EAS 504 ASSIGNMENT-2

By: - Katyayni Shankar Kaushik

General Information: -

The following Lecture was held by Mr. Sharat Chikkerur, he currently holds the position of Principal Data Scientist, in Microsoft. He is an alumnus of University at Buffalo (2005).

Computational Advertising Online Advertising is a multi-billion-dollar industry and works around multiple domains computer science industry (as per the information provided to us through the session conductor), that includes information retrieval, machine learning, auction theory, optimization and distributed computing.

Base Questions: -

1.) What are principal uses of data sciences in this domain?

   The principal use of Data Science in Computational advertising (online advertising) domain can be categorized into multiple categories: -

   - Personalized content for each user.
   - Target Delivery of advertisements (Information Retrieval).
   - Value estimation of advertisements.
   - Dynamic Pricing
   - Recommendation Systems

2.) How are data and computing related methods used in the organizational workflow?
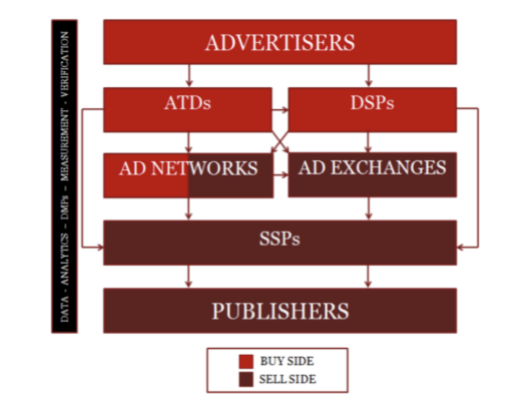
   The above graph can showcase the chain, starting from Advertisers and ending at Publishers. This can be termed as organizational workflow and data and computing related methods are used at each and every step. Following are the key points to be taken into consideration: -

   - Advertisers want to buy advertisements from Publishers (Google, Yahoo, Facebook), who, in turn, have access to public (Data).

- ATD'S, Demand Supply Platforms (DSP's), AD Networks, AD Exchanges, Supply Side Platforms act as middle points of the chain, they play a major role in the organizational workflow. They acquire user data from data aggregators and in turn forms a network and bid for the advertisements online. Thus, data and Machine Learning has a huge role to play in this area.
- Demand side platforms use machine learning the most, since this is where all the advertisements are looked for their value/ relevance, the most, so that they can have a huge value in turn.

3.) What data science related skills and technologies are commonly used in this sector?

The data science related skills and technologies commonly used in this sector are as follows: -
- Machine Learning techniques for value estimation
- Recommender Systems for personalized content
- Game Theory for dynamic pricing
- Information retrieval for targeted delivery of advertisements.
- Distributed Computing to reach out to Web Scale audience.
- Gradient Descent (Mathematical Techniques).
- Technologies are used are as follows: -
  - Vowpal Wobbit (ML library)
  - Generalised Linear Models (value estimation)
  - Latent Dirichlet Allocation (topic modeling)

4.) What are the primary opportunities for growth?

In 2016, Computational Advertising online advertising surpassed Television Advertising and has an annual growth rate of 11% (global statistics). It is said that it will be doubled in/ around 7 years. By 2020, the computational advertising market is estimated to be valued around $260 Billion. In US alone, computational advertising market has a 17% growth rate. 3 years ago, the market was valued at $60 Billion and currently, it is valued at $70 Billion. Computational advertisements are valued to be 2 orders of the Wall Street.

Half of the computational advertisements is done through Search Engines, while in rest of the 50%, majority of the advertisements is done through individual advertisements on websites. This bring in the huge opportunities for anything related to data and has a huge potential in the future (going forward).

Other Questions with respect to this Lecture: -

Demand Side Platforms (DSP): -
        Are at the buying side. It is responsible for optimizing the spends of the advertisers and act like brokers (on the wall street). Thus, DSP's are responsible for picking the advertisement on the ad. exchange, entering the positioning – exiting the position (just like a broker on the stock market). They get revenues based on their performance and thus, have an incentive to be more efficient. Thus, this is where for most of the Machine Learning come into play.

Supply Side Platforms (SSP): -

Work at the selling side of the exchange market, since they understand the advertisement exchange market. These are technology companies that work, similarly, like a investment bank. They help publishers, maximize the revenue. For e.g. if 10 people are bidding for the advertisement, we can set a price to maximize the revenue, we cannot increase the price of an advertisement arbitrarily. If the price is too high, no one would bid and if the price is too low, you would lose the profits.

Ad Networks: -

Can be considered as a place online, where buyers (advertisers) and sellers (publishers) interact with each other and advertisers bid for the advertisements made available by publishers.

Ad Exchange: -

can be considered as an exchange (like stock exchange), where exchange happens between advertisers (sellers) and publishers (buyers). For e.g. Google and Facebook have their own exchange.

Ques: -

Role of Data Science in Computational Advertising Online Advertising: -

Ans: -

Most of the computational advertising done is through Search Engines and Display advertisements shown on Individual websites. Thus, this brings in a huge role of Data.

Companies like, Google, Facebook etc use target advertising (for computational advertising) to capture the target audience, based on keywords like demographics, geographic location, user history etc. Even the economics play a huge role, where in advertisers try to optimize their purchase and publishers try maximize their selling's. Thus, at each and every step, we can find the implementation of Data Science Techniques, like Machine modeling techniques, Information Retrieval, Distributed Computing, Recommender Systems etc. It is used for optimizing the revenue, maximizing the clicks, making recommendations to users, ranking of advertisements, differentiating between real and fake clicks etc.

Stochastic Gradient Descent (SGD): -

There are multiple ways in which gradient descent techniques might differ from each other for e.g. learning rate, weight updation etc. Stochastic Gradient Descent or SGD can be considered as a function of time/ iteration. Since, in SGD, the whole batch of data is not considered, the rate of convergence is much slower than Batch Gradient Descent. Various ways in which SGD is used in ML applications are as follows: -

- Term Frequency – Inverse Document Frequency (TF-IDF): - To weight each word by its relative rarity, upon going from text data to numeric data.
- Probabilistic Topic Modeling (includes Topic Mixture Models): - Since every document belongs to a particular topic and topic determines the words used.

- Vector Embedding methods (like word2vec and GloVec) are also used in SGD.
- LDA (latent Dirichlet Association): - Since every document has the distribution of its own of how the numbers are divided, this problem cannot be solved using batch gradient descent methods, due to the nature of its complexity, thereSfore Stochastic Gradient techniques are used in such scenarios.