



SDS PODCAST

EPISODE 220

FIVE MINUTE

FRIDAY:

DATA SCIENCE

IN RETAIL



This is FiveMinuteFriday episode #220, Data Science in Retail.

Welcome back to SuperDataScience Podcast, ladies and gentlemen, super excited to have you on the show. And today, we're continuing our series of FiveMinuteFriday episodes dedicated to data science in various industries. And today, we're talking about Data Science in Retail.

Now, retail is a massive industry which encompasses any kind of sales that occur when a business sells a product or service to an individual consumer for his or her own use. As you can imagine, this is an immense industry. In fact, according to different and various sources, the retail sector in 2018 alone accounted for about 25 trillion ... That's trillion with a T ... US dollars in sales. And the sector represents about 31% of the world's GDP and employs literally billions of people throughout the globe.

Even just from this information, we can extrapolate that data science has huge applications and potential in retail, potential to disrupt and enhance the various areas of this industry. In fact, according to IBM, 62% of retailers say the use of big data techniques gives them a serious competitive advantage, knowing what your customer wants and when.

Now, in this podcast, we cannot physically possibly go through all of the different applications of data science in retail. That would just require probably days and days of time. But at the same time, we can and we will dive into the top five applications of data science in retail that we have identified for you, guys. I hope you're excited and let's dive straight into it.

Application of data science in retail #1, recommendations. We've all heard of examples of recommendation engines from companies such as Amazon, Netflix and even Google where companies use recommendation engines to help customers explore different products and services that might be useful and beneficial to them.

Now, recommendation engines commonly use one of two techniques, either collaborative or content-based filtering. Now, let's start with content-based



filtering. Content-based filtering is when a recommendation engine knows your prior experiences and prior purchases. And based on that and with different similarities of products, it will recommend some other product to you.

On the other hand, collaborative filtering is when a recommendation engine explores the different behaviors of other users in the platform. And based on their interaction of products, it can recommend something completely different to you. Based on the similarities that you have with other users rather than similarities between products, it's looking at the similarities between users. And then based on other user's past experiences, it can recommend something very different to you which you might still find relevant.

And that's why when you watch Netflix, for example, and you finished watching a movie or a TV show, you get a recommendation for a TV show, a movie that you've never ever even considered before. At the same time, you might actually find it very, very interesting, and you might be into it. That's just an example of a recommendation engine in action.

In terms of statistics, it is estimated that more than 35% of all Amazon sales are generated by their referral engine. That is a massive amount. Basically if Amazon didn't have their referral engine, they would be short on that 35% of their sales. So, it's a massive value add for the Amazon retail business.

Next, application of data science is price optimization. Now, price optimization is extremely crucial now that we live in a hyper-connected global world where you can purchase anything from absolutely any part of the world. The competition is nowadays not just local like before where the consumers were like a captive audience in their city or neighborhood or even country. There are no borders anymore. Now, you can purchase anything from anywhere, so competition is all over the world and



companies have to constantly monitor that they are providing a fair price to their consumers. Otherwise, they'll lose their consumers.

Competitor price data can be collected electronically using algorithms from competitor's websites and so on. And that data plus other information can be combined to provide the best pricing to the consumers at the time. For instance, we all know that airlines vary their cost of travel according to peak season, days of the week, time of day.

And more, like an example as probably closer to many us, is the Uber surge pricing. We know that based on the supply-demand ratio, Uber will change its pricing and they actually have an app of surge pricing. That's another example of data science in action where it allows for companies to maximize their revenue but also for consumers to get a fair price and for the markets to balance out.

Use case of data science in retail #3: Targeted advertising. We're venturing into a world that is more and more data driven, that is more and more customized where everything, all the services and experiences and products are becoming more and more customized to people, because people want to buy things that are relevant to them. People want to invest into services and have experiences that are relevant to them.

The more relevant a company can make their experiences to an individual, the more likely they are going to succeed in making a sale and building a relationship with that person. From that notion, many companies actually propel and power advertising with different data sources that are available to them starting from social networking, to previous searches, to browsing history, behavioral social status, preference and so on.

And here's a distinct example. Already in 2011, Cablevision, an Argentinian based television company, began using demographic data to show different ads to different people watching the same program. As you can imagine, very, very tailored customized experience. Of course, there are ethical considerations which are being discussed by governments and should be



taken into consideration by companies. But in general, that's where the world has headed. What data sources are legally available to companies, companies will use in order to provide tailored experiences to their customers and that is where they will need data science in order to facilitate that.

Number four, predicting trends. Trends are very important because there are a lot of companies to forecast and maintain their inventory, understand how their cash flow is going to change, how like when they need more people working and delivering products and services when they need less. It's basically knowing the future of demand or supply is important for companies in order to be able to manage their operations effectively and make the most for themselves and for consumers as well.

We've got traditional trends such as seasonal trends. And data science, of course, helps there. Here's an example which we saw in the Advanced Tableau Course in SuperDataScience where February sales in retail, like just across the whole retail ... We're looking specifically at the Australian retail section, but this is true for the whole world. The February sales in retail almost always experience a drop.

And that is just like a seasonal trend or seasonal pattern because people use their credit cards to buy lots of Christmas presents and perform lots of spending in December. And February is the time of the year when they have to repay those credit card loans. And that's why sales in retail slow down in February. Knowing that can empower businesses, and data science can help spot patterns and trends like that.

And those are more individual transfer products and services and different kind of events which can be extracted through social media. For instance, natural language processing can be used to extract information from platforms such as Twitter. And machine learning can make sense of it and give businesses edge and can advance your competition.



A specific example here is that Nordstrom Marketing team which follows Pinterest, Instagram and Twitter networks to identify the most popular products. They then use this data to promote good products in their physical stores. As you can imagine, it's much faster to do all of this with machines with natural language processing than with people. It's much more cost effective and will yield much better results in the long run. And that's where data science comes in again.

And finally moving on to use case #5, inventory management and anticipatory shipping. Now, inventory management, we've already touched on upon that when we're talking about trends. Inventory management is an important question for retail businesses and probably for most retail businesses. Starting from your food chains where inventory will go bad and is perishable, even to technology companies where inventory just gets outdated over time.

A great example here of how companies are very focused on their inventory is Dell, one of my favorite examples and this is why. Dell assembles nearly 80,000 computers every single day. Every 24 hours, they assemble about 80,000 computers. And the most interesting part is that none of those computers stays with Dell for longer than 72 hours. Dell sells every single computer, every single piece of equipment that it assembles within 72 hours. It's out of their, not just out of their factory, gets out of their factory within two hours.

But within 72 hours, it's out of the company. And that is one of the main focuses their CEO, Kevin Rollins, actually talks about this, that it is extremely important to turn this inventory over very fast because otherwise it loses value, depreciates, and it becomes outdated in the world of technology. But as you can imagine, that happens worldwide more or less, in any area of the retail industry.

It's very important to manage inventory. Computer technology and data science can really help with that where it facilitate this process. Here's a



very famous example where Amazon has a patented technology of anticipatory shipping. Basically what happens is Amazon will ship a product closer to you to a warehouse that's already closer to you even before you've purchased it. How crazy is that? That is like data science in action.

Instead of waiting for you to buy a certain product and then you will get it in two or three days, or even in one day from their main warehouse somewhere in your region, what they'll do is they will predict that you will buy that product. And even before you buy it, they will ship that product from their main warehouse to a warehouse near you.

And when you do buy it which is already predicted, they obviously don't predict a 100% certain but looks like a high likelihood. There's a high likelihood that you're interested in this. They will recommended it to you and you'll probably buy it. That product will actually arrive to you within an hour or something like that. That is called anticipatory shipping.

Very interesting, I think, SNL ... My brother told me about this. SNL had a funny sketch about that, about Amazon. It was how Amazon will ship things to you even before you know you want them. That's of course taken too extreme, but I think it might be fun to check it out. But the point stands. Data science can really help with inventory management and things like anticipatory shipping.

There we go. That's five use cases of data science in retail. There's plenty, plenty more. If you're in the space of retail, I highly encourage you checking them out. If you're not in the space of retail, I hope these examples were helpful and inspiring for whatever work you are doing with data science. On that note, thank you so much for being here. I look forward to seeing you back here next time. And until then, happy analyzing.