Unit 7: Optimizing E-commerce Systems (5 Hrs.)

Search Engine Optimization (SEO)

- The practice of increasing the **quantity** and ***quality*** of traffic to your website through organic *search engine results*.
- The process of getting traffic from the “free”, “organic”, “editorial” or “natural” search result on search engines.
- Helps to ensure that a site is accessible to a search engine and improves the chances that the site will be found by the search engine.
- Refers to the “industry” that has been created regarding using keyword searching a means of increasing relevant traffic to a website.
- Refers to the process of “optimizing” both the on-page and off-page ranking factors in order to achieve high engine rankings for targeted search terms.

Working Mechanism of Search Engine

- Search engines allow users to search the internet for content using keywords.
- Although the market is dominated by a few, there are many search engines that people can use.
- When a user enters a query into a search engine, a **search engine results page (SERP)** is returned, ranking the found pages in order of their relevance.
- How this ranking is done differs across search engines.
- Search engines often change their algorithms (the programs that rank the results) to improve user experience.
- They aim to understand how users search and give them the best answer to their query.
- This means giving priority to the **highest quality** and most **relevant** pages.

How do search engines work?

There are three key steps to how most search engines work:

- **Crawling** - search engines use programs, called **spiders**, **bots** or **crawlers**, to scour the internet. They may do this every few days, so it is possible for content to be out-of-date until they crawl your website again.
- **Indexing** - the search engine will try to understand and categorise the content on a web page through 'keywords'. Following SEO best practice will help the search engine understand your content so you can rank for the right search queries.
- **Ranking** - search results are ranked based on a number of factors. These may include keyword density, speed and links. The search engine's aim is to provide the user with the most **relevant** result.

- Although most search engines will provide tips on how to improve your page ranking, the exact algorithms used are well guarded and change frequently to avoid misuse.
- But by following search engine optimisation (SEO) best practice you can ensure that:
  - Search engines can easily crawl your website. You can also prompt them to crawl new content.
  - Your content is indexed for the right keywords so it can appear for relevant searches.
  - Your content can rank highly on the SERP.

On Page SEO and Off Page SEO

- Search engine optimization strategy can be divided into two different categories: **on-page SEO and off-page SEO.**
- Both are crucial to the success of an SEO campaign, but they're on completely different sides of the fence.
- On-page SEO focuses on optimizing parts of your website that are within your control, while off-page SEO focuses on increasing the authority of your domain through content creation and earning backlinks from other websites.
- To further understand the difference between the two, you have to understand, at a basic level, how search engine algorithms work. Let's break it down.
- There are two main buckets that search engines look at when evaluating your site compared to other sites on the web.
- **On-page SEO** looks at what your site (or your page) is about
- **Off-page SEO** looks at how authoritative and popular your site is
- Put simply, **what you rank for** is largely determined by on-page factors, while **how high you rank** in the search results is largely determined by off-page factors.

What is On-Page SEO?

- On-page SEO (also known as “on-site” SEO) is the act of optimizing different parts of your website that affect your search engine rankings. Where your website appears in search engine results pages is determined by a number of ranking factors including site accessibility, page speed, optimized content, keywords, title tags, etc. It’s stuff that you have control over and can change on your own website.

What is Off-Page SEO?

- Off-page SEO focuses on increasing the authority of your domain through the act of getting links from other websites.
- A good analogy for how authority works is this. If you have a bathtub with rubber duckies in it (the ducks are your pages), and you start filling the tub with water (links), your duckies are all going to rise to the top.
- This is how a site like Wikipedia ranks for pretty much everything under the sun. It has so much water in its bathtub that if you throw another rubber duck in it, it’s going to float to the top without any other effort.
- There’s a score called “Domain Authority” that calculates how authoritative your website is compared to other sites.

Page Ranks

- **PageRank** (**PR**) is an algorithm used by Google Search to rank web pages in their search engine results.
- PageRank is a way of measuring the importance of website pages.
- According to Google: PageRank works by counting the number and quality of links to a page to determine a rough estimate of how important the website is.
- The underlying assumption is that more important websites are likely to receive more links from other websites.
- Currently, PageRank is not the only algorithm used by Google to order search results, but it is the first algorithm that was used by the company, and it is the best known. As of September 24, 2019, PageRank and all associated patents are expired.

- PageRank is a link analysis algorithm and it assigns a numerical weighting to each element of a hyperlinked set of documents, such as the World Wide Web, with the purpose of "measuring" its relative importance within the set.
- The algorithm may be applied to any collection of entities with reciprocal quotations and references. The numerical weight that it assigns to any given element *E* is referred to as the *PageRank of E* and denoted by PR(E).
- A PageRank results from a mathematical algorithm based on the webgraph, created by all World Wide Web pages as nodes and hyperlinks as edges, taking into consideration authority hubs.
- The rank value indicates an importance of a particular page.
- A hyperlink to a page counts as a vote of support.
- The PageRank of a page is defined recursively and depends on the number and PageRank metric of all pages that link to it ("incoming links"). A page that is linked to by many pages with high PageRank receives a high rank itself.

Google Analytics

- **Google Analytics** is a web analytics service offered by Google that tracks and reports website traffic, currently as a platform inside the Google Marketing Platform brand.
- Google launched the service in November 2005 after acquiring Urchin.
- As of 2019, Google Analytics is the most widely used web analytics service on the web.
- Google Analytics provides an SDK that allows gathering usage data from iOS and Android app, known as *Google Analytics for Mobile Apps*.
- Google Analytics can be blocked by browsers, browser extensions, firewalls and other means.
- Google Analytics has undergone many versions since its inception.
- It is currently on its 4th iteration of the platform which is called GA4.
- GA4, now being the default Google Analytics installation, is the renamed version for the App+Web Property that Google had released in 2019 in a Beta form.
- GA4 has currently replaced **UA**, **Universal Analytics**. One notable feature of GA4 is a natural integration with Google's Big Query a feature previously only available with the enterprise GA 360.
- This move indicates efforts by Google to integrate GA and its free users into their wider cloud offering

- Google Analytics is used to track website activity such as session duration, pages per session, bounce rate etc. of individuals using the site, along with the information on the source of the traffic.
- It can be integrated with Google Ads, with which users can create and review online campaigns by tracking landing page quality and conversions (goals).
- Goals might include sales, lead generation, viewing a specific page, or downloading a particular file.
- Google Analytics' approach is to show high-level, dashboard-type data for the casual user, and more in-depth data further into the report set.
- Google Analytics analysis can identify poorly performing pages with techniques such as funnel visualization, where visitors came from (referrers), how long they stayed on the website and their geographical position.
- It also provides more advanced features, including custom visitor segmentation.
- Google Analytics e-commerce reporting can track sales activity and performance. The e-commerce reports shows a site's transactions, revenue, and many other commerce-related metrics.[

Social Media Analytics

- **Social media analytics** is the process of gathering and analyzing data from social networks such as Facebook, Instagram, LinkedIn and Twitter.
- It is commonly used by marketers to track online conversations about products and companies. One author defined it as "**the art and science of extracting valuable hidden insights from vast amounts of semi-structured and unstructured social media data to enable informed and insightful decision making**."
- There are **three main steps** in analyzing social media: **data identification, data analysis,** and **information interpretation**.
- To maximize the value derived at every point during the process, analysts may define a question to be answered.
- The important questions for data analysis are: "Who? What? Where? When? Why? and How?"
- These questions help in determining the proper data sources to evaluate, which can affect the type of analysis that can be performed.

Data Identification

- Data identification is the process of identifying the subsets of available data to focus on for analysis.
- Raw data is useful once it is interpreted.
- After data has been analyzed, it can begin to convey a message.
- Any data that conveys a meaningful message becomes information.
- On a high level, unprocessed data takes the following forms to translate into exact message: noisy data; relevant and irrelevant data, filtered data; only relevant data, information; data that conveys a vague message, knowledge; data that conveys a precise message, wisdom; data that conveys exact message and reason behind it.
- To derive wisdom from an unprocessed data, we need to start processing it, refine the dataset by including data that we want to focus on, and organize data to identify information.
- In the context of social media analytics, data identification means "what" content is of interest.
- In addition to the text of content, we want to know: who wrote the text? Where was it found or on which social media venue did it appear? Are we interested in information from a specific locale? When did someone say something in social media?

Data Analysis

- Data analysis is the set of activities that assist in transforming raw data into insight, which in turn leads to a new base of knowledge and business value.
- In other words, data analysis is the phase that takes filtered data as input and transforms that into information of value to the analysts.
- Many different types of analysis can be performed with social media data, including analysis of posts, sentiment, sentiment drivers, geography, demographics, etc.
- The data analysis step begins once we know what problem we want to solve and know that we have sufficient data that is enough to generate a meaningful result.
- How can we know if we have enough evidence to warrant a conclusion?
- The answer to this question is: we don't know.
- We can't know this unless we start analyzing the data.
- While analyzing if we found the data isn't sufficient, reiterate the first phase and modify the question. If the data is believed to be sufficient for analysis, we need to build a data model.

Information Interpretation

- The insights derived from analysis can be as varied as the original question that was posed in step one of analysis.
- At this stage, as the nontechnical business users are the receivers of the information, the form of presenting the data becomes important.
- How could the data make sense efficiently so it could be used in good decision making?
- Visualization (graphics) of the information is the answer to this question.
- The best visualizations are ones that expose something new about the underlying patterns and relationships contain the data.
- Exposure of the patterns and understating them play a key role in decision making process.
- Mainly there are three criteria to consider in visualizing data.

Home Assignment

- Recommendation Systems: Collaborative, Content Based
- Use of Recommendation Systems in E-commerce