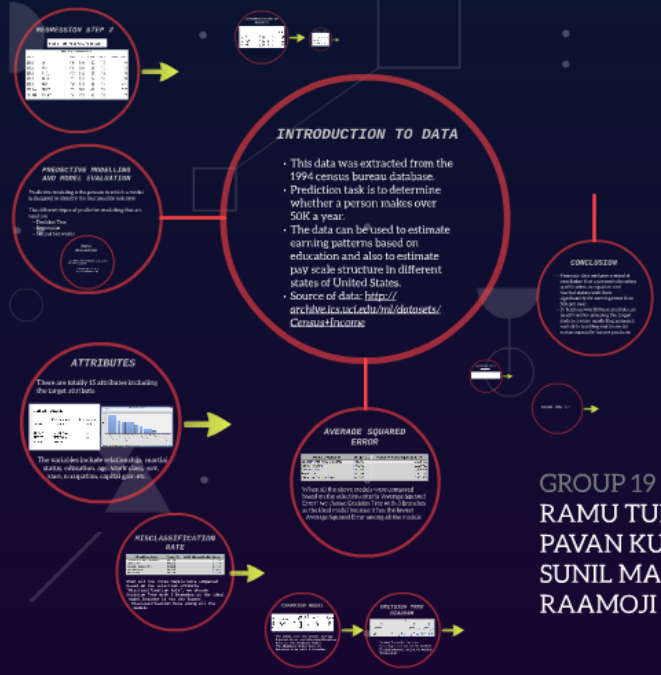
Thecasesolution.com

MIS 6324 BI PROJECT PRESENTATION
1994 US CENSUS DATASET



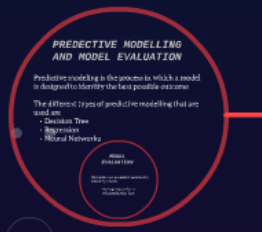
GROUP 19
RAMU TUNGALA
PAVAN KUMAR
SUNIL MARUTHI
RAAMOJI RAO

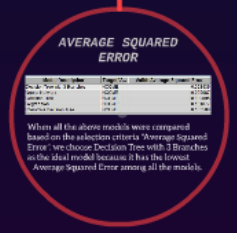
Thecasesolution.com

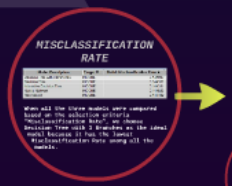
MIS 6324 BI PROJECT PRESENTATION
1994 US CENSUS DATASET

INTRODUCTION TO DATA

- This data was extracted from the 1994 census bureau database.
- Prediction task is to determine whether a person makes over 50K a year.
- The data can be used to estimate earning patterns based on education and also to estimate pay scale structure in different states of United States.
- Source of data: http://archive.ics.uci.edu/ml/datasets/Census+Income







GROUP 19
RAMU TUNGALA
PAVAN KUMAR
SUNIL MARUTHI
RAAMOJI RAO

INTRODUCTION TO DATA

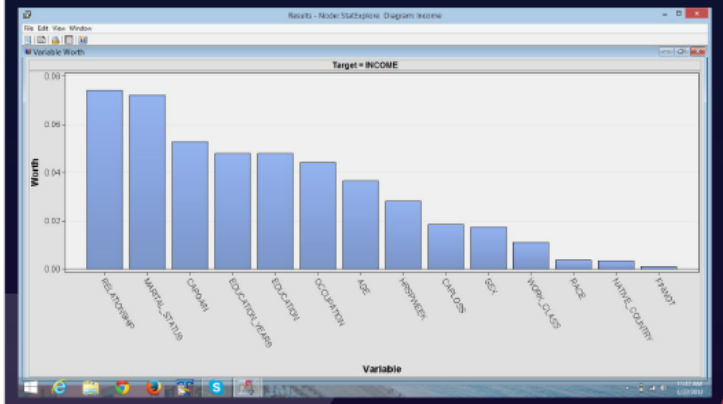
- This data was extracted from the 1994 census bureau database.
- Prediction task is to determine whether a person makes over 50K a year.
- The data can be used to estimate earning patterns based on education and also to estimate pay scale structure in different states of United States.
- Source of data: *http://archive.ics.uci.edu/ml/datasets/Census+Income*

ATTRIBUTES

There are totally 15 attributes including the target attribute

Variable Summary

Role	Measurement Level	Frequency Count
INPUT	INTERVAL	6
INPUT	NOMINAL	8
TARGET	BINARY	1



The variables include relationship, martial status, education, age, work class, sex, race, occupation, capital gain etc .

PREDECTIVE MODELLING AND MODEL EVALUATION

Predictive modeling is the process in which a model is designed to identify the best possible outcome

The different types of predictive modelling that are used are

- Decision Tree
- Regression
- Neural Networks

MODEL EVALUATION

The Models are Evaluated based on the following criteria.

-Average Squared Error
-Misclassification Rate