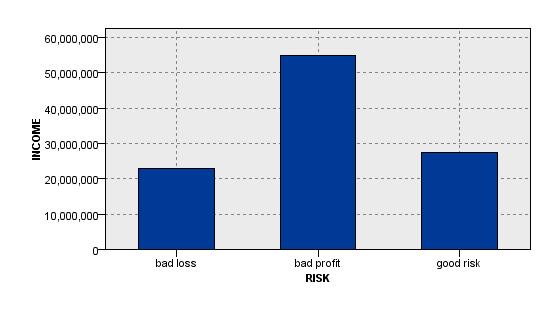
**Problem Statement**

The main issue is the absence of a reliable model to categorize customers' risk based on this data, hampering credit issuance and risk management. The financial institution struggles to accurately assess credit risk for its credit card customers, leading to potential misclassification . The demographic data available for about 2,500 customers—such as age, income, marital status, number of children, and credit behavior—should be used to create a predictive model. Misclassification results in poor lending decisions, financial losses, and ineffective customer management

**Brief summary and Data Audit**

The dataset comprises 4,117 records, highlights both continuous and categorical variables related to customer demographics and financial behaviors. Key continuous variables, such as AGE (ranging from 18 to 50) and INCOME (spanning from 15,005 to 59,944), indicate significant diversity among customers, reflecting various economic backgrounds. The audit underscores the need for careful data preparation, with continuous variables requiring scaling and categorical variables needing encoding or reduction to effectively manage their complexity. Overall, while the dataset offers rich insights for analysis, effective preprocessing is essential to facilitate successful predictive modeling. The bar plot shows that income affected the categories of risk significantly.



**What was done and analysis based on model evaluation**

The data was imported from an Excel file (Risk.xlsx) into a tabular format and specifying the relevant data types, focusing on 11 Fields. The process includes a Feature Selection step to identify the most significant variables for predicting credit risk. A model was then developed based on these selected features, which is evaluated and presented in a table format. The final output, labeled as RISK, categorizes customers into risk levels, followed by an Analysis step to derive insights from the model's performance and predictions.

A diagram of a software flow

Description automatically generated with medium confidence

The evaluation of the decision tree model revealed an overall accuracy of **75.81%**, correctly classifying 3,121 out of 4,117 customer credit risks. However, it misclassified 996 cases, representing **24.19%** of the total. While the model performs well for many customers, the misclassifications indicate areas for improvement. Management can use these insights to refine the model by revisiting feature selection and exploring additional attributes, which will enhance credit risk assessments and risk management strategies.

**A white paper with black text

Description automatically generated**

**How can management make use of the model**

The decision tree model offers a straightforward way for management to assess the credit risk of customers based on key financial and demographic factors. By using this model, management can easily identify which customers are more likely to repay their loans and which may pose a higher risk of default. For instance, the model highlights that customers with no loans or those who are younger with higher incomes are generally safer bets for lending. In contrast, customers with multiple loans or lower incomes are more likely to be a risk. With this information, management can make better decisions about who to approve for credit and tailor their lending strategies accordingly.

Additionally, understanding these risk categories allows management to focus on developing support programs for higher-risk customers, such as financial education or personalized repayment plans, which can improve overall customer relationships and reduce potential losses. Ultimately, this model helps management make informed, data-driven decisions that can enhance the company's financial health

**Appendix**

A graph with blue bars

Description automatically generated

**Decision Tree Visualization**A computer diagram of a computer

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated**Predicted values table**