**Introduction to Machine Learning**

Machine Learning is by far the most used term in technology and computer science these days. But what does it mean? **Does it mean that Machines start learning just like humans? Or something else**. This chapter will help us understand this.

First let us start with a high level definition of machine learning.

**Definition**

*Machine learning is defined as a set of techniques which enables a computer to learn a task which otherwise would have been impossible to program using traditional rule based programming.*

We don’t manually feed rules to the computer but rather lots of good data and computer algorithms are designed in such a way that they extract meaningful patterns or rules from the data. It is closely linked to the way a child learns from lots of examples from their surroundings.

**Example**

Let us take an example where our task is to predict the likelihood of rain on any given day to differentiate between rule based and ML based systems.

**Traditional Rule Based Programming**

In the rule based programming, we can define some set of rules and if on any given day those rules are satisfied, then we can say that it will rain. For example, we may have a rule that if Temperature > 10 degrees Celsius, then the likelihood of rain will go down. Just like temperature, we can have many other parameters which can be manually set to determine the likelihood of rain today.

Hence, the key feature of rule based programming is that we feed the rules ourselves to the system. The human touch is that we find patterns and feed those patterns or rules in the system manually.

**Machine Learning Based system**

On the contrary, ML based systems will get some past data (where we prepare data in such a way that against each list of features, temp was one of them, we will label the right answer, i.e likelihood of raining) and the system will try to extract patterns or rules from that data itself.

The following picture best summarizes this difference between ML and rules based systems:



Now the question arises “Is ML always superior to the Rule based system?”. In the following two sections, we will discuss precisely that:

**ML over Rule based Systems**

The ML should be a preferred option if:

1. The exact rules of the task cannot be clearly known in advance. For example:
   1. Face Classification
   2. Language Translation
2. The number of rules has become really huge.

**Rule Based over ML**

The rule based can also be used over ML when::

1. The exact rules can be known like for any mathematical task for which we have the exact functions like:
   1. Calculating roots of a given polynomial
   2. Calculating eigenvalues and eigenvectors
2. The number of unique rules are really small.

**Exercise.1 Determine which of the systems is more appropriate for the following examples**

Rule based or Machine learning based system would be more suitable for the following situations:

1. You want to build a system which is going to track the sensor data.
2. Build a software which will recognize number plates in a video
3. Autonomous braking system for car
4. Building a game engine.

**Applications of Machine Learning**

There were many tasks in Artificial Intelligence which were very difficult to imagine before the ML era but now we have begun to solve some of the most difficult AI problems. This has been possible not just because of ML algorithms but also because of advances in computing capabilities.

Some of major applications of machine learning are:

1. **Use of ML in Credit Risk Modeling**: Before ML era, there were credit risk models where majorly we had some rules defined to grant a loan application. Now, because of ML, these models can use extensive list of features for ML models from spending and balance to their credit scores etc
2. **Computer Vision Related applications** : Thinking about creating rules to identify a person’s face in an image was really daunting because exceptions cases were huge. Deep learning, a sub field of Machine learning, has been particularly successful in tackling challenges of computer vision. One algorithm which has been particularly helpful in this field is Convolution neural networks by Yann Le Cunn. Some computer vision applications include:
   1. Face Detection
   2. Face Recognition
   3. Object movement tracking
   4. Image Style Transfer
   5. Self driving cars
   6. Pedestrian Detection
   7. Automatic vehicle speed detection
   8. Cancer Detection
3. **Language Related Application**: Just as in Vision, CNN was the breakthrough algorithm, in language, Recurrent Neural Networks (RNNs) were a major development and because of this neural network architecture, a lot of applications have a solution now. Some of the applications include:
   1. Machine Translation
   2. Text summarization
   3. Text sentiment analysis
   4. Text classification
4. **Other Applications**: Apart from the above, ML also has several other applications in the Financial sector, Medical sciences, Robotics Etc. Some other applications include the use of unsupervised learning algorithms to group the similar items such as google news does to produce similar news stories.

**Types of Machine Learning Algorithms**

Earlier, we discussed that Machine learning algorithms take past data (right examples) and try to find the underlying rules or patterns in that data. In Machine learning, we have features and we may have the right answer to supervise the learning process. The features will always be there. The right answer may or may not be available. Basis this, we can categorize the ML algorithms in the following two classes:

1. Supervised ML algorithms
2. Unsupervised Ml algorithms

Let us briefly discuss each of these.

**Supervised Machine Learning**

As we discussed, a machine learning algorithm may or may not have the right answer, Supervises ML refers to those problems where the right answers are available to supervise the learning process. Hence, the name Supervised.

Now, basis the type of the right answer, it can further be divided into the following types:

1. **Regression** : For these supervised learning tasks, the type of right answer is continuous. For example, if we want to predict the price of the stock market or the temperature consumption by the CPU etc, we will use Supervised Regression Algorithms.
2. **Classification**: Here, the type of the right answer will be categorical in nature. For example, if we wish to predict whether there is a human in an image or not, then we will use these types of algorithms.
3. **Ranking** : It may be classified under the classification category but they have some separate algorithms dealing with rankings. For example, to predict the ranking of Football players, we may use these classes of supervised algorithms.

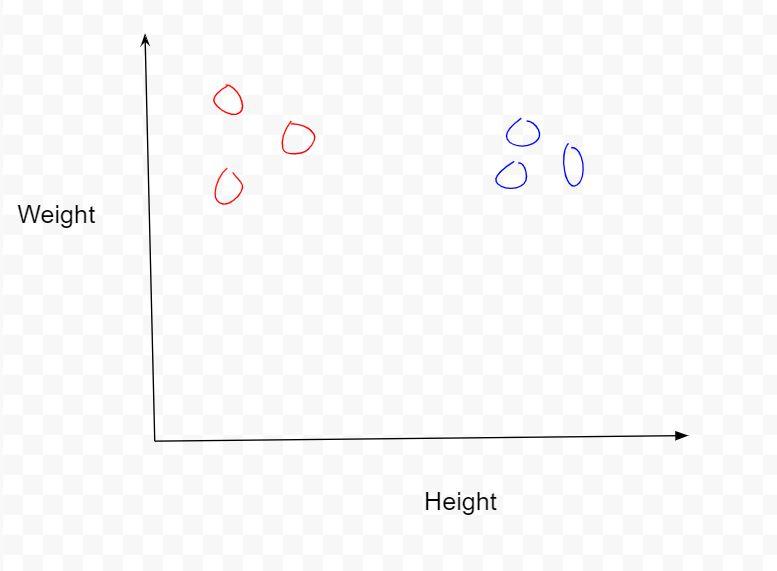
The classification of supervised algorithms which we discussed just now is not exactly the name of algorithms but rather the name of the broad tasks. We may solve one of them with many different algorithms. Which one we choose is decided by the performance of the model.

For example: a classification problem can be solved by Logistic Regression, Decision Tree, Random Forest, Artificial Neural Networks Etc. We have to choose one and usually we choose the one which has the best performance.

**Unsupervised Learning Algorithms**

These are classes of ML algorithms where the right answers are not given but only features are provided. So, usually here the task is to find some pattern in the features itself and then that can be used to label the data or maybe find the similar set of features.

To illustrate this point, let us look at the following diagram:



There are six pairs of height and weight. You can think of these two as features of 6 different persons. Although, we don't have the information on which person belongs to which group by giving a new pair of height and weight, we can visually look at this plot and point out where that person might belong.

This is precisely what is done in Unsupervised algorithms. It is to find structure and pattern in the features without the right answers.

It can also be of several type but let us discuss two of those:

1. **Clustering**: Clustering is about exactly what we discussed. We want to cluster similar looking features in one group and if any new feature comes, we can easily see which group this belongs to.
2. **Dimensionality reduction**: This refers to reducing the dimensions of our feature vector while preserving the maximum information. So, for example, if I have 10 features values which might contain height, weight, eye color etc like body parameters. Based on this, we might have a task to reduce it to 2 dimensional to plot the data effectively because it is hard for us to visualize a 10 dimensional data.

**Exercise.2 Classify the machine learning problem into Supervised or Unsupervised**

Which machine learning type of problem are the following, if supervised, classify it into Regression or Classification:

1. To classify a news article into a coherent group where all other news are similar.
2. To predict who will win the USA election next year.
3. Given a list of 10 features, to reduce them to 3 features.
4. To be able to classify the a given email as spam or not spam.
5. Find closely Related stars based on their characteristics.

**Process of Building a Machine Learning System**

Building a machine system is an iterative process where there are many pieces which need to be repeatedly created and refined over a period of time. Hence, it becomes pertinent to have a framework which makes sure that we are on the right track and does not get lost in the process of building a large system.

In particular, we follow the framework called CRISP-DM process which involves the following steps. Note that it is not a top down approach but rather an iterative process where we can move back and forth between each of the steps to any of the previous steps.

1. Problem Statement (Business Understanding)
2. Data Understanding
3. Data Preparation and Exploratory data analysis
4. Modeling
5. Evaluation
6. Deployment

After evaluating the model usually, we go back to the data preparation stage and try to refine the model to get better results.

This iterative process is best captured by the following diagram:



Let us briefly discuss each of these steps :

1. **Business Understanding**: In this step, the idea is to understand the problem that we are looking at and look at it from the business person’s perspective. Also, in this step, it is to be determined whether the effort is required at all to build a ML system or to just go with rule based systems if possible. For example, if we are looking at building a computer vision system for automatic attendance, then business understanding would entail that we look at the costs associated with it and also determining if building this system is worth the effort
2. **Data Understanding**: Once we have determined that an ML system is something that we have to build, the next step would be to do a little bit of data understanding. Here, we would look at what kind of data is required and how it can be made available. If the organization is maintaining a database, they might look at building a data lake or data warehouse for the system.
3. **Data Preparation and EDA**: In this step, data scientists would start preparing the data for the objective at hand and also do some initial exploratory analysis to get a quick look at the insights that they might find in this initial EDA.
4. **Data Modeling**: They will create a first version of the model in this step using the features which were selected in the above step of EDA. Features which have a high correlation with the target will be considered for the modeling.
5. **Evaluation**: The first version of the model which was created in the previous step will be evaluated using the appropriate methods and then it will be decided whether to continue with this version of model or to refine it.
6. **Deployment**: If the evaluation of the given version passes the test, then it will be used as the final model in production and deployed and finally used to serve the predictions.

**Summary**

This chapter introduced the basics of machine learning where we first defined how a machine learning system works and how this system differs from a traditional rule based programming system. We also looked at some of the cases where Rule based systems might be used over ML. Further, we looked at some of the applications of Machine learning and we hope this made you excited about this field..

Then, we looked at various types of machine learning tasks such as Regression, Classification, Ranking, Clustering, Dimensionality reduction Etc.

We also saw the process of creating a machine learning system. We looked at the 6 step iterative process to create a machine learning ster,

In the next chapters, we will start with the practical aspects of the things we discussed. In particular, we will discuss how to implement Regression algorithms in great detail in Python.