

Naïve Bayes

Aprendizado de Máquina

Prof. Me. Otávio Parraga



Para as questões a seguir, utilizaremos a seguinte base de dados:

ID	Weather condition	Road condition	Traffic condition	Engine problem	Accident
1	Rain	bad	high	no	yes
2	Snow	average	normal	yes	yes
3	Clear	bad	light	no	no
4	Clear	good	light	yes	yes
5	Snow	good	normal	no	no
6	Rain	average	light	no	no
7	Rain	good	normal	no	no
8	Snow	bad	high	no	yes
9	Clear	good	high	yes	no
10	Clear	bad	high	yes	yes

1. Calcule as probabilidades a priori de $P(\text{Accident} = \text{yes})$ e $P(\text{Accident} = \text{no})$.
2. Calcule os likelihoods das features dadas as classes.
3. Em um dia com as seguintes características: Weather condition = rain, Temperature = good, Traffic condition = high e Engine Problem = No, é mais provável ocorra um acidente ou não?
4. Neste caso, faria sentido utilizar o estimador de Laplace? Em quais situações ele deve ser aplicado?
5. Se houvesse mais uma feature, que não possui associação nenhuma com a variável alvo, o modelo sofreria ou não com a adição dela?
6. Se houvesse uma variável numérica no conjunto de dados, poderíamos utilizar ela no modelo? Se sim, como?