İhsan Doğramacı Bilkent University

CS-464

Homework 1

Tuğberk Dikmen

21802480

Question 1)

1.1)

From Box 1, we have a ⅔ chance of picking a fair blue coin and a ⅓ chance of picking an unfair yellow coin.

From Box 2, we have a ½ chance of picking a fair blue coin and a ½ chance of picking an unfair red coin.

The probability of getting heads twice in a row (HH) from a fair coin is $(1/2)^2$ and from the unfair yellow and red coins are $(1/4)^2$ and $(1/10)^2$.

- Total probability:

P(HH) = ½ $*$ ⅔ $* (1/2)^2$ + ½ $*$ ⅓ $* (1/4)^2$ + ½ $*$ ½ $* (1/2)^2$ + ½ $*$ ½ $* (1/10)^2$
P(HH) = 0.15875

1.2)

- Can be solved using Bayes' theorem:

P(Fair|HH) = $\frac{P(HH|Fair)*P(Fair)}{P(HH)}$

P(Fair) = ½ $*$ ⅔ + ½ $*$ ½

P(HH|Fair) = $(1/2)^2$

P(Fair|HH) = 0.91864

1.3)

- Bayes' theorem:

P(Red|HH) = $\frac{P(HH|Red)*P(Red)}{P(HH)}$

P(Red) = ½ $*$ ½

P(HH|Red) = $(1/10)^2$

P(Red|HH) = 0.01575

Question 2)

2.1)

MLE for μ

- Probability Density Function (PDF) of Normal Distribution:

$$f(x;\ \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{\frac{-(x-\mu)2}{2\sigma^2}}$$

- Likelihood Function:

$$L(\ \mu, \sigma) = \Pi\ i\ =\ 1\ to\ n\ f(x_i;\ \mu, \sigma)$$

- Log-Likelihood Function:

$$log\ L(\ \mu, \sigma = \Sigma\ i\ =\ 1\ to\ n\ f(x_i;\ \mu, \sigma)$$

- MLE of μ:

Differentiate the log-likelihood with respect to \(\mu \), set to zero and solve:

$$\hat{\mu}\ MLE = \frac{1}{n}\ \Sigma\ i\ =\ 1\ to\ n\ f(x_i)$$

2.2)

MAP Estimate of μ

- Prior Distribution (Exponential) for μ:

$f(\mu; \lambda) = \lambda e^{-\lambda\mu}$ for μ bigger than or equals to 0

- MAP Estimate:

Multiply likelihood by the prior and maximize:

$$\hat{\mu}\mathrm{MAP} = \mathrm{argmax}\mu\ L(\ \mu, \sigma) * f(\mu; \lambda)$$

- Finding MAP:

Involves solving a more complex equation, requiring calculus, considering both the likelihood and the prior distribution.

2.3)

Probability and Likelihood for a New Data Point

- Given:

μ = 1 and σ= 1

- Probability Density Function (PDF):

$$f(x;\ 1, 1) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{\frac{-(x-1)2}{2\sigma^2}}$$

- Density at x_{n+1} = 1:

Approximately 0.399

- Likelihood of x_{n+1} = 2 :

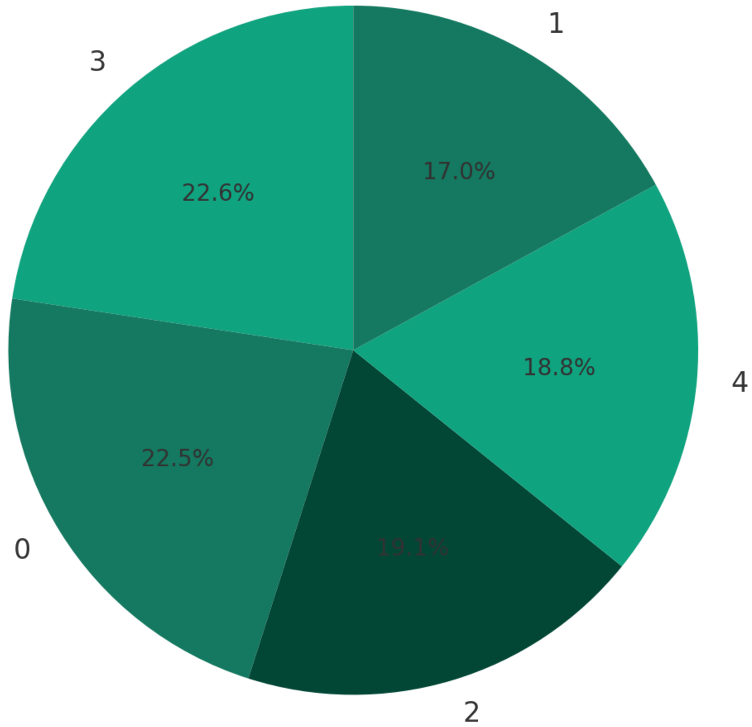
Approximately 0.242

Question 3)

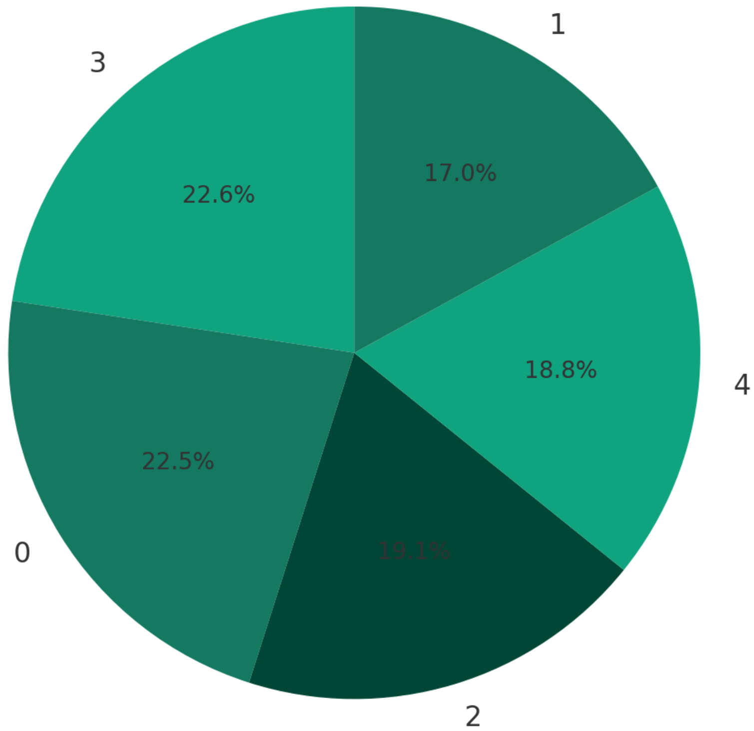
3.1)

1. Class Distribution Percentages:
   - Business: 22.48% in the training set, 24.24% in the test set.
   - Entertainment: 17.03% in the training set, 18.31% in the test set.
   - Politics: 19.12% in the training set, 17.59% in the test set.
   - Sport: 22.60% in the training set, 24.06% in the test set.
   - Tech: 18.76% in the training set, 15.80% in the test set.

Training Set Class Distribution



Training Set Class Distribution



2. Prior Probability of Each Class:

- Business (0): 22.48%
- Entertainment (1): 17.03%
- Politics (2): 19.12%
- Sport (3): 22.60%
- Tech (4): 18.76%

These are the probabilities of a randomly selected document being of each class type.

3. Balance of the Training Set:

- The training set is relatively balanced with each class having a percentage between approximately 17% and 23%. A perfectly balanced set is not essential, but severe imbalance can lead to model bias towards more common classes. The current balance is unlikely to cause significant bias.

4. Occurrences of "alien" and "thunder" in 'Tech' News:

5. Estimators for the Parameters of the Multinomial Naive Bayes Model:

- $\theta_{j|y=k} = \frac{T_{j,y=k}}{\Sigma j=1 \ to \ |V| \ T_{j,y=k}}$ estimates the probability of word j in class y=k.
- $\pi_{y=k} = \frac{N_{yk}}{N}$ estimates the prior probability of class y=k.
- $T_{j,y=k}$ is the number of times word j appears in documents of class y=k, including multiple occurrences.
- N_{yk} is the number of documents in class y=k.
- N is the total number of documents.

3.2 / 3.3 / 3.4)

```
---------------------------------------
Unsmoothed Multinomial Naive Bayes
Accuracy: 0.242

Confusion Matrix
Predicted (COLUMN)    0
Actual (ROW)
0                   135
1                   102
2                    98
3                   134
4                    88
---------------------------------------


Smoothed Multinomial Naive Bayes
Accuracy: 0.975

Confusion Matrix
Predicted (COLUMN)    0   1   2    3   4
Actual (ROW)
0                   131   0   2    0   2
1                     0  96   1    0   5
2                     1   0  96    0   1
3                     0   0   1  133   0
4                     1   0   0    0  87
---------------------------------------


Bernoulli Naive Bayes
Accuracy: 0.966

Confusion Matrix
Predicted (COLUMN)    0   1   2    3   4
Actual (ROW)
0                   132   0   2    0   1
1                     3  96   1    0   2
2                     4   0  94    0   0
3                     0   0   0  134   0
4                     4   2   0    0  82
---------------------------------------
```