**EXPERIMENT - 1**

**AIM:**

The aim of this project is to analyze and explore the **Autism Screening Dataset** to gain insights into the data, identify patterns, and prepare it for further analysis or modeling. The goal is to understand the characteristics of the dataset, handle missing values and duplicates, and visualize the distribution of features to better understand the data structure.

**Procedure:**

1. **Data Loading**:
   * Load the dataset (autism\_screening\_data.csv) into a pandas DataFrame for analysis.
2. **Data Inspection**:
   * Check for missing values using isnull().sum() to identify columns with null values.
   * Clean column names by stripping any leading or trailing spaces using str.strip().
3. **Handling Duplicates**:
   * Identify and count duplicate rows using duplicated().sum().
   * Remove duplicate rows using drop\_duplicates() to ensure data integrity.
4. **Data Exploration**:
   * Display the first few rows (head()) and last few rows (tail()) of the dataset to get a sense of the data.
   * Generate descriptive statistics (describe()) for numerical columns to understand the distribution, central tendency, and spread of the data.
5. **Data Visualization**:
   * Plot histograms for all numerical columns using hist() to visualize their distributions.
6. **Further Analysis (Potential Next Steps)**:
   * Handle missing values (if any) by imputation or removal.
   * Split the data into training and testing sets for predictive modeling.

**OUTPUT :**

