**Leveraging Big Data for Operational Analytics: A Case Study of Facebook**

**Abstract :** With the advent and increased use of the internet, social media has become an integral part of people’s daily routine. In the second quarter of 2020, Facebook has emerged as the biggest social network worldwide, with over 2.9 billion monthly active users which is approximately the 1/5th of the world's total population. Launched in the year 2004, it has grown tremendously since then. With a sudden obsession in social media, the number of people on Facebook has increased enormously, producing a massive amount of data every minute. They are generating almost 500 terabytes of data every day. To store and process this huge amount of data, the concept of big data is used. Big Data refers to large volume of data which is growing exponentially with time. Big Data uses open source data warehousing platforms like hadoop, hive etc. By being massively engaged in its capacity for collecting, storing, and interpreting data, Facebook has placed Big Data at the heart of its operations. This report presents the application of big data on facebook, examining its current as well as future impact.

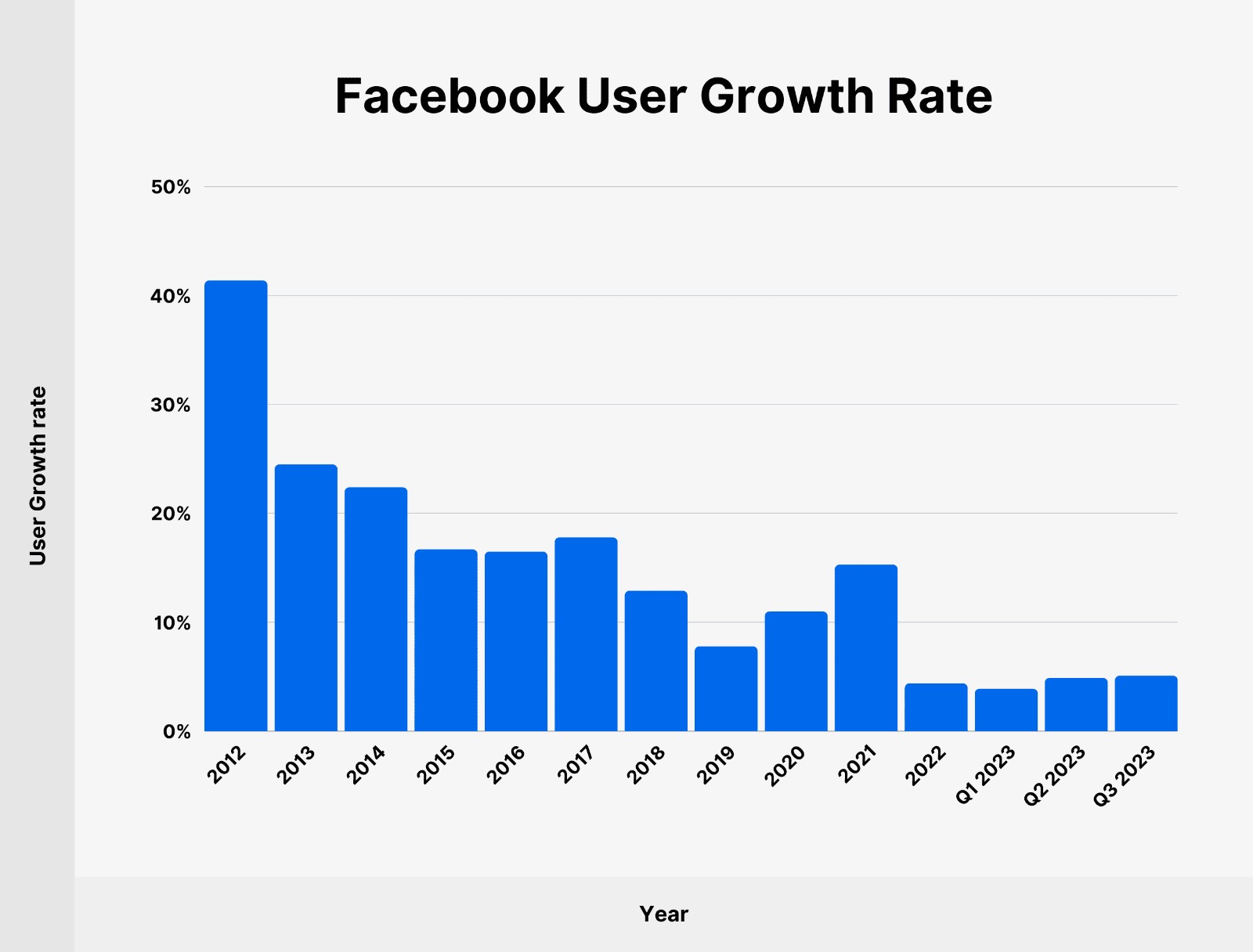
**Keywords:** Big Data, Characteristics, Facebook, Types, Technology.

**INTRODUCTION :**

Arguably the world’s most popular social media network with more than two billion monthly active users worldwide, Facebook stores enormous amounts of user data, making it a massive data wonderland. It’s estimated that there will be more than 183 million Facebook users in the United States alone by October 2019. Facebook is still under the top 100 public companies in the world, with a market value of approximately $475 billion

Every day, we feed Facebook’s data beast with mounds of information. Every 60 seconds, 136,000 photos are uploaded, 510,000 comments are posted, and 293,000 status updates are posted. That is a LOT of data. At first, this information may not seem to mean very much. But with data like this, Facebook knows who our friends are, what we look like, where we are, what we are doing, our likes, our dislikes, and so much more. Some researchers even say Facebook has enough data to know us better than our therapists! Apart from Google, Facebook is probably the only company that possesses this high level of detailed customer information. The more users who use Facebook, the more information they amass. Heavily investing in its ability to collect, store, and analyze data, Facebook does not stop there. Apart from analyzing user data, Facebook has other ways of determining user behavior.

The following image will shows a graph about Facebook user growth rate over the past decade. The y-axis shows user growth rate and the x-axis shows years. The graph shows that Facebook user growth rate has been steadily increasing over the past decade. There are some fluctuations in the growth rate, but overall it has been on an upward trend. According to the graph, the user growth rate peaked at around 50% in 2012 and 2013. Since then, it has gradually declined but has still remained positive. In the third quarter of 2023, the user growth rate was approximately 3%.



**The Role of Big Data in Facebook :**

Big data plays a crucial role in Facebook's operations across various aspects of its platform. Here are some key roles of big data in Facebook's ecosystem:

1. **Personalized User Experience:** Facebook leverages big data analytics to personalize the user experience for its billions of users worldwide. By analyzing vast amounts of user data, including demographic information, interests, behaviors, and interactions, Facebook can tailor content recommendations, friend suggestions, and news feed algorithms to each user's preferences and interests.
2. **Ad Targeting and Optimization:** Big data enables Facebook to deliver highly targeted and relevant advertisements to its users. By analyzing user behavior, interests, and demographic data, Facebook's advertising platform can precisely target ads to specific audience segments, maximizing the effectiveness of ad campaigns and increasing advertising revenue.
3. **Content Optimization and News Feed Ranking:** Facebook uses big data analytics to optimize the ranking of content in users' news feeds. By analyzing factors such as engagement metrics, post popularity, and user interactions, Facebook's algorithms determine which posts are most relevant to each user and prioritize them accordingly in their news feeds.
4. **Platform Reliability and Performance:** Big data analytics helps Facebook monitor and optimize the reliability and performance of its platform. By analyzing data on user activity, server performance, and network traffic in real-time, Facebook can identify and mitigate potential issues before they impact user experience, ensuring the platform remains stable and responsive.
5. **Insights and Decision Making:** Facebook utilizes big data analytics to gain valuable insights into user behavior, market trends, and platform performance. By analyzing large datasets, Facebook can identify patterns, trends, and correlations that inform strategic decisions related to product development, feature enhancements, and business growth initiatives.
6. **Security and Fraud Detection:** Big data analytics plays a crucial role in detecting and preventing security threats and fraudulent activities on the Facebook platform. By analyzing user behavior, login patterns, and content interactions, Facebook's security systems can identify suspicious activities and take proactive measures to protect user data and maintain platform integrity.

**The Relationship Between Facebook and Big Data :**

A report from McKinsey & Co. stated that by 2009, companies with more than 1,000 employees already had more than 200 terabytes of data from their customers’ lives. Facebook tracks each and every activity of a user right from the login time, active hours, photos and videos liked, posts, story updates, games you play, ads you click, searches you make, profiles visited, comments, shares, tags, places visited, current location, pages followed, pages liked, friend requests sent, friend requests accepted, a flashback of posts, stories and many more. One can draw a lot of information from the collected, stored, processed, and analyzed data. Facebook can, at times, understand you better than your therapist. With every passing second, the data and the rate at which the data is being generated keep increasing at an unprecedented rate.

**The Business Relationship Between Facebook and Big Data :**

The major source of earnings for Facebook is advertisements. A wider audience watches Facebook ads, which is why the platform is one of the preferred choices of many social media marketers. Facebook understands users’ behaviour based on the data and then classifies them into certain advertisement categories to be shown. This way, the user might be a little less annoyed by seeing the advertisement of their preferred domains. Also, the media marketing people will be glad to know that their advertisement reaches the target audience. A pretty good business relationship, isn’t it? A few years ago, Facebook introduced the profile picture badge. Users start putting badges showing support to a particular side, be their political interests, sports team, or events happening in their country or worldwide. These interests speak a lot about a person and help to match them to a set of advertisement categories.

**What is Operational Analytics?**

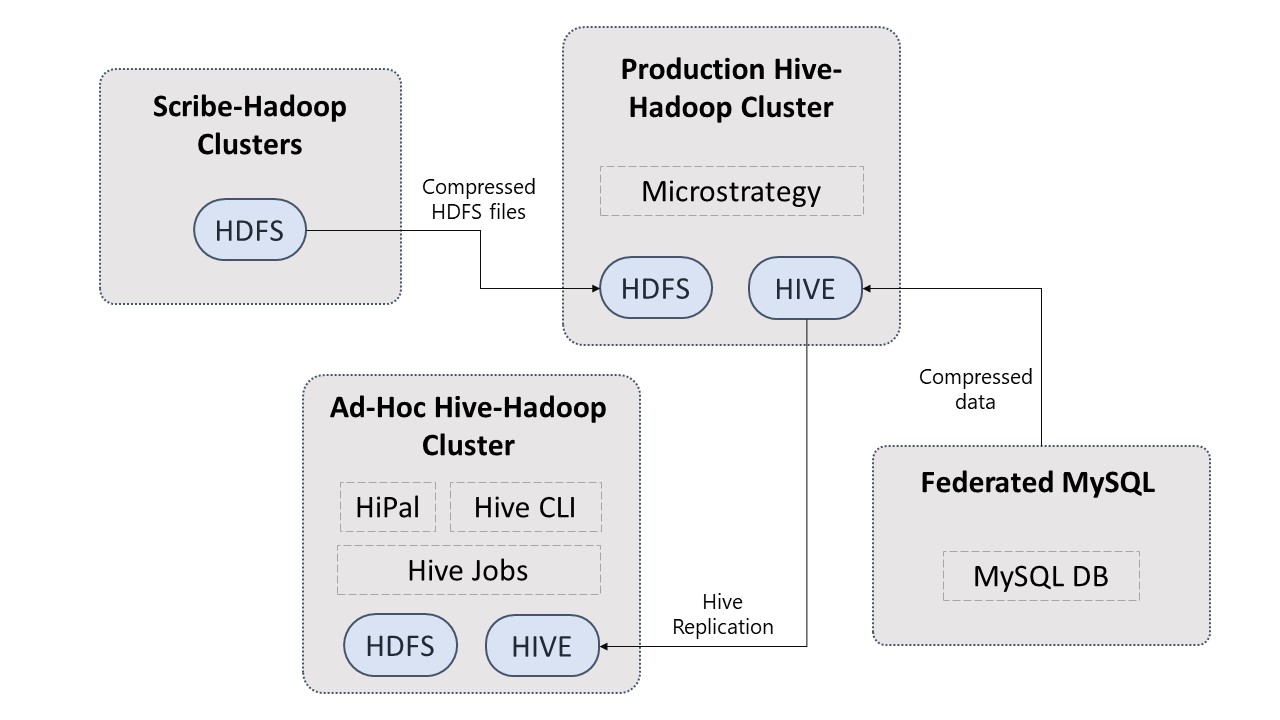
Operational analytics is about using real-time data for daily decisions. Relevant business information flows from many sources into tools that analyze that data and identify problems and opportunities. This actionable data is then used by teams to inform their decision-making. Although it is a subset of business analytics, operational analytics focuses on immediate action. With traditional analytics, IT specialists or analysts mine information stored in data warehouses. Then, using sometimes complex algorithms, they produce weekly, monthly or quarterly reports that executives use to make strategic decisions.

Today, however, companies must keep up with rapid changes in the business landscape. Operational analytics makes insights available in near-real-time for staff who need to address support tickets, repair remote equipment or adjust pricing or sales tactics. In addition, operational analytics empowers predictive analytics by leveraging data mining and AI to connect data points. For example, this approach processes data and sends it to tools such as help desk and social media messaging platforms. This unified effort helps businesses achieve operational excellence.

**Big Data Infrastructure at Facebook :**

Facebook's big data infrastructure is robust and sophisticated, designed to handle the immense volume, velocity, and variety of data generated by its billions of users worldwide. Here are some key components of Facebook's big data infrastructure:

1. **Hadoop:** Facebook relies heavily on Apache Hadoop, an open-source framework for distributed storage and processing of large datasets. Hadoop's distributed file system (HDFS) allows Facebook to store petabytes of data across thousands of servers, while its MapReduce processing framework enables parallel computation for data analysis and batch processing tasks.
2. **Apache Hive:** Facebook uses Apache Hive, a data warehouse infrastructure built on top of Hadoop, to provide a high-level interface for querying and analyzing structured data stored in HDFS. Hive's SQL-like query language (HiveQL) allows engineers and data analysts at Facebook to perform ad-hoc queries and extract insights from large-scale datasets.
3. **Apache Spark:** Facebook employs Apache Spark, a fast and general-purpose cluster computing framework, for advanced analytics and machine learning tasks. Spark's in-memory processing capabilities and rich libraries for machine learning and graph processing enable Facebook to perform complex analytics and derive insights from large-scale datasets efficiently.



**How Does Facebook Use This Data?**

1. **Providing personalized experience:** Facebook leverages big data to provide personalized experiences for its users by analyzing their interactions, interests, and behaviors on the platform. For example, if a user frequently engages with posts related to a specific topic or interacts with certain friends more often, Facebook's algorithms may prioritize content and recommendations related to those interests and connections, creating a more engaging and personalized user experience.
2. **Improving services:** By analyzing large volumes of data, Facebook gains valuable insights into user behavior, preferences, and satisfaction levels, allowing the company to continuously improve its services and features**.** For instance, Facebook uses big data analytics to track user engagement metrics, such as likes, comments, and shares, to identify trends and patterns in user behavior. This information helps Facebook optimize its platform, content delivery, and user interface to enhance user satisfaction and retention.
3. **Ensure the safety and security of the user on the platform:** Big data plays a crucial role in ensuring the safety and security of users on the Facebook platform by enabling the detection and prevention of malicious activities, spam, and abusive content**.** For example, Facebook's automated content moderation systems use big data analytics to detect and remove harmful content, such as hate speech, misinformation, and graphic violence, from the platform, thereby creating a safer and more welcoming environment for users.
4. **Assessing analytics and predicting ad success rate:** Facebook uses big data analytics to assess the performance of advertising campaigns and predict the success rate of ads run on its platform. Additionally, Facebook's predictive analytics models leverage historical ad performance data to forecast the success rate of future ad campaigns, helping advertisers make informed decisions and allocate their advertising budgets more effectively.



1. **Supporting research:** Facebook's vast trove of user data and its big data infrastructure support research efforts across various fields, including social science, public health, and market research. Facebook also collaborates with academic institutions, research organizations, and non-profit entities to provide access to its data for research purposes, contributing to advancements in knowledge and understanding in diverse domains.

**The Ways Facebook using Big Data :**

1. **Analysis of Text:** Facebook analyzes the text content posted by users, including status updates, comments, and messages, using big data techniques to extract insights and understand user sentiments, interests, and preferences. Natural Language Processing (NLP) algorithms are employed to analyze the text, identify keywords, detect sentiment (positive, negative, neutral), and classify content into relevant categories. This analysis enables Facebook to personalize users' news feeds, recommend relevant content, and target advertisements based on the topics and themes that users engage with or express interest in through their text interactions.
2. **Facial Recognition:** Facebook utilizes facial recognition technology powered by big data analytics to automatically identify and tag individuals in photos uploaded to the platform. Advanced machine learning algorithms analyze facial features and patterns in photos to create unique faceprints for each user, enabling accurate recognition and matching of faces across photos. Facial recognition enhances user experience by simplifying the process of tagging friends in photos and enables features such as automatic photo tagging suggestions and personalized photo memories.
3. **Targeted Advertising:** Facebook leverages big data analytics to deliver targeted advertising to its users based on their demographics, interests, behaviors, and interactions on the platform. Through the aggregation and analysis of vast amounts of user data, Facebook's advertising platform enables advertisers to create highly specific audience segments and tailor their ad campaigns to reach users who are most likely to be interested in their products or services. Big data analytics also powers Facebook's ad targeting algorithms, which continuously optimize ad delivery and performance based on real-time feedback and engagement metrics.
4. **Analyzing Likes:** Facebook analyzes user likes—the digital endorsements users give to posts, pages, and ads—to gain insights into users' preferences, interests, and affiliations. Big data analytics algorithms process and analyze users' like activity to identify patterns, trends, and correlations that inform content recommendations, ad targeting, and personalized experiences. For example, if a user frequently likes posts related to fitness and nutrition, Facebook's algorithms may prioritize content and advertisements related to health and wellness in that user's news feed and ad targeting preferences.

**Use Cases:**

1. **Takes You Down the Memory Lane: The Flashback videos**

On the occasion of the platform’s 10th anniversary, Facebook had presented to its users the choice of viewing and sharing a video that unearths the path of their activity on the platform from the date of registration until the existing time. As mentioned above, termed as the “Flashback,” this video is basically a compilation of photos and posts that gained the maximum comments and likes, which are accompanied by nostalgic background music.

These videos are also presented to users on occasions like their “friendversary”, i.e the anniversary of the day they became friends on the platform or on the occasion of the user’s birthday.

1. **The Social Cause: “I Voted” Experiment**

Around the end of 2010, Facebook conducted an extensive social experiment in which it generated a sticker that enabled its users to announce “I Voted” upon their profiles.

This experiment was conducted amidst the 2010 midterm elections and had quite an effective outcome. Upon laying their eyes on the I voted sticker, the users had more chances of voting and also of being more expressive regarding their voting act, as they observe their friends indulging in the action.

As claimed by the Facebook scientists, the sticker — and the peer pressure pertaining to it caused [340,000](https://www.vox.com/2014/11/4/7154641/midterm-elections-2014-voted-facebook-friends-vote-polls) more people to vote in the 2010 midterm elections.

1. **The Tagging Game: Facial Recognition**

Facebook employs image recognition technology that analyzes pictures and videos against other images to detect details like objects and people. It uses the deep learning application "DeepFace" to teach the platform to identify people in photos. Facebook claims DeepFace is more accurate than humans in determining if two different images depict the same person.

In 2010, Facebook introduced facial recognition for the US market through its tag suggestions tool, which would inadvertently tag people in uploaded photos by scanning faces and suggesting names. While users could deactivate it, the feature was enabled by default without explicit consent, leading to privacy controversies.

In December 2017, the tool was rebranded as "face recognition" and Facebook made it easier for users to turn the feature on or off. However, in 2019, amidst concerns over unauthorized collection of biometric data, Facebook made face recognition an opt-in feature in an effort to be more privacy-oriented.

**Challenges and Solutions:**

**Scalability challenges when processing large amounts of data:**

Facebook faces major scalability issues when processing large amounts of data generated by millions of users around the world. As the platform continues to grow and evolve, the volume, speed, and diversity of data pose major challenges to Facebook's processes and operations. To solve these problems, Facebook has invested heavily in creating big data and increasing capacity using technologies such as Hadoop, Hive, and Spark to store, process, and analyze big data efficiently. Additionally, Facebook uses technologies such as data sharing, sharing and replication to distribute data across its systems and ensure high performance and efficiency.

**Ensuring data privacy and security:**

Due to the nature of the user data that Facebook collects and processes, ensuring data privacy and security is important for Facebook. With growing concerns about data privacy and regulatory oversight, Facebook faces significant challenges in protecting users' data from unauthorized access, leakage and improper use. To address these issues, Facebook has implemented stringent security and governance measures, including access, management and monitoring tools, to protect users' information on the platform. Additionally, Facebook complies with data protection laws such as GDPR and CCPA to ensure ethical processing and use of user data.

**Privacy Policy Compliance:**

Due to this massive gold mine of data, advertisers wait like hungry vultures. The 2019 Social Media Marketing Industry Report stated that Facebook is the #1 social platform for marketers.Facebook has always assured its users that information is shared only with their permission and anonymized when sold on to marketers. However, issues still seem to crop up; there have always been high levels of privacy concerns among Facebook users, who ask, “Is Privacy Dead?”. For example, many users complain that Facebook’s privacy settings are not clearly explained or too complicated. It is easy for users to share things unintentionally.

**Impact and Benefits:**

**Increased efficiency and effectiveness:**

By using big data for operational analysis, Facebook has achieved a significant improvement in the efficiency and effectiveness of all aspects of its platform. Real-time monitoring and analysis of users allows Facebook to detect and resolve problems instantly, demonstrating efficiency and reliability. Additionally, predictive maintenance of infrastructure helps reduce downtime and outages and improve overall performance and efficiency.

**Improve user experience and engagement:**

Big data analytics enables Facebook to deliver personalized experiences to its millions of users worldwide, resulting in engagement and satisfaction. By analyzing user behavior, preferences, and interactions, Facebook can customize recommendations, friend suggestions, and news feed algorithms based on each user's preferences, thereby improving user experience and engagement on the platform.

**Cost reduction and optimization:**

Facebook has achieved cost reduction and optimization by effectively using big data analytics. By optimizing server infrastructure, reducing downtime, and improving resource allocation, Facebook can reduce operating costs and increase return on investment in infrastructure and operations. Additionally, by using predictive analytics for resource planning and resource allocation, Facebook can improve resource utilization and avoid unnecessary expenses.

**Future Directions:**

**Emerging Trends in Operational Analytics:**

Looking ahead, Facebook is likely to continue investing in emerging trends in operational analytics, such as real-time analytics, machine learning, and AI-driven automation. These technologies hold the potential to further enhance Facebook's operational efficiency, performance, and user experience, enabling the company to stay ahead of the curve in a rapidly evolving digital landscape.

**Strategies for Continuous Improvement:**

Continuous improvement is essential for Facebook to maintain its leadership position in operational analytics. By fostering a culture of innovation, collaboration, and continuous learning, Facebook can empower its teams to explore new ideas, experiment with new technologies, and drive ongoing improvements in operational efficiency, performance, and user experience.

**Conclusion:**

Now more than ever, organizations are leveraging Big Data analytics to engage with their customers by understanding their behaviour more precisely. Some of the key takeaways from the article are –

Facebook uses all the available user information with its deep neural network models and finds the target audience for a particular advertisement. This helps to serve the users’ advertisements more insightfully.

Due to this, Facebook has emerged as one of the toughest competitors for Google Search Engine and Youtube in the digital marketing race.

With the humungous amount of user data present with Facebook and continuously increasing, there are still a lot of other use cases we might see in the future from Facebook.

However, lately, Facebook has faced much backlash from users over privacy concerns. Facebook is known for tracking different user activities through which it can provide curated advertisements to the target audience. After Apple provided a security update that gives back the control to the user to provide consent to such applications before they can track them, it has been noticed that more than 70 percent of the users have not granted permission to the apps to track their activities. With such security concerns, it will be interesting to see how Facebook can target its users with the right advertisements.