

مقدمة حول تحليل البيانات

تُعتبر عملية تحليل البيانات عملية استكشاف البيانات، وفحصها، لاستخراج المعلومات المفيدة، وتحديد الأنماط، والتوجهات، والعلاقات الموجودة داخلها، حيث أصبحت البيانات في عصر المعلومات الحالي جزءاً أساسياً من عمليات اتخاذ القرار في مختلف المجالات. وتُستخدم تقنيات تحليل البيانات في مجموعة واسعة من التطبيقات، بدءاً من الأعمال التجارية والتمويل، وصولاً إلى العلوم الصحية والبحث العلمي.

1. تعريف تحليل البيانات

-1- يعرف تحليل البيانات بأنه : عملية فحص وتدقيق البيانات وانتقائها، وتحويلها وتخزينها ونمذجتها؛ بهدف استخلاص استنتاجات أو معلومات مفيدة، يمكن أن تكون لمعرفة سبب أو تفسير لأحداث وقعت في الماضي، أو من أجل تحسين أحداث ستحدث في المستقبل والمساعدة في عملية اتخاذ القرارات.

وكذلك **تحليل البيانات** هي عملية الحصول على البيانات الخام، ومن ثم تحويلها إلى معلومات مفيدة لاتخاذ القرارات من قبل المستخدمين، يتم جمع البيانات وتحليلها للإجابة على الأسئلة أو اختبار الفرضيات أو نفي النظريات.

وأيضا **تحليل البيانات** هي أحد مجالات علوم الكمبيوتر التي تركز على استخلاص الأفكار من البيانات الخام، حيث يستخدم محللو البيانات التقنيات الإحصائية، ولغات البرمجة للكشف عن الأنماط والاتجاهات والعلاقات داخل مجموعات البيانات الكبيرة، و الهدف الأساسي هو دعم اتخاذ القرارات المتعلقة بتحليل البيانات المستندة إلى البيانات عبر المؤسسات.

- من هم محللو البيانات؟

هم أفراد مسؤولون عن جمع البيانات وتنظيفها وتحويلها، وتشمل مسؤولياتهم أيضاً تصميم وإنشاء نماذج بيانات قابلة للتطوير وفعالة. ويقوم محللو البيانات بجمع تلك البيانات وتنظيمها من مصادر مختلفة، كما يقومون ببناء نماذج التعلم الآلي، وإجراء الاختبارات الإحصائية وتفسير النتائج وتوصيل النتائج إلى أصحاب المصلحة، من خلال لوحات المعلومات والتقارير وتصورات البيانات، وتمتد مسؤولياتهم أيضاً إلى مشاريع التحليلات الوصفية والتنبؤية والتشخيصية.

2. أهمية تحليل البيانات

في ظل التطور التكنولوجي المستمر، أصبحت كميات البيانات التي ننتجها ونتعامل معها تظهر نمواً كبيراً جداً، فتحليل البيانات يُعتبر أداة حاسمة لفهم هذه البيانات وتحويلها إلى معرفة قابلة للتطبيق، حيث أن البيانات في صورتها الخام ماهي إلا مجموعة كبيرة جداً من الأرقام والحروف والكلمات والصور المختلفة، ودائماً ما توجد بداخل تلك البيانات أنماطاً معينة متكررة لها مدلولات مختلفة فمن خلال تحليل البيانات، يمكن للمؤسسات والأفراد استنباط تلك الأنماط وفهمها وفتح أبواب الفهم للظواهر المعقدة، وإيجاد حلول للتحديات التي قد تواجههم.

• أسباب رئيسية تجعل تحليل البيانات مهمة للشركات:

كما قلنا سابقاً أن عملية تحليل البيانات هي عملية تخزين البيانات الخام وتنظيمها وتحليلها للإجابة على الأسئلة أو اكتساب رؤى مهمة، فتحليل البيانات جزء لا يتجزأ من الأعمال التجارية لأنه يسمح للقيادة بإنشاء استراتيجية قائمة على الأدلة ، وفهم العملاء لاستهداف مبادرات التسويق بشكل أفضل وزيادة الإنتاجية الإجمالية، و الشركات التي تستفيد من تحليلات البيانات تجني ميزة تنافسية لأنها قادرة على إجراء تغييرات أسرع تزيد من الإيرادات وتخفض التكاليف وتحفز الابتكار. وفي عالمنا الرقمي اليوم، تعد القدرة على اتخاذ قرارات تعتمد على البيانات وإنشاء استراتيجية فعالة ، فالتحليل أمراً أساسياً للقيادة الناجحة في أي صناعة.

3. عملية تحليل البيانات

تشمل عملية تحليل البيانات عدة خطوات أساسية، بدءاً من تجميع البيانات ومروراً بتنظيفها وتحليلها، وصولاً إلى تصوير النتائج وتحليلها ومن ثم استخراج الأنماط منها ثم بعد ذلك اتخاذ قرار بشأنها. في هذا السياق، يتم الاعتماد على مجموعة من الأساليب الإحصائية وتقنيات التعلم الآلي لاستكشاف الأنماط والعلاقات الكامنة بداخل البيانات.

- أدوات وبرمجيات تحليل البيانات:

• Excel: أداة قوية لتحليل البيانات والإحصاءات البسيطة.

• Tableau: لتصوير البيانات وإنشاء التقارير التفاعلية.

• Power BI: لتحليل البيانات ومشاركتها عبر الشبكة الداخلية للمؤسسة.

- بعض منصات تحليل البيانات:

1. Google Analytics : لتحليل بيانات الويب وتتبع تفاعل المستخدمين.

2. Apache Hadoop : لتحليل البيانات الضخمة.

3. Apache Spark : لمعالجة البيانات في الوقت الفعلي

- تحليلات البيانات التي تتطلبها بعض المهن:

تستفيد مجموعة واسعة من المهن من مهارات تحليل البيانات لتعزيز قيمة الأعمال، تشمل الأدوار الشائعة ما يلي:

- محلل بيانات : جمع البيانات ومعالجتها وإجراء تحليلات عليها لاستخلاص الرؤى
- محلل استخبارات الأعمال : إنشاء التقارير والتصورات ولوحات المعلومات للتواصل بشأن نتائج تحليلات الأعمال الرئيسية من أجل اتخاذ قرارات أعمال أفضل
- محلل تسويق : تحسين الحملات وأداء المنتج باستخدام التحليلات
- مهندس بيانات : بناء وصيانة خطوط الأنابيب والبنية الأساسية للبيانات ، وتصميم أنظمة قابلة للتطوير وأطر تخزين البيانات
- مدير علوم البيانات : قيادة فريق من علماء البيانات والمحللين
- مهندس التعلم الآلي : تطوير النماذج والخوارزميات لأتمتة المهام التحليلية

4. خطوات عملية تحليل البيانات

تحليل البيانات عبارة عن سلسلة منظمة من الخطوات التي تؤدي من البيانات الخام إلى رؤى قابلة للتنفيذ، فيما يلي إجابات على ما هو تحليل البيانات:

1. جمع البيانات: جمع البيانات ذات الصلة من مصادر مختلفة، وضمان جودة البيانات وسلامتها.
2. تنظيف البيانات: تحديد الأخطاء والقيم المفقودة والتناقضات في مجموعة البيانات وتصحيحها. يعد تنظيف البيانات أمرًا بالغ الأهمية للتحليل الدقيق.
3. تحليل البيانات الاستكشافي: إجراء تحليل أولي لفهم خصائص البيانات وتوزيعاتها وعلاقاتها.
4. تحويل البيانات: إعداد البيانات للتحليل عن طريق ترميز المتغيرات الفئوية، وميزات القياس، ومعالجة القيم المتطرفة، إذا لزم الأمر.
5. بناء النموذج: اعتمادًا على الأهداف، يتم تطبيق أساليب تحليل البيانات المناسبة، مثل الانحدار، أو التجميع، أو التعلم العميق.
6. تقييم النموذج: اعتمادًا على نوع المشكلة، يتم تقييم أداء النماذج باستخدام مقاييس مثل متوسط الخطأ المطلق، ومتوسط الخطأ الجذر التربيعي، وما إلى ذلك.
7. التفسير والتصور: ترجمة نتائج النموذج إلى رؤى قابلة للتنفيذ، تساعد التصورات والجداول والإحصائيات الموجزة في نقل النتائج بشكل فعال.
8. النشر: تنفيذ الرؤى في حلول أو استراتيجيات واقعية، مع ضمان تنفيذ التوصيات المستندة إلى البيانات.

5. أنواع تحليل البيانات

يأتي تحليل البيانات بأشكال مختلفة، ولكل منها غرضها الخاص، وفيما يلي تفصيل واضح للأنواع الرئيسية:

• التحليل الوصفي:

يوضح لك هذا النوع ما حدث بالفعل، يتعلق الأمر بتلخيص البيانات الخام في شيء يسهل فهمه، على سبيل المثال، قد تستخدمه إحدى الشركات لمعرفة مقدار ما باعه كل موظف وما هو متوسط المبيعات، الأمر أشبه بالسؤال: ماذا حدث؟

• التحليل التشخيصي:

بمجرد معرفة ما حدث، يساعد التحليل التشخيصي في تفسير السبب، لنفترض أن المستشفى يستقبل عددًا من المرضى أكبر من المعتاد، من خلال البحث بشكل أعمق في البيانات، قد تجد أن العديد منهم يعانون من نفس الأعراض، مما يساعدك على معرفة السبب، يجب هذا التحليل على السؤال التالي: لماذا حدث ذلك؟

• التحليل التنبئي:

الآن ننتقل إلى المستقبل، يبحث التحليل التنبئي في الاتجاهات من الماضي لمساعدتك على تخمين ما قد يحدث لاحقًا، على سبيل المثال، إذا علم أحد المتاجر أن المبيعات ترتفع عادةً في أشهر معينة، فيمكنه التنبؤ بنفس الشيء للعام التالي، والسؤال هنا هو: ما الذي قد يحدث؟

• التحليل الكمي:

يتضمن التحليل الكمي استخدام التقنيات الرياضية والإحصائية لتحليل البيانات الرقمية.

مثال: يستخدم المحلل المالي التحليل الكمي لتقييم أداء الأسهم من خلال حساب النسب المالية المختلفة وإجراء الاختبارات الإحصائية.

• البحث النوعي:

يركز البحث النوعي على فهم المفاهيم أو الأفكار أو الخبرات من خلال البيانات غير الرقمية مثل المقابلات والملاحظات والنصوص.

مثال: يقوم الباحث بإجراء مقابلات مع العملاء لفهم مشاعرهم وتجاربهم مع منتج جديد، وتحليل نصوص المقابلات لتحديد الموضوعات المشتركة.

6. تطبيقات تحليل البيانات

يُعد تحليل البيانات أداة متعددة الاستخدامات ولا غنى عنها، حيث يمكن استخدامها في مختلف الصناعات والمجالات، وقد جعلتها قدرتها على استخلاص رؤى عملية من البيانات مكوناً أساسياً في عملية اتخاذ القرار وحل المشكلات، والأُن نستكشف بعض التطبيقات المهمة لتحليل البيانات كما يلي :

○ الأعمال والتسويق:

- **أبحاث السوق:** يساعد تحليل البيانات الشركات على فهم اتجاهات السوق وتفضيلات المستهلكين والمنافسة، كما يساعد في تحديد الفرص المتاحة لتطوير المنتجات واستراتيجيات التسعير وتوسيع السوق.
- **التنبؤ بالمبيعات:** يمكن لنماذج تحليل البيانات التنبؤ بالمبيعات المستقبلية استناداً إلى البيانات التاريخية والعوامل الموسمية والعوامل الخارجية، وهذا يساعد الشركات على تحسين إدارة المخزون وتخصيص الموارد.

○ الرعاية الصحية وعلوم الحياة:

- **تشخيص الأمراض:** يعد تحليل البيانات أمراً حيوياً في التشخيص الطبي، بدءاً من تفسير الصور الطبية (مثل التصوير بالرنين المغناطيسي والأشعة السينية) إلى تحليل سجلات المرضى، يمكن أن تساعد نماذج التعلم الآلي في الكشف المبكر عن الأمراض.
- **اكتشاف الأدوية:** تستخدم شركات الأدوية تحليل البيانات لتحديد المرشحين المحتملين للأدوية، والتنبؤ بفعاليتها، وتحسين التجارب السريرية.
- **علم الجينوم والطب الشخصي:** يتيح تحليل البيانات الجينومية إعداد خطط علاج شخصية من خلال تحديد العلامات الجينية التي تؤثر على قابلية الإصابة بالمرض والاستجابة للعلاج.

○ التمويل:

- **إدارة المخاطر:** تستخدم المؤسسات المالية تحليل البيانات لتقييم مخاطر الائتمان، واكتشاف الأنشطة الاحتيالية، ونمذجة مخاطر السوق.
- **التداول الخوارزمي:** يعد تحليل البيانات جزءاً لا يتجزأ من تطوير خوارزميات التداول التي تقوم بتحليل بيانات السوق وتنفيذ الصفقات تلقائياً استناداً إلى استراتيجيات محددة مسبقاً.
- **اكتشاف الاحتيال:** تستخدم شركات بطاقات الائتمان والبنوك تحليل البيانات لتحديد أنماط المعاملات غير المعتادة واكتشاف الأنشطة الاحتيالية في الوقت الفعلي.

○ **التصنيع وسلسلة التوريد:**

- **مراقبة الجودة:** تعمل تحليلات البيانات على مراقبة جودة المنتج والتحكم فيها على خطوط التصنيع، كما تساعد في اكتشاف العيوب وضمان الاتساق في عمليات الإنتاج.
- **تحسين المخزون:** من خلال تحليل أنماط الطلب وبيانات سلسلة التوريد، يمكن للشركات تحسين مستويات المخزون، وتقليل تكاليف النقل، وضمان التسليم في الوقت المناسب.

○ **العلوم الاجتماعية والأكاديمية:**

- **البحث الاجتماعي:** يقوم الباحثون في العلوم الاجتماعية بتحليل بيانات الاستطلاعات والمقابلات والبيانات النصية لدراسة السلوك البشري والمواقف والاتجاهات، ويساعد ذلك في تطوير السياسات وفهم القضايا المجتمعية.
- **البحث الأكاديمي:** يعد تحليل البيانات أمراً بالغ الأهمية للبحث العلمي في الفيزياء والأحياء والعلوم البيئية، فهو يساعد في تفسير النتائج التجريبية واستخلاص الاستنتاجات.

○ **الإنترنت والتكنولوجيا:**

- **محركات البحث :** تستخدم Google خوارزميات تحليل بيانات معقدة لاسترداد نتائج البحث وتصنيفها استناداً إلى سلوك المستخدم وأهميته.
- **أنظمة التوصية:** تستفيد خدمات مثل Netflix و Amazon من تحليل البيانات لتوصية المحتوى والمنتجات للمستخدمين استناداً إلى تفضيلاتهم وسلوكياتهم السابقة.

○ **العلوم البيئية:**

- **نمذجة المناخ:** تحليل البيانات أمر ضروري في علم المناخ، فهو يحلل درجة الحرارة وهطول الأمطار وغيرها من البيانات البيئية، ويساعد في فهم أنماط المناخ والتنبؤ بالاتجاهات المستقبلية.
- **المراقبة البيئية:** تحليل بيانات الاستشعار عن بعد يراقب التغيرات البيئية، بما في ذلك إزالة الغابات، ونوعية المياه، وتلوث الهواء.