AUTOPOCKET

AUTOML PYTHON PACKAGE

Natalia Choszczyk
Filip Langiewicz
Maciej Momot
Mikołaj Rowicki
Krzysztof Tkaczyk

AGENDA

1. INSPIRATION
2. ABOUT
3. TARGET & SPECIALIZATION
4. COMPARISON WITH OTHER TOOLS
5. HOW TO USE IT?
6. PREPROCESSOR
7. MODELLER
8. POSTPROCESSOR
9. OUTPUT EXAMPLE
10. VISUALIZATIONS
11. QUESTIONS

”

Sztuczna inteligencja zastąpi człowieka w bankowości już za 5-10 lat

AUTOPOCKET:

ECONOMY

Framework specializing in financial data, credit scoring and insurance.

EXPLAINABILITY

Detailed explanations using LIME, SHAP and dependency plots.

SIMPLICITY

Very simple API: *doJob()* handles the entire pipeline.

TARGET & SPECIALIZATION

COMPARISON

Package	Best Score (ROC-AUC)	Key Strengths
Autopocket	0.8387	Comprehensive interpretability via LIME and SHAP. Tailored for financial datasets, offering detailed preprocessing and explanations.
MLJAR (Explain)	0.8184	Speed and simplicity. Suitable for rapid prototyping.
AutoGluon	0.851	High performance and ensembling. Designed for larger datasets with less emphasis on interpretability.

Note: Training performed on credit.csv dataset and restricted to tree-based and linear models due to banking regulations (no boostings).

HOW TO USE IT?

1. INSTALLATION

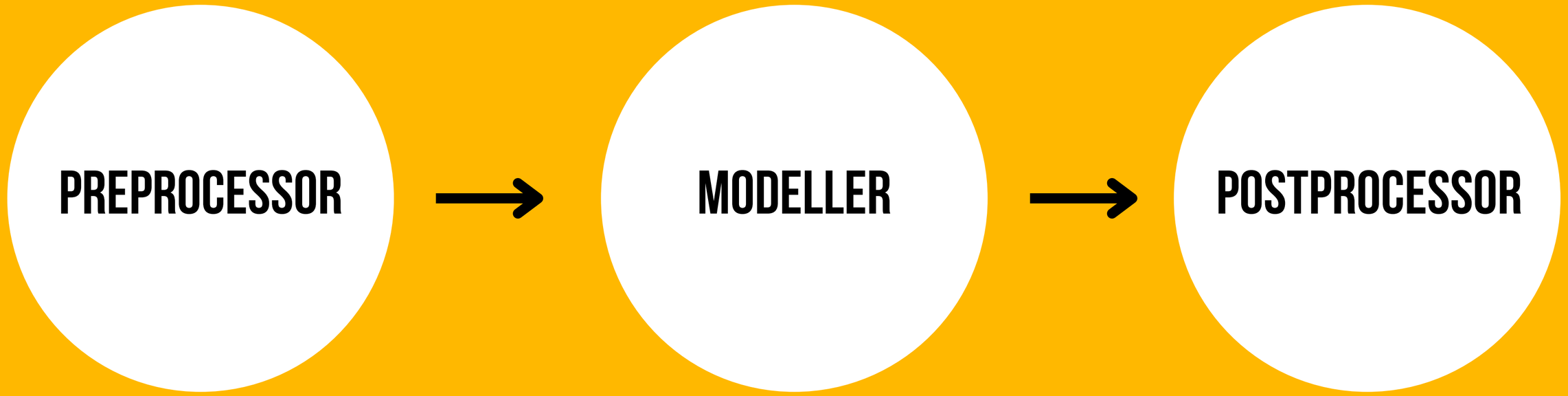
```
pip install autopocket
```

2. USER GUIDE

```
from autopocket import AutoPocketor

AutoPocketor().doJob(
    path = “path/to/your/data.csv”,
    target = “target”
    )
```

3. ENJOY TIME SAVED


PREPROCESSOR
MODELLER
POSTPROCESSOR

PREPROCESSOR

Column type analyzer
Binary column handler
Data imputer
Date handler
Number format fixer
Pattern remover
Redundant column remover
String stripper
String to lower converter
Feature encoder
Feature selector
Outlier handler

MODELLER

LogisticRegression
RandomForestClassifier
RidgeClassifier
DecisionTreeClassifier
LinearRegression
Lasso
Ridge
DecisionTreeRegression
RandomForestRegression
ElasticNet
LassoLarsIC

POSTPROCESSOR

Models Leaderboard

ICE Plotter

Lime Explainer

Lime Plotter

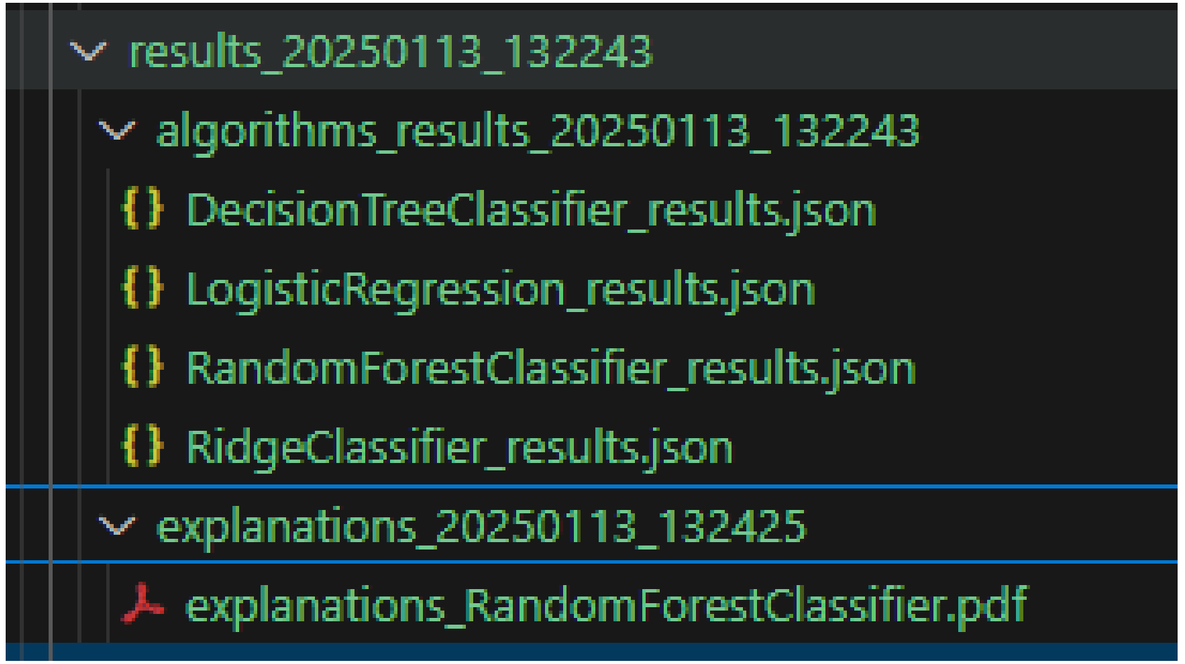
Lime Top Feature Selector

Partial Dependence Plotter

SHAP Explainer

SHAP Plotter

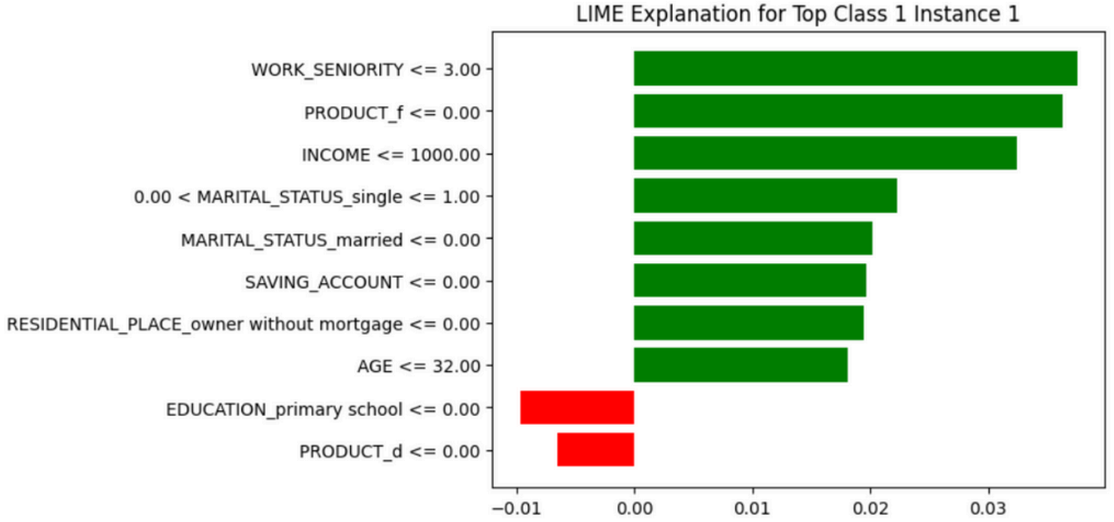
OUTPUT EXAMPLE

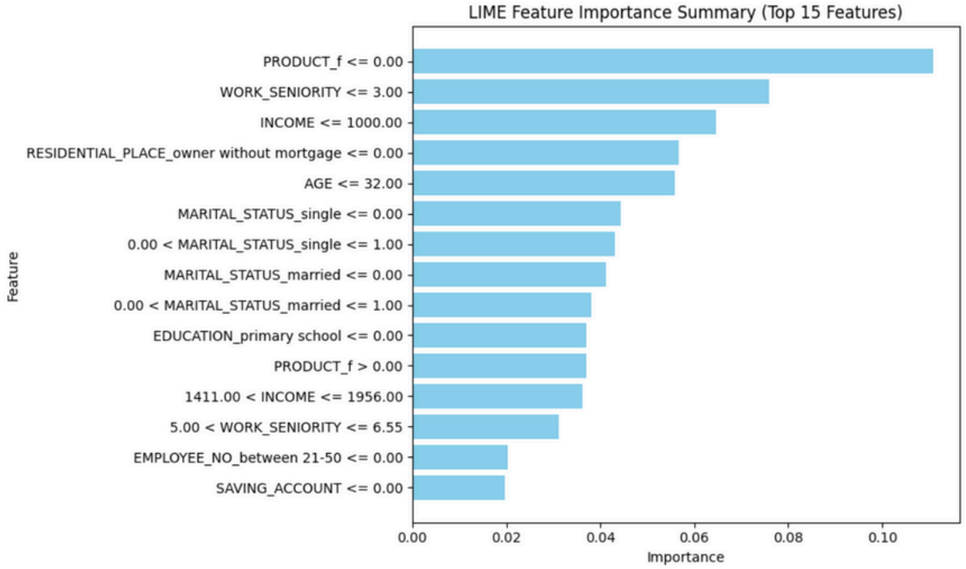


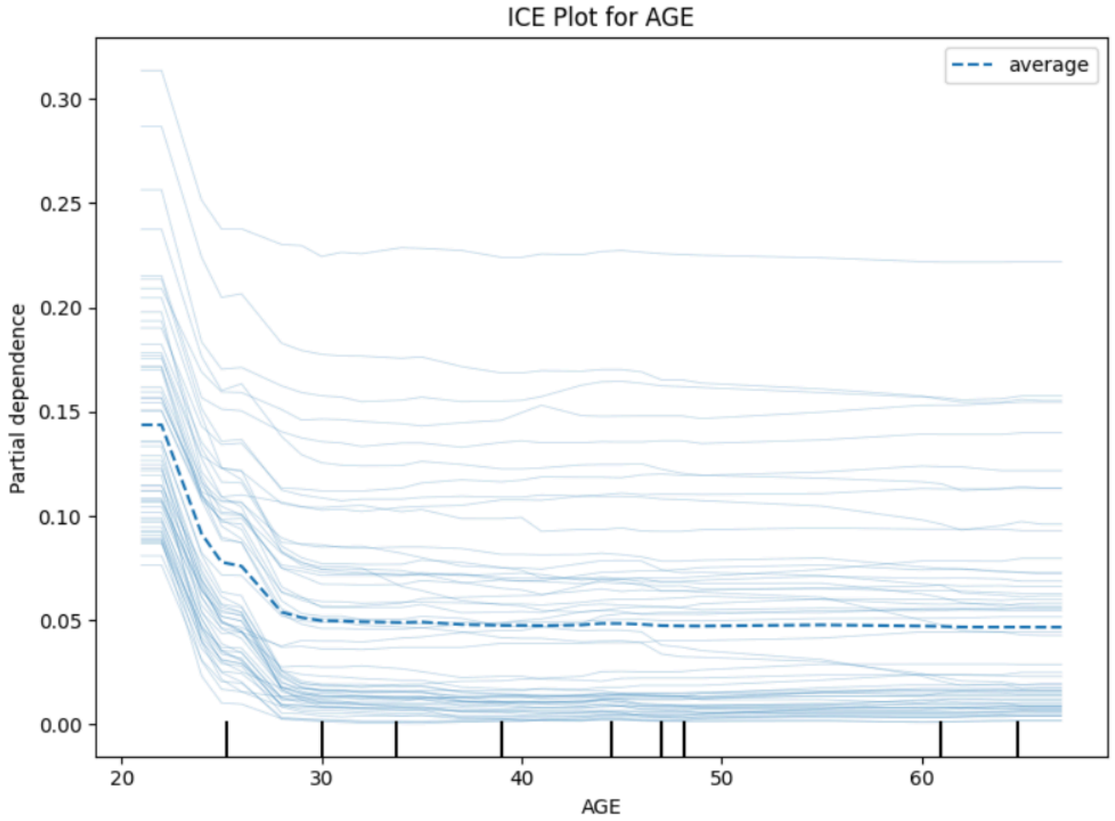

```
Performing preprocessing...
X shape: (5859, 62)
Preprocessing done.

Performing modelling...
Performing binary classification
Fitting dummy estimator
Dummy score (strategy: most_frequent): 0.5 roc_auc
Fitting 4 models
1 / 4  | Fitting: RandomForestClassifier. Best score: 0.838996859781005 roc_auc
2 / 4  | Fitting: LogisticRegression. Using ['lbfgs'] solver.Best score: 0.828229636931441 roc_auc
3 / 4  | Fitting: DecisionTreeClassifier. Best score: 0.7965988540997622 roc_auc
4 / 4  | Fitting: RidgeClassifier. Some models did not converge. Best score: 0.8311035126435768 roc_auc
Saving results to results/algorithms_results
Chosen model: RandomForestClassifier
Modelling done.
```

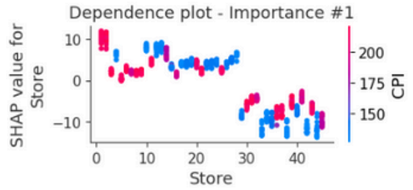
VISUALIZATIONS

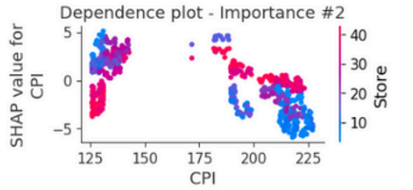


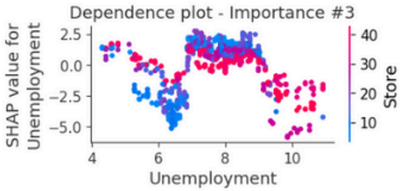


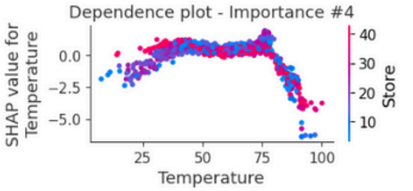


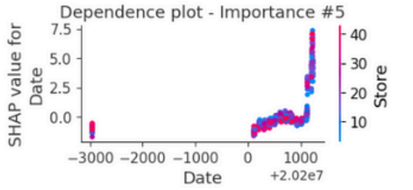
VISUALIZATIONS

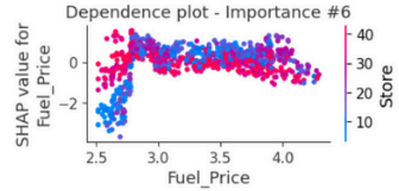


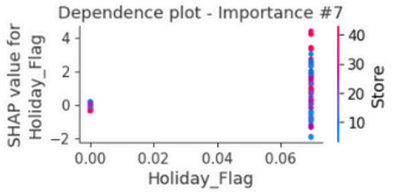


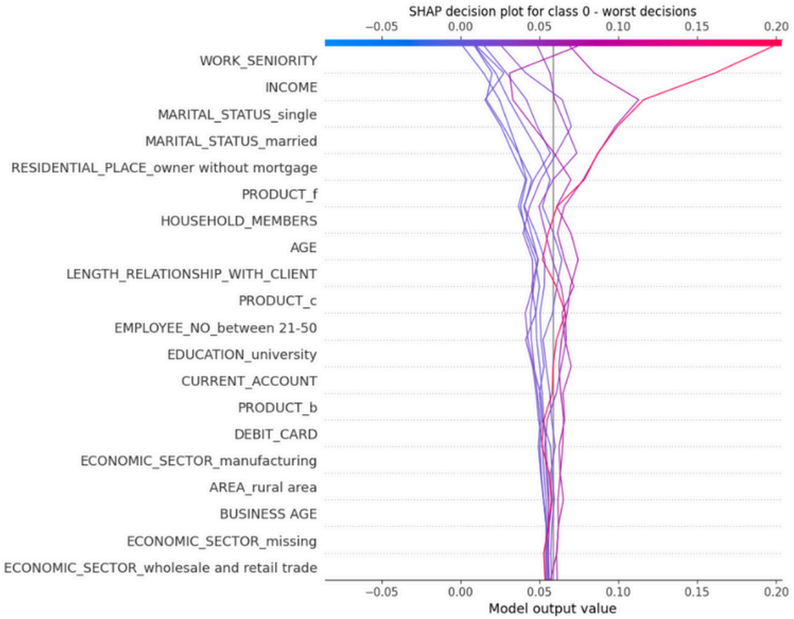












QUESTIONS

THANK YOU