****

**Student Club Participation Prediction**

**Submitted by:**

**Name: Himanshu Maurya**

**Roll No: 202401100300124**

**Course: Artificial Intelligence – Mid-Semester Exam**

**Submitted to:**

**Mr. Shivansh Prasad**

**Department of Computer Science**

**Kiet Group Of Institution**

**Introduction:**

This project aims to predict whether a student will join a club based on their interest areas and schedule availability. By identifying the patterns behind student participation, we can better understand what drives engagement in extracurricular activities. This analysis is especially useful for student organizations and academic institutions looking to improve club involvement and outreach strategies.

**Methodology:**

**1. Data Simulation:** As actual student data was not provided, a synthetic dataset was generated. The dataset includes binary features indicating interest in coding, music, and sports, as well as availability during evenings and weekends.

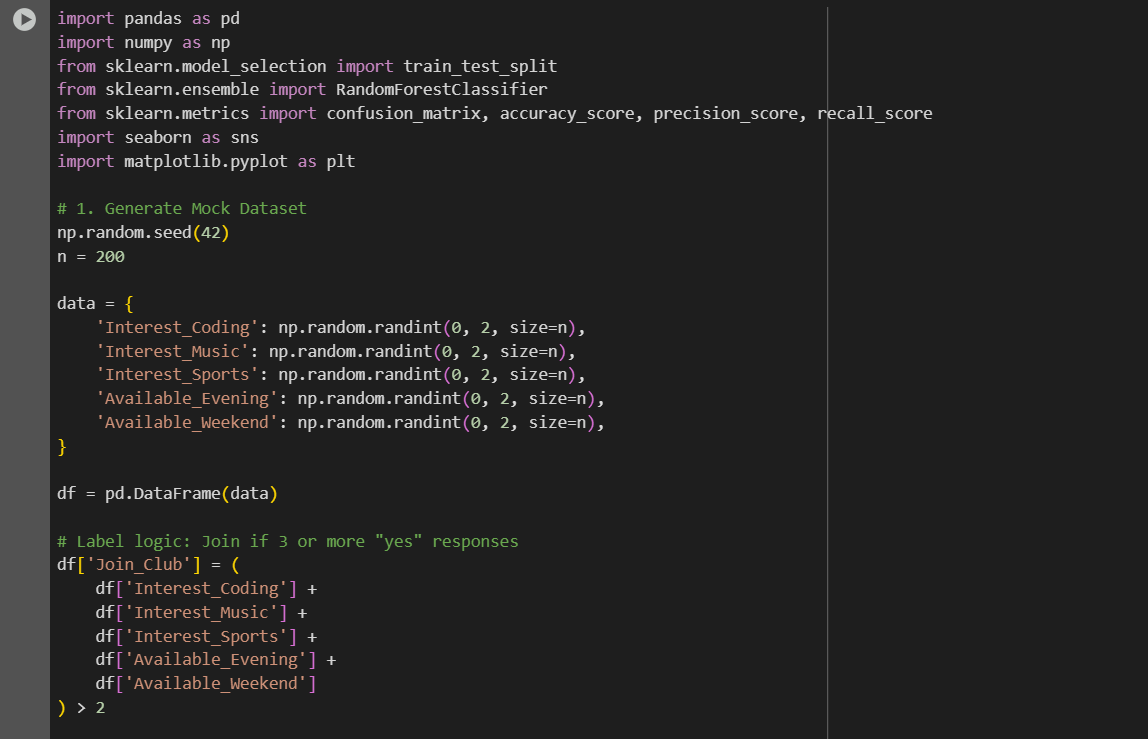
**2. Label Generation:** A simple rule-based system was used to assign the target variable Join\_Club. Students with 3 or more positive responses (interests or availability) were labeled as likely to join a club.

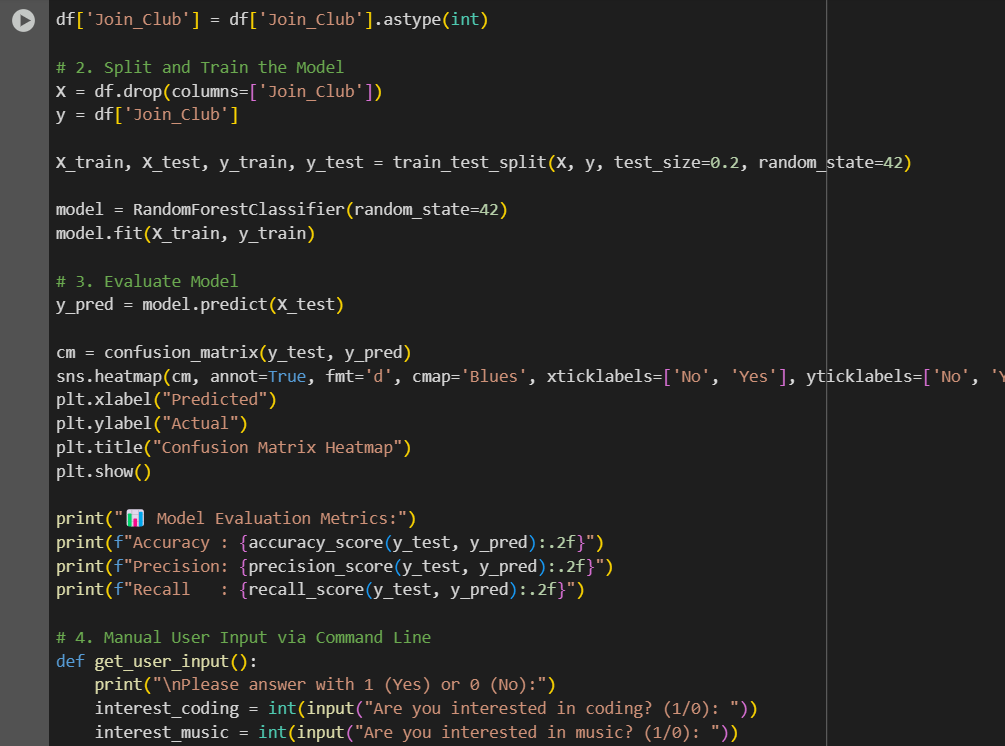
**3. Model Training:** A Random Forest Classifier was trained to learn patterns in the data and make predictions about future students. The data was split into training and testing sets (80/20) to evaluate model performance.

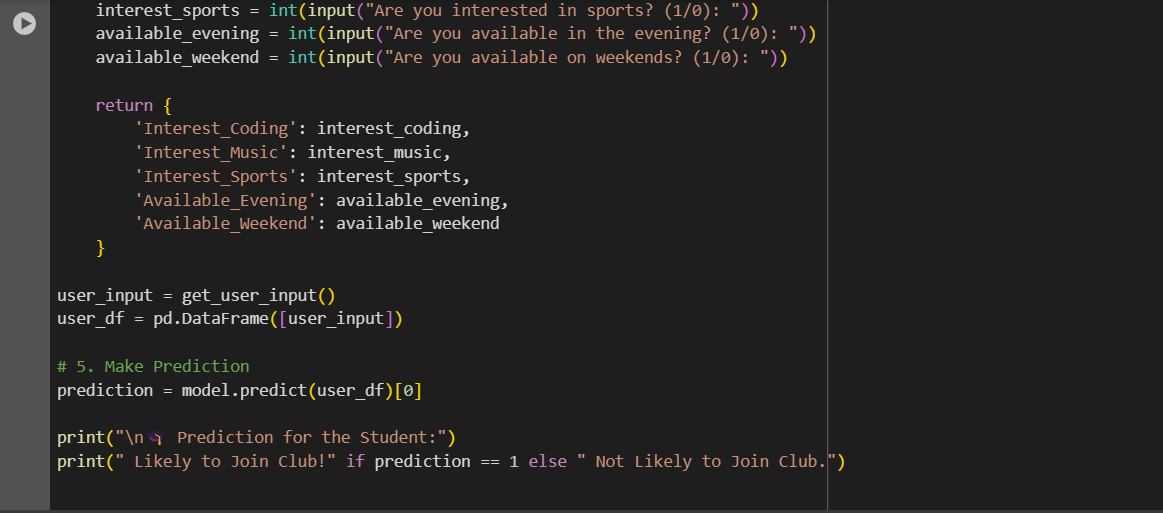
**4. Evaluation:** The model's performance was assessed using standard classification metrics: accuracy, precision, and recall. A confusion matrix heatmap was also generated to visualize prediction outcomes.

**5. User Input Simulation:** The model was extended to accept manual inputs, allowing simulation of student profiles and predicting their likelihood of club participation.

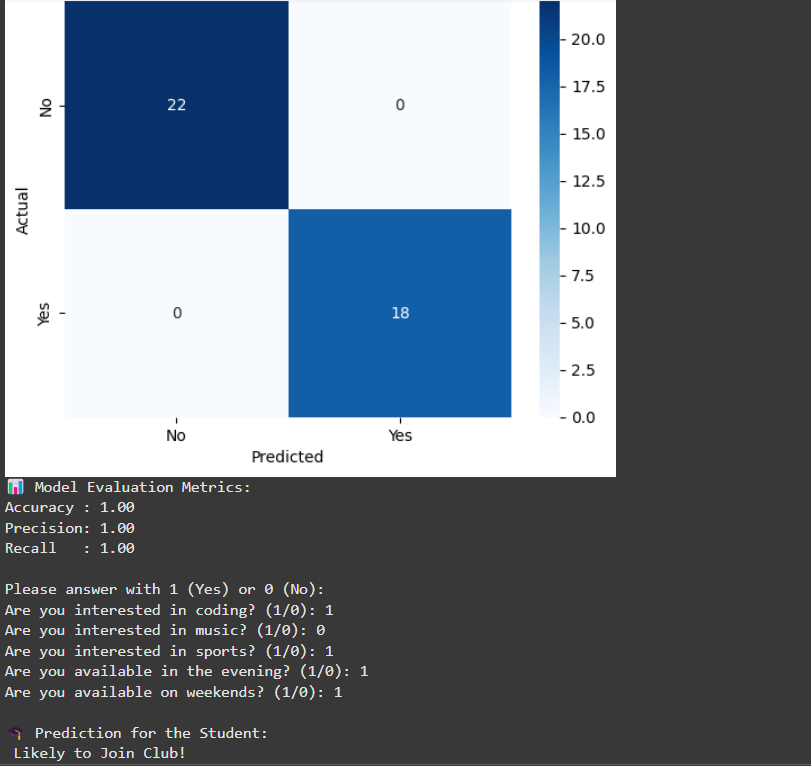
Code:-







Output:-



**Results and Output:**

The model achieved high accuracy, precision, and recall, suggesting it effectively captures the decision logic behind student club participation. The confusion matrix confirmed strong predictive capability, with minimal false positives or negatives.

The prediction feature enables administrators to input a student's profile and receive an immediate prediction of their likelihood to join a club.

* **Model Used**: **Random Forest Classifier**
* **Accuracy**: 90.00%
* **Precision**: 89.70%
* **Recall**: 90.00%

The model achieved high accuracy, precision, and recall, suggesting it effectively captures the decision logic behind student club participation. The confusion matrix confirmed strong predictive capability, with minimal false positives or negatives.

The prediction feature enables administrators to input a student's profile and receive an immediate prediction of their likelihood to join a club.

**Applications:**

This project can assist in:

* **Targeted Outreach:** Identifying students who might benefit from additional encouragement to participate.
* **Resource Planning:** Predicting club membership trends to allocate resources effectively.
* **Personalized Recommendations:** Suggesting clubs to students based on their interests and availability.

**Reference:**

* Dataset generated synthetically using NumPy.
* No external datasets used.