FDP

on

“Smart Systems Design and Application”

Class: BE Computer
Computer Laboratory II
Group A (Mandatory Assignments)

Assignment No:5 :

Implement Naïve Bays to predict the Work_type

Predict WORK_TYPE for query tuple from given dataset using naïve bays approach:

Work Type	Age	Qualification	Experience
Consultancy	middle_age	Ph.D.	medium
Service	youth	MTech.	low
Research	youth	MTech.	low
Service	youth	BTech.	medium
Consultancy	middle_age	MTech.	high
Research	middle_age	Ph.D.	medium
Research	youth	BTech.	medium
Service	middle_age	MTech.	medium
Consultancy	senior	BTech.	high
Research	middle_age	Ph.D.	medium

query={'Age':'middle_age','Qualification':'MTech.','Experience':'medium','Work Type':'?'}

Note: Age & Experience attributes are preprocessed into range as follows:
Age={ if age<30:youth, if 30<=age<40:middle age, if age>40 :senior}
Experience={if Experience<5:low, if 5< Experience<15: medium, if Experience>15:high}

Naïve bays implementation Steps:

1. Calculate prior Probabilities of class to be predicted
2. Calculate conditional probabilities
3. Calculate posterior probability
4. Highest probability among above is predicted class for query tuple.

Mathematical model

- D is a dataset consisting tuples X
- $D=\{X_1,X_2,\ldots,X_n\}$
- Tuple X has attributes {*Work Type, Age, Qualification, Experience*}

Suppose that there are m classes, $C_1,C_2,\ldots,C_m$

Given a tuple X, the classifier will predict that X belongs to class having the highest posterior probability conditioned on X.

$P(Ci|X)>P(Cj|X)$ for $1 \leq j \leq m, j \neq i$

- Thus the *maximum prediction hypothesis* by bays theorem:

$$P(Ci|X) = \frac{P(X|Ci) * P(Ci)}{P(X)}$$

- To predict class label of X:

$$P(X|Ci) * P(Ci) > P(X|Cj) * P(Cj) \text{ for } 1 \leq j \leq m, j \neq i$$

1.Calculate prior Probabilities of class to be predicted

- Prior probability of Work Type : Service = 0.3
- Prior probability of Work Type : Consultancy = 0.3
- Prior probability of Work Type : Research = 0.4

2.Calculate conditional probabilities:

- Prior probability of Work Type : Service =

$$P(Service) = \frac{\#\text{ of occurances 'Service'}}{\#\text{ Total number of tuples}} = \frac{3}{10} = 0.3$$

$$P(middle\_age|Service) = \frac{\#\text{ of occurances of 'middle\_age' \&'Service'}}{\#\text{ of occurances of 'Service'}} = \frac{1}{3}$$

- P(medium | Service)= 0.66
- P(MTech. | Service)= 0.66

Multiplication of above probabilities to give Posterior Probability of Work Type : **Service** = (0.3333*0.3333*0.6666*0.6666)=0.0444

2.Calculate conditional probabilities:

- Prior probability of Work Type : Consultancy = 0.3
- conditional probility P(middle_age | Consultancy)= 0.66
- conditional probility P(medium | Consultancy)= 0.33
- conditional probility P(MTech. | Consultancy)= 0.33

Multiplication of above probabilities to give Posterior Probability of Work Type : **Consultancy** = 0.022

2.Calculate conditional probabilities:

- Prior probability of Work Type : Research = 0.4
- conditional probility P(middle_age | Research)= 0.5
- conditional probility P(medium | Research)= 0.75
- conditional probility P(MTech. | Research)= 0.25

Multiplication of above probabilities to give Posterior Probability of Work Type : **Research** = 0.0375

Predicted Class (Highest posterior Probability)for given query tuple is:

Work Type : **Service** = 0.044

Work Type : **Research** = 0.0375

Work Type : **Consultancy** = 0.022

Service

Demo

Thank you

I am thankful to BOS, Head of Computer Department & Convener of SSD - FDP SIT, Narhe for giving opportunity to present Assignment in this FDP.

……

Kushal P. Birla

kpbirla@kkwagh.edu.in

+919403835023

Assistant Professor

Department of Computer Engineering,

K.K.Wagh Institute of Engineering & Research, Nashik