

Atividade 13 - N8 INART

Luís Otávio Lopes Amorim - SP3034178

23 de Janeiro de 2021

1 Atividade Básica

1. Qual a diferença entre os processos de classificação e regressão?

SOLUÇÃO

A diferença entre os dois processos é que a classificação utiliza os valores de entrada visando caracterizar os valores de entrada entre duas ou mais classes distintas (por exemplo, se uma imagem mostra um gato ou um cachorro).

Por outro lado a regressão funciona como um ajuste de curvas, tentando, dessa forma, encontrar uma função que se aproxime de uma curva fornecida. Um exemplo clássico disso é o problema de previsão de preços de imóveis com base em seu tamanho, sua localização, número de quartos, etc.

2. Acesse o site da UCI que possui um repositório de base de dados para aprendizado de máquina: UCI.

Em seguida, selecione a opção “Regression” no menu à esquerda para filtrar apenas base de dados de problemas envolvendo regressão. Escolha uma base de dados e descreva-a.

SOLUÇÃO

O dataset utilizado, que traz dados sobre o trânsito de São Paulo está disponível aqui: [link](#).

O objetivo dele é caracterizar a velocidade do trânsito utilizando vários parâmetros para isso, sendo eles: hora, número de ônibus parados, número de veículos em excesso, número de vítimas de acidentes, número de veículos ultrapassando o sinal vermelho, número de caminhões de bombeiros nas ruas, número de entregadores de rua (frete), número de acidentes envolvendo entregadores de cargas perigosas, número de locais com falta de energia, número de incêndios, número de alagamentos, número de manifestações, número de defeitos na rede de trólebus, número de árvores caída nas ruas, número de semáforos desligados e número de semáforos intermitentes.

Essa base de dados foi coletada para um estudo de doutorado da Uninove, dos alunos Afonse Pereira e Andrea Martiniano. Algo interessante de se notar é que no artigo disponibilizado por eles, é utilizada a sigmoide como função de ativação,

o que é, pelo menos para mim, inesperado, já que não se trata de um problema de classificação, tipo de problemas que essa função trabalha bem, mas sim, um problema de regressão, portanto creio que uma ativação linear faria mais sentido, nesse caso.

2 Atividade Complementar

4. Considerando os problemas envolvendo aproximação de funções, discorra se há alguma vantagem e/ou desvantagem em se utilizar a função de ativação linear para o neurônio da camada de saída da rede ao invés do uso de funções não lineares, como a sigmoide e a tangente hiperbólica.

SOLUÇÃO

Há tanto vantagens quanto desvantagens, essa afirmativa é, de certa forma óbvia, se partirmos do pressuposto de que há situações em que a utilização da função linear é preferível à sigmoide em alguns casos, mas o oosto é verdade também (por exemplo no dataset iris, o uso de uma função linear não faz sentido). Inclusive, a vantagem está intimamente relacionada com a desvantagem.

A maior vantagem do uso de funções lineares é que essas podem ser utilizadas para problemas de regressão, ou seja, aproximar funções, segundo o texto fornecido pelo professor, sabe-se que utilizar uma função linear no neurônio de saída faz com que seja possível criar qualquer função contínua.

Por outro lado, a desvantagem está no fato de que problemas de classificação se tornam muito mais complexos, ou mesmo impossíveis de serem resolvidos, já que dependendo do número de barreiras de separação entre as classes, a função que descreve essa barreira é muito complexa, ou até mesmo descontínua, assim, o uso de uma função em forma de classificação (sigmoide, ReLu) é mais recomendado.

5. Explique qual a função dos “limiares θ dos neurônios A, B e C” e dos “pesos λ do neurônio de saída Y” na transformação da função sigmoide para composição da função de saída (você vai encontrar essa informação nos últimos parágrafos do texto e em sua última figura).

SOLUÇÃO

Se considerarmos a função de saída de cada neurônio da camada intermediária $g = \sigma(wx + \theta)$, podemos interpretar θ como uma translação dessa função em relação ao eixo x, já o peso λ como esticamento/achatamento na direção y.

Isso é essencial para uma regressão linear pois, assim, a saída da rede neural se torna uma combinação linear de n funções não lineares sendo n o número de neurônios da camada escondida:

$$Y(x_1, x_2, \dots, x_n) = \sum_{i=0}^m [g_i(x_1, x_2, \dots, x_n) \cdot \lambda_i] \quad (1)$$

Sendo $Y(x_1, x_2, \dots, x_n)$ a saída da rede neural, $g_i(x_1, x_2, \dots, x_n)$ a saída do neurônio i da camada escondida, λ_i seu peso e m o número de neurônios na camada escondida. Ou seja, a função Y é uma combinação linear de m funções $g : \mathbb{R}^n \rightarrow \mathbb{R}$. Assim, caso as funções g sejam tais que o conjunto \mathbb{A} formado por elas seja linearmente independente (o que é o caso, devido à natureza da família de funções sigmóides) esse conjunto se torna uma base do conjunto \mathbb{B} das funções $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Com isso, podemos dizer que o conjunto \mathbb{A} gera o conjunto \mathbb{B} , ou seja, podemos escrever qualquer elemento de \mathbb{B} como combinação linear dos elementos do conjunto \mathbb{A} . Sabendo disso, é fácil perceber que a função Y da equação 1 pode ser qualquer função de \mathbb{A} .

Portanto, devido aos vieses θ e aos pesos λ a rede neural de regressão linear é capaz de imitar qualquer função fornecida como entrada, bastando fornecer a ela uma quantidade suficiente de neurônios na camada intermediária e de exemplos de treino.