

```
In [ ]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
sns.set(style="white")
import warnings
warnings.simplefilter("ignore")
```

```
In [ ]: from sklearn.preprocessing import StandardScaler, RobustScaler
from sklearn.pipeline import Pipeline
from sklearn.decomposition import PCA

from sklearn.svm import SVC
from sklearn.neighbors import KNeighborsClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier, ExtraTreesClassifier, AdaBoostClassifier

from sklearn.model_selection import train_test_split
from sklearn.model_selection import GridSearchCV
from sklearn.model_selection import cross_val_score, cross_validate
from sklearn.model_selection import KFold, StratifiedKFold
from sklearn.metrics import classification_report, f1_score, accuracy_score, confusion_matrix

from imblearn.over_sampling import SMOTE
from imblearn.combine import SMOTEENN
from imblearn.pipeline import make_pipeline
```

```
In [ ]: # Load dataset
df=pd.DataFrame.from_csv("glass.data.txt", sep=",", header=None, index_col=None)
# drop first column
df=df[df.columns[1:11]]
# Add columns name
df.columns=["RI", "Na", "Mg", "Al", "Si", "K", "Ca", "Ba", "Fe", "Class"]
df.head(3)
```

```
In [ ]: print(df.info())
```

```
In [ ]: #Let's see the distribution of glass types.
```

```
In [ ]: # glass type histogram
f, ax = plt.subplots(figsize=(5, 2))
sns.countplot(y="Class", data=df, palette='Paired')
sns.despine()
plt.title('Plot 1.5 Glass types histogram')
```

```
In [ ]: df['Class'].value_counts()
```