CHAPTER 1: AN INTRODUCTION TO ALGORITHMS AND MACHINE LEARNING WITH A REAL-WORLD EXAMPLE

In this chapter, we will explore how computers do tasks in the figurative sense. We will also explore how computers can learn to do tasks themselves. This chapter will be the basis for the next chapter where we introduce actual programming.

1.1 Algorithms

Have you ever wondered how computers are able to do what they do? **Computer programs** are the step by step instructions that complete a task. Still, this does not tell specify how the task is done. A **computer algorithm** specifies how a task in a computer program is done. A computer algorithm brings together various tasks and ideas using *AND, OR*, or *NOT*. Using these three phrases, algorithms can be extremely complex or short and simple.

These ideas can be applied to real world as well. For example, let's say you need transportation to school. School in this case is the program. How you get to school is the algorithm. There could be a bus algorithm, a carpool algorithm, a ride service algorithm, or a biking algorithm.

EXAMPLES

Bus algorithm

1. Go to the bus stop.
2. Get in the bus.
3. Get off bus near school.

Biking algorithm

1. Get on bike.
2. Bike to school.

Ride service algorithm

1. Open ride service app on phone.
2. Request a ride.
3. Wait for ride service to arrive.
4. Get into ride service and give them the address to school.

Carpool algorithm

1. Call a friend who is driving to school.
2. Ask for a ride to school.
3. Wait for friend to arrive.
4. Get into car with friend when he arrives.

All of the example algorithms perform the same task but do them a different way. So how do you choose which algorithm is best? Algorithms are rated based on *efficiency* and *correctness.* Efficiency could be measured by how much time it takes a computer to perform the task the algorithm specifies. Still, this does not take into account computer hardware. Instead, efficiency is typically based on a computer analysis of the algorithm called **asymptotic analysis**. Asymptotic analysis is an analysis of algorithms that does not take into account computer hardware but instead analyzes the program based on its own merits. Computer analysts do this by basing analysis off input size and time instead. Each algorithm is represented by a function, and the efficiency is based on the type of function (linear, exponential, logarithmic, etc.) each algorithm represents. Correctness is based off whether the program works or not.

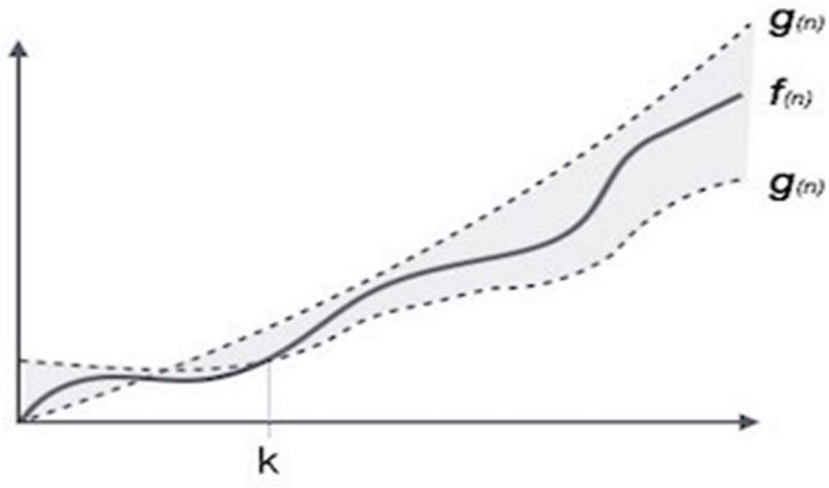

Figure 1: An example of asymptotic analysis.

How to Create an Algorithm for a Simple Task (Pizza!)

1. Identify the task to be completed.

   Example: *Eat dinner*

2. Specify how the task must be completed.

   Example: *Order a cheese pizza from Domino's for delivery to eat for dinner.*

3. Consider any conditions

   Example: *If Domino's is closed, go to Ledo's.*

4. Test the algorithm.

   Example: *Call Domino's and order the pizza*

5. Analyze the algorithm

   Examples: *Was Domino's open? Did someone answer? Was the pizza delivered? Was the pizza just cheese or did it have other toppings? How long did it take for the pizza to arrive?*

6. Make modifications to the algorithm then test again

   Example: *Order more pizza from a different place, using a different ordering method, or asking for different toppings. Then, compare to the previous experience.*

1.2 Machine Learning

While algorithms are impressive , they do have their problems. Ordinary algorithms can perform tasks, but still, someone needs to write the code that those tasks perform. Then, if you want to modify the task, someone will need to change the code again. That is where machine learning comes in. **Machine learning algorithms** are algorithms that rewrite themselves as a sort of improvisation that makes the algorithm perform better. These algorithms are not alive, but just programmed to edit themselves. Sometimes this can lead to code and results that the original programmer cannot even understand. Computers take former data given to them then make modifications for new data based on the analysis of the old data.

Machine learning has been applied to several industries so far, including applications in streaming services and self-driving cars. Machine learning should not be confused with artificial intelligence. Instead, machine learning is a subset of artificial intelligence. **Artificial intelligence** is creating an independent agent that will be able to do tasks based on its own volition. Machine learning is only teaching a computer to learn from its mistakes. A computer that is enable to machine learn is not an independent agent.

Machine learning is divided into three categories: supervised learning, unsupervised learning, and reinforcement learning. **Supervised learning** is where a computer is given input data which then creates a mapping function. This mapping function will be used to predict output variables. In this case, the input data is "training data" meaning we already know what the outputs should be, but we are just testing the accuracy of the computer. **Unsupervised learning** is where a computer is given input variables but output variables do not exist yet. There is no "training data". In this case, the computer is given more responsibility in that the output variables that the computer finds will be the basis of future data analysis. **Reinforcement learning** is like unsupervised learning. In reinforcement learning, there is no "training data" so there are no "correct" output variables. So, the learning agent performs multiple trials to maximize the "reward" or the best result.

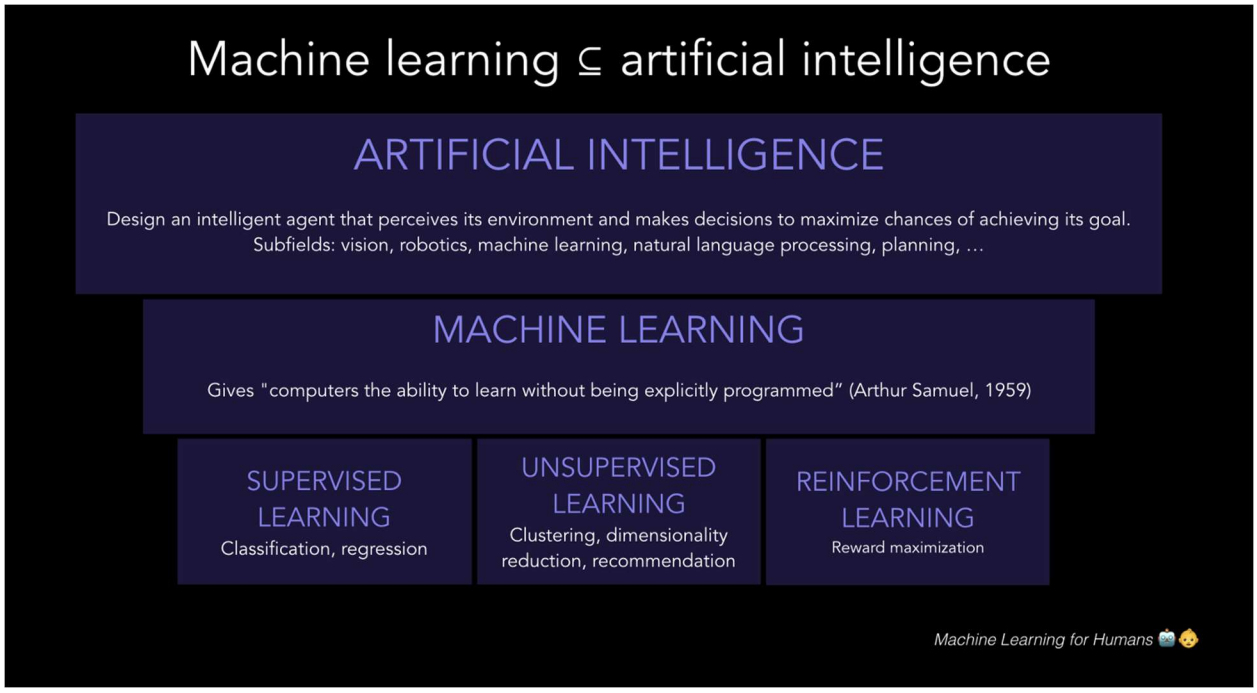

Figure 2: How machine learning and artificial intelligence relate to each other.

1.3 An Example: Netflix's Recommendation System

The streaming service *Netflix* is filled with thousands of television shows and movies. With several different types of movies and shows, it can become complicated for users to choose what to watch. Instead, Netflix uses a recommendation system that predicts what users may like. It is the basis of Netflix homepage. Netflix's homepage is several rows of categories filled with different movies and TV shows. This recommendation system utilizes several algorithms and machine learning methods previously mentioned. These algorithms are used to create each category and format the homepage. Netflix is given the data based on what you watch, how long you watch it, when you watch, what rating you give it, among other things. This data is combined to present a digestible recommendation list for the user. Let's look at the various algorithms that make up Netflix's recommendation system.

- Personalized video ranker (PVR) - This algorithm's task is to take each Netflix category or genre and order it for each user profile. The order is based on the user's previously watched content, so the order of each category is different for each user profile.

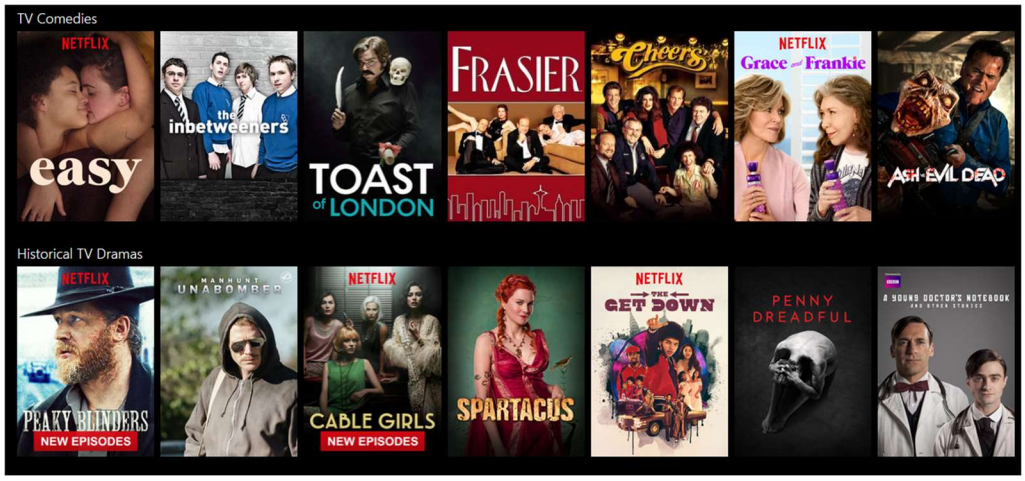

Figure 3: With PVR, content in Netflix categories change for each profile user. So, another user profile may not have any of these shows in their "TV Comedies" or "Historical TV Dramas" categories, or they may be in a different order.

- Top-N Video Ranker - This algorithm predicts the best recommendations for each user profile from the entire catalog. This algorithm is then presented in its own separate category.

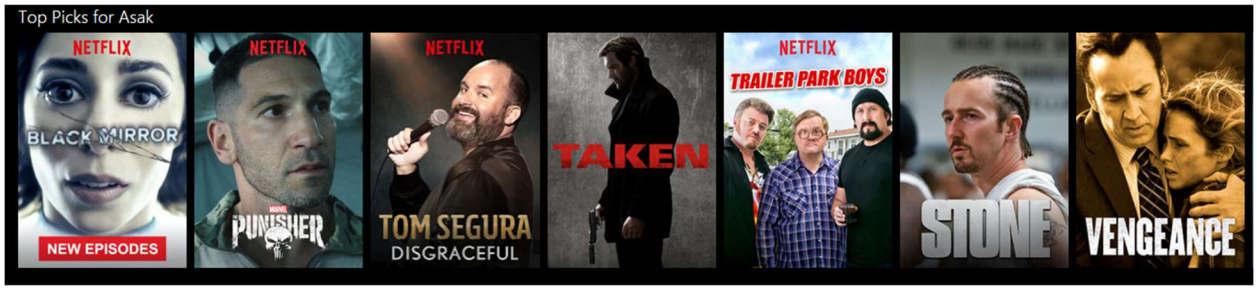

Figure 4: These titles are the best recommendations for this specific user “Asak”.

- Trending Now - This algorithm creates another separate category that chooses titles based on expected or unexpected events. For example, this category will display Christmas movies around Christmas time, or popular titles from a recently deceased actor/director.

- Continue watching - This algorithm takes content from each Netflix user profile that the user has not yet finished. It then organizes the content based on a prediction of whether or not that user will finish each individual title.

- Video-video similarity - This algorithm takes a past title you watched then displays similar content to that title. The content is not ranked in this case, but the past title that is picked is specific to the user.

Figure 5: This category is based on a specific show the user watched in the past. While these titles are not organized in any specific way, the title *Better Call Saul* was used specifically. This can be based on several viewing habits, such as frequency, recency, or the rating of *Better Call Saul.*

- Page generation - This algorithm takes each category from the previously mentioned algorithms then organizes them

From Netflix, you can see how algorithms are combined and used to create a whole program. Surprisingly, Netflix puts a lot of effort into organizing its titles for its users. Still, this is only possible based on machine learning. In the case of Netflix, this is organized based on unsupervised and reinforcement learning. There are no “correct” outputs in this case. Instead, the recommendation system predicts what users will watch next based on previous data. This data is later reinforced based no whether or not the user actually watches the titles given.

Summary

In this chapter, we learned about algorithms. We learned that algorithms are the basis for how computers complete tasks. We learned how algorithms are judged based on correctness and efficiency. We also learned about machine learning and how it relates to algorithms. We learned about the different types of machine learning. We then related these ideas to a familiar system to us all, Netflix's recommendation system.

Works Cited

Maini, Vishal. “Machine Learning for Humans, Part 5: Reinforcement Learning.” *Medium*, Machine Learning for Humans, 19 Aug. 2017, medium.com/machine-learning-for-humans/reinforcement-learning-6eacf258b265.

“What is a computer algorithm?” *HowStuffWorks*, HowStuffWorks, 5 Sept. 2001, computer.howstuffworks.com/question717.htm.

Brogan, Jacob. “Defining Algorithms—a Conversational Explainer.” *Slate Magazine*, 2 Feb. 2016, www.slate.com/articles/technology/future_tense/2016/02/what_is_an_algorithm_an_explainer.html.

“Analysis of Algorithms | Set 1 (Asymptotic Analysis).” *GeeksforGeeks*, 9 Sept. 2016, www.geeksforgeeks.org/analysis-of-algorithms-set-1-asymptotic-analysis/.

“Machine Learning: What it is and why it matters.” *What it is and why it matters | SAS*, www.sas.com/en_us/insights/analytics/machine-learning.html.

“Supervised and Unsupervised Machine Learning Algorithms.” *Machine Learning Mastery*, 21 Sept. 2016, machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/.

Gomez-Uribe, Carlos, and Neil Hunt. “The Netflix Recommender System: Algorithms, Business Value, and Innovation.” *ACM*, Netflix, Inc., 2015, delivery.acm.org/10.1145/2850000/2843948/a13-gomez-uribe.pdf?ip=108.22.201.224&id=2843948&acc=OA&key=4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35%2EE5B8A747884E71D5&CFID=852013797&CFTOKEN=93735469&__acm__=1516218225_5cd4e45830e209ca9eafae84e6383e22.

Maini, Vishal. “A Beginner's Guide to AI/ML – Machine Learning for Humans – Medium.” *Medium*, Machine Learning for Humans, 19 Aug. 2017, medium.com/machine-learning-for-humans/why-machine-learning-matters-6164faf1df12.