Analysis Of The Code

1. **Data Preparation (What steps would you take to prepare your data? Discuss your approach)**

   Ans - I performed the following steps to prepare my data

- I added CustomerChurn.csv path so as to read data from that file
- **Encoding**: I did encoding to convert Categorical data to numeric data. I encoded two columns, 'Gender' and 'Attrition_Flag', to numeric values using the map() function.
- **Scaling**: I performed scaling using StandardScaler to standardize the features
- **Balancing**: Using the SMOTE() function the class was balanced.

2. **Model Hyperparameter Tuning (Which hyperparameters would you tune and why? How would you tune them?)**

   Ans - I have used GridSearchCV with the Random Forest classifier. I have used the n_estimator hyperparameter because It focuses on tuning the number of trees (n_estimators) in the forest. Also, It prints the best parameters and precision score, which helps us select the optimal model.

3. **Choice of Evaluation Metric (Which metric would be suitable for model evaluation and why?)**

   Ans - I have used precision metrics. The precision metric would be suitable for model evaluation because it focuses on measuring the accuracy of positive predictions.

4. **Overfitting avoidance mechanism (Which mechanism (feature Selection/ regularization) would you use and why?)**

   Ans - I have used the regularization mechanism to avoid overfitting.
   Regularization through the C parameter helps us balance the data by fitting the training data properly to avoid Overfitting.

5. **Results analysis**
   a. **Which of the two models (random forest or support vector classifier) would you recommend for deployment in the real-world?**

   Ans – I would suggest both Random Forest and Support Vector Machine classifiers are suitable for deployment in the real world.

   - Random Forest combines the result of multiple trees and accordingly calculates precision.

- Support Vector Machine can be used when a large number of features are involved

b. Is any model underfitting? If yes, what could be the possible reasons?
Ans - No, No model is underfitted as I have used hyperparameter tuning.

.