# Topic(s): Decision Tree & Random Forest

Please ensure you update all the details:

**Topic: Decision Tree And Random Forest**

**Name : Prashanth K C**

**Batch Id:**  05012021-10AM

**Problem-1**

1. **Business Problem**
   1. **Objective :- making profit**
   2. **Constraints (if any):- restriction or limitations on decision variable**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name of feature** | **description** | **type** | **Relevance** |
| Sales | no of sales in company | qualitative ,continues | relevant |
| Comp Price | company price for product | qualitative ,discrete | relevant |
| Income | income of company | qualitative ,continues | relevant |
| Advertising | Advertising given by company | qualitative ,discrete | irrelevant |
| Population | no of population | qualitative ,discrete | irrelevant |
| Price | price of product | quantitative discrete | relevant |
| Shelve Loc | quality of shelving location | qualitative ,discrete | irrelevant |
| Age | no of age | Quantitative , discrete | irrelevant |
| Education | how many are educated | Quantitative ,discrete | irrelevant |
| Urban | located in urban or not | Quantitative , discrete | irrelevant |
| US | located in us or not | Quantitative , discrete | irrelevant |

1. **Work on each feature of the dataset to create a data dictionary as displayed in the below**

**Problem-2**

1. **Business Problem**
   1. **Objective :- minimize classification error , maximize accuracy of the target variable.**
   2. **Constraints (if any):- minimize the computational time**

**2 .Work on each feature of the dataset to create a data dictionary as displayed in the below**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name of feature** | **description** | **type** | **Relevance** |
| Number of times pregnant | No of times Pregnancies | Quantitative , discrete | relevant, it provides useful information |
| Plasma glucose concentration | Blood glucose levels in your body, | Quantitative , continues | relevant |
| Diastolic blood pressure | Blood Pressure in your body | Quantitative , discrete | relevant |
| Triceps skin fold thickness | Thickness of skin | qualitative ,discrete | irrelevant |
| 2-Hour serum insulin | insulin measurement | qualitative ,continues | relevant |
| Body mass index | BMI rate | quantitative ,continues | relevant |
| Diabetes pedigree function | Diabetes Pedigree Function | qualitative ,continues | irrelevant |
| Age (years) | age of individual person | Quantitative , discrete | irrelevant |
| Class variable | over all outcome | Quantitative , discrete | relevant |
|  |  |  |  |

Problem3:

Buisiness Objective : Minimise fraud

Buisiness Constraint : Maximise Convenience to customers

|  |  |  |  |
| --- | --- | --- | --- |
| Name Of Feauture | Description | Type | Relevance |
| Undergrade | Inferior to standard grade | Categorical | Irrelevant, it does not provide useful information |
| Marital Status | Status of a individual person | Nominal | It is not useful |
| Taxable Income | Tax for individual person | Continuous | Its provides useful information for calculations |
| City Population | Population of city | Discrete | Its provides useful information |
| Work Experience | Years of experience | Continous | Ita a useful information |
| Urban | Characterstics of city | Category | Does not provide useful informations |

Problem 4 :

Buisiness Objective : Minimise Fraud Candidate

Buisiness Constraint : Maximise the Profit

|  |  |  |  |
| --- | --- | --- | --- |
| Name of Feature | Description | Type | Relevance |
| Postion of employee | Position of person in an organization | Categorical | It is not providing useful information |
| No of years of experience | Experiecnce year of a person | Continuous | It provides useful information |
| Monthly income of employee | Income of employee | Continuous | It provides useful information |