What is Machine Learning?

- ML is the branch of AI and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.
- ML is the study of computer algorithms that improve automatically through experience by using data.

Machine Learning Methods.

It has three primary categories.

- **Supervised ML:**

  The type of ML in which labeled datasets are used to train models/algorithms, and the model make accurate predictions based on the labeled data.

  - **Use:**
    - Classification of spam messages
    - House prices
    - Is it a cat or dog?
    - Who are the unhappy customers?

  - **Supervised ML Methods:**
    - Neural Networks
    - Naïve Bayes
    - Linear Regression
    - Logistic Regression
    - Random Forests
    - SVM and many more

- **Unsupervised ML:**

  It uses ML algorithms to analyze and cluster unlabeled data. These algorithms find hidden patterns or groupings without the need for human intervention.

  - **Use:**
    - AirBnB
    - Amazon
    - Credit Card Fraud Detection

  - **Unsupervised ML Methods:**
    1. Clustering
       - K-Means
       - Density based spatial clustering of application with noise (DBSCAN)
       - Hierarchical Cluster Analysis (HCA)
    2. Anomaly Detection and novelty detection
       - One-Class SVM
       - Isolation Forest

3. Visualization and dimensionality reduction
    - Principal Component Analysis (PCA)
    - Kernel PCA
    - Locally-Linear Embedding (LLE)
    - t-distributed Stochastic Neighbor Embedding (t-SNE)
4. Association rule learning
    - Apriori
    - Eclat

➢ **Reinforcement ML:**

It is a behavioral ML model that is similar to supervised Learning, but the algorithm is not trained using simple dataset. This model learns as it goes by using trial and error. A sequence of successful outcomes are reinforced to develop the best recommendation or policy for a given problem.

- ## Why we use ML?

➢ Consider how to write a spam filter by using traditional programming techniques.

1. First to notice some words or phrases (such as “4U”, “credit card”, “free”, and “amazing” etc) tend to come up a lot in the subject line. Also some abnormalities in the sender’s name, the email’s body and so on.
2. Write a detection algorithm for each of the pattern that is noticed, and in this way the algorithm would flag the email as a spam if a number of these patterns are detected.
3. Test the program and repeat step 1&2 until it is good enough.

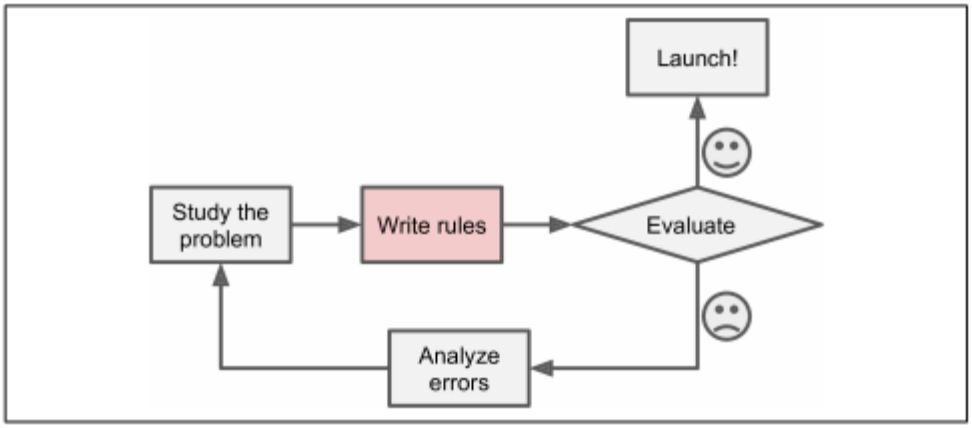

Figure 1-1. The traditional approach

In contrast, the spam filter based ML automatically learns that which words or phrases are good predictors of spam by detecting unusually frequent pattern of words in the spam example.

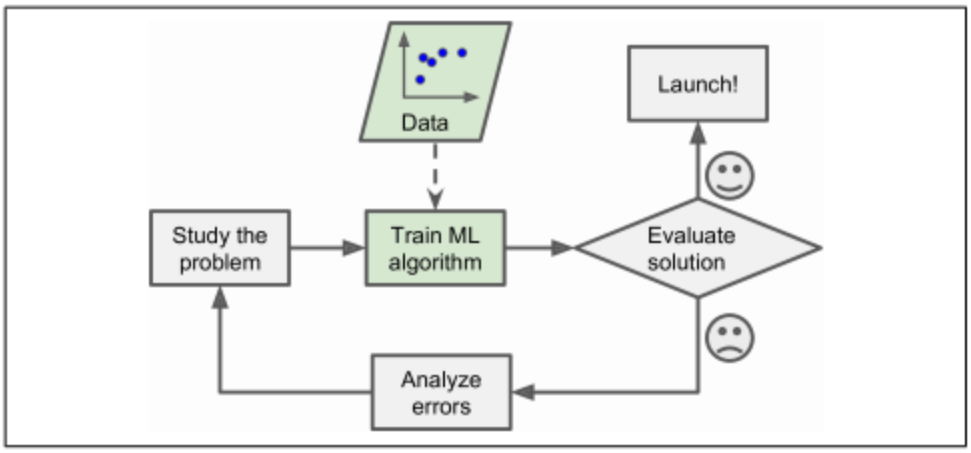

Figure 1-2. Machine Learning approach

If spammers find out that our filter is blocking emails that containing “4U”, they may write it “For U”. So by using traditional programming techniques, we will need to update our model, which will consume time if it keep going for other words or phrases.

In contrast, the spam filter based on ML will automatically learns that “For U” has unusually frequent in the spam and will keep flagging without our intervention.

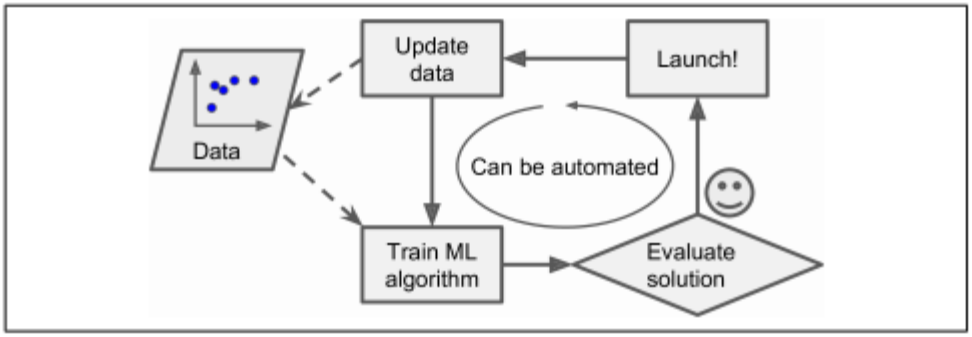

Figure 1-3. Automatically adapting to change

To summarize, Machine Learning is great for:

1. Problems for which the existing solutions require a lot of hand-tuning, or long list of rules, one Machine learning algorithm can often simply the code and perform better.
2. Complex problem for which there is no such good solution at all using a traditional approach, the best ML techniques can find a solution.
3. Fluctuating environment: a ML system can adapt to new data.
4. Getting insights about complex problems and large amount of data.