Performance Metrics in Machine Learning

Once a machine learning model is trained, how do we know if it's actually any good? How do we compare different models? This is where **Performance Metrics** come in.

Evaluating model performance is a key element in the entire machine learning process. There are a large number of performance metrics available, and the choice of which one(s) to use is critical. The appropriate metric depends heavily on:

1. **The type of Machine Learning task:** Regression, Classification, Clustering, etc., all have different goals and thus different ways of measuring success.
2. **The specific business problem:** The definition of "good performance" is often tied to the practical application and the associated costs of different types of errors.

In the following sections, we will discuss some of the most common performance metrics used in Machine Learning, starting with those applicable to Regression problems and then moving on to Classification problems.

* **Performance Measures for Regression Problems**
* **Performance Measures for Classification Problems**

Regression Problems: Predicting Continuous Values

Regression Predictive Modelling

**Regression** is a type of supervised learning task where the goal is to predict a **continuous output variable (y)** based on one or more input variables or **features (X)**. The model learns an approximate mapping function f such that y = f(X).

* **Continuous Output:** The target variable y is a real-valued number, which can be an integer or a floating-point value. These often represent quantities, amounts, sizes, etc.
* **Example:** Predicting the specific dollar value a house might sell for (e.g., $182,000), based on features like its square footage, number of bedrooms, and location.

Key Characteristics of Regression:

* **Predicting a Quantity:** The core task is to estimate a numerical value.
* **Input Types:** Input features (X) can be real-valued (e.g., square footage) or discrete (e.g., number of bedrooms).
* **Multivariate Regression:** When the prediction depends on multiple input features (like the house price example).
* **Time Series Forecasting:** A special case of regression where the input variables are ordered by time, and the goal is often to predict future values based on past values.

Because a regression model predicts a quantity, its performance or "skill" must be reported in terms of the **error** or difference between its predictions and the actual values.

Example: Linear Regression Model

A common algorithm for regression is Linear Regression. It assumes a linear relationship between the input features and the output variable.

A typical Linear Regression model looks like this:

y = a₁x₁ + a₂x₂ + a₃x₃ + ... + b

Where:

* y is the predicted output value (the Target Feature, e.g., House Price).
* x₁, x₂, x₃, ... are the input Features (e.g., Square Foot Area, Number of Bedrooms, Number of Balconies).
* a₁, a₂, a₃, ... are the coefficients or weights assigned to each feature, representing their importance or impact on the output.
* b is the bias term or intercept.

**The core problem** for the Linear Regression learning algorithm is to **find the optimum coefficient values** (a₁, a₂, a₃, ..., b) by learning from the training data.

Regression Metrics Worked out (Example Scenario)

Let's consider predicting house prices based on square footage (x1), number of bedrooms (x2), and number of balconies (x3).

Suppose our trained Linear Regression model learned the following mapping function:

**y = 27\*x₁ + 121\*x₂ + 235\*x₃ + 766**

Now, let's see how well this model predicts prices for some new houses (test data) it hasn't seen during training:

| **x1 (SqFt Area)** | **x2 (No of BRs)** | **x3 (No of Balconies)** | **y-act (Actual House Price)** | **y-pred (Predicted House Price)** | **Difference (Error: y-act - y-pred)** |
| --- | --- | --- | --- | --- | --- |
| 1650 | 3 | 2 | 2,23,000 | 2,25,600 | -2,600 |
| 1320 | 2 | 1 | 1,76,000 | 1,82,000 | -6,000 |
| 1830 | 3 | 2 | 2,53,000 | 2,57,700 | -4,700 |
| 1620 | 3 | 2 | 2,29,000 | 2,31,000 | -2,000 |
| 2680 | 5 | 3 | 4,52,000 | 4,51,200 | 800 |

*(Note: Predicted values calculated using the formula above, likely with some rounding in the example)*

The "Difference" column shows the error for each individual prediction. To evaluate the model overall, we need metrics that aggregate these errors across the entire dataset. We will look at common regression metrics next.