Machine Learning

Brian Wang

October 30, 2017

Machine learning

What is Machine Learning?

Machine learning is a field of computer science that gives computers the ability to learn without being explicitly programmed. [1] The term “Machine learning” was firstly introduced in 1959 by Arthur Samuel, an American pioneer in field of computer science and artificial intelligence. Machine learning is a new revolution from pattern findings, computational learning theory in artificial intelligence; machine learning constructs algorithms that learn from the given data set and from that make prediction and decision.

Top Usesage of Machine Learning in real life

1. Data Security
Internet Security Companies, like Kaspersky Lab and Norton, found that many malwares seem to have similar code structures, then they develop this learning model with previous data base to detect malware. As the learning model stores more and more malware code structures, the model will increase accuracy and identify new malware.

2. Financial Trading
Learning model for financial trading conducts a series of analysis on the past movement of the stock price and relevant information (financial report, Earnings report, new press and domestic news and international news and etc.), the most common strategies are regression and quantitative trading. Some finance companies that use these strategies to generate profits: Two Sigma, D.E. Shaw.

3. Healthcarela
Machine learning algorithms can process more information and spot more patterns than their human counterparts. One study used computer assisted diagnosis (CAD) when to review the early mammography scans of women who later developed breast cancer, and the computer spotted 52% of the cancers as much as a year before the women were officially diagnosed. Additionally, machine learning can be used to understand risk factors for disease in large populations. The company Medecision developed an algorithm that was able to identify eight variables to predict avoidable hospitalizations in diabetes patients.[2]

4. Marketing Personalization
Everyone who surf on the internet will experience this. When you visit online shop and look at the product but don’t buy it, then you might see digital ads across the web for exactly the product you just saw. Companies collect the data you have created (what computer scientist called ‘cookies’) when you surf online, then they generate analysis through learning model and come up with you really want in past 7 days or a month. Finally, they will make recommendation for you; these recommendations usually show up on digit ads and emails. The same thing you also experience in Netflix channels and Amazon. These recommendations does not come up random!

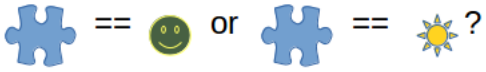
5. Fraud Dection
Machine learning is getting better and better at spotting potential cases of fraud across many different fields. PayPal, for example, is using machine learning to fight money laundering. The company has tools that compare millions of transactions and can precisely distinguish between legitimate and fraudulent transactions between buyers and sellers.[3]

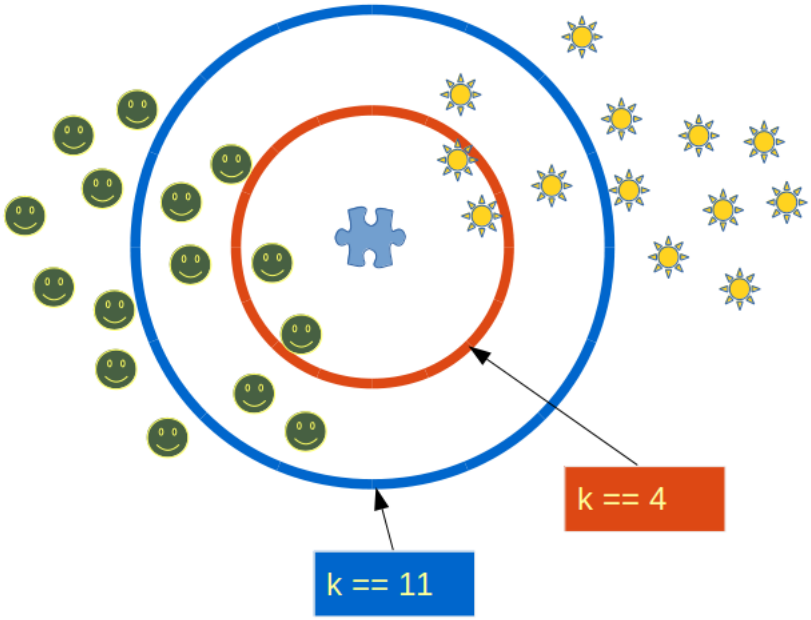
6. Onine Search
Machine learning is the core of all search engine, it is what makes you search experience so efficient. Every time you search something and click something on the search engine, the data will be collected and use to improve future search result. For example, when you search " renting an apartment in Berkeley" on Google, a list of result will show, and you click the fifth result. Then, the Machine learning model will store these data, if millions of people search the same thing and they all click the fifth result, the fifth result is likely to move up.

7. Smart Car
IBM recently surveyed top auto executives, and 74% expected that we would see smart cars on the road by 2025. A smart car would not only integrate into the Internet of Things, but also learn about its owner and its environment. It might adjust the internal settings - temperature, audio, seat position, etc. - automatically based on the driver, report and even fix problems itself, drive itself, and offer real time advice about traffic and road conditions.[4].

Machine Learning in R

There are many complicated and useful functions and packages in R, but in this post I am going to focus on the simplest and the most intuitive function, *k-nearst neighbor*. K nearest neighbors is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions). [5]. Intuitively, it is just use the given data





Simple example