Pattern Recognition
Course Code: CSE313/423
by
Dr Sudarshan Nandy

Module -II

Statistical Patten Recognition:

Bayesian Decision Theory, Classifiers, Normal density and discriminant functions

Basic of probability

- Probabilty denotes “chance”
- Laws that governs the chance of occurance of an event is maintained in the theory of the probability
- How much chance of and event -----> is known as probabilty.
- Random Experiment: when an occurance of the event can not be predicted.
- Outcome: The result of an random experiment
- Event: Phnomenon which occurs in a random experiment

Basic of probability

- Mutually Exclusive: if two or more events are not occuring simultaneously or togather then events are mutually exclusive. They occurs one at a time.

  Example: events A , B, C, D are mutually exclusive, only if A, or B, or C or D can occur.

- Exhustive: it is the a group of events from a random experiment from which at least one event has occured.
- Equally Likely/ Equaly Probable :If after taking into consideration all relevant evidance, none of them can be expected in preference to another.

Basic of probability

- Probability:

  if a random experiment has n possible outcomes, which are mutually exclusive, exhaustive and equally likely and m of these are favourable to an event A then the probality of A is P(A) and is calculated as follows:

  P(A) = m/n = Number of outcomes favourable to the Event / total number of mutually exclusive event.

Basic of probability

- What is the process of calculation?

  Step1 : count the number of outcomes from an experiment where the outcomes must satisfy: a. Mutually Exclusive b. Exhaustive c. Equaly likely ----->(n)

  Step2 : Count the numbers of Event in favour of the Event for which you are caculating the probability.------>(m)

  Step3 : Divide Step2 by Step1 which gives the probability of the event.

- But the calculation is always 0<P(A) <1

Basic of probability

- **Example 1:Two coins are tossed. Find the probability of getting both heads or both tails.**

  1. Events are ---> HH, HT, TH and TT and hence n=4

  2. Both head or tail means HH or TT then m = 2

  3. So the probability : P(HH) or P(TT) = 2/4 = 0.50

- **Example 2: Two dice with faces marked 1,2,3,4,5,6 are thrown simultaneously and the points on the dice are multiplied together. Find the probability that the product is 12.**

  1. for two dice if thrown simultaneously then n = 6*6 = 36

  2. now the products will be 12 for the following event- (2,6), (3,4), (4,3), (6,2)

  Hence , m = 4

  3. So, P = 4/36 = 1/9

Basic of probability

- **A bag contains 6 white balls and 4 black balls. One ball is drawn. What is the probabilty that it is white?**

1. There are 6 favourable event possible, hence n =10
2. The number of outcomes is, m =6
3. Hence, P(W) = 6/10

Basic of probability

- **Multiplication Rule:**

  Probability is to find the chance of an event.

- In an **independent event**, two events does not effect each other.

  Ex: rolling a six sided normal die for 5 and 3

  Event is independent .

  So, P(5) = 1/6 and P(3) = 1/6

- Now, to understand the probability of the 5 and 3 , we need to go through the multiplication rule in the probability:

  So, $P(\text{ A and B}) = P(A \cap B) = P(AB) = P(A)*P(B)$

  Hence, P(5 and 3)= 1/6 * 1/6 = 1/36

- **P(AB)** is also known as a compound probability. This is actually means probability of occurance of A as well as B envent.

Basic of probability

- The **Dependent events** on other side are those where occurance of an event effects the another event.
- A famous example of this is the drawing the king and then queen from the deck of cards. In this experiment the queen is drawn without putting the king back.
- So, P(KING) = no of king / total number of cards = 4/52

  P(QUEEN) = no of queen / (total number of cards – King) = 4/51
- Now the probability of any dependent events is caculated as follows:

$$P(A \cap B) = P(AB) = P(A) * P(B|A)$$

Basic of probability

- So,

  P(King and Queen) = P(King) * P(Queen | King)

  = 4/52 *4/51= 16/2652

- Hence the Conditional probability where we consider that the probability of occurance A , assuming that B has already occured and it is denoted as P(A | B).

$$P(A|B)=\frac{P(AB)}{P(B)} \qquad P(B|A)=\frac{P(AB)}{P(A)}$$

Basic of probability

- **Given P(A) = 1/2 , P(B) = 1/3 and P(AB) = 1/4 then state whether the events A and B are (i) mutualy exclusive, (ii)exhaustive (III) Equally Likely/ Equaly Probable (iV)independent.**

  Ans:

- i) No, here P(AB) = 1/4 and $P(AB) \neq 0$

  II)So, here P(A+B) = P(A) +P(B) -P(AB) = 7/12 and if $P(A+B) \neq 1$

  then the events are not exhaustive.

  III) no, it is not equaly likely because P(A) and P(B) are different.

  IV) Here, $P(AB) \neq P(A).P(B)$

  So, A and B are not independent.

Basic of probability

- **IF TWO EVENTS A AND B ARE MUTUALY EXCLUSIVE THEN THE PROBABILITY OF OCCURANCE:**

  P(A or B) = P(A) + P(B)

- **Proof**: suppose there are n outcomes from a random experiment. Events occur for this experiments are mutually exclusive. So,

  Favourable to Event A is = m_1 ;

  Favourable to Event B is = m_2 ;

  Hence, P(A) = m_1/n

  and P(B) = m_2/n

  So, number of case favourable to either A or B = (m_1+m_2)

- Hence, the probability for either A or B is = P(A or B) = $P(A \cup B)$ = P(A+B)

  P(A+B) = (m_1+m_2) / n = (m_1/n) +(m_2/n) = P(A)+P(B)

- Hence the probabilty of the occurance of any events from series of events which are mutually exclusive, is actually

$$P(A_1+A_2+A_3+\ldots\ldots)=P(A_1)+P(A_2)+\ldots..+P(A_k)$$

Basic of probability

- Given that $P(A) = 3/8$ and $P(B) = 5/3$ and $P(A+B) = 3/4$ find $P(A|B)$ and $P(B|A)$. Are A and B independent?

Bayes theorem

- In probability theory and statistics, Bayes' theorem (alternatively Bayes' law or Bayes' rule), named after Reverend Thomas Bayes
- It basically describe the probability of an Event, and the knowledge on the condition related to the event is known.
- This is mathematically expressed as follows:

$$P(A|B)=\frac{P(B|A)P(A)}{P(B)}$$

- Where, P(A|B) is a conditional probability. This is the calculation of probability A , considering the B event has already happend. P(A) and P(B) is the probability of A and B event.

Bayes theorem

Proof:

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \quad if\ P(B)\neq 0$$

and

$$P(B|A)=\frac{P(B\cap A)}{P(A)} \quad if\ P(A)\neq 0$$

$$Now,$$

$$P(A\cap B)=P(B\cap A)$$

$$\Rightarrow P(A|B)P(B)=P(B|A)P(A)$$

$$\Rightarrow P(A|B)=\frac{P(B|A)P(A)}{P(B)}$$

Naive Bayes classifier

- This is based on Bayes theorem.
- Naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable.
- For example, a fruit may be considered to be an apple if it is red, round, and about 10 cm in diameter. A naive Bayes classifier considers each of these features to contribute independently to the probability that this fruit is an apple, regardless of any possible correlations between the color, roundness, and diameter features.
- Naive Bayes classifiers can be trained very efficiently in a supervised learning environment.
- Work well in the complex environment.
- It is work very well even with a small dataset.

Naive Bayes classifier

- According to the Bayes theorem:

$$P(A|B)=\frac{P(B|A)*P(A)}{P(B)}$$

- The features of the patterns is representing a class which is represented as follows:

$$L=(x_1,x_2,x_3,\ldots..,x_n)$$

- These features are independent from each other and can be treated as a list of independent event. It is also representing the instance and hence probability or the chance of Kth class, denoted as L_K, is represnted by all the attributes or features in the pattern:

$$P(L_k|x_1,x_2,x_3,..)$$

Naive Bayes classifier

- According to the Baye's Theorem, the conditional probability can be decomposed as:

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

- Now consider that the all features

$$(x_1,x_2,x_3,\dots x_n)\in X$$

are mutualy exclusive and conditional on the category L_k. Under this assumption,

$$P(L_k|x_1,x_2,x_3,..)=P(x_i|L_k)$$

Naive Bayes classifier

- So, now the Baye's theorem can be re-written as:

$$P(L_k|x_1,x_2,x_3,..)=\frac{P(L_k)P(x_i|L_k)}{P(x_1)P(x_2)P(x_3)...P(x_n)}$$

$$=\frac{P(L_k)\prod_{i=1}^{n}P(x_i|L_k)}{P(x_1)P(x_2)P(x_3)...P(x_n)}$$

$$=\frac{P(L_{k=yes})\prod_{i=1}^{n}P(x_1|L_{k=yes})P(x_2|L_{k=yes})P(x_3|L_{k=yes})}{P(x_1)P(x_2)P(x_3)...P(x_n)}$$

- Now as the denominator is constant hence, it is canceled and then it is always

$$P(L_k|x_1,x_2,x_3,..)\propto P(L_k)\prod_{i=1}^{n}P(x_i|L_k)$$

Naive Bayes classifier

- Now, the Naive bayes classifier comes with a decision rule which is based on the maximum probable likelihood and it is represented as follows:

$$D = argmax_k P(L_k) \prod_{i=1}^{n} P(x_i | L_k)$$

Naive Bayes classifier

Sl No	**A1=Outlook**	**A2=Temperature**	**A3=Humidity**	**A4=Windy**	**Target=Play Golf**
P0	Rainy	Hot	High	False	No
P1	Rainy	Hot	High	True	No
P2	Overcast	Hot	High	False	Yes
P3	Sunny	Mild	High	False	Yes
P4	Sunny	Cool	Normal	False	Yes
P5	Sunny	Cool	Normal	True	No
P6	Overcast	Cool	Normal	True	Yes
7	Rainy	Mild	High	False	No
8	Rainy	Cool	Normal	False	Yes
9	Sunny	Mild	Normal	False	Yes
10	Rainy	Mild	Normal	True	Yes
11	Overcast	Mild	High	True	Yes
12	Overcast	Hot	Normal	False	Yes
13	Sunny	Mild	High	True	No

Naive Bayes classifier

Outlook

	Yes	No	P(yes)	P(no)
Sunny	2	3	2/9	3/5
Overcast	4	0	4/9	0/5
Rainy	3	2	3/9	2/5
Total	9	5	100%	100%

Temperature

	Yes	No	P(yes)	P(no)
Hot	2	2	2/9	2/5
Mild	4	2	4/9	2/5
Cool	3	1	3/9	1/5
Total	9	5	100%	100%

Humidity

	Yes	No	P(yes)	P(no)
High	3	4	3/9	4/5
Normal	6	1	6/9	1/5
Total	9	5	100%	100%

Wind

	Yes	No	P(yes)	P(no)
False	6	2	6/9	2/5
True	3	3	3/9	3/5
Total	9	5	100%	100%

Play		P(Yes)/P(No)
Yes	9	9/14
No	5	5/14
Total	14	100%

Naive Bayes classifier

- Problem = P(yes|today=(outlook$_{Sunny}$,Temp$_{hot}$,Humid$_{normal}$,Wind$_{false}$))
- Solution ->

  P(yes|today=(outlook$_{Sunny}$,Temp$_{hot}$,Humid$_{normal}$,Wind$_{false}$)) =P(yes)*P(Sunny|yes)*P(hot|yes)*P(normal|yes)*P(False|yes)
- P(yes|today) = (9/14)*(2/9*2/9*6/9*6/9) =0.0141
- P(no|today)=(5/14)*(3/5*2/5*1/5*2/5)=0.0069
- P(yes|today)+P(no|today) = 1
- P(yes|today) = 0.0141/(0.0141+0.0069)=0.671
- P(no|today) = 0.0069/(0.0069+0.0141) = 0.323
- D = argmax(P(yes|today), P(no|today)) = argmax(0.671, 0.323) =P(yes|today) => 0.671

Naive Bayes classifier

- Problem = P(yes| today=($outlook_{Rainy}$,$Temp_{hot}$,$Humid_{high}$,$Wind_{true}$))

Code

```
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import KFold,train_test_split,cross_val_score

from sklearn.model_selection import train_test_split

from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import make_scorer, accuracy_score,precision_score

df=pd.read_csv('/root/Desktop/Class/pattern/code/Iris.csv')
print(df.head())

df.drop(columns="Id",inplace=True)


print(df.head())
corr_matrix=df.corr()
fig,ax=plt.subplots(figsize=(15,10))
ax = sns.heatmap(corr_matrix,
        annot=True,
        linewidths=0.7,
        fmt=".3f",
        cmap="Accent_r")
X=df.iloc[:,0:4].values
y=df.iloc[:,4].values

X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.5,random_state=42)
gaussian = GaussianNB()
gaussian.fit(X_train, y_train)
Y_pred = gaussian.predict(X_test)
accuracy_nb=round(accuracy_score(y_test,Y_pred)* 100, 2)
#acc_gaussian = round(gaussian.score(X_train, y_train) * 100, 2)

accuracy = accuracy_score(y_test,Y_pred)

print('accuracy_Naive Bayes: %.3f' %accuracy)
```