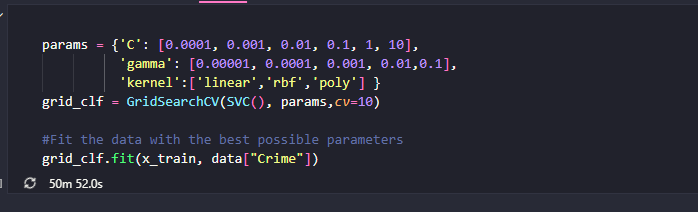


Hyperparameter tuning

<https://www.geeksforgeeks.org/hyperparameter-tuning/>

<https://www.vebuso.com/2020/03/svm-hyperparameter-tuning-using-gridsearchcv/>

<https://www.youtube.com/watch?v=5nYqK-HaoKY>



Time consuming if continue to train. It takes 90 minutes in average

Bayes

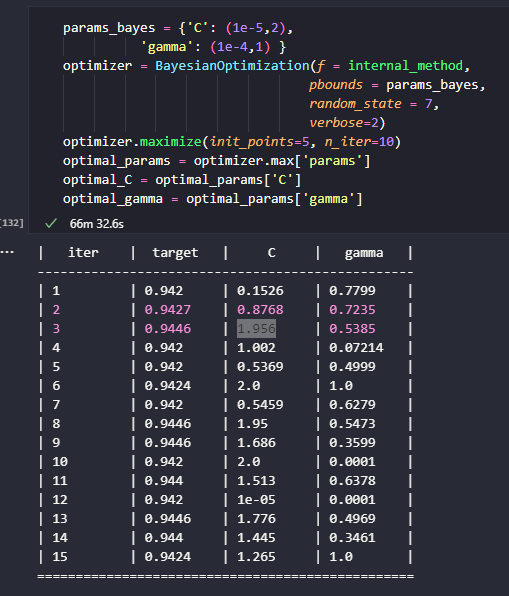
<https://github.com/bayesian-optimization/BayesianOptimization>

<https://www.run.ai/guides/hyperparameter-tuning/bayesian-hyperparameter-optimization>

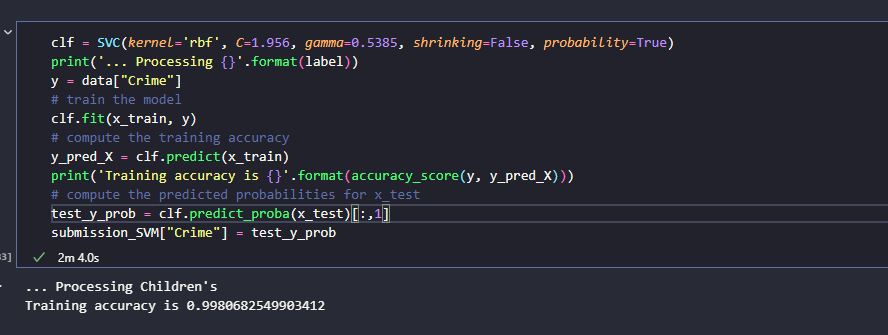
<https://github.com/bayesian-optimization/BayesianOptimization/blob/master/examples/constraints.ipynb>

Bayesian Optimization: bayes\_opt or hyperopt

<https://www.analyticsvidhya.com/blog/2021/05/bayesian-optimization-bayes_opt-or-hyperopt/>

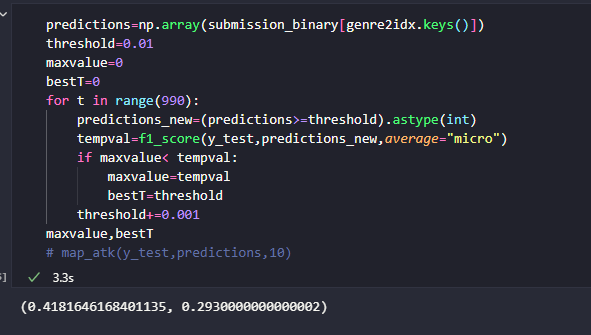
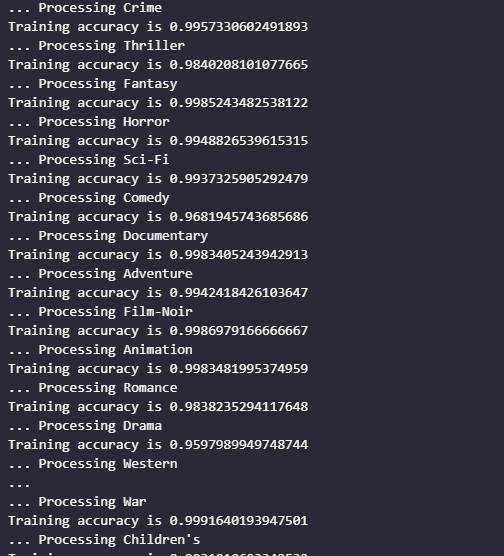


Use optimal value



Use SMOTE technique for imbalanced data

Result: oversampling



Not using SMOTE though still results in good accuracy in training but not good in testing -> overfitting

->Remedy

Use BorderlineSMOTE

-> acc increase



Bagging

2 first models combined -> 

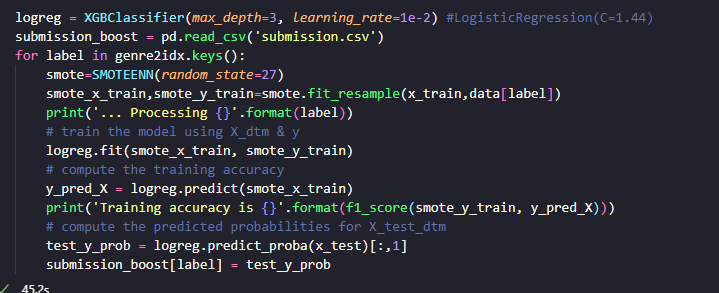


3 first models combined ->



Using Gradient Boosting

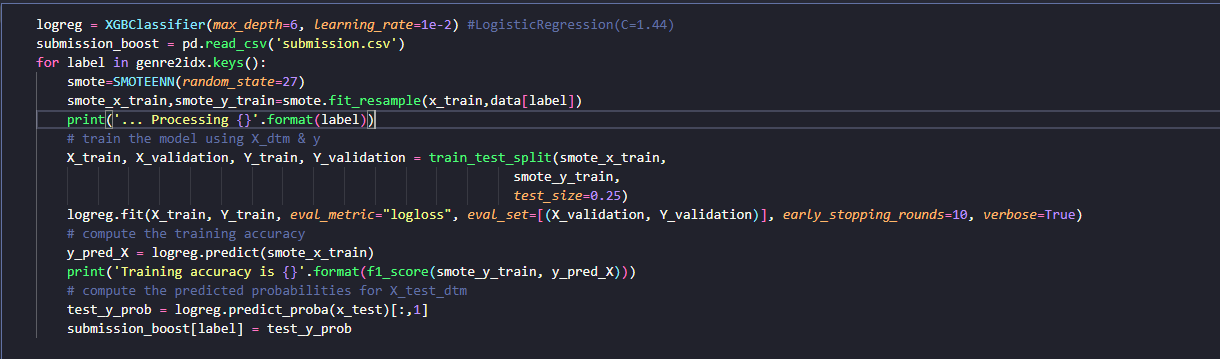
<https://www.youtube.com/watch?v=StWY5QWMXCw>

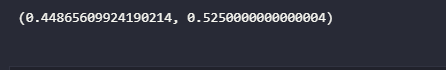


→ 

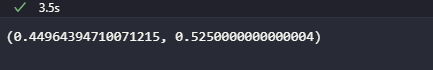


→





Change drop to 0.8



Để C và gamma = 





→

Không phải movie vào cx có rating