First data :
Adult Data Set

https://archive.ics.uci.edu/ml/datasets/adult

Abstract: Predict whether income exceeds $50K/yr based on census data. Also known as "Census Income" dataset.

Data Set Characteristics : Multivariate
Number of Instances: 48842
Train data : 32561
Teating data 16281
Number of Attributes: 14
Associated Tasks : Classification

Attribute Information:

Listing of attributes : >50K, <=50K.

1. Age : continuous.

2. workclass : Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked.

3. fnlwgt : continuous.

4. education : Bachelors, Some-college, 11th, HS-grad, Prof-school, Assoc-acdm, Assoc-voc, 9th, 7th-8th, 12th, Masters, 1st-4th, 10th, Doctorate, 5th-6th, Preschool.

5. education-num : continuous

6. marital-status : Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse.

7. occupation : Tech-support, Craft-repair, Other-service, Sales, Exec-managerial, Prof-specialty, Handlers-cleaners, Machine-op-inspct, Adm-clerical, Farming-fishing, Transport-moving, Priv-house-serv, Protective-serv, Armed-Forces.

8. relationship : Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried.

9. race : White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Black.

10. sex : Female, Male.

11. capital-gain : continuous.

12. capital-loss : continuous.

13. hours-per-week : continuous.

14. native-country : United-States, Cambodia, England, Puerto-Rico, Canada, Germany, Outlying-US(Guam-USVI-etc), India, Japan, Greece, South, China, Cuba, Iran, Honduras, Philippines, Italy, Poland, Jamaica, Vietnam, Mexico, Portugal, Ireland, France, Dominican-Republic, Laos, Ecuador, Taiwan, Haiti, Columbia, Hungary, Guatemala, Nicaragua, Scotland, Thailand, Yugoslavia, El-Salvador, Trinadad&Tobago, Peru, Hong, Holand-Netherlands.

possible problems : Our aim is identify whether income exceeds $50K/y . but we need preprocessing for data . sincce data have massing value and Qualitative data. In addition this data is an imbalance data . we can find important feature (feature selection like lasso regression) . next find a proper model to predict income (like logistic regression).
According to the customer's financial ability, we can choose products that are more suitable for customers(e.g advertisement . insurance)

Second data :

Auto MPG Data Set

https://archive.ics.uci.edu/ml/datasets/auto+mpg

Abstract: Revised from CMU StatLib library, data concerns city-cycle fuel consumption

Data Set Characteristics : Multivariate
Number of Instances: 398
Number of Attributes: 8
Associated Tasks : regression

Attribute Information:

1. mpg: continuous

2. cylinders: multi-valued discrete

3. displacement: continuous

4. horsepower: continuous

5. weight: continuous

6. acceleration: continuous

7. model year: multi-valued discrete

8. origin: multi-valued discrete

9. car name: string (unique for each instance)

possible problems : we also need to preprocessing for data . and we can try to find main influence variables associate with fuel consumption(feature selection) . it can help us refinement car in the future .