

الذكاء الاصطناعي وعلاقته بالبيانات الضخمة

Artificial intelligence and its relationship to big data

د. مروة زهواني

محرر التطبيقات الكمية والنوعية للارتقاء الاقتصادي، الاجتماعي والبيئي

بالمؤسسات الجزائرية، جامعة غرداية /الجزائر

zahouani.marwa@univ-ghardaia.dz

تاريخ القبول: 2026/06/ 30

تاريخ الاستلام: 2024/05/ 24

ملخص:

يركز الذكاء الاصطناعي (AI) على دراسة وتطوير الآلات والبرامج الذكية إضافة إلى جمع المعرفة، التخطيط بذكاء، التعلم، التواصل، الإدراك والتعامل مع الأشياء بصفة ذكية. ومن خلال أبحاث تكنولوجيا المعلومات والاتصالات ذات الصلة التقنية والمتخصصة للغاية، فالبيانات الضخمة هي عمل ضخمة، لكن امتلاكها والقدرة الحاسوبية اللازمة للمعالجة لا يكاد يكفي لتوليد نتائج ذات مغزى. فاستخراجها والتعلم الآلي خلق قيمة لقادة الأعمال والممارسين فأصبحت تعتبر مورد كامل للمديرين التنفيذيين في شتى المجالات الذين يتطلعون إلى تحقيق نتائج حقيقية تصل إلى المحصلة النهائية.

توصلت هذه الدراسة الى أن الذكاء الاصطناعي يسمح لمستخدمي البيانات الضخمة بأتمتة وتعزيز المهام التحليلية الوصفية والتنبؤية المعقدة التي يؤديها البشر، تقليل المهام الصعبة والشاقة التي تأخذ وقتا وحيزا كبيرا، إطلاق العنان للذكاء الاصطناعي في البيانات الضخمة يكون له تأثير كبير على الدور الذي تلعبه البيانات في تحديد كيفية عملنا وكيفية سفرنا وكيفية إدارة أعمالنا.

كلمات مفتاحية: بيانات ضخمة، ذكاء اصطناعي، تعلم آلي، تكنولوجيا معلومات

Abstract:

Artificial intelligence (AI) focuses on developing intelligent machines and software for knowledge gathering, planning, learning, communication, perception, and manipulation. However, big data acquisition and processing power are insufficient for meaningful results.

Extraction and machine learning have become valuable resources for business leaders and practitioners.

This study found that artificial intelligence allows users of large data to automate and enhance the complex descriptive and predictive analytical tasks performed by humans, reduce difficult tasks that take time and space, unleash artificial intelligence in big data has a significant impact on the role that data play in determining how we work, how we travel and how we manage our business.

Keywords: Artificial intelligence, big data, machine learning, information technology.

مقدمة:

تزداد أهمية الذكاء الاصطناعي والبيانات الضخمة في مجالات بحثية واسعة، غالبًا ما يُنظر إليها على أنها أساس لبحوث ريادة الأعمال، مثل الاقتصاد، السياسة، الابتكار، الإدارة وعلم النفس. وبالمثل، تعمل ثورة البيانات هذه أيضًا على تعطيل مجالات التطبيق المرتبطة بها، بما في ذلك الصناعة وإدارة الأعمال والابتكار.

وفي أيامنا هذه، أصبحت البيانات الضخمة مصدرًا مهمًا للفرص بين تطبيقات الذكاء الاصطناعي، حيث يتم إنشاء كميات هائلة جدًا من البيانات بوتيرة متسارعة، وتوفر المصادر العديدة للبيانات تنوعًا كبيرًا لها، إذا تم تسخير كل هذه البيانات بذكاء. حيث يمكن أن تدرك حقًا فكرة عصر المعلومات، وجمع المعلومات القابلة للتنفيذ من البيانات بعد إجراء المعالجة والتحليلات الذكية على البيانات المتاحة.

وعليه أدت الزيادة السريعة في البيانات الرقمية منذ أواخر التسعينيات إلى دفع الباحثين في علوم الكمبيوتر وفي البحوث الأخرى بمختلف الميادين إلى تغيير الطريقة التي ينظرون بها إلى العالم ويحللون بها. فبالنسبة للبيانات الضخمة المتنامية بشكل كبير هناك قضية مهمة تتمحور حول طرق جديدة لجمع البيانات والوصول إليها ومشاركتها وتخزينها وتحليلها وتقديمها، حيث يتم تحليل البيانات الضخمة بشكل عام باستخدام أساليب الذكاء الاصطناعي.

وبناء على ما سبق نطرح الاشكالية التالية:

فيما تتمثل البيانات الضخمة؟ وما علاقتها بالذكاء الاصطناعي؟

للإجابة على هذه الإشكالية تم الاعتماد على المنهج الوصفي، وذلك من خلال وصف الأطر النظرية لمفهوم البيانات الضخمة وعلاقتها بالذكاء الاصطناعي، كما تم تقسيم هذه الورقة البحثية إلى مجموعة العناصر التالية:

- تعريف الذكاء الاصطناعي؛
- تعريف البيانات الضخمة (Big Data) ونشأتها؛
- خصائص وتطبيقات البيانات الضخمة؛
- مصادر وأنواع البيانات الضخمة؛
- العوامل المساهمة في النمو وتطوير البيانات الضخمة؛
- الذكاء الاصطناعي وعلاقته بالبيانات الضخمة.

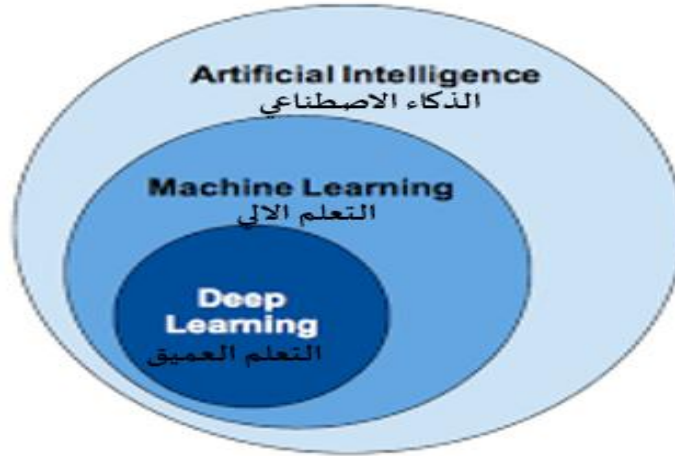
2. تعريف الذكاء الاصطناعي:

إن أول من صاغ ووضع مصطلح الذكاء الاصطناعي في 1956 م، هو جون مكارثي John McCarthy العالم الأمريكي، حيث عرفه بأنه «the science and engineering of making intelligent machines»، أي هو علم وهندسة صناعة الآلات الذكية، وخاصة برامج الحاسوب الذكية، وهو فرع من فروع علوم الحاسوب الذي يهدف إلى إنشاء الآلات الذكية. والذكاء الطبيعي كمفهوم يصعب تعريفه بدقة، ويمكن اعتباره الجزء الحسابي الذي يعطينا القدرة على تحقيق الأهداف في العالم من حولنا، وبقدرة الخالق تختلف درجات الذكاء بين الناس، وكذلك الحيوانات وبعض الآلات. (ال قاسم، بدون سنة نشر، صفحة 4)

فيعتبر تقنية مزدهرة للعديد من التطبيقات الذكية في مختلف المجالات كالمركبات ذاتية القيادة (السيارات ذاتية القيادة والطائرات بدون طيار)، التشخيص الطبي، الرعاية الصحية عن بُعد، أنظمة الأمن السيبراني، الخدمات المصرفية للذكاء الاصطناعي، معالجة الصور واللغة الطبيعية المعالجة في الرؤية الحاسوبية والتصنيف التحويلي في الاتصالات اللاسلكية. وعليه يتفرع فرعين رئيسيين من الذكاء الاصطناعي: التعلم الآلي (ML) والتعلم العميق (DL)، والشكل رقم (01) يوضح أدناه فروع الذكاء الاصطناعي، حيث يشير ML إلى القدرة

- على التعلم واستخراج أنماط ذات مغزى من البيانات، ويعتمد أداء الخوارزميات والأنظمة المستندة إلى **ML** بشكل كبير على الميزات التمثيلية. في غضون ذلك، تستطيع **DL** حلها أنظمة معقدة من خلال التعلم من التمثيلات البسيطة. وفقًا لـ **DL** له ميزتان رئيسيتان:
- تُظهر القدرة على تعلم التمثيلات الصحيحة ميزة واحدة لـ **DL**؛
 - تسمح **DL** للنظام بتعلم البيانات بطريقة عميقة، ويتم استخدام طبقات متعددة بشكل متسلسل لتعلم تمثيلات ذات مغزى بشكل متزايد. (Pham , Nguyen , & Huynh-The, 2020, p. 4)

الشكل رقم (01): يوضح فروع الذكاء الاصطناعي



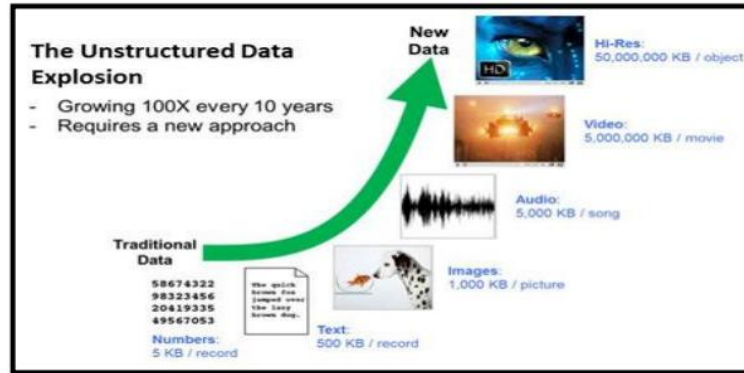
المصدر: (Naseeb, 2020, p. 1)

يعد كل من الذكاء الاصطناعي والبيانات الضخمة تقنيات رائدة في حد ذاتها. حيث تلقت البيانات الضخمة بالذكاء الاصطناعي لخلق إمكانيات جديدة، وعندما يلتقيان يكمل الاثنان بعضهما البعض، مما يساعدنا في تحليل وتنفيذ مجموعات البيانات الكبيرة بطرق فريدة وغير مستكشفة. وعليه توفر البيانات الضخمة حلولاً لجمع البيانات وتخزينها بطريقة قوية، بينما يتعلم الذكاء الاصطناعي، من خلال تقنيات التعلم الآلي، من مجموعات البيانات لاتخاذ قرارات أفضل في المستقبل.

3. تعريف البيانات الضخمة (Big Data):

تقترن الزيادة في البيانات التي تنتجها الشركات والأفراد والعلماء والجهات الفاعلة العامة وظهور الويب والشبكات وتوزيع المستشعرات والهواتف الذكية وغيرها من الكائنات المتصلة بأدوات الكمبيوتر ووجهات نظر جديدة لـ "التحليل". يصعب أحياناً إدارة هذه المعلومات الزائدة لأننا لم نعد نتحدث عن بضعة غيغا بايت (10^9 بايت) نتحدث الآن أكثر عن تيرابايت (10^{12} بايت)، بيتابايت (10^{15} بايت) و **exa bytes** (10^{18}) بايت و **bytes zetta** (10^{21})، لذلك يصبح من الضروري استخدام أساليب ووسائل جديدة لتسهيل إدارة هذه البيانات وهذا ما يعرف بالبيانات الضخمة "BIG DATA". (HADJARI , BENBACHIR , & BOUKHATEM , 2017, p. 17)

الشكل رقم(02): يمثل انفجار البيانات غير المهيكلة



المصدر: (HADJARI, BENBACHIR , & BOUKHATEM , 2017, p. 18)

إن مصطلح البيانات الضخمة هو عبارة عن النمو الأسّي للبيانات، أو بمعنى آخر هي جميع المراحل التي تنطوي عليها عملية استخراج المعلومات المفيدة من مجموعات ضخمة من البيانات الخام. (Allemand, 2018)

كما تشير البيانات الضخمة إلى حجم كبير جداً من البيانات غير المتجانسة في كثير من الأحيان والتي تحتوي على العديد من الأشكال والتنسيقات (النص، وبيانات المستشعر، والصوت، والفيديو، والبيانات على المسار، وملفات السجل، وما إلى ذلك)، بما في ذلك التنسيقات غير المتجانسة، البيانات الضخمة لها طبيعة معقدة تتطلب تقنيات قوية وخوارزميات متقدمة لمعالجتها وتخزينها من خلال تطبيقات نماذج الذكاء الاصطناعي. وبالتالي، لا يمكن معالجتها باستخدام أدوات مثل نظم إدارة قواعد البيانات التقليدية. (Saadna , 2020, p. 4)

1.3 نشأة البيانات الضخمة: لخصتها (HADJARI, BENBACHIR , & BOUKHATEM , 2017, p. 17):

يعد استخدام مصطلح **BIG DATA** لتأهيل دراسة البيانات بكميات كبيرة، أي تتطلب المنافسة بين العديد من أجهزة الكمبيوتر الحالية القوية حديثاً جداً، ولم يتم تعميمه إلا في حوالي 2009، وأول بحث لهذا الموضوع يعود إلى أواخر التسعينيات، ومع ذلك، فإن فكرة البيانات الكبيرة يجب أن تكون نسبية وفقاً للأوقات؛

منذ آلاف السنين، استخدم الإنسان بيانات مباشرة من الطبيعة لوصف الظواهر التي لاحظها، وقبل بضع مئات من السنين، تم وضع نظريات الرياضيات والفيزياء وإضفاء الطابع الرسمي عليها لإعطاء مجموعة من الصيغ لعدة عقود، فأصبحت محاكاة الظواهر المعقدة ممكنة بفضل تطور التقنيات وعلوم الكمبيوتر، ثم تأتي الحسابات العددية والمحاكاة؛

ظهر مصطلح "**BIG DATA**" أخيراً في أكتوبر 1997 في المكتبة الرقمية لـ **ACM Association for Computing Machinery** (جمعية آلات الحوسبة)، وخلال مؤتمر في أكتوبر 1997 تناول التحديات التكنولوجية المستقبلية، استخدم اثنان من الباحثين من وكالة ناسا مايكل كوكس ودايفيد ايليسوورث **David Ellsworth** مصطلح "**BIG DATA**" لأول مرة لوصف تحدي حقيقي وتحديات السنوات القادمة في مقال يشرحون ويحددون هذه الفكرة بوضوح: "البيانات الكبيرة هي حجم هائل من البيانات، وهذا يعني جميع البيانات التي تم جمعها دائماً؛"

سمح المؤتمر الثاني الذي عقد في أكتوبر 1999، حول موضوع تصور البيانات، من قبل **IEEE** (معهد مهندسي الكهرباء والإلكترونيات) لجمهور أوسع بالوصول إلى هذا المفهوم الجديد وطرح أسئلة حول تأثيراته. وفي العقد الأول من القرن الحادي والعشرين، عندما ظهر الإكسابايت (10^{18} بايت) في القياس الكمي للبيانات التي يتم إنتاجها سنوياً، تم تسليط الضوء على قيمة **BIG DATA**، أولاً بالنسبة لفوائد البحث في مجالات الفيزياء أو البيولوجيا أو العلوم الاجتماعية، وأدى ظهور مواقع **Facebook**،

Linkedin، Amazon و Twitter في العقد الأول من القرن الحادي والعشرين، وبشكل أكثر تحديداً اعتباراً من عام 2005، إلى الكشف عن الإمكانيات الكاملة للبيانات العامة المتاحة على الإنترنت؛

في عام 2006: تم إنشاء **Hadoop: Hadoop** بحيث يمكن للأنظمة الجديدة إدارة انفجار بيانات الويب، حيث يعد **Hadoop** مجانياً للتنزيل والاستخدام والتحسين، وهو وسيلة مفتوحة تماماً لتخزين ومعالجة البيانات التي "تسمح بمعالجة متوازية وموزعة لأحجام كبيرة من البيانات على خوادم قياسية وغير مكلفة، وقادرة على تخزين البيانات ومعالجتها، والتطور إلى ما لا نهاية".

4. **خصائص البيانات الضخمة:** يعرف معظم علماء وخبراء البيانات، البيانات الضخمة بمفهوم **Vs** التي تتمثل في بعض الخصائص التي تتميز بها البيانات الضخمة كما يلي: (رشوان، 2018، صفحة 28)

✓ **الحجم (Volume):** تتمثل في حجم البيانات المستخرجة من مصدر ما، وهو ما يحدد قيمة وحجم البيانات لكي تصنف من ضمن البيانات الضخمة، وبحلول العام 2020 أصبح الفضاء الإلكتروني يحتوي على ما يقرب من 40000 ميجابايت من البيانات الجاهزة للتحليل واستخلاص المعلومات؛

✓ **التنوع (Variety):** هي البيانات المستخرجة، والتي تساعد المستخدمين سواء كانوا باحثين أو محليي على اختيار البيانات المناسبة لمجال بحثهم وتتضمن بيانات مهيكلة وغير مهيكلة مثل: الصور ومقاطع وتسجيلات الصوت والفيديو والرسائل القصيرة وسجلات الكلمات وبيانات الخرائط، وتتطلب وقتاً وجهداً لتهيئتها في شكل مناسب للتجهيز والتحليل؛

✓ **السرعة (Velocity):** هي سرعة إنتاج واستخراج البيانات لتغطية الطلب عليها حيث تعتبر السرعة عنصراً حاسماً في اتخاذ القرار بناء على هذه البيانات، وهو الوقت الذي نستغرقه من لحظة وصول هذه البيانات إلى لحظة الخروج بالقرار بناء عليها؛

✓ **الموثوقية (Veracity):** هي موثوقية مصدر البيانات، ومدى دقتها وصحتها وحدائتها، حيث أن هناك مدير تنفيذي من بين كل ثلاثة مدراء لا يثقون في البيانات التي تعرض عليها لاتخاذ القرار؛

✓ **القيمة (value):** للاستفادة من البيانات الضخمة نحتاج إلى متخصصين يمتلكون الخبرات والمهارات الكافية للتعامل مع هذه البيانات وتحليلها التحليل المناسب، وفي هذه الحالة تعتبر المعلومات ذات قيمة؛

- ✓ ذات قيمة متغيرة (**variability**): يقصد أن نفس المعلومات أو نفس البيانات يمكن أن تعني عدة أشياء، واستنادا إلى السياق الذي وردت فيه يمكن تحديد قيمتها الحقيقية وتحليلها تحليلًا مناسبًا؛
- ✓ متعددة المظاهر (**visualization**): إن استخدام البيانات الضخمة يجب تحليلها وإظهارها بأشكال مختلفة تناسب مع طبيعة استخدامها، وتأخذ أشكال متعددة مثل: الاحصاءات والارقام والأشكال الهندسية وغيرها.

1.4 تطبيقات البيانات الضخمة: من بين تطبيقات البيانات الضخمة اختصرها المرجع التالي في: (without, -, pp. 10-11)

- ✓ تحليل ملفات التتبع لاستنتاج السلوك في الماضي والمستقبل: آثار المعاملات التجارية، وتتبع أنشطة الأشخاص المتنقلين، وآثار الوصول إلى خوادم الويب، واستقطاعات سلوك المستهلكين للمسافرين، واستخدام الإنترنت، ينتج عن هذا التحليل بشكل عام نظام توصية لاستهداف العملاء المحتملين بشكل أفضل أو لزيادة أداء الشركة، حيث تم إجراء هذا النوع من التحليل في البداية دون اتصال على كميات كبيرة جدًا من البيانات، ولكن هناك طلبًا متزايدًا على المعالجة عبر الإنترنت لتدفقات البيانات المستمرة؛
- ✓ تحليل إشارات الإخراج من العديد من أجهزة الاستشعار في منشأة صناعية، إلى جانب الوصول إلى قواعد بيانات المواقف السابقة والمواصفات الفنية، من أجل تحديد أماكن الفشل المستقبلي. يجب إجراء هذا التحليل بشكل عام في الوقت الفعلي من الإشارات التي تم التقاطها عند التثبيت قيد الاستخدام؛
- ✓ تحليل الشبكات الاجتماعية، أي تحليل الرسوم البيانية، من أجل استنتاج، على سبيل المثال: العلاقات، التأثيرات بين الأفراد أو مجموعات الأفراد. يُعد تحليل الرسوم البيانية الكبيرة جزءًا محددًا من البيانات الضخمة غالبًا ما يُستشهد به كمثال، وهو ذو أهمية خاصة لبعض عمالقة الويب.

5. مصادر البيانات الضخمة: تعد البيانات الضخمة أيضًا غلافًا لأنواع مختلفة من البيانات الدقيقة. أدناه، ندرج خمسة مصادر

رئيسية للبيانات ذات الحجم الكبير: (Haas, Pentland, & Gerard, 2014, p. 322)

- ✓ البيانات العامة: هي البيانات التي عادة ما تحتفظ بها الحكومات والمنظمات الحكومية والمجتمعات المحلية التي يمكن تسخيرها لتطبيقات الأعمال والإدارة واسعة النطاق. تتضمن أمثلة هذه البيانات تلك المتعلقة بالنقل واستخدام الطاقة والرعاية الصحية التي يتم الوصول إليها بموجب قيود معينة من أجل حماية الخصوصية الفردية؛

✓ البيانات الخاصة: هي البيانات التي تحتفظ بها الشركات الخاصة والمنظمات غير الهادفة للربح والأفراد والتي تعكس معلومات خاصة لا يمكن بسهولة نسبها من مصادر عامة. على سبيل المثال، تشمل البيانات الخاصة معاملات المستهلكين، وعلامات تعريف التردد اللاسلكي التي تستخدمها سلاسل التوريد التنظيمية، وحركة سلع الشركة ومواردها، وتصفح مواقع الويب، واستخدام الهاتف المحمول، من بين عدة أشياء أخرى؛

✓ استنفاد البيانات: يشير إلى البيانات المحيطة التي يتم جمعها بشكل سلمي، والبيانات غير الأساسية ذات القيمة المحدودة أو الصفرية لشريك جمع البيانات الأصلي. تم جمع هذه البيانات لغرض مختلف، ولكن يمكن إعادة دمجها مع مصادر البيانات الأخرى لإنشاء مصادر جديدة للقيمة. عندما يتبنى الأفراد ويستخدمون تقنيات جديدة (مثل الهواتف المحمولة)، فإنهم يولدون البيانات المحيطة بمنتجات ثانوية لأنشطتهم اليومية. قد يقوم الأفراد أيضًا بنقل المعلومات بشكل سلمي أثناء ممارسة حياتهم اليومية (على سبيل المثال، عند إجراء عمليات شراء، حتى في الأسواق غير الرسمية، أو عند وصولهم إلى الرعاية الصحية الأساسية، أو عندما يتفاعلون مع الآخرين). مصدر آخر لاستنفاد البيانات هو سلوك البحث عن المعلومات، والذي يمكن استخدامه لاستنتاج احتياجات الأشخاص أو رغباتهم أو نواياهم. يتضمن ذلك عمليات البحث على الإنترنت أو الخطوط الهاتفية الساخنة أو أنواع أخرى من مراكز الاتصال الخاصة؛

✓ بيانات المجتمع: هي استخلاص للبيانات غير المهيكلة وخاصة النصوص في الشبكات الديناميكية التي تلتقط الاتجاهات الاجتماعية. تتضمن بيانات المجتمع النموذجية تعليقات المستهلكين على المنتجات، وأزرار التصويت (مثل، "أجد هذه المراجعة مفيدة")، وموجزات **Twitter**، يمكن بعد ذلك استخلاص بيانات المجتمع هذه من أجل معنى استنتاج الأنماط في البنية الاجتماعية؛

✓ بيانات التقدير الذاتي: هي أنواع من البيانات التي يكشف عنها الفرد من خلال قياس الإجراءات والسلوكيات الشخصية. على سبيل المثال، الشكل الشائع لبيانات التقدير الذاتي هو ذلك الذي يتم الحصول عليه من خلال الأساور التي تراقب التمرين والحركة، والبيانات التي يتم تحميلها بعد ذلك إلى تطبيق الهاتف المحمول ويمكن بعد ذلك تتبعها وتجميعها. في علم النفس، لدى الأفراد "تفضيلات محددة" لما يرغبون في القيام به مقابل "التفضيلات المكشوفة"، حيث يتم استنتاج تفضيل الفعل أو السلوك.

على سبيل المثال، قد يشتري الفرد مصابيح كهربائية موفرة للطاقة بهدف توفير الكهرباء، ولكن بدلاً من ذلك، أبق الأضواء مضاءة لفترة أطول لأنهم يستخدمون الآن طاقة أقل. تساعد بيانات التقدير الذاتي هذه على سد العلاقة بين علم النفس والسلوك. يمكن لعلماء العلوم الاجتماعية من مختلف المجالات، مثل علم النفس أو التسويق أو السياسة العامة، الاستفادة من بيانات التفضيل المنصوص عليها والضمنية لاستخدامها في أبحاثهم.

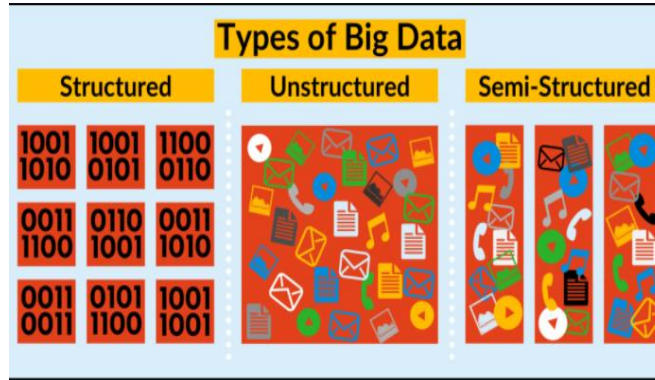
1.5 أنواع البيانات الضخمة: هناك ثلاثة أنواع من البيانات الضخمة، وسنشرحها فيما يلي: (Abhinav , 2020, pp. 3-4)

✓ المهيكلة (المنظمة): هي أحد أنواع البيانات الضخمة، ونعني بالبيانات المنظمة البيانات التي يمكن معالجتها وتخزينها واسترجاعها بتنسيق ثابت. يشير إلى المعلومات عالية التنظيم التي يمكن تخزينها والوصول إليها بسهولة وسلسلة من قاعدة بيانات عن طريق خوارزميات محرك البحث البسيط، كتنظيم جدول الموظف في قاعدة بيانات الشركة حيث ستكون تفاصيل الموظف، ومناصبهم الوظيفية، ورواتبهم... الخ، موجودة بطريقة منظمة؛

✓ غير المهيكلة (غير منظمة): تشير البيانات غير المنظمة إلى البيانات التي تفتقر إلى أي شكل أو هيكل محدد على الإطلاق. هذا يجعل معالجة وتحليل البيانات غير المهيكلة أمراً صعباً للغاية ويستغرق وقتاً طويلاً. البريد الإلكتروني هو مثال على البيانات غير المهيكلة؛

✓ شبه مهيكلة (منظمة): هو النوع الثالث من البيانات الضخمة. تتعلق البيانات شبه المنظمة بالبيانات التي تحتوي على كل من التنسيقات المذكورة أعلاه، أي البيانات المنظمة وغير المنظمة. على وجه الدقة، يشير إلى البيانات التي على الرغم من عدم تصنيفها ضمن مستودع معين (قاعدة بيانات)، إلا أنها تحتوي على معلومات حيوية أو علامات تفصل العناصر الفردية داخل البيانات. ولتوضيح ذلك ندعم شرحنا بالشكل الموالي:

الشكل رقم (03): يوضح أنواع البيانات الضخمة



المصدر: (Allen, 2020, p. 2)

6. العوامل المساهمة في النمو وتطوير البيانات الضخمة:

تغطي البيانات الضخمة تقنيات وأدوات لجمع ومعالجة وتحليل واستخراج المعرفة المفيدة من البيانات المهيكلة وغير المهيكلة بكميات كبيرة تم إنشاؤها بسرعة عالية من خلال مصادر مختلفة، بواسطة أكثر من 2 مليار شخص وملايين المؤسسات الذين يعيشون حياتهم ويقومون بأعمالهم عبر الإنترنت، ومن خلال ملايين أجهزة الاستشعار وأجهزة الاتصال التي ترسل البيانات وتستقبلها عبر الإنترنت، وعليه ترتبط منظمة التعاون الاقتصادي والتنمية (OECD) بنمو البيانات بالعوامل التالية: (Makrufah, himova, & Aybeniz S., 2017, p. 6)

- (1) الانتشار المتزايد للأجهزة الذكية وتطبيقات الاتصالات الذكية (ICT)؛
 - (2) التخفيض الكبير في تكاليف الوصول إلى الإنترنت على مدى السنوات العشرين الماضية؛
 - (3) التخفيض الكبير في تكاليف التخزين. انخفضت تكاليف التخزين إلى النقطة التي يمكن عندها عمومًا الاحتفاظ بالبيانات لفترات طويلة إن لم يكن إلى أجل غير مسمى؛
 - (4) أصبحت أدوات معالجة البيانات قوية ومتطورة وواسعة الانتشار وغير مكلفة بشكل متزايد، مما يجعل البيانات سهلة البحث والربط.
7. الذكاء الاصطناعي وعلاقته بالبيانات الضخمة:

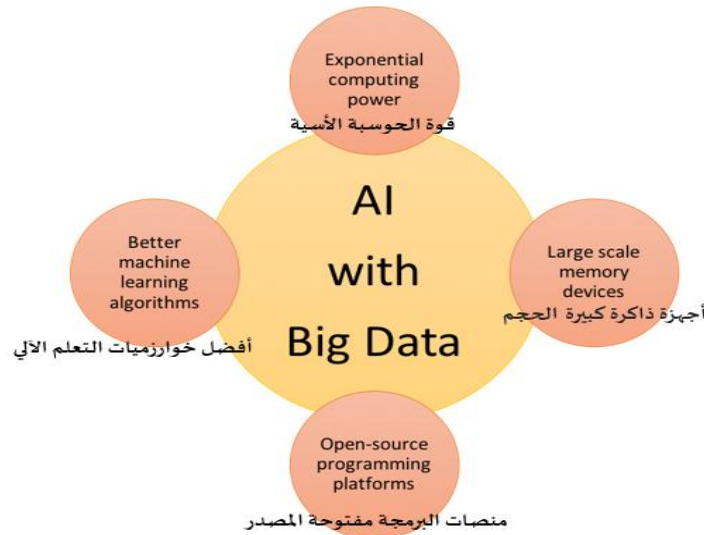
أثبتت البيانات الضخمة قدرتها على دعم كافة المجالات، فمن المحتمل أن توفر هذه الأخيرة عددًا من الحلول الواعدة للمساعدة في مواجهة العضلات الكبيرة من خلال دمجها مع تحليلات الذكاء الاصطناعي، فمثلا ساعدت البيانات الضخمة الكثير من الأطباء على فهم COVID-19 من حيث تتبع تفشي المرض، وهيككل الفيروس، وعلاج الأمراض، وتصنيع اللقاحات. لذلك يمكن للبيانات الضخمة المرتبطة بالأدوات الذكية القائمة على الذكاء الاصطناعي بناء نماذج محاكاة معقدة باستخدام تدفقات البيانات، وسيساعد ذلك محللو البيانات، في إعداد تقارير أفضل، كما تدعم البيانات الضخمة أيضًا التنبؤ المستقبلي للحالات المراد دراستها كالتنبؤ بالأمراض، أسعار الأسهم بالأسواق المالية... الخ، من خلال قدرتها على تجميع البيانات للاستفادة من كميات كبيرة من البيانات للكشف المبكر. علاوة على ذلك، يمكن أن تساعد تحليلات البيانات الضخمة من مجموعة متنوعة من مصادر العالم الحقيقي. (Quoc-Viet, Dinh C, Thien, Won-Joo, & Pubudu N. , 2020, p. 4)

اليوم، تلعب تقنيات التعلم الآلي وخاصة نماذج التعلم العميق، جنبًا إلى جنب مع أجهزة الكمبيوتر القوية دورًا مهمًا في تحليل البيانات الضخمة. يمكن لأساليب التعلم العميق الاستفادة من القوة التنبؤية للبيانات الضخمة في مجالات مثل محركات البحث والطب وعلم الفلك، على عكس مجموعات البيانات التقليدية المستخدمة في نهج جمع البيانات الذي كان خاليًا من الضوضاء، غالبًا ما تكون البيانات الضخمة غير مكتملة بسبب أصولها المتباينة. تجلب البيانات الضخمة إمكانات تحويلية وفرصًا كبيرة لمختلف الحقول، تتطلب خوارزميات التنقيب عن البيانات النموذجية وجود جميع البيانات في الذاكرة الرئيسية، وهذه صعوبة تقنية واضحة للبيانات الضخمة والتي تنتشر عبر مواقع مختلفة. بالإضافة إلى ذلك، تحتاج أساليب التنقيب عن البيانات إلى التغلب على التباين وعدم التجانس وعدم اليقين وعدم اكتمال البيانات الضخمة أيضًا. يعتبر التعلم العميق والبيانات الضخمة بمثابة الصفقات الكبيرة والأسس للابتكار الأمريكي والثورة الاقتصادية. حتى في الحكومة والمجتمع، تظهر البيانات الضخمة كعلاج مفيد لحل بعض المشكلات. (Sohangir, Dingding , & Pomeranets, 2018, p. 7)

يوجد حاليًا نقلة نوعية أساسية في مجال تكنولوجيا المعلومات. في حين أن سلوك الأنظمة المعقدة المستخدمة ليتم برمجتها بشكل مشترك من قبل خبراء المجال وتكنولوجيا المعلومات أو على الأقل تكييفها من البيانات عن طريق افتراضات النماذج الإحصائية المختارة بعناية، فإن التعلم الآلي يثبت حاليًا أنه متفوق في العديد من مجالات التطبيق، لذلك تنشر معظم الشركات البيانات الضخمة والذكاء الاصطناعي لبناء مجموعات بياناتها الحالية ولتطوير آلات يمكنها التفكير بنفسها، لكن البيانات الضخمة هي في الواقع المادة الخام للذكاء

الاصطناعي. لذلك، عندما تلتقي البيانات الضخمة بالذكاء الاصطناعي، فإن لديهم القدرة على تحويل كليهما بالطريقة التي يتم بها تنظيم البيانات والطريقة التي تتعلم بها الآلات. والشكل أدناه يلخص علاقة الذكاء الاصطناعي والبيانات الضخمة:

الشكل رقم (04): يوضح العلاقة بين الذكاء الاصطناعي والبيانات الضخمة لخلق إمكانيات جديدة



المصدر: (JOSHI, 2018)

خاتمة:

المعلومات هي المورد الرئيسي وعامل مهم في العالم الحديث، حيث يشير التطور السريع لهذا المجال في العديد من البلدان المتقدمة والنامية إلى القدرات المحتملة للبيانات الضخمة. ومع ذلك، تعد البيانات الضخمة بطبيعتها مجالا شاسعا للغاية وديناميكيا ومتطورا باستمرار، ولديه القدرة على تغيير الأبحاث بشتى ميادينها، كما تشير البيانات الضخمة إلى أنه يجب علينا تغيير الطريقة التي نجمع بها البيانات ونخزنها ونحللها، مع ضمان إطار أخلاقي يحترم اللوائح الدولية.

النتائج: على ضوء ما سبق في هذه الورقة، تم التوصل إلى بعض النتائج وهي أن:

- الذكاء الاصطناعي يسمح لمستخدمي البيانات الضخمة بأتمتة وتعزيز المهام التحليلية الوصفية والتنبؤية المعقدة التي يؤديها البشر؛
- تقليل المهام الصعبة والشاقة التي تأخذ وقتنا وحيزا كبيرا؛
- إطلاق العنان للذكاء الاصطناعي في البيانات الضخمة حيث يكون له تأثير كبير على الدور الذي تلعبه البيانات في تحديد كيفية عملنا وكيفية سفرنا وكيفية إدارة أعمالنا.

التوصيات: بناء على النتائج التي تم التوصل إليها من خلال هذه الورقة يمكن اقتراح التوصيات التالية:

- ضرورة إتاحة البيانات الضخمة وتطبيقها على نطاق أوسع وعدم حصرها في مجال محدد؛
- الإفصاح والشفافية للبيانات والمعلومات خاصة بالنسبة إلى الدول النامية؛
- تحفيز الدول النامية على استخدام البيانات الضخمة بدلا من الكلاسيكية، والتأقلم مع التطور التكنولوجي وإدراج الذكاء الاصطناعي في شتى المجالات.

قائمة المراجع:

- Allen, R. (2020). What are the Types of Big Data? Retrieved 2022, from selecthub: <https://www.selecthub.com/big-data-analytics/types-of-big-data-analytics/>
- Naseeb, C. (2020, Nov 20). Understanding Artificial Intelligence, Machine Learning, Deep Learning and Data Science. Consulté le 2022, sur medium.com/analytics-vidhya: <https://medium.com/analytics-vidhya/understanding-artificial-intelligence-machine-learning-deep-learning-and-data-science-e7b48e6cb9cf>
- Sohangir, S., Dingding , W., & Pomeranets, A. (2018). Big Data: Deep Learning for financial sentiment analysis . Journal of Big Data, 5(3).
- Abhinav , R. (2020). What is Big Data – Characteristics, Types, Benefits & Examples. upGrad, - (-).
- Allemand, S. (2018). Introduction aux technologies et applications Big Data.

- Haas, G., Pentland, & Gerard , G. (2014). BIG DATA AND MANAGEMENT. Academy of Management Journal, 57(2).
- HADJARI , I., BENBACHIR , M., & BOUKHATEM , F. (2017). Big DATA: Conceptions, architectures, fonctionnements et applications.
- HADJARI, I., BENBACHIR , M., & BOUKHATEM , F. (2017, 06 14). Big DATA : Conceptions, architectures, fonctionnements et applications. 2-96.
- JOSHI, N. (2018, JULY 16). When Big Data met AI. Retrieved 2022, from allerin: <https://www.allerin.com/blog/when-big-data-met-ai>
- Makrufah, S., himova, H., & Aybeniz S. , A. (2017). About Big Data Measurement Methodologies and Indicators, I.J. Modern Education and Computer Science, , 10, p6. -, 10.
- Pham , Q.-V., Nguyen , D., & Huynh-The, T. (2020). Artificial Intelligence (AI) and Big Data for Coronavirus (COVID-19) Pandemic: A Survey on the State-of-the-Arts . w.v, w.n.
- Quoc-Viet, P., Dinh C, N., Thien, H.-T., Won-Joo, H., & Pubudu N. , P. (2020, April 21). Artificial Intelligence (AI) and Big Data for Coronavirus (COVID-19) Pandemic: A Survey on the State-of-the-Arts. NOT PEER-REVIEWED.
- Saadna , y. (2020). Cours Big data et deep learning.
- without, n. (-). définitions et objectifs du cours.

المؤلفات باللغة العربية:

- عبد الرحمن محمد سليمان رشوان . (2018). دور تحليل البيانات الضخمة Data Big في ترشيد اتخاذ القرارات المالية والإدارية في الجامعات الفلسطينية- دراسة ميدانية . مجلة الدراسات الاقتصادية والمالية، 11(1).
- فهد ال قاسم. (بدون سنة نشر). الذكاء الاصطناعي. بدون طبعة.