Demonstration:

Sample Input:

A sample from dataset with various features related to a student's profile, including Marital status, Application mode, Age at enrolment, and economic indicators like Unemployment rate, Inflation rate, and GDP.

Predicted Output:

The model predicts the outcome as 'Dropout' based on the given input.

Model Performance Metrics:

- **Accuracy:** The overall accuracy of the model is 77%, meaning the model correctly predicted the outcome for 77% of the cases.
- **Classification Report:** This report provides detailed information on precision, recall, and F1-score for each class ('Dropout', 'Enrolled', 'Graduate'). It shows that the model performs relatively well for 'Dropout' and 'Graduate', but less so for 'Enrolled'.

Work Done:

Exploratory Data Analysis (EDA):

- Explored the dataset to understand its structure and characteristics.
- Identified patterns and trends in the data.

Data Preprocessing:

- Handled missing values in the 'Debtor' column.
- Preprocessed the data to make it suitable for model training.

Model Training:

- Trained a RandomForestClassifier to predict student outcomes.
- Obtained initial model performance metrics.

Conflicts:

Missing Data in 'Debtor' Column:

- Encountered challenges in handling missing data in the 'Debtor' column.
- Plan to impute missing values using a suitable method to enhance the model's performance.

4. **Model Improvement:**

- Fine-tune the predictive model to improve accuracy and overall performance.

- Consider trying different algorithms or ensemble methods.

```
1. Demonstration:
   - Sample input: Marital status                                  4.00
Application mode                                  12.00
Application order                                  1.00
Course                                             8.00
Daytime/evening attendance                         1.00
Previous qualification                             1.00
Nacionality                                        1.00
Mother's qualification                             3.00
Father's qualification                             1.00
Mother's occupation                                6.00
Father's occupation                                6.00
Displaced                                          1.00
Educational special needs                          0.00
Debtor                                             0.00
Tuition fees up to date                            1.00
Gender                                             0.00
Scholarship holder                                 1.00
Age at enrollment                                 30.00
International                                      0.00
Curricular units 1st sem (credited)                0.00
Curricular units 1st sem (enrolled)                6.00
Curricular units 1st sem (evaluations)             7.00
Curricular units 1st sem (approved)                0.00
Curricular units 1st sem (grade)                   0.00
Curricular units 1st sem (without evaluations)     0.00
Curricular units 2nd sem (credited)                0.00
Curricular units 2nd sem (enrolled)                6.00
Curricular units 2nd sem (evaluations)            16.00
Curricular units 2nd sem (approved)                1.00
Curricular units 2nd sem (grade)                  10.00
Curricular units 2nd sem (without evaluations)     0.00
Unemployment rate                                 11.10
Inflation rate                                     0.60
GDP                                                2.02
```

```
Name: 1255, dtype: float64
   - Predicted output: Dropout
   - Accuracy: 0.77
   - Classification Report:
              precision    recall  f1-score   support

     Dropout       0.84      0.77      0.80       316
    Enrolled       0.53      0.32      0.40       151
    Graduate       0.78      0.94      0.85       418

    accuracy                           0.77       885
   macro avg       0.72      0.67      0.68       885
weighted avg       0.76      0.77      0.76       885


2. Work Done:
   - Performed exploratory data analysis (EDA) on the dataset.
   - Conducted data preprocessing and handled missing values.
   - Trained a preliminary model using RandomForestClassifier.

3. Conflicts:
   - Encountered challenges in handling missing data in the 'Debtor' column.
   - Plan to mitigate by imputing missing values using a suitable method.

4. Next Steps:
   - Fine-tune the predictive model for better accuracy.
   - Incorporate additional features for a more comprehensive analysis.
   - Improve visualizations for better interpretation.
```