School of Computer Science Engineering and Technology

Course- BTech

Course Code-

Year- 2023-2024

Date – 27-07- 2023

Type- Core

Course Name- Statistical Machine learning

Semester- odd

Batch- ALL

https://www.kaggle.com/datasets/camnugent/california-housing-prices

Download the dataset from the above link.

Question 1:

Find out the maximum likelihood parameters values for california_housing ['latiyude',total_rooms','house_median_age','total_bedrooms'] data, each column individually.

Question 2:

Find log-likelihood values for 50, 75, and 80. For California_housing ['latitude'] and find the Max likelihood values.

Question 3:

Estimate the density of each column of california_housing ['latitude',total_rooms','house_median_age','total_bedrooms'] data, For the bins [5, 10, 15, 20]. And write the observations.

Solution:

1.

```
import pandas as pd
a=pd.read_csv('/content/sample_data/california_housing_train.csv')
p=a['households']
μ=np.mean(p)
б=np.std(p)#sigma value
б²=np.var(p)#sigma square value
```

2.

```
z=np.array(d['total_rooms''])
a=np.mean(z)
b=np.std(z)#sigma value
c=np.var(z)#sigma square value
x=50
d=np.log(np.sqrt(2*3.14))
e=np.log(b)
f=(x-a)**2
g=2*(c**2)
h=f/g
i=-d-e
print(i-h)
```

3. import matplotlib.pyplot as plt

```
Plt.hist(d['total_rooms'']),5)
```