### 1. Feature

Assignment-1

Sudharsan.R

22CSR207

A feature is an input variable used in making predictions.

**Example**: The size of a house.

### 2. Label

A label is the output variable that the model is trying to predict.

**Example**: The price of the house.

### 3. Prediction

A prediction is the value the model estimates based on the input features.

**Example**: The estimated price of a house.

### 4. Outlier

An outlier is a data point that significantly differs from other data points.

**Example**: A house that costs $10 million in a neighborhood where most houses cost $200,000.

### 5. Test Data

Test data is used to evaluate the model's performance.

**Example**: Data to check if the house price prediction model works well.

### 6. Training Data

Training data is used to teach the model.

**Example**: Data showing house sizes and their prices.

### 7. Model

A model is a mathematical representation used to make predictions.

**Example**: An equation predicting house prices based on size.

### 8. Validation Data

Validation data is used to tune and validate the model during training.

**Example**: Data used to adjust the house price prediction model before final testing.

### 9. Hyperparameter

A hyperparameter is a setting used to control the training process.

**Example**: The learning rate in a neural network.

### 10. Epoch

An epoch is one complete pass through the training data.

**Example**: Training the house price model on all data once.

### 11. Loss Function

A loss function measures how well the model's predictions match the actual results.

**Example**: The difference between predicted and actual house prices.

### 12. Learning Rate

The learning rate controls how much the model's parameters are updated during training.

**Example**: How quickly the model learns house prices.

### 13. Overfitting

Overfitting is when a model performs well on training data but poorly on new data.

**Example**: A model that predicts training house prices perfectly but fails on new data.

### 14. Underfitting

Underfitting is when a model is too simple to capture the data patterns.

**Example**: A model that predicts all house prices as the average price.

### 15. Regularization

Regularization is a technique to prevent overfitting by simplifying the model.

**Example**: Adding a penalty for large weights in the house price model.

### 16. Cross-Validation

Cross-validation is a method to evaluate the model by splitting data into training and testing sets multiple times.

**Example**: Testing the house price model on different data splits.

### 17. Feature Engineering

Feature engineering is creating new features to improve the model.

**Example**: Creating "price per square foot" from house size and price.

### 18. Dimensionality Reduction

Dimensionality reduction reduces the number of features to simplify the model.

**Example**: Combining similar features in the house price model.

### 19. Bias

Bias is the error from overly simplistic models.

**Example**: Predicting house prices with a simple linear model.

### 20. Variance

Variance is the error from overly complex models that fit the training data too closely.

**Example**: A model that captures noise in house price data.

|  |  |  |  |
| --- | --- | --- | --- |
| Anime-name | No. of episodes | Rating | Budget |
| DEMON-SLAYER | 150 | 9.1 | 200 crores |
| ATTACK ON TITAN | 139 | 9.8 | 369.3 crores |
| BLACK-CLOVER | 350 | 9.0 | 150 crores |