

# School of Computer Science Engineering and Technology

Course- BTech

Type- Core

Course Code-

Course Name- Statistical Machine learning

Year- 2023-2024

Semester- odd

Date – 5-10- 2023

Batch- ALL

Question 1:

<https://www.kaggle.com/code/liananapalkova/visualising-a-high-dimensional-dataset/input>

- a) Read the data with pandas and find features and target variables
- b) Find target variable
- c) Use PCA to reduce the dimension
- d) Find the accuracy before PCA and after PCA with SVM

Question 2:

<https://www.kaggle.com/datasets/kellistephenson/increasing-breast-cancer-awareness>

- a) Read the data with pandas and find features and target variables
- b) Find target variable
- c) Apply logistic regression and SVM find the accuracy
- d) And find Uncertainty of both models for 1-50 samples