Insurance

Automated Underwriting

The original data had 15 variables.

5 Irrelevant variables were removed.

The data was finally prepared with the remaining 10 variables.

Missing values were negligible in comparison to the whole dataset and were removed.

The current dataset has the following variables:

1. Age
2. Workclass
3. Education
4. Marital Status
5. Occupation
6. Relationship
7. Race
8. Sex
9. Hours per week
10. Income

Age and Hours per week being the numeric variables. Outliers have been removed according to the requirement.

Dummy variables have been created for the other remaining factor variables.

The final dataset contains dummy variables and the other two numeric variables.

Further Approach:

Modelling will be done considering these variables.

Segments can be created on the basis of Age group, Income and Occupation.

Based on these segments further sub-categories can be made on the other remaining variables.

This will help in assessing the risk at different levels and also help in targeting the customer for the right product.

This will further help in automated underwriting of the proposals based on the customer segment and help in processing it faster with reduced expenses and claims.