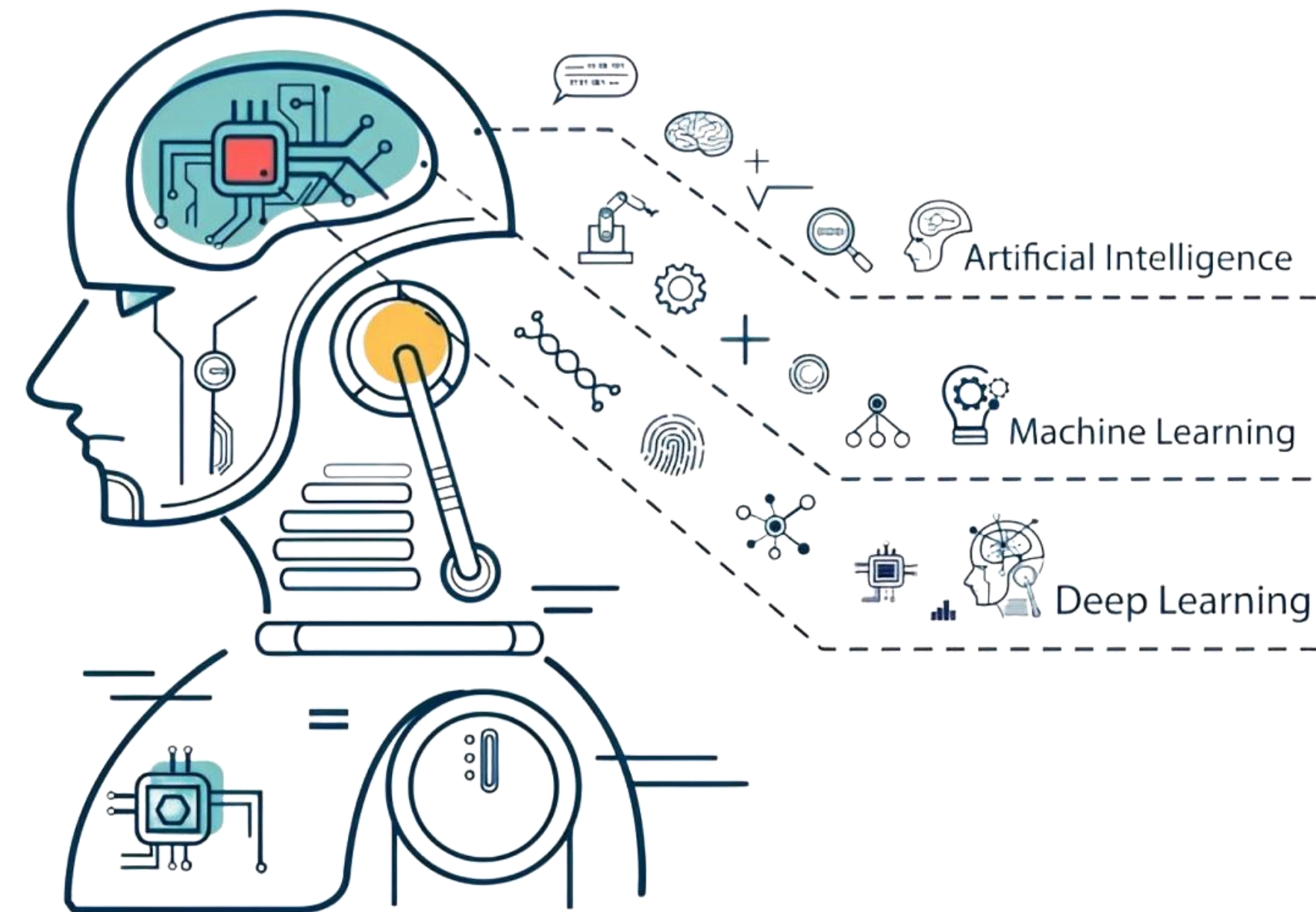


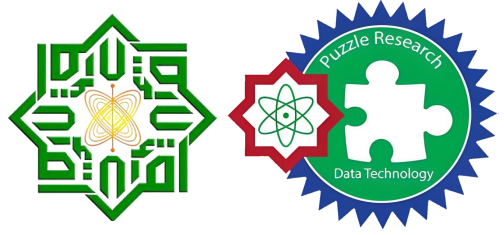
CAMPRODATIC

Machine Learning Ramadhan 1445 H / 2024

Ensemble Learning : Stack Models Concept, Method & Tutorial

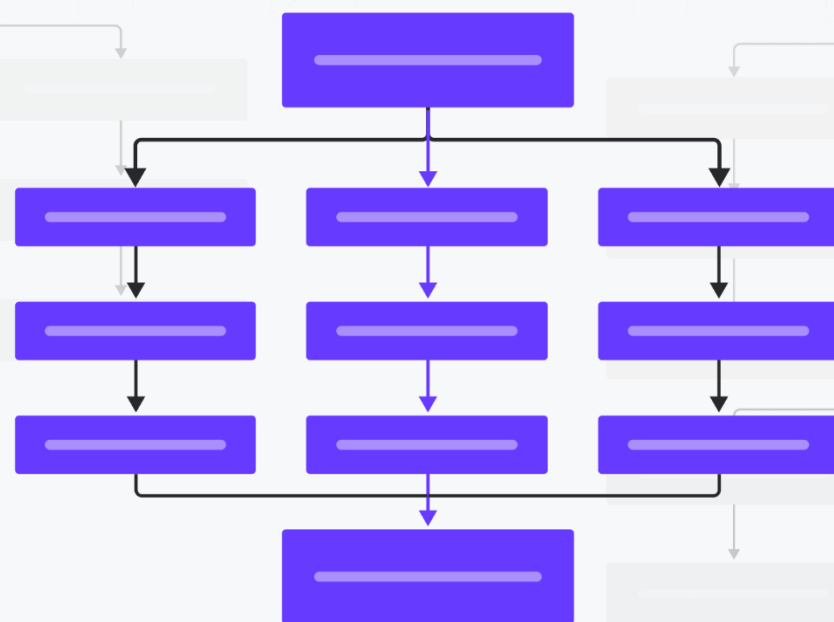
Muhammad Fauzi Fayyad
6 Ramadhan 1445 H/ 17 March 2024



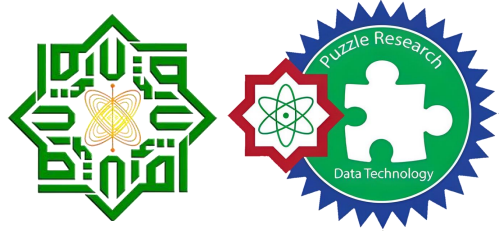


Concept of Ensemble Learning

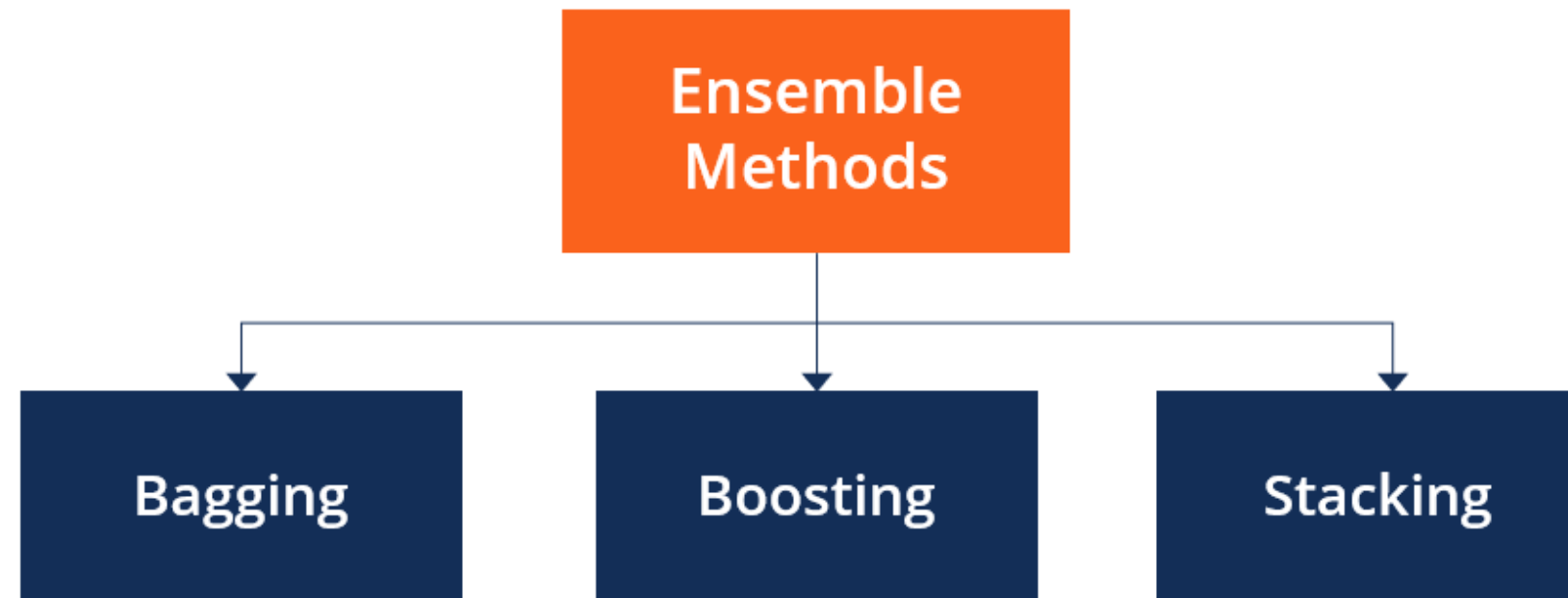
Ensemble learning



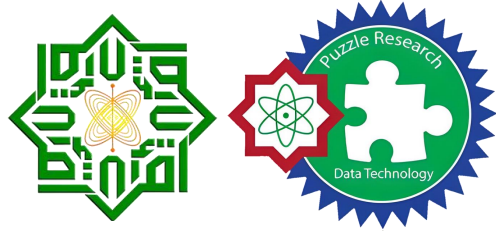
- Ensemble learning was first introduced in 1979 by Dasarathy and Sheela's research. This research became the initial foundation of ensemble learning.
- Ensemble learning combines several ML algorithms to improve modelling performance. maximising the collective knowledge of various models, resulting in better prediction performance.
- Ensemble learning is divided into parallel and sequential ensemble.



Ensemble Methods

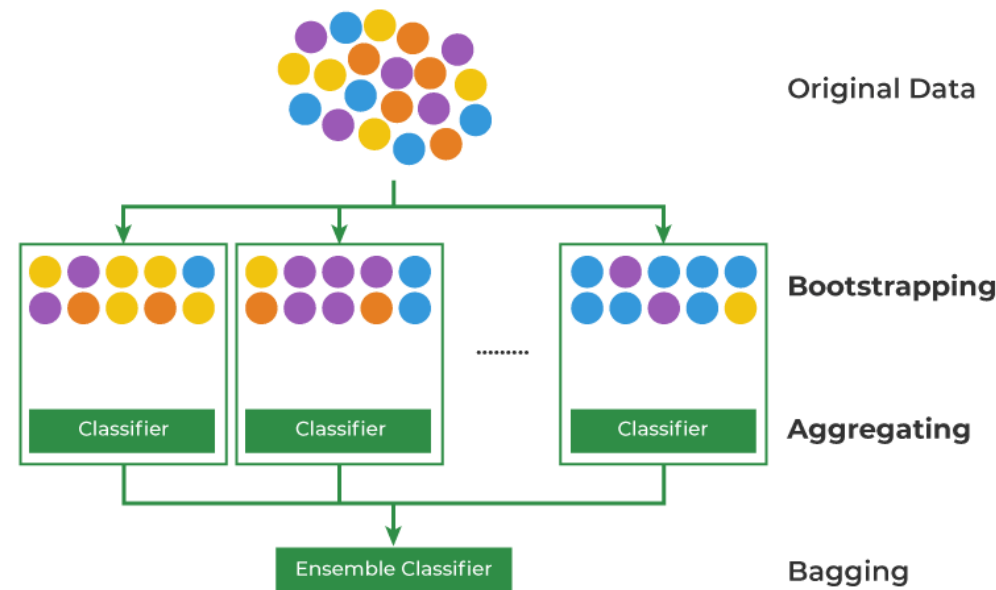


- Bagging (Bootstrap Aggregating) is the process of building multiple models on **bootstrap samples** and **aggregating** their predictions through voting or averaging.
- Boosting is the process of **sequentially** training **weak learners**, with a focus on correcting errors made by previous learners.
- Stacking is combining multiple base models to form a **meta-model**.

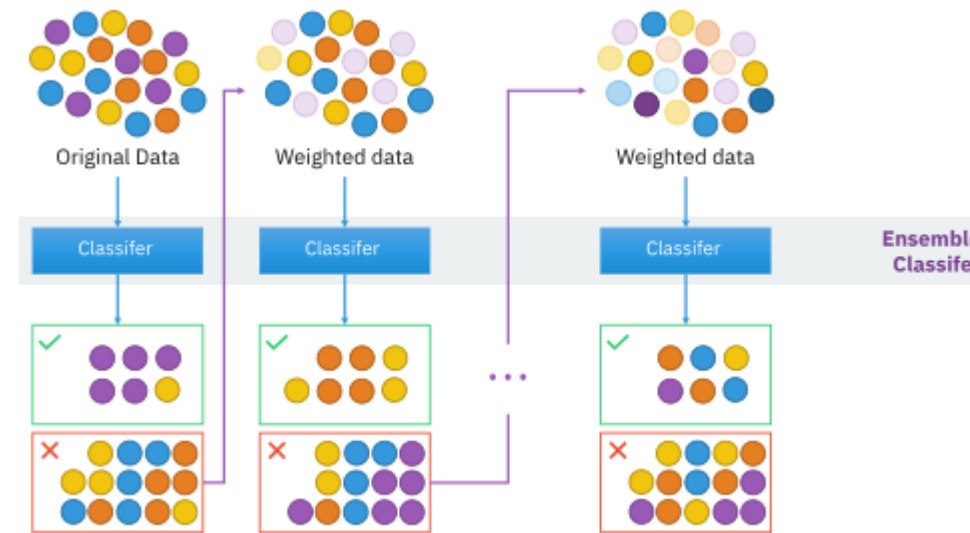


Perbedaan Ensemble Learning

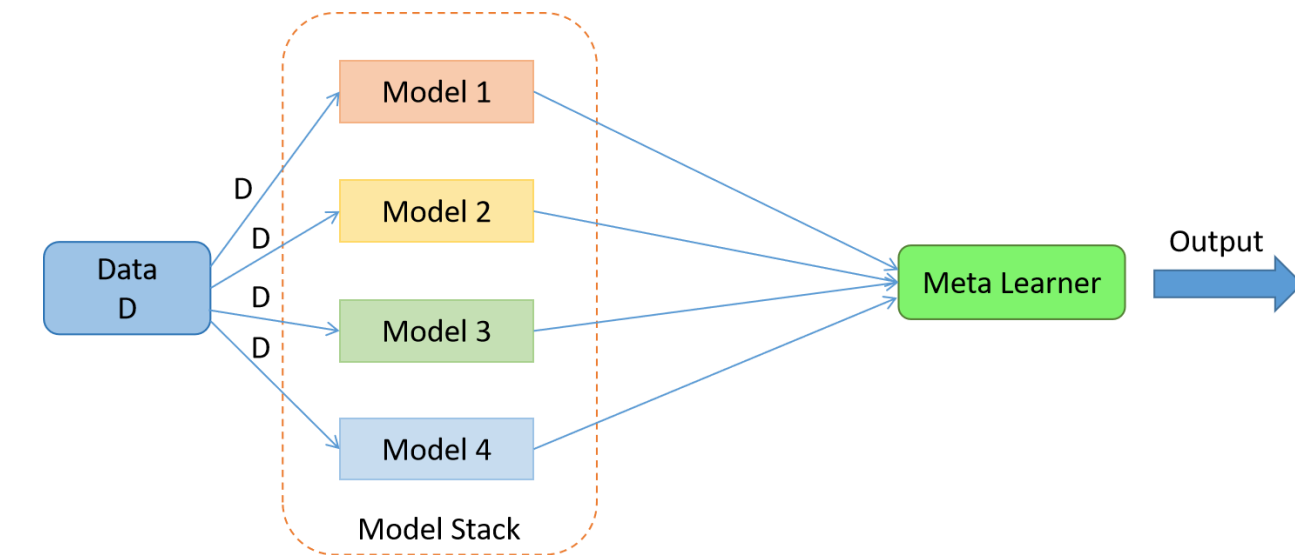
1. Bagging



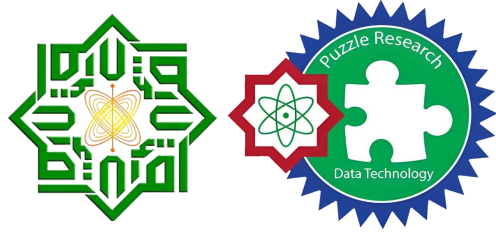
2. Boosting



3. Stacking

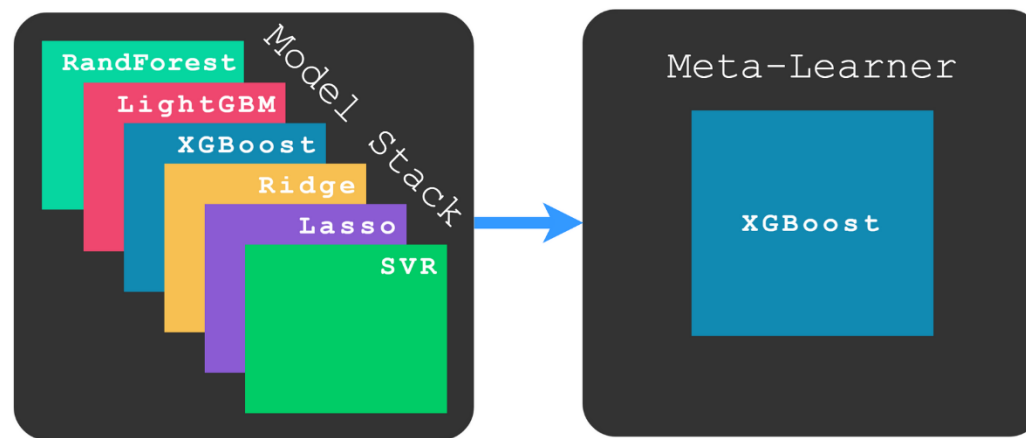


- The training process is performed in **parallel**
- It is comprised of one and the same algorithm
- Example Random Forest, and Extra Tree Algorithm
- The training is performed **sequentially**
- It is comprised of one and the same algorithm
- Example of AdaBoost, Gradient Boosting, XGBoost and CatBoost Algorithm
- **The base learner output** becomes the input for the meta learner
- It is composed of more than 1 different algorithm
- Stack model made up of **Base learner** and **Meta Learner**



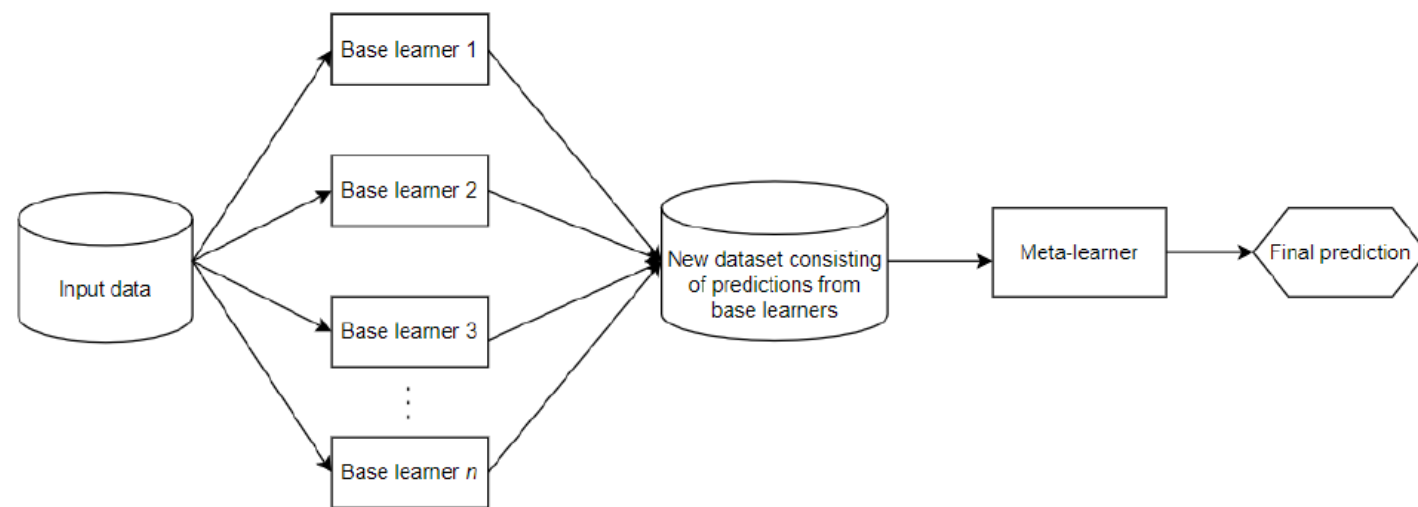
Konsep Stacking

Model Stacking with Layers:
A Machine Learning Ensemble Technique



- Stacked generalisation (Stacking) is an ensemble learning framework that trains various ML algorithms by combining predictions from two or more ensemble members.
- Stacking involves model building using several base algorithms, called **level-0 models**, and a meta-learning algorithm to **aggregate the predictions** from those base models. These meta-models are referred to as **level-1 models**.
- Meta-learning is a part of machine learning where algorithms are trained using the output of other ML algorithms and provide more accurate predictions based on the predictions made by other base classifiers.

Alur Proses Stacking



The stacking framework consists of several steps:

1. Training the base model

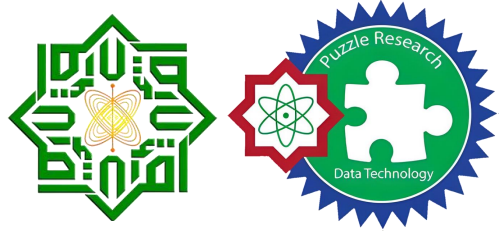
Several algorithms will be trained as base models using training data.

2. Creating a New Dataset

The predicted results (output) of the base model are used to create a new dataset. The predicted target labels from the base model are taken as new features, while the actual target labels are taken as labels in the new dataset.

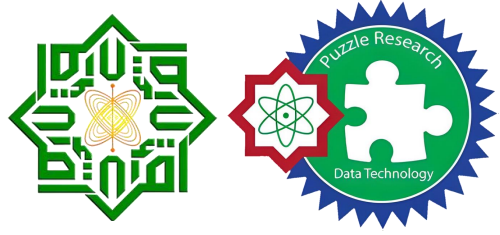
3. Training the Meta-Learner model

The algorithm selected as the Meta-Learner model will be trained using the new dataset.

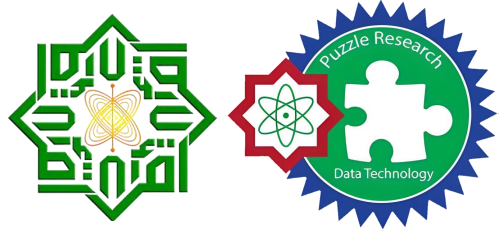


How to Evaluate the Performance of Classification Machine Learning Models

- After training the stacking model. Next, the model will be evaluated to measure the generalisation ability of the model in classifying a given data. Some metrics to measure the performance of the classifier model are Accuracy, Precision, Recall and F1-Score.
- Accuracy metric is one of the techniques to evaluate the performance of a classification algorithm based on the correct grouping of data. The **higher the Accuracy value**, the **better** the resulting classifier model.



Implementation of Ensemble Stacking Models with Python



References :

- Rincy, T. N., & Gupta, R. (2020, February). Ensemble learning techniques and its efficiency in machine learning: A survey. In *2nd international conference on data, engineering and applications(IDEA)* (pp. 1-6). IEEE.
- Mienye, I. D., & Sun, Y. (2022). A survey of ensemble learning: Concepts, algorithms, applications, and prospects. *IEEE Access*, 10, 99129-99149.

Thank You

