**Statistical Learning Lab**

**Assignment - 5**

**Subset Selection Regularization and Dimension Reduction**

**Create an R markdown file and submit the pdf with all the code snippets, plots, results and explanation.**

1. Load the dataset “Cellphone.csv”. Display first few rows of the dataset.
2. Perform preliminary analysis to show how the variables are related to each other. Use scatter plot, box plot etc. to visualize how different variables impact the “Price” variable, which is the response variable.
3. Perform best subset selection on this dataset, (dropping the Product\_id variable). Explain the results.
4. Create a plot with Cp on y-axis and number of variables on the x-axis. Determine the lowest Cp and report how many variables are included in the lowest Cp model.
5. Plot the best subset selection output and explain the plot.
6. Use principal component regression on the same dataset with 5 components and 7 components. How much variability is explained by these two models?
7. Perform Lasso on the model and explain the results.