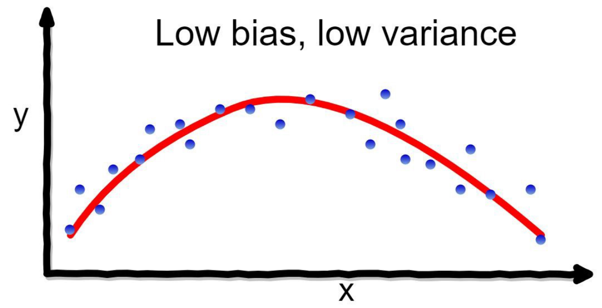
Ensemble Learning Relation With Bias and variance

Ashish Patel Follow
May 16, 2019 · 5 min read

Ensemble learning series…!!!



Hey Folks, After so long i am coming back again here with most important and demanding topic of machine learning which will give you how to design proper model for production.So, I thought to write whole series of Ensemble learning. after reading this series you know

- **What is Ensemble learning and relation with Bias and variance.**
- **Ensemble learning types with practical example.**
- **We try to cover deep learning with ensemble learning integration in this series. So Stay tuned.**

In this article, We discuss the bias and variance relation with ensemble learning. So let's Begin….!!!

Ensemble learning Series — (Happy Learning…!!!)

1. **Ensemble Learning Relation With Bias and variance**
2. **Ensemble Learning- The heart of Machine learning**
3. **Bagging — Ensemble meta Algorithm for *Reducing variance***
4. **Boosting- Ensemble meta Algorithm for *Reducing bias***
5. **Stacking -Ensemble meta Algorithms for *improve predictions***
6. **Ensemble learning impact on Deep learning**

. . .

1. Why we do Ensemble Learning?

As we know that we have many reason to do Ensemble learning(Meta Machine learning).

- In the group of Algorithm, Several model are simple and take less computation power and some are complex and taking more computation.
- When we put the model for production, that time accuracy and computation time is most important concern.
- Imagine that you have train a model with high accuracy But for real time application has no use. and in another scenario We have simple model(Simple Algorithm) may have less accuracy and not fit on data properly. In this cases, we compromise with accuracy and computation time.
- To Solve this problem, We can train a model on different algorithms(Weak Learners) to get a average result(Confidence index) from them. which will help you to implement real time application with very nice accuracy.

. . .

2. Reasons to use ensemble

1. **The dataset is too large or small —** If dataset is too large or small we have to use sampling to choose sample to take average of the result.
2. **Complex(Non-linear) data —** Real time dataset is mostly in non-linear fashion. so when we train a single model which cannot define the class boundary clearly and model become under-fit. That case we have to take different sub sample and take average of different model.
3. **High Confidence —** when we train a model with multiple classes and get high correlated output these situation lead the High Confidence. So, In this case most of the model predict the same class which lead that high confidence.

. . .

3.How to Create Ensemble System

- All models should have a difference of population. we have to split dataset in subset in such a way, every subset has less correlation between each other. So, The result different classification model will create which can give u independent result.
- Every model should have to give different hypothesis results. Our expected result should be very in each model. this will help to generalized ensemble system.
- Based on the model category we can visualize the data to get different perspective view, It should be either (linear or non-linear) or (supervised or Unsupervised).

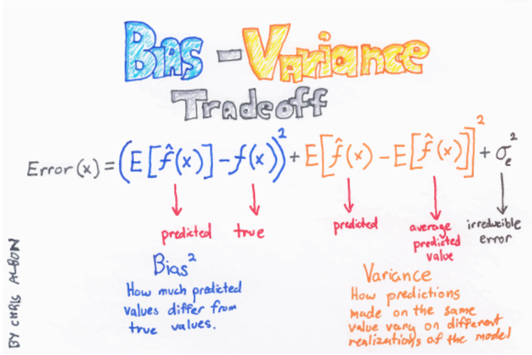
. . .

4.Quantification of Performance

- Calculate the performance of model by calculating the difference between Input and Output.

$$Err(x) = (E[\bar{f}(x)] - f(x))^2 + E[\bar{f}(x) - E[\bar{f}(x)]]^2 + \sigma_\epsilon^2$$

$$Err(x) = Bias^2 + Variance + Irreducible Error$$


Bias - Variance
Tradeoff
Error(x) = (E[f̂(x)] − f(x))² + E[f̂(x) − E[f̂(x)]]² + σₑ²
predicted
true
predicted
average predicted value
irreducible error
Bias²
How much predicted values differ from true values.
Variance
How predictions made on the same value vary on different realizations of the model
BY CHRIS ALBON

Courtesy — Chris Albon

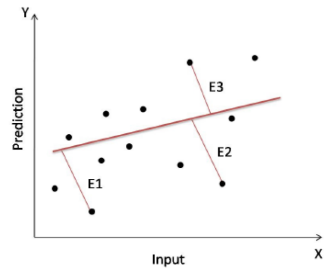
- Error Show’s that basic three component : **BIAS, VARIANCE, IRREDUCIBLE ERROR**
- We can’t handle Irreducible error but we can handle Bias and variance.

* * *

5. Bias Variance trade-off

Bias Error : (High Bias Under-fitting)

- **Bias is average of difference between predicted and Actual result. High Bias means We are getting low performance.**

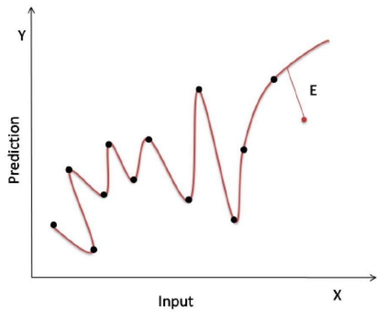


Bias Error

- When we train a linear on complex data we can see the error with with black dot and red line distance of error E1,E2,E3. Simple meaning of this our model is not fitting properly thus under-performing, for this case we have to use Complex(polynomial) Model to avoid this scenario.

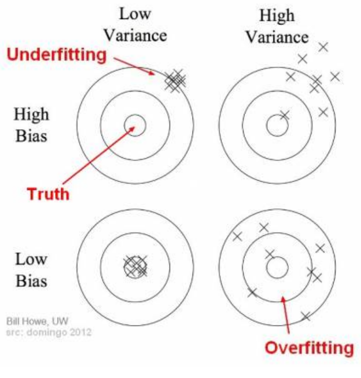
Variance Error : (High Variance over-fitting)

- **Quantify the difference of predicted value in the same as observation on that time model is over-fitting.**
- Train a model show high variance near **100% Accuracy** on training data. when we check the model on present data this fail to predict correct result.
- In two condition this happen: **1) Less training data 2) Complex model on simple data.**



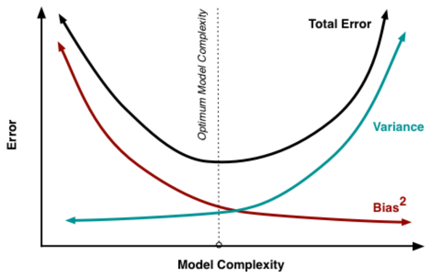
Variance Error

- In above figure, We can see that our model perform well on black training example. But, when we predict the value of Red dot we see the error difference it’s variance error.
- The problem of over-fitting can be solved by increasing the number of training instances or by choosing the correct classifier for prediction.



Courtesy — “Understanding The Bias-Variance Trade-off.” Towards Data Science.

- In above Figure, Imagine that Center of the **Circle(O)** contain the **Actual value** and **Cross (X)** represented **predicted value.**
- **When We have High bias increase model complexity so we get low bias. We expect the result low bias and low variance.**



- **Managing Bias and variance in balance way here Ensemble learning is come in the picture.**

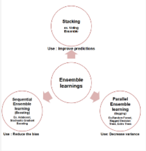
I hope you are enjoying this article. Thanks for reading…!!!

References :

Ensemble Learning- The heart of Machine learning

Ensemble learning Series…!!!

medium.com


Stacking
ex. Voting Ensemble
Use : Improve predictions
Ensemble learnings
Sequential Ensemble learning (Boosting)
Parallel Ensemble learning (Bagging)
Use : Reduce the bias
Use: Decrease variance

Single estimator versus bagging: bias-variance decomposition - scikit-learn 0.21.1 documentation

In regression, the expected mean squared error of an estimator can be decomposed in terms of bias, variance and noise…

scikit-learn.org

Machine Learning | Ml Research Lab | Ensemble Learning | Bias Variance Tradeoff | Deep Learning

425

WRITTEN BY

Ashish Patel

Follow

Data Scientist | Kaggle Kernel Master | Deep learning Researcher

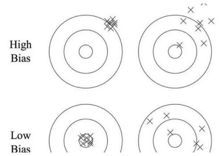
ML Research Lab

Follow

The vision of the ML Research Lab is to provide best technical tutorial to ML aspirant and Researcher to gain the Knowledge of Machine Learning, Deep Learning, Natural Language Processing, Statistics and Computer Vision.

More From Medium

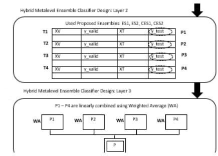
Also tagged Bias Variance Tradeoff



Bias, Variance, and Tradeoff

Suraj
May 9 · 6 min… 4

Also tagged Ensemble Learning


Hybrid Metalevel Ensemble Classifier Design: Layer 2
Hybrid Metalevel Ensemble Classifier Design: Layer 3

Meta Ensemble Self-Learning Model with Optimization

… MLearning.ai
May 6 · 4 min… 5

Also tagged Machine Learning

Top 5 Resources For Aspiring Data Scientists

Bex… For Tech
May 14 · 8 mi… 31

Learn more.

Medium is an open platform where 170 million readers come to find insightful and dynamic thinking. Here, expert and undiscovered voices alike dive into the heart of any topic and bring new ideas to the surface. Learn more

Make Medium yours.

Follow the writers, publications, and topics that matter to you, and you'll see them on your homepage and in your inbox. Explore

Share your thinking.

If you have a story to tell, knowledge to share, or a perspective to offer — welcome home. It's easy and free to post your thinking on any topic. Write on Medium