



# Assignment 3

EAS 504

Anish Shah



eBay is a multinational e-commerce corporation that facilitates consumer-to-consumer and business-to-consumer sales through its website.

Base Questions:

- What are the principal uses of data sciences in this domain?
  - A. Ranking Items, predicting fraud transactions, inventory management, coupons and rewards, SEO, shipping, machine translation, search analysis, pricing analysis, and left-hand navigation bar predictions are some of the principal uses of data sciences in this domain.
- How are data and computing related methods used in the organizational workflow?
  - A. The data is generally collected from the user behavior and buying patterns. This data is generally collected from user browsing cookies. As an e-commerce company the data comes in, in a structured format that users enter while exhibiting items in the market place. This data is generally stored in a database attached to each user in which his/hers browsing patterns/previous search queries are stored. As the data is huge, clusters of databases with low latency is used in e-commerce.
- What data science related skills and technologies are commonly used in this sector?
  - A. Skills such as neural networks, deep learning techniques such as Long Short-Term Memory (LSTM), A/B testing to assess the difference of product on purchase behavior, GRU models are employed to assess query understanding, statistics concepts such as significance tests are used in this sector. Technologies include data parallel processing tools such as Spark and native MapReduce jobs.
- What are the primary opportunities for growth?
  - A. Global E-commerce market is already worth approx. \$22 trillion and is growing rapidly. Electronics, Clothing apparel are the most common things that are bought online. But the market is still kind of unexplored in the domain of FMCG (Fast Moving Consumer Goods) for which people still trust their regular brick and mortar stores, this FMCG industry is biggest opportunity for growth as Forbes expects this individual industry to be worth \$25 billion till 2025 in United States. Encouraging more people and creating a loyal customer base is the key to make a strong foothold here.

Use of Recommender systems and Pricing models in E-commerce:

Mainly there are two kind of recommender systems used in e-commerce industry, time-based recommender systems and customer-based recommender systems.

- Time based recommender systems:  
Recommendation happen based on the time component. For example, recommending air-conditioner during summers or winter coats during winters.
- Customer based recommender systems:

These recommender systems focus more on learning from customer behavior and recommending products based on those learnings. Typical kinds of recommendations have a workflow in following pattern – ‘x% of people who bought product A also bought product B’. That is applying learnings from customer behavior.

Pricing Models:

Previously the models didn't have any interesting reasoning behind pricing. A product is priced based on production, operational costs, etc. with some margin. But now mostly dynamic pricing model is used where there is reasoning and has good impact on the number of sales and overall profit with maintaining loyal user base. We can use dynamic pricing models in order to optimize the inventory, where there are patterns of buying few products simultaneously and also in the case of giving discounts to products whose life period is at its end and we have large inventory of those items.

Case Studies:

Some of the popular case where data science is extensively used in e-commerce:

1. Ranking items:

- Once the query is given and relevant items are pulled from a database, it is extremely important to rank them appropriately since users in most cases tend to pick whatever they want to buy by first few items that are displayed. So in order to optimize the overall throughput or number of conversions from viewing to actually buying an item it is important to use Machine Learning algorithms to rank those items based on query to display in order.

2. Left-hand navigation bar predictions:

- Depending on the kind of query that was issued, the criteria of filtering products varies by query. Therefore, using machine learning to predict the probability of each filter for a query gives us relevant filters that can be applied for a problem. For example, if we query television, appropriate filters could be type, screen size etc. But not type of engine or fuel type (for cars).

Apart from the principal uses mentioned above there are other ML problems which could be solved using data science mentioned by the speaker:

Cold start problem, query understanding, image search, voice search, conventional search, handwritten text enabled search, item attribute extraction, title generation.

From this we can say that data science plays a huge role in the day-to-day working in this industry.