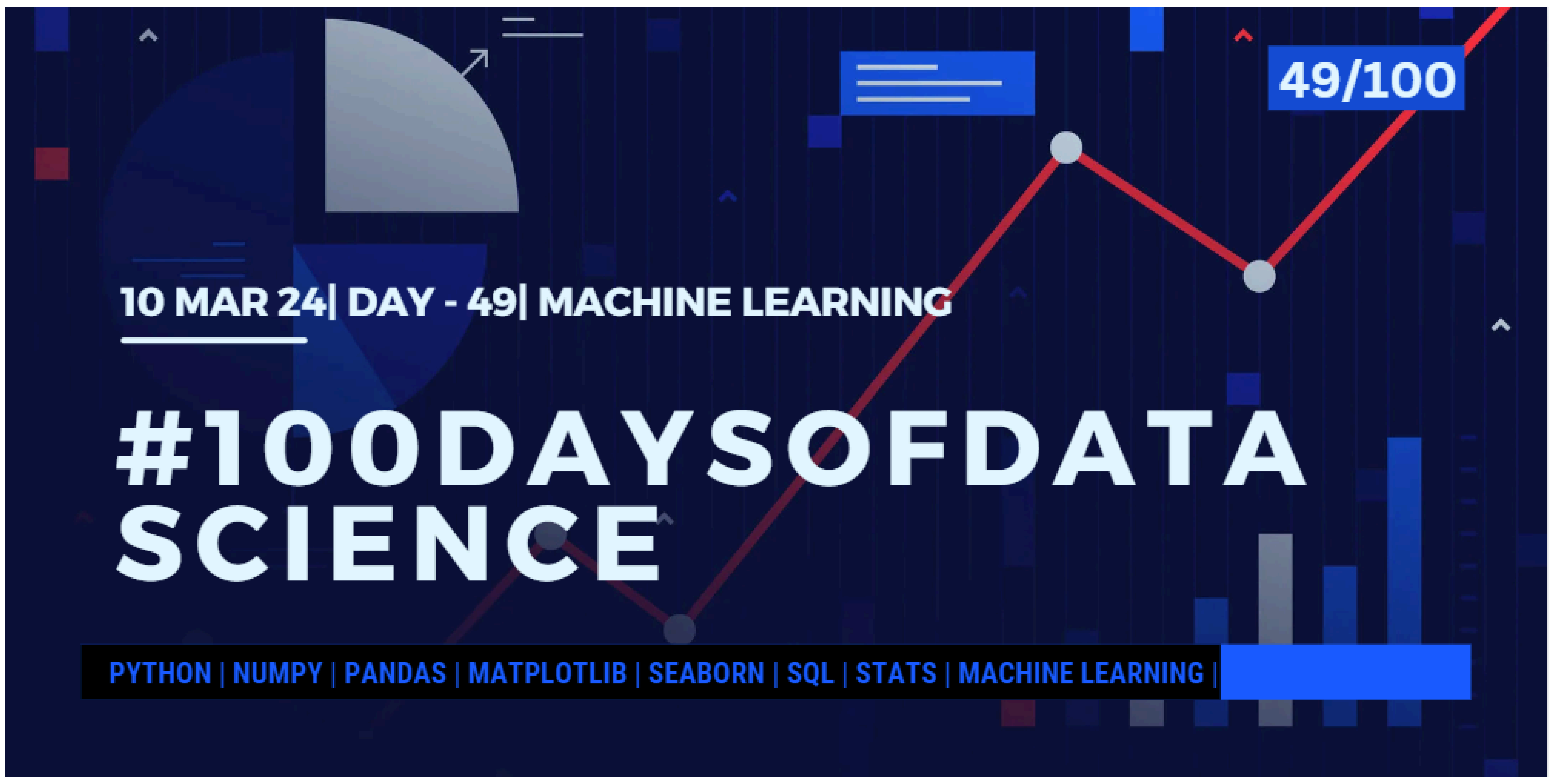
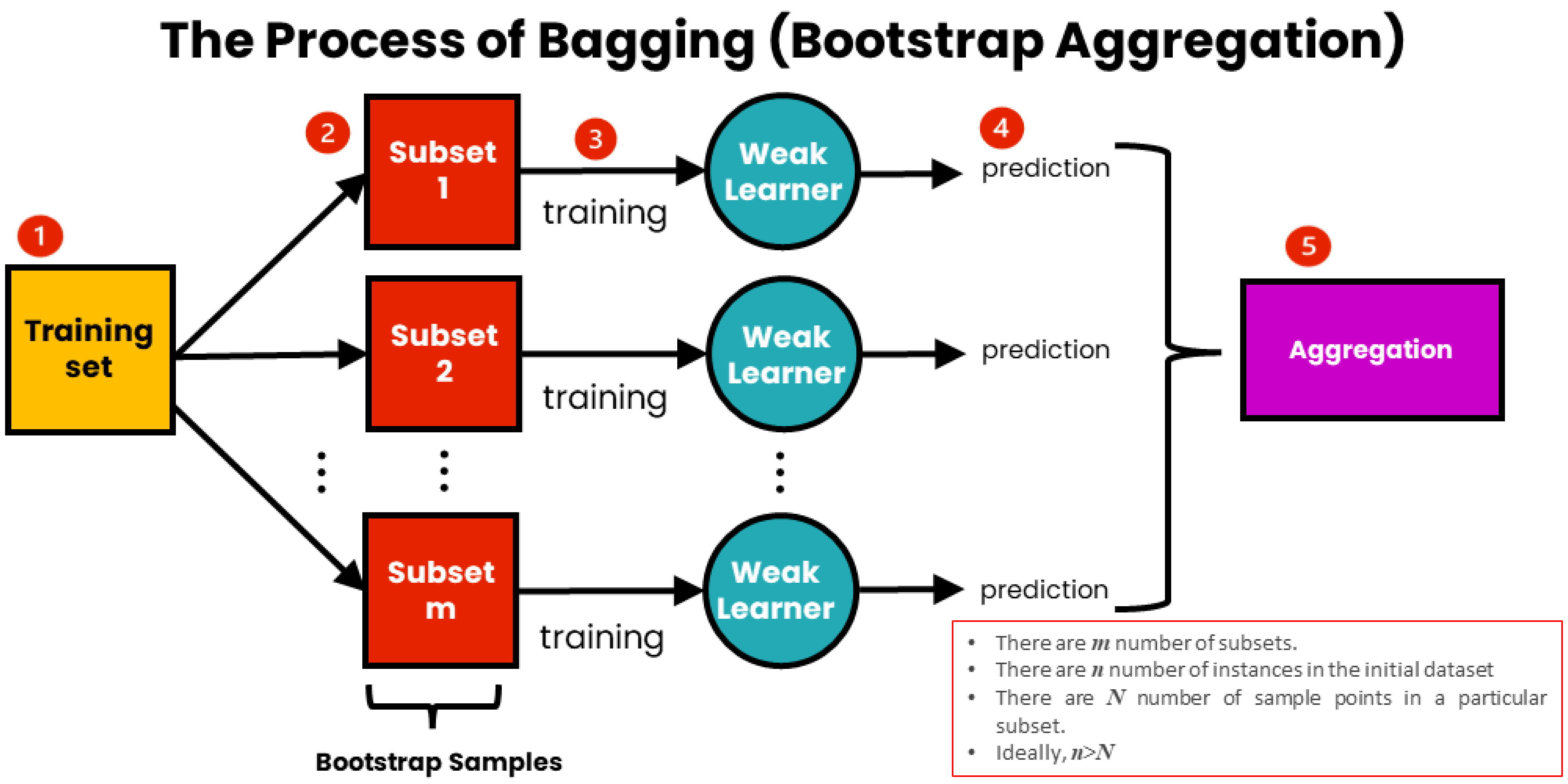



## Ensemble Modeling

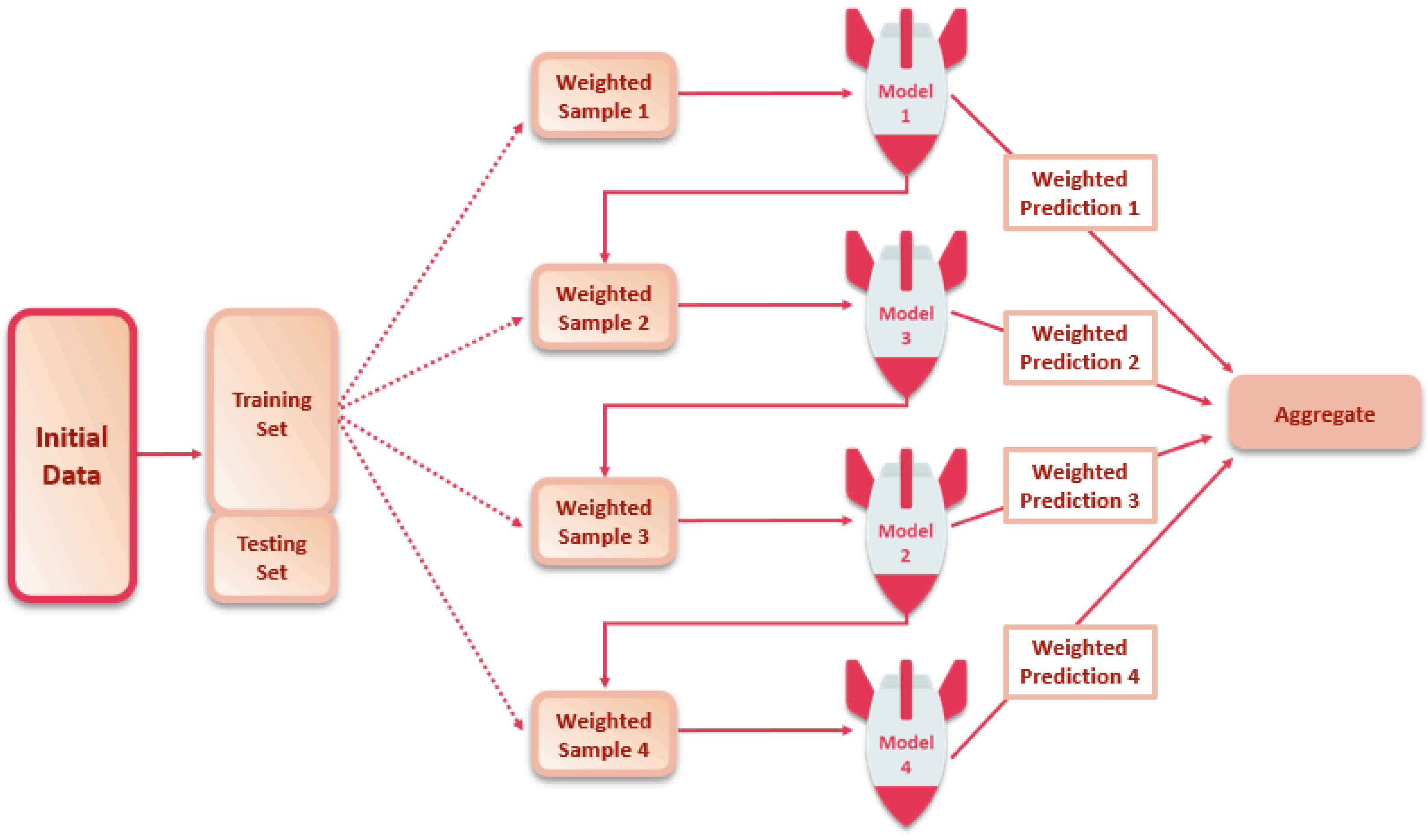
b&b.png

- **Pruning -**
  - It is opposite of spliting that means in this we reverse thed data and go to the previous node to avoid overfitting such that the model gives accurate results
- Tuning Techniques for decision tree
  1. Using Hyper parameter in the same model/pruning the model(It won't perform much change)
  2. Building multiple decision tree(**Ensemble Modeling) - In ensemble modelling we work on the fundamental of 1 decision tree is a weak learner and multiple decision tre is a strong learner
- **Ensemble Modeling:**
  - Ensemble modeling is a process where multiple diverse models are created to predict an outcome, either by using many different modeling algorithms or using different training data sets.
  - The ensemble model then aggregates the prediction of each base model and results in once final prediction for the unseen data.
  - Ensemble learning is a machine learning technique that enhances accuracy and resilience in forecasting by merging predictions from multiple models.
  - It aims to mitigate errors or biases that may exist in individual models by leveraging the collective intelligence of the ensemble.
- Types of ensemble modeling:
  - **Bagging(Parallel)**
  - **Boosting(Sequential)**

- **Bagging(Parallel)**
  - Bagging algorithm are parallel procesing technique which allow us to build and ensemble of DT and average out the results
  - We create bags of sample from the training data while randomly sampling the observation and make sure that 60% of trainign data goes in each bag that means n decision tree are built upon n bags of samples after which we pass the test data which goes as an input and gives us the predictions if it is a classifier we take the mode value and if it is a regressor we take the mean value
  - Types of bagging algorithm the proces of sampling differs in both
    - **Extra Tree Classifier**
    - **Random Tree/Forest Classifier(Bootstrap aggregation)**
      - **Advantages:**
    - **Reduced variance:** By averaging predictions from multiple models trained on different subsets of data, bagging reduces the impact of noise and outliers in the training data, leading to a more stable and robust model.
    - **Parallelizable:** Bagging models can be trained independently on different data subsets, making it suitable for parallel processing and faster training on large datasets.
    - **No need for strong learners:** Bagging works well even with weak learners as the averaging process reduces their individual errors.
      - **Disadvantages:**
      - **May not improve bias:** If the base models all suffer from the same bias, bagging will not eliminate it.
    - **Increased variance in some cases:** Bagging can introduce additional variance if the base models are highly unstable.
    - **Computationally expensive:** Training multiple models can be computationally expensive compared to training a single model.



- **Boosting(Sequential)**
  - Boosting algorithm help us to convert a weak learner into a strong learner
  - It is a sequential process where a modle builts an additive DT through multiple iterations
  - In each iteration we populate the bag of training data and builts the model on that bag it test the model internally on the remaining training data to identify and highlight misclassifier data point or error in that model
  - In the next iteration the bag will populate in such a manner the misclassified data poits of the previous iteration swill definitely go into training along with some random oberservation this process this process keeps on repeating to built an additive model until either of the stopping criteria is achieve
  - Stopping criteria
    1. **Maximum number of the estimator**
    2. **Zero conversion**
  - The model form the last iteration is consider as a final model which is use to predict test data points (False predition is misclassifier)
  - Types of boosting algorithm the proces of sampling differs in both
    - **AdaBoost Clasifier**
    - **XG Boost Classifier(Extreme Gradient Boost) [Not in sklearn]**
    - **Gradient Boost Classifier**
  - **Advantages:**
    - **Lower bias:** Boosting algorithms sequentially train models, focusing on correcting the errors of previous models. This can lead to a more accurate final model with lower bias.
    - **Can handle complex problems:** Boosting can effectively handle complex problems where simple models struggle to learn the underlying relationships.
  - **Disadvantages:**
    - **Overfitting risk:** Boosting algorithms are more prone to overfitting if not tuned carefully.
    - **Less interpretable:** Due to the sequential nature of the learning process, boosting models can be less interpretable compared to bagging.
    - **Computationally expensive:** Boosting requires training models sequentially, which can be computationally expensive compared to bagging.




***Made with ❤ by Zahid Salim Shaikh***

In [ ]: