Faculty of engineering - Shoubra Benha University

Research Article / Research Project / Literature Review

in fulfillment of the requirements of

Department	**Engineering Mathematics and Physics**
Division	**----------**
Academic Year	**2019-2020 Preparatory**
Course name	**Computer**
Course code	**ECE001**

Title: -

Big data

By:

	Name	Edu mail	B.N
1	هشام ابراهيم جوده ابراهيم	hesham196146@feng.bu.edu.eg	1017

Approved by:

Examiners committee	Signature
Dr.Ahmed Bayoumi	
Dr.Shady Elmashad	
Dr. Abdelhamid Attaby	

Research objectives

Discuss the importance of Big data

and its applications.

Create a website about this topic.

Upload It on GitHub:

https://github.com/heshamgoda/html_project

And publish it on GitHub Pages:

https://heshamgoda.github.io/html_project/

Abstract

Big data is a new driver of the world economic and societal changes. **Big Data Analytics** poses a grand challenge on the design of highly scalable algorithms and systems to integrate the **data** and uncover **large** hidden values from datasets that are diverse, complex, and of a massive scale.

Table of contents

Divide your research into sections or subjects, mention each section first page at this table

Introduction

What exactly is Big Data?

To really understand big data, it's helpful to have some historical background. Here is Gartner's definition, circa 2001 (which is still the go-to definition): Big data is data that contains greater variety arriving in increasing volumes and with ever-higher velocity. This is known as the three Vs.

Put simply, big data is larger, more complex data sets, especially from new data sources. These data sets are so voluminous that traditional data processing software just can't manage them. But these massive volumes of data can be used to address business problems you wouldn't have been able to tackle before.

The History of Big Data

Although the concept of big data itself is relatively new, the origins of large data sets go back to the 1960s and '70s when the world of data was just getting started with the first data centers and the development of the relational database.

Around 2005, people began to realize just how much data users generated through Facebook, YouTube, and other online services. Hadoop (an open-source framework created specifically to store and analyze big data sets) was developed that same year. NoSQL also began to gain popularity during this time.

The development of open-source frameworks, such as Hadoop (and more recently, Spark) was essential for the growth of big data because they make big data easier to work with and cheaper to store. In the years since then, the volume of big data has skyrocketed. Users are still generating huge amounts of data—but it's not just humans who are doing it.

With the advent of the Internet of Things (IoT), more objects and devices are connected to the internet, gathering data on customer usage patterns and product performance. The emergence of machine learning has produced still more data.

While big data has come far, its usefulness is only just beginning. Cloud computing has expanded big data possibilities even further. The cloud offers truly elastic

scalability, where developers can simply spin up ad hoc clusters to test a subset of data.

Literature Review

How Big Data Works?

Big data gives you new insights that open up new opportunities and business models. Getting started involves three key actions:

- **Integrate**

Big data brings together data from many disparate sources and applications. Traditional data integration mechanisms, such as ETL (extract, transform, and load) generally aren't up to the task. It requires new strategies and technologies to analyze big data sets at terabyte, or even petabyte, scale. During integration, you need to bring in the data, process it, and make sure it's formatted and available in a form that your business analysts can get started with.

- **Manage**

Big data requires storage. Your storage solution can be in the cloud, on premises, or both. You can store your data in any form you want and bring your desired processing requirements and necessary process engines to those data sets on an on-demand basis. Many people choose their storage solution according to where their data is currently residing. The cloud is gradually gaining popularity because it supports your current compute requirements and enables you to spin up resources as needed.

- **Analyze**

Your investment in big data pays off when you analyze and act on your data. Get new clarity with a visual analysis of your varied data sets. Explore the data further to make new discoveries. Share your findings with others. Build data models with machine learning and artificial intelligence. Put your data to work.

Results and discussion

Big Data Use Cases

Product Development	Companies like Netflix and Procter & Gamble use big data to anticipate customer demand. They build predictive models for new products and services by classifying key attributes of past and current products or services and modeling the relationship between those attributes and the commercial success of the offerings. In addition, P&G uses data and analytics from focus groups, social media, test markets, and early store rollouts to plan, produce, and launch new products.
Customer Experience	The race for customers is on. A clearer view of customer experience is more possible now than ever before. Big data enables you to gather data from social media, web visits, call logs, and other sources to improve the interaction experience and maximize the value delivered. Start delivering personalized offers, reduce customer churn, and handle issues proactively.
Machine Learning	Machine learning is a hot topic right now. And data—specifically big data—is one of the reasons why. We are now able to teach machines instead of program them. The availability of big data to train machine learning models makes that possible.
Predictive Maintenance	Factors that can predict mechanical failures may be deeply buried in structured data, such as the year, make, and model of equipment, as well as in unstructured data that covers millions of log entries, sensor data, error messages, and engine temperature. By analyzing these indications of potential issues before the problems happen, organizations can deploy maintenance more cost effectively and maximize parts and equipment uptime.

BIG DATA

What exactly is Big Data?

really understand big data, it's helpful to have some historical background. Here is Gartner's definition, circa 2001 (which is still the go-to definition): Big data is data that contains greater variety arriving in increasing volumes and with ever-higher velocity. This is known as the three Vs.

Put simply, big data is larger, more complex data sets, especially from new data sources. These data sets are so voluminous that traditional data processing software just can't manage them. But these massive volumes of data can be used to address business problems you wouldn't have been able to tackle before.

```html
<html>
<head>
<ul>
<li><a href="index.html">Big Data Definition</a></li>
<li><a href="history.html">Big Data History</a></li>
<li><a href="how_it_work.html">How Big Data Works</a></li>
<li><a href="uses_of_big_data.html">Big Data Uses</a></li>
<li><a href="challenges.html">Big Data Challenges</a></li>
</ul>
<h1>BIG DATA</h1>
</head>
<body>
<h2>What exactly is Big Data?</h2>
<img src="Big-data.jpg" width=400>
<p>
really understand big data, it's helpful to have some historical background. Here
is Gartner's definition, circa 2001 (which is still the
go-to definition): Big data is data that contains greater variety arriving in
increasing volumes and with ever-higher velocity.
This is known as the three Vs.
</p>
<p>
Put simply, big data is larger, more complex data sets, especially from new data sources. The
that traditional data processing software just can't manage them. But these massive volumes o
business problems you wouldn't have been able to tackle before.
</p>
</body>
</html>
```

- Big Data Definition
- Big Data History
- How Big Data Works
- Big Data Uses
- Big Data Challenges

BIG DATA

Big Data Use Cases

Product Development	Companies like Netflix and Procter & Gamble use big data to anticipate customer demand. They build predictive models for new products and services by classifying key attributes of past and current products or services and modeling the relationship between those attributes and the commercial success of the offerings. In addition, P&G uses data and analytics from focus groups, social media, test markets, and early store rollouts to plan, produce, and launch new products.
Customer Experience	The race for customers is on. A clearer view of customer experience is more possible now than ever before. Big data enables you to gather data from social media, web visits, call logs, and other sources to improve the interaction experience and maximize the value delivered. Start delivering personalized offers, reduce customer churn, and handle issues proactively.
Machine Learning	Machine learning is a hot topic right now. And data—specifically big data—is one of the reasons why. We are now able to teach machines instead of program them. The availability of big data to train machine learning models makes that possible.
Predictive Maintenance	Factors that can predict mechanical failures may be deeply buried in structured data, such as the year, make, and model of equipment, as well as in unstructured data that covers millions of log entries, sensor data, error messages, and engine temperature. By analyzing these indications of potential issues before the problems happen, organizations can deploy maintenance more cost effectively and maximize parts and equipment uptime.

```
<table border=black cellpadding=20>
<tr>
 <th width="20%">Product Development</th>
<td>
<p>
 Companies like Netflix and Procter & Gamble use big data to anticipate customer demand. They build predictive models for new
 products and services by classifying key attributes of past and current products or services and modeling the relationship
 between those attributes and the commercial success of the offerings. In addition, P&G uses data and analytics from focus groups,
 social media, test markets, and early store rollouts to plan, produce, and launch new products.
</p>
</tr>
<tr>
 <th>Customer Experience</th>
<td>
<p>
 The race for customers is on. A clearer view of customer experience is more possible now than ever before. Big data enables you to
 gather data from social media, web visits, call logs, and other sources to improve the interaction experience and maximize the value
 delivered. Start delivering personalized offers, reduce customer churn, and handle issues proactively.
</p>
</tr>
<tr>
 <th>Machine Learning</th>
<td>
<p>
 Machine learning is a hot topic right now. And data—specifically big data—is one of the reasons why. We are now able to teach
 machines instead of program them. The availability of big data to train machine learning models makes that possible.
</p>
</tr>
<tr>
 <th>Predictive Maintenance</th>
<td>
<p>
 Factors that can predict mechanical failures may be deeply buried in structured data, such as the year, make, and model of equipment,
 as well as in unstructured data that covers millions of log entries, sensor data, error messages, and engine temperature. By analyzing
 these indications of potential issues before the problems happen, organizations can deploy maintenance more cost effectively
 and maximize parts and equipment uptime.
</p>
</tr>
</table>
```

Conclusions

While big data holds a lot of promise, it is not without challenges.

First, big data is…big. Although new technologies have been developed for data storage, data volumes are doubling in size about every two years. Organizations still struggle to keep pace with their data and find ways to effectively store it.

But it's not enough to just store the data. Data must be used to be valuable and that depends on curation. Clean data, or data that's relevant to the client and organized in a way that enables meaningful analysis, requires a lot of work. Data scientists spend 50 to 80 percent of their time curating and preparing data before it can actually be used.

Finally, big data technology is changing at a rapid pace. A few years ago, Apache Hadoop was the popular technology used to handle big data. Then Apache Spark was introduced in 2014. Today, a combination of the two frameworks appears to be the best approach. Keeping up with big data technology is an ongoing challenge.