

Objetivo do curso: compreensão básica de modelos de aprendizado de máquina.

1 O que é Machine Learning?

É uma área relativamente antiga que é estudada há várias décadas

O que chamou atenção recentemente foram as performances em algumas tarefas interessantes.

Uma delas foi a análise de imagens.

Convolutional Neural Network foi usado em escala pela primeira vez em 2012, o que mostrou sua performance. Desde lá, sua performance apenas melhorou. O **DeepLearning** está passando a performance humana em alguns casos na análise de imagens.

Outra área que o *Machine Learning* foram os jogos. As máquinas conseguem entender e performar melhor que um humano nos jogos. Um dos melhores jogadores de *Go* perdeu para um humano.

Como realmente funciona?

Nós vamos ensinar uma máquina a aprender.

Damos exemplos para ela (amostras de dados), e o que gostaríamos que ela concluísse com os dados.

Para treinar, damos vários exemplos: esse conjunto de dados x_1 resulta nesse y_1 , para x_2 , y_2 , e assim por diante.

Para cada x_i que dermos para ele, ele deve retornar o respectivo y_i ; Porém, se passarmos um conjunto que não foi contemplado, ele deve ser capaz de prever qual será o resultado.

Queremos que ele aprenda os parâmetros do modelo matemático e prever o que foi pedido.

1.1 Regressão Lógica

Queremos um training set que consiga aprender um modelo e consegue prever o resultado dado um conjunto de dados.

Para fazer o *learning* temos um algoritmo que é feito com vários parâmetros e *learning* significa que gostaríamos de inferir quais parâmetros desse modelo são consistentes com nossos dados de treinamento.

Vamos considerar um dos algoritmos mais básicos: **Logistic Regression**

O objetivo do **Machine Learning** é que, dado N exemplos, data x e outcome y , gostaríamos de construir modelo preditivo que é capaz de prever y dado x .

Linear Predictive Model: X_{i1} é o primeiro componente do vetor X , X_{i2} o segundo e assim por diante até X_{iM} .

Vamos multiplicar cada componente do vetor X por um parâmetro e somamos um bias:

$$(b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \dots + (b_M \times x_{iM}) + b_0$$

Isso é um mapeamento dos dados X_i para um número Z_i .

Muitas vezes é melhor dar uma chance se vai chover ou não em um dia do que afirmar algo. Para fazer isso, usamos uma **Logistic Function** notado por σ :

$$p(y_i = 1|x_i) = \sigma(z_i)$$

z_i : multiplicação dos parâmetros dos dados X com os parâmetros b_1, b_2 até b_M

Essa função, a **Sigmoid Function** $p(y_i = 1|x_i) = \sigma(z_i)$, sempre está entre 0 e 1. Quando z_i é grande, como 5 ou 6, a função converte ele para um número perto de 1. Quando é pequeno, -1, -2, -4, converte para perto de 0.

Sigmoid Function é uma maneira de converter previsões para uma perspectiva probabilística.

Os parâmetros b dizem o quão importante as variáveis são para a predição.

É um modelo bem simples; é apenas uma combinação linear de multiplicação das variáveis observados pelos parâmetros associados, somando-os, mapeando eles para uma variável z_i e, então, executando-os por meio de uma função Sigmoid Function.

O coração do machine learning é: temos um modelo paramétrico que é caracterizado por um conjunto de parâmetros que queremos aprender. A maneira em que fazemos o aprendizado é ter um conjunto de dados e, para esses dados, temos parâmetros X e resultado Y . Gostaríamos de aprender os parâmetros do nosso modelo de tal forma que as previsões do modelo sejam consistentes com os dados do treinamento.

O que queremos dizer com *Learning* é inferir os parâmetros B_0 até B_M que nos forneçam saídas de mapeamento de X para Y consistente com os dados.

Os conceitos básicos da **Logistic Regression** são bastante usados em Deep Learning.