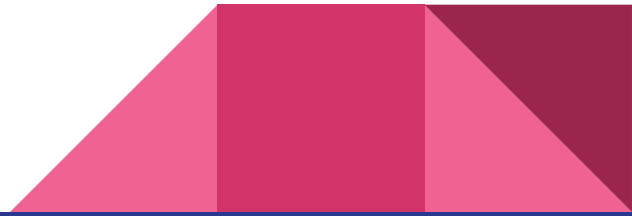
Boosting Algorithm

Regression

What is Boosting algorithm

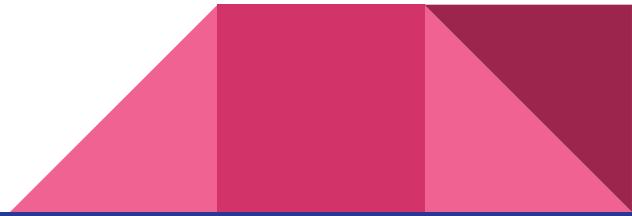
Combining multiple weak learning algorithm to improve prediction accuracy is called as boosting algorithm.

Boosting is an ‘ensemble learning’ technique that combines multiple weak models to build a strong model.



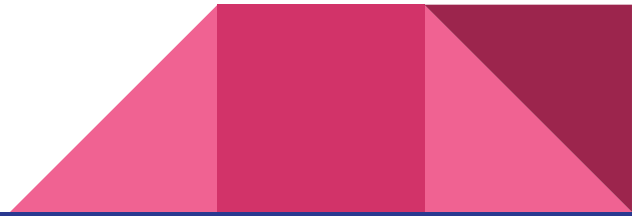
Working of Boosting algorithm

- First the initial dataset is used to create a model and is evaluated, Next model is built on the first model aiming to reduce errors in the first model. Similarly models are added to focus on more accuracy.
- This process will be continued until either the complete dataset is predicted correctly or till the iteration reaches the predefined number.



Adaboost(adaptive boosting) algorithm

- This is an ensemble learning technique that improves classification accuracy by combining multiple decision trees.
- This algorithm assigns higher weights to the incorrectly predicted instances in each iteration ensuring the subsequent models focus on these instances.
- sklearn.ensemble.AdaBoostRegressor



Gradient boost

- Gradually reduces the residual error for each subsequent models.
- Optimizes a loss function to reduce the error.
- sklearn.ensemble.GradientBoostingRegressor

XG(Extreme Gradient) boosting

- Optimized Gradient boosting algorithm.
- Uses regularization technique to prevent overfitting.
- Handles both categorical and numerical data.
- xgboost

LG(Light Gradient) boosting algorithm

- To prevent overfitting by reducing tree depth in Decision tree.
- Histogram of distribution is used to bucket the data to bins
- In LG boosting, decision tree will be grown only one leaf at a time from the whole tree.
- lightgbm