

Academic Task Number: 02
Date of allotment: 04-Apr-2025
Date of submission: 04-Apr-2025

Course code: CSI323
Course Title: Predictive Analysis
Maximum Marks: 20 Practical +10 theories

Objective:

Develop a predictive model in IBM SPSS Modeler to determine whether a student will be placed based on their academic and professional background.

- **Target (Dependent) Variable:**
 - status (Binary Classification: "Placed" or "Not Placed")
- **Independent (Predictor) Variables:**
 - gender (Nominal)
 - ssc_p (Scale) – Secondary Education Percentage
 - ssc_b (Nominal) – Secondary Education Board
 - hsc_p (Scale) – Higher Secondary Percentage
 - hsc_b (Nominal) – Higher Secondary Board
 - hsc_s (Nominal) – Higher Secondary Stream
 - degree_p (Scale) – Degree Percentage
 - degree_t (Nominal) – Degree Type
 - workex (Nominal) – Work Experience (Yes/No)
 - etest_p (Scale) – Employability Test Score
 - specialisation (Nominal) – MBA Specialization
 - mba_p (Scale) – MBA Percentage

A. Data Preparation

1. Load the Dataset: Import the dataset into IBM SPSS Modeler.
2. Handle Missing Values: Check for missing values in attributes like hsc_p, degree_p, etc.
3. Data Transformation: Convert categorical variables (gender, ssc_b, etc.) into dummy variables if needed.

B. Model Building

1. Select Modelling Technique:
 - Decision Tree (C5.0, CHAID, CART)
 2. Train the Model: Use the dataset to train the model with status as the target variable.
 3. Evaluate Performance: Use Accuracy, Precision, Recall, and ROC Curve to assess performance.
- **Key Insights:**
 - The importance of work experience (workex) in getting placed.
 - The impact of academic performance (ssc_p, hsc_p, degree_p, mba_p) on placement.
 - The significance of employability test scores (etest_p).
 - The effect of MBA specialization (specialisation) on placement chances.
 - The influence of gender (gender) and education board (ssc_b, hsc_b) on placements.

Dataset link - :

<https://docs.google.com/spreadsheets/d/1Fn2fLETCznkr8q0GBx3495itk4kLzHzJ/edit?usp=sharing&ouid=103368472363855442930&rtpof=true&sd=true>