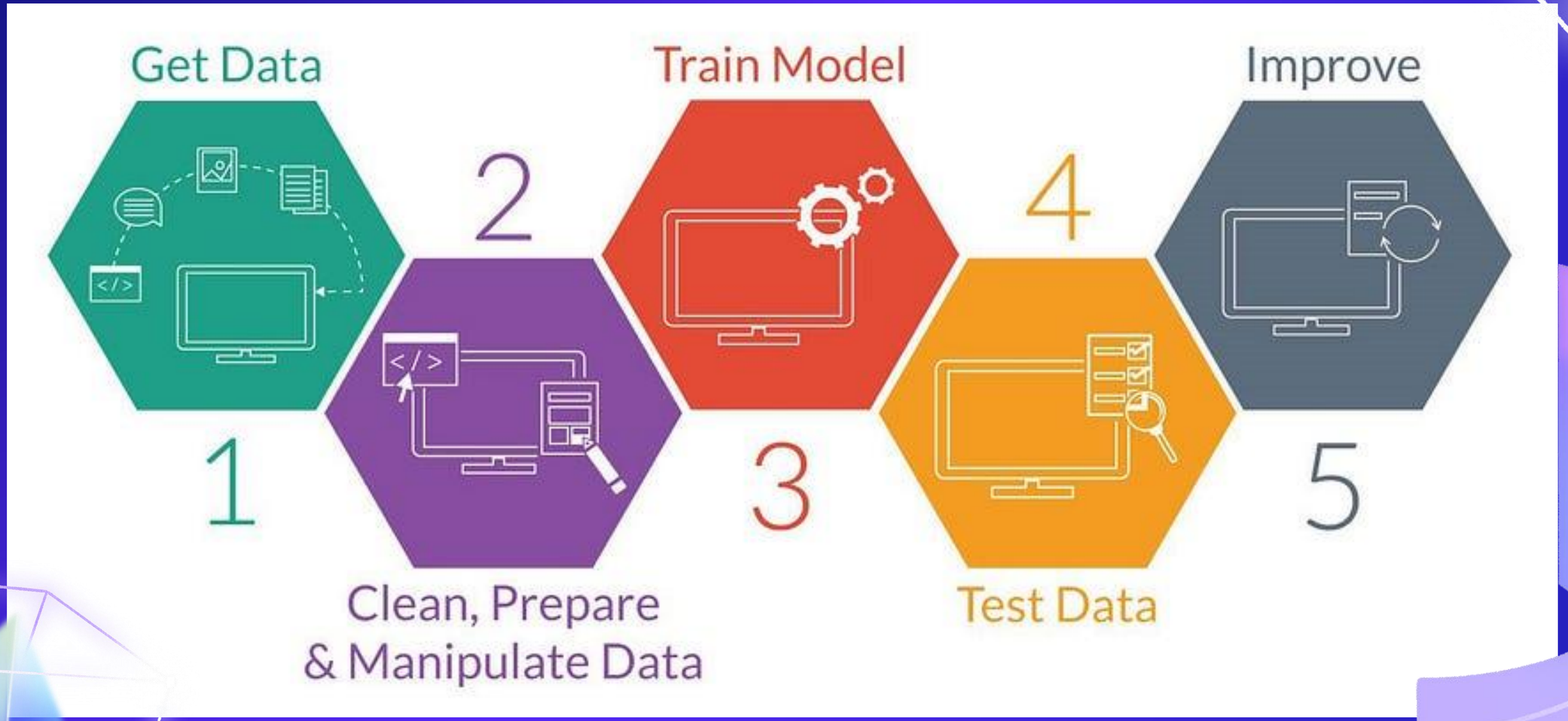
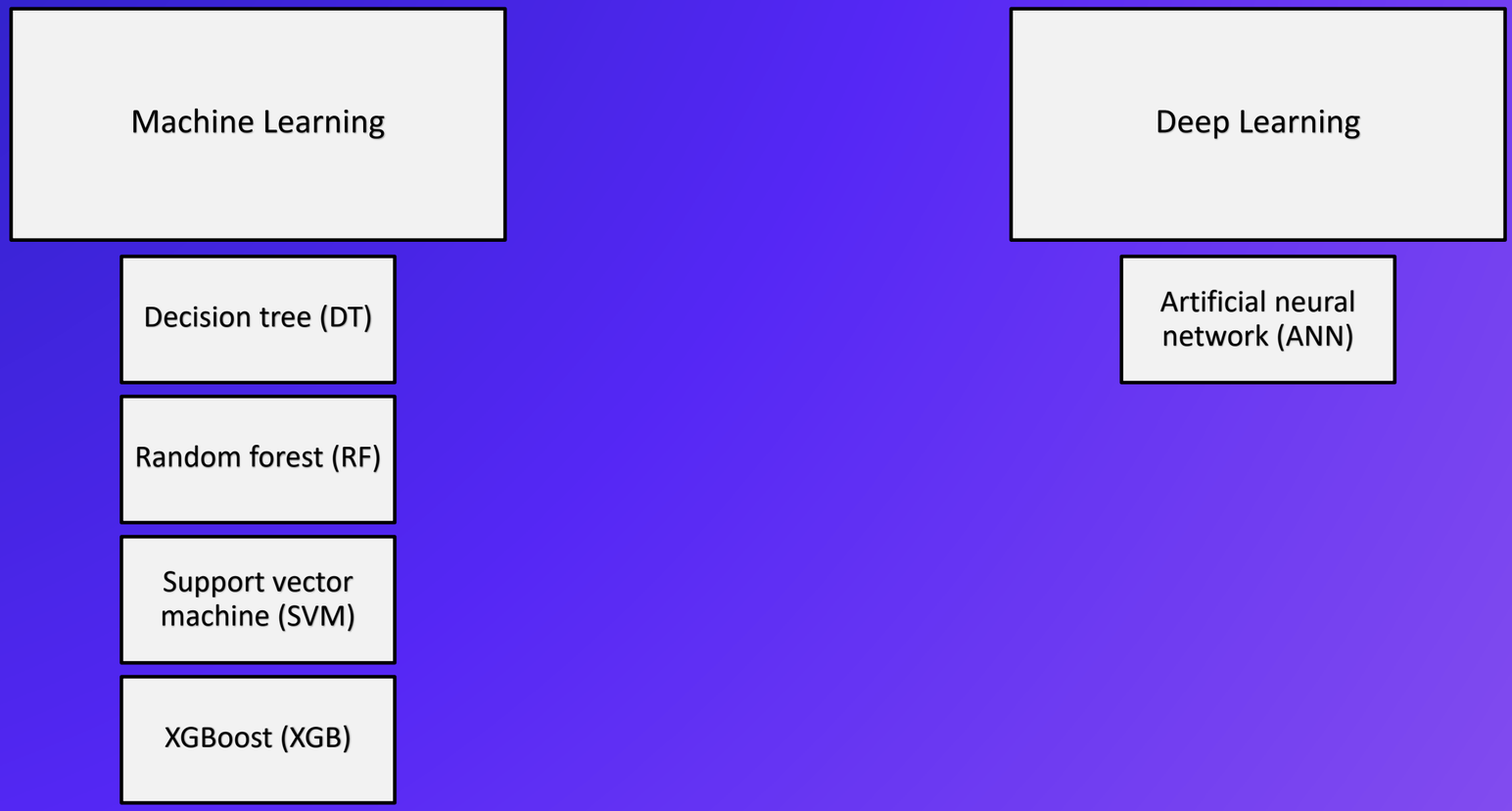

MACHINE LEARNING
Random Forest algorithm
By Kittithat Kongtaworn (P'Junior)

ML Workflow



Ref: https://medium.com/nerd-for-tech/the-ideal-workflow-for-your-machine-learning-project-9df1a7125b17

Algorithms for summer camp



Algorithms for summer camp

Decision tree algorithm

Advantages

1. Interpretability
2. Versatility
3. Feature Selection
4. Non-parametric
5. Robust to Outliers
6. Computational Efficiency

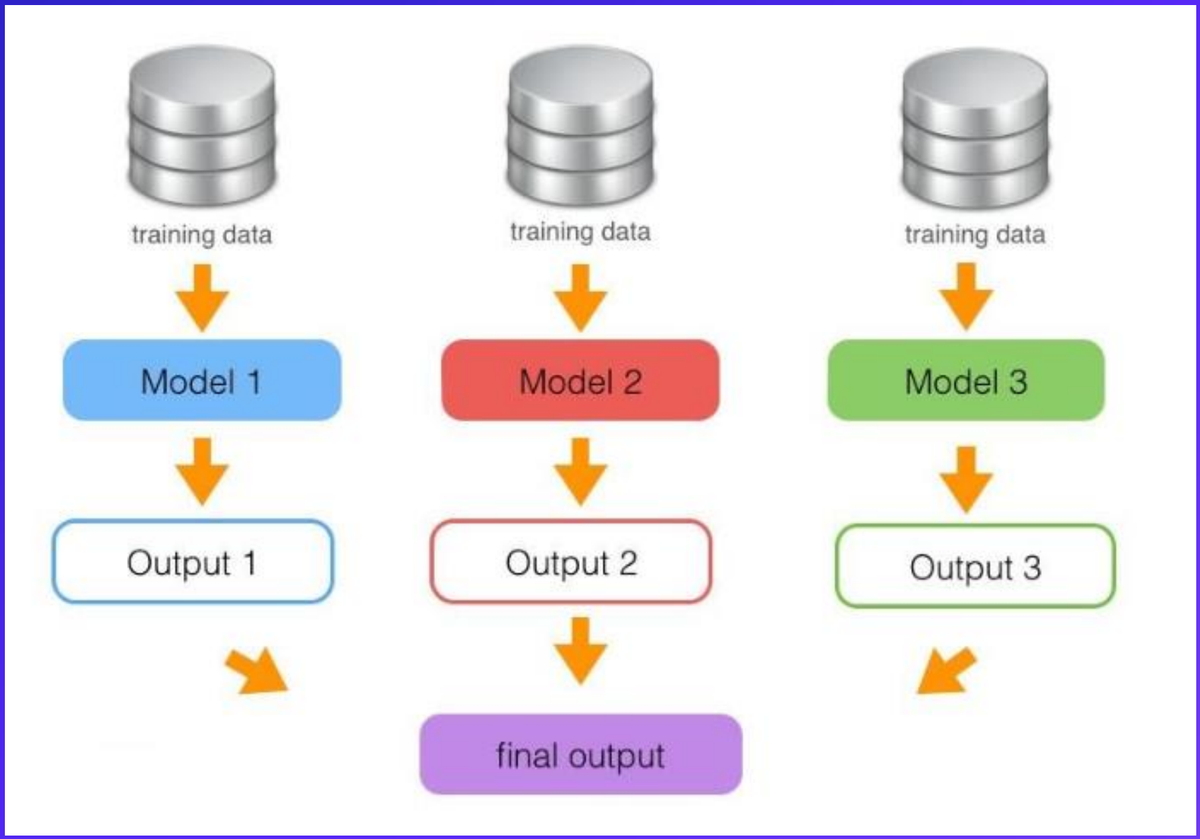
Disadvantages

1. High Variance
2. Overfitting
3. Instability
4. Limited Expressiveness
5. Greedy Nature

Ref: https://medium.com/@theclickreader/decision-tree-regression-explained-with-implementation-in-python-1e6e48aa7a47

Random forest algorithm

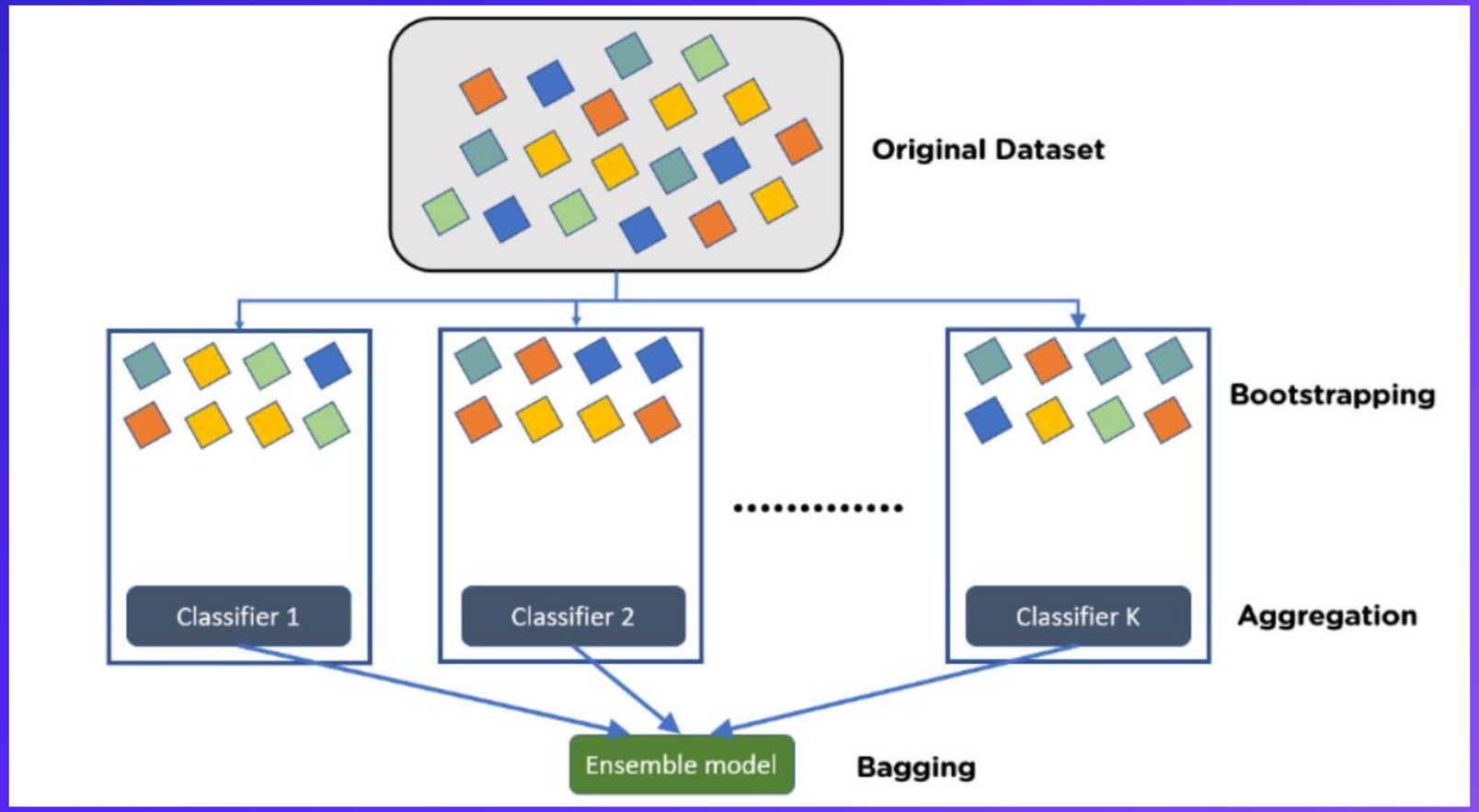
What is Ensemble model?



- Bootstrap Aggregating (Bagging)
  Ex. Random Forest
- Boosting
  Ex. XGBoost

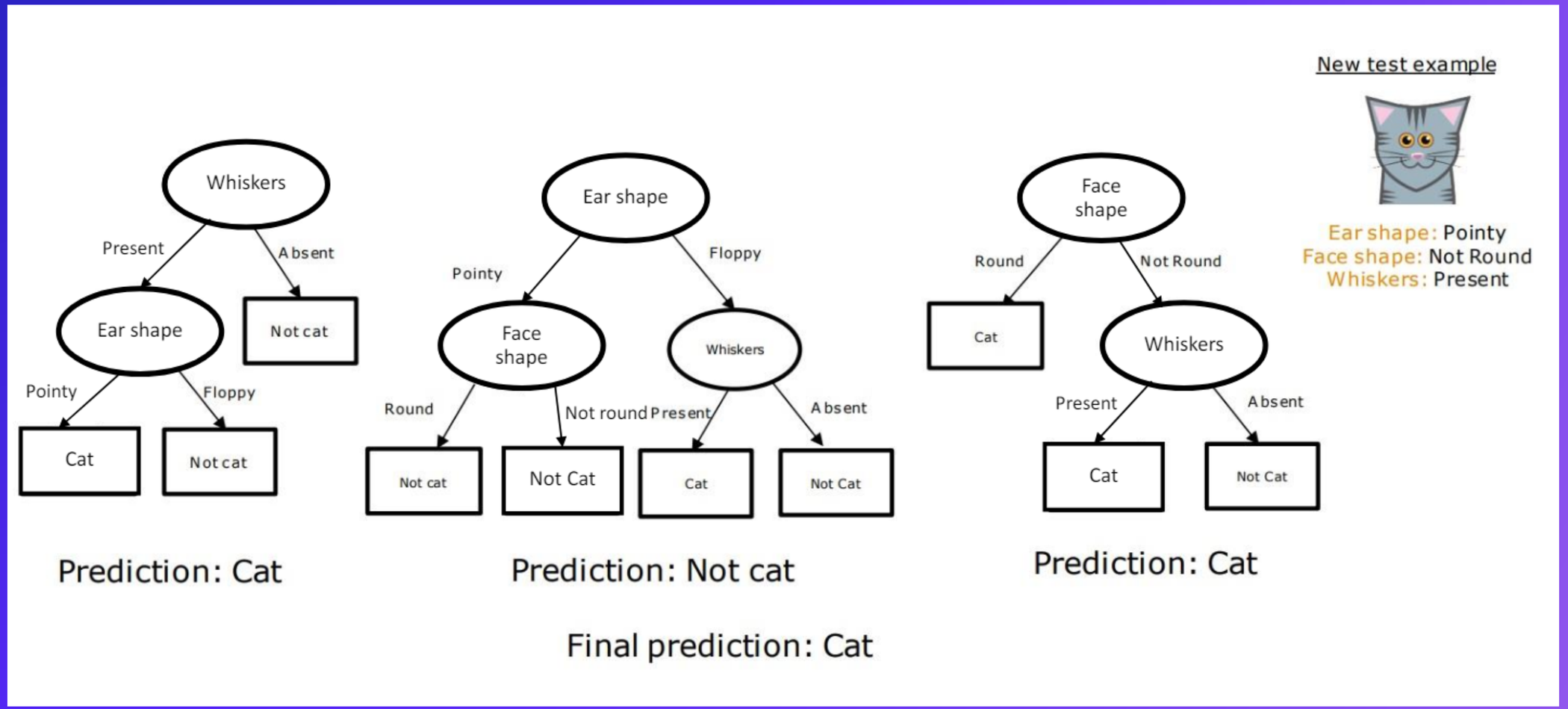
Random forest algorithm

Bootstrap Aggregating (Bagging)

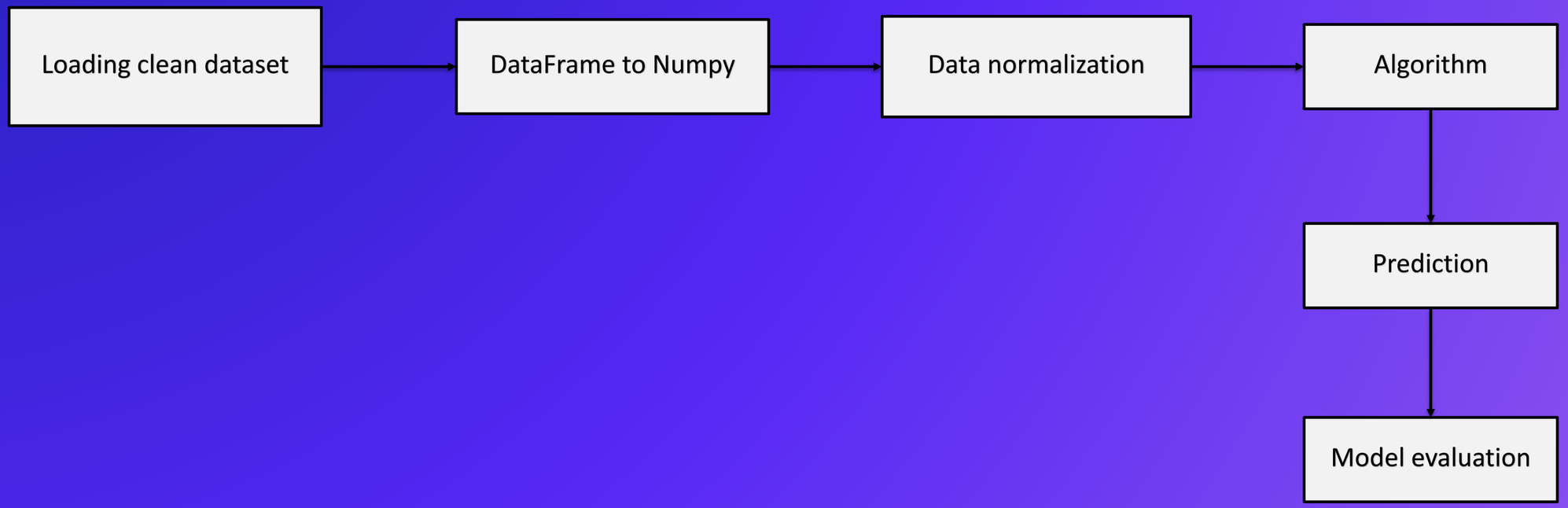


Random forest algorithm

What is a Random forest?



Step in python



Model evaluation for regression

Mean Absolute Error (MAE)	$MAE = \frac{1}{n} * \sum \lvert prediction - actual \rvert$
Mean Squared Error (MSE)	$MSE = \frac{1}{n} * \sum (prediction - actual)^2$
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{n} * \sum (prediction - actual)^2}$
Mean Absolute Percentage Error (MAPE)	$\left(\frac{1}{n}\sum \frac{\lvert Actual - Forecast \rvert}{\lvert Actual \rvert}\right) * 100$
R^2	$R^2 = 1 - \left(\frac{\Sigma(y - \hat{y})^2}{\Sigma(y - \bar{y})^2}\right)$

Lab for decision tree

THANK YOU!