TWO POPULAR SERVICES THAT USE MACHINE LEARNING ALGORITHMS

Without realising it, we are confronted daily with applications that contain very diverse algorithms. But what exactly is an algorithm? An **algorithm** consists of a series of instructions that, when executed correctly, can solve a problem or complete a task. An algorithm is therefore often described as a roadmap. When an algorithm is executed by a computer, we speak of a **computer algorithm**. **AI algorithms** refer to the group of algorithms used within the field of artificial intelligence (AI). When algorithms learn a model based on data, we speak of **machine learning algorithms**.

When you read or hear about algorithms, you often hear the words model and technique. A **model** describes the knowledge and information acquired by the algorithm that is used to make decisions and/or predictions. **Machine learning technique** refers to the many strategies used within the field of machine learning. An example of such a technique is linking the preferences of listeners to the preferences of other listeners who are listening to similar music. By using this technique, a machine learning algorithm can present the listener recommendations.

Now all terms have been explained to avoid confusion, this brAInfood will take a closer look at machine learning algorithms using two concrete examples: Netflix and navigation applications. Are you curious how these two services use machine learning algorithms and why? Then continue reading.

Knowledge Centre Data & Society (2021). Two popular algorithms for machine learning. brAInfood of the Knowledge Centre Data & Society. Brussels: Knowledge Centre Data & Society.

brAInfood of the Knowledge Centre
Data & Society

RECOMMENDER SYSTEM

With Netflix, you can watch whatever you want from a large database of movies and series at any time. The longer it takes to make a choice, the more likely you are to switch off Netflix. So it is crucial for Netflix that you can make the right choice quickly. To help you choose between the many options, Netflix uses a **'hybrid recommender system'**, a combination of two types of algorithms that gives you recommendations on which movie or series to stream:

- **'Collaborative filtering'**: with this algorithm, users are grouped based on similar behaviour and the recommendations are created based on the characteristics of these user groups.
- **'Content-based':** this algorithm uses the information of an item to recommend similar items to a specific user.

DATA

To create recommendations, Netflix uses the following **data**:

- **All interactions you have with the platform** (e.g. viewing history, ratings of videos, what devices you watch on, what recommendations you ignore);
- **Data of other users with similar interests and tastes**;
- **Information about the videos** (e.g. titles, actors, genre, year of release);
- **No demographic information** of its customers (e.g. age, gender).

PERSONALISATION

The homepage of the platform consists of a large matrix with multiple rows, each containing different movies/ series. Each row is **personalised** in three ways:

- **The themes** of the rows (e.g. 'romantic comedies');
- **Movies/series** within that theme that could be **of interest to you**;
- **The order** in which you see the movies/ series.

Below, we describe some **common rows and their algorithms**.

NETFLIX

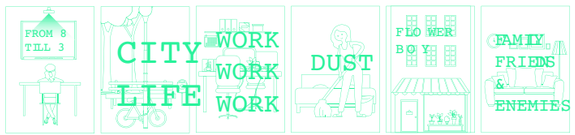
POPULAR ON NETFLIX

The **Personalized Video Ranker** (PVR) ensures that within a subset (e.g. **a genre**) the order of the recommendations are personally relevant to you. For example, you and I might see different titles in the **'Popular on Netflix' row** because the PVR arranges the titles based on what suits each viewer.

TRENDING NOW

The **Trending ranker** identifies two types of trends: recurring (e.g. more romantic movies around Valentine's Day) and occasional trends (e.g. more nature documentaries after a natural disaster). Together with personalisation based on your viewing history, these trends are presented to you in the '**Trending Now' row**.

CONTINUE WATCHING

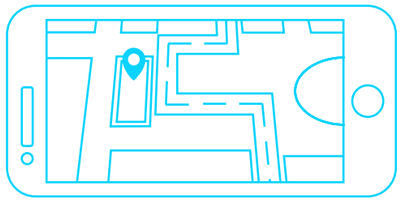


The **Continue Watching ranker** sorts the movies/series that you have recently watched based on your expectation of whether you want to continue watching this movie/series or not. It does this by looking at when you turned off the movie/series, and whether you have watched other titles since then and on which devices.

THEME ROWS

The **Page Generation algorithm** determines which theme rows appear on your start page so that they are relevant and diverse to you. For example, if you watch different types of movies/series at different times of the day or if multiple people use the same account, the algorithm takes this into account.

NAVIGATION APPS



Navigation applications (e.g. Google Maps, Apple Maps, Waze) help millions of people to get from point A to point B. Navigation apps must not only guide their users to their destination, they must also be able to **calculate the quickest route** or a **route tailored to the user's means of transport or other preferences**.

DATA

In order to offer a range of routes, navigation apps use **data** from various sources, such as:

- **Live traffic information** from users' devices;
- **Historical traffic data**;
- **Information on speed limits and roadworks**;
- **Information on the type, shape and quality of the road**.

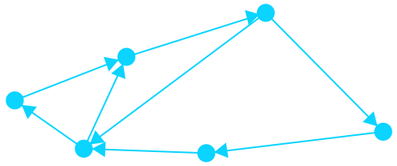
For this last group of data, Google Maps uses information that it collects with Google Street View or from satellite data. These data are processed using various machine learning models to make predictions about traffic flows.

DIJKSTRA'S ALGORITHM

The application searches for the shortest time path. To make this calculation, very advanced algorithms are used, but many of these are based on **Dijkstra's concept**.

Dijkstra's algorithm, developed by Dutch computer scientist Edsger Dijkstra in 1959, cleverly determines the **shortest path from a given starting point to each node of a network**.

A GRAPH



The Dijkstra algorithm cannot be applied to a map directly, because it contains too much information. It will therefore work with **a representation of the map that consists of nodes and connections between these nodes**. This representation is called a graph in mathematics.

There are all kinds of specific types of graphs possible, often depending on the properties of the connections. For example, the connections can be directional (e.g. one-way); in this case we speak of a **directed graph**. The connections can also be non-directional; this is called an **undirected graph**.

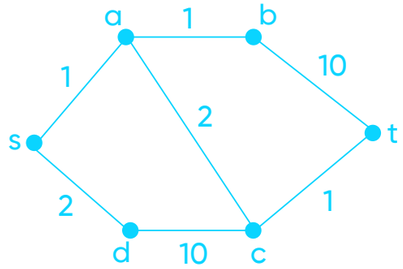
NETWORK VALUE

Imagine that your current position is a node, just like your final destination. Between those two different positions there are several other nodes, which are connected to each other. In the graph, **each connection has a value**. This value represents a 'distance' or a 'cost', which stands for the effort required to get to the other node. In navigation applications, for example, this may include **the actual distance or the amount of traffic**.

GREEDY METHOD

The Dijkstra algorithm works with a 'greedy method'. This method goes **step by step through all the nodes in the network**. To determine the next step each time, the algorithm will look for the connection with another node that immediately gives the greatest benefit. This often does not lead directly to an optimal final solution, but eventually the algorithm will propose the **global optimal solution from starting point to final destination**.

EXAMPLE



In this simple representation, the algorithm will first go from 's' to 'a', and then from 'a' to 'b'. However, the shortest path will be the sequence 's-a-c-t'.

Imagine that a navigation app needs to calculate a route between Brussels (starting point) and Antwerp (destination). If the application were to work purely with the Dijkstra method, the algorithm would start searching in all directions around Brussels. That would be an inefficient effort and would cost too much time and computing power.
Navigation apps therefore work with **more advanced algorithms**, based on the concept of Dijkstra, allowing for a **specific search in the direction of the destination**.

Sources Netflix

Hunt, N., & Gomez-Uribe, C. A. (2016). The Netflix Recommender System: Algorithms Business Value and Innovation. journal, ACM Transactions on Management Information Systems (TMIS), 6(4). [paper]

Netflix. Hoe het aanbevelingssysteem van Netflix werkt. [website]

Goltsman, K. (2017). Under the Hood of Netflix Recommender System. [blogpost of DataScience Foundation]

Sources navigation applications

Fitro, A., P Sulistio Ilham, A., B Saeful, O., & Frendianata, I. (2018). Shortest Path Finding in Geographical Information Systems using Node Combination and Dijkstra Algorithm. International Journal of Mechanical Engineering and Technology 9(2), 2018. pp. 755–760. [paper]

Lanning, D. R., Harrell, G. K., & Wang, J. (2014). Dijkstra's algorithm and Google maps. Proceedings of the 2014 ACM Southeast Regional Conference on - ACM SE '14, 1. [paper]

Lau, J. (2020). Google Maps 101: How AI helps predict traffic and determine routes. [blogpost of Google]

Byrne, M. (2015). The Simple, Elegant Algorithm hat Makes Google Maps Possible. [blogpost of vice]

Velden, L. (Department of Mathematics, Technical University of Munich) (2014). Der Dijkstra-Algorithmus. The classic among shortest path algorithms. [website]

Bitesize Data Science (2019). Navigation Apps: The Algorithms that Get You from Here to There. [blogpost of Bitesize Data Science]