المحور الثالث: تحليلات الأعمال وتصوير البيانات (BA & DV)

+

المحور الخامس: تنقيب البيانات (Data Mining)

I. مقدمة:

✓ **تذكير ببنية نظام ذكاء الأعمال:**

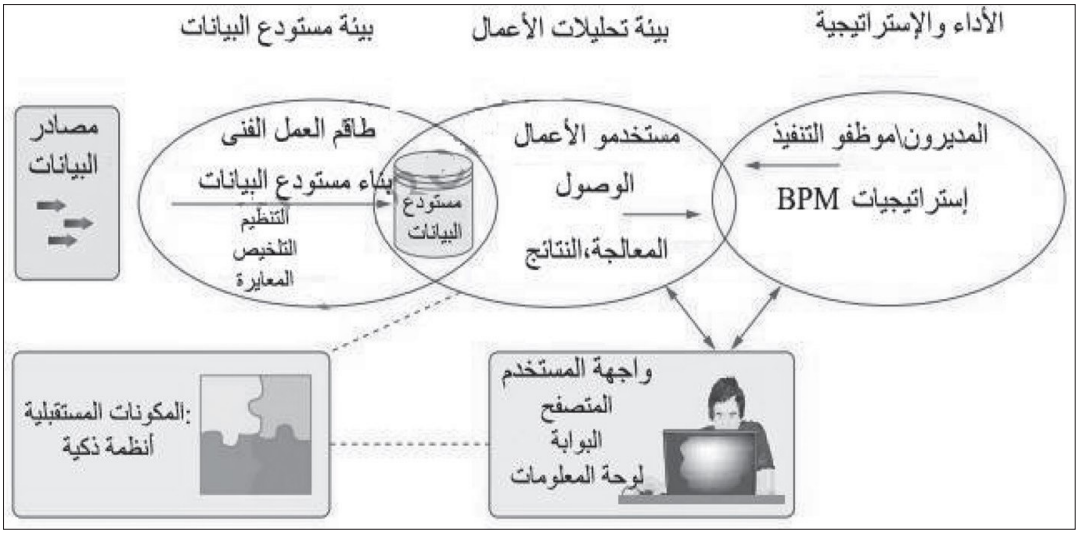
- مستودعات البيانات مع بيانات مصدرها؛
- **تحليلات الأعمال:** مجموعة أدوات معالجة وتحليل البيانات الموجودة في مستودع البيانات؛
- نظام إدارة الأعمال لمراقبة وتحليل الأداء؛
- واجهة المستخدم أو المتصفح (لوحة المعلومات).

✓ مصطلح التحليلات جاء ليحل محل "تقنيات دعم القرار المحوسبة"؛

✓ العديد من المفكرين يستعملون مصطلح "تحليل الأعمال" للدلالة على "ذكاء الأعمال"، فقد نجدهما في عدد من المراجع مترادفان؛

✓ كما أن العديد من المراجع تستعمل مصطلح "تحليل الأعمال" للدلالة على مصطلح "تحليل البيانات"؛

شكل يوضح بنية نظام ذكاء الأعمال



المصدر: صورة مأخوذة من الأنترنيت

II. تعريف تحليل الأعمال (أو التحليل الذكي للأعمال):

✓ "عملية وضع قرارات قابلة للتطبيق أو توصيات لاتخاذ إجراءات بناء على رؤى متولدة من البيانات التاريخية"؛

✓ وفقا لمعهد بحوث العمليات وعلوم الإدارة (USA) (INFORMS): "التحليلات تمثل مزيجا من تقنيات الحاسوب وتقنيات علم الإدارة والإحصاءات لحل مشكلات حقيقية"؛

III. فوائد التحليل الذكي للأعمال:

- ✓ التشخيص المستمر لوضعية المنظمة (كل الوظائف)؛
- ✓ التوفير المستمر للحلول الذكية التي تسمح باتخاذ الإجراءات المناسبة لتصحيح الأوضاع؛
- ✓ بناء سيناريوهات مستقبلية لتطوير المنظمة؛
- ✓ الحفاظ على القدرات التنافسية للمنظمة وتحسينها؛
- ✓ التحكم في عامل الزمن وجعله ميزة تنافسية للمنظمة،
- ✓ الرفع من كفاءة المنظمة في تحقيق أهدافها،
- ✓ .....

IV. النظام البيئي للتحليلات:

- ✓ الهدف من دراسة النظام البيئي لتحليل الأعمال: تقديم تصنيف للأنواع المختلفة من المشاركين في هذه السيرورة + إعطاء رؤيةً أوسع لكيفية تلاقي اللاعبين المختلفين في هذا المجال====الهدف هو الوصول إلى **نظام بيئي متصل البيانات** (Dtat Ecosystem)؛
- ✓ **النظام البيئي متصل البيانات:** بيئة تجمع البيانات من أنظمة الأعمال المختلفة وتوفر منصة لإدارة البيانات والتطبيقات لدعم اتخاذ القرارات
- ✓ يوجد أحد عشر (11) نوعًا مختلفًا من اللاعبين في النظام البيئي للتحليلات؛
- ✓ يتم تجميع مكونات النظام البيئي التحليلي في ثلاث فئات:

1. **موردو التقنية:** يضم 06 متدخلين، وتأتي إيراداتهم من تقديم التقنية؛ والحلول والتدريب لمنظمات مستخدمي التحليلات؛ حتى يتمكّنوا من توظيف هذه التقنيات بأكثر الطرق فعاليةً وكفاءةً؛
2. **مسرعات التحليلات:** تضم 04 متدخلين، تعمل أجهزةٌ التسريع مع مورّدي التقنية والمستخدمين؛
3. **منظمات مستخدمي التحليلات:** وهو جوهر النظام البيني وهو العنصر الأكثر أهمية؛

IV.1. الفئة الأولى: موردو التقنية: وتضم:

1. موردو البنية التحتية لتوليد البيانات:

- ✓ مجموعة جديدة من الشركات التي تمكّن من توليد وجمع البيانات التي يمكن استخدامها لتطوير رؤى تحليلية للمنظمة؛
- ✓ تشمل هذه المجموعة شركات متخصصة في إنشاء بنية تحتية لجميع البيانات من مصادر مختلفة؛
- ✓ من أهم المكونات الناشئة لمثل هذه البنية التحتية هي **أجهزة الاستشعار (أساس النظام البيئي للأنترنيت)**. حيث تقوم أجهزة الاستشعار بجمع كمية هائلة من البيانات بمعدل أسرع. وقد تم اعتمادها من قبل قطاعات مختلفة. مثل: الرعاية الصحية، والرياضة، والطاقة. **على سبيل المثال:** البيانات الصحية التي يتم جمعها تستخدم أجهزة الاستشعار بشكلٍ عام لتتبّع الحالة الصحية للمستخدمين. وبعض أجهزة الاستشعار الرئيسية التي تقوم بجمع المعلومات الصحية هي: **Alivecor, Google, Fitbit..**

- ✓ وبالمثل: فإنَّ صناعة الرياضة تستخدم أجهزة استشعار لجمع البيانات من اللاعبين والميدان لتطوير الإستراتيجيات وتحسين اللعب الجماعي. ومن أمثلة الشركات التي تنتجُ أجهزة الاستشعار ذات الصّلة بالرياضة: **Sports Sensors**....؛
- ✓ تؤدي أجهزة الاستشعار دورًا رئيسًا في إنترنت الأشياء (**IoT: Internet Of Things**). وتُعَدُ جزءًا أساسيًا من الكائنات الذكية. وهي **تجعلُ الاتصال بين آلةٍ وأخرى ممكنًا**. والجهات الفاعلة الرئيسة في البنية التحتية لإنترنت الأشياء: هي إنتل (**Intel**) ومايكروسوفت (**Microsoft**) وجوجل (**Google**).: وآي بي إم (**IBM**). وسيسكو (**Cissco**): وسمارت بين (**Smartbin**): وسيكو برودكتس (**SIKO products**): وأوميجا إنجنيرنج (**Omega Engineering**)، آبل (**Apple**) وسي أي بي (**SAP**).

2. موردو البنية التحتية لإدارة البيانات:

تضم هذه المجموعة جميعَ المنظمات الرئيسية التي توفر الأجهزة والبرامج التي تفسّر الأساس الجوهري لجميع **حلول إدارة البيانات**. ومن الأمثلة الواضحة على ذلك:

- ✓ كبار الجهات التي توفر البنية التحتية **لحوسبة قواعد البيانات**: **Oracle IBM, Dell, HP**,....؛ **وموزّدو حلول التخزين** مثل: **EMC** (اشترتها **Dell**)، **Netapp**...،
- ✓ الشركات التي تقدّم **منصات الأجهزة والبرمجيات المحلية** مثل: **IBM, Oracle**...،
- ✓ **موزّدو حلول البيانات الذين يقدّمون أنظمةً مستقلةً لإدارة قواعد البيانات للأجهزة والبرامج** مثل: عائلة **SQL server** من مايكروسوفت (**Microsoft**)؛
- ✓ ويندرجُ تحت هذه المجموعة أيضاً **موزّدو البرامج المتكاملة المتخصّصة** مثل: **SAP**، كما تشتملُ هذه المجموعة على منظمات أخرى. مثل موزّدي أجهزة قواعد البيانات ومقدّمي الخدمات والدامجين والمطورين وغيرهم؛

3. موردو مستودعات البيانات:

- ✓ توفر الشركات التي تركز على مستودعات البيانات **التقنية والخدمات التي تهدف إلى دمج البيانات من مصادر متعددة**. وبالتالي تمكينُ المنظمات من استخلاص وتقديم قيمة من أصول البيانات الخاصة بها؛
- ✓ وتوفر العديدَ من الشركات في هذا المجال الأجهزة الخاصة بهم لتوفير مستودعات البيانات بكفاءة واسترجاعها ومعالجتها. تعد شركاتٌ مثل: **IBM , Oracle et Teradata**من اللاعبين الأساسيين في هذا المجال؛
- ✓ تشملُ التطورات الأخيرة في هذا المجال **أداء التحليلات على البيانات مباشرةً في الذاكرة**؛

4. موردو البرمجيات الوسيطة:

- ✓ بعدما كان هدف من إنشاء مستودعات البيانات على مستوى المنظمات هو "جلب جميع مخازن البيانات إلى منصة واحدة في المنظمة"، أصبح الهدف هو إعطاء مغزى لهذه البيانات من خلال تحليلها، وأصبح بالتالي هذا الأخير صناعة في حد ذاتها؛
- ✓ يتمثل الهدف العام لقطاع البرمجيات الوسيطة في **توفير أدوات سهلة الاستخدام للتقارير أو التحليلات الوصفية** والتي تشكّل جزءًا أساسيًا من ذكاء الأعمال أو التحليلات المستخدّمة في المنظمات؛
- ✓ من أمثلة المنظمات في هذا المجال: **Microstrategy, Plum**....

5. موردو خدمة البيانات:

- ✓ يتم توليدٌ الكثير من البيانات التي تستخدمها المنظمة للتحليلات داخليًا من خلال عملياتها، لكن هناك العديدٌ من مصادر البيانات الخارجية التي تلعب دورًا رئيسًا في اتخاذ القرار بالمنظمات؛
- ✓ تشمل أمثلة مصادر البيانات هذه بيانات يتمَّ جمعها من قبل أطراف أخرى (أطراف ثالثة)، والتي يمكن أن تمد المنظمة بمعلوماتٍ مفيدة في صُنع القرار؛
- ✓ وقد انتهزت العديدٌ من المنظمات الفرصة لتطوير آليات جمع وتكامل وتوزيع البيانات المتخصّصة من أجل بيعها؛
- ✓ تركز هذه المنظمات، عادةً، على قطاع صناعي مُحدَّد وتستندٌ إلى علاقاتها القائمة في هذه الصناعة من خلال منصاتها المتخصصة وخدماتها لجمع البيانات. فمثلاً؛ توفر شركة **Nielsen** (وهي *شركة* أمريكية للمعلومات والبيانات وقياس السوق، مقرها الرئيسي نيويورك) مصادر بيانات لعملائها حول سلوك الشراء بالتجزئة للعملاء. وتقوم **Google** بتجميع بيانات لمواقع الإنترنت الفردية وتقوم بعمل مُلخص متاح من خلال خدمات تحليلات جوجل؛ وهناك مئات الشركات الأخرى التي تقوم بتطوير منصات وخدمات متخصصة لجمع وتكامل ومشاركة هذه البيانات مع عملائها.

6. مطورو البرامج التي تركز على التحليلات:

- ✓ قامت المنظمات في هذه الفئة بتطوير برامج التحليلات للاستخدام العام مع البيانات التي تمَّ جمعها في مستودعات البيانات أو المتاحة من خلال إحدى المنصات الأساسية المشار إليها سابقًا (بما في ذلك البيانات الضخمة)؛
- ✓ تشمل هذه الفئة كذلك المخترعين والباحثين في الجامعات وغيرها من المنظمات التي تعمل على تطوير الخوارزميات وابتكار أنواع مُحدَّدة من تطبيقات التحليلات؛
- ✓ يمكننا تحديدٌ كبار اللاعبين في هذا المجال باستخدام أنواع التحليلات الثلاثة: الوصفية، التنبؤية والتوجيهية (سيتم التطرق لهذه المستويات من التحليل لاحقا)؛

التحليلات الوصفية (إعداد التقارير): تتضَمَّن مجموعةٌ أدوات: **SQL Server BI** التابعة لشركة مايكروسوفت، برامج متخصصة من إعداد من شركاتٍ أخرى مثل (**Tableau**) لتصوير البيانات، كما تُقدّم **SAS** أداة التحليلات المرئية...

التحليلات التنبؤية: قد تبَنّت العديدٌ من شركات البرمجيات الإحصائية، مثل **SAS** و**SPSS**, التحليلات التنبؤية، كما طوّرت إمكانيات البرامج؛ وكذلك ممارساتٌ الصناعة لاستخدام أساليب التنقيب في البيانات والأساليب الإحصائية الكلاسيكية للتحليلات. ومن أمثلة الأدوات المُستخدّمة للتحليلات التنبؤية: **IBM SPSS Modeler** من **IBM** و**Entreprise Miner** من **SAS**... وكثيرٌ من الشركات الأخرى التي قد تبيع برامجها على نطاق واسع أو تستخدمها في ممارسات الاستشارات الخاصة بها.

التحليلات التوجيهية: يُقدَّم مُورّدو البرمجيات في هذه الفئة أدواتٍ النمذجة والخوارزميات؛ من أجل تحسينٍ أمثل للعمليات التي عادةً ما تُسمّى برمجيات علم الإدارة/ بحوث العمليات: فعلى سبيل المثال لدى **IBM** برامج البرمجة الخطية الكلاسيكية: والمختلطة والبرمجة الصحيحة.

IV.2. الفئة الثانية: مسرعات التحليلات: وتضم:

1. مطورو التطبيقات:

- ✓ تستخدم المنظمات في هذه المجموعة معرفتها الصناعية وخبرتها التحليلية وحلولها المتاحة للبنية التحتية للبيانات ومستودعات البيانات والبرامج الوسيطة ومجمّعي البيانات ومُقدّمي برامج التحليل الآلي لتطوير حلول مُخصّصة لصناعة معينة؛
- ✓ وبالتالي، تتيح هذه المجموعة الصناعية إمكانية استخدام تقنية التحليلات في صناعة معينة؛
- ✓ المعرفة التقنية بالصناعة يسمح للمنظمات التي تعمل على تطوير التطبيقات باعتماد هيكلة مناسبة للحلول التي تبحث عنها المنظمات المعنية؛
- ✓ على سبيل المثال، تقدّم شركة **Cerner** حلول السجلات الطبية الإلكترونية لمقدّمي الخدمات الطبية، وتشمل عروضها الآن العديد من تقارير التحليلات والمرئيات (البيانات في شكل صور ورسومات)، كذلك، تُقدّم شركة **IBM** محرك كشف الاحتيال لصناعة التأمين الصحي، حيث تعمل مع شركة تأمين لاستخدام منصة تحليلات **Watson** الشهيرة في مساعدة مُقدّمِي الخدمات الطبية وشركات التأمين في التشخيص وإدارة المرض.

2. محللو صناعة التحليلات والمؤثرون فيها:

تشمل هذه المجموعة ثلاث أنواع من المنظمات أو المحترفين:

النوع الأول: مجموعة المنظمات المهنية: تُقدّم المشورة إلى موردي صناعة التحليلات والمستخدمين، تشمل خدماتهم تحليلات تسويقية وتغطية التطورات الجديدة وتقييم تقنيات مُحدّدة وتطوير التدريب وغيرها من الخدمات، ومن أمثلة هذه المجموعة منظمات مثل مجموعة جارتنر ومعهد مستودعات البيانات.، و**Forrester** و**Mckinsey** والعديد من المنشورات العامة والفنية ومواقع الويب التي تغطي صناعة التحليلات.

النوع الثاني: مجموعة منظمات مهنية تقدّم أيضاً بعضًا من الخدمات نفسها، ولكنها قائمةٌ على أساس منظم. فعلى سبيل المثال: يركّز معهد بحوث العمليات وعلم الإدارة **INFORMS**، وهي شركةٌ مهنية، على تعزيز التحليلات، وتركّز مجموعة المصالح الخاصة لدعم القرار والتحليلات، وهي مجموعةٌ فرعيةٌ تابعةٌ لرابطة نظم المعلومات، أيضاً على التحليلات. كما أنَّ لدى معظم المورّدين الرئيسيين (مثل: تيراداتا و**SAS**) مجموعات المستخدمين الخاصة بهم،

تشجّع هذه الكيانات استخدامَ التحليلات وتمكّن من مشاركة الدروس المستفادة من خلال منشوراتها ومؤتمراتها؛

النوع الثالث: تشمل المختصِين في مجال التحليلات؛ هي ما نسميه سفراء التحليلات أو المؤثرين أو القائمين بالدعاية، وقد أبدى هؤلاء المحلّلون حماسهم للتحليلات من خلال حلقاتهم الدراسية وكتبهم ومنشوراتهم العلمية؛

3. المعاهد الأكاديمية ووكالات الاعتماد:

دور الجامعات في إعداد الإطارات والتطبيقات في مجال تحليل الأعمال + برامج الشهادات التي تمنحُ شهادة الخبرة في برامج مُحدّدة، والتي يقدّمها كل موردي التقنية الأساسية تقريبًا [أي بي إم (**IBM**): ومايكروسوفت (**Microsoft**) وأوراكل (**Oracle**): وإس إيه إس (**Sas**) وغيرها من المنظمات. وتضمّن هذه الشهاداتٌ أن يتمتع الموظفون المحتملون الجُّدد بمستوّى معين من المهارات التقنية؛

4. المنظمون وصناع السياسات:

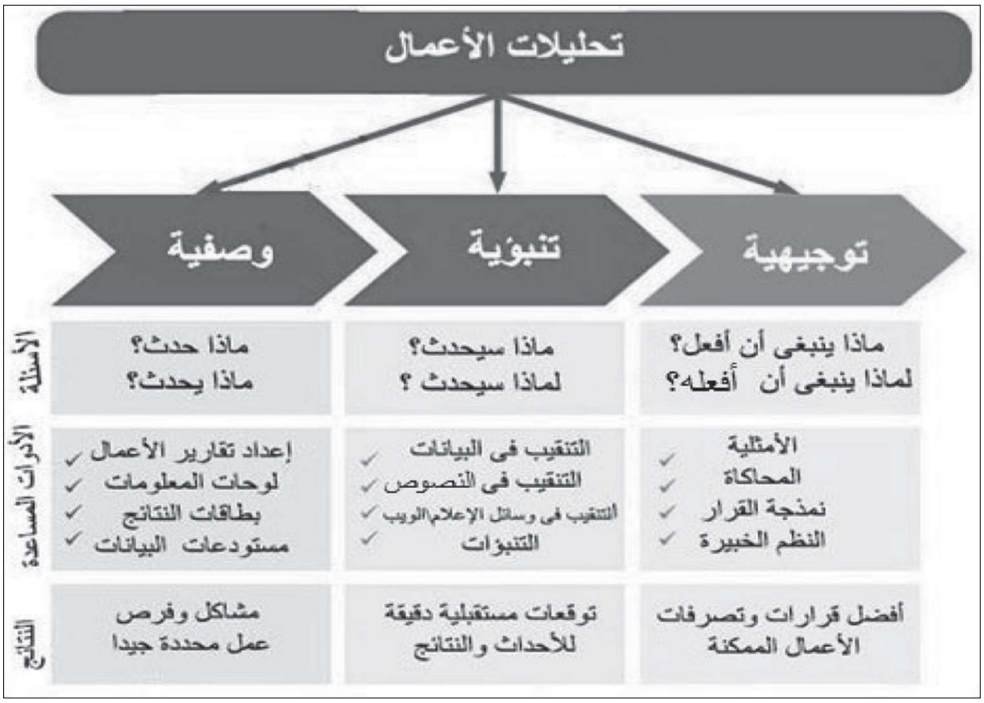
- ✓ يعتبر اللاعبون في هذا المكوّن مسؤولين عن تحديد القواعد واللوائح الخاصة بحماية موظفي وعملاء وحاملي أسهم منظمات التحليلات، حيث يتطلّب جمعٌ ومشاركةٌ بيانات العميل قوانينَ صارمةً لتأمين الخصوصية.
- ✓ تقوم عدة منظمات في هذا المجال بتنظيم عملية نقل البيانات وحماية حقوق المستخدمين، فعلى سبيل المثال: تَنظّم لجنة الاتصالات الفيدرالية (FCC) الاتصالات بين الدول والاتصالات الدولية؛
- ✓ بالمثل. فإن لجنة التجارة الفيدرالية (FTC) هي المسؤولة عن عرض الممارسات التجارية غير المشروعة المتعلقة بالبيانات؛
- ✓ ومن ناحية أخرى؛ يُوجَّد وكالةٌ فيدراليةٌ غير نظامية تُسمَّى المعهدُ الوطني للمعايير والتقنية (NIST). تساعد على تقدِّم البنية التحتية للتقنية. كما أنَّ هناك العديدَ من المنظمات الأخرى في جميع أنحاء العالم التي تنظم أمن البيانات وتدفع عجلة صناعة التحليلات. وهذا مكونّ مهم في النظام البيني؛

IV.3. الفئة الثالثة منظمات مستخدمي التحليلات:

كل المنظمات مهما كان نوع نشاطها

V. مستويات تحليل الأعمال:

شكل يوضح مستويات ذكاء الأعمال



المصدر: راميش شاردا وآخرون، 2020، ذكاء الأعمال والتحليلات وعلم البيانات، منظور إداري، ترجمة محمد بن عايض القرني، مركز البحوث والدراسات، معهد الإدارة العامة، المملكة العربية السعودية، ص 58

أولا، التحليلات الوصفية:

- ✓ معرفة ما يحدث في المنظمة + فهم بعض الاتجاهات والأسباب الكامنة وراء بعض الوضعيات؛
- ✓ من أجل ذلك: ضرورة توفر جميع البيانات ذات الصلة في شكل يتيح إعداد التقارير والتحليلات المناسبة؛

✓ وتقوم البيانات الوصفية بِوَصّف بنية البيانات وبعض معانيها، مما يُسهم في استخدامها الفعّال أو غير الفعال.

✓ **تقنيات التحليل الوصفي للبيانات في مستودع البيانات (OLAP et OLTP):**

- للتحليلات الوصفية فرعين رئيسيين هما: الإحصاءات ومعالجة التحليلات عبر الانترنيت (OLAP)، حيث يعد OLAP هو المصطلح المستخدّم لتحليل وتوصيف وتلخيص البيانات المهيكلة المخزّنة في قواعد البيانات التنظيمية (غالباً ما يتم تخزينها في مستودع البيانات أو في سوق البيانات)؛ بينما تساعد الإحصائيات في توصيف البيانات إمّا على متغير واحد في كلّ مرة أو متغيرات متعددة معا باستخدام إما أساليب وصفية أو استنتاجية.
- OLAP: On-Line Analytical Processing Systems هو نظام يدعم تحليل البيانات متعدد الأبعاد بسرعات عالية على أحجام ضخمة من البيانات. هذه البيانات بشكل عام من مستودع بيانات أو سوق بيانات أو أي مخزن بيانات آخر. يساعد OLAP في تحليل البيانات التاريخية وفهمها وهو مفيد لمختلف وظائف المنظمة.
- **بنية نظام المعالجة التحليلية عبر الأنترنيت (OLAP):**

1. مستودع البيانات؛
2. ELT: أدوات الاستخراج، التحويل والتحميل؛
3. خادم المعالجة التحليلية عبر الأنترنيت: الجهاز الأساسي الذي يقوم بتشغيل نظام المعالجة التحليلية عبر الإنترنت (OLAP). يستخدم أدوات الاستخراج والتحويل والتحميل (ETL) لتحويل المعلومات في قواعد البيانات العلائقية وتجهيزها لعمليات المعالجة التحليلية عبر الإنترنت(OLAP) .
4. قاعدة بيانات المعالجة التحليلية عبر الانترنيت: هي قاعدة بيانات منفصلة تتصل بمستودع البيانات؛
5. مكعبات المعالجة التحليلية عبر الانترنيت: مكعب البيانات هو نموذج يمثل مجموعة متعددة الأبعاد من المعلومات، ومعظم مكعبات البيانات يكون لها أكثر من ثلاثة أبعاد. مكعبات المعالجة التحليلية عبر الإنترنت (OLAP)هي مكعبات ثابتة لأنه بمجرد تصميمها لا يمكنك تغيير الأبعاد والبيانات الأساسية. على سبيل المثال، إذا قمت بإضافة بُعد المستودع إلى مكعب يتضمن أبعاد المنتج والموقع والوقت، فيجب عليك إعادة تشكيل المكعب بالكامل .
6. أدوات المعالجة التحليلية عبر الانترنيت؛

- **كيف يعمل نظام المعالجة التحليلية عبر الانترنيت (OLAP):**

1. يجمع خادم المعالجة التحليلية عبر الإنترنت (OLAP) البيانات من مصادر بيانات متعددة، بما في ذلك قواعد البيانات العلائقية ومستودعات البيانات.
2. بعد ذلك، تقوم أدوات الاستخراج والانتقال والتحميل (ETL)بتنظيف البيانات وتجميعها وحسابها مسبقًا وتخزينها في مكعب المعالجة التحليلية عبر الإنترنت (OLAP) وفقًا لعدد الأبعاد المحددة.
3. يستخدم محللو الأعمال أدوات المعالجة التحليلية عبر الإنترنت (OLAP) في الاستعلام وإنشاء تقارير من البيانات متعددة الأبعاد في مكعب المعالجة التحليلية عبر الإنترنت.(OLAP)

- **ما هي أنواع المعالجة التحليلية عبر الإنترنت(OLAP) ؟**

1. **المعالجة التحليلية متعددة الأبعاد عبر الإنترنت(MOLAP) :** تتضمن المعالجة التحليلية متعددة الأبعاد عبر الإنترنت (MOLAP)إنشاء مكعب بيانات يمثل بيانات متعددة الأبعاد من مستودع بيانات .يخزن نظام المعالجة التحليلية متعددة الأبعاد عبر الإنترنت (MOLAP) البيانات المحسوبة مسبقًا في المكعب الزائدي. يستخدم مهندسو البيانات المعالجة التحليلية متعددة الأبعاد عبر الإنترنت (MOLAP) لأن هذا النوع من تقنية المعالجة التحليلية عبر الإنترنت (OLAP) يوفر تحليلًا سريعًا .
2. **المعالجة التحليلية العلائقية عبر الإنترنت(ROLAP) :** بدلاً من استخدام مكعب بيانات، تسمح المعالجة التحليلية العلائقية عبر الإنترنت (ROLAP) لمهندسي البيانات بإجراء تحليل بيانات متعدد الأبعاد على قاعدة بيانات علائقية. بمعنى آخر، يستخدم مهندسو البيانات استعلامات SQL للبحث عن معلومات محددة واسترجاعها بناءً على الأبعاد المطلوبة. المعالجة التحليلية العلائقية عبر الإنترنت (ROLAP) هي مناسبة لتحليل البيانات الشاملة والتفصيلية. ومع ذلك، فإن المعالجة التحليلية العلائقية عبر الإنترنت (ROLAP) يكون أداؤها في الاستعلامات بطيئًا مقارنةً بالمعالجة التحليلية متعددة الأبعاد عبر الإنترنت .(MOLAP)
3. **المعالجة التحليلية الهجينة عبر الإنترنت (HOLAP) :** تجمع المعالجة التحليلية الهجينة عبر الإنترنت (HOLAP) بين MOLAP و ROLAP لتوفير أفضل ما في كلا البنيتين. تسمح المعالجة التحليلية الهجينة عبر الإنترنت (HOLAP) لمهندسي البيانات باسترجاع النتائج التحليلية بسرعة من مكعب البيانات واستخراج المعلومات التفصيلية من قواعد البيانات العلائقية .

- **ما عمليات المعالجة التحليلية عبر الإنترنت(OLAP) ؟**

1. **تجميع:** في عملية التجميع، يقوم نظام المعالجة التحليلية عبر الإنترنت (OLAP) بتلخيص البيانات الخاصة بسمات محددة. بمعنى آخر، يعرض بيانات أقل تفصيلاً. على سبيل المثال، يمكنك عرض مبيعات المنتجات وفقًا لمدن نيويورك وكاليفورنيا ولندن وطوكيو. توفر عملية التجميع عرضًا لبيانات المبيعات استنادًا إلى دول، مثل الولايات المتحدة والمملكة المتحدة واليابان .
2. **التنقيب:** التنقيب هو عكس عملية التجميع. يدخل محللو الأعمال إلى أسفل التسلسل الهرمي للمفاهيم ويستخرجون التفاصيل التي يحتاجونها. على سبيل المثال، يمكنهم الانتقال من عرض بيانات المبيعات حسب السنوات إلى عرضها حسب الأشهر.
3. **التقطيع:** يستخدم مهندسو البيانات عملية التقطيع لإنشاء عرض ثنائي الأبعاد من مكعب المعالجة التحليلية عبر الإنترنت .(OLAP) على سبيل المثال، يقوم مكعب المعالجة التحليلية متعددة الأبعاد عبر الإنترنت (MOLAP) بفرز

البيانات وفقًا للمنتجات والمدن والأشهر. من خلال تقطيع المكعب، يمكن لمهندسي البيانات إنشاء جدول يشبه جدول البيانات يتكون من منتجات ومدن لشهر معين .

4. **التكعيب:** يستخدم مهندسو البيانات عملية التكعيب لإنشاء مكعب فرعي أصغر من مكعب المعالجة التحليلية عبر الإنترنت .(OLAP) إنها تحدد الأبعاد المطلوبة وتنشئ مكعبًا أصغر من المكعب الأصلي.
5. **التمحور:** تتضمن عملية التمحور تدوير مكعب المعالجة التحليلية عبر الإنترنت (OLAP) على طول أحد أبعاده للحصول على منظور مختلف لنموذج البيانات متعدد الأبعاد. على سبيل المثال، يحتوي مكعب المعالجة التحليلية عبر الإنترنت (OLAP) ثلاثي الأبعاد على الأبعاد التالية على المحاور المعنية:

- المحور — X المنتج
- المحور — Y الموقع
- المحور — Z الوقت

عند المحور، يحتوي مكعب المعالجة التحليلية عبر الإنترنت (OLAP) على التكوين التالي:

- المحور — X الموقع
- المحور — Y الوقت
- المحور — Z المنتج

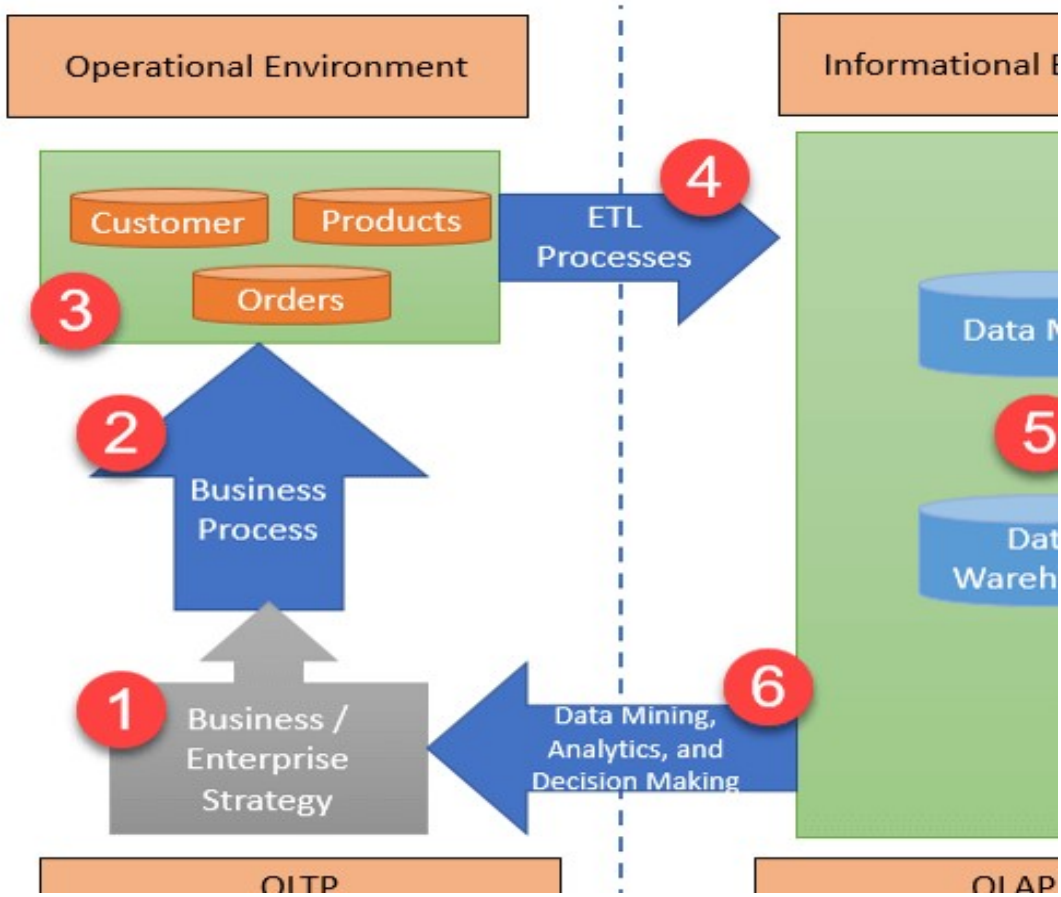
- **معالجة المعَامَلات عبر الانترنيت OLTP:**

1. يُستَخْدَم مصطلح **OLPT** (On-Line Transactional Processing) (وهو نظام معالجة المعَامَلات عبر الإنترنت) في نظام المعَامَلات المسؤول بشكل أساسيٌ عن التقاط وتخزين البيانات المتعلّقة **بوظائف العمل اليومية**. مثل:**ERP, CRM** ...، وعليه، فهو موجود في المستوى التشغيلي للمنظمة لمعالجة المعاملات اليومية للمنظمة؛
2. ويتعامل نظام **OLTP** مع احتياجات الأعمال الحساسة وأداء معاملات الأعمال اليومية بطريقة أوتوماتيكية وتشغيل التقارير فوريًا والتحليل الروتيني. (**مثال:** شخص يريد سحب مبلغ مالي من حسابه؟ **OLTP** يسمح بالتأكد الفوري من وجود رصيد + من ألا يتجاوز السقف المسموح به.....)

3. **OLTP** est une catégorie de traitement de données axée sur **des tâches orientées transactions**. **(www.oracle.com)**
4. **OLTP** implique généralement l'insertion, la mise à jour et/ou la suppression de petites quantités de données dans une base (opérations bancaires, achats de biens, billets, réservations). **(www.oracle.com)**

5. يعتمد كلّ من OLAP و OLTPبشكلٍ كبيرٍ على بعضهما؛ إذ يستخدمOLAP البيانات التي تم التقاطها بواسطة OLTP ويقوم هذا الأخير آليًا بإجراء عمليات الأعمال التي تتم إدارتها بواسطة القرارات التي يدعمها الأول؛

شكل يوضح الفرق بين OLP et OLTP



https://www.guru99.com/what-is-oltp.html

✓ تصوير البيانات (Data visualization):

- يرتبط تصويرٌ البيانات ارتباطًا وثيقا بمجالات الرسومات البيانية للمعلومات "وتصوير المعلومات والتصوّر العلمي والرسومات البيانية الإحصائية".
- حتى وقتٍ قريب، كانت الأشكال الرئيسة لتصوير البيانات المتاحة في كلّ من تطبيقات ذكاء الأعمال تتضمّن الجداول والرسوم البيانية؛ إضافةً إلى الأنواع الأخرى من العناصر المرئية المستخدّمة لإنشاء بطاقات الأداء ولوحات البيانات.
- يتم تصوير البيانات بالاعتماد على خوارزميات متخصصة....

ثانيا، التحليلات التنبؤية:

✓ ماذا سيحدث؟ ولماذا سيحدث؟

✓ **التحليلات التنبؤية هي فرع من التحليلات المتقدمة** التي تقدم تنبؤات بشأن الأحداث والسلوكيات والنتائج المستقبلية. وهي تستخدم تقنيات إحصائية - بما في ذلك خوارزميات تعلم الآلة(التعلم الآلي للآلة) والنمذجة التنبؤية المتطورة - لتحليل البيانات الحالية والتاريخية وتقييم احتمالية حدوث شيء ما، حتى لو لم يكن هذا الأمر على رادار الأعمال .

(https://www.sap.com/mena-ar/products/technology-platform/cloud-analytics/what-is-predictive-analytics.html)

التعلم الآلي للآلة: هو مجموعة فرعية من الذكاء الاصطناعي .(AI) وهو يركز **على تعليم أجهزة الكمبيوتر للتعلم** من البيانات والتحسين مع الخبرة - بدلا من برمجتها صراحة للقيام بذلك. في تدريب الآلة، يتم كتابة الخوارزميات التي تمكن من العثور على الأنماط والارتباطات في مجموعات البيانات الكبيرة وعلى اتخاذ أفضل القرارات والتنبؤات على أساس ذلك التحليل.

✓ باستخدام البيانات والمعالجة الكافيتين للتحليلات الوصفية - يمكن أن تبدأ أدوات تحليلات الأعمال في إنشاء نماذج تنبؤية تستند إلى الاتجاهات والسياق التاريخي. ومن ثم يمكن استخدام هذه النماذج للاسترشاد بها في اتخاذ القرارات المستقبلية المتعلقة بالخيارات التجارية والتنظيمية؛ **(www.oracle.com)**

✓ **التنقيب في البيانات (Data mining):**

1. التنقيب في البيانات هي تقنية لإعداد تحليلات الأعمال والتحليلات التنبؤية، بصفة خاصة؛
2. بصفة عامة؛ يُعَد التنقيب في البيانات طريقة لتطوير المعلومات الاستخبارية (أي: المعلومات أو المعرفة القابلة للتطبيق) من البيانات التي تعنى المنظمة بجمعها وتنظيمها وتخزينها؛
3. التنقيب في البيانات هو: تقنية تهدف إلى استنتاج المعرفة من كميات هائلة من البيانات، تعتمد على الخوارزميات الرياضية والتي تعتبر أساس التنقيب عن البيانات وهي مستمدة من العديد من العلوم مثل علم الإحصاء والرياضيات والمنطق وعلم التعلم، والذكاء الاصطناعي والنظم الخبيرة، وغيرها من العلوم الذكية وغير التقليدية؛
4. وعليه، إنَّ التنقيبَ في البيانات ليس نظامًا جديدًا؛ بل هو تعريفٌ جديدٌ لاستخدام العديد من التخصّصات؛ إذ يتم وَضْعِ بيانات الاستخراج بشكلٍ مُحكّم عند تلاقي العديد من التخصصات، والتي تشمل الإحصاء والذكاء الاصطناعي وتعلم الآلة وعلوم الإدارة ونظم المعلومات وقواعد البيانات، إذ إنَّ عملية التنقيب في البيانات تسعى جاهدةٌ إلى التحسين في استخراج المعلومات والمعارف المفيدة من قواعد البيانات الكبيرة عن طريق استخدام التقدُّم في جميع التخصّصات السابقة؛
5. هو خطوة واحدة في اكتشاف المعرفة (أحد خطوات اكتشاف المعرفة القابلة للتطبيق) داخل قواعد البيانات، حيث تستخدم أساليب التحليل مثل: الشبكات العصبية Neural Networks أو الخوارزميات الجينية Genetic Algorithmsأو أشجار القرارات Decision Trees والتقنية والنماذج الهجينة Hybrid Models لتحديد الأنماط والعلاقات في مجموعات البيانات.

6. يعرف التنقيب في البيانات على أنه عملية تحليل كمية بيانات (عادة ما تكون كمية كبيرة) لإيجاد علاقة منطقية تلخص البيانات بطريقة جديدة تكون مفهومة ومفيدة لصاحب البيانات (المنظمات)؛

1. عمليا، قدمت مؤسسة IGI Global (مؤسسة نشر علمي) 82 تعريفا لمفهوم التنقيب في البيانات؛

2. **وعموما، التنقيب في البيانات هو خطوة أساسية لتطبيق أساليب ذكية بهدف الكشف عن أنماط البيانات المثيرة للاهتمام والمخبأة في مجموعات البيانات الكبيرة؛**

3. تستخدمُ المنظمات مجموعةً واسعة من تقنيات التنقيب في البيانات للحصول على فهم أفضل لعملائها ولعملياتها ولحل المشكلات التنظيمية المعقدة؛

4. ظهر التنقيب في البيانات (Data mining) في أواخر الثمانيات وأثبت وجوده كأحد الحلول الناجحة لتحليل كميات ضخمة من البيانات، وذلك بتحويلها من مجرد معلومات متراكمة وغير مفهومة (بيانات) إلى معلومات قيمة يمكن استغلالها والاستفادة منها بعد ذلك؛

5. تقنيات التنقيب في البيانات جعل المنظمة غير مضطرة للتخلص من جميع البيانات المتعلقة بنشاطها (بيانات العملاء مثلا)، لأن هذه التقنية جعلت الحصول على المعرفة الضرورية القابلة للتطبيق والمفيدة للمنظمة ممكن جدا حتى ولوكان حجم البيانات كبير جدا جدا جدا (خاصة وأن تكلفة الاحتفاظ الالكتروني بالبيانات والمعلومات جعل هذه العملية غير مكلفة).

6. **لماذا زيادة الاهتمام بالتنقيب في البيانات؟**

- زيادة حِدَّة المنافسة على النطاق العالمي تقودها الاحتياجات المتغيرة باستمرار للعملاء؛
- الاعتراف العام بالقيمة غير المستغلة المخفية من مصادر البيانات الكبيرة؛
- توحيد وتكامل سجلات قواعد البيانات، مما يتيح رؤية واحدة للعملاء، الموردين...
- توحيد قواعد البيانات ومستودعات البيانات الأخرى في مكان واحد في شكل مستودع البيانات؛
- الزيادة المتسارعة في تقنيات المعالجة ومستودعات البيانات؛
- الانخفاض الكبير في تكلفة الأجهزة والبرمجيات الخاصة بمستودعات البيانات ومعالجتها؛
- ....

7. **كيف يعمل التنقيب في البيانات:**

- من خلال استخدام البيانات الحالية ذات الصّلة والتي تم الحصولُ عليها من داخل وخارج المنظمة؛ يستطيع التنقيبُ في البيانات أن يبني نماذج لاكتشاف الأنماط بين السّمات أو الخصائص المعروضة في مجموعة البيانات؛
- **هذه النماذجُ هي: التمثيلات الرياضية** التي توضح (العلاقات/ الارتباطات الخطية البسيطة أو العلاقات المعقّدة وغير الخطية) بين البيانات المتوفرة في مستودع البيانات،
- تحَدّد هذه النماذج (التمثيلات الرياضية) أربع (04) أنواع رئيسية من أنماط العلاقات أو الارتباطات بين البيانات:

1. **تحديد الترابط في سمات أو خصائص** مجموعات مشتركة من الأشياء (العملاء أو المنتجات) (**الأنماط التفسيرية**)؛
2. **التنبؤات** والتي تُوضّح طبيعة الأحداث المستقبلية لأحداث معينة بناءً على ما حدث في الماضي، مثل: التنبؤ بالحرارة القصوى في يوم معين.
3. **التجميع** ويعملُ على تحديدِ المجموعات الطبيعية للأشياء استنادًا إلى خصائصها المعروفة مثل: تعيين العملاء في شرائح مختلفة بناءً على التركيبة السكانية وسلوكيات الشراء السابقة؛
4. **العلاقات التسلسلية**، وتعمل على كَشْف علاقات الأحداث الزمنية، مثل التنبؤ بأن عميلا مصرفيًا موجودًا لديه بالفعل حساب فحص سيفتح حساب توفير متبوعًا بحساب استثمار في غضون عام؛

13. تقنيات التنقيب في البيانات:

هناك العديد من طرق التنقيب عن البيانات، من أهمها:

1. تحليل الارتباط؛
2. الخوارزميات الجينية (Genetic Algorithms): تحاكي الجينات البشرية في عملها (الجينات البشرية) (مثلا في البحث عن الحل الأمثل أو عن البيانات المناسبة)، ظهرت سنة 1975 على يد العالم جون هولاند في جامعة ميتشيغان .
3. الشبكات العصبية.
4. التحليل الاحصائي.
5. شجرة القرارات.
6. .....

14. تحليلات النص والتنقيب فيه:

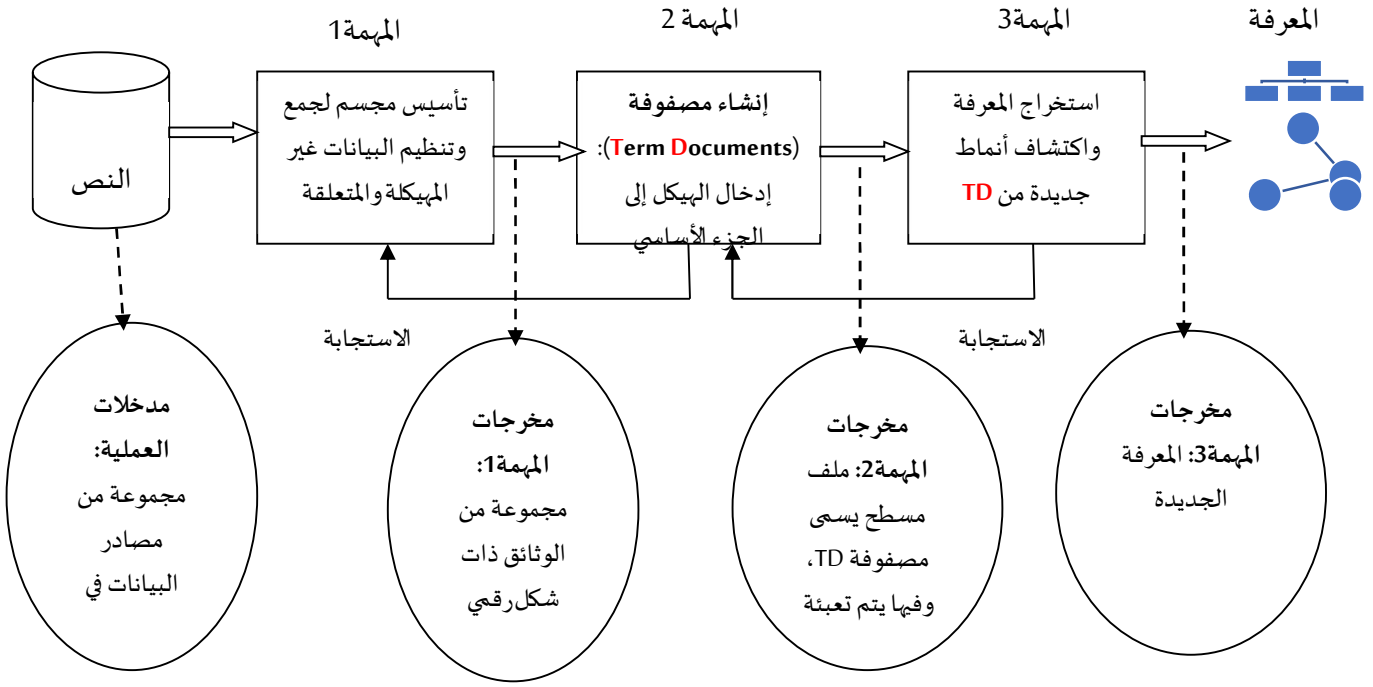
- اكتشاف بيانات من خلال نصوص (تغريدات) موجودة في الواب أو في وسائل التواصل الاجتماعي أو هو اكتشاف المعرفة من خلال قواعد البيانات النصية؛
- الشركات التي تستفيد بكفاءة وبشكل فعال من **مصادر البيانات النصية** يكون لديها المعرفة اللازمة لاتخاذ قرارات أفضل، مما يجعلها تتميز في المنافسة عن تلك الشركات التي تتخلف عن الركب.
- **هل يوجد فرق بين تحليل النص والتنقيب فيه؟** في العموم يستعملان مرادفان لبعضهما، إلا أنه يوجد فرق جوهري بين المصطلحين، فالأول يشمل الثاني: تحليل النص = استرجاع المعلومات + التنقيب في النص
- **الهدف من تحليلات النصوص والتنقيب في النص**: هو تحويل البيانات النصية غير المهيكلة (الأقل تنظيما) إلى معلومات قابلة للتطبيق من خلال تطبيق معالجة اللغات الطبيعية (**NLP**)والتحليلات؛
- **مثال عن التنقيب في النص:** الطلب من العملاء إبداء آرائهم حول منتج معين دون تقييدهم بنموذج معين لذلك، حيث تعتبر هنا المعلومات المحصلة في كل نص مكتوب غير خاضعة لنمط معين أو هيكلة معينة، وعليه يتم اللجوء إلى تقنيات التنقيب في النص من أجل الاستفادة من آراء الزبائن؛

معالجة اللغات الطبيعية (NLP: Natural Language Processing): أحد المكونات المهمة في التنقيب في النصوص، وهي عبارة عن مجال فرعي من الذكاء الاصطناعي وعلوم لغات الحاسوب. كما أنها تدرس مشكلة "فهم" لغة الإنسان الطبيعية، من وجهة نظر تحويل الصور البشرية (كالوثائق النصية) إلى تمثيلات أكثر رسمية (في شكل بيانات رقمية ورمزية) تسهل على برامج الحاسب إجراء المعالجات المطلوبة ببراعة.

معالجة اللغات الطبيعية(NLP) : عبارة عن تقنية تعلّم الآلة تمكّن أجهزة الكمبيوتر من تفسير اللغة البشرية ومعالجتها وفهمها.

- وتعد إدارة علاقات العملاء (CRM)هي أحد المجالات البارزة التي تظهر فيها فوائد معالجة اللغات الطبيعية (NLP). وبوجه عام، فإن (CRM) يسعى إلى تقدير العملاء من خلال فهم أفضل واستجابة فعالة لاحتياجاتهم الفعلية والمتوقعة، ويُعد تحليل المشاعر أحد المجالات المهمة في إدارة علاقات العملاء (CRM)، حيث يكون لمعالجة اللغات الطبيعية (NLP)تأثير كبير في تحليل المشاعر. وتحليل المشاعر هو تقنية تستخدم لكشف الآراء المؤيدة وغير المؤيدة تجاه منتجات وخدمات معينة باستخدام عدد كبير من مصادر البيانات النصية (ملاحظات العملاء في شكل منشورات على شبكة الإنترنت).
- **عملية التنقيب في النص ثلاثية الخطوات والمهام:**

شكل يوضح الخطوات والمهام الثلاث للتنقيب في النص



المصدر: راميش شاردا وآخرون، 2020، ذكاء الأعمال والتحليلات وعلم البيانات، منظور إداري، ترجمة محمد بن عايض القرني، مركز البحوث والدراسات، معهد الإدارة العامة، المملكة العربية السعودية، ص 452.

- **المهمة1:** الغرض الرئيس من نشاط المهمة الأول هو جمع كل الوثائق المتعلقة بالسياق (مجال الاهتمام) الذي تتم دراسته. وقد تتضمن هذه المجموعة الوثائق النصية ورسائل البريد الإلكتروني وصفحات الويب والملاحظات القصيرة. بالإضافة إلى البيانات النصية المتوفرة، ويمكن أيضًا تسجيل التسجيلات الصوتية باستخدام خوارزميات التعرف على الكلام وجعلها جزءًا من مجموعة النصوص.

وبمجرد تجميعها، يتم تحويل الوثائق النصية وتنظيمها بطريقة تجعلها كلها في **نفس الشكل التمثيلي** (مثل الملفات النصية **ASCII**)، قابلة للمعالجة بالحاسب.

- **المهمة2:** وفي هذه المهمة، يتم استخدام الوثائق الرقمية والمنظمة (المجموعة) لإنشاء مصفوفة (**TD**) (تسمى أيضا **TDM**) حيث تمثل الصفوف في الوثائق. أما الأعمدة فتمثل المصطلحات:

	مخاطر الاستثمار	إدارة المشروع	هندسة البرمجيات	التطوير	SAP
وثيقة1	1			1	
وثيقة2		1			
وثيقة3			3		1
وثيقة4		1		2	
وثيقة5	1			1	
.....					

ملاحظة هامة جدا: عدد تكرارات المصطلح قد لا يدل دائما على أهميته في السياق (المجال) المدروس.

- **المهمة3:** استخراج أنماط جديدة (معرفة جديدة) في سياق المشكلة المحددة التي يتم تناولها بالاعتماد على مصفوفة (**TDM**)؛

أما الطرق الرئيسية لاستخلاص المعرفة الجديدة فهي:

1. **التصنيف:** تصنيف المعلومات المستخرجة حسب الموضوع مثلا، حيث تكون **الأصناف محددة مسبقا** (معرفة جديدة ناتجة عن هذا التصنيف)؛
2. **التجميع:** تجميع المعلومات المستخرجة حسب الموضوع مثلا، في مجموعات ذات مغزى ولكنها **غير محددة مسبقا** (معلومات مجمعة تعطي معرفة جديدة)؛
3. **الاقتران (الارتباط):** لاكتشاف علاقات تجذب الاهتمام بين المتغيرات في قواعد البيانات الكبيرة. وتتمثل الفكرة الرئيسية في إنشاء قواعد الارتباط (أو حل مشكلات سلة السوق) في تحديد المجموعات المتكررة التي تتماشى معًا (معرفة جديدة من خلال الوصول إلى علاقات الارتباط الموجودة)؛
4. **تحليل الاتجاهات:** الاتجاه العام الذي تأخذه سلوكيات الزبائن مثلا (سلوك الزبائن، دوافع الشراء...)

ثالثا، التحليلات التوجيهية (الارشادية):

✓ ما لذي ينبغي فعله؟ لماذا ينبغي فعله؟

تعتمد التحليلات التوجيهية على **النمذجة** بشكل أساسي (البرمجة الخطية، المحاكاة، نماذج الاستشراف...)؛

قائمة المصادروالمراجع:

1. راميش شاردا وآخرون، 2020، ذكاء الأعمال والتحليلات وعلم البيانات، منظور إداري، ترجمة محمد بن عايض القرني، مركز البحوث والدراسات، معهد الإدارة العامة، المملكة العربية السعودية.
2. oracle.com
3. https://aws.amazon.com/ar/what-is/olap
4. https://www.guru99.com/what-is-oltp.html
5. أياد زعرور، 2015، التنقيب في البيانات والتوقعات الاقتصادية، المعهد الوطني لإدارة، لبنان.
6. Sofiane AOUAG, 2024, Cours Business Intelligence, Département Informatique Faculté des Mathématique et de l'Informatique, Université de Batna II.