آليات تطبيق نظم الذكاء الاصطناعي في بيئة البيانات الضخمة:

دراسة وصفية تحليلية

إعـــــــــداد

د/ ياره ماهر محمد قناوي

أستاذ مساعد بقسم المكتبات والمعلومات

جامعة المنيا

ymaher57@yahoo.com

1- تمهيد:

في ظل الابتكارات السريعة للتقنيات الرقمية يتزايد حجم البيانات الرقمية بشكل كبير (Klein, 2017) وبالتالي يتم إنشاء قدر كبير من البيانات متعددة المصادر مثل شبكات التواصل الاجتماعي والهواتف الذكية وأجهزة الاستشعار، وهذا القدر الهائل من البيانات التي لا تستطيع قواعد البيانات العلائقية وتقنيات التحليل تخزينها ومعالجتها يُسمى بالبيانات الضخمة لذلك فإنه يتطلب تطوير أدوات وتقنيات التحليل الجديدة حتى يتسنى اكتشاف الأنماط بين مجموعات البيانات الكبيرة وسرعان ما يتم إنتاج البيانات الضخمة من العديد من المصادر بتنسيقات مختلفة، ومن الآن فصاعدا يجب أن تتمكن الأدوات التحليلية الجديدة من اكتشاف الارتباطات بين البيانات المتغيرة بسرعة لاستغلالها بشكل أفضل.

كما تواجه تقنيات المعالجة التقليدية مشاكل في التعامل مع كمية هائلة من البيانات حيث أنه من الضروري تطوير طرق فعالة لتحليل البيانات في حل المشكلات التي تواجه البيانات الضخمة، سمحت العديد من أُطر البيانات الضخمة مثل سبارك (Spark) وهادوب (Hadoop) بتوزيع الكثير من البيانات وتحليلها (oussous et al, 2018) بالإضافة إلى ذلك تم تقديم أنواع مختلفة من تقنيات الذكاء الاصطناعي مثل التعلم الآلي Machine Learning (ML) والأساليب القائمة على البحث لتقديم نتائج أسرع وأكثر دقة لعملية تحليل البيانات الكبيرة وقد أدى الجمع بين أدوات البيانات الضخمة وتقنيات الذكاء الاصطناعي إلى خلق فرص جديدة في تحليل البيانات الضخمة.

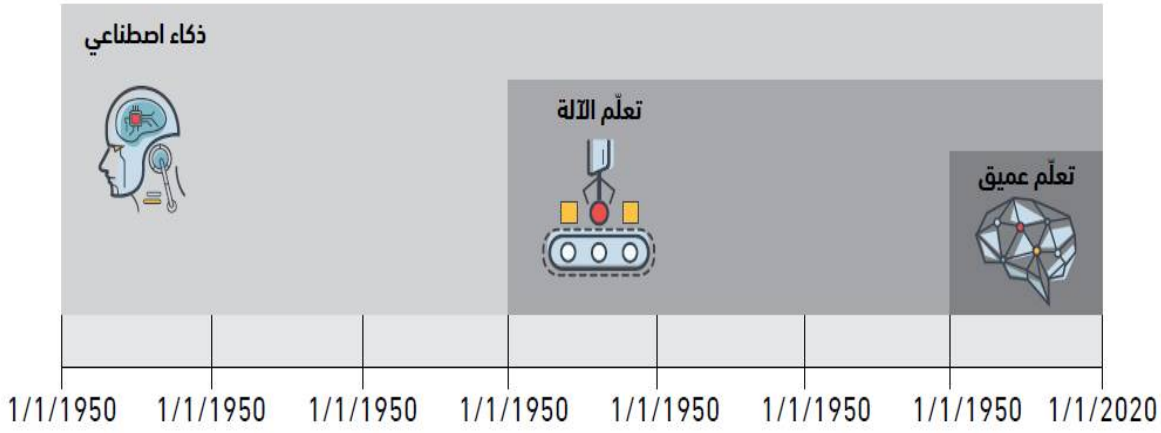
2- مصطلحات الدراسة:

1- الذكاء الاصطناعي Artificial Intelligence

هو مجموعة من التقنيات الفنية المتنوعة تحاكى الذكاء البشري وتستثمر قوة الدماغ الرقمي، والتعلم العميق فرع متقدم من تعلم الآلة الذي هو فرع متطور من الذكاء الاصطناعي عبر السنين.

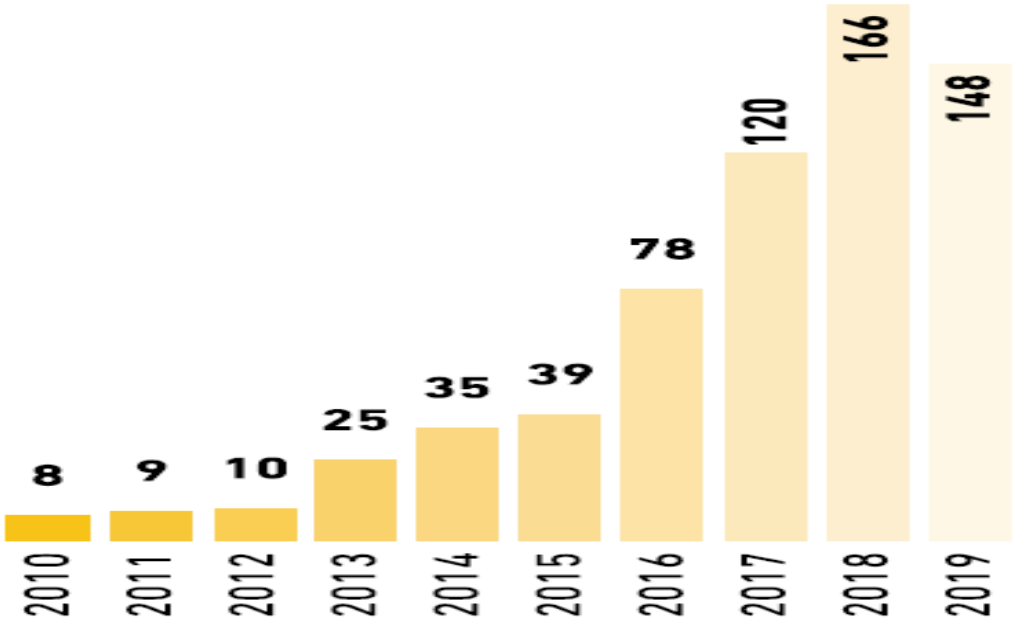
ويمكن تعريف الذكاء على أنه مجال واسع من علوم الحوسبة تجعل الآلات تبدو كأن لديها ذكاء بشري على الرغم من ذلك لسنا قادرين حتى الآن على إنشاء ذكاء اصطناعي يرقى إلى مستوى الذكاء البشري وتبرز أهمية التقنية بقيام شركات التقنية العملاقة (آبل، جوجل، مايكروسوفت، فيسبوك، أمازون) بتأسيس أو الاستحواذ على 635 شركة ذكاء اصطناعي منذ عام 2010 لاستخدامها في منتجاتهم [1].

ويوضح شكل رقم (1) التطور التاريخي للذكاء الاصطناعي حتى الوصول إلى التعلم العميق.



وازداد الدخل الإجمالي العائد من برمجيات الذكاء الاصطناعي من (9.5) بليون دولار في عام 2018 إلى (15) بليون دولار في عام 2019 ومن المتوقع أن يتصاعد العائد إلى حوالي (119) بليون دولار بحلول عام 2025 [2].

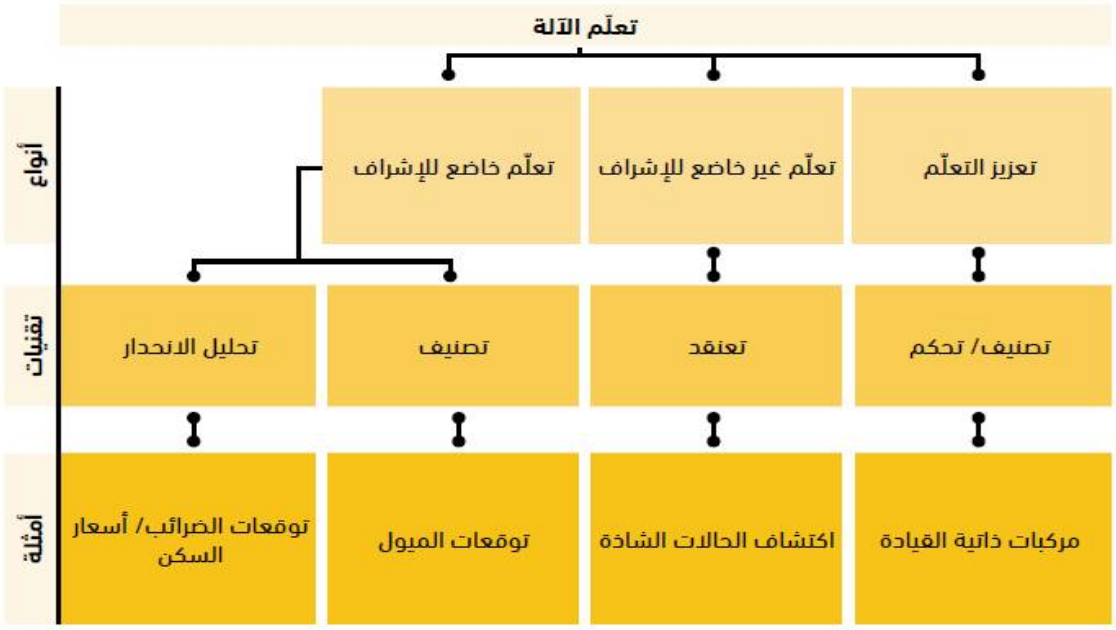
ويوضح شكل رقم (2) نمو وتطور الذكاء الاصطناعي في شركات التقنية العملاقة



2- تعلم الآلة Machine Learning:

يتم تعلم الآلة من خلال تمكين الحواسيب التعلم من البيانات ومخرجات الأنظمة ويعرف على أنه «تقنية لتحليل البيانات والتعلم منها ثم تطبيق ما تعلمته لاتخاذ القرار ويعد تعليم الآلة نوعا من أساليب التنقيب عن البيانات الذي يشمل تقنيات عديدة منها: التعلم العميق والشبكات العصبية الاصطناعية ومعالجة اللغة الطبيعية [3].

ويوضح الشكل (3) تعلم الآلة (الأنواع والتقنيات وأمثلة عليها)



3- منصات البيانات الضخمة Big data Platforms:

تعد معالجة البيانات الآنية والتحليلات التفاعلية منصات مختلفة للبيانات الضخمة تقوم منصات معالجة البيانات بإجراء عمليات حسابية مكثفة وتستغرق وقتاًطويلاً لمعالجة البيانات ويعد أباتشي هادوب (Apache Hadoop) أكثر أنظمة معالجة البيانات شيوعا ويُعد كلاً من نظام الملفات الموزعة هادوب (HDFS) مفاوض آخر للمورد (YARN) ونموذج البرمجة الموزعة Map Reduce هي الوحدات النمطية المختلفة لمنصة هادوب التي تعمل عبر سلسلة قيمة البيانات الضخمة مثل: التجميع، التخزين، الإدارة (Bordo, S& Hassn, 2016).

ويوضح الجدول رقم(1) المصطلحات ذات الصلة بالذكاء الاصطناعي مع البيانات الضخمة [4]

م	المصطلح	التفسير
1	خوارزمية Algorithm	إجراء منهجي ينتج (من خلال عدد محدد من الخطوات) إجابة سؤال أو حل مشكلة.

2	الشبكة العصبية الاصطناعية Artificial Neural Network (ANN)	نموذج اصطناعي خطي حوسبي يستند إلى البنية العصبية للدماغ القادرة على التصنيف والتنبؤ واتخاذ القرارات والتصور من خلال دراسة الأمثلة لتقليد النظام العصبي البشري.
3	تحليلات البيانات الضخمة Big data Analytics	استخدام تقنيات تحليلية متقدمة للبيانات الضخمة هيكلية وشبه هيكلية وغير هيكلية ومن مصادر مختلفة.
4	معالجة اللغة الطبيعية Natural Language Processing (NLP)	فرع من الذكاء الاصطناعي متعدد التخصصات لعلوم الحاسب واللغويات والخوارزميات لتحليل ومعالجة وفهم اللغة البشرية.
5	شبكات عصبية متكررة Recursive Neural Network	نوع من الشبكات العصبية العميقة التي تشكلت من خلال تطبيق نفس مجموعة الأوزان بشكل متكرر على هيكل ما لتنبؤ مهيكل.
6	تحليل الانحدار Regression Analysis	في النمذجة الإحصائية مجموعة من العمليات الإحصائية لتقدير العلاقات بين المتغيرات لتبيان اتجاهات البيانات.
7	ميكنة العمليات روبوتيًا Robotic process Automation (RPA)	استخدام برمجيات روبوتية لأتمته مجموعة من العمليات الإحصائية لتقدير العلاقات بين المتغيرات لبيان اتجاهات البيانات.
8	بيانات مهيكلة Structural data	بيانات يمكن تحديدها ضمن ملف إلكتروني مثل قاعدة بيانات علائقية منظمة في صفوف (سجلات) وأعمدة (حقول)
9	تعلم آلة خاضع للإشراف Supervised Machine Learning	فرع من تعلم الآلة يتم فيها تغذية الخوارزمية بمجموعة من بيانات «التدريب» التي تحتوي على ملاحظات.
10	التعلم غير الخاضع للإشراف Unsupervised learning	فرع من تعلم الآلة يتم فيها تغذية الخوارزمية بمجموعة من بيانات «التدريب» خالية من الملاحظات.
11	بيانات غير مهيكلة Unstructured Data	بيانات لا يمكن تحديدها في مواقع ثابتة أي نص حر في مستندات وتقارير أعمال ومقالات ووسائل التواصل الاجتماعي بيانات معالجة النصوص وملفات PDF ورسائل البريد الإلكتروني ومنتديات الإنترنت والمدونات وصفحات الويب.
12	وكيل (مساعد افتراضي) تشاتبوت	برامج المساعدة الافتراضية التي تتعامل مع المستخدمين باللغة الطبيعية يستخدم دلالات الألفاظ (السيمانتيقا) والتعلم

	Virtual Agent assistant chat bot.	العميق ومعالجة اللغة الطبيعية ونماذج التنبؤ وأتمته المهام من ملاحظة سلوكيات البشر.

3- مبررات اختيار الدراسة وأهميتها:

1- البيانات الضخمة والذكاء الاصطناعي مُكملان بعضهما البعض، فإن البيانات الضخمة تصبح عديمة القيمة بدون برامج لتحليلها.

2- يُعد التنقيب عن البيانات، التحليل الإحصائي، التعليم الآلي، والشبكات العصبية وما إلى ذلك من تقنيات التحليل المختلفة أدوات لإتخاذ قرار أفضل وأسرع بشأن مجموعة البيانات الضخمة.

3- زيادة قدرة تقنيات الذكاء الاصطناعي في التعامل مع مزيد من البيانات بشكل أسرع وزيادة الاهتمام بأنواع أخرى من البيانات وسيصبح التعلم الآلي المرئي (الصوت والفيديو) أشكالاً أكثر أهمية بالنسبة للذكاء الاصطناعي في بيئة البيانات الضخمة.

4- تحظى البيانات الضخمة والذكاء الاصطناعي باهتمام كبير في الوقت الحاضر وتنتشر تطبيقاتهما في كافة أنشطتنا اليومية حيث يتم استخدام الذكاء الاصطناعي لتسهيل جمع وتنظيم البيانات الضخمة.

5- دور الذكاء الاصطناعي والتعلم الآلي في تزويد مؤسسات المعلومات بتحليل ذكي لتلك البيانات والحصول على تفسيرات منظمة لمجموعة واسعة من البيانات غير المهيكلة.

6- يعد هذا البحث الأول في اكتشاف الذكاء الاصطناعي في عصر البيانات الضخمة بالإضافة إلى عدم وجود دراسات تناولت آليات تطبيق الذكاء الاصطناعي في بيئة البيانات الضخمة البيانات .

7- تُعد البيانات الضخمة مع الذكاء الاصطناعي موردًا اقتصاديًا حيويًا ذو أهمية مجتمعية وتسمح نظم الذكاء الاصطناعي للبيانات الضخمة بتحسين الأداء.

8- اهتمام الشركات والمؤسسات والمستشفيات في تنفيذ وتوسيع تحليلات البيانات وقدرات الذكاء الاصطناعي في المستقبل لزيادة حجم البيانات بواسطة أجهزة الحاسب والهواتف الذكية و أجهزة إنترنت الأشياء.

4- مشكلة الدراسة

تتكون البيانات الضخمة من مجموعات بيانات معقدة يمكن أن تتضمن بيانات مهيكلة وبيانات وسائل التواصل الاجتماعي ويمكن للذكاء الاصطناعي التعلم والتكيف بطريقة لم تتمكن أجهزة الحاسب الآلي التقليدية القيام بذلك الأمر، وبدأ جمع كمية هائلة من البيانات ولكن عندما تم صياغة البيانات الضخمة كان حجم البيانات كبير جدا وكان من الصعب تفسيرها الأمر الذي ترتب عليه تطوير خوارزميات الذكاء الاصطناعي وإنشاؤها لاستخلاص رؤى وتتمثل مشكلة الدراسة في فهم كيفية عمل الذكاء الاصطناعي والبيانات الضخمة معاً وفي التعامل مع هذا الكم الهائل من البيانات ويفتح هذا البحث أفقًا لتطوير الحلول التي من شأنها معالجة التحديات المتعلقة بالذكاء الاصطناعي

وتحليلات البيانات الضخمة، كما توجد عدة دراسات حول تحليلات البيانات الضخمة ولكن لا توجد دراسات حول تطبيق آليات الذكاء الاصطناعي في بيئة البيانات الضخمة الأمر الذي دفع الباحثة لتناول هذه الدراسة.

وتكمن مشكلة الدراسة في الإجابة على التساؤلات التالية:

1- ما أنواع تقنيات الذكاء الاصطناعي المطبقة على تحليلات البيانات الضخمة؟

2- ما مزايا تقنيات الذكاء الاصطناعي في بيئة البيانات الضخمة؟

3- كيف يتم استخدام الذكاء الاصطناعي في إطار البيانات الضخمة؟

4- ما متطلبات تنفيذ الذكاء الاصطناعي مع البيانات الضخمة؟

5- ما المجالات الرئيسية لتصنيف الذكاء الاصطناعي في بيئة البيانات الضخمة؟

6- ما آليات تطبيق نظم الذكاء الاصطناعي في بيئة البيانات الضخمة؟

7- ما المؤسسات الأكثر إنتاجية في نشر كل ما يتعلق بالبيانات الضخمة والذكاء الإصطناعي؟

8- ما التحديات الرئيسية والقضايا المفتوحة لتحليلات البيانات الضخمة مع الذكاء الاصطناعي؟

5- أهداف الدراسة:

1- التعرف على تقنيات الذكاء الاصطناعي التي تعتمد على البيانات الضخمة.

2- دراسة مزايا تقنيات الذكاء الاصطناعي في بيئة البيانات الضخمة.

3- التعرف على آليات تطبيق نظم الذكاء الاصطناعي في بيئة البيانات الضخمة.

4- التعرف على كيفية تطبيق نظم الذكاء الاصطناعي في إطار البيانات الضخمة.

5- دراسة متطلبات تنفيذ الذكاء الاصطناعي مع البيانات الضخمة؟

6- تحليل واستكشاف الدراسات البحثية المتعلقة بالذكاء الاصطناعي في بيئة البيانات الضخمة.

7- حصر المؤسسات الأكثر إنتاجية في نشر كل ما يتعلق بالبيانات الضخمة والذكاء الاصطناعي.

8- مناقشة القضايا المفتوحة لتقديم اتجاهات بحثية جديدة في مجالات تحليل البيانات الضخمة مع تقنيات الذكاء الاصطناعي.

6- منهج الدراسة وأدواتها:

* نظرًا لطبيعة الدراسة الحالية وتحقيقًا لأهدافها استخدام المنهج الوصفي التحليلي الذي يسعى إلى تجميع البيانات من الميدان حول الظاهرة وتفسيرها وتحليلها بغرض الكشف عن الأوضاع القائمة بالفعل لمحاولة النهوض بها.

* اعتمدت الدراسة بشكل رئيسي على قائمة مراجعة حول (آليات تطبيق نظم الذكاء الاصطناعي في بيئة البيانات الضخمة، وقد تم اشتقاق أبعاد وعبارات القائمة في ضوء الأبحاث والدراسات مع ما يتناسب مع أهداف الدراسة الحالية

وقد تم استخدام الأساليب الإحصائية الوصفية التالية:

1. التوزيعات التكرارية.

2. النسب المئوية.

3. الرسوم البيانية.

وتضم قائمة المراجعة العناصر التالية (أنواع تقنيات الذكاء الاصطناعي الشائعة في تحليلات البيانات الضخمة، خوارزميات التعلم تحت الإشراف في البيانات الضخمة، مزايا تقنيات الذكاء الاصطناعي مع البيانات الضخمة، كيفية عمل الذكاء الاصطناعي في إطار البيانات الضخمة- المجالات الرئيسية لتطبيق الذكاء الاصطناعي. متطلبات تنفيذ الذكاء الاصطناعي مع البيانات الضخمة- التحديات والقضايا المفتوحة).

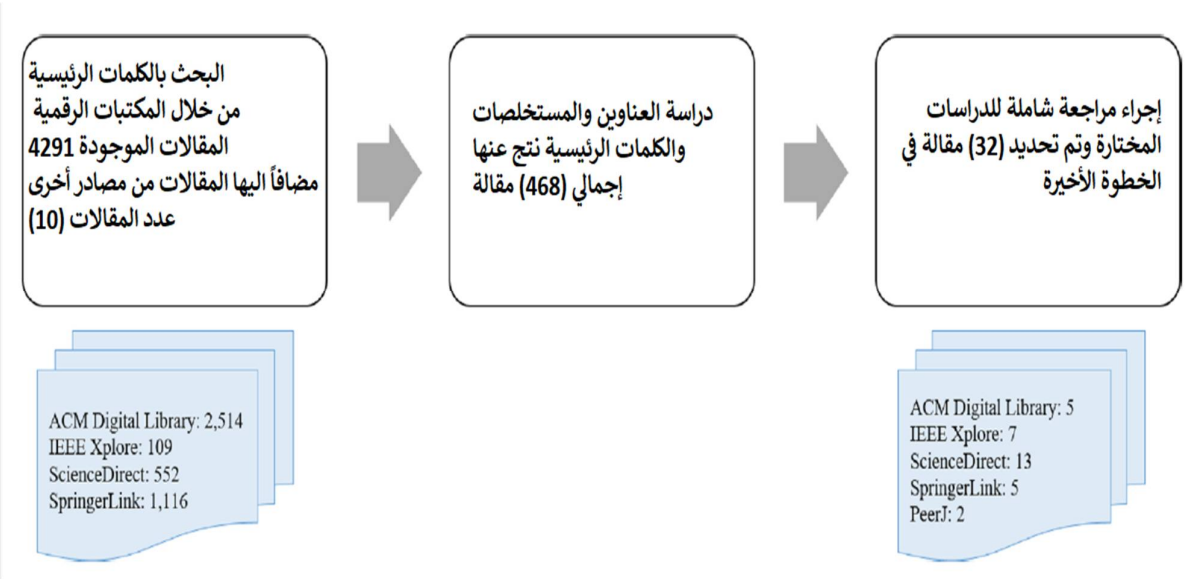
وقد حُكمت قائمة المراجعة من جانب أحد المتخصصين في المجال:

1- أ.د. هاشم فرحات سيد .أستاذ علم المعلومات- جامعة الملك سعود.

2- أ.د. أحمد فرج . أستاذ المكتبات والمعلومات ورئيس قسم المكتبات والمعلومات ،جامعة أسيوط.

1/6 أدوات جمع البيانات

بعد تحديد تساؤلات الدراسة يتم تطبيق بعض المعايير لاختيار الدراسات النهائية كما في الشكل (4) عملية اختيار الدراسات ذات الصلة.



* تم استخدام بعض قواعد البيانات Science Direct, I EEE explore, springer link وتم استبعاد رسائل الماجستير والدكتوراه وأبحاث المؤتمرات بغير اللغة الإنجليزية من الدراسة وتم البحث بالكلمات الرئيسية التالية في الفترة من 2016-2021.

* "الذكاء الاصطناعي" و"تحليلات البيانات الضخمة".

* "التعلم الآلي" و"تحليلات البيانات الضخمة".

* "الشبكات العصبية" و "تحليلات البيانات الضخمة".

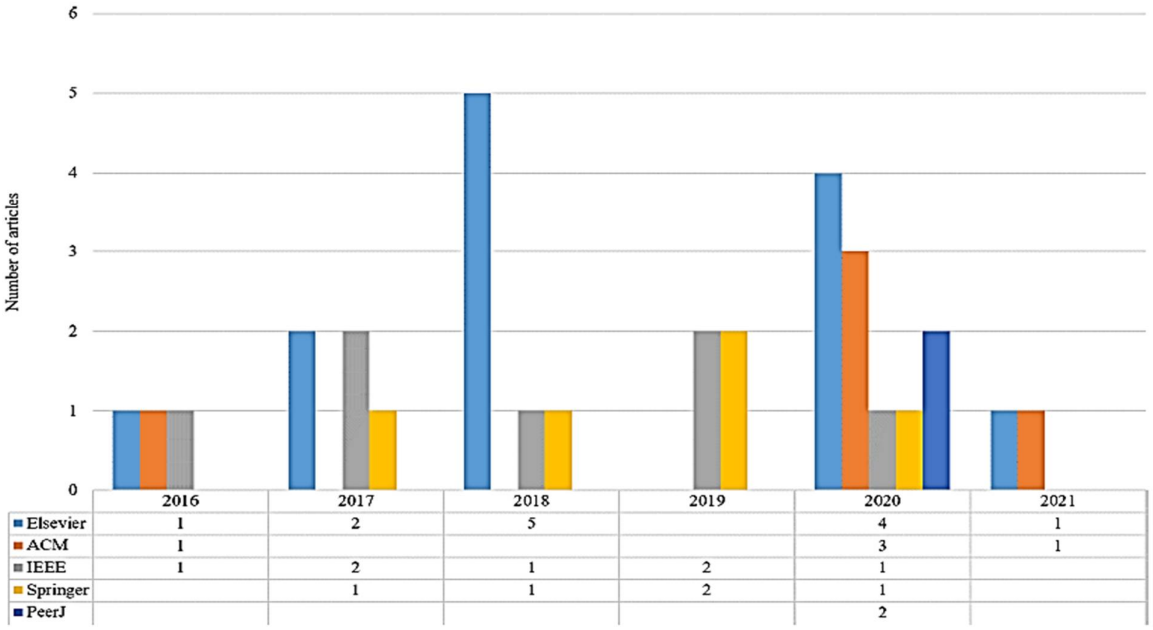
وترتب على نتائج البحث (291) ورقة علمية من خلال استراتيجية البحث بالكلمات الرئيسية وتم إتباع الخطوات التالية:

1. تم استبعاد التسجيلات المتكررة ومراعاة بعض المعايير لاختيار الدراسات عالية الجودة.

2. تم دراسة العناوين والمستخلصات والكلمات الرئيسية نتج عنها إجمالي (468) مقالة.

3. تم إجراء مراجعة شاملة للدراسات المختارة لتأكيد هذه الدراسات وتم تحديد (32) مقالة في الخطوة الأخيرة.

كما يوضح الشكل (5) توزيع المقالات في قواعد البيانات وتاريخ النشر.



* تركز الدراسة على اكتشاف التطورات الأخيرة في مجال البيانات الضخمة المتعلقة بالذكاء الاصطناعي في المنشورات الأكاديمية العلمية وبالبحث في قواعد بيانات Web of science (wos)، قاعدة Thomson Reuters والمجلات مفتوحة المصدر بالبحث حول موضوع: "البيانات الضخمة" "والذكاء الاصطناعي" في الفترة الزمنية من 2016-2021 تم العثور على (107) من المنشورات .

ويوضح الجدول رقم (2) التوزيع النوعي للمواد ذات الصلة بالبيانات الضخمة والذكاء الاصطناعي.

م	نوع الوثيقة	عدد التسجيلات	النسبة
1	مقالات	62	%57.94
2	مراجعات علمية	20	%18.69
3	المواد التحريرية	19	%17.67
4	ورقة عمل	4	%3.74
5	فصل من كتاب	1	%0.93
6	مراجعة لكتاب	1	%0.93

7	أنواع أخرى	7	%6.54

يتضح من الجدول السابق أن أكثر الأنواع استخدامًا من المواد هي «المقالات العلمية» حول موضوع الذكاء الاصطناعي والبيانات الضخمة بإجمالي عدد (62) بنسبة 57.4% كما يوضح الجدول رقم (3) أعلى 10 مقالات تم الاستشهاد بها في الفترة من 2010-2021

العنوان	المؤلفون	مصدر العنوان	عدد مرات الاستشهاد بالمقالة
1- An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data.	Lei, Y, Jia, f (2016)	IEEE on industrial Electronics	(150)
Big Date in Smart Farming-review.	Wolfert, s(2017)	Agricultural Systems	(104)
3- Artificial Intelligence and Big data.	O'leary, D.E (2013)	IEEE intelligent systems	(166)
4- A survey on indexing techniques of big data: Taxonomy and performance evaluation.	Gani, A& Siddiqa, A (2016)	Knowledge and information systems.	(65)
5- Artificial Intelligence in Medicine and cardiac Imaging.	Dilsizian, S.E (2014)	Current Cardiology Reports	(40)
6- Big data analytics for Prostate Radiotherapy.	Coates, J.S (2016)	Frontiers in oncology	(16)
7- 5G smart diabetes: towards personalized diabetes with healthcare Big data a clouds.	Chen, M& Yang, J (2018)	IEEE Communication Magazine	(15)
8- Using Big data and deep learning: toward a mobile system.	Kiral-Kornek, I, (2018)	Ebio medicine	(15)

9- Birth industry 5.0: Making a sense of big data with artificial intelligence.	Ozdemir, V & Hekim, N (2018)	Omics- A journal of integrative Biology	(14)
10- Big data and Machine Learning in Plastic surgery A new frontier in surgical innovation.	Kanevsky, J (2016)	Plastic and Reconstructive surgery.	(14)

7- آليات تطبيق الذكاء الاصطناعي في بيئة البيانات الضخمة: مراجعة علمية

أولاً: الدراسات العربية وتم تقسيمها إلى فئتين هما:

* دراسات حول تحليلات البيانات الضخمة:

1) الكشكي، ناصر أبو زيد (2021)، تحليلات البيانات الضخمة في المؤسسات الأكاديمية، دراسة استشرافية بالتطبيق على مؤسسات التعليم العالي المصرية.

تناولت الدراسة واقع الجامعات المصرية ومدى استعدادها لتبني التحليلات الأكاديمية في إدارة الجامعات وتقترح نموذجًا قابلاً للتطبيق عن طريق إنشاء "وحدة التحليلات الأكاديمية" واستخدمت قائمة المراجعة والمقابلة العلمية كأدوات لجمع البيانات وقد خلصت الدراسة إلى مجموعة من النتائج أهمها:

جميع الجامعات عينة الدراسة غنية بالبيانات الضخمة التي يمكن استثمارها بصورة رائعة عند تحليلها وتوصي الدراسة بتبني وزارة التعليم العالي منهج «الإدارة القائمة على البيانات» وتصميم التحليلات الأكاديمية في كل المؤسسات التابعة لها.

2) أحمد، نجوى شكري (2021). إدارة البيانات الضخمة في دار الكتب القومية: دراسة استكشافية

هدفت الدراسة إلى بيان درجة الوعي بمفهوم البيانات الضخمة وخصائصها في مكتبة دار الكتب القومية والكشف عن استعدادات البنية التحتية والإدارية والتنظيمية لإدارتها وكان من أهم نتائج الدراسة: توافر العديد من الاستعدادات الخاصة بالبنية التحتية والإدارية داخل مكتبة دار لكتب القومية وتوصي الدارسة بإيجاد الاستراتيجيات اللازمة التي من شأنها أن تدعم وتعزز استخدام البيانات الضخمة والاستفادة منها داخل مكتبة دار الكتب القومية.

3- محمد، أسماء حسين (2021). واقع استفادة المكتبات الجامعية من تحليلات البيانات الضخمة Big data في تطوير خدمات المعلومات واتخاذ القرار: دراسة حالة لمكتبة كلية الهندسة

هدفت الدراسة إلى تحديد واقع الاستفادة من تحليلات البيانات في مكتبة كلية الهندسة في تطوير خدمات المعلومات المقدمة للمستفيدين ودعم اتخاذ القرار ومن أهم النتائج التي توصلت إليها

الدراسة: تعدد مجالات الاستفادة من تحليلات البيانات الضخمة، حيث ساعدت تحليلات البيانات الضخمة على التنبؤ باحتياجات المستفيدين المتجددة من المكتبة وتوصي الدراسة بالتأكد من وجود مصداقية البيانات الضخمة قبل تحليلها وتنظيمها.

4) الذبياني، علي أحمد (2020). دور مؤشرات الأداء الرئيسية في تحليلات البيانات الضخمة.

هدفت الدراسة إلى معرفة الدور الذي يمكن أن تحقق مؤشرات الأداء الرئيسية في تحليلات البيانات الضخمة واتبعت الدراسة المنهج الوصفي لإجراء هذه الدراسة وتوصلت الدراسة إلى عدد من الاستنتاجات أهمها: يمكن لتكنولوجيا البيانات الضخمة وأدواتها المساعدة في الاستفادة من جمع وتخزين وتوزيع إدارة كمية هائلة من البيانات التي تنظمها المنظمة وفقا لمؤشرات بطاقة الأداء المتوازن.

5) قناوي، يارة ماهر (2020). دور تحليلات البيانات الضخمة في إنترنت الأشياء: دراسة تحليلية مقارنة.

هدفت الدراسة إلى التعرف على علاقة تحليلات البيانات الضخمة بإنترنت الأشياء، وفوائد تحليلات البيانات الضخمة مع إنترنت الأشياء، وعرض لأشهر منصات البيانات الضخمة وإنترنت الأشياء مثل منصات MAPR, Hadoop, Kaa ودراسة المتطلبات الرئيسية لمعالجة تحليلات البيانات الضخمة في إنترنت الأشياء واعتمدت الدراسة على المنهج المقارن وتوصلت إلى عدة نتائج أهمها، تعدد الأمان والخصوصية من المشكلات الرئيسية التي تواجه إنترنت الأشياء مع البيانات الضخمة بنسبة 97.2%، وتوصي الدراسة بضرورة إجراء المزيد من الدراسات حول خصوصية المواقع في عصر إنترنت الأشياء والبيانات الضخمة.

6- عبد القوي، أماني (2021). متطلبات إدارة البيانات الضخمة في مراكز المعلومات بالمملكة العربية السعودية: دراسة حالة.

هدفت الدراسة التعرف على متطلبات إدارة البيانات الضخمة في المركز الوطني للمعلومات الصحية واستخدمت المنهج الوصفي (دراسة الحالة) وتوصلت الدراسة إلى عدة نتائج أهمها: عدم وجود خطة استراتيجية خاصة بالبيانات الضخمة في المركز واستحداث إدارة خاصة تهتم بشؤون البيانات الضخمة في المركز، والاستفادة من أدوات إدارة البيانات الضخمة مفتوحة المصدر مثل: هادوب Hadoop مع مراعاة إمكانات المركز ونوع البيانات التي يضمها.

<u>ثانيًا: دراسات حول الذكاء الاصطناعي</u>

1) فرج، أحمد (2021). تقنيات الذكاء الاصطناعي وتأثيراتها على فاعلية محتوي وخدمات مؤسسات المعلومات.

يتناول العرض ملامح تطور تقنيات الثورات الصناعية الأربع وصولاً لتطبيقات الذكاء الاصطناعي ويستهدف العرض المرئي محاولة إلقاء الضوء على الرؤى والآفاق المستقبلية للتوجهات الجارية لتطبيقات الذكاء الاصطناعي وبيان تأثيراتها على فعالية إدارة المحتوي الرقمي وخدمات مؤسسات المعلومات، كما يستهدف العرض بيان سبل توظيف تقنيات الذكاء الاصطناعي في إدارة المحتوى الرقمي.

2) متولى عبد الله حسين (2021). تقنية التعرف على الوجوه كأحد تطبيقات الذكاء الاصطناعي في المكتبات ومؤسسات المعلومات إعادة قراءة لعلم الفراسة.

يتناول هذا العرض المرئي تقنية التعرف على الوجوه Face Recognition Technology التي تعتمد على الذكاء الاصطناعي ويتم في هذا العرض تناول التعريف بتقنية التعرف على الوجوه وطريقة عملها مع التركيز على المكتبات ومؤسسات المعلومات واستخدمت الدراسة منهج البحث الاستكشافي للحصول على البيانات والمعلومات ومن نتائج الدراسة: استخدام هذه التقنية يساعد مؤسسات المعلومات في الارتقاء بجودة الخدمات المقدمة بها وتحقيق درجة أعلى من تأمين مجموعة المكتبة والمعلومات الخاصة بالمستفيدين.

3) سميع، ليلى سيد (2021). تحليل المشاعر (الآراء) المعتمد على الذكاء الاصطناعي لتقييم فعالية خدمات المعلومات.

تناقش الدراسة أسلوب تحليل المشاعر أو ما يعرف بالتنقيب عن الآراء كأحد تطبيقات الذكاء الاصطناعي التي تتعلق بمعالجة اللغة الطبيعية وسبل تطبيقه كمؤشر لتقييم خدمات المعلومات من خلال قياس رضاء المستفيدين من خدمات المعلومات اعتمادا على التحليل الآلي لتعليقات المستفيدين المنشورة عبر وسائل التواصل الاجتماعي، وقد أظهرت نتيجة التحليل المبدئي للعينة الاستكشافية قدرة ذلك الأسلوب في تقديم تحليل تفصيلي موضوعي لتعليقات المستفيدين، كما أظهرت الدراسة مجموعة من التحديات المتعلقة بطبيعة اللغة العربية ولهجاتها عند معالجة اللغة الطبيعية.

4) ناجي، إهداء صلاح (2021). تطبيقات نظم الذكاء الاصطناعي في تحليل المحتوى وعمليات التكشيف: دراسة تطبيقية لنظم معالجة اللغة الطبيعية.

هدفت الدراسة إلى التعرف على تطبيقات نظم الذكاء الاصطناعي في مجال تحليل المحتوي وعمليات التكشيف ورصد منصات ومكتبات الذكاء الاصطناعي التي يمكن الاستفادة منها في مجال تحليل المحتوى مع العمل على تطبيق أحد نظم معالجة اللغة الطبيعية واستخدامها في تحليل مصادر المعلومات لتحقيق الإفادة منها واعتمدت الدراسة على المنهج الوصفي التحليلي.

5) غلوم، حسين فولاز (2021). دور الذكاء الاصطناعي في دعم وتطوير خدمات المكتبات الأكاديمية في دولة الكويت: إدارة المكتبات بجامعة الكويت أنموذجًا.

هدفت الدراسة إلى معرفة الدور الذي تقوم به تقنيات الذكاء الاصطناعي في دعم وتطوير مستقبل خدمات المكتبات الأكاديمية واعتمدت الدراسة على المنهج الوصفي والمقابلة الشخصية مع أصحاب القرار والمسئولين في إدارة المكتبات بجامعة الكويت لمعرفة مدى استخدام تطبيقات الذكاء الاصطناعي في دعم الخدمات الإلكترونية في مكتباتها.

وأوضحت نتائج الدراسة أن هناك قصور وضعف في استخدامات الذكاء الاصطناعي في مكتبات جامعة الكويت بشكل ملحوظ، كما أشارت نتائج الدراسة أن هناك قصور وضعف في استخدامات الذكاء الاصطناعي كان من أبرزها قلة وعي العاملين، وقلة الميزانية والتي لا تتناسب مع التكلفة الباهظة لتطبيقات الذكاء الاصطناعي وتوصى الدراسة بوضع خطط واستراتيجيات واضحة من

قبل المسئولين وأصحاب القرار عن كيفية تطبيق واستخدام أفضل لتطبيقات الذكاء الاصطناعي في المكتبات الأكاديمية في دولة الكويت.

6) العريشي، جبريل بن حسن (2021). استخدام البيانات الضخمة والذكاء الاصطناعي في مواجهة جائحة كورونا . هدفت الدراسة إلي تسليط الضوء علي العديد من التحديات والقضايا المرتبطة باستخدام تقنيات البيانات الضخمة والذكاء الاصطناعي في تطوير الحلول التي أسهمت في مكافحة كوفيد ١٩ وتوصلت الدراسة إلي عدة نتائج من أهمها : ظهر دور البيانات الضخمة والذكاء الاصطناعي في بناء نماذج تنبؤيه تفيد في الإنذار المبكر باحتمال انتشار الفيروس وهو ما يفيد الحكومات والهيئات الصحية في مراقبة تفشي الفيروس مستقبلاً وتوصي الدراسة بضرورة وضع آليات للتخفيف من قواعد خصوصية البيانات التي قد تحول دون التطور السريع لخوارزميات الذكاء الاصطناعي في وقت انتشار الأوبئة.

ثانيًا: الدراسات الأجنبية وتم تقسيمها إلى عدة فئات:

1. دراسات حول إدارة البيانات الضخمة.

2. دراسات حول تقنيات تحليل البيانات الضخمة.

3.دراسات حول منصات البيانات الضخمة.

4. دراسات حول تطبيقات تحليلات البيانات الضخمة.

5. دراسات تجمع بين الذكاء الاصطناعي والبيانات الضخمة.

6. دراسات حول التعلم الآلي والتعلم العميق.

7. دراسات حول إنشاء البيانات المهيكلة .

8. دراسات حول نظم الرعاية الصحية الذكية باستخدام التعلم الآلي.

1- دراسات حول إدارة البيانات الضخمة:

* قدم **(Siddiqa et al, 2016)** لمحة أساسية حول تقنيات إدارة البيانات الضخمة وتقديم تصنيف مفصل بناءً على التخزين، والمعالجة والأمان وتمت مناقشة مقالات مختلفة في كل فئة بالإضافة إلى مزايا الطرق المقترحة في هذه الدارسة.

* قام المؤلفون **(Tsai et al, 2015)** بمراجعة العديد من الدراسات المتعمقة بتحليل البيانات الضخمة التقليدية والحديثة، يعتبر إجراء اكتشاف المعرفة والتنقيب عن البيانات الذي يتضمن التحليل والمخرجات أساسًا لهذه الدراسات حيث تم مناقشة العديد من تقنيات استخراج البيانات والبيانات الضخمة مثل التجميع والتصنيف، وتم اقتراح بعض القضايا المفتوحة والاتجاهات البحثية المستقبلية لتوفير طرق فعالة وتوصلت نتائج الدراسة إلى أن جميع الدراسات السابقة قامت بالتركيز على فئة التعلم الآلي لتقنيات الذكاء الاصطناعي ولم تهتم بدراسة فئات الذكاء الاصطناعي الأخرى مثل الذكاء الحسابي.

2- دراسات حول تقنيات تحليل البيانات الضخمة:

* قام **(Chani et al, 2019)** بمراجعة مناهج تحليلات الوسائط الاجتماعية في أربع فئات:

الشبكات العصبية، الحوسبة الضبابية، سرب الاستخبارات، التعلم العميق وقام المؤلفون بتقسيم التقنيات التي تمت مراجعتها بناءً على مقاييس الجودة الخاصة بهم ومع ذلك لا يوجد إجراء منهجي لاختبار المقالات المتعلقة بهذا المجال.

* وركز كلاً من **(Mittal, S et al, 2019)** على دراسة تقنيات التعلم الآلي لتحليل البيانات الضخمة ويتم مراعاة ثلاث فئات رئيسية والتي تشمل التعلم الخاضع للإشراف، والتعلم غير الخاضع للإشراف والتعلم المعزز ومع ذلك لا توجد طريقة واضحة لاختيار المادة ولم يتم تقييم الدراسات بناءً على معايير الجودة.

* قدمكلاً من **(Athmaja, H, Kavitha, 2017)** مراجعة منهجية قائمة على الأدبيات لنهج تحليلات البيانات الضخمة وفقا لآليات التعلم الآلي.

* قدم **(sivarajah et al, 2017)** عمل آخر يتعلق بتقنيات تحليل البيانات الضخمة وصنف المؤلفون هذه التقنيات إلى ثلاث مجموعات رئيسية بما في ذلك التحليلات الوصفية والتنبؤية والتعليمية ومع ذلك هناك بعض الثغرات في تحليل البيانات وعملية اختيار الدراسة.

* قدم **(Qiu et al, 2016)** عرضا موجزًا لتقنيات التعلم الآلي حيث قام بالتركيز على بعض أساليب التعلم الحديثة مثل: التعلم العميق، التعلم الموازي، التعلم النشط، التعلم عن طريق الخوارزميات كجزء من التعلم الآلي وركزت الدراسة فقط على تقنيات التعلم الآلي ولم يتم إجراء مقارنة فيما يتعلق بالطرق المقترحة.

3- دراسات حول منصات البيانات الضخمة:

* قام **(oussous et al, 2018)** بدراسة تأثير تحديات البيانات الضخمة وأدوات عديدة لتحليلها حيث تم مناقشة الأدوات المستخدمة في معالجة البيانات الضخمة فضلاً عن ذلك تنقسم تحديات تحليلات البيانات الضخمة إلى ست فئات: إدارة البيانات الضخمة، تنقية البيانات الضخمة، جمع البيانات الضخمة، البيانات الضخمة غير المتوازنة، تحليلات البيانات الضخمة، والتعلم الآلي للبيانات الضخمة.

* ناقش **(NiColalde et al, 2018)** الجهود البحثية نحو تقنيات معالجة البيانات الضخمة حيث ناقش المؤلفون بعض التحديات ذات الصلة مثل: تخزين البيانات، وتحليلها، اكتشاف المعرفة، قابلية التوسع وتصور البيانات وأمن المعلومات.

4- دراسات حول تطبيقات تحليلات البيانات الضخمة:

* درس **(Vaishya et al, 2020)** التطبيقات الرئيسية للذكاء الاصطناعي للوقاية من جائحة فيروس كورونا (كوفيد-19) وتضم سبعة تطبيقات للذكاء الاصطناعي لجائحة كوفيد 19:

1. الكشف علن المرض 2. مراقبة علاج المريض 3. تتبع جهات الاتصال

4. توقع الحالات والوفيات 5. إنتاج الأدوية 6. تقليل أعباء العمل

7. الوقاية من الأمراض

* ناقش **(Phamcav et al,2020)** تطبيق تقنيات الذكاء الاصطناعي والبيانات الضخمة لإدارة وتحليل الحجم الهائل من العيانات الناشئة عن فيروس كورونا وأويتم التعرف في خمس فئات لمراجعة تقنيات البيانات الضخمة المختارة والتي تشمل التوقع بتفشي كوفيد 19- تتبع انتشار الفيروس، التشخيص والعلاج، اكتشاف الأدوية.

5- دراسات تجمع بين الذكاء الاصطناعي والبيانات الضخمة

* دراسة **(Chen, M, 2018)** عن تحليلات البيانات الضخمة المتعلقة بالمجال الطبي وإنترنت الأشياء فضلاً عن التقنيات التي من الممكن تبنيها وتطبيقها.

* وتركز دراسة أخرى **(Jing, y, 2018)** في مجال العلوم الطبية على التعلم العميق باستخدام البيانات الضخمة في نظام الهاتف المحمول كشفت نتائج الدراسة أنه يتم دمجه مع الأجهزة العصبية والتي تزود نظام تخدير في الوقت الحقيقي ومخصص للمريض مع أداء موثوق به مدى الحياة.

* وأجريت أبحاث مماثلة **(li,y & zhang, 2018)** على مرض السكري بالإضافة إلى شبكات 5G.

* دراسة **(Griffen et al, 2018)** عن البيانات الضخمة والذكاء الاصطناعي في مجال الكيمياء الطبية وركزت الدراسة على تأمين تبادل المعرفة كمهمة حاسمة للملكية الفكرية للكيميائي الطبي باستخدام التقنيات المناسبة للبيانات الضخمة والذكاء الاصطناعي والبرامج ذات الصلة.

درس كلاً من **(Ozemir, V.H, 2018)** مفاهيم الثورة الصناعية الخامسة 5.0 من خلال مراجعة البيانات الضخمة، والذكاء الاصطناعي، إنترنت الأشياء، نظام التشغيل الآلي (الروبوت) وشبكة الواي فاي واسعة النطاق، وأوصت الدراسة أن الثورة الصناعية الخامسة 5.0 يمكنها إضفاء الطابع الديمقراطي على إنتاج مشاركة المعرفة للبيانات الضخمة.

* دراسة **(syeda, M,2018)** عن دور البيانات الضخمة والتعلم الآلي في اتخاذ قرار التشخيص في علم الأشعة حيث كان الهدف هو دعم متخصص التصوير بتحسين دقة وكفاءة التشخيص والعلاج باستخدام المعدات والطرق الحديثة وتم استخدام وتحليل البيانات الضخمة المدروسة في نظرية SP وهي تبسيط ودمج المفاهيم عبر الذكاء الاصطناعي وفقًا لإدارة وعمليات المعرفة.

* دراسة **(Zhuang, F, 2017)** في علوم وهندسة المواد وهي اكتشاف المواد الجديدة((NOMAD)ركز البحث على مبادرات وفتح آفاقًا جديدة لعلوم التعدين باستخدام البيانات الضخمة.

* دراسة **(Kiral- Korrek, I, 2017)** في مجال أبحاث الأورام الطبية وبالتحديد العلاج الإشعاعي للبروستاتا أظهرت نتائج الدراسة أن البيانات الضخمة أفضل مع دمج أنواع البيانات غير المهيكلة للمعايير السريرية.

* اقترح **(Lei et al, 2016)** طريقة تعلم من مرحلتين للتشخيص الذكي للآلات مما يقلل من الحاجة للعمالة البشرية ويجعل تشخيص الأخطاء الذكي يتعامل مع البيانات الضخمة بسهولة أكبر.

* درس **(Gani et al, 2016)** تقنيات الفهرسة الحالية للبيانات الضخمة حيث اقترح تصنيفا لتقنيات الفهرسة تضم الفئات التالية: الذكاء غير الاصطناعي، الذكاء الاصطناعي وطرق فهرسة الذكاء الاصطناعي التعاونية.

* قدم بحث**(Ahmed, f,2021**) بعنوان «الذكاء الاصطناعي في الطب وتصوير القلب، يركز على تسخير البيانات الضخمة وممارسات الحاسب المتقدمة لتشخيص واستخدام العلاج.

* دراسة **(Kanasky et al, 2016)** عن البيانات الضخمة والتعلم الآلي فيما يتعلق بمجال الجراحة التجميلية من حيث الابتكارات في العلوم الطبية وكشفت نتائج الدراسة أن جراح التجميل الحديث للتعلم الآلي يُحدد كحقل فرعي من الذكاء الاصطناعي ويدعم معالجة المشاكل المختلفة في الجراحات التجميلية.

* في بحث آخر يتناول البيانات الضخمة حتى المعرفة في الذكاء الاصطناعي 2.0 يتناول الفرص والتحديات مع التطورات التكنولوجية الناشئة للذكاء الاصطناعي في إعدادات البيانات الضخمة وأظهرت نتائج الدراسة أن التعلم الآلي القائم على البيانات (جزء من الذكاء الاصطناعي**(Abbass, H.F, 2016)**

* بحث **(O, Leary, 2013)** عن الأنظمة الذكية وقام بدراسة بعض الاهتمامات الأساسية واستخدامات الذكاء الاصطناعي في البيانات الضخمة.

6- دراسات حول التعلم الآلي والتعلم العميق:

* طور المؤلفون **(Tian, et al, 2020)** اكتشاف في السلوك غير الطبيعي لتجنب تسرب البيانات الضخمة في الشبكات الحضرية واقترح طرقا مختلفة للتعليم العميق لتحليل الانحرافات بين الإجراءات الواقعية وطبيعة الأنشطة اليومية ووفقًا لنتائج التجربة يمكن للطريقة المقترحة معرفة النمط الطبيعي للسلوك وتحديد الأنشطة غير الطبيعية بدقة عالية.

* قدم كلاً من **(Hammou, l & Mouline, 2020)** تقنية فعالة جديدة لتحليل المشاعر وقد حاول المؤلفون اعتماد مكتبة fast text مع متغيرات الشبكة العصبية المتكررة (RNN) لتمثيل البيانات النصية وتصنيفها وتم اقتراح نظام موزع يعتد على التعليم الآلي لتحليلات البيانات الضخمة.

* تتزايد حركة تصفح الإنترنت بسرعة في عصر البيانات الضخمة للوسائط المتعددة لذلك تعد معالجة البيانات والتحميل الزائد من التحديات الرئيسية في هذا السياق اقترح

(Wang et al, 2021) نموذج تيار هجين لحل هذه التحديات لتحليل الفيديو الذي يحتوي على وحدات معالجة البيانات وتصنيفها وتقليل حمل البيانات تم تطوير نسخة معدلة من طريقة الشبكة العصبية الإلتفافية (CNN).

وأكدت نتائج الدراسة أن النموذج المقترح يقلل من تحميل البيانات، ويتحكم في حجم إدخال الفيديو ويقلل من الحمل الزائد على الشبكة وتم ملاحظة أن النموذج يتمتع بأداء جيد للنمو المستمر لبيانات الوسائط المتعددة الكبيرة مقارنة بالنماذج التقليدية الأخرى.

* قدم **(Kannan et al, 2019)** نهجًا تنبؤي فيما يتعلق ببيانات قيمة العملات باستخدام آلة ناقلات الدعم ويتم إجراء التحليل الدلالي في المرحلة الثانية ويتم رد فعل المجتمع على حل المشكلات باستخدام التحليل التنبؤي.

* اقترح **(Feng et al, 2019)** العديد من طرق البحث عن البيانات، طرق التعلم العميق والتنبؤ بالبيانات الجنائية اكتشفت المؤلفون حقائق وأنماط مختلفة مثيرة للاهتمام من البيانات الجنائية لمجموعات بيانات شيكاغو وسان فرانسيسكو وفيلادلفيا.

* اقترح **(Cracillo et al, 2018)** منصة جديدة لاكتشاف الاحتيال (Scarff) تستخدم المنصة المقترحة أدوات البيانات الضخمة من نوع Kafka, Spark, and Cassadra بالإضافة إلى تقنية التعلم الآلي لمعالجة البيانات وتشير نتائج الدراسة أن الكفاءة، الدقة، قابلية التوسع في إطار العمل المقدم مرضية.

<u>7- دراسات حول البيانات المهيكلة :</u>

* اتخذت هيكلة البيانات مناهج متعددة وضع كلاً من **(Philip Hayes&steven, H, 2016)** نظاماً لاستخدامه في خدمة رويترز الإخبارية للمساعدة في تصنيف المقالات الإخبارية يصنف الناتج المقالات الإخبارية غير المهيكلة إلى حوالي (700) فئة بالتعرف على أكثر من (1700) اسم شركة بدقة تصل إلى 85% وبدأ الباحثون في إنشاء تحليل وجهات النظر غير المهيكلة المتوفرة في المدونات ورسائل تويتر والنصوص الأخرى.

* وفي تطبيق آخر بدأت الشركات في دراسة تأثير قضايا البيانات غير المهيكلة لسمعة الشركة استعرض **(O, Leary, D.E, 2012)** وزملاؤه كيف تقوم بعض الشركات بتحليل مجموعة من الأنواع المختلفة من البيانات لتوفير مراقبة مستمرة لمجموعة من الأنشطة بما في ذلك إنشاء مقاييس مهيكلة تقسيمات لسمعة الشركات والمنتجات والتحقق من قضايا البيانات المالية في الكشف عن الاحتيال.

*قام باحثو الذكاء الاصطناعي بمراعاة إنشاء التطبيقات التي تحلل البيانات غير المهيكلة بحيث يمكن استخدام المعلومات الناتجة مباشرة لفهم عملية ما أو للتفاعل مع التطبيقات الآخر على سبيل المثال وجد كلاً من **(Bollen, J, 2011)** أن توقعات سوق الأسهم الخاصة بمتوسط مؤشر Dow Jones قد تحسنت من خلال مراعاة وجهات النظر لسوق الأوراق المالية هذا مفهوم غير منظم ولكنه يعتمد على البيانات المنظمة التي تم إنشاؤها من جوجل.

* قام **(Bohen, J, 2011)** وزملاؤه بفحص رسائل تويتر ووجدوا أن علامات التصنيف والرموز التعبيرية كانت مفيدة في تأكيد وجهات النظر بمجرد هيكلة البيانات تبدأ الشركات البحث عن البيانات لتطوير رؤى حول أنواع البيانات الضخمة.

<u>8- دراسات حول نظم الرعاية الصحية الذكية باستخدام التعلم الآلي:</u>

* اقترح كلا من **(Ahmed, F & Keshtor, 2021)** الشبكة العصبية الالتفافية (CNN) العميقة على صور الأشعة السينية لتحديد فيروس كورونا (كوفيد19) حققت الطريقة المقترحة دقة تصنيف البيانات بنسبة 90.64%.

* اقترح **(Ahmed, I, 2021)** إطارًا جديدًا لتحليل فيروس كورونا (كوفيد19) والتنبؤ به باستخدام تحليلات البيانات الضخمة وإنترنت الأشياء حيث تم تطوير الإطار المقترح على أساس الشبكات العصبية ووفقًا لنتائج التجربة يتمتع الإطار المقترح بأداء جيد بدقة تصل إلى 99% مقارنة بأساليب التعلم الآلي التقليدية.

* اقترح **(Ragab & Attallah, 2020)** نظامًا جديدًا للتشخيص بمساعدة الحاسب (CAD) معتمدا على تقنيات الذكاء الاصطناعي وكشفت نتائج الدراسة أن دمج كافة العناصر في التشخيص بمساعدة الحاسب يتم التميز بدقة بين صور توضيح فيروس كورونا وغير الواضحة للفيروس كوفيد 19.

* واقترح **(Al-Kabbany, 2020)** نموذجًا يعتمد على الشبكات العصبية العميقة لتحليل عمليات المسح الطبي لفيروس كورونا كوفيد-19.

* طور **(alzubi, 2020)** تقنيات جديدة للبيانات الضخمة وطرق التعلم الآلي لتحديد مرض السكري ويتم استخدام منصة MAP Reduce لدمج جزء صغير من البيانات بكفاءة أظهرت نتائج الدراسة أن الشبكة العصبية توفر دقة عالية ومعدل خطأ أقل.

* قدم **(Nair, Shetty, 2018)** نظامًا جديدًا للتنبؤ بالحالة الصحية من خلال تطبيق نماذج التعلم الآلي على تحليلات البيانات الضخمة حث تم إنشاء النظام باستخدام Apache Spark ونشره في البيئة السحابية يرسل المستخدم المواصفات الصحية الخاصة به ويتنبأ النظام بالحالة الصحية للمستخدم في الوقت الفعلي.

* اقترح **(Kaur, p.s, 2018)** نموذجًا جديدًا مبتكرًا لأنظمة معلومات الرعاية الصحية الذكية باستخدام خوارزميات التعلم الآلي يشمل النموذج على أربع طبقات: تتعامل طبقة مصدر البيانات مع مصادر البيانات غير المتجانسة، طبقة تخزين البيانات لعملية تحسين تخزين البيانات وإدارتها واستخدام تقنيات الأمان والخصوصية مثل إخفاء البيانات، التحكم الدقيق في الوصول للبيانات وأخيرًا تعد طرق التعلم الآلي المستخدمة في التشخيص المبكر للمرض أفضل طبقة لتحسين دقة النموذج المقترح باستخدام الحوسبة الضبابية.

<u>اتجاهات الدراسة السابقة وأهميتها للموضوع:</u>

1- لم تقدم العديد من المقالات دراسات كاملة حول تقنيات الذكاء الاصطناعي في سياق البيانات الضخمة وهذا يتفق مع نتائج الدراسة **(Mittal, s 2019).**

2- لم تفحص بعض الدراسات إجراءات اختيار الدراسة كما توجد بعض الثغرات في تحليل البيانات وهذا يتفق مع نتائج دراسة **(Sivarajah et al, 2017).**

3- ركزت غالبية الدراسات على تقنيات التعلم الآلي ولم يتم إجراء مقارنة فيما يتعلق بالطرق الأخرى المقترحة مثل (التعلم الموازي- التعلم النشط) وهذا يتفق مع نتائج دراسة **(Qiu et al, 2016).**

4- جميع الدراسات السابقة قامت بالتركيز على فئة التعلم الآلي لتقنيات الذكاء الاصطناعي ولم تتم دراسة فئات الذكاء الاصطناعي الأخرى مثل (الذكاء الحسابي) وهذا يتفق مع نتائج دراسة **Tsai) etal, 2016).**

5- قام المؤلفون بتقييم التقنيات التي تمت مراجعتها بناءً على مقاييس الجودة الخاصة بهم، ومع ذلك لا يوجد إجراء منهجي لاختيار المقالات المتعلقة بهذا المجال وهذا يتفق مع نتائج دراسة **(Chani etal, 2019).**

6- أغلب الدراسات توصلت إلى أن التعلم الآلي القائم على البيانات جزء من الذكاء الاصطناعي ويحول الذكاء الاصطناعي 2.0 المقترح البيانات الضخمة إلى معرفة مهيكلة وهذا يتفق مع نتائج دراسة (**Ozdemir, v, 2018- Abass, H, 2016).**

7- على الرغم من أن فترة الدراسة تتراوح من 2016-2021 إلا أن غالبية الدراسات نشرت في عام 2018 وهذا يتفق مع نتائج دراسة **(Kulakli, A, 2020)** دفعت الأسباب السابقة الباحثة إلى إعداد هذه الدراسة حول آليات تطبيق الذكاء الاصطناعي في بيئة البيانات الضخمة للتغلب على تلك المشكلات.

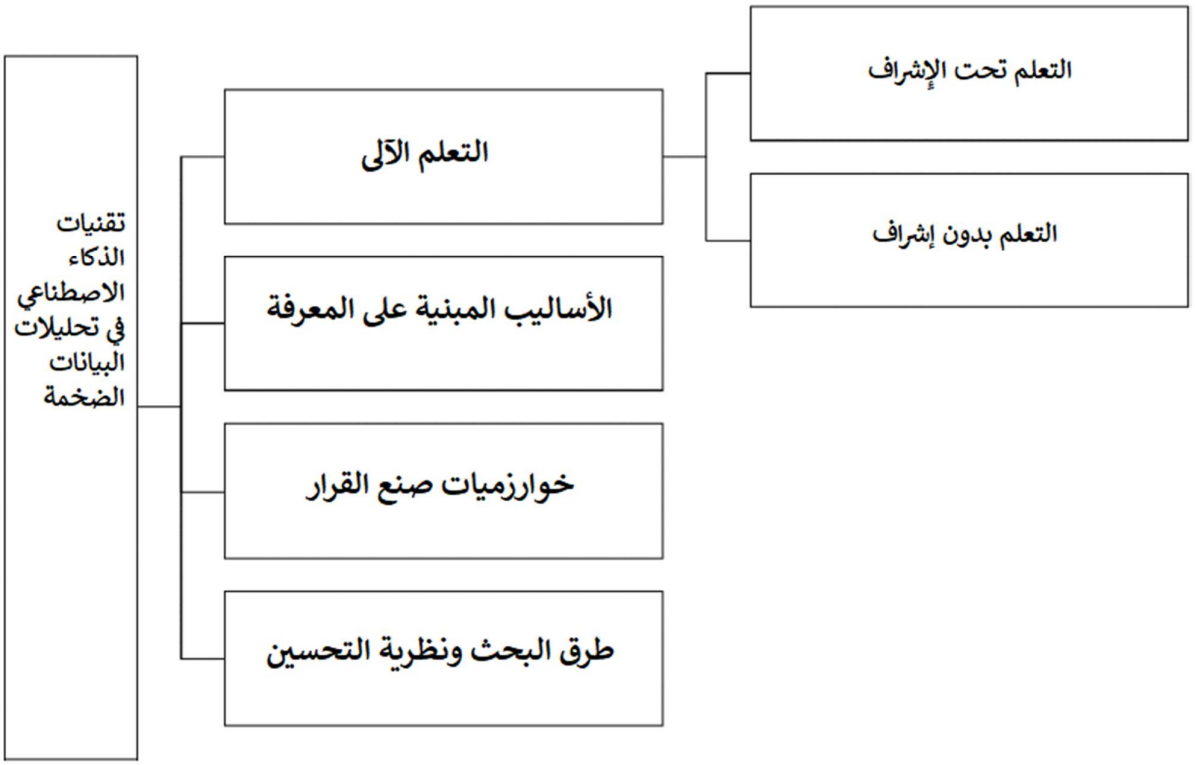
أولاً: الجانب النظري

1) آليات تحليل البيانات الضخمة القائمة على الذكاء الاصطناعي:

يتم إجراء تصنيف ومراجعة حول دراسات تحليلات البيانات الضخمة بناءً على الحقول الفرعية للذكاء الاصطناعي، يحتوي التصنيف المقدم على أربع فئات رئيسية:

التعلم الآلي- الأساليب القائمة على المعرفة والاستدلال، خوارزميات اتخاذ القرار، طرق البحث

كما يوضح بالشكل رقم (6) خطة تصنيف الذكاء الاصطناعي في بيئة البيانات الضخمة.



مصدر الشكل : (Kotsiantis, Z,2007)

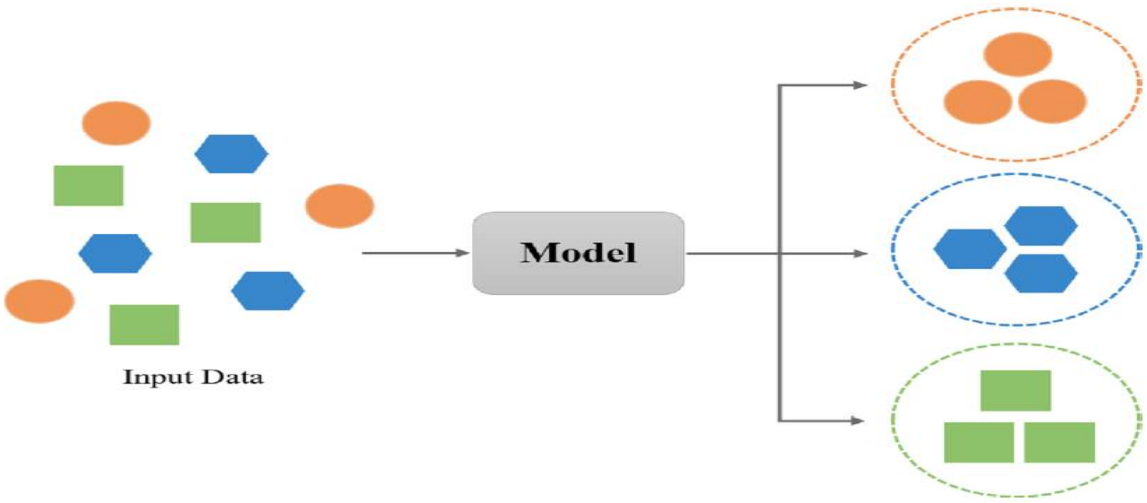
أ- آليات التعليم الآلي:

يمكن تقسيم خوارزميات التعلم الآلي إلى فئتين رئيستين هما: التعلم تحت الإشراف، والتعلم بدون إشراف.

- التعلم تحت الإشراف: Supervised Learning

يُعد الهدف من خوارزمية التعلم تحت الإشراف هو توقع العنونة الصحيحة لبيانات الإدخال المقترحة حديثًا من خلال هذه الطريقة يتم تقديم مجموعة من المدخلات والمخرجات وإيجاد العلاقة فيما بينها أثناء تدريب النظام.

ويوضح الشكل (7) نظرة عامة على التعلم تحت الإشراف



ويعرض الجدول رقم (4) آليات تحليلات البيانات الضخمة في التعلم تحت الإشراف.

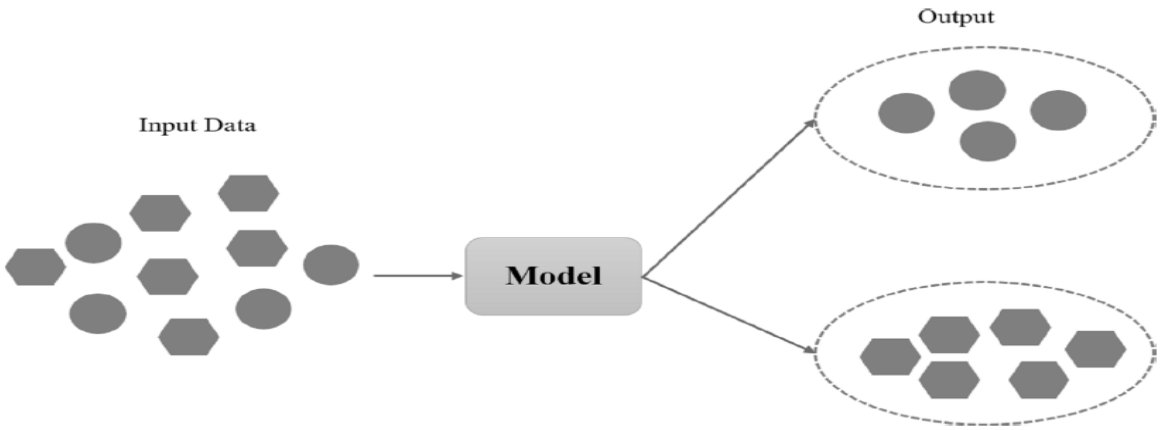
م	تقنيات الذكاء الاصطناعي	قابلية التوسع	الكفاءة	الدقة	الخصوصية
1	تصنيف المجموعة على أساس الخوارزميات العشوائية.	✓	✓	✓	×
2	دعم آلة النقل	×	✓	✓	×
3	الشبكة العصبية (LSTM)	×	✓	✓	×
4	تحليل الانحدار/ شجرة القرار	×	✓	✓	×
5	الشبكة العصبية المتكررة (RNN)	✓	×	✓	×
6	خوارزميات الشبكة العصبية.	×	×	✓	×
7	التعلم العميق DL	×	×	✓	×
8	مجموعة من آلات التعلم المتطرفة	✓	✓	✓	×
9	تحسين مصنف آلة ناقلات الدعم باستخدام CGWO	×	×	✓	×

وبتحليل بيانات الجدول السابق يتضح أن الكفاءة والدقة من أهم العوامل في تطبيق آليات التعلم تحت الإشراف بينما حظيت قابلية التوسع والخصوصية باهتمام أقل من الباحثين.

- التعلم بدون إشراف Un supervised Learning

التعلم غير الخاضع للإشراف هو فرع من تعلم الآلة يتم فيها تغذية الخوارزميات بمجموعة من بيانات التدريب خالية من الملاحظات.

يوضح الشكل (8) نظرة عامة على التعلم بدون إشراف



مصدر الشكل (Banchhor, S, 2020).

ويوضح الجدول (5) آليات تحليلات البيانات الضخمة غير الخاضعة للإشراف

م	تقنيات الذكاء الاصطناعي	قابلية التوسع	الكفاءة	الدقة	الخصوصية
1	التكتل القائم على Centroid	✓	✓	✓	×
2	آلة ناقلات الدعم	×	✓	✓	×
3	حقول Markov العشوائية	✓	×	✓	×
4	أنظمة ضبابية	✓	✓	✓	×

تم عرض آليات تحليل البيانات الضخمة التي تعتمد على الذكاء الاصطناعي وتم وصف أهم تقنيات تحليلات البيانات الضخمة المبنية على الذكاء الاصطناعي حتى عام 2021 فإن التعلم الآلي، الطرق القائمة على المعرفة والاستدلال، خوارزميات اتخاذ القرار، طرق البحث ونظرية المفاضلة هي أربع فئات رئيسية من الذكاء الاصطناعي لتحليلات البيانات الضخمة ويمكن القيام بالمهام المتكررة بمساعدة ذكاء الآلة، ويعد تقليل الخطأ وتحسين درجة الدقة مزايا أخرى لتحليلات البيانات الضخمة التي تعتمد على الذكاء الاصطناعي.

2) كيف تعمل البيانات الضخمة والذكاء الاصطناعي معًا:

البيانات الضخمة والذكاء الاصطناعي مُتشابكان بشكل معقد سيكون الأول أمر مجديًا إذا تم تنفيذه بشكل صحيح، بينما يعتمد الأخير على الأول ليكون ناجحًا بالإضافة إلى القدرة على اتخاذ قرارات أكثر ذكاء والتحكم بشكل أفضل في المخاطر للتعامل مع الأنشطة المختلفة ومعالجتها (Rahupathi, w, 2014).

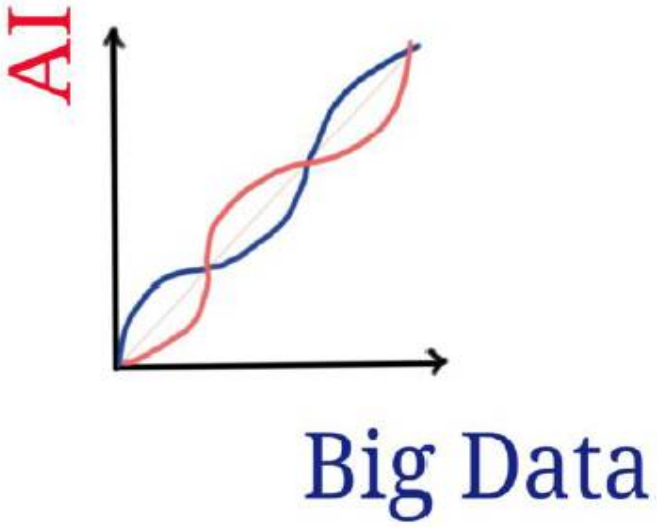
لم يتم اكتشاف البيانات الضخمة والذكاء الاصطناعي بشكل كامل نظرًا لأنها لا تزال كبيرة جدا وقيد التنفيذ من قبل مؤسسات مختلفة

ويوضح الجدول (6) الذكاء الاصطناعي قبل ومع البيانات الضخمة

م	الذكاء الاصطناعي قبل البيانات الضخمة	الذكاء الاصطناعي مع البيانات الضخمة
1	توفر كميات بيانات محدودة (ميجا بايت).	توفر كميات بيانات متزايدة باستمرار (زيتا بايت).
2	عينات محدودة الحجم	عينات ضخمة تسمح بدقة نمذجة أعلى تحليل بيانات كبيرة في ميللي ثانية.
3	عجز عن تحليل بيانات كبيرة في ميللي ثانية.	تحليل بيانات كبيرة في ميللي ثانية.
4	مصادر بيانات محددة.	مصادر بيانات متعددة ومختلفة.
5	غالبا ما تستند إلى بيانات مهيكلة.	تستند إلى بيانات ميهكلة وشبه مهيكلة وغير مهيكلة.

مصدر الجدول: (sultan, A, 2016)

ويوضح الشكل (9) العلاقة بين الذكاء الاصطناعي والبيانات الضخمة



1/2 إدارة البيانات:

يمكن وصف إدارة البيانات بأنها عملية تتضمن جمع البيانات والتحقق منها، تخزينها، تأميتها، معالجتها وتعديلها للتأكد من سهولة الوصول إليها وموثوقيتها وضمان سرعة انتشارها للعديد من المستخدمين، تعد هذه العملية أكثر تعقيدًا ويترتب عليها ارتفاع سرعة البيانات (في الوقت الحقيقي تقريبا مما يترتب عليه إحراز تقدم كبير في مجال نظم الحاسب (Balandin, S, 2013) يمكن للذكاء الاصطناعي المساعدة في هذه العملية من خلال القيام بالسماح بالاقتراحات المخصصة بناءً على التعلم الآلي وباستخدام البيانات في الوقت الفعلي لتحسين تجربة المستخدمين، يمكن أن يساعد الذكاء الاصطناعي في تحديد المعلومات من خلال تدفق البيانات عبر تصنيف معالجة اللغة الطبيعية (NLP) وإنشاء العلاقات، ويمكن لأنظمة التعلم الذاتي عرض مقاطع فيديو ثابتة أو متحركة لاكتشاف المعلومات الجديدة والتعامل معها.

2/2 إدارة الأنماط:

تحتاج كافة أنواع الشركات والمؤسسات إلى البحث باستمرار عن أنماط الاهتمام التي من شأنها إبلاغها بما قد يحدث لاتخاذ إجراءات تم تحديدها ومعالجتها من أجل عمليات التنفيذ يمكن أن يساعد الذكاء الاصطناعي من خلال تحديد مجموعة واسعة من الإشارات، الأحداث، والاتجاهات المتوقعة وغير المتوقعة التي يمكن أن تؤدي إلى التشخيص المبكر لمرض السرطان بمجرد دمجها مع التحليلات، يمكن للذكاء الاصطناعي أن يكشف عن إمكانات الإجابة بالتفصيل (Who, Health,2011) كما يمكن تحديد وتعلم آليات وتحديات القرار والقيام بإجراءات إضافية لتقديم نتائج أسرع وأفضل.

3/2 إدارة السياق:

تتغير مزايا البيانات المجمعة في ضوء السياق الذي يتم فيه تفسير البيانات والنتائج التي يمكن تطبيق هذه البيانات عليها (Larson, E.C, 2012).

حيث يعد استخراج المعلومات ذات الصلة من البيانات أمرًا بالغ الأهمية لفهم مجموعة البيانات نفسها وتساهم المعرفة والعلاقة بين العالم الداخلي والخارجي في فهم السياق ويمكن للذكاء

الاصطناعي (AI) تسهيل عمليات الحاسب المعقدة التي تستخدم «مجالات» البيانات مع سياق واحد (Noctale, V, 2012).

يتمكن الذكاء الاصطناعي استخراج الاختلاف الدقيقة الخاصة بالسياق لمتابعة تفسير البيانات في سياقات مختلفة سواء كان الذكاء الاصطناعي يتفاعل مع المستخدم أم لا وتحتوي تطبيقات معالجة اللغة الطبيعية (NLP) علي عدة أشكال يمكن استخدامها لمراجعة الاتصالات والتفاعلات البشرية.

4/2 إدارة القرار:

تمتلك إدارة القرار كافة جوانب تطوير والحفاظ على آليات اتخاذ القرار التلقائية التي تستخدمها المؤسسات للتعامل فيما بينها داخليًا وخارجيًا، حيث أن برنامج صنع القرار له تأثير كبير على إدارة الشركة لأعمالها لتصبح صانع قرار فعال ومؤثر (Seera, M, 2014) يمكن للذكاء الاصطناعي (AI) أن يلعب دوراً هاماً في تغذية المعلومات واستخدام المهارات في بيئة ديناميكية ومتطورة بشكل مستمر (Daji, H.H, 2012) سيساعد الذكاء الاصطناعي في توسيع نطاق الموارد الرئيسية مثل البيانات الضخمة التي تتزايد بسرعة لمعالجة أهداف العلم وتطبيق مهامها على توقعات العملاء، ستؤدي زيادة استخدام الذكاء الاصطناعي في التفاعلات البشرية إلى تعزيز تجارب المستخدم وفعالية حل المشكلات وتقديم التوصيات حول التحقق من خيارات القرار وتحليل نتائج القرارات والتنبؤ بها فضلا عن تتبع المخرجات بشكل فعال.

5/2 إدارة إجراءات العمل:

إدارة الإجراءات هي عملية وضع الخطط وتنظيم أنشطة العمل والإجراءات لكافة الأشخاص، الروبوتات، العمليات، التطبيقات والتقنيات التي تستخدمها الأعمال التجارية حيث أن الهدف الرئيسي هو تنظيم، ترتيب وتنسيق أي نشاط وتصميم خطط المشروع حسب طلب الرؤساء (Boarouis, A, 2013).

سيقوم الذكاء الاصطناعي بتحليل هذه الأنشطة ومحاولة ربطها بخطوات اتخاذ القرار السابقة أو تعديل المعايير الخاصة أو التوصية بوضع إجراء جديد.

يعد الذكاء الاصطناعي جزءًا مهمًا يساعد على تتبع وتقييم تصرفات وسلوكيات المديرين سيعتبر الذكاء الاصطناعي السلوك المعدل في المستقبل قبل التنفيذ مما يضمن تحقيق النتائج المرجوة بشكل كمي.

6/2 إدارة الأهداف:

تتضمن إدارة الأهداف ممارسة تحديد ومراقبة الأهداف لتوفير المدخلات فضلاً عن المساعدة في تقييم النجاح وتقديم اقتراحات لكافة الأصول لتحسين الأداء (Bouran's, A, 2013) سيساعد الذكاء الاصطناعي في توجيه البشر، الروبوتات، التطبيقات والبيانات الأساسية القابلة للتطوير من خلال التعديل الآلي للأهداف التي تتعقب الظروف المتطورة.

7/2 إدارة المخاطر:

تتضمن إدارة المخاطر تحديد وتقييم وضع الأولويات والحد منها عن طريق الاستخدام المنظم والذكي للموارد لتقليل المخاطر ويشمل ذلك استخدام البيانات الضخمة لتتبع التكرار بشكل مستمر والتعرف على المشاكل والتهديدات الناشئة، قد يدعم الذكاء الاصطناعي مؤسسات المعلومات في تحديد والاستجابة عند حدوث المواقف في الأحداث، الاتجاهات، سجلات النظام والمدخلات الشخصية (مثل مواقع التواصل الاجتماعي) للتهديدات المحتملة بالإضافة إلى ذلك يمكن تحديد المشاكل في وقت مبكرا مثل السلوك الثقافي. (Tsanas,A, 2012)

جدول (7) يوضح متطلبات تنفيذ الذكاء الاصطناعي مع البيانات الضخمة

مصدر الجدول (IREA,2019)

(1) المتطلبات التقنية	(2) المتطلبات التنفيذية (السياسات واللوائح اللازمة)	(3) المتطلبات التنظيمية	(4) المتطلبات الوظيفية والسياسات واللوائح اللازمة)
* **متطلبات الأجهزة** (الشبكات الذكية - العدادات الذكية لتجميع كميات كبيرة من برامج البيانات عالية الجودة والدقة. * البرامج الخاصة بتقنيات الذكاء الاصطناعي (التعلم الآلي- التعلم العميق). * النظام السحابي (إذا لم يتم تخزين البيانات محليا). * كميات هائلة من البيانات الدقيقة **الخبرة البشرية.** * تطوير علماء البيانات	* تقييم تأثير الذكاء الاصطناعي على الوظائف. * خلق فرص جديدة للعمل. * السماح بالوصول العام إلى البيانات بحيث يمكن لأي شخص استخدام أو تطوير التقنيات الرقمية. * تمكين تمويل الأبحاث وتطوير تطبيقات الذكاء الاصطناعي.	* تحديد قواعد خصوصية البيانات للمستخدمين. * إنشاء حوافز للمشاركة في المشاريع التجريبية كمقدمي البيانات. * تحديد بروتوكولات الأمن السيبراني. والبروتوكولات الخاصة بالتشغيل البيني للبيانات الضخمة. * ضمان تمثيل	* **مشغلو النظام** * اعتماد نهج مبتكر لتشغيل النظام من خلال تعزيز التعاون بين مشغلي أنظمة التوزيع ونقل البيانات. * القائمون على جمع المعلومات عبر الإنترنت. * شركاء تكنولوجيا المعلومات والاتصالات. - تطوير حلول الذكاء الاصطناعي لدمج الطاقة المتجددةحدة.VER

خوارزميات التعلم الآلي والتحسين المستمر للنماذج التي يمكن تطبيقها. * فهم التقنيات الرقمية والعمل مع علماء البيانات لتطبيق تقنيات الذكاء الاصطناعي. * التعاون مع مشغلي النظام (خبراء تكنولوجيا المعلومات والاتصالات ICT أو علماء البيانات.		الخوارزميات واللوائح أو تكييفها عند الضرورة.	

4- التحديات والقضايا المفتوحة

ويوضح الجدول (8) التحديات والحلول المقترحة

التحديات	الحلول المقترحة
1- الحوسبة الضبابية: * قدر كبير من البيانات تحتاج إلى تحليل في الوقت الفعلي. * توفير قدر كبير من الحوسبة والتخزين والاتصالات محليا.	* إجراء المزيد من الدراسات لتحليلات بيانات إنترنت الأشياء الكبيرة من خلال بنية الحوسبة الضبابية.

2- معالجة كمية كبيرة من البيانات: * إنتاج البيانات الضخمة من مصادر متعددة وغير متجانسة. * عدم تكامل البيانات. * عدم تنسيق البيانات. * كمية هائلة من البيانات غير المهيكلة وغير الكاملة وغير الدقيقة. * صعوبة تخزين البيانات بالطرق التقليدية.	تنفيذ تقنيات ذكاء اصطناعي مختلفة لتحليل هذه الكميات الضخمة من البيانات في الوقت الفعلي. * التحقق من كفاءة وقابلية التوسع في خوارزميات التحليلات الذكية المطبقة على البيانات الضخمة.
3- الأمان والخصوصية: * صعوبة توافر طريقة آمنة للتعامل مع البيانات الضخمة المجمعة من أنظمة وبيئات مختلفة. * معظم البيانات التي تم جمعها في جميع أنحاء العالم مملوكة من عدد محدود من الشركات. * مجموعات بيانات متحيزة لتغذية خوارزميات التعلم الآلي.	* يجب التعامل مع المسائل المتعلقة بعملية الأمان الخاصة بتحليلات البيانات الضخمة في مجالات مختلفة مثل: حماية أجهزة إنترنت الأشياء من الهجمات تقنيات الذكاء الاصطناعي الآمنة. التواصل الآمن مع الأنظمة الخارجية. * زيادة الدراسات التي تركز على القضايا الأمنية لتحليلات البيانات الضخمة حيث تعد التحديات والتدابير الأمنية خطة بحث مثيرة للاهتمام في المستقبل. * حماية البيانات من خلال التنقيب فيها والحفاظ على الخصوصية.
4- المعايير والمقاييس النوعية * استخدام سمات جودة مختلفة للتحقق من صحة البيانات المقدمة.	* زيادة الدراسات التجريبية حول البنية التحتية الذكية ومعايير الجودة.
5- جودة البيانات * كميات كبيرة من البيانات شبه	* الحاجة إلى المزيد من الدراسات حول مشاكل جودة وتنسيق البيانات غير المهيكلة وشبة المهيكلة. * استخدام الذكاء الاصطناعي لتحديد البيانات المغلوطة

المهيكلة وغير المهيكلة. * تحديات في تنسيق وجودة البيانات. * وجود بيانات مغلوطة مع البيانات الضخمة.	وتصحيحها.
6- القيود المفروضة على خوارزميات الذكاء الاصطناعي * مجموعة الخوارزميات المتاحة غير قياسية. * تفتقر الخوارزميات إلى التوثيق. * تركز تقنيات الذكاء الاصطناعي على تطبيقات الآلة الواحدة في ظل البيانات الضخمة. * صعوبة استخدام الخوارزميات الجينية «الوراثية» في بيئة Map Reduce.	* الحاجة إلى ذكاء اصطناعي قابل للتطوير لمجموعة من الآلات مبنية على منصات Hadoop, Map Reduce. * زيادة الاهتمام بالدراسات التي تدور ق حول تصميم وتنفيذ استخدام الخوارزميات الجينية على برنامجHadoop.

مصدر هذا الجدول : (Rahmani, A, &x AZhir, 2021).

ثانيًا: الجانب التطبيقي

1- أنواع تقنيات الذكاء الاصطناعي الأكثر شيوعاً المستخدمة في تحليلات البيانات الضخمة

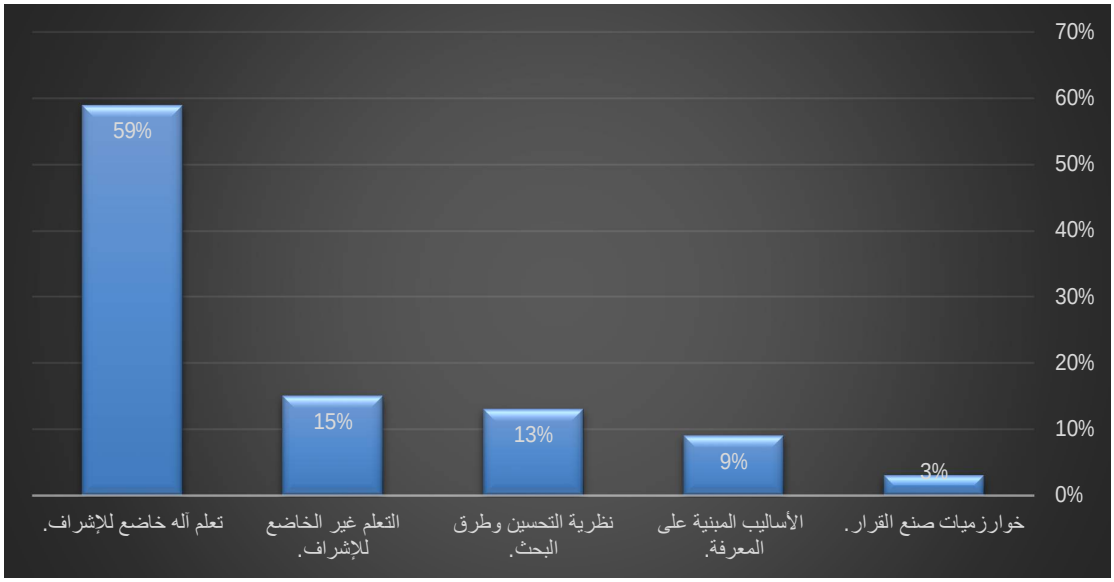
السؤال الأول: ما أنواع تقنيات الذكاء الاصطناعي الشائعة المستخدمة في تحليلات البيانات الضخمة؟ وجاءت النتائج كما هي موضحة بالجدول (9)

جدول (9) يوضح أنواع تقنيات الذكاء الاصطناعي الشائعة المستخدمة في تحليل البيانات الضخمة؟

م	بدائل الاستجابة	نسبة تكرار البديل
1	خوارزميات صنع القرار.	%3
2	الأساليب المبنية على المعرفة.	%9
3	نظرية التحسين وطرق البحث.	%13
4	التعلم غير الخاضع للإشراف.	%15
5	تعلم آله خاضع للإشراف.	%59

يتضح من الجدول السابق أن التعلم الخاضع للإشراف يأتي في الترتيب الأول بنسبة 59% كأحد تقنيات الذكاء الإصطناعي لأكثر شيوعاً في تحليلات البيانات الضخمة ، بينما يأتي في الترتيب الثاني التعلم غير الخاضع للإشراف بنسبة 15%

الشكل (10) يوضح أنواع تقنيات الذكاء الاصطناعي الشائعة المستخدمة في تحليلات البيانات الضخمة



2- مزايا تقنيات الذكاء الاصطناعي في بيئة البيانات الضخمة

السؤال الثاني : ما مزايا تقنيات الذكاء الاصطناعي في بيئة البيانات الضخمة؟

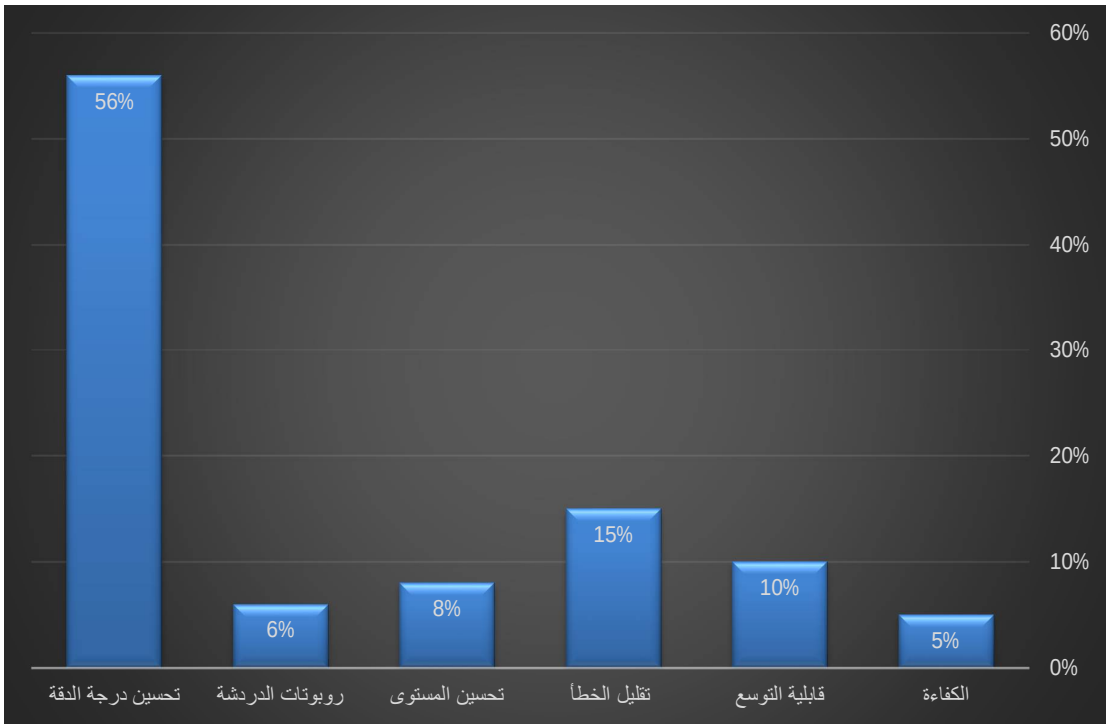
وجاءت النتائج كما هي موضحة بالجدول (10)

جدول (10) يوضح مزايا تقنيات الذكاء الاصطناعي في بيئة البيانات الضخمة

م	بدائل الاستجابة	نسبة تكرار البديل
1	الكفاءة	%5
2	قابلية التوسع	%10
3	تقليل الخطأ	%15
4	تحسين المستوى	%8
5	روبوتات الدردشة	%6
6	تحسين درجة الدقة	%56

يتبين من الجدول السابق من مزايا تقنيات الذكاء الإصطناعي مع البيانات الضخمة تحسين درجة الدقة بنسبة 56% ، ثم يأتي في الترتيب الثاني تقليل الخطأ بنسبة 15%

الشكل (11) يوضح مزايا تقنيات الذكاء الاصطناعي في بيئة البيانات الضخمة



3- كيفية عمل البيانات الضخمة والذكاء الاصطناعي معاً

السؤال الرابع: كيف تعمل البيانات الضخمة والذكاء الاصطناعي معاً؟

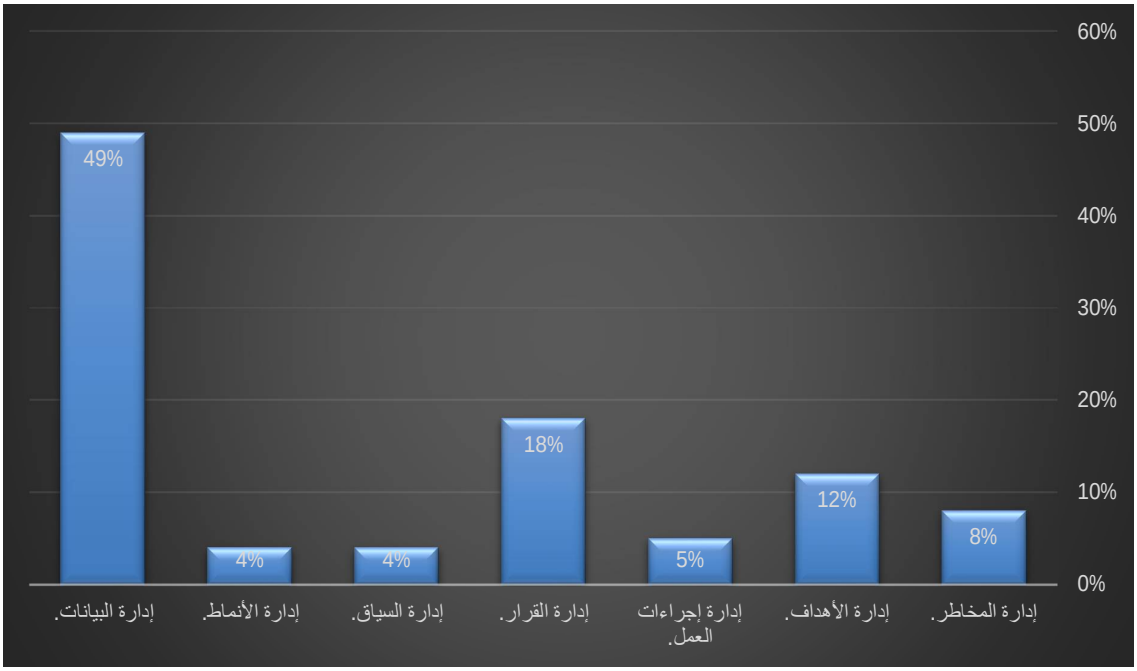
وجاءت النتائج كما هي موضحة بالجدول (11)

جدول (11) كيفية عمل البيانات الضخمة والذكاء الاصطناعي معاً

م	بدائل الاستجابة	نسبة تكرار البديل
1	إدارة المخاطر.	%8
2	إدارة الأهداف.	%12
3	إدارة إجراءات العمل.	%5
4	إدارة القرار.	%18
5	إدارة السياق.	%4
6	إدارة الأنماط.	%4
7	إدارة البيانات.	%49

تبين من نتائج الدراسة من أهم مؤشرات عمل الذكاء الإصطناعي مع البيانات الضخمة إدارة البيانات في الترتيب الأول بنسبة 49% مما يترتب عليها ارتفاع سرعة البيانات (في الوقت الحقيقي تقريباً) ويأتي في الترتيب الثاني الحفاظ على آليات إتخاذ القرار بنسبة 18%

الشكل (12) يوضح كيفية عمل البيانات الضخمة والذكاء الاصطناعي معاً



4-متطلبات تنفيذ الذكاء الاصطناعي مع البيانات الضخمة

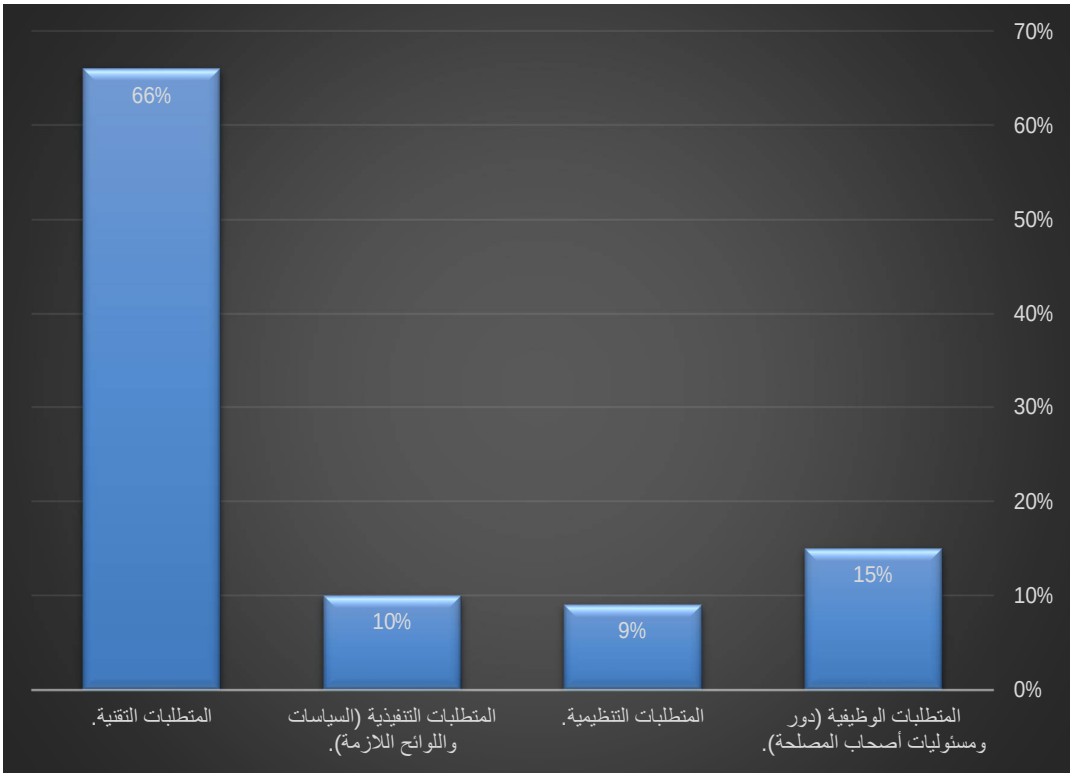
السؤال الرابع: ما متطلبات تنفيذ الذكاء الاصطناعي مع البيانات الضخمة؟ وجاءت النتائج كما هي موضحة بالجدول (12)

جدول (12) متطلبات تنفيذ الذكاء الاصطناعي مع البيانات الضخمة

م	بدائل الاستجابة	نسبة تكرار البديل
1	المتطلبات الوظيفية (دور ومسئوليات أصحاب المصلحة).	%15
2	المتطلبات التنظيمية.	%9
3	المتطلبات التنفيذية (السياسات واللوائح اللازمة).	%10
4	المتطلبات التقنية.	%66

تبين من نتائج الدراسة أن المتطلبات التقنية في الترتيب الأول كأحد متطلبات تنفيذ الذكاء الاصطناعي مع البيانات الضخمة بنسبة 66% وتشمل المتطلبات التقنية (متطلبات الأجهزة – تقنيات الذكاء الاصطناعي (التعلم الآلي – التعلم العميق) – نظم الحوسبة السحابية) ، وبينما يأتي في الترتيب الثاني المتطلبات الوظيفية (مشغلو النظام – القائمون على جمع المعلومات - شركات تكنولوجيا المعلومات والاتصالات) بنسبة 15%

الشكل (13) يوضح متطلبات تنفيذ الذكاء الاصطناعي مع البيانات الضخمة



5- المجالات الرئيسية للذكاء الاصطناعي مع البيانات الضخمة

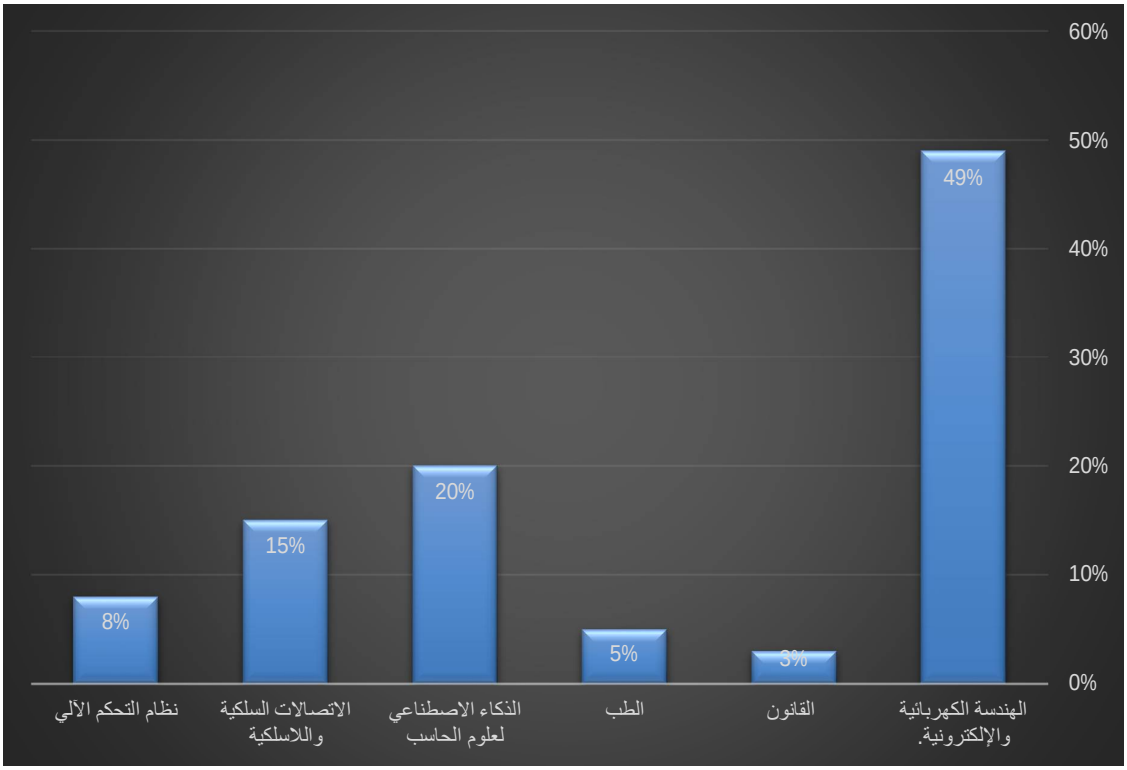
السؤال الخامس: ما المجالات الرئيسية للذكاء الاصطناعي مع البيانات الضخمة؟ وجاءت النتائج كما هي موضحة بالجدول (13)

جدول (13) المجالات الرئيسية في الذكاء الاصطناعي مع البيانات الضخمة

م	بدائل الاستجابة	نسبة تكرار البديل
1	الهندسة الكهربائية والإلكترونية.	%49
2	القانون	%3
3	الطب	%5
4	الذكاء الاصطناعي لعلوم الحاسب	%20
5	الاتصالات السلكية واللاسلكية	%15
6	نظام التحكم الآلي	%8

تبين من نتائج الدراسة أن (الهندسة الكهربائية والإلكترونية) من المجالات الرئيسية لتطبيق الذكاء الاصطناعي مع البيانات الضخمة في المرتبة الأولى بنسبة 49%:

الشكل (14) يوضح المجالات الرئيسية في الذكاء الاصطناعي مع البيانات الضخمة



6- آليات تطبيق الذكاء الاصطناعي في بيئة البيانات الضخمة

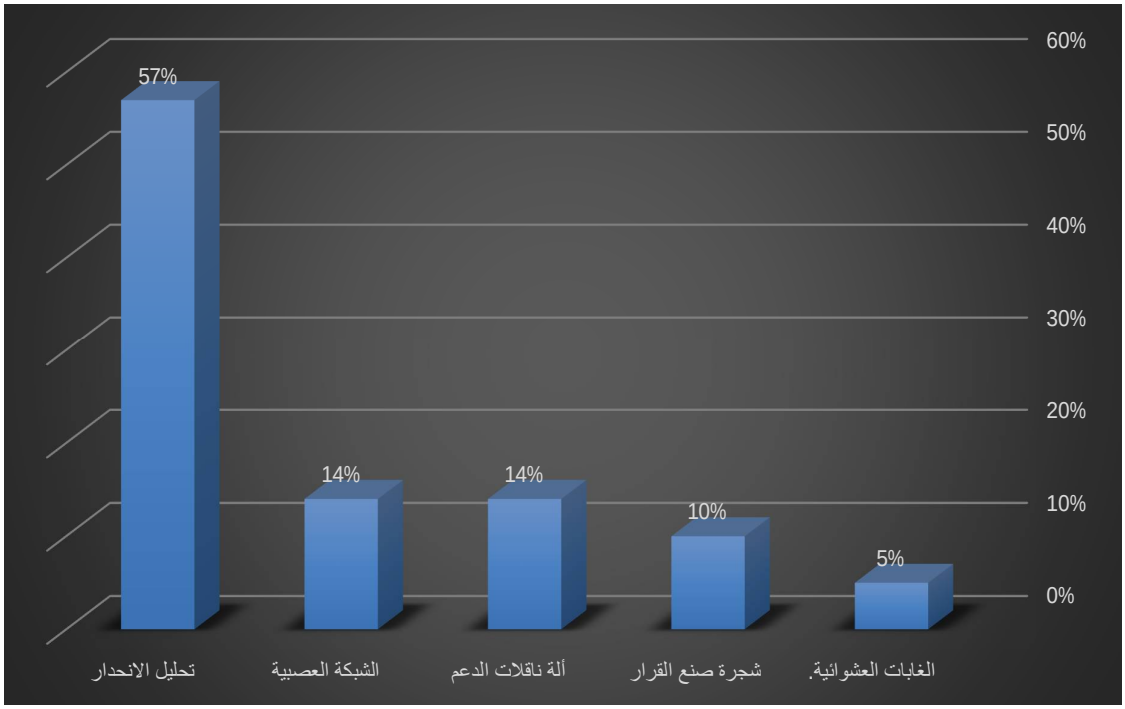
السؤال السادس: ما آليات تطبيق الذكاء الاصطناعي (خوارزميات التعلم تحت الإشراف) في بيئة البيانات الضخمة ؟ وجاءت النتائج كما هي موضحة بالجدول (14)

جدول (14) يوضح خوارزميات التعلم تحت الإشراف في البيانات الضخمة

م	بدائل الاستجابة	نسبة تكرار البديل
1	الغابات العشوائية.	%5
2	شجرة صنع القرار	%10
3	آلة ناقلات الدعم	%14
4	الشبكة العصبية	%14
5	تحليل الانحدار	%57

يتضح من الجدول السابق أن تحليل الإنحدار أكثر خوارزميات التعلم تحت الإشراف إستخداماً في بيئة البيانات الضخمة في الترتيب الأول بنسبة 57%

الشكل (15) يوضح آليات تطبيق الذكاء الاصطناعي (خوارزميات التعلم تحت الإشراف) في بيئة البيانات الضخمة



7- المؤسسات الأكثر إنتاجية في نشر كل ما يتعلق بالبيانات الضخمة والذكاء الاصطناعي

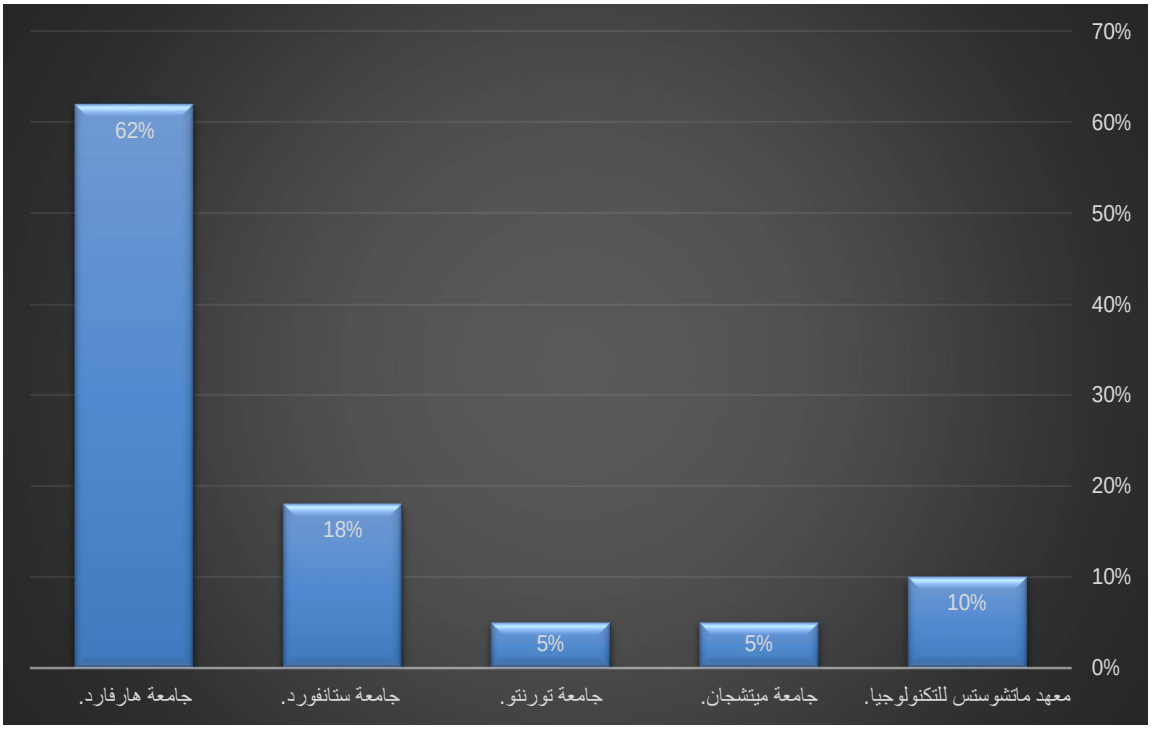
السؤال السابع: ما المؤسسات الأكثر إنتاجية في نشر كل ما يتعلق بالبيانات الضخمة والذكاء الاصطناعي؟ وجاءت النتائج كما هي موضحة بالجدول (15)

جدول (15) المؤسسات الأكثر إنتاجية في نشر كل ما يتعلق بالبيانات الضخمة والذكاء الاصطناعي

م	بدائل الاستجابة	نسبة تكرار البديل
1	معهد ماساتشوستس للتكنولوجيا.	%10
2	جامعة ميتشجان.	%5
3	جامعة تورنتو.	%5
4	جامعة ستانفورد.	%18
5	جامعة هارفارد.	%62

يتضح من نتائج الدراسة أن المؤسسات الأكثر إنتاجية في نشر كل ما يتعلق بالبيانات الضخمة والذكاء الإصطناعي هي جامعات أممريكا الشمالية بالترتيب (جامعة هارفارد – جامعة أستانفورد – معهد ماساتشوستس للتكنولوجيا) تأتي في الصدارة جامعة هارفارد بنسبة 62% لذا فإن الدول الرائدة في النشر هي الولايات المتحدة الأمريكية

الشكل (16) يوضح المؤسسات الأكثر إنتاجية في نشر كل ما يتعلق بالبيانات الضخمة والذكاء الاصطناعي



8- التحديات الرئيسية والقضايا المفتوحة

السؤال الثامن:

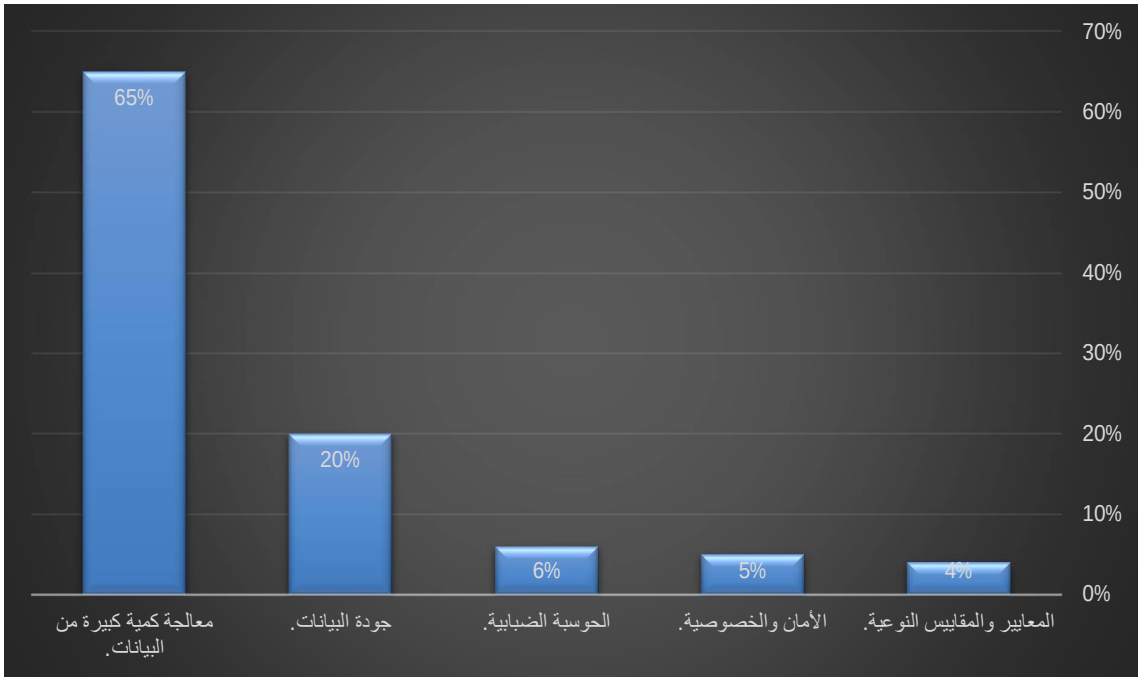
ما التحديات الرئيسية والقضايا المفتوحة؟ وجاءت النتائج كما هي موضحة بالجدول (16)

جدول (16) التحديات الرئيسية والقضايا المفتوحة

م	بدائل الاستجابة	نسبة تكرار البديل
1	المعايير والمقاييس النوعية.	%4
2	الأمان والخصوصية.	%5
3	الحوسبة الضبابية.	%6
4	جودة البيانات.	%20
5	معالجة كمية كبيرة من البيانات.	%65

تبين من نتائج الدراسة أن أكثر تحديات الذكاء الإصطناعي مع البيانات الضخمة (معالجة كمية كبيرة من البيانات في الترتيب الأول بنسبة 65% ، ثم يأتي في الترتيب الثاني جودة البيانات بنسبة 20%

الشكل (17) التحديات الرئيسية والقضايا المفتوحة



النتائج والتوصيات:

أولاً: النتائج

1. هناك العديد من الدراسات التي تتعلق بتحليلات البيانات الضخمة ولكن تفتقر هذه الدراسات على المراجعة الكاملة والمنهجية لآليات تطبيق الذكاء الاصطناعي مع البيانات الضخمة.

2 يفتح هذا البحث أفقًا لتطوير الحلول التي من شأنها معالجة التحديات لتحدياتالمتعلقة بالذكاء الاصطناعي وتحليلات البيانات الضخمة.

3. تبين من نتائج الدراسة أن أهم تقنيات الذكاء الاصطناعي الشائعة في تحليل البيانات الضخمة هي خوارزمية التعلم الخاضع للإشراف في الترتيب الأول بنسبة 59% ويأتي التعلم غير الخاضع للإشراف في الترتيب الثاني بنسبة 15%.

4. تبين من نتائج الدراسة أن المؤسسات الأكثر إنتاجية في نشر كل ما يتعلق بالبيانات الضخمة والمواد ذات الصلة بالذكاء الاصطناعي هي جامعات أمريكا الشمالية مثل جامعة هارفارد في الترتيب الأول بنسبة 62% ثم جامعة ستانفورد في الترتيب الثاني بنسبة 18% تليها في الترتيب الجامعات الصينية والأسترالية والأوروبية لذلك فإن الدول الرائدة هي الولايات المتحدة الأمريكية.

5- من مزايا تقنيات الذكاء الاصطناعي مع البيانات الضخمة يأتي في الترتيب الأول تحسين درجة الدقة بنسبة 56% ثم تقليل الخطأ في الترتيب الثاني بنسبة 15%.

6. تعد إدارة البيانات من أهم مؤشرات عمل البيانات الضخمة والذكاء الاصطناعي معًا في الترتيب الأول بنسبة 49% بينما تعد إدارة السياق في الترتيب الأخير بنسبة 4%.

7. يتضح من نتائج الدراسة أن الهندسة (الكهربائية والإلكترونية) من المجالات الرئيسية للمقالات ذات الصلة في مجال الذكاء الاصطناعي مع البيانات الضخمة بنسبة 49%.

8. من متطلبات تنفيذ الذكاء الاصطناعي مع البيانات الضخمة يأتي في الترتيب الأول (المتطلبات التقنية) بنسبة 66% ثم يأتي في الترتيب الثاني المتطلبات الوظيفية بنسبة 15%.

9. معالجة كمية كبيرة من البيانات من أهم تحديات تطبيق الذكاء الاصطناعي مع البيانات الضخمة في الترتيب الأول بنسبة 65% ثم جودة البيانات في الترتيب الثاني بنسبة 20%.

ثانيًا: التوصيات

1. زيادة الاهتمام بالدراسات التي تتعلق بتطبيق آليات الذكاء الاصطناعي مع البيانات الضخمة.

2. من الضروري إعادة تعريف مفهوم الذكاء الاصطناعي والمصطلحات ذات الصلة لتعكس الطبيعة المتغيرة لتطوير الذكاء الاصطناعي وتطبيقه في عصر البيانات الضخمة.

3. يجب أن يكون هناك نهج متفق عليه لجمع وهيكلة وتمثيل البيانات من خلال خوارزميات التعلم الآلي والتعلم العميق.

4. ضرورة تنفيذ تقنيات ذكاء اصطناعي مختلفة لتحليل هذه الكميات الضخمة من البيانات في الوقت الفعلي.

5. زيادة الدراسة التي تركز على القضايا الأمنية لتحليلات البيانات الضخمة مع الذكاء الاصطناعي.

6. زيادة الدراسات التجريبية حول البنية التحتية الذكية ومعايير الجودة.

7. هناك الحاجة إلى مزيد من الدراسات حول مشاكل جودة البيانات وتنسيقات البيانات غير المهيكلة وشبه المهيكلة.

قائمة مراجعة حول (آليات تطبيق نظم الذكاء الإصطناعي

في بيئة البيانات الضخمة : دراسة وصفية تحليلية

تصف هذه الدراسة أولاً المفاهيم ذات الصلة بالبيانات الضخمة والذكاء الإصطناعي وتحلل الخصائص والمتطلبات التقنية والتنظيمية والسياسات اللازمة للذكاء الإصطناعي في بيئة البيانات الضخمة ودراسة التحديات والقضايا المفتوحة

1- ما أنواع تقنيات الذكاء الإصطناعي الشائعة المستخدمة في تحليل البيانات الضخمة ؟

1- خوارزميات صنع القرار (3%)

2- الأساليب المبنية على المعرفة (9%)

3- نظرية التحسين وطرق البحث (13%)

4- التعلم غير الخاضع للإشراف (15%)

5- تعلم آله خاضع للإشراف (59%)

التقنية الشائعة التي يستخدمها الباحثون لتحليل البيانات الضخمة هي التعلم الخاضع للإشراف بنسبة 59%

2- ما خوارزميات التعلم تحت الإشراف في البيانات الضخمة ؟

1- الغابات العشوائية (5%)

2- شجرة صنع القرار (10%)

3- ألة ناقلات الدعم (14%)

4- الشبكة العصبية (14%)

5- تحليل الإنحدار (57%)

3- ما مزايا تقنيات الذكاء الإصطناعي في بيئة البيانات الضخمة ؟

1- الكفاءة 2- قابلية التوسع

3- الخصوصية 4- تقليل الخطأ

5- تحسين المستوى 6- روبوتات الدردشة

7- تحسين درجة الدقة (56%)

4- كيف تعمل البيانات الضخمة والذكاء الإصطناعي معاً؟

1-إدارة المخاطر 2-إدارة الأهداف

3- إدارة إجراءات العمل 4- إدارة القرار

5- إدارة السياق 6- إدارة الأنماط

7- إدارة البيانات (49%)

5- ما المجالات الرئيسية في الذكاء الاصطناعي مع البيانات الضخمة ؟

1- نظام التحكم الآلي **2- القانون**

3- الطب **4- الذكاء الإصطناعي لعلوم الحاسب**

5- الإتصالات السلكية واللاسلكية **6- الهندسة الكهربائية والإلكترونية (49.53%)**

6- ما المؤسسات الأكثر إنتاجية في نشر كل ما يتعلق بالبيانات الضخمة والذكاء الإصطناعي ؟

1- معهد ماتشوستس للتكنولوجيا **2- جامعة ميتشجان**

3- جامعة تورنتو **4- جامعة ستانفورد**

5- جامعة هارفارد (62%)

7- ما متطلبات تنفيذ الذكاء الإصطناعي مع البيانات الضخمة ؟

1- المتطلبات الوظيفية (دور ومسئوليات أصحاب المصلحة)

2- المتطلبات التنظيمية

3- المتطلبات التنفيذية (السياسات واللوائح اللازمة)

4- المتطلبات التقنية 66.42%

8- ما التحديات الرئيسية والقضايا المفتوحة؟

1- المعايير والمقاييس النوعية **2- الآمان والخصوصية**

3- الحوسبة الضبابية **4- جودة البيانات 42%**

5- معالجة كمية كبيرة من البيانات 67%

المراجع

أولاً: المراجع العربية

1. **أحمد، نجوى شكري (2021).** إدارة البيانات الضخمة في دار الكتب والوثائق القومية: دراسة استكشافية، المجلة العلمية للمكتبات والوثائق والمعلومات، مج3، ع7 يوليو- ص ص 145-229.

2- الذبياني، علي بن أحمد (2020). دور مؤشرات الأداء الرئيسية في تحليلات البيانات الضخمة، مجلة دراسات المعلومات (جمعية المكتبات والمعلومات السعودية)، ع24يناير، ص ص 85-103.

3- سميع، ليلى سيد (2021). تحليل المشاعر (الآراء) المعتمد على الذكاء الاصطناعي لتقييم فعالية خدمات المعلومات، الندوة العلمية الأولي: تطبيقات الذكاء الاصطناعي في العلوم الاجتماعي والإنسانيات (الخصائص – البرمجيات- آليات التنفيذ) 29-30 ديسمبر) ص25.

4- عبد القوى، أماني (2021). متطلبات إدارة البيانات الضخمة في مراكز المعلومات بالمملكة العربية السعودية: دراسة حالة، المؤتمر السنوي الثاني والثلاثون للاتحاد العربي للمكتبات والمعلومات (مستقبل مؤسسات المعلومات العربية في ظل التقنيات الذكية رؤية استراتيجية 2050)، ص371.

5- العريشي ، جبريل بن حسن (2020). استخدام البيانات الضخمة والذكاء الاصطناعي في مواجهة جائحة كورونا، المجلة العربية للدراسات الأمنية، مج36، ع2، ص ص 249- 264.

6- غلوم، حسين فولاز (2021). دور الذكاء الاصطناعي في دعم وتطوير خدمات المكتبات الأكاديمية في دولة الكويت: إدارة المكتبات بجامعة الكويت أنموذجا، المؤتمر السنوي الثاني والثلاثون للاتحاد العربي للمكتبات والمعلومات، القاهرة 15-16 ديسمبر. ص ص 13-30.

7- فرج، أحمد (2021). تقنيات الذكاء الاصطناعي وتأثيراتها على فاعلية محتوى وخدمات مؤسسات المعلومات، الندوة العلمية الأولى: تطبيقات الذكاء الاصطناعي في العلوم الاجتماعية والإنسانية (الخصائص- البرمجيات- آليات التنفيذ)، 29-30 ديسمبر، ص25.

8- قناوي، يارة ماهر (2020). دور تحليلات البيانات الضخمة في إنترنت الأشياء: دراسة تحليلية مقارنة، المجلة الدولية لعلوم المكتبات والمعلومات،مج6، ع2 يونيو 2020، ص ص 73-110.

9- الكشكي، ناصر أبو زيد (2021). تحليلات البيانات الضخمة في المؤسسات الأكاديمية: دراسة استشرافية بالتطبيق على مؤسسات التعليم العالي المصرية، مجلة بحوث في علم المكتبات والمعلومات، ع27 سبتمبر، ص ص 9-66.

10- متولى، عبد الله حسين (2021). تقنية التعرف على الوجوه كأحد تطبيقات الذكاء الاصطناعي في المكتبات ومؤسسات المعلومات إعادة قراءة لعلم الفهرسة، الندوة العلمية الأولى: تطبيقات الذكاء الاصطناعي في العلوم الاجتماعية والإنسانيات (الخصائص- البرمجيات- آليات التنفيذ)، 29-30 ديسمبر.

11- محمد، أسماء حسين (2021). واقع استفادة المكتبات الجامعية من تحليلات البيانات الضخمة Big data في تطوير خدمات المعلومات واتخاذ القرار: دراسة حالة لمكتبة كلية الهندسة، المجلة العلمية للمكتبات والوثائق والمعلومات، مج3، ع6 يوليو، ص ص 111-144.

12- ناجي، إهداء صلاح (2021). تطبيقات نظم الذكاء الاصطناعي في تحليل المحتوى وعمليات التكشيف: دراسة تطبيقية لنظم معالجة اللغة الطبيعية، الندوة العلمية الأولى، تطبيقات الذكاء الاصطناعي في العلوم الاجتماعية والإنسانيات (الخصائص، البرمجيات، آليات التنفيذ)، 29-30 ديسمبر.، ص 33

ثانيًا: المراجع الأجنبية

1- **Ahmed, F.B (2021).** A deep Learning approach for covid-19 8 viral Pneumonia screening with X-ray images. Digital government: Research and Practice 2 (2), pp 1-12.

2- **Ahmed, I & Jeon, G (2021).** A frame work for pandemic prediction using big data analytics. Big data research100190. Do I 10.1016I. bdr. 202) 100190.

3- **Abbas, H. G (2016).** A review of theoretical and practical challenges of trusted Autonomy in big data "IEEE Access, vol. 4, pp. 2808-2830. http://doi-ovg/100/109/access.2016.257/058.

4- **Athmajas, S.H (2017).** A survey of Machine Learning algorithms for big data analytics in: 2017 international conference on innovations in information, embedded and communication systems (ICIECS).

5- Al-Kabbany & El.bana (2020). Amulti- Task pipeline with specialized in streams for classification and segmentation of infection manifestation n Covid-19 cans, peerj computer science 6:e303 Do I 10.7717) Peerj-CS.303.

6- Alzubi, A. A (2020). Big data analytic diabetics using MapReduce and classification techniques. The journal of super comptuiing 76 (6), pp 4328-7337.

7- Balandin, E.B (2013). «main trends in M health use scenarion» journal on selected topics in Nano Electronics and computing, Vol,4, No, 1, pp 64-70.

8- Banchhor, C.S (2020). Integrating Guckoo Search-Grey wolf optimization and correlative naïve bays classifier with Map reduce model

for big data, data & knowledge Engineering 127: 10/788 do I 10.1016/ J-Datak. 2019.10/788.

9- Bourouis, A.Z (2013). «M-health: s kindisease analysis system using smart phone's camera, Procedia computer science, vol. 19,pp 1116-1120.

10- Borodo S.M (2016). Big data platform and techniques Indonesian Journal of Electrical Engineering and computer science 1(1). PP. 191-200.

11- Carcillo, F (2018). Scraff: a scalable framework for streaming credit card fraud detection with spark, information fusion 41: 182-194 Do I 10.1016/J.inffus. 2017.09-005.

12- chen, M& yang (2018). «5G smart diabetes: Toward personalized diabetes diagnosis with Health care big date clouds, IEEE communication Magazine, Vol. 56, no. 4, pp 16-23.

13- Darji, H.H & shan (2012). concepts of distributed and parall EL Database, International Journal of computer science and information technology & Security, vol. 2, No. 6. P 1150.

14- Bollen, J & Mao (2011). Twitter mood as a stock market predictor, computer, vol. 44, No . 10, pp. 91-94.

15- Feng, M & zheng. (2019). Big date analytics and Mining for effective Visualization and trends forecasting of crime date. IEEE Access 7: 106111-106123.

16- Griffen, E. A(2018). Can we accelerate Medicinal chemistry by augmenting the chemist with Big date and artificial intelligence? Drug discovery today, Vol 23, no. 7, pp 1373-1384.

17- Hammou, B. A. (2020). Towards a real-time processing framework based on improved distributed recurrent neural network variants with fast text for social big date analytics, information processing & management 57(1). P/02/12. Do I10.1016/ J.IPM. 2019.

18- IREA (2019). Artificial Intelligence and big data: innovation landscape Brief. www. irena-org, pp1-23.

19- Jing, X& Bain (2018). Deep Learning form drug Design: an Artificial intelligence paradigm for drug discovery in the Big data a Era, AAPs Journal, vol. 20, No, 3. http://doi.org/10.1208/s/2248.618-0210.0.

20- Kannan, N.S .(2019). Predicative big date analytic on demonetization data using support vector machine, cluster computing 22(6)-pp 1472-14720.

21- Kaur, P & sharma. (2018). Big data and Machine Learning based secure health care frame work. Procedia computer science 132-pp 1049-1059.

22- Klein,S (2017). The world of big data and IoT solutions in Microsoft's azure IoT suite. Pp 3-13.

23- Kiral, J& Kronok (2018). Epileptic seizure production using big date and deep learning: tow toward a mobile system, Biomedicine, Vol, 22, pp 103-111.

24- Larson, E.C (2012). Spiro smart: using a Microphone to Measure lung function on a mobile phone, in proceedings of the 2012 conference (BMEICON), pp1-4.

25- Lei, y.f (2016). An intelligent fault diagnosis Method using unsupervised feature learning forward Mechanical Big data, IEEE transactions on Industrial Electronics, vol. 63, Vol. 63, No. 5, pp 3/37. 3/47.

26. Li, Y& Zhang (2018). Ultra- DenseHet Nets meet big data: green frameworks, techniques and approaches, IEEE communications Magazine, vol. 56, No. 6, pp 56-63.

27- Mittal, S.S (2018). Big date analytics using machine Learning techniques in: 2019 9th international conference on cloud computing. Date science & engineering. Piscata way: IEEE.

28- Nair LR & Shetty (2018). Applying Spark based Machine learning Model on streaming big data for health status prediction. Computer& electrical engineering 65, pp 393-399.

29- Natale, V.M (2012). Menitoring sleep with a smartphone accelerometer, sleep and biological Rhthms, vol 10, No. 4, pp 287-292.

30- Nicolable, F.C (2018). Big data analytics in IOT: challenges, open research issues and tools, in: world conference on information systems and technologics, cham: springer.

31- O'Leary, D (2013). Artificial intelligence and big data, IEEE Intelligent systems, vol 28, no. 2, pp96-99.

32- O'Leary, D.E (2012). Knowledge discovery for continuous Financial Assurance using Multiple types of digital information in Data mining information age publishing, pp 103-122.

33- oussous, A. B.(2018). big data technologies: a survey. Journal of kings saud university- computer and information science 30 (4), pp 431-448.

34- Ozdemir, V.N. (2018). Birth of industry 5.0: Making sense of big data with Artificial intelligence, the internet of things and Next- Generation technology policy, Journal of integrative Biology, Vol. 21, No, PP 65-76.

35- Philip, H & Steven. (2016). A system for content – based indexing of a database of news stories. Proc2nd conf. Innovation applications of Artificial intelligences for the advancement of At, pp 49-64.

36- Ragab, D.A .(2020). Fust – CAD : Coronavirus (Covid-19) diagnosis based on fusion of CNNS and hands crafted features. Peerj computer science 6: 306 . doi 10.77/7/peerj-cs.306.

37- Rahmani, A & Azhir (2021). Artificial Intelligence approaches and Mechanisms for big date analytics, as systematic study ,peerj computer science. Sci. 7: e488, pp 1-28.

38- Siddiqa, A & Abaker (2016). A survey of big data management taxonomy and state –of the- art. journal of network and computer applications 71, pp 151-166.

39- Sultan, Aboabkr (2019) الذكاء الاصطناعي مع البيانات الضخمة والحوسبة الإدارية . http://www.researchgate.net/publicatin/3409/3409/4289. فرص وتهديدات

40- Sayda, T.M (2018). Role of big data and machine learning in diagnostic precision support in Radiology, journal of the American college of Radiology, vol. 15, No. 3, pp 569-576.

41- Tsai, C.W (2015). Big data analytics: a survey. Journal of big data 2 (1), pp 1-32.

42- Tian, Z& luo (2020). User and Entity behaviour analysis under urban big data. ACM transactions son data science (3), pp 1-19.

43- Tsanas, A& Little (2012). Novel speech signal processing a algorithms for high – accuracy classification of parknson's disease , IEEE, vol. 59, No.5, pp 1246-1271.

44- Vaishya, R(2020). Artificial intelligence (AI) applications for covide-19 pandemic, Diabetes & Metabolic syndrome: clinical Research & reviews 14 (4), pp 337-339.

45- wang, S.C (2020). A Hybrid big data analytical approach for analyzing customer patterns through an integrated supply chain Network, journal of industrial information integration 20: 100177.

46- Zhuang, Y.F (2017). Challenges and opportunities: from big data to knowledge in AI 2.0 frontiers of information technology & Electronic Engineering, vol. 18, No. 1, pp 3-14.

الهوامش:

1) http://pdfs.semantic.scholar.org/du869836/bsda.of

2) http://assistant.google.com

3) http:///www.computer.world.com/article/3/82207/cuso-date.storage-goes-from-im-to-2-cents-per-gigiabyte.html.

4) http:///www.IBM.com/waston

آليات تطبيق نظم الذكاء الاصطناعي في بيئة البيانات الضخمة:

دراسة وصفية تحليلية

إعـــــــــداد

د/ ياره ماهر محمد قناوي

أستاذ مساعد بقسم المكتبات والمعلومات

جامعة المنيا

ymaher57@yahoo.com

المستخلص:

تصف هذه الدراسة أولاً المفاهيم ذات الصلة بالبيانات الضخمة والذكاء الاصطناعي ، وتحلل المتطلبات الوظيفية والتقنية للذكاء الاصطناعي في بيئة البيانات الضخمة ، كما تهدف الدراسة الى التعرف على آليات تطبيق تقنيات الذكاء الاصطناعي التي تعتمد على البيانات الضخمة والتعرف على كيفية تطبيق الذكاء الاصطناعي في إطار البيانات الضخمة ، وتسعى الدراسة إلى تحليل واستكشاف المجالات الرئيسية المتعلقة بالذكاء الاصطناعي مع البيانات الضخمة ومناقشة القضايا المفتوحة لتقديم اتجاهات بحثية جديدة في مجال تحليلات البيانات الضخمة مع الذكاء الاصطناعي ، وقدمت الباحثة مراجعة علمية شاملة عن آليات تطبيق الذكاء الاصطناعي في بيئة البيانات الضخمة واعتمدت الدراسة على المنهج الوصفي التحليلي من خلال قائمة مراجعة تضم العناصر التالية : (أنواع تقنيات الذكاء الاصطناعي الأكثر شيوعا في تحليلات البيانات الضخمة، خوارزميات التعلم تحت الإشراف في البيانات الضخمة، مزايا تقنيات الذكاء الاصطناعي مع البيانات الضخمة، كيفية عمل الذكاء الاصطناعي في إطار البيانات الضخمة- المجالات الرئيسية لتطبيق الذكاء الاصطناعي - متطلبات تنفيذ الذكاء الاصطناعي مع البيانات الضخمة- التحديات والقضايا المفتوحة) وتوصلت الدراسة إلى عدة نتائج من أهمها : هناك العديد من الدراسات التي تتعلق بالبيانات الضخمة ولكن تفتقر هذه الدراسات على المراجعة الكاملة والمنهجية لآليات تطبيق الذكاء الاصطناعي مع البيانات الضخمة ، من أكثر تحديات الذكاء الاصطناعي مع البيانات الضخمة يأتي في الترتيب الأول معالجة كمية كبيرة من البيانات بنسبة 65% وتوصي الدراسة بزيادة الاهتمام بالدراسات التي تتعلق بتطبيق آليات الذكاء الاصطناعي مع البيانات الضخمة وضرورة تنفيذ تقنيات ذكاء اصطناعي مختلفة لتحليل هذه الكميات الضخمة من البيانات في الوقت الفعلي.

الكلمات الدالة:

الذكاء الاصطناعي – تحليلات البيانات الضخمة – التعلم الآلي – التعلم العميق – التعلم الآلي الخاضع للإشراف – التعلم الآلي بدون إشراف

The Mechanisms OF Applying Artificial intelligence systems in a big data environment: An analytical descriptive study

preparation

Dr. Yara Maher Mohamed Kenawy
Ass.Prof, Department
of Library and Information
Minia University
ymaher57@yahoo.com

Abstract:

This study first describes the concepts related to big data and artificial intelligence, and analyzes the functional and technical requirements of artificial intelligence in the environment of big data. The study seeks to analyze and explore the main areas related to artificial intelligence with big data and discuss open issues to present new research trends in the field of big data analytics with artificial intelligence. Through a checklist that includes the following items: (Types of artificial intelligence techniques are most common in massive data analyzes, learning algorithms under the supervision of large data, the advantages of artificial intelligence techniques with huge data, how to make artificial intelligence within the framework of massive data - the main areas of artificial intelligence application - requirements for the implementation of artificial intelligence with huge data - Open challenges and Among the most important: There are issues) and the study reached several results many studies related to big data, but these studies lack a complete and systematic review of the mechanisms of applying artificial intelligence with big data. One of the most challenges of artificial intelligence with big data comes in the first place, processing a large amount of data by 65%. The study recommends increasing Interest in studies related to the application of artificial intelligence mechanisms with big data and the necessity of implementing various artificial intelligence techniques to analyze these large amounts of data in real time.

Key words:

Artificial Intelligence - big Data Analytics - Machine Learning - Deep Learning - Supervised Learning - unsupervised Learning.