PROJECT 1 REPORT
MACHINE LEARNING 2
ECE 780

KIRTHANA KAKOLLU
01088074

This project is about dimensionality reduction carried out by six methods using real application datasets of SELDI spectrum data of ovarian cancer, madelon data, & SMK CAN 187 data.

The six methods are

1. Fisher score
2. Principal Component Analysis
3. Sequential Forward Selection
4. PCA and Sequential Forward Selection
5. Lasso Regularization
6. Stability Selection

λ used is 0.3

For SELDI Spectral Data	Accuracy
Fisher score	88.88%
Principal Component Analysis	97.22%
Sequential Forward Selection	94.44%
PCA and Sequential Forward Selection	95.83%
Lasso Regularization	94.44%
Stability Selection	97.22%

For Madelon	Accuracy
Fisher score	60.66%
Principal Component Analysis	61.24%
Sequential Forward Selection	61.82%
PCA and Sequential Forward Selection	61.47%
Lasso Regularization	60.33%
Stability Selection	61.2%

For SMK CAN 187	Accuracy
Fisher score	66.66%
Principal Component Analysis	77.77%
Sequential Forward Selection	74.90%
PCA and Sequential Forward Selection	76.19%

Lasso Regularization	72.37%
Stability Selection	73.56%