

## Motivation

This study falls within the design and the development of a **decision support** and **expert system**. The main objective of this study is focused on the **automatic selection** and **parametrization** of ML models. The major goal is to achieve an optimal performance for a given task while providing the **rationale** traceability behind a recommendation or decision. The designed system is particularly aimed at the provision of explanations of such rationale traceability and promising trend analysis of the area of big industrial data. The empirical studies are hence carried out with respect to the big data analysis for industry 4.0 actors (engineers and researchers).

## 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.

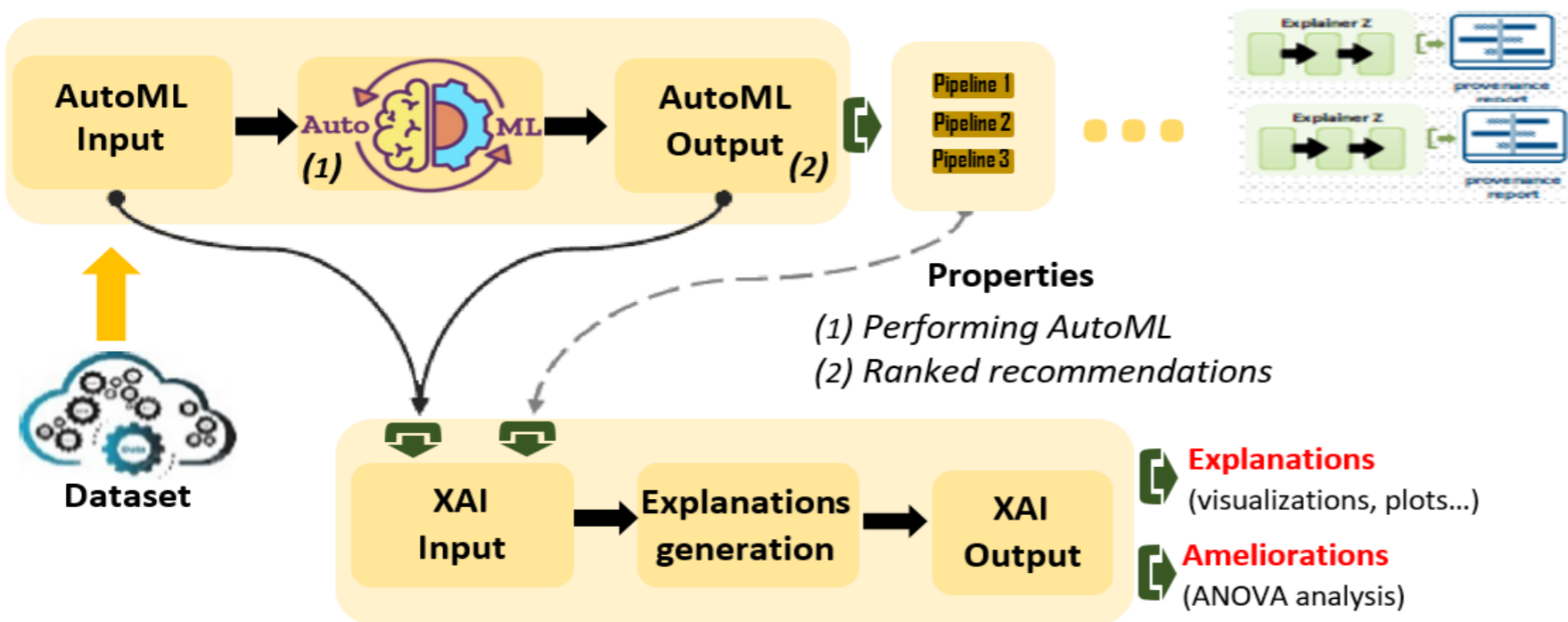
**Explainable AI (XAI)** provide a set of tools and frameworks to better understand and interpret the predictions of a machine-learning model.

## Proposed system

### Recommender module (AMLBID)

### Comparative Analysis

### Reporting & Trust building



Understanding

Diagnosis

Refinement

### Tasks

- Recommendation Properties
- Model sum. & Classification stats.
- Features Importance & Dependence
- What-if analysis & Interaction effects
- Decision path
- Recommendation refinement

### Data Input

### Explanatory module (AMLExpainer)

### Interactive dash generation

## Recommender module

## Explainer module

## AMLBID package

AMLBID 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 machine learning models. It attempts to automate the process of algorithm selection, the tuning of hyperparameters, and traceability in supervised machine learning.

```

1 from AMLBID.recommender import AMLBID_Recommender
2 from AMLBID.explainer import AMLBID_Explainer
3 from AMLBID.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=AMLBID_Recommender.recommend(Data,
10                                           metric="Accuracy",
11                                           mode="Recommender_Explainer")
12
13 model.fit(X_train, Y_train)
14
15 #Generate the interactive explanatory dash
16 Explainer = AMLBID_Explainer.explain(model, config,
17                                     X_test, Y_test)
18 Explainer.dash()

```

## Perspectives

- Expand AMLBID to support classifiers of distributed machine-learning libraries.
- Upload the AMLBID to the PyPI/Conda-forge package index to facilitate its distribution and use.

[1]. Garouani, M.; Ahmad, A.; Bouneffa, M.; Lewandowski, A.; Bourguin, G. and Hamlich, M. (2021). *Towards the Automation of Industrial Data Science: A Meta-learning based Approach*. In Proceedings of the 23rd International Conference on Enterprise Information Systems - Volume 1: ICEIS, , pages 709-716. DOI: 10.5220/0010457107090716

[2]. Garouani, M.; Hamlich, M.; Ahmad, A.; Bouneffa, M.; Lewandowski, A.; Bourguin, G. and (2021). *Towards an automatic assistance framework for the selection and configuration of machine-learning-based data analytics solutions in industry 4.0*. The Fifth International Conference on Big Data and Internet of Things (BDIoT'21), Rabat, Morocco.