

Babu Banarasi Das University



CASE STUDY ON Predicting Student Churn Using IBM SPSS Modeler

SUBMITTED TO: MR. AYUSHMAN BHADURIA

SUBMITTED BY: Chandra Bhan Maurya

Agenda:

To analyze student academic and behavioral data in order to develop a predictive model that identifies students who are likely to drop out (churn) from the institution.

Outcomes:

A working predictive model that helps the institution detect at-risk students early, enabling timely academic or counseling interventions to reduce dropout rates.

Required Tools:

IBM SPSS Modeler for data processing and model building, MS Excel for dataset handling.

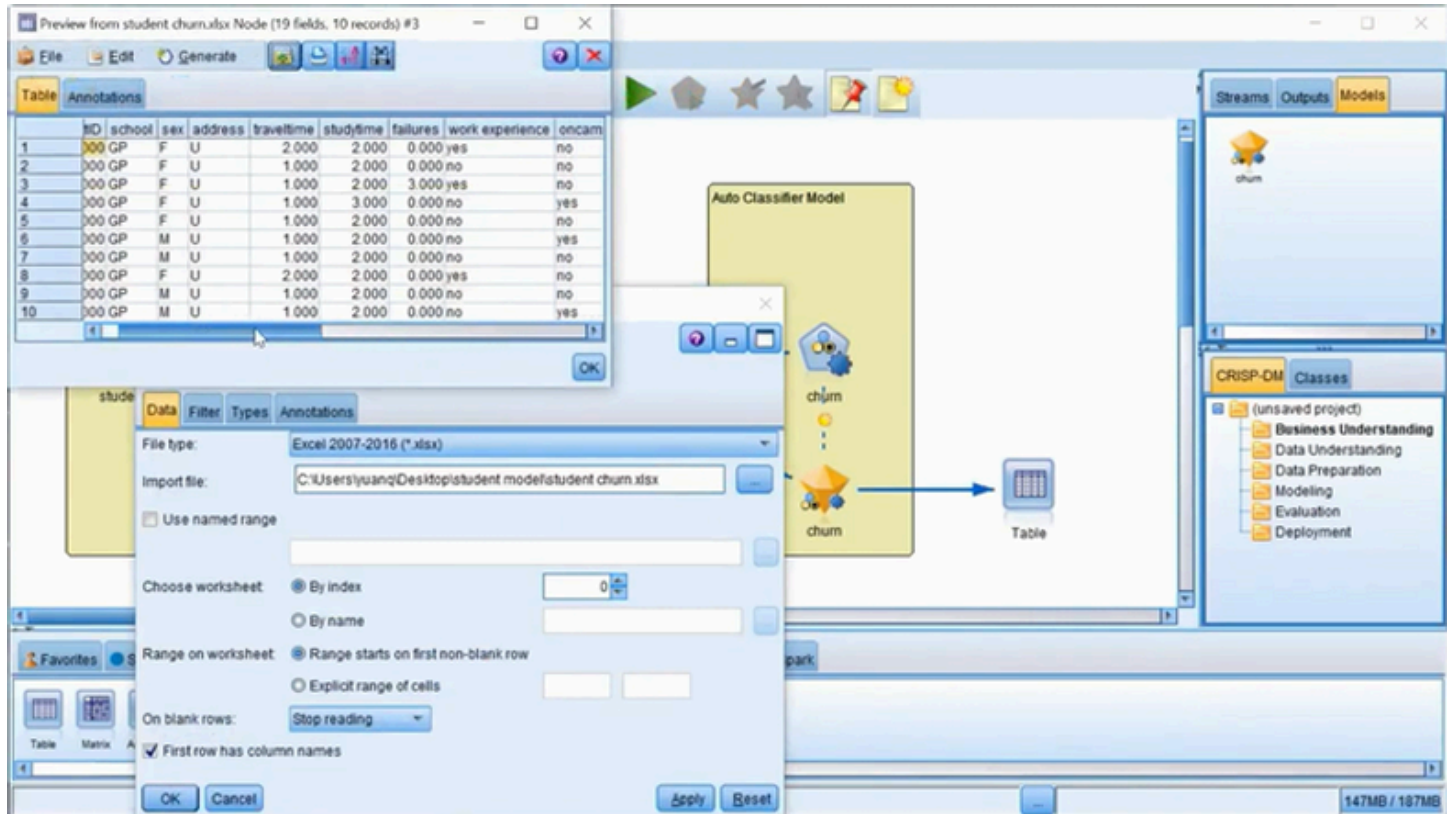
Working:

Collect student data → Import and preprocess it in IBM SPSS Modeler → Build and evaluate predictive models → Select the most accurate model → Interpret results and recommend intervention strategies.

Steps to Perform the Project

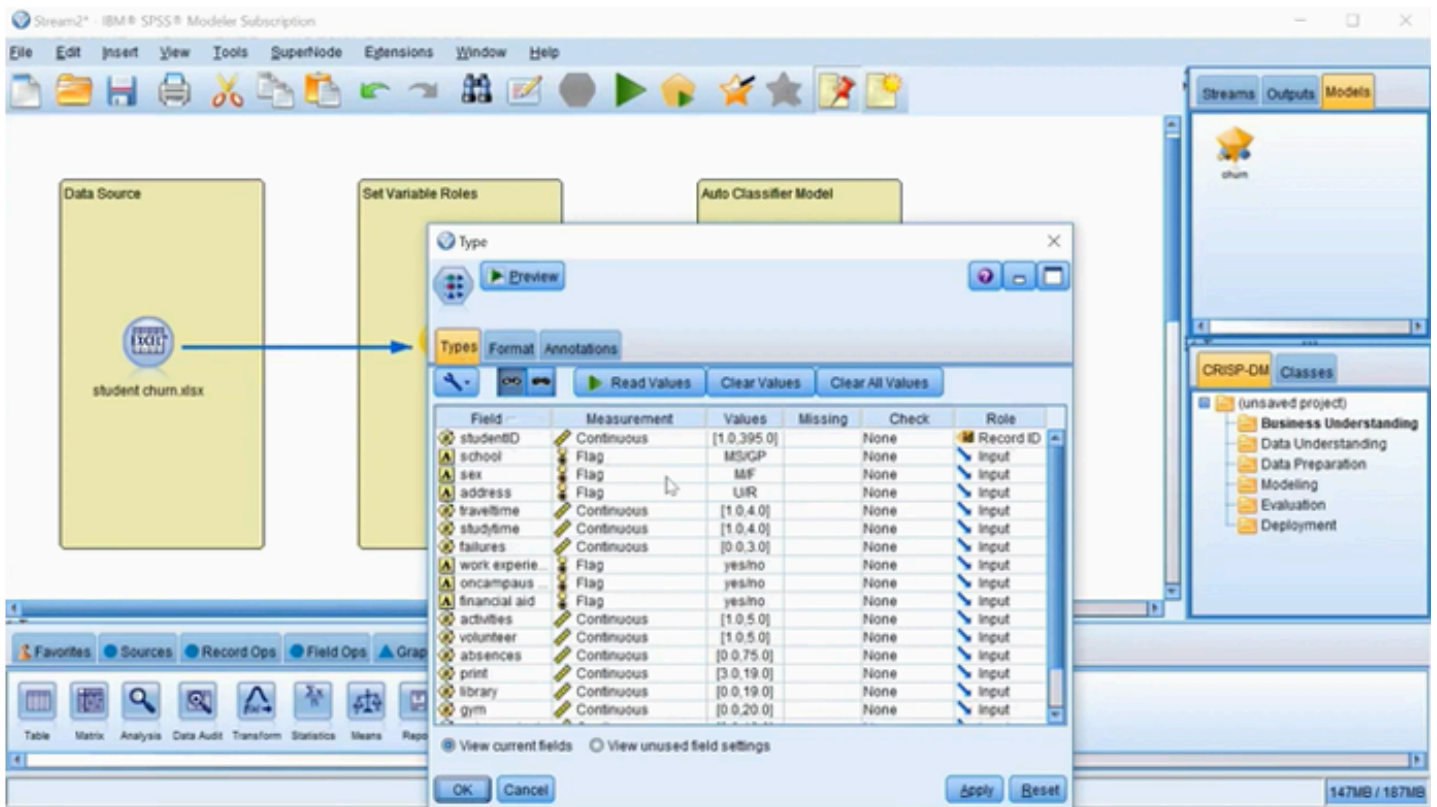
Step 1: Import the Dataset:

Open IBM SPSS Modeler and bring in the student dataset containing academic, attendance, and demographic records



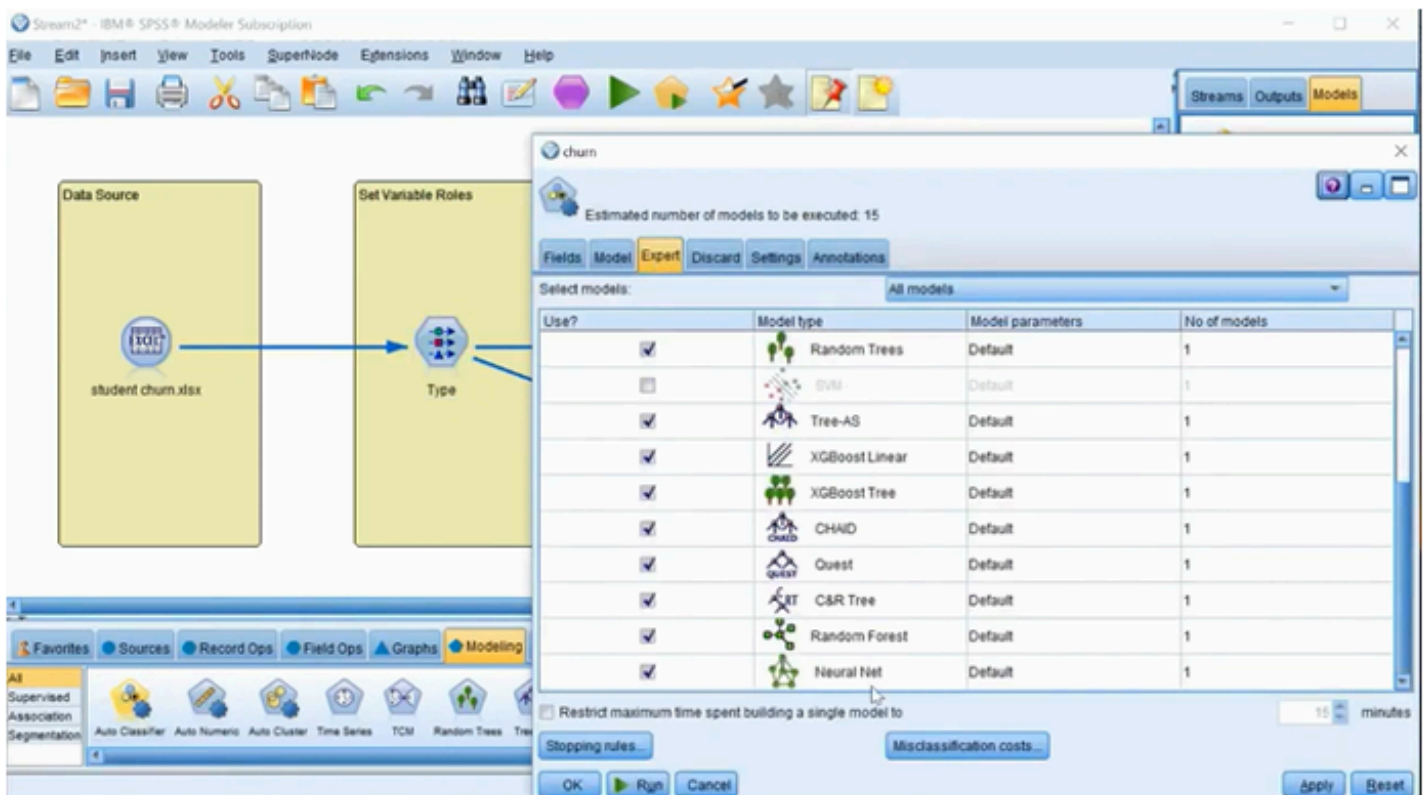
Step 2:

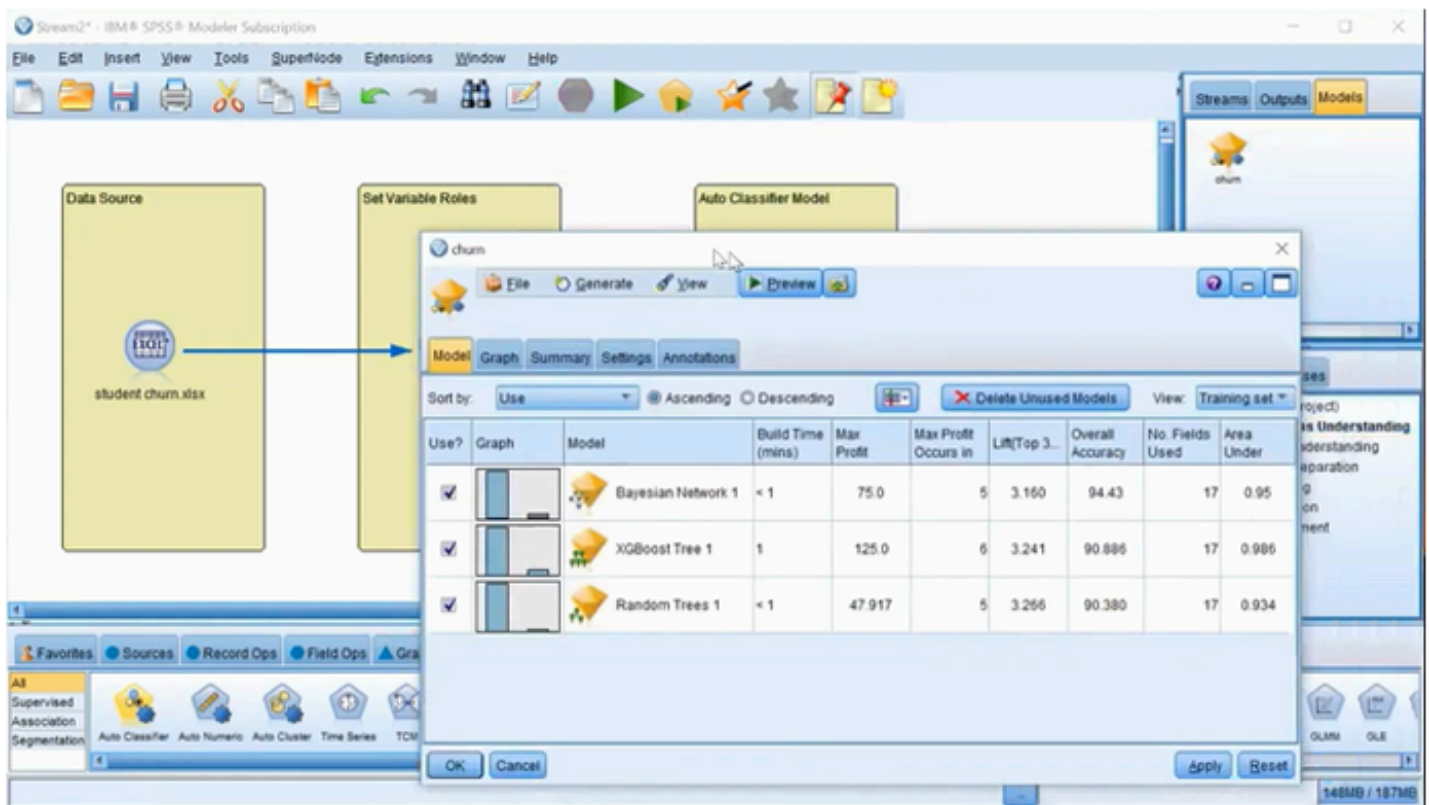
Set Variable Roles: Assign Student Status as the Target variable and mark other related fields such as marks, attendance, activities, etc., as Input variables.



Step 3: Select the Model:

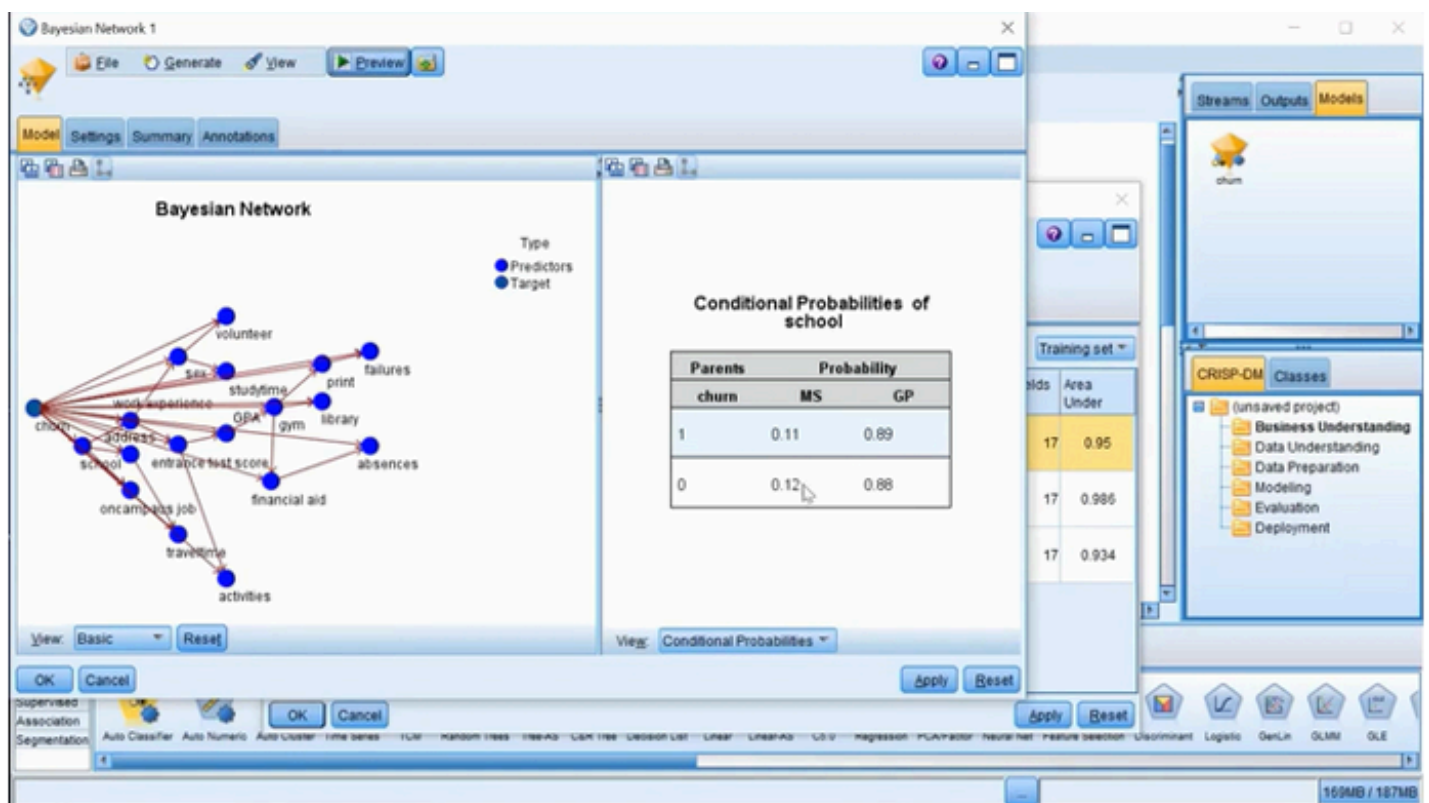
Drag and place the Auto Classifier node to automatically test and compare multiple classification algorithms.





Step 4: Run the Model:

Execute the Auto Classifier to generate the best performing model based on accuracy and classification performance.



Step 5: Evaluate and Interpret Results:

Review charts, accuracy tables, and predictor importance to understand which factors influence student churn.

The screenshot displays the IBM SPSS Modeler interface. On the left, a data table titled 'Table (21 fields, 395 records) #2' is visible, containing columns for 'library', 'gym', 'entrance test score', 'GPA', 'churn', '\$X^2-churn', and '\$XFC-churn'. The 'churn' column is highlighted in yellow. In the center, an 'Auto Classifier Model' is shown, with a flow from a 'churn' node to a 'Table' node. On the right, a 'Streams' panel shows a 'churn' node. Below the 'Streams' panel, a 'CRISP-DM' diagram is visible, showing the project structure: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment. The bottom of the interface features a toolbar with various modeling tools, including 'Auto Classifier', 'Auto Numeric', 'Auto Cluster', 'Time Series', 'TCM', 'Random Trees', 'Tree-AS', 'C&R Tree', 'Decision List', 'Linear', 'Linear-AS', 'C5.0', 'Regression', 'PCA/Factor', 'Neural Net', 'Feature Selection', 'Discriminant', 'Logistic', 'GenLin', 'GLM', and 'OLE'.

Conclusion:

The project successfully demonstrated how student churn can be predicted using IBM SPSS Modeler. By defining the target variable and using the Auto Classifier model, we identified the key academic and behavioral factors that influence whether a student continues or drops out. The predictive results help institutions take early action, such as providing counselling, academic support, or attendance monitoring, to reduce dropout rates and improve overall student retention.