

## How (I think) Netflix gathers evidence while you browse

By Kim Falk

In this article, excerpted from my book [Practical Recommender Systems](#), I explain the concept of “evidence” using Netflix.

In Ron Zacharski’s *A Programmers Guide to Data Mining* there is a great example that illustrates the difference between ratings and evidence. He shows the explicit rating of a guy called Jim. Jim states he is a vegan, and enjoy French films, but in Jim’s pocket is a rental receipt for Marvel’s *The Avengers* and one for 12-pack of Pabst Blue Ribbon Beer. Which should you use to recommend things? We think it is an easy choice.

There is no substitute for good evidence. The best recommender is only as good as the evidence provided.

Everything on Netflix’s landing page is personalized, the row headlines as well as their content, with the exception of the banner, which is a campaign that everybody sees (Emperically tested by asking two others). The rows are different for each user; the headlines will range from familiar categories like "Comedies" and "Dramas" to highly tailored slices such as "Imaginative Time Travel Movies from the 1980s". Figure 1 shows my personalized Netflix’s page. On my page, the first couple of rows contain recently added content, suggestions and popular content.

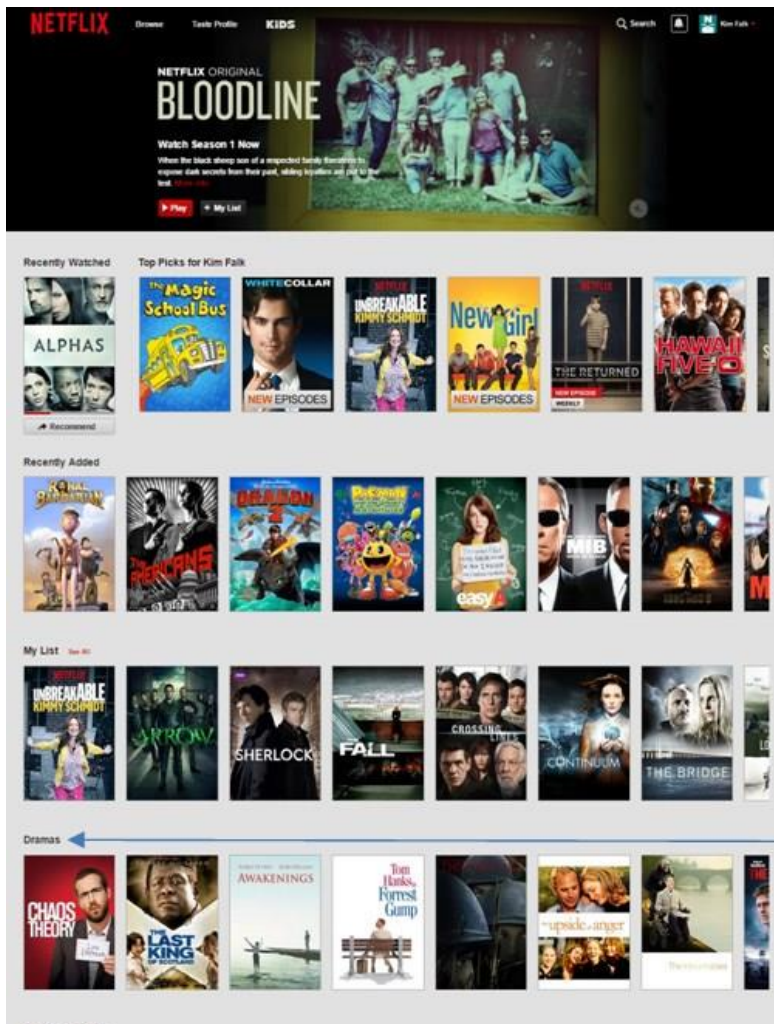


Figure 1: My personal Netflix front page

The first personalized row title is Drama. This means that between the genres or row headlines, they think that drama is the one that matches my interests more. The Drama row contains a list of content that Netflix considers interesting to me within category of Drama. Figure 2 illustrates what part of Netflix content is shown in my drama list.

Hovering the mouse over the row, I can make it scroll sideways, showing other content in the drama genre. If I scroll over the content of the drama row, what does this say about me?

For source code, sample chapters, the Online Author Forum, and other resources, go to <http://www.manning.com/falk/>

It could mean that I am writing an article, and I am trying to make screenshots, but most likely, it suggests that I enjoy dramas, and I want to investigate the list further.

If I see a movie that looks interesting, I can place my mouse over it and I'm shown its details. Moreover, if those details are intriguing I might click to go to the page of the film, and if it still sounds good I will either add it to "my list" or start watching it.

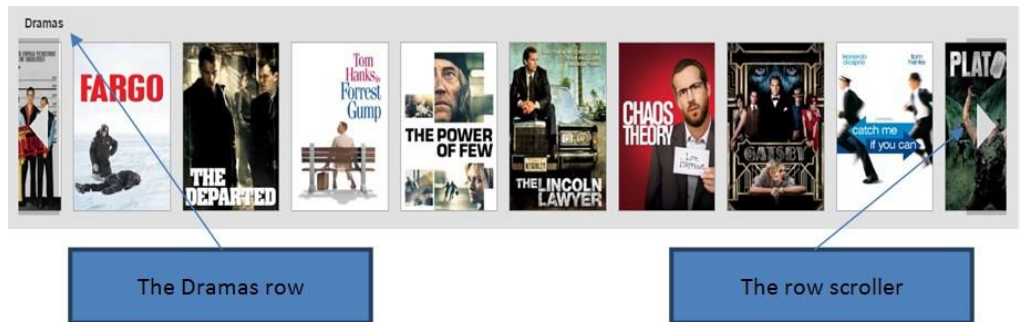


Figure 2: Netflix dramas row

For an ecommerce site, the act of watching a film is equivalent to buying something. But for a streaming site, this is not where it ends. If a visitor starts watching a film, it is a positive event, but if he stops after three minutes and never restarts it again, then it shows that the visitor did not like what he saw. However, if he restarts the movie again later or within some specific timeframe, it might actually mean that he liked it even more than if he had watched it in the first go.

Often the purpose of ecommerce sites is to make people buy products, even if the product the customer is actually buying might not be exactly what he wants. That depends on the site's affiliation to the product. If Amazon sells you a bad t-shirt, you would say that the brand of the t-shirt is bad, and go back to Amazon and buy another t-shirt. However if you buy a t-shirt from a Gap website and it is bad quality, then you will probably go to another site. That is a bit different from subscription-based sites. Mofibo, a Danish e-book streaming service, provides recommendations as inspiration and discovery, but with a catch: it is important to Mofibo that the reader knows what kind of book it is before starting to read it. Because Mofibo pay a fee every time the reader opens a book, not per page, but per book, that means that while Mofibo wants you to read as much as possible; they want to minimize the amount of books you do it in.

### ***The evidence Netflix collects***

If we try to imagine what goes on behind the curtains at Netflix, and try to imagine what data they collect, we'll find that often the purpose is not what it seems.

Let's imagine it is Saturday night at Jimmie's place. Jimmie has user id = 1234, and after nuking some popcorn, he opens Netflix and does the following:

- Scrolls the Drama(id: 2) row.
- Hovers the mouse over a movie (id: 41335) to get details.
- Clicks to get more details on movie (id: 41335)
- Starts watching movie (id:41335)

While he watches the movie, we will go try to imagine what just happened on the Netflix server.

Table 1 shows some of the events that could be collected from a user doing what we described above, along with interpretations of what it could mean. Moreover, a column containing an event name, is added to connect this table to the log described further down.

**Table 2.1: Examples of evidence from Netflix**

<i>Event</i>	<i>Meaning</i>	<i>Event Name</i>
Scrolling a themed row	User is interested in the theme; here "dramas"	genreView
Placing the mouse over a film (requesting overview of content)	User is interested in the movie (which is a drama and therefore also shows interest in this category)	Details
Clicking on the film (requesting details of content)	User is more interested in the movie	moreDetails
Adding the film to his list	User intends to watch the movie	AddToList
Starting to watch the film	User "purchases" the movie	PlayStart

As the table illustrates, all these events are evidence to the system, since they uncover interests of the users. The following table shows how the evidence is probably recorded from a user who did the following steps:

For source code, sample chapters, the Online Author Forum, and other resources, go to <http://www.manning.com/falk/>

**Table 2: Example of how Netflix probably logs evidence.**

<i>userId</i>	<i>contentId</i>	<i>Event</i>	
1234	2	genreView	2015-06-07 20:01:00
1234	41335	Details	2015-06-07 20:02:21
1234	41335	FutherDetail	2015-06-07 20:02:30
1234	41335	PlayStart	2015-06-07 20:03:01

There will probably be a long list of other columns such as device type, location, speed, and weather, which can all be used to better understand the user's context. Actually, I would also venture that the number of log events for this scenario would be much more, but let's keep the example simple. The list of events types are probably much longer as well.

Now that we have a general idea about what the evidence is, we can start looking into the implementation of an evidence collector. An evidence collector is used to collect data like the one we looked at above. Just to ensure that you are not thinking that this can only work with media streaming sites, let's look at another scenario.

In a previous work, I had a colleague who would spend all his breaks surfing the internet for garden tractors. Actually, he could spend time on anything that ran on gasoline that you use in a garden. Let's imagine this former colleague had a small break and he opened up his favorite (imaginary) site call, super power garden tools.

He would do the following:

- Select the Garden tractor category
- Click on a green monster that can pull trees up
- Click on specifications to see how big trees it can hull up
- Buy green monster.

These events would be the same as the others, only the importance of buying an expensive product might be more significant than starting to watch the movie, but I hope you get the picture.