Course: MMO: Introduction to Machine Learning

Video Number:01-04

Estimated Length: 5 min

|  |  |  |
| --- | --- | --- |
|  | **Script text or talking points** | **Visuals / Actions on Screen** |
| 1 | There are three general areas of machine learning: supervised learning, unsupervised learning and reinforcement learning. Most people who are just getting into machine learning start with supervised learning.  Supervised learning tends to be the most time consuming as it requires the person building the model to identify and label relevant data. In most cases, this model builder needs a good understanding of the problem domain to be able to best identify what data will give the result they are looking for and be able to recognize what are reasonable results.  Within supervised learning there are two areas: classifications and regressions. Classifications take data and put them into known classes. For example, given a picture of a tree, is it a pine tree, a maple tree, a palm tree or an oak tree? Regressions are for data predictions, usually time series data, where you want to make a prediction on a future event. An example of this is predicting the number of COVID cases at a future date based on several factors given the history of the spread of the disease. | Slide 1 |
|  | Unsupervised learning takes the bulk of the labeling process out of the equitation and is used to find insights into the data you may not be able to make on your own. Because of this the work of preparing the data tends to be less labor intensive.  There are a lot of uses for unsupervised learning. You may have a collection of customers that you want to divide into segments that make sense based on complex histories with each costumer, but in turn the data may be too complex for your to envision how to best split them up. Clustering can find commonalities in the data and divide customers into some predefined number of groups that reflect the data.  Unsupervised learning can also be used find outliers in data. This is called anomaly detection and is commonly used to do things like understand failures in a system.  There is also expectation maximization that can find patterns, or glean insight, from large amounts of data. An example of this is text mining that may categorize emails into groups of commonalitie.  There are many other cases of unsupervised learning. Anytime you want some insight into the data where you yourself don’t know the answer upfront, you are likely using some type of unsupervised learning. | Slide 2 |
|  | Finally, there is reinforcement learning. This is the process of creating a system where the model can get better over time using a reward / penalty system. Such a system is always attempting to maximize the reward and has a wide variety of uses.  One example would be a system that looks at the history of energy usage in a computer and uses that information over time to devise ways to use as little energy as possible. In this case the actions and the resulting energy use would provide both the reward and the penalty.  It is also used commonly in game AIs. The reason why AIs can now beat chess grandmasters is that they are no longer based on a static set of rules, but instead looking at rewards and penalties for possible moves. This allows them to completely overcome strategies, such as playing illogically, that stimmed rules based Ais in the past.  Even the recommendations we get from online shopping and browsing are a type of reinforcement learning where the model looks at your purchasing over time and recent browsing history to try and create increasingly accurate recommendations for you.  Any time you see a system that appears to get better or change over time, you are likely looking at some sort of reinforcement learning. |  |