Institute for Advanced Studies
in Basic Sciences
GavaZang, Zanjan, Iran

Concepts of bias and variance in machine learning

Prof. Majid Ramezani

Report for Project <4>

Faeze Ahmadi

October 2025

My Simple Understanding of Bias and Variance in Machine Learning

First of all, what is machine learning?

Machine learning means that instead of being programmed step by step, the computer learns patterns automatically from data.

For example, if you show a model a thousand pictures of cats and dogs and tell it which is which, it learns to recognize whether a new picture is a cat or a dog.

In simple terms:

Machine learning is about teaching computers to “learn from experience (data)” rather than being given explicit instructions.

Bias and Variance

When we build a machine learning model, it tries to find a pattern in the data.

Sometimes it learns too simply (it doesn’t capture enough of the real pattern), and sometimes it learns too specifically (it memorizes the training data instead of generalizing).

This is where two important concepts come in: Bias and Variance.

Bias

Bias means how much the model makes simplifying and possibly wrong assumptions about the data.

For example:

imagine we build a model to predict a student's grade based on the hours they study.

If the model is too simple (for instance, it just fits a straight line), it may not capture the real relationship.

In this case, we say the model has high bias — it thinks too simply and cannot learn the true patterns in the data.

This problem is called Underfitting.

Variance

Variance means how sensitive the model is to the training data.

For example:

if the model is too complex, it tries to perfectly fit every single training point.

As a result, when it sees slightly different new data, it becomes confused and performs poorly.

In this case, we say the model has high variance — it depends too much on the training data.

This problem is called Overfitting.

A Simple Analogy

We can think of this in our own studying habits.

If we study too simply and rigidly, skipping over important ideas, our bias is high — we will struggle when faced with new or tricky questions.

On the other hand, if we over-focus on every small detail and spend too much time on minor issues, our variance is high — we'll run out of time and energy, and it will hurt our overall performance.

The goal, both in studying and in machine learning, is to find a balance between bias and variance.

References:

- Reconciling modern machine-learning practice and the bias-variance trade-off — Mikhail Belkin, Daniel Hsu, Siyuan Ma, Soumik Mandal (2019). https://arxiv.org/abs/1812.11118
- Rethinking Bias-Variance Trade-off for Generalization of Neural Networks — (2020) https://proceedings.mlr.press/v119/yang20j/yang20j.pdf