(Narahara & Kobayashi., 2018)	Presencial	2018
	(ReadyAI, 2019)	On-line	2019
	(Sperling & Lickerman, 2012)	Presencial	2012
	(Tang, 2019)(Tang et al., 2019)	Presencial	2019
	(Vachovsky et al., 2016)	Presencial	2016
	(Van Brummelen, 2019) (Van Brummelen & Abelson, 2018) (Van Brummelen et al., 2019)	Presencial	2019
	(Zhu, 2019)	Presencial	2019
	(Zimmermann- Niefeld et al., 2019a)(ZimmermannNiefeld et al., 2019b)	Presencial	2019
Presente estudo	(Bhatia e Lao, 2020) (Bhatia, 2020)	Presencial	2020
	(Bilstrup et al., 2020)	Presencial	2020
	(Chittora e Baynes, 2020)	--	2020
	(Google, 2021)	On-line	2021
	(Gresse von Wangenheim et al., 2020)	Presencial/On-line	2020
	(Mike et al., 2020)	Presencial	2020
	(Norouzi et al., 2020)	Presencial	2020
	(Rodríguez-García et al. 2021) (Rodríguez-García et al. 2020)	Remoto	2021
	(Van Brummelen et al., 2020)	Remoto	2020
	(Wan et al., 2020)	Presencial	2020

Apêndice 2: Questionário de avaliação do curso.



Avaliação do curso ML4Teens

Você está sendo convidado(a) a participar da avaliação do curso ML4Teens que tem como objetivo ensinar os conceitos básicos de Machine Learning com foco em classificação de imagens para alunos do ensino médio.

O curso é voltado a iniciantes no nível de "use/modify". Ele foi projetado para ser aplicado de forma remota tanto em forma de turma com instrutor (via Moodle) ou usando o material H5P de forma online individual (a ser futuramente disponibilizado pelo site da Computação na Escola).

O teste faz parte do TCC de Jacqueline Cardozo do SIN/INE/UFSC dentro da iniciativa Computação na Escola do INCoD/INE/UFSC. O tempo médio esperado para realização da avaliação do curso é de 1 hora.

Para visualizar as aulas do ML4Teens, basta acessar o Moodle do curso:

<https://moodle.ufsc.br/course/view.php?id=82751>

Solicitamos que neste primeiro momento, não compartilhe os materiais do curso. Uma versão final, aprimorada com base em seus comentários, ficará publicamente disponível.

Depois que você analisar o curso disponibilizado no Moodle, favor responder o questionário de avaliação.

Qualquer dúvida, favor entrem em contato com Jacqueline: jacquocard@gmail.com

Muito obrigada pela sua ajuda! O seu feedback é essencial para que possamos melhorar a qualidade dos nossos cursos!

 ml4teenscomputacao@gmail.com (not shared) [Switch account](#) 

* Required

Seu nome *

Your answer

Você é: *

- aluno de graduação
- professor no ensino médio
- professor no ensino superior
- Other:

Utilidade

Acho o curso útil para ensinar Machine Learning no ensino médio. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Acho que em sua forma atual, o curso atende objetivos de aprendizagem de ML no ensino médio. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Adequação funcional

Você acha que está faltando algum conteúdo no curso? *

Sim

Não

Se está faltando algum conteúdo no curso, qual?

Your answer

Você acha que tem algo desnecessário no curso? *

Sim

Não

Se você acha que tem algo desnecessário no curso, o que seria?

Your answer

Se você acha que tem algo desnecessário no curso, o que seria?

Your answer

Você nota algum erro em relação ao conteúdo do curso? *

Sim

Não

Se você nota algum erro em relação ao conteúdo do curso, o que seria?

Your answer

Usabilidade

Você acha que o curso possui elementos ambíguos ou difíceis de entender? *

- Sim
- Não

Se você acha que o curso possui elementos ambíguos ou difíceis de entender, o que seria?

Your answer

O design visual do curso é atraente. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Acho que o design visual (fonte, cores, imagens) são compreensíveis. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Acho que alunos iniciantes no ensino médio conseguem alcançar os objetivos de aprendizagem. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

O curso adota uma estratégia instrucional (tipos de atividades etc.) adequada para um curso online. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Achei o curso desnecessariamente complexo. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Acho que é fácil alcançar os objetivos de aprendizagem com esse curso. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Experiência de aprendizagem

Acho que o curso ensina o conteúdo de forma eficiente (permite aos alunos alcançar os objetivos com o esforço, tempo e recursos adequados). *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Acho que o conteúdo fornecido é suficiente para concluir todas as atividades propostas. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Acho que vídeos e animações ajudaram na compreensão do conteúdo. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Acho que o curso é adequadamente desafiador aos alunos. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

O curso não se torna monótono. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Eu recomendaria o uso deste curso para outros professores/instrutores. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Achei o curso fácil de entender. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Acho que a organização do conteúdo ajudará os alunos a aprender com este curso. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Há algo interessante no curso que captura a atenção dos alunos no início do curso. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

conteúdo é relevante. *

- Concordo totalmente
- Concordo parcialmente
- Não concordo nem discordo
- Discordo parcialmente
- Discordo totalmente

Pontos fortes e fracos

O que você mais gostou no curso?

Your answer

O que você menos gostou no curso?

Your answer

Mais alguma sugestão?

Your answer

Apêndice 3: Artigo científico originado do presente estudo.

ML4Teens - Introduzindo *Machine Learning* para alunos do Ensino Médio

Jacqueline Cardozo, Christiane Gresse von Wangenheim

Departamento de Informática e Estatística, Universidade Federal de Santa Catarina,
Florianópolis, SC, Brasil

j.cardozo@grad.ufsc.br, c.wangenheim@ufsc.br

***Abstract.** Understanding Machine Learning has become increasingly important for societies that aim to train High School students with the proposed challenges to face the digital transformations. There are few initiatives to teach ML for this audience. The present work aims to help K-12 students, more specifically High School students, to understand the fundamentals of Machine Learning through an online course.*

***Resumo.** O entendimento sobre Machine Learning tem se tornado cada vez mais importante para as sociedades que almejam formar estudantes do Ensino Médio com as competências necessárias para enfrentar os impactos das transformações digitais. Existem poucas iniciativas de ensino de ML voltadas para esse público. O presente trabalho visa ajudar os alunos da Escola Básica, mais especificamente estudantes do Ensino Médio, a entenderem fundamentos de Machine Learning por meio de um curso on-line.*

1. Introdução

A Inteligência Artificial (IA) tem cada vez mais se tornado parte do cotidiano das pessoas. Assim como a Educação Básica abrange as áreas clássicas do conhecimento, como linguagens e matemática, a alfabetização em IA e computação tem se tornado uma grande questão a ser indagada [Burgsteiner et al., 2016]. Essa discussão do que os alunos da Educação Básica devem aprender sobre computação tem impulsionado o desenvolvimento de diversas diretrizes curriculares [Marques et al., 2020a]. A Sociedade Brasileira de Computação (SBC) enfatiza a importância do ensino de pensamento computacional e IA para alunos do Ensino Médio, provendo a base necessária para entender grandes questões éticas envolvidas em IA [SBC, 2017]. O Ministério da Educação (MEC) enfatiza que grande parte das futuras profissões envolverá, direta ou indiretamente, computação e tecnologias digitais, fazendo-se necessário abordar o tema na Educação Básica a fim de formar cidadãos com os valores e habilidades necessárias para enfrentar os impactos das transformações digitais [MEC, 2017].

No mundo todo, a população jovem se desenvolve em meio à tecnologia - computadores, celulares e outros dispositivos - que permitem que eles usem estes dispositivos digitais para se comunicar, se divertir, realizar pesquisas on-line, etc. [Resnick, 2012]. Embora tenham muito contato com o mundo digital, essa geração de jovens, em grande parte, não possui real fluência no uso de tais tecnologias para desenvolver seus próprios artefatos

[Resnick, 2012], assim como também usar essas ferramentas virtuais para entender e transformar o mundo [SBC, 2017].

O entendimento de Machine Learning (ML) pode encorajar estudantes a considerar carreiras nas áreas de computação e promover uma preparação dos alunos para o ensino superior [Marques et al., 2020a] O Ensino Médio é um ótimo momento para que os estudantes entrem em contato com IA, pois estão maduros o suficiente para encarar problemas mais complexos [Estevez et al., 2019].

Enquanto iniciativas de ensino de IA estão sendo propagadas no Ensino Médio dos Estados Unidos, China e Grã Bretanha [Marques et al., 2020a], no Brasil há ainda poucos trabalhos sobre a aplicação de cursos de programação com o objetivo de ensinar programação fora do ensino superior [Scaico et al., 2010][Marques et al., 2020b]. O ensino de IA é sugerido apenas como itinerário formativo - parte flexível do currículo ou atividade optativa - no novo currículo de Educação Básica brasileiro definido pela Base Nacional Comum Curricular [MEC, 2017]. A SBC defende que o ensino de IA provê a base necessária para uma discussão sobre a interação Homem-Máquina e suas implicações éticas [SBC, 2017], sendo condizente com a previsão do BNCC quanto à necessidade de preparar os alunos para enfrentar os impactos das transformações digitais.

Com base nos resultados de uma revisão sistemática de literatura [CARDOZO, 2022] identificou-se que atualmente existem muito poucos cursos e iniciativas educacionais especificamente voltados a esse estágio educacional e menos ainda em português voltado a escolas brasileiras. A análise também mostrou que as UIs para o público do Ensino Médio seguem uma tendência de utilizar aprendizagem supervisionada, tarefas de classificação e visão computacional como domínio de aplicação.

O objetivo geral deste projeto é o de desenvolver um curso on-line para ensinar *ML* em escolas brasileiras de Ensino Médio. O curso será projetado de forma que permita a sua adoção no formato interdisciplinar, inserida em conteúdo programado do Ensino Médio.

2. O Curso ML4teens

O curso de ML4Teens 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.

2.1. Análise de Contexto

O Ensino Médio brasileiro corresponde ao período de três anos que finaliza a Educação Básica. A faixa etária dos estudantes do Ensino Médio tipicamente é de 15 a 18 anos. Os estudantes desta faixa são fluentes em língua portuguesa e possuem conhecimento básico em pelo menos uma língua estrangeira, como o inglês ou espanhol. Espera-se que os estudantes do Ensino Médio consolidem os conhecimentos adquiridos no ensino fundamental, obtendo preparo básico para o mercado de trabalho em termos das disciplinas “tradicionalis” [MEC, 2017]. O Ensino Médio é a etapa dos estudos em que os jovens sentem mais “peso”, exigindo determinação e esforço para estudar. O percentual de jovens estudantes que buscam

um futuro em áreas STEM costuma ser baixo. A promoção de orientação, informação e estímulo por parte dos professores nas escolas pode ser uma estratégia importante para apoiar a decisão de mais alunos buscarem aperfeiçoamento nessas áreas [Atlas das Juventudes, 2021]. A grande maioria dos alunos não têm conhecimento de computação e/ou inteligência artificial. Alguns talvez aprendam em cursos extracurriculares, já que neste momento este conteúdo ainda não faz parte do BNCC. Em 2018, o Brasil contava com 28.673 escolas que ofertavam o Ensino Médio. Aproximadamente 60% dessas escolas possuem disponibilidade de infraestrutura e recursos tecnológicos como laboratório de informática e internet [INEP, 2019]. O acesso à internet está disponível em 95,1% das escolas que ofertam Ensino Médio.

2.2. Objetivos de Aprendizagem

Os objetivos de aprendizagem do curso ML4Teens foram desenvolvidos com base nas competências de alfabetização de IA de Long e Magerko (2020) e as diretrizes do AI4K12 [Touretzky et al. 2019]. Os objetivos foram selecionados considerando o contexto educacional e a base necessária para conhecimento dos fundamentos de IA/ML e introduzindo conceitos para a realização da tarefa de classificação de imagens da atividade final do curso em concordância com a multidisciplinaridade junto ao tema de Botânica [MEC, 2017].

Tabela 1. Objetivos de aprendizagem do curso ML4Teens.

Categoria	Objetivo de aprendizagem	Referência(s)
OA1. Conceitos básicos de ML	Identificar exemplos de Machine Learning e diferenciá-lo da aprendizagem humana	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]
OA2 .Redes neurais	Compreender a estrutura de uma rede neural e descrever como suas partes formam um conjunto de funções que computam uma saída capaz de identificar padrões em dados	3-A-ii 3-5, 3-B-i 6-8, 3-B-ii 3-5 [AIK12, 2020]
OA3. Gerenciamento de dados	Preparar um conjunto de dados usado para treinar um modelo de ML considerando o tamanho do conjunto de dados, a forma com que os dados foram coletados e rotulados, além de sua qualidade (equilíbrio, balanceamento, viés)	3-C-ii 9-12 [AI4K12, 2020]; 11, 12 [Long e Magerko, 2020]; 1A-DA-05 [CSTA, 2017]
OA4. Treinamento de modelo de ML	Treinar um modelo de ML para classificação/predição usando um algoritmo de aprendizagem supervisionada com dados reais e ajustando os parâmetros de treinamento	3-A-ii 9-12, 3-A-iii 9-12 [AI4K12, 2020]
OA5. Processo de ML	Compreender e aplicar as etapas envolvidas no Machine Learning e suas práticas e desafios	3-A-iv 9-12 [AI4K12, 2020]; 9 [Long e Magerko, 2020]
OA6. Ética de ML	Identificar e descrever diferentes questões éticas acerca de ML (privacidade, viés introduzido por características dos dados de treinamento, etc.)	3-C-iii 6-8 [AI4K12, 2020]; 3A-AP-24 [CSTA, 2017]; 13, 16 [Long e Magerko, 2020]
OA7. Impactos do IA/ML	Identificar prós e contras de IA e ML para atividades cotidianas e opções de carreira atuais e futuras	2-IC-21 [CSTA, 2017]; 6 [Long e Magerko, 2020]
OA8. Biologia/ Botânica	EM13CNT206 / Discutir a importância da preservação e conservação da biodiversidade [MEC, 2017] por meio do reconhecimento de espécies de árvores nativas de Santa	[MEC, 2017]

2.3. Design do Curso

Alinhado ao contexto e aos objetivos de aprendizagem, o curso foi projetado para iniciantes para ensinar a tarefa de classificação de imagens utilizando *Deep Learning*. Escolheu-se a tarefa de classificação de imagens por ser uma das mais simples e introdutória e *Deep Learning* por ser uma das mais adotadas atualmente.

Prevê-se uma aplicação interdisciplinar deste curso, pela falta de inclusão explícita da área de conhecimento de IA/ML atualmente na BNCC. Assim, o curso será aplicado à área de conhecimento de biologia referente ao tópico de biodiversidade, políticas ambientais e ecossistemas [MEC, 2017] reforçando a aprendizagem de conceitos de botânica. Os alunos serão estimulados a exercitarem o entendimento sobre as diferentes espécies de árvores endêmicas e nativas de Santa Catarina a fim de realizar um projeto de classificação de imagens de seis espécies de árvores presentes no campus da Universidade Federal de Santa Catarina. Espera-se que os estudantes consigam identificar os diferentes tipos de espécies nas imagens para que possam catalogá-las corretamente.

A partir da aplicação do curso dentro do Ensino Médio, levando em consideração que ainda não faz parte da BNCC, projetou-se um curso de curta duração, que pode ser aplicado a distância (considerando a situação da pandemia Covid-19). Ele também poderá ser aplicado na forma presencial, em sala de aula, sem a necessidade da assistência de professores graduados na área de computação, uma vez que o curso fornecerá todo o conteúdo necessário para o entendimento do tema, assim como avaliação de aprendizagem dos alunos.

2.4. Plano de Ensino

O plano de ensino foi desenvolvido levando em consideração principalmente as diretrizes de AI4K12 [Touretzky et al. 2019] e as competências de Long e Magerko (2020) para levantar os objetivos do curso, levando em conta o estágio escolar dos alunos do Ensino Médio. Também foram adicionadas atividades extras com o objetivo de estimular os estudantes a modificarem os códigos por trás do projeto de classificação de imagens, assim como a introdução do conteúdo necessário para entender o básico de algoritmos e uso de bibliotecas como o Fastai [Howard, 2020]. Para as atividades práticas, foi introduzida a ferramenta VisualJupyter, também conhecido como Vulcan [Franz, 2021].

Tabela 2. Plano de ensino do curso ML4Teens.

Aula	Conteúdo	Objetivos de Aprendizagem	Método Instrucional	Material Instrucional	Avaliação
Introdução					
1. (15min)	Motivação, Aplicações de IA no cotidiano e Conceito de ML	OA1, OA7, OA8	Apresentação interativa, Vídeos, Demonstrações	Slides 1, "Rápido, desenhe!"	Quizzes
Conceitos Básicos					
2.	Aprendizagem	OA1, OA3,	Apresentação,	Slides 2,	Quizzes

(30min)	Supervisionada e Classificação de imagens	OA4, OA5	Demonstrações, Vídeos, Atividade no Vulcan	Vulcan (Franz, 2021)	
3. (30min)	Redes neurais	OA1, OA2	Apresentação, Vídeos	Slides 3, "Digit Recognizer"	Quizzes
Processo de ML					
4.1. (15min)	Preparação dos Dados (Limpeza, Classificação e Pré-processamento)	OA3, OA5, O8	Apresentação, Atividades	Slides 4, Vulcan (Franz, 2021)	Quizzes
4.2. (30min)	Treinamento, Avaliação do Modelo (Aprendizado por Transferência e Ajuste Fino)	OA4, OA5, O8	Apresentação, Atividade no Vulcan	Slides 4, Vulcan (Franz, 2021)	Quizzes, rubrica (Salvador, 2021)
4.3. (15min)	Predição	OA5, O8	Apresentação, Atividade no Vulcan	Slides 4, Vulcan (Franz, 2021)	Quizzes, rubrica (Salvador, 2021)
Implicações Éticas e Impactos					
5. (15min)	Responsabilidade, Defeitos, Preconceito e Desinformação	OA6	Apresentação, Vídeos	Slides 5, "Máquina Moral"	Quizzes
6. (15min)	Impactos e Oportunidades	OA7	Apresentação, Vídeos	Slides 6	Quizzes

2.5. Material Didático

O material didático desenvolvido é composto por apresentações interativas (Figura 1), vídeos com conteúdo imersivo, como a demonstração do Vulcan (Figura 2) e introduzindo conceitos mais desafiadores ou que demandam mais exemplos (Figura 3).



Figura 1. Ilustração com exemplos de uso de IA no cotidiano.

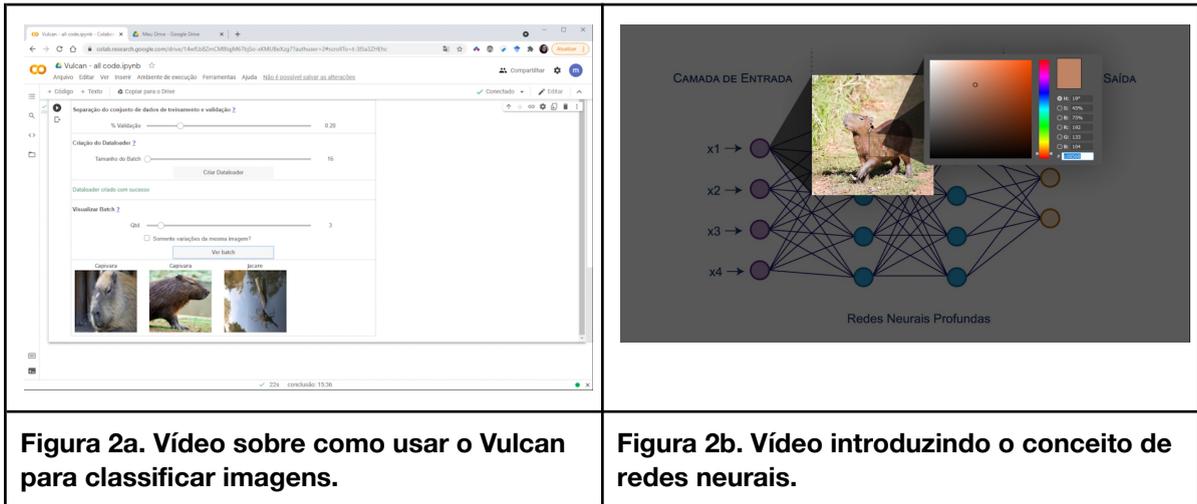


Figura 2a. Vídeo sobre como usar o Vulcan para classificar imagens.

Figura 2b. Vídeo introduzindo o conceito de redes neurais.

O design do curso foi desenvolvido com foco no público jovem, dispondo de ilustrações vetoriais, personagens exclusivos, memes e *emojicons*. O curso também possui narrativas como a de Zëd (Figura 4), o alienígena que tem como objetivo mapear a vida terrestre, que convida os estudantes a desenvolverem os modelos de classificação de imagens de capivaras e jacarés e posteriormente, da classificação de árvores nativas e endêmicas. Esse personagem foi definido com base numa enquete com representantes do público alvo.



Figura 3. Zëd personifica o popular meme baseado no anime *Brave of the Sun Fighbird* (Romano, 2018).

As aulas também possuem quizzes interativos (Figura 5) intercalados aos slides. Os quizzes possuem tipos variados, como “arraste e solte” onde o estudante é convidado a completar termos em frases relacionados ao conteúdo da aula; perguntas com resposta única; múltipla escolha e respostas abertas (Tabela 4).



Figura 4a: Exemplo de quiz do tipo “escolha única”.

Figura 4b. Exemplo do recurso externo “Rápido, desenhe!”

Alguns links externos com atividades interativas são disponibilizados para que os alunos explorem recursos fora do curso. Um exemplo é a aplicação “Rápido, desenhe!” (Figura 6), onde os estudantes podem brincar com um modelo treinado para reconhecer desenhos simples, dando a oportunidade aos alunos de entrarem em contato com dados rotulados.

O curso on-line foi disponibilizado via moodle UFSC. Ele é dividido em aulas formatadas na plataforma HP5, que permitem a apresentação de todo material do curso, assim como a parte interativa (vídeos e quizzes) e será disponibilizado em 2022 no site de cursos da iniciativa Computação na Escola/INCoD/INE/UFSC.

2.6. Avaliação da Aprendizagem

A avaliação da aprendizagem dos alunos do curso é realizada por meio de quizzes (Tabela 3) interativos em cada módulo do curso. Para a avaliação de desempenho a partir do trabalho final criado como resultado da aprendizagem no Jupyter Notebook (Tabela 5), é usado a rubrica do modelo de avaliação de Salvador (2021).

Tabela 3. Plano de avaliação referente aos objetivos e competências.

ID	Critério de avaliação	Avaliação	Fonte
Introdução			
A1	Diferenciar entre como pessoas e computadores aprendem	Quizzes	3-A-i 3-5 [AI4K12, 2020]
A2	Reconhecer que existem muitos jeitos diferentes de pensar e desenvolver máquinas “inteligentes”. Identificar uma variedade de tecnologias que usam IA, incluindo sistemas cognitivos, robótica e ML	Quizzes	3 [Long & Magerko, 2020]
A3	Distinguir entre artefatos tecnológicos que usam ou não IA	Quizzes	1 [Long & Magerko, 2020]
A4	Reconhecer e descrever exemplos de como computadores raciocinam e fazem decisões	Quizzes	8 [Long & Magerko, 2020]
Conceitos Básicos			
A5	Entender como aprendizagem supervisionada identifica padrões em dados rotulados	Quizzes	3-A-ii 3-5 [A4IK12, 2020]

A6	Ilustrar a estrutura de uma rede neural e descrever como suas partes formam um conjunto de funções que computam uma saída	Quizzes	3-B-i 6-8 [AI4K12, 2020]
A7	Demonstrar como pesos são designados em uma rede neural para produzir o comportamento desejado de entrada e saída	Quizzes	3-B-ii 3-5 [AI4K12, 2020]
Processo de ML			
A8	Ilustrar o que acontece em cada etapa necessária ao usar ML para construir um classificador ou preditor	Quizzes, rubrica (Salvador, 2021)	3-A-iv 9-12 [AI4K12, 2020]
A9	Entender as etapas envolvidas em Machine Learning e suas práticas e desafios	Quizzes, rubrica (Salvador, 2021)	9 [Long & Magerko, 2020]
Implicações Éticas e Impactos			
A10	Entender que os dados não devem ser considerados pelo seu valor bruto e requerem interpretação; Entender como os conjuntos de treinamento podem afetar os resultados de um algoritmo	Quizzes	13 [Long & Magerko, 2020]
A11	Comparar os prós e contras associados às tecnologias de computação que afetam as atividades cotidianas das pessoas e as opções de carreira	Quizzes	2-IC-21 [CSTA, 2017]
A12	Identificar tipos de problemas que IA se sobressai e problemas que são mais difíceis para a mesma. Usar essa informação para determinar quando é apropriado usar IA e quando usar habilidades humanas	Quizzes	5 [Long & Magerko, 2020]

Os quizzes foram formulados de acordo com os critérios de avaliação e objetivos de aprendizagem e são disponibilizados imediatamente após a exposição do conteúdo. O sistema utilizado para a administração do curso permite a inserção de questionários de diversos tipos, como Múltipla Escolha, onde o aluno pode selecionar mais de uma resposta correta dentre algumas opções, Escolha Única, Arrastar e Soltar para completar frases etc.

Tabela 4. Quizzes inseridos como atividades interativas utilizando a ferramenta H5P (<https://h5p.org/>).

Id	Pergunta	Tipo	Critério de Avaliação
Introdução			
Q1	Usa inteligência artificial?	Múltipla Escolha	A3
Q2	Ligue a tarefa ao método ML em que ela se encaixa:	Arrastar e soltar	A2
Q3	Selecione exemplos de como os computadores raciocinam para tomam decisões:	Múltipla Escolha	A4
Conceitos Básicos			
Q4	Arraste o termo ao espaço correto sobre os processos de aprendizagem supervisionada:	Arrastar e soltar	A5
Q5	Como os humanos aprendem melhor do que os computadores?	Escolha Única	A1
Q6	Qual deve ser o valor dos pesos w_1 e w_2 para que o resultado indique para o banco se ele deve conceder o empréstimo ou não, considerando a tabela abaixo?	Respostas Abertas	A7

Q7	Arraste o termo de acordo com a sua descrição correta sobre a estrutura de uma rede neural:	Arrastar e soltar	A6
Q8	Arraste o termo de acordo com a sua descrição correta sobre as funções usadas em redes neurais:	Arrastar e soltar	A6
Processo de ML			
Q9	Podemos considerar que esse conjunto de dados possui qualidade nos seguintes quesitos:	Escolha Única	A9
Q10	Sobre a etapa de Transfer Learning, arraste os termos para os espaços corretos:	Arrastar e soltar	A9
Q11	Calcule o desempenho do modelo de acordo com as seguintes informações:	Respostas Abertas	A9
Q12	Junte o termo a sua respectiva descrição sobre a etapa de ajuste fino:	Arrastar e soltar	A9
Q13	Combine a etapa do processo de ML com a sua respectiva descrição:	Arrastar e soltar	A9
Implicações Éticas e Impactos			
Q14	Ligue a categoria do problema a sua respectiva descrição sobre os principais problemas éticos envolvidos no desenvolvimento em ML:	Arrastar e soltar	A9
Q15	Baseado em tudo que você aprendeu sobre preconceito, feedback loops e desinformação, você considera que os modelos de machine learning:	Escolha Única	A10
Q16	Complete com o termo relacionado a sua descrição correta sobre os tipos de problemas resolvidos por humanos e máquinas:	Arrastar e soltar	A12

Para a avaliação do desempenho em relação às principais atividades práticas do curso do processo de ML é adotado a rubrica de Salvador (2021), alinhados aos objetivos de aprendizagem do curso ML4Teens e a ferramenta Vulcan de Franz (2021). Essa rubrica (Tabela 5) pode ser utilizada manualmente pelo instrutor em aplicações em turmas do curso. A sua aplicação de forma automatizada depende ainda de uma integração do modelo de avaliação do Salvador (2021) dentro do ambiente visual Vulcan.

Tabela 5. Rubrica de Salvador (2021) alinhada aos objetivos do curso ML4Teens.

ID	Critério	Níveis de desempenho		
		Baixo - 0 pt	Aceitável - 1 pt.	Aceitável - 1 pt.
Preparação de dados (OA3)				
C1	Quantidade de imagens	Menos de 5 imagens por categoria	6 de 10 imagens por categoria	Mais de 10 imagens por categoria
C2	Distribuição do conjunto de dados	Quantidade de imagens por categoria varia muito	Quantidade de imagens por categoria varia pouco	Todas as categorias possuem a mesma quantidade de imagens
Preparação de dados/Botânica (OA3/OA8)				
C3	Rotulagem das imagens	Menos de 20% das imagens rotuladas corretamente	De 20% a 99% das imagens rotuladas corretamente	Todas as imagens rotuladas corretamente
Treinamento de modelo de ML/Transfer Learning e Fine-Tuning (OA4)				

C4	Treinamento - <i>Transfer Learning</i>	O modelo não foi treinado (<i>transfer learned</i>)	O modelo foi treinado com os parâmetros padrão	O modelo foi treinado com parâmetros ajustados (arquitetura, época e taxa de aprendizagem)
C5	Treinamento - Ajuste Fino	O modelo não sofreu o ajuste fino	Foi feito <i>unfreeze</i> das camadas e melhor taxa de aprendizagem não encontrada ou modelo não treinado	Foi feito <i>unfreeze</i> das camadas, a melhor taxa de aprendizagem foi encontrada e o modelo sofreu o ajuste fino
Avaliação e interpretação do desempenho de um modelo de ML (Transfer Learning e Fine-Tuning) (OA5)				
C6	Interpretação de acurácia	Categorias com baixa acurácia não identificadas	Categorias com baixa acurácia identificadas e interpretação incorreta em relação ao modelo	Categorias com baixa acurácia identificadas corretamente e interpretação correta em relação ao modelo
C7	Interpretação da matriz de confusão	Classificações incorretas não identificadas e interpretação incorreta em relação ao modelo	Classificações incorretas identificadas e interpretação incorreta em relação ao modelo	Classificações incorretas identificadas e interpretação correta em relação ao modelo
C8	Ajustes/melhorias feitas	Sem novas iterações de desenvolvimento	Uma nova iteração com alterações no conjunto de dados e/ou parâmetros de treinamento	Diversas novas iterações com alterações no conjunto de dados e/ou parâmetros de treinamento
Avaliação e interpretação do desempenho de um modelo de ML/Testar e aperfeiçoar programas (OA5)				
C9	Testes com novos objetos	Nenhum novo objeto testado	1-2 novos objetos testados	Mais de dois novos objetos testados
C10	Interpretação dos testes	Interpretação errada	---	Interpretação correta

3. Avaliação do Curso

Para avaliar a qualidade do curso ML4Teens, foi realizada uma avaliação preliminar com o objetivo de analisar a qualidade do curso desenvolvido em termos de utilidade, adequação funcional, experiência de aprendizagem e usabilidade do ponto de vista dos especialistas no contexto do ensino de *Machine Learning*, por meio de questionário desenvolvido com base na ISO/IEC 25010 (2011), ISO/IEC 9241 (1998), TAM [Davis, 1989] e SUS [BROOKE, 1996]. A avaliação do curso foi feita por meio de teste de usuário disponibilizado a um painel de especialistas. Este teste tem como objetivo avaliar a percepção dos participantes do painel em relação aos critérios de qualidade definidos. O teste consiste em analisar o curso, assim como seus materiais, vídeos e links externos, por meio da plataforma Moodle, disponibilizada pela Universidade Federal de Santa Catarina.

A partir dos fatores de qualidade foi projetado o questionário de avaliação publicado no Google Forms. Dados demográficos são coletados para a análise das diferentes perspectivas dos membros do painel em relação a adequação do curso. O formulário é dividido em cinco sessões, com perguntas que utilizam a escala Likert de 5 pontos e respostas abertas para permitir sugestões de melhorias e apontamento de erros. O

questionário é finalizado com perguntas de respostas abertas sobre os pontos fortes e fracos do curso.

3.2. Execução da Avaliação

A avaliação do curso foi realizada por uma painel de especialistas, representados por professores do Ensino Médio, do ensino superior, de ambos e alunos da graduação. Devido a falta de competências em *ML* de alunos do Ensino Médio atualmente, foram convidados alunos da graduação e pós-graduação para integrarem o painel. Ao todo, foram convidados 12 participantes por email, dos quais todos participaram na pesquisa (100% taxa de resposta). Foi definido um período de coleta de avaliações de duas semanas para dar tempo necessário para os membros do painel conseguirem realizar as atividades do curso, que tem como duração média de duas horas e meia. A avaliação ocorreu durante as últimas semanas de Novembro de 2021.

3.3. Análise da Avaliação

Os dados foram analisados de acordo com os fatores de qualidade definidos. A análise e inferência sobre as respostas da avaliação estão sintetizadas a seguir.

3.3.1. Utilidade do curso ML4Teens

Em relação a utilidade do curso, todos os participantes concordam que ele é útil para o ensino de *ML* para estudantes do Ensino Médio e atende aos objetivos de aprendizagem propostos.

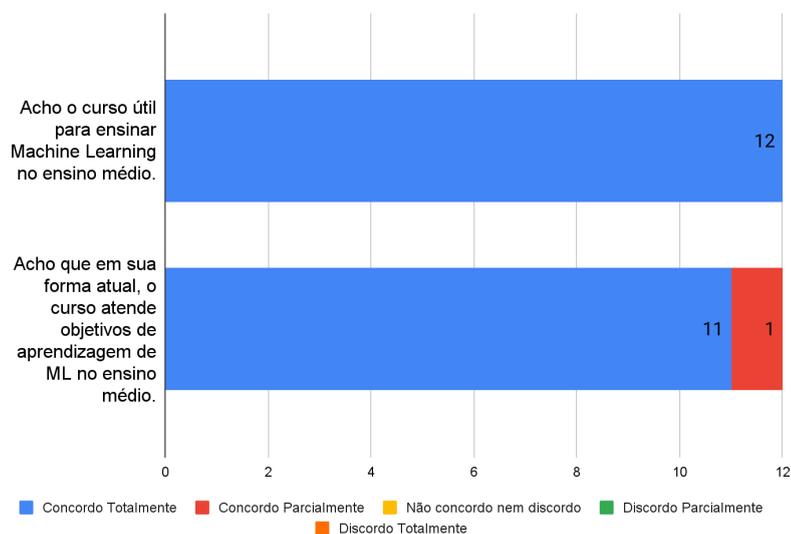


Figura 5: Frequência das respostas referente a experiência da aprendizagem.

3.3.2. Adequação funcional

Em relação a funcionalidade do curso, de forma geral os participantes o acharam adequado. Poucos participantes relataram que há elementos faltantes no curso e alguns erros foram reportados e corrigidos. Sobre tópicos desnecessários e difíceis de entender, apenas a aula sobre Redes Neurais foi citada, sendo que a maioria participantes demonstraram

preocupação em relação à complexidade do conteúdo ministrado, particularmente sobre os conceitos de Perceptron e Redes Neurais Convolucionais, o que pode desmotivar os alunos mais iniciantes. Como solução a esses pontos, a aula sobre Redes Neurais foi atualizada para possuir um conteúdo mais simplificado e abstraído e ainda permitindo que estudantes interessados possam acessar os conteúdos mais complexos na forma de aulas extras.

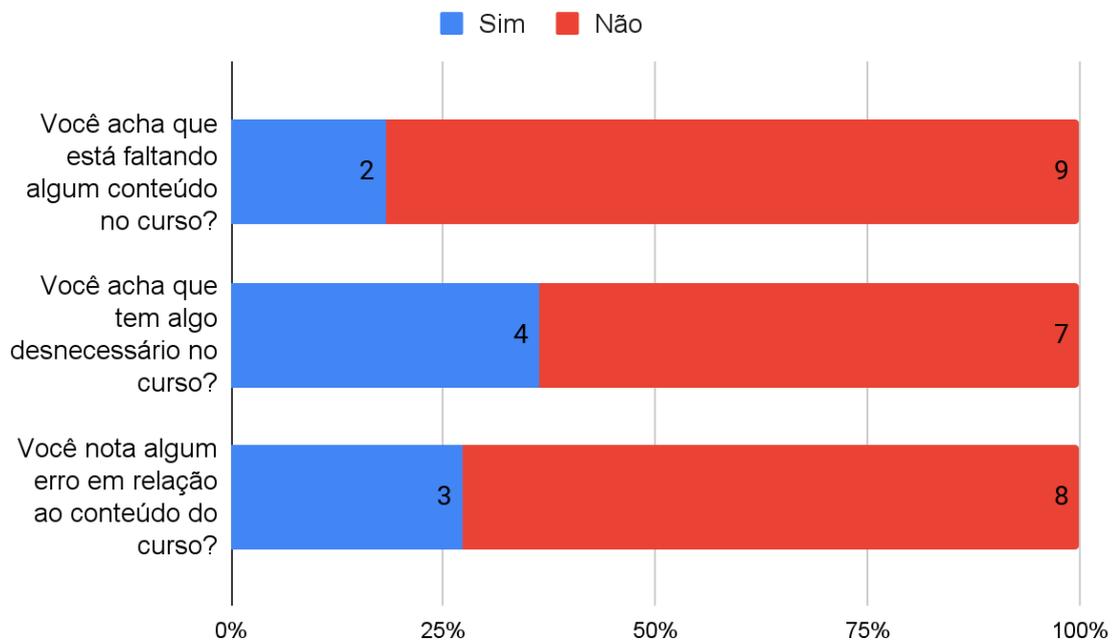


Figura 6: Resultados em relação a adequação funcional e usabilidade do curso.

3.3.3. Usabilidade

De modo geral, houve boa aceitação em relação ao design e estrutura do curso, havendo algumas respostas em discordância, focadas na percepção de complexidade do conteúdo ministrado na aula de Redes Neurais citada anteriormente.

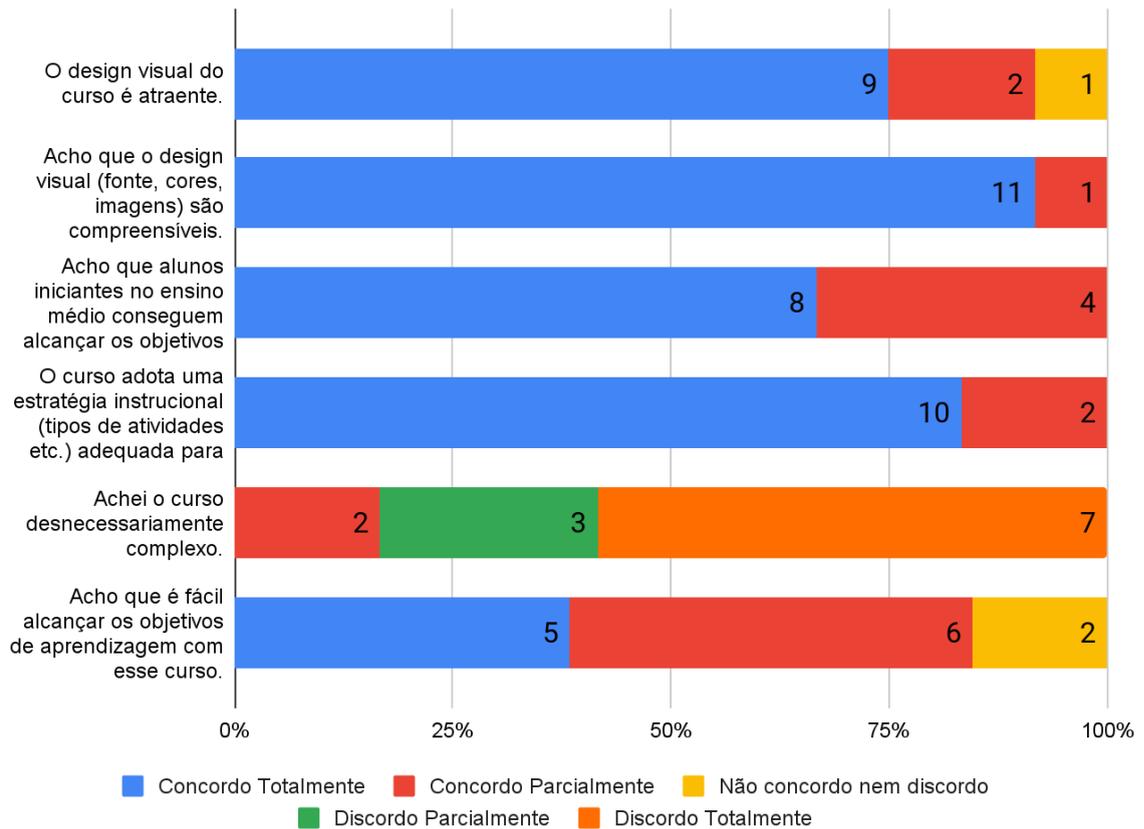


Figura 7: Resultados em relação a usabilidade do curso.

3.3.4. Experiência de aprendizagem

Os resultados em relação à experiência de aprendizagem mostram que os participantes reconhecem que o curso ensina de forma eficiente, se mantém adequadamente desafiador para o público alvo, além de ser bem organizado, relevante e fácil de entender. Todos participantes relatam que os vídeos e animações foram úteis para a compreensão do conteúdo, o que demonstra que esse tipo de material didático pode ter grande valor em cursos on-line. É também possível inferir que o curso possui potencial para ser aplicado em escolas uma vez que todos os respondentes recomendariam o uso do curso para outros professores.

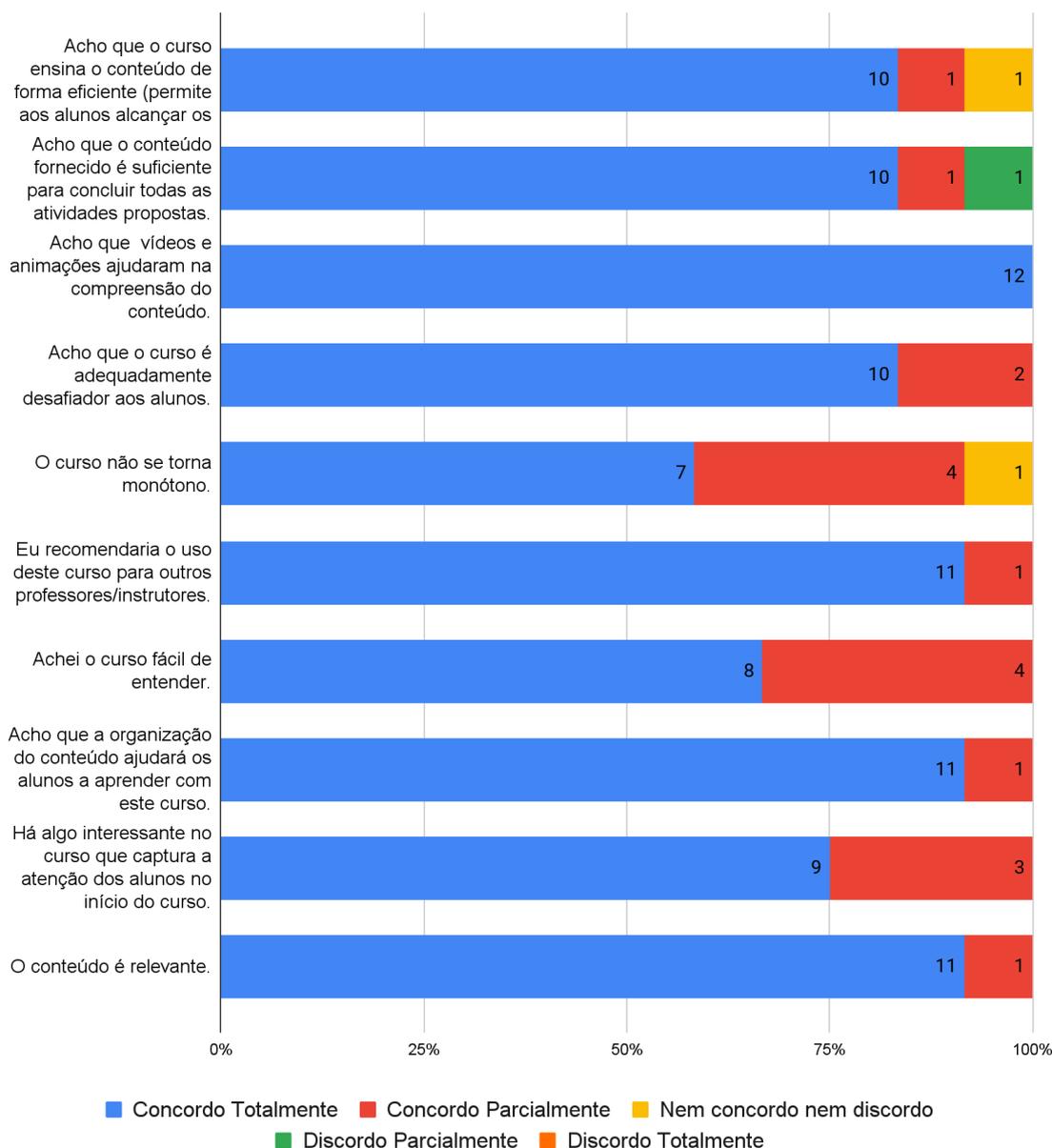


Figura 8: Frequência das respostas referente a experiência da aprendizagem.

3.3.5. Discussão

Os resultados da avaliação indicam inicialmente que o curso ML4Teens pode fornecer uma experiência de aprendizagem útil e interessante, tem boa usabilidade e se adequa funcionalmente aos objetivos de aprendizagem, indicando ser uma opção instrucional de valor para o ensino de ML e projetos de classificação de imagens para alunos do Ensino Médio. O design atrativo, exemplos e os vídeos que fornecem uma forma mais interativa de expor o conteúdo foram citados como os pontos fortes do curso, além do uso do Vulcan dentro do Jupyter Notebook como ferramenta de desenvolvimento de atividades práticas.

Por meio da análise da avaliação do curso, é possível inferir também que abrir a caixa preta para ensinar o funcionamento de redes neurais ainda é um desafio que precisa ser

explorado, especialmente em cursos introdutórios onde os estudantes têm pouco ou nenhum conhecimento prévio dos processos de ML. No ML4Teens, foi utilizado o recurso de caminhos alternativos como materiais extras para que estudantes que se sintam confortáveis possam explorar mais a fundo temas que foram considerados complexos pelo painel de especialistas.

É importante notar também que não houve coleta de dados sobre a avaliação do desempenho dos aplicantes, visto que a aplicação utilizada para desenvolvimento das atividades não possui ainda esta funcionalidade. Trabalhos futuros com avaliação do desempenho são importantes para verificar se o conteúdo do curso é suficiente para o desenvolvimento das atividades de desenvolvimento de modelos de ML.

Ameaças à Validade. Os resultados obtidos nesta avaliação devem ser interpretados com cautela, levando-se em consideração potenciais ameaças à sua validade. Devido a falta de uma aplicação no contexto real com estudantes do Ensino Médio (fora do escopo do presente trabalho), os resultados são limitados a fornecer apenas uma primeira indicação sobre a qualidade e adequação do curso ML4Teens. Embora os participantes do painel de especialistas foram escolhidos com base no perfil que correspondesse aos usuários alvo, a ausência de feedback de estudantes do Ensino Médio pode afetar os resultados obtidos. Sem participantes do público alvo, há possíveis lacunas de mais conteúdo e explicações que podem ter sido relevadas pelos participantes do painel devido ao seu conhecimento prévio no assunto. A partir disso, é necessário realizar estudos futuros com maior número de participantes com o perfil específico do público alvo.

Devido a natureza do teste do projeto e limitações práticas, o curso foi realizado em ambiente controlado para fins de avaliação e a análise do desempenho dos participantes nas atividades práticas não foram coletadas. Mais estudos de avaliação em contextos educacionais são, portanto, necessários para confirmar os resultados. Visando neutralizar a ameaça de possíveis problemas na definição da medição em si, os questionários foram desenvolvidos decompondo sistematicamente o objetivo da avaliação em itens de questionário que adotam a abordagem GQM (BASILI et al., 1994). Mais estudos sobre a avaliação de desempenho dos alunos em relação aos quizzes e atividades no Jupyter podem ser realizados para identificar lacunas, adicionar e corrigir conteúdo e adequar a complexidade dos quizzes.

4. Conclusão

O presente estudo apresenta o desenvolvimento do curso ML4Teens para o ensino de *ML* com foco em classificação de imagens 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. O material instrucional do curso ML4Teens foi desenvolvido, levando-se em conta principalmente as diretrizes de AI4K12 [Touretzky et al. 2019] e as competências de Long e Magerko (2020). Apresentações, vídeos e quizzes interativos foram criados, assim como todo o conteúdo necessário para que os usuários possam realizar as atividades propostas utilizando a ferramenta Vulcan no Jupyter Notebook. Uma avaliação preliminar do curso foi realizada por meio de um painel de

especialistas composto por professores do Ensino Médio e graduação e alunos da graduação. Os resultados obtidos nesta avaliação indicam que o curso ML4Teens pode ter valor no ensino de *ML* com foco em classificação de imagens para alunos do Ensino Médio.

Desta forma, espera-se criar uma contribuição importante para o ensino de *ML* nas escolas brasileiras, buscando facilitar a introdução dos processos de *ML* para alunos do Ensino Médio de forma interativa e on-line. Destaca-se também o caráter inédito do presente curso, visto que atualmente existem muito poucos cursos especificamente voltados a esse estágio educacional e menos ainda em português voltado a escolas brasileiras.

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. Será possível verificar também, por meio da análise do *feedback* deste público, quais são os pontos fortes do curso para que este possa inspirar futuras iniciativas semelhantes de ensino de *ML* para jovens estudantes e inspirar futuras gerações a ingressar na área de tecnologia. Para dar continuidade, também sugere-se a criação de novos cursos voltados aos estágios de *modify* e *create*, levando o aluno a criação dos seus próprios modelos de classificação de imagens com *ML*.

Referências

AI4K12. 2020a. Disponível em: <https://github.com/touretzkyds/ai4k12/wiki>. Acesso em: 17 maio. 2021.

AI4K12. Draft Big Idea 3 - Progression Chart. 2020b. Disponível em: https://drive.google.com/file/d/1QL6I_I5cdNTVnYBIZ3_Lxur2DgFjmG_d/view

Atlas das Juventudes. Evidências Para A Transformação Das Juventudes (2021). Disponível em:

<<https://atlasdasjuventudes.com.br/wp-content/uploads/2021/06/ATLAS-DAS-JUVENTUD-ES-COMPLETO.pdf>> Acesso em: 21 de Jun de 2021.

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, United States, 2016.

Cardozo, J. Desenvolvimento de um Curso On-line para o Ensino de Machine Learning voltado à Classificação de Imagens no Ensino Médio. 2021. Trabalho de Conclusão de Curso. (Graduação em Sistemas de Informação) – Universidade Federal de Santa Catarina. 2022.

Estevez, J. G., Garate, G., Graña, M. Gentle Introduction to Artificial Intelligence for High-School Students Using Scratch. IEEE Access, 7, 2019.

Howard, J., Gugger, S. Deep Learning for Coders with Fastai and Pytorch. Sebastopol : O'Reilly Media, Incorporated, c2020.

INEP. Dados Do Censo Escolar. 2019. Disponível em:

<http://portal.inep.gov.br/artigo/-/asset_publisher/B4AQV9zFY7Bv/content/dados-do-censo

-escolar-noventa-e-cinco-por-cento-das-escolas-de-ensino-medio-tem-acesso-a-internet-mas-
-apenas-44-tem-laboratorio-de-ciencias/21206> Acesso em: 22 de Jun de 2021.

Franz, A. C. M. Desenvolvimento de uma Ferramenta Visual de Classificação de Imagens para o Ensino de Machine Learning no Ensino Médio. 2021. Trabalho de Conclusão de Curso. (Graduação em Sistemas de Informação) – Universidade Federal de Santa Catarina. 2021.

Long, D. Magerko, B. What is AI Literacy? Competencies and Design Considerations. In Proc. of the CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, 2020.

Marques, L. & Gresse von Wangenheim, C. & Hauck, J. Teaching Machine Learning in School: A Systematic Mapping of the State of the Art. *Informatics in Education*, 19(2), 283-321. 2020a.

Marques, L. & Gresse von Wangenheim, C. & Hauck, J. 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, 2020b.

MEC. Base Nacional Comum Curricular. Brasília: MEC. 2017. Disponível em: <http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518_versaofinal_site.pdf>. Acesso em: 24 de out de 2020.

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

Romano, A. Vox. 2018. Disponível em: <<https://www.vox.com/2018/5/15/17351806/is-this-a-pigeon-anime-butterfly-meme-explained>> Acesso em: 3 de Nov de 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. *Revista Brasileira de Informática na Educação*, 21, 2010.

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: 24 de out de 2020.

Salvador, G. C. Desenvolvimento de um Modelo de Avaliação de Aprendizagem de Machine Learning voltado a Classificação de Imagens no Ensino Médio. 2021. Trabalho de Conclusão de Curso. (Graduação em Sistemas de Informação) – Universidade Federal de Santa Catarina. 2021.

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.

Weintrop, D. Block-based Programming in Computer Science Education. *Communications of the ACM*, 62(8), 2019.