Please let me introduce you to one of the most prominent buzzwords in the field of machine learning today,

Recommender Systems (Heading)

So what is it? Basically it is the one responsible for:

* Filling your Yout\*be feed with videos they think you would like
* Making Bukalap\*k, Tokp\*d, and your other favorite e-commerces gobble you up with things to buy, by telling you “others also bought this”, or “viewed by others”.
* Filling your Netfl\*x feed so that you keep procrastinating everything and just marathon your favorite series
* And many, many more. It’s basically almost everywhere

So why is it so important?

With so many datas and product lining today, companies need to serve their market targets smartly. They can’t just let customers traverse their extensive catalogue and lose money on the first 10 seconds. So they need to create an automated salesman which can accurately know what a customer really needs.

How is it really done? (Or in geeky language: “How do you model it?”)

Basicallly there are two ways you can build the algorithm:

1. Collaborative Filtering – (User Centered)
2. Content-Based Recommendation – (Item Centered)

Okay to understand the difference between those two, let’s say we are going to build a recommender system for books, and of course, our goal is to build a system which can recommend which book a certain user will like, but hasn’t read.

1. Collaborative filtering

This algorithm can be called user-centered because it is based on grouping users with similar behavior.

For example, there are two users: Jona and Joni.

* Jona has read and likes book A, B, and C.
* Joni has read and likes book B, C, and D.

Because both likes mostly the same books, B and C, the system can recommend books which Joni likes to Jona, vice versa. Meaning, Jona can be recommended book D, and Joni can be recommended book A.

1. Content-Based Recommendation

This algorithm can be called item centered because it is based on grouping items with similar attributes.

Let’s say there are three books. The Magic, The Hero, and The Love, and for this algorithm we need specific attributes that the books need to have. Let’s say we lay it like this

|  |  |  |  |
| --- | --- | --- | --- |
|  | Romance Point | Action Point | Fantasy Point |
| The Love | 0,9 | 0,05 | 0,6 |
| The Hero | 0,1 | 0,9 | 0,75 |
| The Magic | 0,7 | 0,1 | 0,8 |

We can see that in the table, “The Magic” and “The Love” are quite similar because its attributes are distributed in pretty much the same way: high in romance and fantasy, but low in action. Also, “The Hero” is the one pretty different from the others. So we can say that given someone likes “The Love”, he would be recommended to “The Magic” way more than to “The Hero”.

This is a very rough introduction, and very sorry to say that I still couldn’t include the math, but I very much hope that you understand, so that you can at least show off to your friends while watching the next recommendation at your youtube feed.