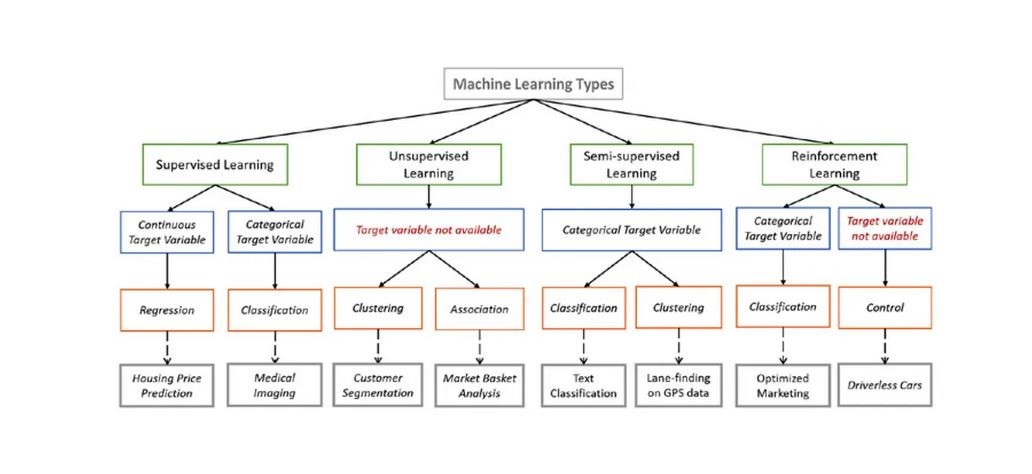
**(ML) Section 4 Assignment 2 Main Page- Mathematical Models for Machine Learning**

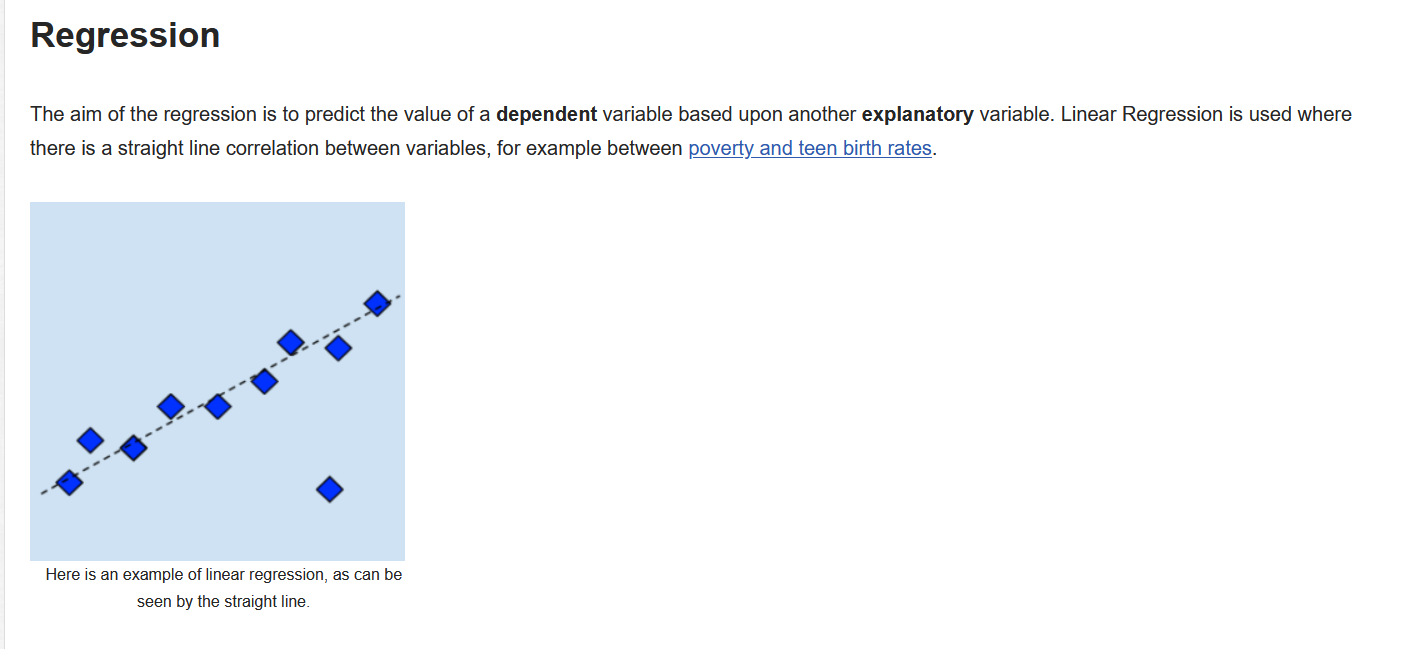


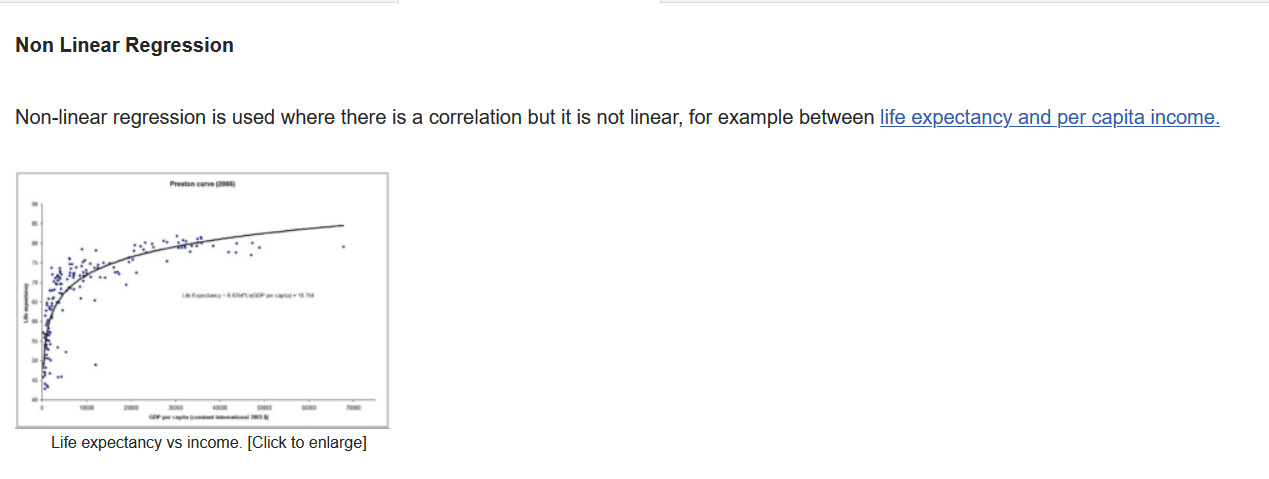
The image above was shown in the last reading and the three main types of learning (supervised, unsupervised, semi-supervised (hybrid)) were discussed. Now we will talk about the data types that it best uses, the mathematical model that is used and an example of what kind of problems it can solve.

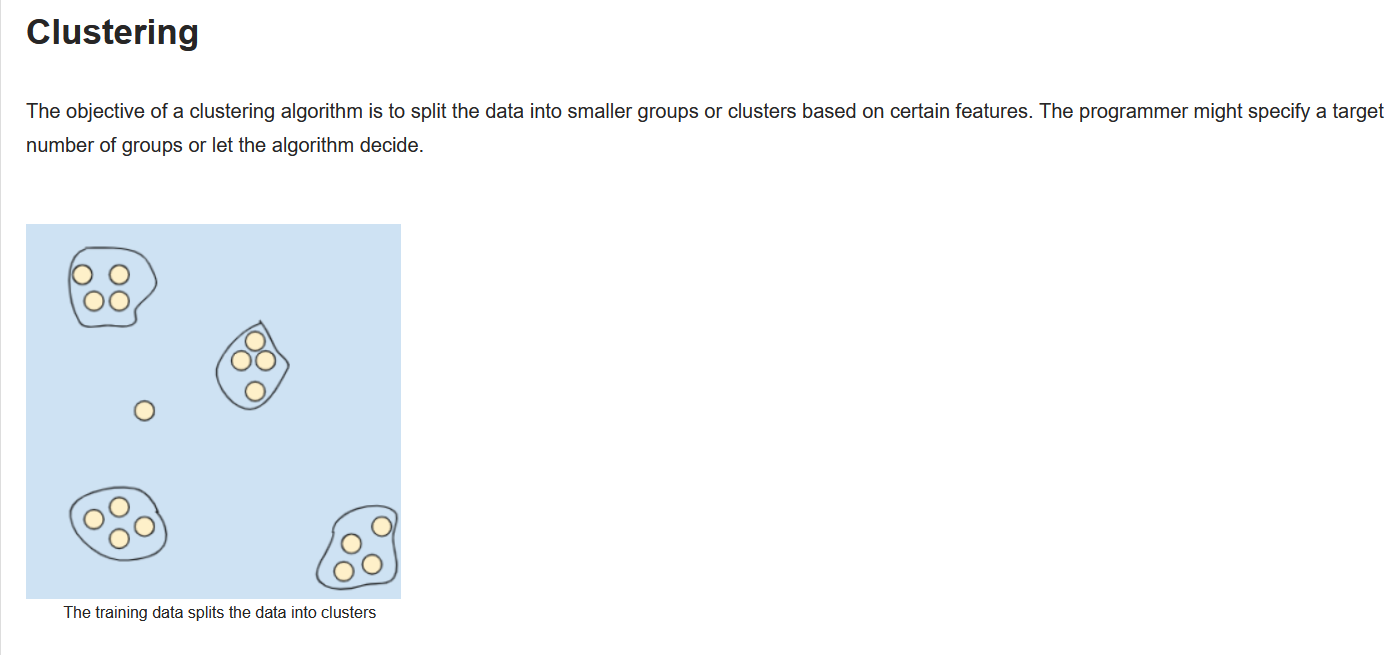
Supervised learning can use either continuous (analogue) data or categorical data. Continuous data is something like time or temperature. Categorical data is categories like price or weight. Depending on the outcome, you can use regression which compares the data to each other (cost/time) or classification (what it is or is not).

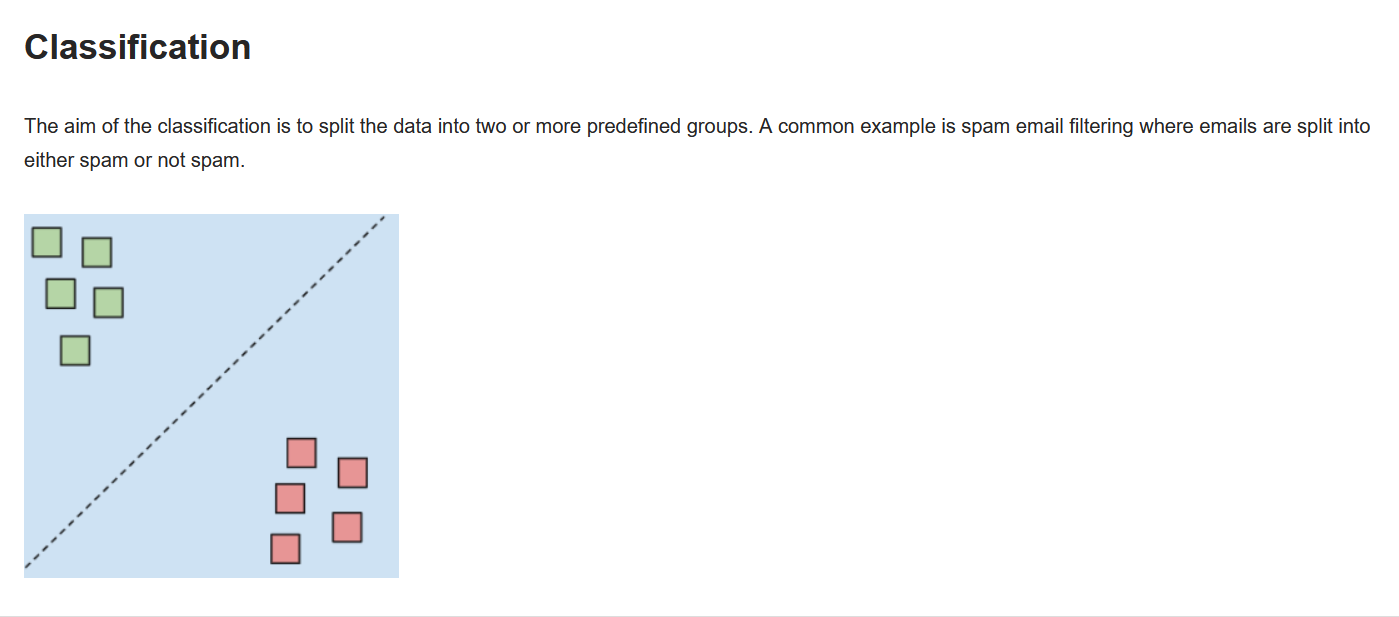
Unsupervised learning can use the target data what ever it is (images, categorical) and either cluster it or associate it. For example, it can identify what kinds of consumers buy Porsches in order to target that demographic or it can decide is a stock or company is a bargain for investors.

Read through these main Mathematical models



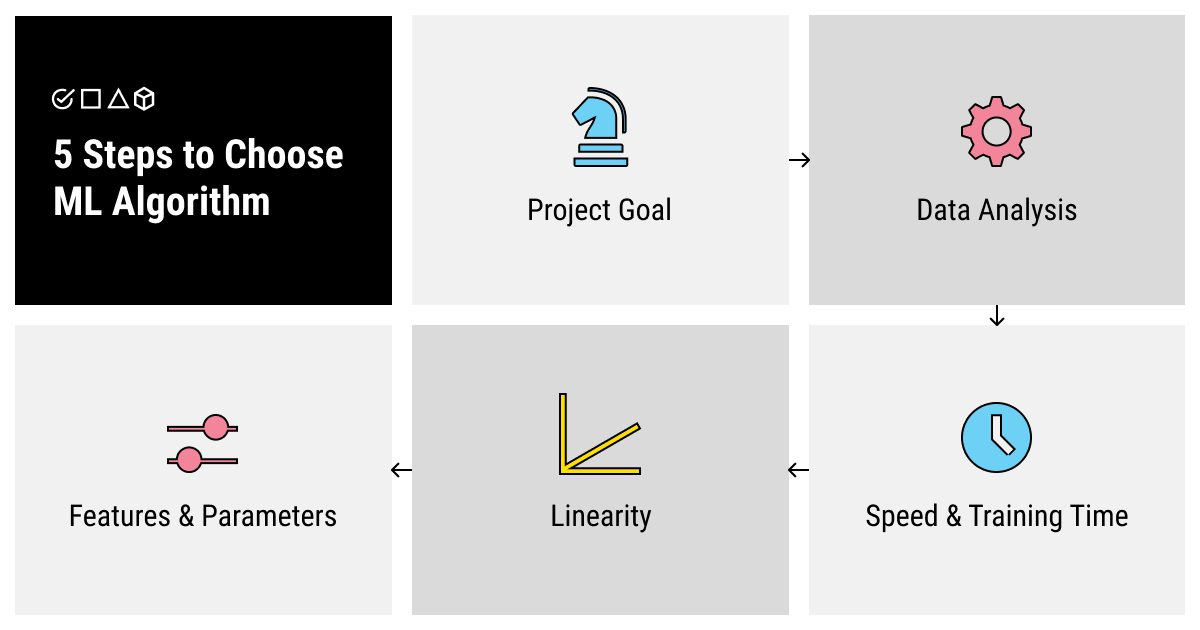






These are some of the simplest models for more complex models take a look at this [Wiki site.](https://machine-learning.paperspace.com/wiki/machine-learning-models-explained).

5 Simple Steps to Choose the Best Machine Learning Algorithm That Fits Your AI Project Needs (from [How to Choose a Machine Learning Algorithm.](https://labelyourdata.com/articles/how-to-choose-a-machine-learning-algorithm))



5 steps to choose and ML algorithm

Learning about the different types of machine learning algorithms is not enough to understand how to choose the one that fits your specific purpose. So let’s stick to an incremental method and see how exactly you can approach this problem.

**Step 1. Understand Your Project Goal**

As it has already become apparent, each machine learning algorithm was designed to solve a specific problem. So, first of all, you should consider the type of project that you’re dealing with.

Answer this question: what kind of an output do you need? Do you need an algorithm for prediction based on the previous data? Turn to supervised forecasting algorithms. Are you looking for an image recognition model that will work with poor-quality photos? Dimensionality reduction in combination with classification will help you with it. Do you need to teach your model to play a new game? A reinforcement algorithm will be your best bet.

**Step 2. Analyze Your Data by Size, Processing, and Annotation Required**

When you’ve answered the question of what type of output you need, ask yourself what input do you have. What is your data like? Is it raw, just collected from wherever, and requires processing? Is it biased, dirty, and unstructured? Or do you already have a big annotated dataset on your hands? Do you have enough data or is additional collecting (or even collecting from scratch) required? Do you need to spend time preparing your data for the training process or are you good to go?

Insufficient, poor-quality, unprocessed data usually doesn’t lend itself to great training of a supervised algorithm. You should decide if you want to spend time and resources on preparing the best data you can before starting the training process. If not, you can opt for unsupervised algorithms but keep in mind the limitations of such a choice.

**Step 3. Evaluate the Speed and Training Time**

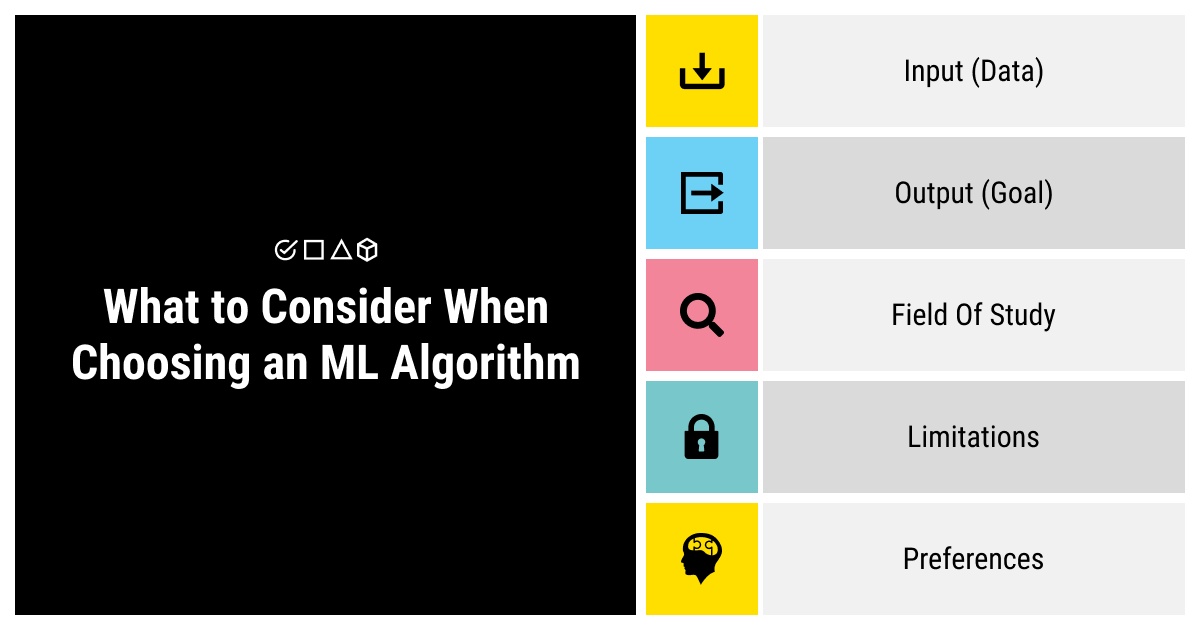
Here’s another question for you to answer that can help you understand what type of machine learning algorithm you need. Do you need it fast even if it means lower quality of training (and, respectively, predictions)? More and higher-quality data lead to better training. Can you allocate the required time for proper training?

**Step 4. Find Out the Linearity of Your Data**

Another important question is what the environment of your problem is like? Linear algorithms (such as linear regression or support vector machines) are simpler and faster to train. However, they are not usually used for more complex problems as they deal with linear data. If the data is multifaceted, multidimensional, and has many intersecting correlations, linear algorithms might not be sufficient for your task.

**Step 5. Decide on the Number of Features and Parameters**

Finally, how complex and accurate your final AI model should be? Don’t forget that longer training usually leads to better, more accurate performance when the AI model is deployed. You can specify more features and parameters for your model to interpret if you have time to let it train longer. So giving your algorithm more time to learn may be a good investment into your future output accuracy and interpretability.



What to consider when choosing an ML algorithm

Choosing a machine learning algorithm is obviously a complex task, especially if you don’t have extensive experience in this field. However, learning about the types of algorithms and the tasks that they were designed to solve and answering a set of questions might help you solve this problem. Try to outline as much as you can about:

* Your input (the data: is it collected/sufficient/processed/annotated?)
* Your output (what goal do you pursue?)
* Your field of study (how linear or complex the data is?)
* Your limitations (can you spare time and resources?)
* Your preferences (what features do you absolutely need for success?)

Learning more about machine learning algorithms, their types (from supervised and unsupervised to semi-supervised and reinforcement learning), and answering these questions might lead you to an algorithm that’ll be a perfect match for your goal.

**TO DO:**

For each of the problems below – determine what the data sets needed would be, if you would use a supervised or unsupervised learning model and what the type of mathematical model you would use.

1. Spam identification. Most of our email inboxes also have an unsolicited, bulk, or spam inbox, where our email provider automatically filters unwanted spam emails.  But how do they know that the email is spam? They use a Machine Learning model to identify all the spam emails based on common characteristics such as the email, subject, and sender content.
2. Recommender systems are one of the most characteristic and ubiquitous machine learning use cases in day-to-day life. These systems are used everywhere by search engines, e-commerce websites (Amazon), entertainment platforms (Google Play, Netflix), and multiple web & mobile apps. Prominent online retailers like Amazon and eBay often show a list of recommended products individually for each of their consumers. These recommendations are typically based on behavioral data and parameters such as previous purchases, item views, page views, clicks, form fill-ins, purchases, item details (price, category), and contextual data (location, language, device), and browsing history.
3. From Alexa and Google Assistant to Cortana and Siri, we have multiple virtual personal assistants to find accurate information using our voice instruction, such as calling someone, opening an email, scheduling an appointment, and more. These virtual assistants use Machine Learning algorithms for recording our voice instructions, sending them over the server to a cloud, followed by decoding them using Machine Learning algorithms and acting accordingly.
4. Managing an increasing number of online customer interactions has become a pain point for most businesses. It is because they simply don’t have the customer support staff available to deal with the sheer number of inquiries they receive daily. Machine learning algorithms have made it possible and super easy for chatbots and other similar automated systems to fill this gap. This application of machine learning enables companies to automate routine and low priority tasks, freeing up their employees to manage more high-level customer service tasks.