

Worksheet: Model Evaluation & Tuning

This worksheet demonstrates loading the Breast Cancer dataset, training a model, and optimizing it using GridSearchCV.

1. Data Loading & Setup

```
import pandas as pd
from sklearn.datasets import load_breast_cancer
from sklearn.model_selection import train_test_split

data = load_breast_cancer()
df = pd.DataFrame(data.data, columns=data.feature_names)
df['target'] = data.target
```

2. Model Training & Grid Search

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import GridSearchCV

rf = RandomForestClassifier()
param_grid = {'n_estimators': [50, 100], 'max_depth': [None, 10, 20]}
grid_search = GridSearchCV(rf, param_grid, cv=5)
grid_search.fit(X_train, y_train)
```

3. Performance Reports

Classification Report:

	precision	recall	f1-score	support
0	0.98	0.94	0.96	63
1	0.96	0.99	0.98	108
accuracy			0.97	171
macro avg	0.97	0.96	0.97	171
weighted avg	0.97	0.97	0.97	171

Mean Squared Error:

0.0292