UNIT-I (Supervised Learning Algorithms)

Introduction to Machine Learning: Machine Learning is the “science of programming computers so they can learn from the data”.

In other words, machine learning is the field of study that gives “computers the ability to learn without being explicitly programmed”.

A computer program is said to be learned from experience **E** with respect to given Task **T** and some performance measure **P** and that performance measure **P** improves with the Experience E.

For example- spam filter program - that can learn to flag from some given examples (Training Set) of spam emails. Here each containing example is called a training instance(sample).

Here T- flag spam for emails

E- Training data

P- Correctly specified email such as Accuracy is one of performance measures.

Why Machine Learning is required:

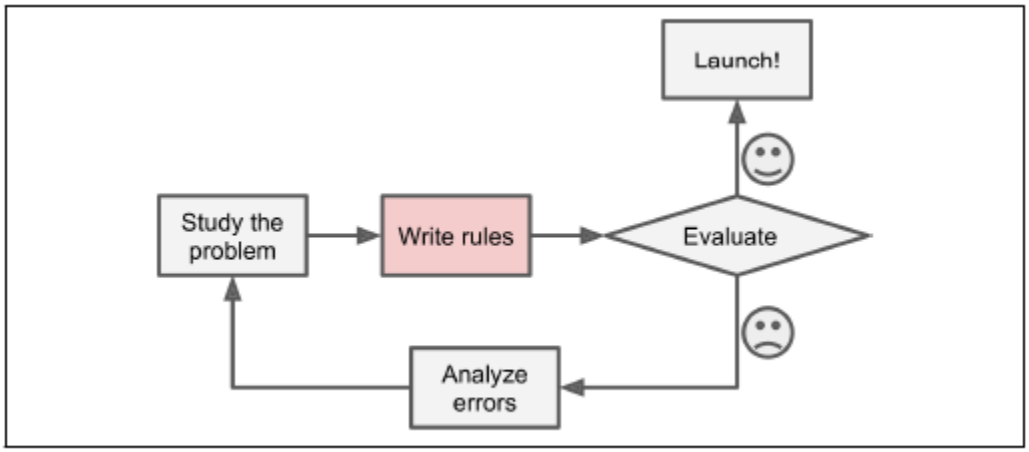

Figure 1-1. The traditional approach

A. Traditional programming for some problems is very complex because you need to write so much rules.(These rules are very hard to maintain).
B. Traditional Programs may work for one data set but fail for another dataset.
C. Machine learning programs are much shorter and easier to maintain and most likely more accurate than traditional ones.
D. If there is any update in spam detection then you need to update the entire algorithm in a traditional approach.

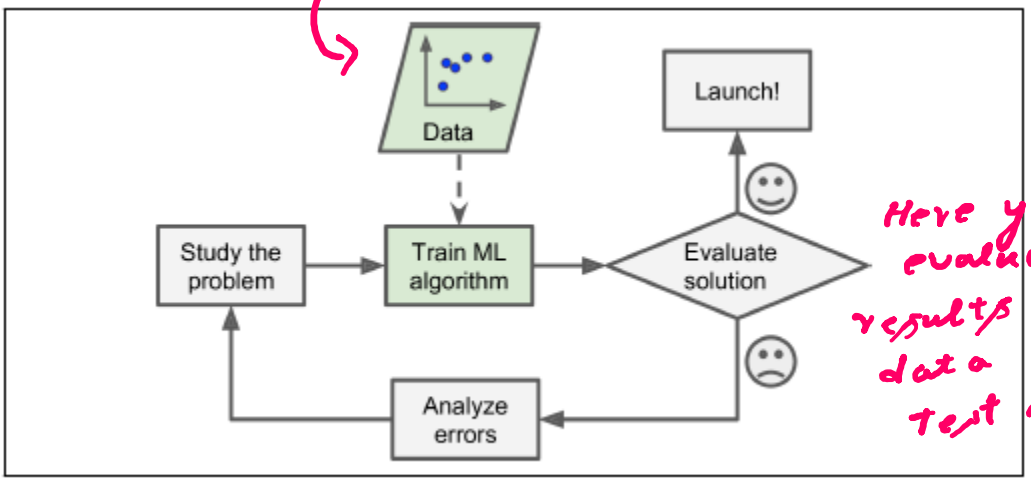

Figure 1-2. Machine Learning approach

Advantages of Machine Learning: -

1. **Easily identifies trends and patterns-** Machine learning can review large volumes of data and discovers specific trends and patterns that would not be apparent to humans.
2. **No Human intervention is needed-** Since in ML, giving machines the ability to learn, it lets them make predictions and improve the algorithm on their own.
3. **Continues Improvements-** They keep improving in accuracy and efficiency. Prediction and accuracy is increased with the growing data.
4. **Handling multidimensional and multivariate data:**
5. **Wide Applications-**

Disadvantages of Machine Learning: -

1. **Data Acquisition:** ML requires large dataset to train, unbiased data and good quality data.
2. **Time and Resources:** ML needs enough time to let the algorithm learn and develop, with considerable amount of accuracy and relevancy.
3. **Interpretation of Result:** Another challenge is to accurately interpret the results generated by these algorithms.
4. **High error susceptibility:** ML is autonomous but highly suspicious to errors.

   Poor training dataset leads to biased prediction which is coming from the biased training data.

   These errors must be unnoticed from a long period of time.