**Assignment - 02**



**Group Members:**

**Name Roll No**

Zuraiz Ahmed 231370227

Abdul Rehman 231370197

Ibrahim 231370216

Ahmad Ali 231370208

**Submitted To:**

Mam Sumbal Fatima

**Submission Date:**

2 - Feb - 2025

**Group Members & Their Tasks**

1. **Zuraiz Ahmed (Roll No: 231370227)**

* Introduction to Machine Learning.
* Explanation of Supervised Learning with an example.

1. **Abdul Rehman (Roll No: 231370197)**

* Explanation of Overfitting and Underfitting.
* Practical Calculation of Error Metrics (Absolute Error, MAE, MSE, RMSE).

1. **Ahmad Ali (Roll No: 231370208)**

* Explanation of Unsupervised Learning with an example.
* Explanation of Reinforcement Learning with an example.

1. **Ibrahim (Roll No: 231370216)**

* Explanation of the need for training and testing data in ML.
* Final conclusion and summary of the assignment.

**Part 1: Basics**

1. **Machine Learning**

Machine Learning (ML) is a type of technology where computers learn from data and improve their performance without being explicitly programmed. Instead of following fixed rules, ML models find patterns in data and make predictions or decisions based on what they have learned.

**Traditional Programming and Machine Learning**

* We give the computer step-by-step rules (like if-else conditions) to follow.
* The computer learns patterns from data and makes its own rules.

**Example**

* To detect spam emails, we write rules like If the email has lottery or free money mark it as spam.
* The computer looks at many spams and non-spam emails and learns patterns to decide what is spam.

1. **Three Types**
2. **Supervised Learning**

In supervised learning, the model is trained on labeled data (data with correct answers). It learns by mapping input to the correct output.

**Example**

Predicting house prices based on features like area, number of bedrooms, and location.

1. **Unsupervised Learning**

In unsupervised learning, the model is trained on data without labels. It finds hidden patterns in data

**Example**

Grouping customers based on their shopping behavior.

1. **Reinforcement Learning**

In reinforcement learning, the model learns by interacting with an environment and receiving rewards or penalties.

**Example**

A robot learning to walk by trying different movements and adjusting based on what works best.

1. **Features and Labels**

* **Features:** The input variables (independent variables) used to make predictions.
* **Labels:** The output variable (dependent variable) that we want to predict.

**Example**

If we are predicting a student's exam score based on study hours

* **Features:** Number of study hours, previous scores, sleep hours.
* **Label:** Exam score.

1. **Training & Testing**

* We use training data to help the model learn patterns by giving it examples with correct answers. The model studies this data and tries to find relationships between inputs and outputs.
* After training, we use testing data (which the model has never seen before) to check how well it can make predictions on new information.
* This is important because if a model only memorizes the training data, it may not work well with new data. Testing helps ensure the model understands patterns instead of just remembering answers.

1. **Overfitting & Underfitting**
2. **Overfitting**

The model learns too much from training data and memorizes it instead of understanding patterns. It performs well on training data but poorly on new data.

1. **Underfitting**

The model is too simple and cannot capture important patterns, leading to poor performance on both training and test data.

**How to avoid**

* **Overfitting:** Use more data, remove unnecessary complexity, and use regularization techniques.
* **Underfitting:** Use a better model with enough complexity and ensure good-quality training data.

**Part 2: Practical Calculations**

1. **Absolute Error**

|  |  |  |  |
| --- | --- | --- | --- |
| Day | Actual Price ($) | Predicted Price ($) | Absolute Error |
| 1 | 150 | 145 | 5 |
| 2 | 152 | 149 | 3 |
| 3 | 148 | 151 | 3 |
| 4 | 155 | 157 | 2 |
| 5 | 149 | 146 | 3 |
| 6 | 160 | 158 | 2 |
| 7 | 162 | 159 | 3 |
| 8 | 155 | 153 | 2 |
| 9 | 158 | 154 | 4 |
| 10 | 150 | 151 | 1 |

1. **Mean Absolute Error**
2. **Mean Squared Error**
3. **Root Mean Squared Error**

**Part 3: Metrics of Regression**

1. **Absolute Error**

Absolute Error is the difference between the actual and predicted values. It shows how far off a prediction is.

In stock price prediction, it helps understand how much the predicted prices deviate from actual prices.

1. **Mean Absolute Error**

MAE is the average of all absolute errors. It tells us, on average, how much the predictions are off from the actual values.

If MAE = 2.8, it means, on average, our stock price predictions are $2.8 away from the actual price.

1. **Mean Squared Error**

MSE is the average of squared errors.

Squaring the errors gives more importance to larger errors. This helps in detecting big mistakes.

If there are large errors in prediction, they will significantly increase the MSE value.

1. **Root Mean Squared Error**

RMSE is the square root of MSE. It is used to bring the error back to the same unit as the original data.

MSE gives squared values, which can be difficult to interpret.

RMSE gives an error value in the same unit as the original values.

RMSE = 3.0 means, on average, our stock price predictions are $3 off from the actual price.

1. **Which Metric**

* **MAE:** Easy to understand but does not give importance to large errors.
* **MSE:** Good for penalizing large errors but difficult to interpret.
* **RMSE:** Best choice because it gives an error in the same unit as stock prices.

**Conclusion**RMSE is the most useful metric for stock price predictions because it is in the same unit as actual prices and penalizes large errors.