# Q1. Overfitting and Underfitting

## Overfitting

A model **learns the training data too well**, including noise and outliers, and fails to generalize to new data.

#### Consequences

- Very high training accuracy
- Poor test/validation performance

#### Mitigation

- More training data
- Regularization (L1, L2)
- Simpler model
- Cross-validation
- Early stopping
- Dropout (for neural networks)

---

## Underfitting

A model is **too simple** to capture the underlying pattern in the data.

#### Consequences

- Poor performance on both training and test data

#### Mitigation

- Use a more complex model

- Add more relevant features
- Reduce regularization
- Train for more epochs

---

## Q2. How can we reduce overfitting? (Brief)

- Increase training data
- Reduce model complexity
- Apply **regularization**
- Use **cross-validation**
- Use **early stopping**
- Use **dropout** (deep learning)
- Feature selection or pruning

---

## Q3. Explain underfitting and its scenarios

**Underfitting** occurs when a model cannot learn patterns in the data.

### Scenarios

- Using **linear regression** for non-linear data
- Very few training features
- Excessive regularization
- Insufficient training time

- Shallow neural networks

---

# Q4. Bias–Variance Tradeoff

### Bias

Error due to **oversimplified assumptions**.

### Variance

Error due to **sensitivity to training data**.

### Tradeoff

- High bias → underfitting
- High variance → overfitting

📌 The goal is to **balance bias and variance** to minimize total error.

---

# Q5. Detecting overfitting and underfitting

### Indicators

| Condition | Training Error | Validation Error |
|---|---|---|
| Overfitting | Low | High |
| Underfitting | High | High |
| Good Fit | Low | Low |

### Detection Methods

- Learning curves
- Cross-validation scores

- Performance gap between training and test data

---

# Q6. Bias vs Variance comparison

| Aspect | Bias | Variance |
|---|---|---|
| Definition | Error from wrong assumptions | Error from sensitivity to data |
| Model complexity | Too simple | Too complex |
| Problem | Underfitting | Overfitting |
| Example | Linear regression | Decision tree (unpruned) |

## Examples

- **High bias model:** Linear regression on complex data
- **High variance model:** Deep decision tree

---

# Q7. Regularization in Machine Learning

## Definition

Regularization adds a **penalty term** to the loss function to discourage complex models and prevent overfitting.

---

## Common Regularization Techniques

### 1. L1 Regularization (Lasso)

- Adds absolute value of coefficients
- Produces sparse models (feature selection)

Loss+λ∑∣w∣\text{Loss} + \lambda \sum |w|Loss+λ∑∣w∣

---

**2. L2 Regularization (Ridge)**

- Adds squared coefficients
- Shrinks weights smoothly

Loss+λ∑w2\text{Loss} + \lambda \sum w^2Loss+λ∑w2

---

**3. Elastic Net**

- Combination of L1 and L2

---

**4. Dropout**

- Randomly removes neurons during training (DL)

---

**5. Early Stopping**

- Stops training when validation loss increases

---