Machine Learning – Day 6 Notes

Understanding Underfitting and Overfitting

---

1. Initial Assumption

I believed that **mistakes in Machine Learning mainly come from weak or simple models**.

I learned that this is only part of the picture.

2. Key Realization

A model can fail in **two fundamentally different ways**:

1. By **not learning enough**
2. By **learning too much**

Both lead to poor performance, but for opposite reasons.

3. Two Types of Model Learning Failures

A. Underfitting

When the model learns too little

Underfitting happens when:

- The model is too simple
- Important patterns in the data are ignored
- The model fails to capture relationships that actually matter

Impact:

- Poor performance on both training and new data
- Predictions are overly general and inaccurate

Real-life example:

- Predicting house prices using only the number of rooms
- Predicting exam scores using only attendance

Core insight:
The model lacks the capacity or information to learn meaningful patterns.

B. Overfitting

When the model learns too much

Overfitting happens when:

- The model memorizes training data
- Noise is mistaken for real patterns
- Performance is excellent on training data but poor on new data

Impact:

- High training accuracy
- Low real-world reliability

Real-life example:

- Memorizing answers instead of understanding concepts
- A model that fits historical stock data perfectly but fails on future prices

Core insight:
The model learns patterns that do not generalize to real situations.

4. Why This Changes How Performance Issues Are Viewed

When a model performs poorly, the issue may not be:

- The algorithm choice
- The lack of model complexity

Instead, the problem may be:

- Too simple learning (underfitting)
- Too specific learning (overfitting)

Understanding *how* the model is learning becomes more important than *which* model is used.

5. Balance Is the Goal

Effective Machine Learning aims to:

- Capture meaningful patterns
- Ignore noise and randomness
- Generalize well to unseen data

This balance lies **between underfitting and overfitting**.

6. Shift in Perspective

Earlier Thinking

- Poor results mean the model is weak

Updated Thinking

- Poor results may indicate:
    - Insufficient learning
    - Excessive memorization

Diagnosis matters more than assumptions.

7. Final Takeaway

Machine Learning performance issues are often **learning issues**, not algorithm issues.

Data → Learning Behavior → Performance

Understanding whether a model is:

- Ignoring patterns
- Or memorizing noise

is key to improving results.

Good models don’t learn everything —
they learn what matters.