



- 1. Estabelecer quais os procedimentos**, mesmo que seja na forma genérica, para atingir o resultado esperado
- 2. A primeira etapa é o pré-processamento:** Por exemplo, um conjunto de dados é composto por figuras 28x28 pixels em escala de cinzas, isto é, cada pixel possui um valor entre 0 e 255. Para utiliza-las no TensorFlow, o valor dos pixels é dividido pelo valor máximo, 255, de tal forma que os pixels passam a ter valor entre 0 e 1.
- 3. A segunda etapa é separar o dataset** entre conjunto de treino e teste. Essa divisão é fundamental para garantir o processo de generalização da técnica de IA. Utilizar o mesmo conjunto de treino e teste irá contaminar sua solução e consiste em grave falha na metodologia para solucionar o problema
- 4. A terceira etapa é construir a rede neural artificial**, ou seja, definir sua topologia - quantas camadas compõe a rede, qual o tamanho de cada camada, quais funções matemáticas são utilizadas, etc.
- 5. O quarto passo é o treinamento em si.** De forma geral, os frameworks possuem diferentes algoritmos que podem ser utilizados para realizar essa etapa e muitas vezes as diferenças, além do desempenho, estão na complexidade e quantidade de memória utilizada por estes algoritmos
- 6. O quinto passo é avaliar os resultados obtidos** que podem ser feitos por meio de métricas estatísticas como acurácia, especificidade, etc, ou por meio de funções específicas de perda que medem quão diferente a solução obtida pela rede neural é da solução esperada. Se o resultado for satisfatório de acordo com seus requisitos o processo é finalizado, caso contrário, deve alterar a topologia e parâmetros da rede neural para iniciar um novo processo de treinamento

Suposição: Objetivo de criar uma IA capaz de classificar figuras em diferentes categorias, sendo necessário extrair informações destas figuras. Para isso o TensorFlow auxiliará neste processo

Utilização do TensorFlow

Frameworks para IA

1. Definição de Framework

- São abstrações de software que provêm funcionalidade genérica para um determinado contexto e/ou aplicação
 - São estruturas de código que contêm funções básicas para suportar uma determinada estrutura
 - Vão além do código e dispõe como deve ser organizado os dados, parâmetros e até mesmo a lógica/sequência de execução de modelos e ferramentas
- Fica à cargo do programador incluir as funções e informações complementares para que um objetivo seja alcançado

TensorFlow

- Framework utilizado para diferentes aplicações em IA
 - O principal foco é no processo de treinamento e inferência de redes neurais profundas (Deep Neural Networks)
 - No geral, a maior parte das aplicações em Tensor Flow envolvem imagem e/ou vídeo, que são artefatos compostos por uma grande quantidade de informações
 - Essas grande quantidade torna o processamento extremamente custoso em arquiteturas de processamento sequencial
 - Hardwares e arquiteturas de processamento / computação paralela são voltadas para o desenvolvimento de soluções em IA.
- Criado em 2015 pela equipe Google Brain (Google)
 - Em paralelo ao desenvolvimento do framework, Google passa a desenvolver a partir de 2015 as unidades de processamento Tensor (Tensor Processing Unit, TPU)
 - São circuitos integrados de aplicação específica (Application-Specific Integrated Circuit, ASIC) utilizados para acelerar o processamento de aplicações de IA
 - Google continuou o processo de desenvolvimento deste hardware lançando em 2016 a primeira versão, em 2017 a segunda, 2018 a terceira e 2021 a quarta geração de TPUs.
 - Hoje as principais placas de vídeo da nVidia utilizam na arquitetura a quarta geração de tensor cores (512 tensor cores no modelo RTX 4090)