**Application of Big Data in Ecommerce**

ARPITA VERMA ARJITA VERMA

Computer Department Computer Department

NATIONAL P G COLLEGE NATIONAL P G COLLEGE

***Abstract:* Today the consumption of data is rapidly growing around the world with large investment in big data analytics hardware ,Software and Services. Due to the large data sets available, Deep-Learning, the sub-set of Artificial Intelligence is recently emerging the hottest technology trend. Large Companies such as Facebook , Google , Amazon, Intel, Microsoft are heavily investing in big data. Ecommerce business is making accurate strategic decisions on how to operate their online empires . Big data analytics plays a vital role in ecommerce for the competitive business environment. In this we study various ways to use big data in ecommerce. Ecommerce business is using trends such as pridictions,optimising pricing and forecasting demands to gain a competitive advantage. Now ecommerce business is making accurate strategies decisions on how to operate their online empires.**

***Keywords:* Ecommerce , Big Data analytics and business intelligence.**

**INTRODUCTION**

**Big data** is a term used to refer to [data sets](https://en.wikipedia.org/wiki/Data_set) that are too large or complex for traditional [data-processing](https://en.wikipedia.org/wiki/Data_processing) [application software](https://en.wikipedia.org/wiki/Application_software) to adequately deal with. Data with many cases (rows) offer greater [statistical power](https://en.wikipedia.org/wiki/Statistical_power), while data with higher complexity (more attributes or columns) may lead to a higher [false discovery rate](https://en.wikipedia.org/wiki/False_discovery_rate). Big data challenges include [capturing data](https://en.wikipedia.org/wiki/Automatic_identification_and_data_capture), [data storage](https://en.wikipedia.org/wiki/Computer_data_storage), [data analysis](https://en.wikipedia.org/wiki/Data_analysis), search, [sharing](https://en.wikipedia.org/wiki/Data_sharing), [transfer](https://en.wikipedia.org/wiki/Data_transmission), [visualization](https://en.wikipedia.org/wiki/Data_visualization), [querying,](https://en.wikipedia.org/wiki/Query_language) updating, [information privacy](https://en.wikipedia.org/wiki/Information_privacy) and data source. Big data was originally associated with three key concepts: volume, variety, and velocity. Other concepts later attributed with big data are veracity (i.e., how much noise is in the data) and value.

Current usage of the term "big data" tends to refer to the use of [predictive analytics](https://en.wikipedia.org/wiki/Predictive_analytics), [user behavior analytics](https://en.wikipedia.org/wiki/User_behavior_analytics), or certain other advanced data analytics methods that extract value from data, and seldom to a particular size of data set. "There is little doubt that the quantities of data now available are indeed large, but that's not the most relevant characteristic of this new [data ecosystem](https://en.wikipedia.org/w/index.php?title=Data_ecosystem&action=edit&redlink=1)."Analysis of data sets can find new correlations to "spot business trends, prevent diseases, combat crime and so on.”  Scientists, business executives, practitioners of medicine, advertising and [governments](https://en.wikipedia.org/wiki/Government_database) alike regularly meet difficulties with large data-sets in areas including [Internet search](https://en.wikipedia.org/wiki/Web_search_engine), [fintech](https://en.wikipedia.org/wiki/Fintech" \o "Fintech), [urban informatics](https://en.wikipedia.org/wiki/Urban_informatics), and [business informatics](https://en.wikipedia.org/wiki/Business_informatics). Scientists encounter limitations in [e-Science](https://en.wikipedia.org/wiki/E-Science) work, including [meteorology](https://en.wikipedia.org/wiki/Meteorology), [genomics](https://en.wikipedia.org/wiki/Genomics), [connectomics](https://en.wikipedia.org/wiki/Connectomics), complex physics simulations, biology and environmental research.

Big data can be described by the following characteristics:

**Volume**

The quantity of generated and stored data. The size of the data determines the value and potential insight, and whether it can be considered big data or not.

**Variety**

The type and nature of the data. This helps people who analyze it to effectively use the resulting insight. Big data draws from text, images, audio, video; plus it completes missing pieces through data fusion.

**Velocity**

In this context, the speed at which the data is generated and processed to meet the demands and challenges that lie in the path of growth and development. Big data is often available in real-time. Compared to small data, big data are produced more continually. Two kinds of velocity related to Big Data are the frequency of generation and the frequency of handling, recording, and publishing.

**Veracity**

It is the extended definition for big data, which refers to the data quality and the data value. The [data quality](https://en.wikipedia.org/wiki/Data_quality) of captured data can vary greatly, affecting the accurate analysis.

Data must be processed with advanced tools (analytics and algorithms) to reveal meaningful information. For example, to manage a factory one must consider both visible and invisible issues with various components. Information generation algorithms must detect and address invisible issues such as machine degradation, component wear, etc. on the factory floor.

**E-Commerce**

Ecommerce, also known as electronic commerce or internet commerce, refers to the buying and selling of goods or services using the internet, and the transfer of money and data to execute these transactions. Ecommerce is often used to refer to the sale of physical products online, but it can also describe any kind of commercial transaction that is facilitated through the internet.

Whereas e-business refers to all aspects of operating an online business, ecommerce refers specifically to the transaction of goods and services.

The [history of ecommerce](https://www.shopify.com/blog/69521733-proceed-to-checkout-the-unexpected-story-of-how-ecommerce-started) begins with the first ever online sale: on the August 11, 1994 a man sold a CD by the band Sting to his friend through his website NetMarket, an American retail platform. This is the first example of a consumer purchasing a product from a business through the World Wide Web—or “ecommerce” as we commonly know it today.

Since then, ecommerce has evolved to make products easier to discover and purchase through online retailers and marketplaces.  Independent freelancers, small businesses, and large corporations have all benefited from ecommerce, which enables them to sell their goods and services at a scale that was not possible with traditional offline retail.

**Problems In E-Commerce**

There are several technical and non-technical disadvantages of E=Commerce such as the lack of system security and reliability.

Network bandwidth causes a major issues in E-Commerce.

The software development industry is still evolving and keep changing.

It is difficult to ensure the security of online transactions, and many more,

Since , today we are stepping into a digital world,more and more things are getting digitized and so are the people,

Since people are getting more and more comfortable to these e-commerce websites, so some people are misusing this digitization. As some facts say that you can easily launch his/her website in just 5 to 10 minutes.This compromise the security of people using the e-commerce websites. As people establish fake sites and trap customers for money which is becoming a major e-commerce business nowadays. People are launching fake websites and winning the customers to use their sites by posting interesting advertisement and offers and thereby tapping them. This has becom a huge problem now.

**Solution to the problem**

Since, it is highly affecting the customers i.e., public and people are getting frightened of using the websites or doing online transactions. The solution to this could be by controlling the websites that are getting launched i.e.,by having a control over which site is going to launch and which is not. This could be easily controlled by the government by simply using certain criteria which is needed to be fulfilled by a company before launching their site. The government can simply manage this data of all the companies launching their websites by storing it in big data. Government can also give a certificate of authenticity which a website show on their homepage so as to let the customers know that the website is original website not the fake one.