تقنية البيانات الضخمة



إعداد محمود عبدالسلام



سلسلة كتيبات تعريفية (العدد 16) موجّه إلى الفئة العمرية الشابة في الوطن العربي



صندوق النقـد الـعربـي ARAB MONETARY FUND

تقنية البيانات الضخمة

سلسلة كتيبات تعريفية (العدد 16) موجّه إلى الفئة العمرية الشابة في الوطن العربي



إعداد محمود عبدالسلام

صندوق النقد العربي أبوظبي – الإمارات العربية المتحدة 2021

© صندوق النقد العربي 2021 حقوق الطبع محفوظة

لا يجوز نسخ أو اقتباس أي جزء من هذ الكتيب أو ترجمته أو إعادة طباعته بأي صورة دون موافقة خطية من صندوق النقد العربي إلا في حالات الاقتباس القصير بغرض النقد والتحليل، مع وجوب ذكر المصدر.

الآراء الواردة في هذا الكُتيب لا تمثل بالضرورة وجهة نظر صندوق النقد العربي، وتبقى معبرة عن وجهات نظر مؤلف الكتيب

توجه جميع المراسلات إلى العنوان التالي:

الدائرة الاقتصادية

صندوق النقد العربي

ص.ب. 2818 - أبو ظبى - دولة الإمارات العربية المتحدة

هاتف: +97126171552

فاكس: +97126326454

البريد الإلكتروني: Economic@amfad.org.ae

الموقع الإلكتروني: http://www.amf.org.ae

يستهدف الكتيب الفئة العمرية الشابة والمواطن العادي في مجتمعنا العربى للتعريف بتقنية البيانات الضخمة واستخداماتها المختلفة.

Big Data Technology تقنية البيانات الضخمة

44	1 1
	ىمحبو پا

4	تمهيد:
5	أو لاً : مقدمة عن البيانات
9	ثانياً : تطور أنظمة المعالجة للبيانات
13	ثالثاً : البيانات الضخمة وخصائصها
20	رابعاً : مزايا وإمكانيات البيانات الضخمة
24	خامساً: تطبيقات ونماذج من الواقع لاستخدام البيانات الضخمة
26	سادساً · الخلاصة و التو صبات

تمهيد:

صندوق النقد العربي مؤسسة مالية عربية إقليمية تأسست عام 1976 ومقرها أبو ظبي بدولة الإمارات العربية المتحدة، بدأت ممارسة نشاطها في عام 1977، ويبلغ عدد الدول الأعضاء فيها 22 دولة عربية. يسعى الصندوق إلى أن يكون مؤسسة مالية عربية رائدة في مجال دعم الإصلاحات الاقتصادية والمالية والنقدية للوصول إلى الاستقرار المالي في المنطقة العربية.

في ظل الثورة الصناعية الرابعة التي يشهدها العالم ونظراً للتطور والنمو السريع الذي يشهده العالم في قطاع تقنية المعلومات وظهور العديد من التقنيات الحديثة ومنها تقنية البيانات الضخمة التي أصبح لها أهمية كبيرة في مجال تقنية المعلومات بشكل عام وبالنسبة للاقتصاد والقطاع المالي بشكل خاص، فقد تم إعداد هذا الكتيب والموجه للفئة العمرية (15-25) عاماً، بهدف تعريفهم بتقنية البيانات الضخمة، وتسليط الضوء على أهميتها واستخداماتها. ذلك في إطار حرص صندوق النقد العربي على تحقيق رؤيته أن يكون منصة توعوية داعمة لصانعي القرار في الدول العربية وشريكاً استراتيجياً وداعماً أساسياً لهذه الدول في مجال التحول المالي الرقمي والاستفادة من التقنيات الحديثة لتحقيق الأهداف التنموية و دعم الاقتصادات العربية.

يعطي الكتيب نبذة مبسطة عن البيانات وأهميتها والتعريف بتقنية البيانات الضخمة وتوضيح مجالات استخدامها في العديد من القطاعات والوقوف على أهم الفرص والتحديات لهذه التقنية.

أولاً: مقدمة عن البيانات

بدايةً وقبل التطرق للموضوع الأساسي وهو البيانات الضخمة يجب أن نفهم أو لاً: ما هي البيانات؟ وما الفرق بينها وبين المعلومات والمعرفة؟

البيانات، يمكن تعريفها بأنها مجموعة من الحروف، أو الكلمات، أو الأرقام، أو الرموز، أو الصور المتعلقة بموضوع ما. والبيانات في حد ذاتها ليس لها معنى أو قيمة، وهي الصورة الخام للمعلومة (1)، مثال بيانات الموظفين وصور هم.

أما المعلومات، فهي البيانات التي تم تحليلها ومعالجتها، بحيث أصبحت تحمل معنى وقيمة، ويمكن الاستفادة منها في اتخاذ القرارات، مثل الحصول على إجمالي عدد الموظفين ومعدل أعمار هم ومعدل سنوات الخبرة، ... الخ.

والمعرفة، هي عملية تحليل المعلومات المختلفة وربطُها ببعضها البعض ووجود فَهم واضح لها ومزجها مع الخبرة، وهو ما يوضحه الشكل التالي.



شكل رقم (1) الفرق بين البيانات والمعلومات والمعرفة

المصدر: من إعداد مؤلف الكتيب.

خلاصة القول: البيانات تعني المواد الخام، والمعلومات هي بيانات منظمة، والمعرفة هي معلومات تم إدراكها.

 $^{^{(1)}}$ Sanders, John. (2016). "Defining Terms: Data, Information and Knowledge", Pages (1-3)

كيف يمكن قياس حجم البيانات ؟

يمكن قياس حجم البيانات باستخدام مجموعة من وحدات القياس المختلفة تبدأ بأقل شيء وهي البت (Bit)، والبايت (Byte) والكيلو بايت (Kilobyte)، وصولاً لأعلى شيء متعارف عليه وهو دوميجمقر وبايت (Domegemegrottebyte).



البت (Bit): هي أصغر وحدة تخزين ممكنة، تمثل مرور أو عدم مرور تيار كهربائي، وقيمة البت تكون 1 أو 0 حيث يمثل 1 مرور تيار كهربائي، و 0 عدم مروره (0).

وحدة قياس سعة التخزين وهي البايت (Byte) ومضاعفاتها مبينة، كما هو موضح في الجدول التالي.

جدول رقم (1) أنواع وحدات قياس حجم البيانات

الحجم	الاختصار	وحدة القياس
0 أو 1	b	البت Bit
8 بت	В	البايت Byte
1024 بایت	kB	كيلوبايت Kilobyte
1024 كيلو بايت	MB	میجابایت Megabyte
1024 ميجابايت	GB	جيجابايت Gigabyte
1024 جيجابايت	TB	تيرابايت Terabyte
1024 تيرابايت	PB	بيتابايت Petabyte
1024 بيتابايت	EB	إكسابايت Exabyte
1024 اكسابايت	ZB	زيتابايت Zettabyte
1024 زيتابايت	YB	يوتابايت Yottabyte
1024 يوتابايت	XB	زينوتابايت Xenottabyte
1024 زينوتابايت	SB	شايلينوبايت Shilentnobyte
1024 شايلينوبايت	DB	دوميجمقر وبايت Domegemegrottebyte

⁽²⁾ Weiss, A. (2018). "Big Data Shocks: An Introduction to Big Data for Librarians and Information Professionals" , Chicago: American Library Association , page 19

Ward, B. (2004). "How Linux Works: What Every Superuser Should Know?".

أساليب تخزين البيانات

تختلف وتتنوع أساليب تخزين البيانات ما بين:

- 1. بيانات هيكلية، يتم تخزينها بطريقة منظمة، مثل البيانات الموجودة في قواعد البيانات وتتميز بسهولة المعالجة والتحليل.
- 2. بيانات غير هيكلية، تمثل النسبة الأكبر من البيانات، وتكون موجودة ومخزنة بطريقة غير منظمة مثل البيانات الموجودة على الشبكة العالمية للمعلومات (الإنترنت) خاصة منها مواقع التواصل الاجتماعي وتطبيقات الهواتف الذكية ولا يمكن معالجتها بسهولة.
- بيانات شبة هيكلية، تخزن بهيئة وترتيب معين مُختلف عن قواعد البيانات (⁴).

شكل رقم (2) أساليب تخزين البيانات

بیاتات شبه هیکلیهٔ SEMI-STRUCTURED DATA	البيانات غير هيكلية UNSTRUCTURED DATA	بیاتات هیکلیهٔ STRUCTURED DATA
تعد على أنها نوع من البيانات المهيكلة ولكنها ليست مبينة في جداول أو قواحد بيانات، ولكنها تخزن في ملفات بشكل منظم مثل : JSON,HTML and XML	تمثل النسبة الأكبر من البيانات، مثل البيانات التي تنتجها المواقع الإلكترونية يوميا من نصوص وصور وفيديوهات وخاصة مواقع التواصل الاجتماعي، وكذلك تطبيقات الأجهزة الذكية.	بيانات تم تنظيمها في جدول أو قاعدة بيانات. مثل البيانات الخاصة بالموظفين أو عملاء البنك أو معاملات المؤسسة المختلفة.
Classe Colleged (children)	abook WhatsApp App	

المصدر: من إعداد مؤلف الكتيب.

'n

⁽⁴⁾ مركز الاحصاء، "مفاهيم عامة حول البيانات الكبيرة"، أدلة المنهجية والجودة، دليل رقم (13)، أبوظبي، صفحة (4).

مصادر البيانات

بالنسبة لمصادر البيانات، فتنقسم إلى ثلاثة مصادر رئيسة (٥):

1-المصدر الأول: البيانات التي تنتجها المؤسسات

البيانات التي يتم إنتاجها من قبل الشركات والمؤسسات خلال عَملها بشكل يومي، ويتم تخزينها في قواعد بيانات أو ملفات تكون خاصة بالشركة أو المؤسسة.



2-المصدر الثاني: البيانات التي ينتجها الإنسان

البيانات الناتجة عن الإنسان سواءً منها المستندات الموجودة على أجهزة الحاسب الآلي أو تلك المخزنة في الأجهزة المحمول المختلفة كالرسائل النصية وتطبيقات الأجهزة الذكية، إضافة إلى تلك البيانات التي يدرجها الإنسان في الشبكة العالمية للمعلومات مثل البريد الإلكتروني ومواقع التواصل الاجتماعي، من مشاركة للصور والمحادثات والتعليقات وتكون أغلبها بيانات غير منظمة.



⁽⁵⁾ Ghotkar, M. and Rokde, P. (2016). "Big Data: How it is Generated and its Importance", pages 2-4.

3-المصدر الثالث: البيانات التي تنتجها الآلات

هي جميع البيانات التي يتم إنتاجها تلقائياً من دون تدخل الإنسان، مثل بيانات كاميرات المراقبة وأجهزة الاستشعار والأقمار الاصطناعية.



ثانياً: تطور أنظمة المعالجة للبيانات

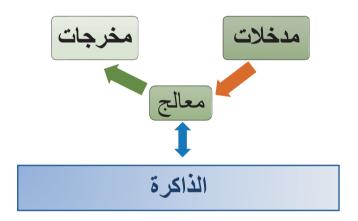
لقد تطورت طرق وآليات معالجة البيانات، حيث مرت بعدة مراحل ارتباطأ بالتطور الذي شهدته البنية الهيكلية سواءً على صعيد تطور الأجهزة ومكوناتها ومواصفاتها الفنية وآلية ربطها ببعضها البعض أو وفقاً للتطور في أنظمة التشغيل نفسها. يمكن تلخيص هذا التطور إلى أهم ثلاث مراحل على النحو التالى:

المرحلة الأولى: الأنظمة التسلسلية أو المتتالية (Serial Systems)

مع ظهور أول حاسب آلي عام 1945 (6)، كانت الأجهزة تعتمد بشكل أساسي في تنفيذ المهام على استخدام "معالج مركزي" Central Processing) ((CPU) حيث يتم تنفيذ تعليمات البرامج ومعالجة البيانات في الذاكرة الإلكترونية بشكل تسلسلي، من هنا تأتي تسميتها بهذا الإسم، علماً وأن ذلك كان يتطلب وقتاً كبيراً جداً لتنفيذ التعليمات بالتتابع واحدة تلو الأخرى.

الرسم البياني التالي يوضح آلية عمل الأنظمة التسلسلية:

شكل رقم (3) الأنظمة التسلسلية أو المتتالية (Serial Systems)

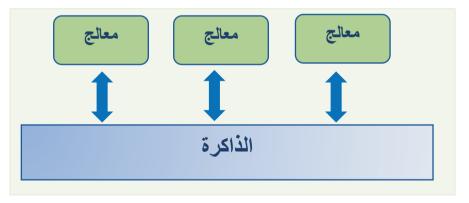


⁽⁶⁾ Mercier, M. (2019). "Contribution to High Performance Computing and Big Data Infrastructure Convergence", Université Grenoble Alpes, Oct.

المرحلة الثانية: الأنظمة المتوازية (Parallel Systems)

بسبب الوقت الكبير المطلوب لتنفيذ البرامج ومعالجة البيانات، فضلاً عن عدم القدرة على تنفيذ أكثر من مهمة في نفس الوقت، فقد ظلت فكرة زيادة كفاءة أجهزة الحاسب الآلي وتعزيز مواصفاتها لمنحها قدرة أعلى في الأداء وتحسين زمن المعالجة والاستجابة، هي الشغل الشاغل للمختصين، وبالفعل نجحوا في ذلك، عن طريق زيادة القدرة على الأداء المتزامن، أي قدرة الجهاز على أداء أكثر من مهمة في نفس الوقت وذلك عن طريق زيادة عدد المعالجات وحجم الذاكرة، وهو ما يعرف بالأنظمة المتوازية، التي تسببت في طفرة في الكثير من المجالات خصوصاً منها الفضاء والطيران. والتي يمكن تجسيدها بالرسم التالى:

شكل رقم (4) الأنظمة المتوازية (Parallel Systems)

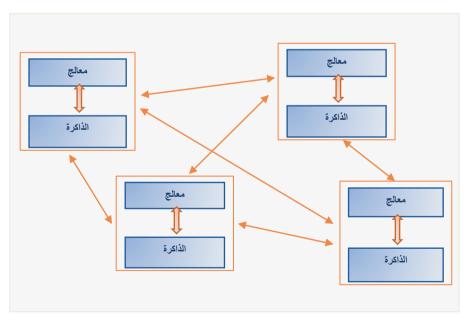


المرحلة الثالثة: الأنظمة الموزعة (Distributed System)

مع بداية الثمانينات التي عرفت تطور الشبكات والقدرة على ربط أكثر من جهاز في شبكة واحدة، فقد ظهر بُعدٌ جديد في التفكير في اتجاه آخر وهو تنفيذ تلك المهام الكبيرة التي مازالت تحتاج إلى وقت معالجة كبير وتتطلب مساحة تخزينية وذاكرة أكبر بشكل مشترك، متزامن وموزع بين أكثر من جهاز واحد عرف بتقنية الأنظمة الموزعة.

كما يبين الشكل التالي، فالأنظمة الموزعة عبارة عن أنظمة تعمل على العديد من الأجهزة مرتبطة مع بعض بشبكة، أي أنها أنظمة تشغيل تظهر للمستخدم وكأنه يعمل على نظام مركزي، إلا أنه في الحقيقة يعمل على عدة وحدات معالجة مركزية تعمل على مستوى التوزيع.

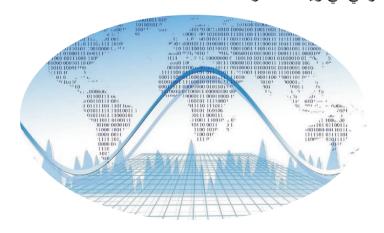
شكل رقم (5) الأنظمة الموزعة (Distributed Systems)



ثالثاً: البيانات الضخمة وخصائصها

ثورة البيانات وتضخمها

العديد من الخبراء والمختصين ينظرون إلى البيانات بأنها "نفط المستقبل" ذلك مع ظهور مجموعة من العلوم الجديدة، مثل التنقيب في البيانات التي تهتم بالصورة الخام البيانات ومعالجتها وتحويلها في شكل يمكن الإستفادة منه في مجالات المعرفة والذكاء الاصطناعي وفروعه المختلفة، من تعلم الآلة والتعلم العميق. حيث يمكن القول أن البيانات هي الأساس أو الوقود لتلك العلوم التي لم تكن تحظى بكثير من النجاح في تلك الفترة، مثل ما هو الحال في الوقت الحاضر، ذلك لأن حجم البيانات ومصادر ها التي كانت موجوده آنذاك لم تكن بنفس الكمية والعدد الموجود اليوم، نتيجة للتحول الرقمي والاعتماد على الخدمات الإلكترونية والأنظمة الذكية واستخدام مواقع التواصل الاجتماعي بشكل أساسي ويومي، والتوجه نحو الشراء الإلكتروني، علماً وأن حجم البيانات تضخم بشكل غير مسبوق كما تبرزه البيانات التالية حيث تشير الإحصاءات الماضية وهي في زيادة مستمرة (7).



⁽⁷⁾ WEF, (2019). "Why big data keeps getting bigger", available at: https://www.weforum.org/agenda/2019/07/why-big-data-keeps-getting-bigger

Big Data Technology تقنية البيانات الضخمة

• يقوم 444,694 مستخدم بالتقديم على وظائف في اللينكدان (LinkedIn).

في

الدقيقة

الواحدة

- يتم إرسال 666,667,41 رسالة على الواتساب.
 - يضاف 319 مستخدم جديد على تويتر.
- عقد ما يقارب من 333,208 اجتماع على منصة زوم.
 - ، تحميل 500 ساعة من الفيديو هات على اليوتيوب.
- إرسال 150,000 مشاركة وتحميل 147,000 صورة على الفيسبوك. مع وجود نحو 4.5 بليون مستخدم على الشبكة العالمية للمعلومات، تشير التقديرات إلى أنه سيتم إنشاء 463 "إكسابايت" (Exabyte) من البيانات يومياً على مستوى العالم أي ما يعادل 212،765،957 قرص (DVD).

المصدر: المنتدى الاقتصادي العالمي - 2020

علماً وأنه يمكن متابعة إحصاءات الشبكة العالمية للمعلومات آنياً من خلال الموقع: www.internetlivestats.com

التحديات والصعوبات

بالرغم من توفر البيانات وزيادة أحجامها بشكل كبير جداً بما يساعد علوم الذكاء الاصطناعي والتنقيب في البيانات وتعلم الآلة على تحقيق أهدافها، إلا أنها مازالت تواجه تحديات وصعوبات من نوع آخر، نتيجة لهذا الكم الضخم من البيانات. فقد ظهرت العديد من التحديات في استيعابها ومعالجتها بالطرق والأدوات التقليدية، يتمثل أهمها فيما يلى:

- تنوع واختلاف البيانات ووجودها في صور غير منظمة بالشكل الكافى.
 - تنوع المصادر.
 - حجم البيانات الهائل والمتزايد.
 - صعوبة التوسع والتمدد.
 - التكلفة الكبيرة في تخزين البيانات ومعالجتها وتحليلها.
 - عامل السرعة والوقت في الحصول على المعلومات.

نشأة البيانات الضخمة

نتيجة للتحديات والصعوبات التي ذكرناها، وجد المختصون والمهتمون بتلك العلوم أنفسهم أمام خيارين، إما إهمال الحجم الهائل من البيانات أو التغلب على تلك التحديات والصعوبات وايجاد حل لها. حيث كان

BIG DATA

الحل في الأنظمة الموزعة، نتيجة لقدرتها على معالجة البيانات الضخمة على أكثر من جهاز في نفس الوقت. بما ساعد على ظهور تقنية البيانات الضخمة، التي تمثل مصطلح جديد يصف الكم الهائل من البيانات والمعلومات من المصادر المختلفة، التي لا يمكن تخزينها ومعالجتها بالطرق التقليدية وهي واحدة من أهم التقنيات

تعريف البيانات الضخمة

المستقبلية

تختلف وتتنوع التعريفات والمفاهيم لهذا المجال مابين الخبراء والشركات والمنظمات المتخصصة، حيث يعرف معهد ماكنري العالمي البيانات الضخمة، أنها مجموعه من البيانات التي يفوق حجمها القدرة على معالجتها باستخدام أدوات قواعد البيانات التقليدية، من التقاط ومشاركة ونقل وتخزين وإدارة وتحليل، في غضون فترة زمنية مقبولة (8).

تعرف شركة جارتنر المتخصصة في أبحاث واستشارات تقنية المعلومات البيانات الضخمة على أنها "الأصول المعلوماتية كبيرة الحجم وسريعة التدفق وكثيرة التنوع التي تتطلب طرق معالجة مجدية اقتصادياً ومبتكرة من أجل تطوير البصائر والمساعدة على اتخاذ القرارات" (9).

⁽⁸⁾ مركز الاحصاء، "مفاهيم عامة حول البيانات الكبيرة"، أدلة المنهجية والجودة، دليل رقم (13)، أبوظبي، صفحة (4).

⁽⁹⁾ Laney, D. et al. (2013). "Big Data Means Big Business", Gartner Inc.

حسب شركة (IBM) تنشأ البيانات الضخمة عن طريق كل شيء من حولنا. في كل الأوقات، كل عملية رقمية وكل تبادل في وسائل التواصل الاجتماعي ينتج البيانات الضخمة، التي تتناقلها الأنظمة وأجهزة الاستشعار والأجهزة النقالة. البيانات الضخمة لها مصادر متعددة تختلف من حيث السرعة والحجم والتنوع. لكي نستفيد من البيانات الضخمة، نحتاج إلى معالجة مثالية وقدرات تحليلية ومهارات (10).

حسب الإتحاد الدولي للاتصالات (ITU) يشير مصطلح البيانات الضخمة إلى مجموعات البيانات التي تتميز أنها فائقة حجماً وسرعة أو تنوعاً، بالقياس إلى أنواع مجموعات البيانات معهودة الاستخدام (11).

الخلاصة: تشكل البيانات الضخمة مجالاً هندسياً يوفر مجموعة من الأدوات والتقنيات التي تفوق في حجمها قدرات وإمكانيات الطرق والأدوات التقليدية في وقت قياسي.

⁽¹⁰⁾ Miele, S. and Shockley, R. (2013). "Analytics: The real-world use of big data", IBM Global Services

 $^{^{(11)}}$ ITU, (2013). "Big Data today: normal tomorrow", ITU Technology Watch Report

آلية عمل البيانات الضخمة

لتبسيط التعريفات والمصطلحات السابقة، دعونا نتعرف على إطار عمل البيانات الضخمة، لنفترض أن شركة لديها بيانات عملاء حجمها يزيد عن واحد تيرا بايت، الذي يمثل الحد الأدنى لحجم البيانات لكي تصنف كبيانات ضخمة. أرادت هذه الشركة أن تحدد قائمة بأسماء العملاء وأرقام هواتفهم ممن يقيمون في منطقة معينة لاستهدافهم في حملة إعلانية. حسب الطريقة التقليدية، هذا الأمر يمكن من خلال الرسم التالى:

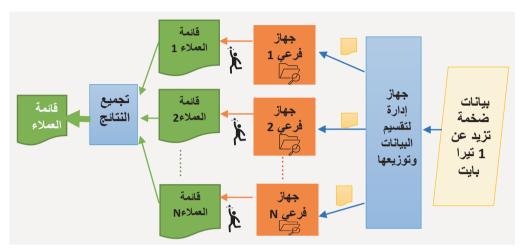
شكل رقم (6) الطريقة التقليدية في معالجة البيانات



المصدر: من إعداد مؤلف الكتيب.

معالجة الأمر بالطريقة التقليدية يستغرق وقتاً طويلاً جداً، ربما يصل إلى أيام للحصول على النتائج المطلوبة، أو قد يصعب ذلك. فلنا أن نتخيل الفرق إذا أعطينا مجموعة كتب لشخص واحد للبحث عن الصفحات التي بها كلمة معينة أو أن تستعين بمجموعة أشخاص ونعطي كل منهم كتاب للبحث، فبالطبع الطريقة الثانية هي الأسرع، وهي نفس فكرة عمل البيانات الضخمة. يجسد الرسم التالي آلية عمل تقنية البيانات الضخمة بشكل مبسط. كل مكون ووظيفه بها مجموعة من الوظائف والمكونات والتفاصيل التي تحتاج إلى مختصين لفهمها.

شكل رقم (7) آلية عمل تقنية البيانات الضخمة في معالجة البيانات



المصدر: من إعداد مؤلف الكتيب.

خصائص البيانات الضخمة

هل كل البيانات ذات أحجام كبيرة هي بيانات ضخمة؟

لقد تم وضع مجموعة من الخصائص والمحددات لتصنيف البيانات على أنها بيانات ضخمة ومتعارف عليها باسم (V_S) بدأت بثلاث خصائص وصولاً إلى 10 خصائص نظراً لأنها تبدأ بحرف الV، وهي:

- الحجم (Volume): هو العامل الرئيس في وصف البيانات أنها ضخمة بحيث يتعدى حجمها واحد تير ا بايت.
 - التنوع (Variety): تنوع البيانات ما بين مهيكلة وغير مهيكلة.
 - السرعة (Velocity) إنتاج معدلات مرتفعة من البيانات في كل لحظة.
 - الدقة / الموثوقية (Veracity) : أن تكون البيانات موثوقة وصحيحة.
- القيمة (Value): القدرة على تحويل جميع أنواع البيانات إلى بيانات ذات فائدة

- التمثيل البصري (Visualization): إمكانية إظهار وعرض البيانات بشكل أفضل بحيث يمكن للشخص الذي يقرأها أن يفهمها بشكل سريع.
- التباين / التغير (Variability): عدد الاختلافات في البيانات نتيجة للتغير سواءً في بنية البيانات أو المعنى أو الشكل.
- الثغرة الأمنية (Vulnerability): الحفاظ على أمن وخصوصية البيانات.
- الجودة / المصداقية (Validity): أن تكون مصادر البيانات دقيقة وأن تكون البيانات موثوقة للاستخدام المقصود.
 - التقلب (Volatility) : مدة صلاحية البيانات وطول مدة تخزينها.

Volume الحجم Veracity الموثوقية القيمة Volatility Vulnerability الثغرة الأمنية خصائص السائات الضّخمة (Vs) Variety Velocity التثوع Validity Variability التباين

شكل رقم (8) ملخص خصائص البيانات الضخمة (Vs)

رابعاً: مزايا وإمكانيات البيانات الضخمة

بشكل عام تقدم تقنية البيانات الضخمة العديد من المزايا والإمكانيات التي تتلخص فيما يلي:

- توفير الوقت والجهد والتكلفة العالية المطلوبة في معالجة وتحليل البيانات الضخمة.
- قدرتها على تحليل بيانات وسائل التواصل الاجتماعي لإستخراج الأراء المختلفة مما يتيح إمكانية تحليل انطباعات الجمهور تجاه قضايا معينة لقياس مدى رضاهم من عدمه.
 - التحليلات التنبؤية.
 - تحلیل وتقییم المخاطر المالیة.
- تصنيف المستخدمين وتحديد الفئة المستهدفة لموضوع معين مما يساعد على تحسين أداء وفاعلية الحملات التسويقية والإعلانية.
 - تحسين خدمة وإدارة العملاء.
 - کشف الاحتیال.

أخيراً، لأن البيانات هي الشريان الحيوي لكثير من العلوم كما ذكرنا سابقاً، فإن إمكانية استيعاب البيانات بأحجام أكبر مما سبق ومعالجتها في وقت أسرع سيزيد من الاستفادة وتحقيق الأهداف المرجوة من تلك العلوم، خصوصاً في مجال الذكاء الاصطناعي.



مجالات استخدام البيانات الضخمة

نظراً للإمكانيات والمزايا التي توفرها تقنية البيانات الضخمة، فإن مجالات ومساحات التطبيق للبيانات الضخمة كبيرة جداً ويمكن الاستفادة منها في العديد من المجالات الحياتية والقطاعات ومنها الاقتصاد والمالية والإستثمار والتسويق والطاقة والتعليم والصحة والبيئة والاتصالات والبحث العلمي وغيرها من مجالات. سوف نتناول في هذا الجزء بعض الأمثلة عن كيفية الاستفادة من تلك التقنية:

في القطاع العام والحكومي

يمكن أن تستخدم البيانات الضخمة في دعم واتخاذ القرارات، من خلال محتويات مواقع التواصل الاجتماعي للوقوف على أراء المواطنين وردود أفعالهم حول قرار ما أو نظام معين، مطبق أو ترغب الحكومة تشريعه، بما يساعد على اتخاذ القرار المناسب لكل حالة.

القطاع الخاص

- يمكن للشركات والمؤسسات من خلال تحليل البيانات الناتجة عن وسائل التواصل الاجتماعي، لتحديد جمهورها والتنبؤ بنتائج حملة التسويق والمبيعات.
- كذلك يمكن للحركات تحسين العمليات التجارية وفهم احتياجات العملاء
 والتنبؤ بها وكذلك تطوير منتجات جديدة وفقاً لإحتياجات العملاء.
- أيضاً سهولة فهم ظروف السوق وتحليل سلوكيات الشراء بالتالي زيادة المبيعات.
- إمكانية التحكم في السمعة عبر الشبكة العالمية للمعلومات، حيث يمكن للمؤسسات والشركات تحليل مواقع التواصل الاجتماعي ومعرفة ما يُقال عنها والعمل على تحسين سمعتها.

التعليم

- الاستفادة من تحليل البيانات الصادرة من مواقع التعليم عبر الشبكة العالمية للمعلومات ومن المواد التعليمية المختلفة، لتحسين تعلم الطلاب ورفع كفاءتهم.
- السرعة في تشخيص الاحتياجات في التعلم أو المتاعب التي تواجه الطلاب أثناء عملية التعلم.
- الاستفادة من تلك البيانات في قياس أداء الطلاب والمساعدة في تصميم
 بيئات تعلم متخصصة وفق احتياجاتهم.
- الاستفادة من تحليل بيانات مواقع التواصل الاجتماعي في معرفة ردود
 الفعل الفردية والجماعية لمجموعة من القضايا التعليمية.

الصحة

- من خلال تحلیل البیانات من السجلات الطبیة یمکن اکتشاف مسببات أمراض معینة مما یساعد علی إیجاد علاج لها.
 - التنبؤ بالأوبئة والأمراض لمنعها.
- المساعدة على تقديم العلاج الملائم بمساعدة المعلومات والبيانات
 الهائلة التي تجمعها الأجهزة الذكية وأجهزة الاستشعار.
- استيعاب وتحليل البيانات الضخمة من أجهزة الاستشعار الحيوية اللاسلكية للمرضي مثل بيانات ضربات القلب وضغط الدم ومعدلات الأوكسجين والاستفادة منها في التعرف على الأنماط الدقيقة والخفية حول الأمراض، خصوصاً منها المزمنة مثل أمراض القلب والسكري والانسداد الرئوي، مما يساهم في فهم أكثر لمسببات تلك الأمراض وتحسين البحث العلمي.

مجال الأحوال الجوية

في ظل وجود العديد من أجهزة الاستشعار المتوفرة اليوم، والملحقة
 بأجهزة الهواتف الذكية، فإن لدينا بيانات أكثر من أي وقت مضى وبدقة
 عالية للمساعدة على استقراء الأحوال الجوية.

القطاعات المالية والاستثمارية

- تمكين البنوك من تحقيق عوائد أكبر من خلال جمع بيانات أشمل عن العملاء وتحليلها وإمكانية ترويج كل منتج وفقاً لمتطلبات كل عميل مما يرفع نسبة تفاعل واستخدام العملاء للخدمات والمنتجات المصرفية، والحفاظ على استدامة العملاء في ظل المنافسة الكبيرة وتعزيز مبادرة "اعرف عميلك".
- تسهیل إدارة المخاطر من خلال سهولة تحلیل البیانات من مصادر متنوعة ومختلفة.
 - تحدید الفرص الاستثماریة.
 - منع الاحتيال.
 - تمكين وتعزيز الاستفادة من تقنية البلوك تشين.

المجال الأمنى

- تسهيل عملية تحليل البيانات الضخمة التي توفر ها كاميرات المراقبة.
 - تحلیل البیانات من السجلات الاجرامیة والتنبؤ بالجریمة.
- تحلیل البیانات مواقع التواصل الاجتماعي بما یساعد على منع العملیات
 الإرهابیة و مكافحة الجریمة.

أمن المعلومات

- o تساعد في تحديد الهجمات الخبيثة ومنع الاحتيال.
 - بحث وتحليل البرامج الضارة.
 - فرز وتصنيف الرسائل الإلكترونية

خامساً: تطبيقات ونماذج من واقع استخدام البيانات الضخمة

هناك العديد من الأمثلة والتطبيقات التي نجحت في الاستفادة من تقنية البيانات الضخمة في عدة مجالات من أهمها:

التعليم

يعتبر تطبيق دايلنجو لتعلم اللغات واحد من التطبيقات والأنظمة التي استطاعت الاستفادة من البيانات الكبيرة عن طريق تخزين البيانات للمستخدمين وتفاعلاتهم وتحليلها باستمرار لمعرفة نجاحات المستخدمين وإخفاقاتهم وأنشطتهم، والعمل على التحسين المستمر لذلك، يعتبر اليوم واحداً من التطبيقات المعروفة في تعليم اللغات.

المبيعات

استطاعت شركة أمازون معالجة ملايين العمليات كل يوم والاستفاده من تحليل السلوك الاستهلاكي لعملائها وتحديد احتياجتهم واقتراح المنتجات والبضائع عليهم مما ساهم في زيادة المبيعات والأرباح.

الترفيه

استطاعت منصة نتفليكس من استغلال وتحليل بيانات المستخدمين في تحديد المواد المناسبة واقتراحها عليهم حسب الوقت المناسب لهم.

النقل والمواصلات

استفادت شركة أوبر من تحليل بيانات الركّاب في تقديم خدمات وعروض تناسب احتياجتهم وقدراتهم المالية.

الصحة

نجحت شركة مايكروسوفت بالتعاون مع أحد المعاهد الهندية في حصولها على كم بيانات ضخم من سجلات، لأكثر من مليون شخص، في مساعدة الأطباء في التعرف بشكل دقيق على أسباب انتشار أمراض العيون، إضافة إلى مساعدة الأطباء على تحديد الوقت المطلوب للمريض حتى تستقر حالته الصحية ويعالج بصره، وكذلك القيام بالعمليات الجراحية للعيون بنجاح.

الخدمات المالية

نجحت شركة أمريكان اكسبريس في استغلال قاعدة بيانات تحتوي على كمية هائلة من البيانات لمالكي البطاقة الإئتمانية من مختلف أنحاء العالم وفهم سلوك المستهلك بشكل أعمق والتنبؤ بالمعاملات المستقبلية.

الصناعة

تعتمد شركة جنرال الكتريك على ابتكار منتجات وطرح حلول رقمية ترتبط باحتياجات ومتطلبات المستهلكين.

الدعاية والتسويق

يعتبر الفيسبوك واحداً من أكبر مصادر البيانات الضخمة وقد نجح باستغلالها والاستفاده منها في تصميم خوار زميات تعرض للمستخدمين إعلانات لبضائع وخدمات تناسب احتياجاتهم واهتماماتهم.

شكل رقم (10) تطبيقات من و اقع استخدام البيانات الضخمة



سادساً: الخلاصة والتوصيات

البيانات، نظراً لأهميتها، يطلق عليها إسم "نفط المستقبل" وهي الأساس للحصول على المعلومات والمعرفة، كما تعتبر الغذاء والوقود للعلوم الحديثة مثل الذكاء الاصطناعي وتعلم الآلة، ودور تقنية البيانات الضخمة هو تجميعها ومعالجتها وتجهيزها بالشكل المطلوب وبأقصى سرعة.

ماز ال مجال البيانات الضخمة يواجه مجموعة من التحديات والصعوبات التي يجب العمل عليها وإيجاد حلول مناسبة لها للاستفادة منها بشكل بأفضل، منها:

- التأكد من جودة وصحة البيانات.
- صعوبة الحصول على البيانات من مصادرها، خصوصاً منها الحكومية.
 - التحديات المتعلقة بأمن وسرية المعلومات وسوء استغلالها.
- عدم فهم وإدراك بعض المؤسسات والأفراد للإمكانيات والمزايا التي تتبحها البيانات الضخمة.
 - التكاليف المالية لبدء وتمويل مشاريع البيانات الضخمة.
 - نقص المهارات والخبرات نظرا لحداثة المجال.
- عدم وجود قوانين وتشريعات تنظم البيانات الضخمة وخصوصاً المسؤولية والمساءلة.
 - الحصول على البيانات في الوقت المناسب.
 - الحفاظ على ملكية وحساسية البيانات وسبل التعامل معها من حيث التخزين وقابلية النشر أو المشاركة.



ربما من أهم هذه التحديات، المخاطر المرتبطة بالمستخدم البسيط خصوصاً فيما يتعلق بحريته الشخصية، ما يدعو بضرورة الحفاظ على البيانات والمعلومات الشخصية قدر الإمكان، بحيث لا تكون سلعة تستغلها الشركات والتي أصبحت تعرف عن الأشخاص أكثر مما يعرفون عن أنفسهم.

بالرغم من تلك التحديات والصعوبات إلا أن مجال البيانات الضخمة يعتبر حديث العهد، إلا أن له مستقبل واعد، والدليل على ذلك أن العديد من الدراسات والاستطلاعات تتوقع نمو سوق البيانات الضخمة بشكل كبير جداً في المستقبل، بحجم يصل إلى ما يزيد عن 90 مليار دولار بحلول عام 2025. (12)

إضافة إلى أن الكثير من المحللين يتوقعون أن يرتفع الطلب على البيانات الضخمة مستقبلاً، بسبب دورها في سرعة اتخاذ القرار. كما يتوقع أن ينمو دورها في البحث العلمي وفي مجال الأمن وغيرها من المجالات المهمة، وسوف تصبح ركيزة أساسية في المنافسة وزيادة الإنتاجية والابتكار.

سوق البيانات الضخمة يشهد حالياً نمواً كبيراً، كما توفر العديد من الفرص الوظيفية مع توقعات بفرص أكثر في المستقبل، ويرجع ذلك إلى أن العديد من الشركات الكبيرة، مثل "جوجل" و"فيس بوك" و"أمازون" والعديد من الشركات الأخرى الرائدة في مجال التقنية تقوم حالياً باستثمار مبالغ ضخمة في جذب أصحاب المهارات في هذا المجال والمجالات ذات العلاقة خصوصاً الذكاء الاصطناعي وتعلم الآلة.

حسب الدراسات والمؤشرات، فإن أهم القطاعات التي سوف تشهد نمواً في الاستفادة من البيانات الضخمة هي: الرعاية الصحية والقطاع العام وتجارة التجزئة والاتصالات والتسويق والدعاية.

مصادر البيانات وحجمها ضخم وفي نمو متسارع مما يزيد من أهمية تلك التقنية وزيادة الإقبال عليها في ظل زيادة الإدراك بأهمية البيانات، التي تدرك جميع الحكومات العربية والمؤسسات والشركات أهميتها وتعمل بالتالي على توفير كافة السبل والفرص ومنها تقديم المزيد من الدعم للشركات الناشئة ورواد الأعمال لتحقيق الإستفادة القصوى منها، لأن الاستثمار في البيانات الضخمة له أثر اقتصادي واجتماعي من حيث النمو الاقتصادي وتوفير الوظائف.

27

⁽¹²⁾ حكومة دولة الامارات العربية المتحدة، (2020). " تقرير فرص المستقبل".

قائمة المصادر:

المصادر باللغة الانجليزية

- Sanders, J. (2016). "Defining Terms: Data, Information and Knowledge", 10.1109/SAI.2016.7555986. Pages (1-3).
- Weiss, A. (2018). "Big Data Shocks: An Introduction to Big Data for Librarians and Information Professionals", Chicago: American Library Association, page 19.
- Ward, B. (2004). "How Linux Works: What Every Superuser Should Know?".
- Ghotkar, M. and Rokde, P. (2016). "Big Data: How it is Generated and its Importance". pages 2-4.
- Mercier, M. (2019). "Contribution to High Performance Computing and Big Data Infrastructure Convergence", Université Grenoble Alpes, Oct.
- WEF, (2019). "Why big data keeps getting bigger", available at: https://www.weforum.org/agenda/2019/07/why-big-data-keeps-getting-bigger.
- Laney, D. and Le Hong, H. and Lapkin, A. (2013). "Big Data Means Big Business", Gartner, Inc.
- Miele, S. and Shockley, R. (2013). "Analytics: The real-world use of big data", IBM Global Services.
- ITU, (2013). "Big Data today: normal tomorrow", ITU Technology Watch Report.

المصادر باللغة العربية

- مركز الاحصاء، "مفاهيم عامة حول البيانات الكبيرة"، أدلة المنهجية والجودة، دليل رقم (13)، أبوظبي، صفحة (4).
 - حكومة دولة الامارات العربية المتحدة، (2020). " تقرير فرص المستقبل".

الصور المستخدمة في الكتيب متوفرة في الموقع: https://pixabay.com

للحصول على مطبوعات صندوق النقد العربي يرجى الاتصال بالعنوان التالي:

صندوق النقد العربي

شبكة المعر فة

ص.ب. 2818

أبو ظبى - الإمارات العربية المتحدة

هاتف رقم: 6215000 (+9712)

فاكس رقم:6326454 (+9712)

البريد الإلكتروني: <u>Publications@amfad.org.ae</u>

• متوفرة إلكترونياً بموقع الصندوق على الشبكة الإلكترونية:

https://www.amf.org.ae



Arab Monetary Fund Building Corniche Street Abu Dhabi, United Arab Emirates P.O Box 2818 www.amf.org.ae