Week1 Assignment

Collection and analysis of a Data for Amazon

Name: Sravanthi Mummadi

Instructor: Yevgeniy Frenkel

**Pick a company or product that you interact with regularly.**

I would choose an E-commerce company Amazon. I interact with this company at least twice a

week on an average.

**What kinds of data does this company collect?**

Amazon is a leader in collecting data. It collects data like products data like product options, product categories,

product description orders data like order details, options, option groups

customers data like what product does a customer viewed, when did they viewed, what did they

Purchased, shipping address, phone number, how many days did he searched the product

Review information like how many days did he took to write the review,

Prime member information

Website login information like mobile/mobile app/mobile browser/laptop/tablet.

**How is this data used (or possibly used)?**

**Personalized Recommendation System**

It uses data purchase and customer data. It analyzes what items customer purchased previously and what is in customer online shopping cart or on your wish list, which products you reviewed and rated and what items you search for most. This information is used to recommend additional products that other customers purchased when buying those same items.

For example, if a customer added a DVD to your online shopping cart, similar movies purchased by other customers are also recommended for you to purchase. In this way, Amazon uses the power of suggestion to encourage you to buy on impulse as a means of further satisfying your shopping experience and spending more money. This method generates 35% of the company’s sales annually.

**Anticipatory Shipping Model**

Amazon’s anticipatory shipping model uses big data for predicting the products customer are likely to purchase, when customer may buy them and where you might need the products. The items are sent to a local distribution center or warehouse so they will be ready for shipping once you order them. Amazon uses predictive analytics to increase its product sales and profit margins while decreasing its delivery time and overall expenses.

**Supply Chain Optimization**

Because Amazon wants to fulfill your orders quickly, the company links with manufacturers and tracks their inventory. Amazon uses big data systems for choosing the warehouse closest to the vendor and/or customer, the customer, to reduce shipping costs by 10 to 40%. Additionally, graph theory helps decide the best delivery schedule, route and product groupings to further reduce shipping expenses.

**Price Optimization**

Big data is also used for managing Amazon’s prices to attract more customers and increase profits by an average of 25% annually. Prices are set according to customer activity on the website, competitors’ pricing, product availability, item preferences, order history, expected profit margin and other factors. Product prices typically change every 10 minutes as big data is updated and analyzed. As a result, Amazon typically offers discounts on best-selling items and earns larger profits on less-popular items.

For example, the cost of a novel on the *New York Times Best Sellers* list may be 25% less than the retail price, while a novel not on the list costs 10% more than the same book sold by a competitor.

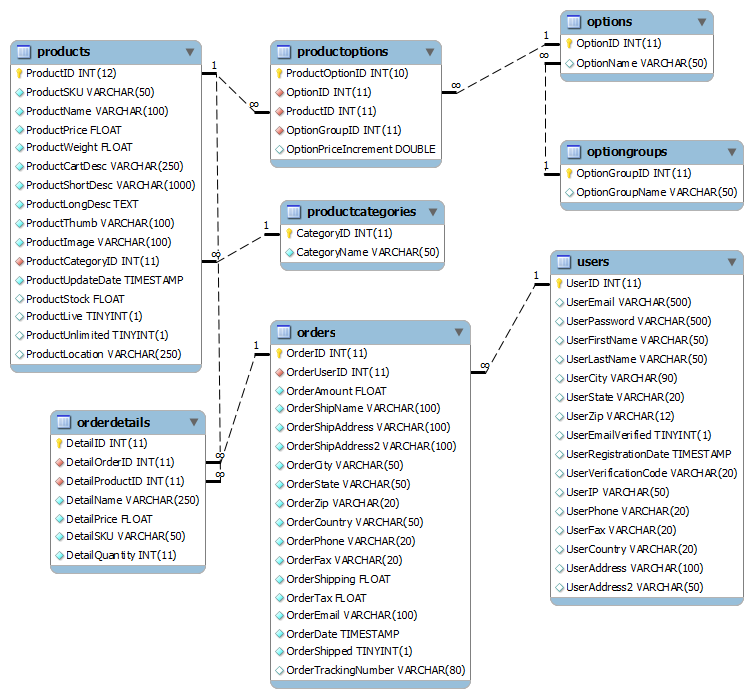
**Amazon Web Services**

Through Amazon Web Services (AWS), Amazon’s cloud computing service introduced in 2006, companies can create scalable big data applications and secure them without using hardware or maintaining infrastructure. Big data applications like clickstream analytics, data warehousing, recommendation engines, fraud detection, event-driven ETL and Internet-of-Things (IoT) processing are through cloud-based computing. Companies may benefit from Amazon Web Services by using them to analyze customer demographics, spending habits and other pertinent information to more effectively cross-sell company products in ways like Amazon. In other words, these retailers can use Amazon to stalk customers, as well.

**What features are likely powered by this data?**

The company collects many variables like orders, products, options, users, order details, etc. recorded in the E-commerce business for future analysis and predictions. In each section, there are many variables, but the most important table is the product table. because it relates to all other tables in the database and related fields.

**Schema diagram for E-commerce Database**



**Given above, suggest a new data-driven feature or service for the company**

I don’t think amazon is not leaving any data. It is using everything to make more money. I read an article, it says some people are ordering items from amazon and they are claiming that they did not get that item, some of them are returning products after using some days.

Though it could not be a biggest concern, they should make item delivery to customer directly, if the price of the item more than $500 something like this make that delivered photo as extra feature and store it in the database. If at all customer claims they would know it is a fraud or not by checking photo feature. I think This could help the business for fraud detection.

Recommendation system used by amazon sometimes shows unnecessary things as well. I think we can make it by collecting more useful customer information.

In one Prime account more 5-6 people will order with different payment method and different address. It could be customer friends or family. If amazon restrict it some limit number of address and users would increase the chance of taking more connections. As amazon has millions of customers. In that case there would be chance of taking more prime memberships like at 2 in 10.it would increase the number of memberships.

**What kind of data would you need for this new feature?**

For the first one I would collect delivered picture from mail service.

And for second one I would collect customer information like age

**What type of machine learning algorithm would you use?**

I would use facial detection algorithm for the first one and for second one clustering based on age groups. This might give the relevant results.

**Are there any privacy considerations?**

We had one professor (Applied Data Scientist) from amazon, according to him, amazon is not using customers private data other than address and phone number. I think there are no privacy considerations.

**Discuss monetization strategy**

Website monetization is the process of converting existing traffic being sent to a website into revenue. The most popular ways of monetizing a website are by implementing pay per click (PPC) and cost per impression (CPI/CPM) advertising.

Nowadays, monetization is very common, many people are creating websites by their own. it is very easy to create at website also, just we have to register the website from godaddy.com. “WIX” is providing many options for modify and enhance our website. it is also one kind of business strategy to improve sales and profits. After creating website, we can add as many items from amazon as we want. We have to market our website by posting and sharing in social media apps. If somebody bought something through these websites, they get some commission(dollars) for each transaction Amazon encouraging monetization strategy.

**References:**

<https://www.quora.com/What-is-a-monetization-strategy>

<https://websitesetup.org/33-ways-to-monetize-website/>

<https://www.quora.com/What-kind-of-data-does-Amazon-collect-from-its-customers>