

Context

- Automated selection & parametrization of machine learning algorithms.
- Guided Hyperparameters optimization.
- Optimal performance of ML models for a given classification task.
- Explainability of the recommended models.
- Application to the Industry 4.0.
- Empirical study on manufacturing data for validation and usability purposes.

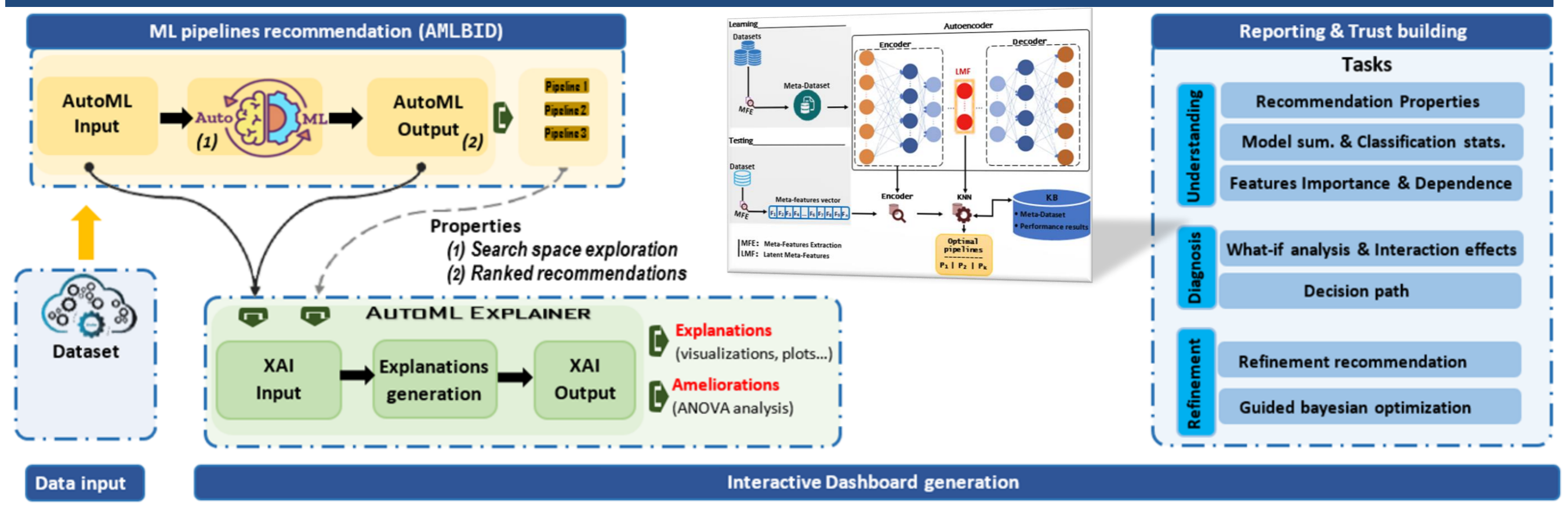
Key concepts

Automated Machine Learning (AutoML) Auto ML is often used to help domain experts, who typically have limited ML expertise, in order to generate and build high quality models to better meet their specific business needs.

Meta-learning refers to the algorithms that are concerned with their own learning process as well as learning across a series of related prediction tasks.

Explainable AutoML (XAutoML) provide a set of tools and frameworks to better understand and interpret the predictions of a machine-learning model.

Proposed assistance system



Recommender module

Suggested configurations

Recommendation 1: RandomForestClassifier Expected accuracy: 0.97917 [Export Pipeline](#)

Recommendation 2: GradientBoostingClassifier Expected accuracy: 0.97826 [Export Pipeline](#)

Gradient Boosting Classifier

Class sklearn.ensemble.GradientBoostingClassifier

(loss, learning_rate, n_estimators, subsample, criterion, min_samples_split, min_samples_leaf, min_weight_fraction_leaf, max_depth, min_impurity_decrease, min_impurity_split, init, random_state, max_features, verbose, max_leaf_nodes, warm_start, validation_fraction, n_iter_no_change, tol, cc_p, alpha)

Gradient boosting classifiers are a group of machine learning algorithms that combine many weak learning models together to create a strong predictive model. Gradient boosting models are becoming popular because of their effectiveness at classifying complex datasets, and have recently been used to win many Kaggle data science competitions.

[Learn more >>](#)

Recommended model configuration

Hyperparameter	Value
n_estimators	50
min_impurity_decrease	0.0
max_features	sqrt
learning_rate	1.0
loss	deviance
random_state	324089

Explainer module

What-If Analysis

Select Observation: Select from list or pick at random

Input: 1218

Rank: Index

Feature	Value	Rank
T4	69	0
T74	210	1
T3	0	2

Prediction: label probability

label	probability
0	32.9 %
1	0.0 %
2	67.1 %

Contributions Table: How has each feature contributed to the prediction?

Reason	Effect
Average of population	95.49%
T4U = 7	-24.56%
T4 = 69	-1.84%

Contributions Plot: How has each feature contributed to the prediction?

Depth: 6, Sorting: Importance

Decision Path: Visualizing individual predictions process

Select Observation: 1218

Decision path: Visualizing individual predictions process

AMLBIID package

AMLBIID is a self-explainable AutoML system in the form of a Python-package. The system proposes a transparent and justified analysis to discover the most suitable model for optimal performance among multiple ML models. It attempts to automate the process of the algorithms selection, the tuning of hyperparameters, and traceability in supervised ML.

```

1 from AMLBIID.recommender import AMLBIID_Recommender
2 from AMLBIID.explainer import AMLBIID_Explainer
3 from AMLBIID.loader import *
4
5 #Load dataset
6 Data, X_train, Y_train, X_test, Y_test = load_data("Dataset.csv")
7
8 #Generate the optimal configurations
9 model, config = AMLBIID_Recommender.recommend(Data,
10                                               metric="Accuracy",
11                                               mode="Recommender_Explainer")
12 model.fit(X_train, Y_train)
13
14 #Generate the interactive explanatory dash
15 Explainer = AMLBIID_Explainer.explain(model, config,
16                                       X_test, Y_test)
17 Explainer.dash()

```

Results

- Garouani, M., Ahmad, A., Bouneffa, M., et al. Using meta-learning for automated algorithms selection and configuration: an experimental framework for big industrial data. *Journal of Big Data* 9, 57 (2022). <https://doi.org/10.1186/s40537-022-00612-4>
- Garouani, M., Ahmad, A., Bouneffa, M., et al. Towards big industrial data mining through explainable automated machine learning. *The International Journal of Advanced Manufacturing Technology* (2022). <https://doi.org/10.1007/s00170-022-08761-9>
- Garouani, M., Ahmad, A., Bouneffa, M., et al. AMLBIID: An Automated Machine Learning tool for Big Industrial Data. *SoftwareX* (2021) 100919, <https://doi.org/10.1016/j.softx.2021.100919>
- Click or Scan the QR Code to explore all results and publications



ACKNOWLEDGMENTS

The contributors gratefully acknowledge the University of the Littoral Côte d'Opale, School of engineering's and business sciences and technics, and CNRST Morocco for the financial support, and the CALCULCO computing platform, supported by SCOSI/ULCO for the computational facilities.