

Obtendo previsões pontuais

Transcrição

Já estimamos nosso modelo de regressão, agora precisamos obter previsões. Mas antes, faremos uma breve recapitulação dos conteúdos da aula anterior.

Utilizamos a técnica de machine learning de modelos de aprendizagem supervisionada. Separamos os dados em treino e teste, e uma vez que nós temos os resultados reais, podemos mensurar a eficiência do modelo e gerar métricas de comparação.

Lembrando de regressão linear é um procedimento utilizado com variáveis contínuas, não são variáveis qualitativas. Podemos utilizar modelos de classificação, e assim os palpites podem estar certos ou errados. Criamos a series com `Y`, criamos um dataframe com as variáveis explicativas; aplicamos o método `train_test_split` e `random_state` para coletarmos dados de maneira aleatória.

Em seguida, importamos o `LinearRegression()` e o instanciamos na variável `modelo`. Feito isso, utilizamos o método `fit()` e calculamos a nossa primeira métrica e criamos as previsões sobre os dados de teste e finalizamos com a comparação entre o que foi previsto e o que realmente aconteceu. Tivemos como resultado `0.69` de explicação das variações.

Na célula "Dados de Entrada", coletaremos o primeiro registro de `X_test`, e armazenaremos isso na variável `entrada`. Agora podemos gerar uma previsão pontual utilizando `modelo`.

```
modelo.predict(entrada)[0]
```

[COPIAR CÓDIGO](#)

Obteremos como resultado um consumo médio estimado de cerveja.

Podemos criar alguns estimadores simples, que podem ser utilizados nos mais variados contextos. Como por exemplo estimar quanto será o consumo de cerveja no fim de semana, utilizando as variáveis `fds` , `chuva` e `temp_max` . Essa será nossa entrada, depois, basta inserir o modelo.

Teremos:

```
temp_max=40
chuva=0
fds=1
entrada=[[temp_max, chuva, fds]]

print('{0:.2f} litros'.format(modelo.predict(entrada)[0]))
```

[COPIAR CÓDIGO](#)

Teremos como resultado 38742.53 litros.