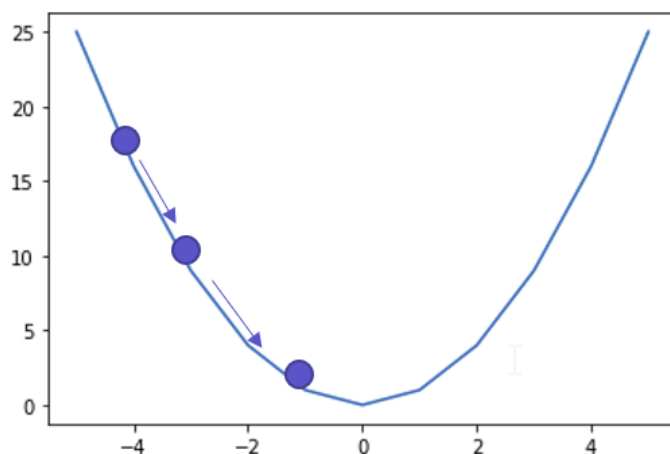


Inteligência Artificial E Machine Learning

SEGUNDO SEMESTRE DE 2022

Prof: Arthur Rocha
prof.arthur.rocha@gmail.com

Questão 01



No algoritmo do gradiente descendente. Como devemos atualizar os parâmetros do modelo?

α Step	θ_n Parâmetros do modelo no instante n	$F(\theta)$ Função de custo	∇ Gradiente
------------------	---	--------------------------------	-----------------------

- A. $\theta_n = \theta_n - \alpha \nabla F(\theta)$
- B. $\alpha_n = \alpha_n - \theta \nabla F(\theta)$
- C. $\theta_{n+1} = \theta_n - \alpha \nabla F(\theta)$
- D. $\theta_{n+1} = \theta_n + \alpha \nabla F(\theta)$
- E. $\theta_{n+1} = \theta_n - \nabla F(\theta)$

Questão 02

Uma loja de roupas (que também as fabrica) está planejando lançar uma nova linha de camisas e pretende aumentar as vendas dessa nova linha. Na linha anterior, eles tiveram boas vendas, mas perceberam que uma parte dos clientes não estava muito satisfeita com os tamanhos disponíveis (P, M, G). Por sorte, essa loja tem vendedores muito espertos que coletaram os dados de tamanhos e medidas um pouco mais precisas desses clientes insatisfeitos (e dos satisfeitos também). Com uma boa base de dados de medidas de clientes, qual poderia ser uma alternativa mais viável para a nova linha? (**Escolha uma alternativa e explique porque**).

- A. Permanecer exatamente como está, afinal as vendas já foram boas.
- B. Utilizar os dados de medidas coletados para:
 - 1. Segmentar os clientes pelo tamanho, agrupando-os em 3 a 5 possíveis tamanhos (usando algum algoritmo não-supervisionado de clustering como a k-means), onde cada tamanho vai refletir uma média das medidas desses grupos identificados.
 - 2. E depois fabricar todos os tamanhos igualmente para cobrir a demanda da clientela.
- C. Utilizar os dados de medidas coletados para:
 - 1. Segmentar os clientes pelo tamanho, agrupando-os em 3 a 5 possíveis tamanhos (usando algum algoritmo não-supervisionado de clustering como k-means), onde cada tamanho vai refletir uma média das medidas desses grupos identificados.
 - 2. E depois fabricar cada tamanho de camisa em quantidades levando em consideração a possível demanda de cada grupo (proporção de pessoas em cada grupo).
- D. Utilizar algum algoritmo de aprendizado de máquina supervisionado para conduzir o processo criativo e de fabricação das camisas.
- E. Utilizar um algoritmo de aprendizado de máquina supervisionado para a partir dos dados rotulados (cliente satisfeito / insatisfeito) aprender a reconhecer a insatisfação de possíveis novos clientes acerca da qualidade da roupa.

Questão 03

No seu trabalho, você foi alocado para um projeto que envolve Visão Computacional. É um problema relativamente complexo e tentativas anteriores de utilizar algoritmos de machine learning clássicos não foram muito bem sucedidas. O objetivo é realizar a detecção e classificação de uma variedade de objetos em imagens. Pela sua expertise, a empresa decidiu te alocar nesse projeto e utilizar Deep Learning para uma prova de conceito. Em um primeiro momento, custo computacional não é um grande problema (afinal, é uma prova de conceito), mas uma boa acurácia de classificação é necessária. Que tipo de modelo de Aprendizado de Máquina você experimentaria primeiro?

- A. Redes Neurais Recorrentes
- B. Redes Neurais Convolucionais
- C. Support Vector Machines
- D. Redes Neurais Perceptron Multicamadas
- E. Ensembles - Random Forests, Gradient Boosting...

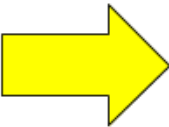
Questão 04

Um treinador esportivo de alta performance está realizando um estudo com seus atletas. Ele está coletando de todas as atividades físicas realizadas e do desempenho físico final de cada atleta (como uma variável quantitativa contínua). Depois de ter coletado muitos dados de seus vários atletas executando diversas variações de treinos, esse treinador gostaria de criar um modelo que explicasse com maiores detalhes qual os melhores treinos para alcançar os resultados desejados. Que modelo seria esse dentre as opções. (O treinador está buscando explicabilidade).

- A. Random Forests
- B. Redes Neurais
- C. Regressão Linear
- D. Regressão Logística
- E. Deep Learning

Questão 05

Com relação ao pré-processamento de dados para input de modelos de Machine Learning.



Color
Red
Red
Yellow
Green
Yellow

Red	Yellow	Green
1	0	0
1	0	0
0	1	0
0	0	1

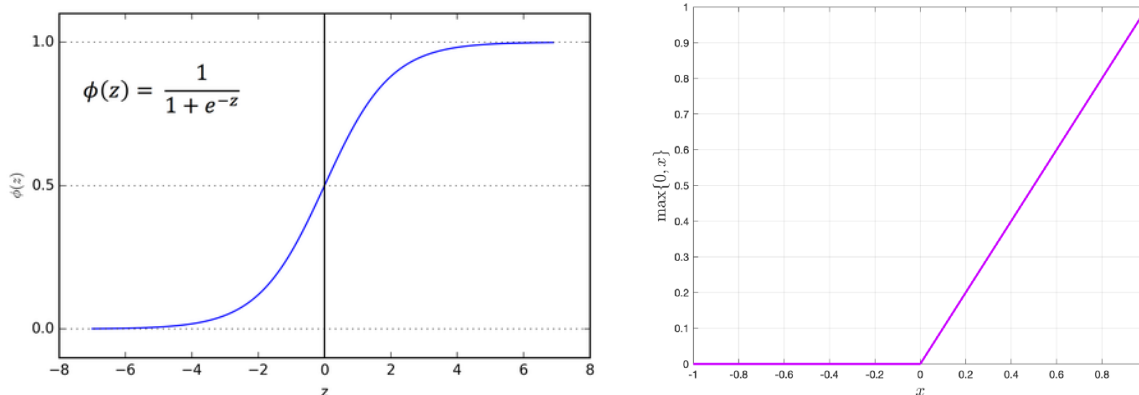
O processo de transformação de variável categórica para binária numérica demonstrado na figura acima é chamado de?

Questão 06

Com relação ao split de dados para treinamento de modelos de machine learning, é **INCORRETO** afirmar:

- A. O conjunto de teste é necessário para simularmos os dados que não conhecemos os rótulos (como se o nosso modelo já estivesse rodando em produção).
- B. O conjunto de validação (dev) é utilizado não para treinarmos os parâmetros do nosso modelo, mas sim para ajuste dos hiperparâmetros.
- C. O conjunto de treinamento é utilizado para ajustar os parâmetros internos do nosso modelo.
- D. O momento correto de realizar o split de dados é após todas as transformações de pré-processamento (como normalização), pois assim, conseguimos capturar o máximo de informação dos nossos dados e criamos um modelo melhor.

Questão 07



Tendo em vista as funções de ativação descritas graficamente acima:

1) : $\phi(z) = \frac{1}{1 + e^{-z}}$

2) : $f(x) = \max(0, x)$

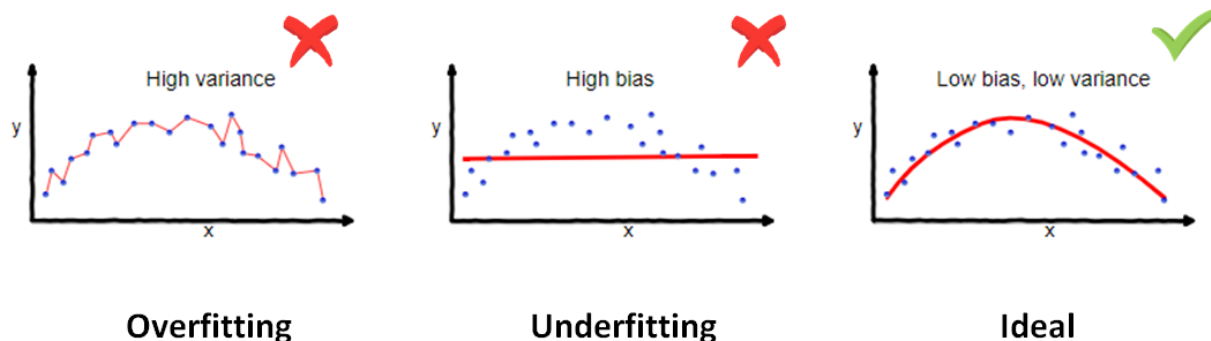
O que é correto afirmar?

- A. Trata-se de duas funções lineares: Sigmóide e ReLU
- B. Trata-se de duas funções não-lineares: Tangente hiperbólica e Sigmóide
- C. Trata-se de pelo menos uma função linear: Tangente hiperbólica e Unidade Linear
- D. Trata-se de duas funções não-lineares: Sigmóide e ReLU
- E. Trata-se de duas funções lineares: Sigmóide e Degrau

Questão 08

Como se chama o algoritmo de treinamento de redes neurais artificiais que (em conjunto com o gradiente descendente) permite o ajuste de sinapses (parâmetros do modelo) em redes de múltiplas camadas, onde a contribuição marginal do erro de classificação/regressão é repassado sucessivamente desde a camada de saída até a camada de entrada?

Questão 09



Na imagem acima, temos uma descrição visual de situações indesejadas para o treinamento do nosso modelo (Underfitting e Overfitting). Sobre essas situações e os conceitos de Bias (Viés) e Variance (Variância) o que é **INCORRETO** afirmar?

- A. A situação de Overfitting é indesejada, pois embora o nosso modelo tenha se ajustado "perfeitamente" aos nossos dados de treinamento, ele terá um desempenho ruim quando aplicado a outros dados. Dizemos nesse caso que o nosso modelo "memorizou" o conjunto de treinamento.
- B. A situação de Underfitting é indesejada, pois o nosso modelo é simples demais e por isso não é capaz de ter um desempenho satisfatório nem mesmo nos nossos dados de treinamento.
- C. Quando estamos treinando o nosso modelo, podemos utilizar estratégias como interrupção prematura do treinamento (usando um conjunto de validação, por exemplo) ou regularização para tentar evitar que o modelo entre em situação de Overfitting.
- D. Quando estamos em Overfitting ou Underfitting, o problema é causado pela qualidade ruim dos dados, ou pela pequena quantidade de dados disponíveis
- E. Quando estamos em Overfitting, às vezes, aumentar nossa base de dados pode ajudar a diminuir esse problema. Já no Underfitting, o problema é que nosso modelo é simples demais para aprender algo útil, então, temos que refazê-lo mais complexo, capaz de assimilar as nuances do nosso dataset.