

Smote On Random Forest Classifier

Date	15 July 2024
Team ID	740091
Project Title	Thyroid disease classification using machine learning
Maximum Marks	5 Marks

```
[ ] from imblearn.over_sampling import SMOTE
    from imblearn.under_sampling import RandomUnderSampler
    from imblearn.pipeline import Pipeline
    from sklearn.ensemble import RandomForestClassifier
```

```
[ ] rus = RandomUnderSampler()
```

```
[ ] smote = SMOTE(sampling_strategy='auto')
```

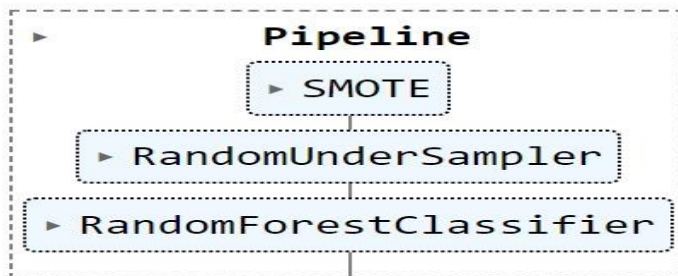
```
[ ] rf_classifier = RandomForestClassifier()
```

```
[ ] pipeline = Pipeline([
    ('smote', smote),
    ('rus', rus),
    ('rf', rf_classifier)
])
```

```
[ ] pipeline.fit(x_train,y_train)
```



```
[ ] pipeline.fit(x_train,y_train)
```



```
[ ] y_pred1 = pipeline.predict(x_test)
```

```
[ ] accuracy = accuracy_score(y_test,y_pred1)
    print("Accuracy:",accuracy)
```



Accuracy: 0.5511383482043629

```
[ ] #evaluating model
plt.figure(figsize=(5,5))
sns.histplot(y_test_pred,bins=20,kde=True)
plt.title("prediction error")
```

↔ Text(0.5, 1.0, 'prediction error')

