Basic ML terminology

Let's break down the basic terminology of Machine Learning in an intuitive way, often using the house price prediction example from before to illustrate.

1. Data

At the heart of machine learning is **data**. This is the raw information that the machine learning algorithm will learn from. Data typically comes in the form of a table (like a spreadsheet), where rows represent individual observations or examples, and columns represent different attributes or pieces of information about those examples.

2. Features (Input Variables, Independent Variables)

- **What they are:** These are the **input characteristics** or **attributes** of your data that you use to make a prediction or find a pattern. Think of them as the "clues" or "pieces of information" available for each data point.
- **Example (House Price Prediction):**
    - Number of bedrooms
    - Number of bathrooms
    - Square footage
    - Age of the house
    - Locality/Neighborhood

3. Labels (Output Variables, Dependent Variables, Target)

- **What they are:** This is the **output** or the "answer" that you are trying to predict or learn from your features. It's the variable you want the machine learning model to be able to determine.
- **Example (House Price Prediction):**
    - The **actual selling price** of the house.
- **Example (Spam Detection):**
    - "Spam" or "Not Spam" for an email.
- **Example (Image Recognition):**
    - "Cat," "Dog," "Bird" for an image.

4. Examples (Data Points, Observations)

- **What they are:** A single row in your dataset, representing a complete set of features and its corresponding label (if available). Each example is one instance of the data the model learns from.
- **Example (House Price Prediction):**
    - One specific house with its number of bedrooms, bathrooms, square footage, age, and its final selling price is a single example.

5. Model (Algorithm, Learner)

- **What it is:** This is the **algorithm** or the mathematical function that learns the relationship between the features and the labels. It's the "brain" of your machine learning system. The goal of the model is to find patterns in the input data (X) that allow it to predict the output data (Y).
- **Example (House Price Prediction):**
    - A **Linear Regression model** might learn a formula like:
        - Price = (Coeff_1 * Bedrooms) + (Coeff_2 * SqFt) + (Coeff_3 * Age) + Intercept
    - Here, Coeff_1, Coeff_2, etc., are parameters that the model learns during training.
- **Think of it as:** The "rule" or "formula" that your real estate agent (from the supervised learning example) developed in their head to estimate prices.

6. Training (Fitting)

- **What it is:** This is the process where the machine learning model **learns** from the provided labeled data (features and their corresponding labels). During training, the model adjusts its internal parameters (like the coefficients in our linear regression example) to minimize the difference between its predictions and the actual labels.
- **Process:** You feed the model a large dataset of *known* examples (e.g., historical house sales where both features and prices are known). The model iteratively refines its internal logic based on this data.
- **Analogy:** The real estate agent studying past sales records and trying to figure out the best way to estimate prices.

7. Training Data

- **What it is:** The specific subset of your overall data that is used to **train** the machine learning model. It contains both features and their corresponding labels.
- **Why split?** Typically, you split your entire dataset into training and testing sets. The training data is used for the learning process.

8. Prediction (Inference)

- **What it is:** Once a model has been trained, it can be used to make an educated guess or **estimate** for new, unseen data. You provide the model with features for which you *don't know* the label, and the model uses its learned patterns to produce an output.
- **Example (House Price Prediction):**
    - You input the features of a *new* house (bedrooms, bathrooms, sq ft, age) into your trained model. The model then outputs its **predicted price** for that house.

9. Test Data

- **What it is:** A separate subset of your overall data that is **not used during training**. It's kept aside to evaluate how well the trained model performs on data it has never seen before.
- **Why use it?** This helps to ensure that your model hasn't just memorized the training data (overfitting) but can actually generalize to new, real-world examples.

10. Accuracy/Performance Metrics

- **What they are:** Ways to measure how well your trained model is performing. Different types of ML problems use different metrics.
- **Examples (Regression - like House Prices):**
    - **Mean Squared Error (MSE):** Measures the average squared difference between predicted and actual values. Lower is better.
    - **Root Mean Squared Error (RMSE):** The square root of MSE, easier to interpret as it's in the same units as the label.
    - **R2 (R-squared):** Indicates how much of the variance in the label can be explained by the features. Higher is better.
- **Examples (Classification - like Spam Detection):**
    - **Accuracy:** The percentage of correctly classified examples.
    - **Precision:** Of all items the model labeled as positive, how many were actually positive?
    - **Recall:** Of all actual positive items, how many did the model correctly identify?
    - **F1-Score:** A balance between precision and recall.

These are the fundamental terms you'll encounter constantly when delving into machine learning. Understanding them is crucial for comprehending how ML systems work.