Hands-on: scikit-learn for Biomedical Data

Pipeline Creation

```python
from sklearn.pipeline import Pipeline
pipe = Pipeline([
    ('scaler', StandardScaler()),
    ('selector', SelectKBest()),
    ('clf', LogisticRegression())
])
```

Model Selection

```python
from sklearn.model_selection import
    GridSearchCV
gs = GridSearchCV(pipe, params,
    cv=StratifiedKFold(5),
    scoring='roc_auc')
```

🧪 Practice Tasks

- Load biomedical dataset
- Handle missing values
- Scale features appropriately
- Deal with class imbalance

- Build classification pipeline
- Perform nested CV
- Generate evaluation report
- Plot ROC and PR curves