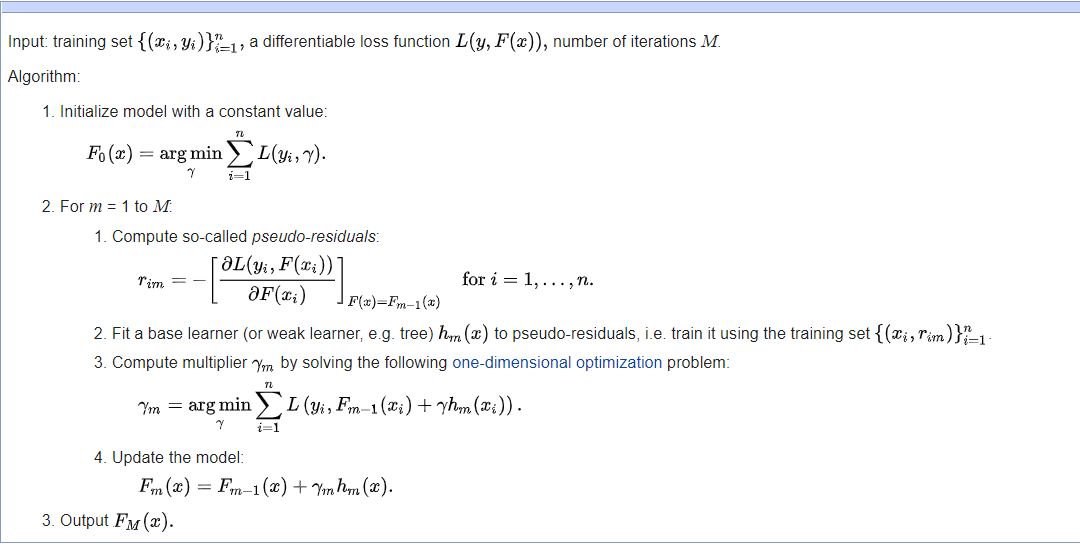
**Gradient Boosting**

**Algorithm Functioning:**

* Initialize a model with constant value 🡺 X as features and Y as target, predicting Y^
* Calculate the residual 🡺 Y - Y^ = R1
* Build a model to predict the residual 🡺 X as features and R1 as target, predicting R1^
* Calculate residual again 🡺 R1 – R1^ = R2
* Update the model 🡺 Y^ = Y^ + theta (R2) [theta is the learning rate]
* Build a model to predict R2 🡺 X as features and R2 as target, predicting R2^
* Calculate the residual again 🡺 R2 – R2^ = R3
* Update the model 🡺 Y^ = Y^ + theta (R3)
* Repeat for number of decision trees.



**Bias(Underfitting) – Variance(Overfitting) Error:**

* Gradient Boosting is a sequential model where the Basic model will have high bias and eventually while adding base learners to it, bias will decrease.
* Adding more number of trees might lead to high variance.

**Hyperparameters:**

* Number of trees, Depth of a tree and learning rate.

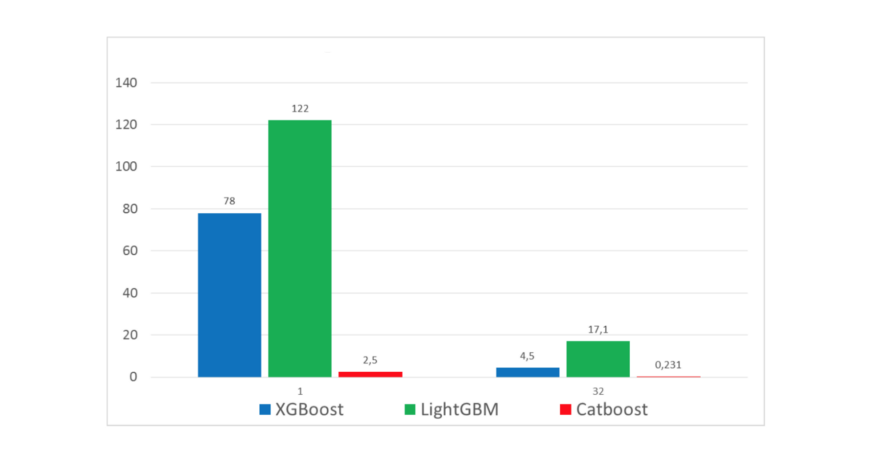
**Pros and Cons:**

* Since its sequential model, consumes more time to train
* Hard to fit – Might face overfitting problem
* When the hyperparameters are tuned, performs better than Random forest (Bagging).
* Works well even with small amount data

**Cat-Boost**

**Advantages Over other Boosting Algorithms:**

* Gives best result with default parameter itself.
* Handles Categorical features better than other boosting algorithms
* Training time – CPU it doesn’t differ much but in GPU, it is faster than XGBoost and LightGBM with increase in size of the data.
* Predicting time – It is faster than other boosting algorithms.



**Working of the Algorithms - Different to other Boosting techniques:**

* Have special techniques to handle categorical features.
* Building tree in symmetric manner.
* Uses Bayesian estimators to avoid overfitting

**Special Features:**

* Feature Importance
* Feature Interaction – How two features interact to contribute in prediction
* Missing value support
* Overfitting detector – It stops training when CV error starts increasing

**Hyperparameters:**

* Learning\_rate
* Depth
* L2\_regularization
* Bagging\_temperature

**Limitations:**

* When the dataset has many numerical features, CatBoost takes more time to train than Light GBM.