Experiment: 1

# Aim:

Data preprocessing methods on student and labor datasets.

# Software Required:

Open source Weka 3.8, Python.

# Description:

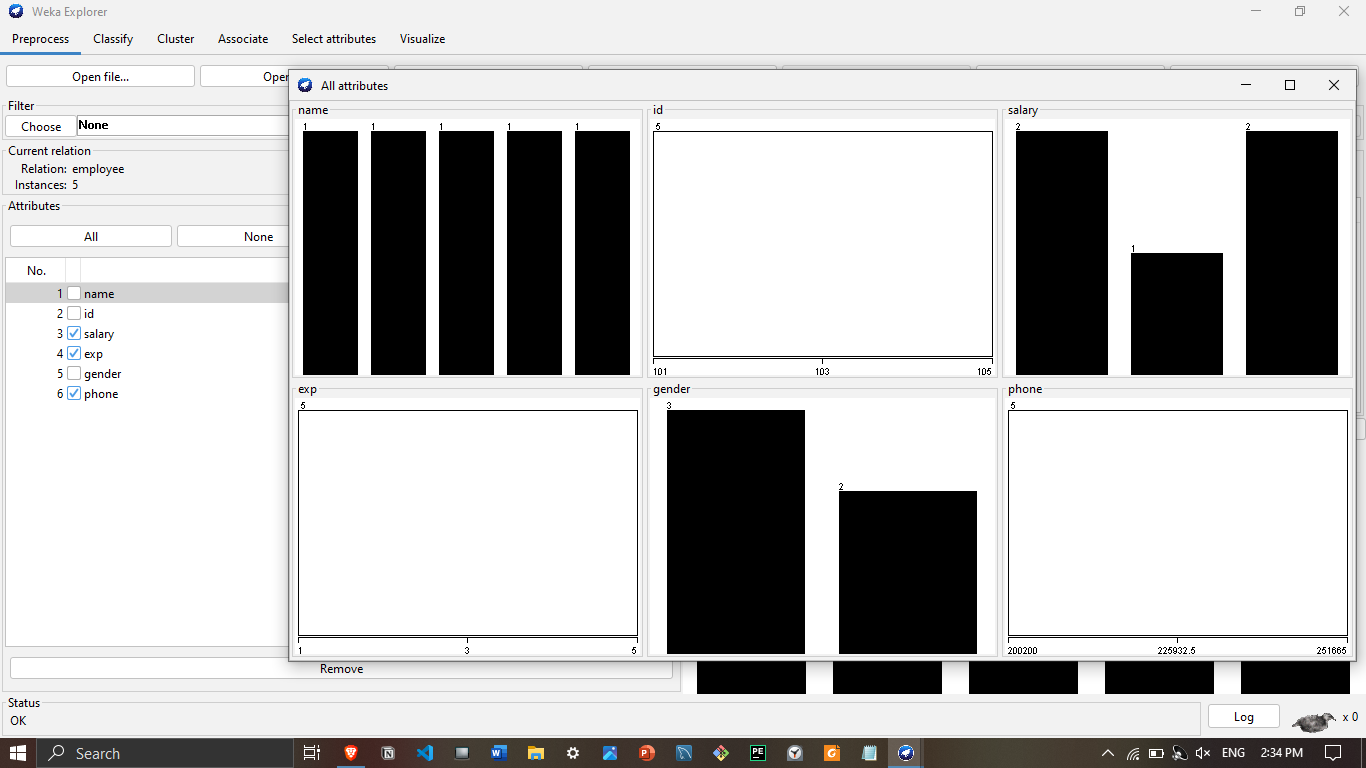
A training dataset for an Employee Table is created using Notepad, saved in ARFF format. Weka Explorer is then opened, the ARFF file is loaded, and data preprocessing is performed through user interaction in the Weka GUI.

# Implementation:

# Code:

|  |
| --- |
| employee\_data = """@relation employee@attribute name {x,y,z,a,b}@attribute id numeric@attribute salary {low,medium,high}@attribute exp numeric@attribute gender {male,female}@attribute phone numeric@datax,101,low,2,male,250311y,102,high,3,female,251665z,103,medium,1,male,240238a,104,low,5,female,200200b,105,high,2,male,240240 |

**Output:**



# Learning Outcomes:

1. **Data Prep Skills:** Proficiency in creating datasets using Notepad.
2. **Weka Explorer Proficiency:** Competence in loading datasets for efficient analysis.
3. **ARFF Format Understanding:** Grasp the importance of saving datasets in ARFF format.
4. **Attribute Handling Expertise:** Manage diverse dataset attributes effectively.
5. **Data Preprocessing Competence:** Gain practical skills in Weka Explorer for preprocessing.