LightGBM Parameters:

לפי האתר:

LightGBM – Leaf Wise tree growth algorithm

לכן,

Num\_leaves – needs to be less than 2^(max\_depth)

Min\_data\_in\_leaf – bigger number leads to a more conservative model

Max depth – limit the max depth of a single tree.

Min\_gain\_to\_split: the minimum amount of loss in training that results from adding a split point. (default is 0)

Min\_sum\_hessian\_in\_leaf: hessian is the second derivative of the objective function evaluated for each observation in a leaf. [this could help](https://stats.stackexchange.com/questions/317073/explanation-of-min-child-weight-in-xgboost-algorithm)

Grow less trees if you decrease the **num\_iterations** (number of trees)

If you change the number of trees change the **learning rate** as well!

Early stopping is important – helps to prevent overfitting ! (parameter name: early\_stopping\_round)

For better accuracy:

Use large max\_bin,

small learning rate with large num\_iterations

use large num\_leaves

use bigger training date

try dart

Deal with overfiting

Use small max\_bin

Use small num\_leaves

Use min\_data\_in\_leaf and min\_sum\_in\_leaf

Use bagging by set – bagging\_fraction and bagging\_freg

Use feature sub-sampling by set feature-fraction (I think it’s the same as colsample by tree)

Lambda\_l1, lambda\_l2, min\_gain\_to\_split, max\_depth, extra\_trees , path\_smooth

I think that most of the hyperparameters are unnecessary …