

**School of Social Sciences**  
**BSc (Hons.) Data Sciences & Analytics**  
**Semester VII**  
**Course: Applied Econometrics**

**PRACTICAL – 6**

**Applying Regularization Techniques – LASSO, RIDGE & Elastic Net**

Customer Personality Analysis is a detailed analysis of a company's ideal customers. It helps a business to better understand its customers and makes it easier for them to modify products according to the specific needs, behaviours, and concerns of different types of customers. Customer personality analysis helps a business to modify its product based on its target customers from different types of customer segments. For example, instead of spending money to market a new product to every customer in the company's database, a company can analyse which customer segment is most likely to buy the product and then market the product only to that segment.

The dataset provides the following details with respect to a Company's customers

1. ID: Customer's unique identifier
2. Year\_Birth: Customer's birth year
3. Education: Customer's education level
4. Income: Customer's monthly household income
5. Kidhome: Number of children in customer's household
6. Teenhome: Number of teenagers in customer's household
7. Recency: Number of days since customer's last purchase
8. Amt\_spent: Amount spent on the company's products in the last two years (in thousands of Rs.)
9. NumDealsPurchases: Number of purchases made with a discount
10. AcceptedCmp1: 1 if customer accepted the offer in the 1st campaign, 0 otherwise
11. AcceptedCmp2: 1 if customer accepted the offer in the 2<sup>nd</sup> campaign, 0 otherwise
12. NumWebPurchases: Number of purchases made through the company's web site
13. NumWebVisitsMonth: Number of visits to company's web site in the last month
14. NumStorePurchases: Number of purchases made directly in stores
15. Complain: 1 if customer complained in the last 2 years, 0 otherwise

**Prepare a summary report with your inferences:**

1. Provide insights to the Company through the customer profile analysis using suitable exploratory data analysis for all the above variables. Also provide your specific analysis on customer satisfaction, campaign effectiveness, and customer's preferred mode of shopping. As part of data cleaning, if you handled/omitted any missing values, do indicate it in your report.
2. Create a model to assess the interrelationship between the amount spent on the company's products based on explanatory variables in the dataset.
3. Comment on multicollinearity.
4. Apply LASSO, Ridge and Elastic Net regularization techniques to estimate the model parameters. Compare and analyse on which is a better choice of model.
5. Obtain the cross validated MSE plot.
6. Comment on the goodness of fit. Also draw your conclusions based on the RMSE.
7. Analyse the residual diagnosis, identify violations and suggest measures to handle the same.

**Note:**

- i. Upload the summary report in the shared drive folder.
- ii. Create a folder with your name and upload.