وفرت البيانات أو البيك ديتا هي إحدى العوامل الأساسية المؤثرة بالثورة الصناعية الرابعة، ووفرة البيانات سوف تفتح الباب أمام تطبيقات جديدة بمجال الذكاء الإصطناعي أو والتعلم الآلي، أو تقريبا كل مجال ممكن نفكر فيه من الصناعة إلى الزراعة إلى النقل، بمعنى آخر البيانات سوف تكون النفط القادم للإقتصاد العالمي. من المتوقع أن الطلب على المهارات المرتبطة بمعالجة البيانات سوف يزداد بالسنوات القادمة، بما إنه علوم البيانات والتعلم الآلي هي مجالات جديدة نسبيا لم يتم دمجها بعد بمخططات التعليم الجامعي ببرامج الماجستير والبكالوريوس، من خلال تجربتي لاحظت إنه لا يوجد الكثير من المعرفة حول علم البيانات، بما في ذلك بين العاملين بمجال التكنولوجيا. البيانات أو الديتا؟ هي كميتها إلى من المعلومات المسجلة رقميا عن نوع العالم؟ من أهم المصادر وسائل التواصل الاجتماعي مثل فيسبوك لندن إنستغرام واتساب كمان في تطبيقات مثل أوبر وأمازون بتساعد على تجميع الكثير من المعلومات والبيانات عنا إضافة لهالشيء أو إنترنت الأشياء. حيخلي كثير أجهزة تكون موصلة على الإنترنت. يقدر إنه بنهاية العام 2020 أكثر من 40, 000 جهاز حيكون موصول على الإنترنت مثل مثلا جهاز اللي بنستعمله للتحكم بالحرارة بغرفة معينة. وهذه المعلومات حتؤدي إلى خلق كثير من البنات اللي ممكن نحللها، خلينا نشرح ماذا نعني بعلم البيانات أو الداتا سينس. علم البيانات لديه الكثير من التعريفات، يعني هلا إذا رحنا على محرك جوجل وحطينا بيتا سينس أو علم البيانات حنلاقي كثير نتائج وممكن نضيع إذا جربنا فعل ا نفهم ماذا نعني بعلم البيانات. لازم نعرف إنه علم البيانات يتضمن تحليل وصفي وتنبئي وإلزامي. التحليل الوصفي هو لما ناخد كمية كبيرة أو هائلة من البيانات ونستخرج معرفة من هذه البيانات، ونقدر نعرض هذه المعرفة من خلال عروض أو من خلال بطريقة مبسطة وسهلة للمستخدم الأخير. تحليل التنبئي هو لما نستخدم البيانات الموجودة حاليا للتربء بالمستقبل مثل ا أنا عندي بيانات بتفرجيني للمرضى اللي مصابين بأورام، حميدة أو خبيثة في استخدم هذه البيانات الحاليه الموجوده عندك. للتنبؤ بأنواع الأمراض للمرضى الجدد اللي حيجوا يزوروا مستشفى معين. تحليل الإلزامي هو لما نكون قادرين نبني منتجات تعتمد على التحليل الوصفي أو التنبؤي في علم البيانات مثل ا إذا فكرنا بتطبيق مثل. يستخدم تحليل وصفي للبيانات عن الحركة الجسدية اللي قمنا فيها وبيعرضها لإلنا، وهو يقع تحت شق التحليل الزامي لعلم البيانات. كعالم بيانات أو داتا سينتس، وهو فن استخدام النظريات العلمية لاستخراج الأفكار والمعرفة من البيانات ليش فن لأنه يتضمن المزج بين العديد من المجالات اللي هي مختلفة عن بعضها البعض. وآخر مجال هو الخبرة الموضوعية أو خبرة المجال. منعني إنه بدنا نكون على ما بيانات، يعني مثلا نقدر نستخدم مثل أو آر أو ثاني لازم يكون عنا. ذكاء الاصطناعي هي عبارة عن أي تقنية تمكن الكمبيوتر من محاكاة الذكاء البشري من خلال استخدام المنطق، تعلم الآلة هي مجموعة فرعية من الذكاء الإصطناعي تتضمن تقنيات إحصائية. تمكن الآلات من تحسين أداء المهام من خلال التجربة. ويندرج تحت التعلم الآلي التعلم العميق أو الديبلورنج. التعلم العميق هو فئة فرعية من التعلم الآلي تسمح للآلات بتدريب نفسها على أداء مهام مثل التعرف على الصور أو الكلام باستخدام من خلال استخدام تقنيات الشبكات العصبية أو اللي حنشرحها بوحدة قادمة. علم البيانات مرتبط بهيدا المجالات المختلفة، ومن خلال أول تطبيق راح نبدء فيه هو أكثر تطبيق ممكن تكونوا بتستخدموا بحياتكم اليومية وهو تطبيق بحث الواب أو الويب سورج على محركات مثل جوجل مثل ا. لما أنا أعمل بحث معين عن مطعم عم بيطلع لي نتائج فقط خص ة بالأكل وبالمطاعم ممكن تكون لينكس لمطاعم ممكن تكون مقالات عن مطاعم موجودة على الإنترنت ما عم يطلعلي مقالات عن ثياب مثل ا هذه القدرة على ربط البحث أو الكلمات اللي حطيتها بالبحث، بمضمون معين أو نتائج معينة موجودة على الإنترنت، هيدي ممكنة من خلال علم البيانات. شيء ثاني اللي أكيد لاحظتوه هو إنه النتائج اللي بتطلع، لكن مرتبة والنتائج اللي ممكن تطلع لي أنا كشخص مختلفة جدا عن أه النتائج اللي ممكن ت تطلع لشخص موجود تحدي أو جنبي إنه عم يعمل نفس البحث، لأنو جوجل بيعرف وين أنا موجوده وبيعرف أنو مطاعم بالصين مثلا مش ممكن تكون مفيدة لإلي أو لبحسة بهالوقت المحدد. تاني الشي وتاني السبب هو إنه جوجل عنده الكثير من النتائج والبحث الموجودة قبل اللي عملها الكثير من الأشخاص، فبالنهاية المعلومات أو النتائج اللي عم تطلعلي أنا. هي فعلا خصها كثير بالبروفايل تبعي، البحث اللي قام فيها العديد من الأشخاص الآخرين، خلينا نشرح مبدء اللي بتستخدموا جوجل بترتيب الصفحات، بمحرك البحث تبعها، نحنا أكيد ما بنعرف كل العوامل والمؤثرة بترتيب بيت جانك، ونوعيتها بيرتبطوا ب آ النتائج اللي بتعطينا إياهن. خلينا نقول إنه أنا رحت على محول جوجل، هذا محول جوجل وحطيت مطاعم، شو الأشيا اللي أثرت بهالنتائج؟ أكيد في معلومات متل وين أنا موجودة، حالي ا دخلت بالنتائج وأكيد خصهم بالكلمة المفتاحية اللي أنا حطيتها اللي هي مطاعم أو أحد الكلمات الرديفة مثل مثل ا قهوة أو كافيه أو شي، في كمان عوامل أخرى دخلت بإعادة الترتيب. كيف بنعرف والصفحة سي صفحة واحدة فقط. هذا Used to fear .أهمية الصفحة؟ خلينا نقول الصفحة. بينما الصفحة بين في صفحتين

Summarized by © lakhasly.com

الشي بيعني إنه بالجراخ تبع الويب. هيدا أول عامل بس أحيانا في عامل تاني بيدخل اللي هو نوعية الصفحة، يعني ممكن يكون في حالة أنو يكونوا بس صفحتين عم يربطوا. وتلات صفحات عم يربطوا لبيب. وصفحة وحدة عم تربط لها سي والسل. إنه طلعلنا بهي النتيجة التانية وليس النتيجة الأولى. ليش هيدا الشي ممكن؟ القيادة بتستخدم تحليل الصور لتقدر تروح من نقطة من خلال تحليل الصور السيارة. ممكن تعرف إذا شي معين أمامها هو عبارة عن حائط أو عبارة عن حاجز أو عن إنسان أو عن طريق مفتوحة، إضافة لهيدا الشي سيارات الذاتية القيادة بتستخدم بيانات برامج تحديد المواقع لتقدر تتجنب الزحمة وتختار بوقت معين. الطريق اللي هي الأقل زحمة للوصول من النقطة ألف إلى النقطة ب، كيف السيارات الذاتية القيادة بتشتغل في رابط تحت، فشوفوا إذا مش شايفين هالشيء قبل التطبيق الآخر اللي بحب أحكي عنه واللي هو كثير مهم هو الطب الوقائي أو الطب الوقائي هو عبارة عن تقديم الأدوية الدقيقة الوقائية عن طريق الجمع بين العديد من البيانات منها البيانات الجنمية مثل ا والبيانات المتعلقة بنموات الحياة والبيئة مثلا. أنا عايشة ببلد معين في نسبة تلوث عليه بالمياه أو بالهوا، أنا حكون عرضة لأمراض أكثر من حدا غيري عايش ببلد أنطف بكثير، إذا أنا جايي من عائلة فيها أنواع معينة من الأمراض أو تبعها عرضة أكثر لأنواع معينة من الأمراض، ممكن أنا أكون عرضة لهالأمراض، فهاي المعلومات لما تكون متوفرة للطبيب للطبيب ممكن يساعدني على اتخاذ تدابير وقائية لتجنب هذه الأمراض بعدين. تطبيق الآخر كمان اللي بحب أحكي عنه هو بمجال الزراعي أو ما يسمى. أو. حالي ا بعض الشركات بتستخدم بيانات الأقمار الإصطناعية للتنبؤ بخصوبة التربة، ولما نكون نحنا قادرين نعمل هذا الشي ممكن ندفع المزارعين إنه يروحوا على أماكن معينة ويزرعوا فيها ويتجنبوا الأماكن الأخرى اللي ما حتنبت فيها زراعاتهم بنفس الشكل. فهيدا الشي بيأدي إلى توفير العديد من الاستثمارات والأموال، وبيأدي إلى نتائج أفضل، تطبيق آخر بالمجال الزراعي، بحب أحكي عنه هو استخدام الحساسات أو السانسور، لنقدر نعرف إذا نبتة معينة بحاجة لمي أو لأ، في المجال الزراعي. تطبيق آخر مهم جدا لعلم البيانات هو تحسين الأمن الإلكتروني أو هيرنسنكس هايبر ساكوريتي. هذا الشي بأدي لتحسين وتطوير الكشف عن الحسابات المزيفة والإحتيال عبر الإنترنت. مثال على الشي اللي هو الشي اللي بتقوم فيه فيس بوك مثلا، كيف فيسبوك؟ بتقدر تكتشف الحسابات المزيفة؟ إذا حساب ما عمل لوجان لفترة شهر أو حساب عم يعمل لوجان من؟ أماكن معينة كتير كتير مختلفة عن الأماكن اللي بيتواجد فيها المستخدم الأساسي، الشخص، أي نوع من مع الأشخاص، أو التواصل مع الأشخاص أو اللي كان يتواصل معن الشخص الأساسي من خلال كل هالعوامل من خلال كل هالبيانات ومقارنتها بالبيانات الأساسية على المستخدم الأساسي. وتوقف هيدي الاكاونت؟ تطبيق الآخر لعلم البيانات هو استخدام البيانات لتوفير الإنترنت للجميع ممكن ما تكونوا تعرفوا إنه في أماكن كثيرة بأفريقيا وبالشرق الأقصى لآسيا ما فيها نسبة أشخاص على الإنترنت عالية، يعني أقل من نص الأشخاص الموجودين بهالأماكن هنه عندن أكاونت على الإنترنت أو يقدروا يقدروا يفوتوا على الإنترنت. طيب إذا نحنا بدنا نبني بنية تحتية للإنترنت مثل مثلا أو شي الوصفي، التحليل الوصفي هدفه إجراء تحليل بصري للبيانات المسجلة من أجل إستخراج الأفكار والدلالات وتحديد الأنماط والإرتباطات، ماذا يعني هذا الشيء؟ يعني مثلا إذا أنا عم حلل أعمار الأشخاص اللي عم يستعملوا انستغرام وبدي أقارن هذه الأعمار بأعمار الأشخاص اللي عم يستخدموا فيسبوك، أعمار معينه مثلا من عمر؟ 15 لعمر 25 من عمر 25 لعمر 40 بعد عمر ال40 هودي ثلاث فئات عمريه أنا حددتها فيي. شوف عدد الأشخاص إلي بينتموا لكل هالفئات بإنستغرام وعدد الأشخاص إلي بينتموا بفيس بوك. و واستكشافها كثير. كثير، إذا بس تطلعنا بالبيانات من دون ما نعرضها، وهذا الهدف من تحليل الوصفي بشكل عام من أكثر الأدوات المستخدمة بمجال استكشاف البيانات وعرضها تبلغ. وما بلوتليب تبلوه هو أداة لازم ندفع لها نستخدمها مندفع لايسنس حوالي 1000 دولار بالسنة لنقدر نستخدمها بتتطلب أي نوع من أنواع البرمجة تجيبلو ما بلوتليب هنه لابريز أو مكتبات متوفرين من خلال آر وبايتون. وبشكل ومش بحاجة إنا ندفع بس لازم نعمل شوية كودينج لنقدر نستكشف البيانات ونعرضهن من خلال استخدام هذا. من بعد ما نستكشف ونعرض البيانات، ممكن أحيان ا تكون بحاجة إنه نعمل تحليلات تنبؤية أو. وأسهل طريقة لنلاقي هالشيء، أو يعني أفضل طريقة لنعمل هالشي هو إنه نستخدم تقنيات التعلم الآلي أو ال. فحنشرح كثير عن تعلم الآلي، فما حخوض بالتفاصيل هلأ. ببعض الحالات ممكن تكون نحنا مهتمين إنه نبني منتج معين، الوصفة أو التنبؤ، وهذا الشي بنسميه ديتا برودا. نموذج تعلم آلة جديد اسمه أو الشبكات العصبية. هي تقنية للتعلم الآلي أو ال مستوحاة من بنية ونشاط دماغ البشري. وهي تستخدم بشكل أساسي لأغراض التصنيف أو الكلاسيفكشن. أو عقد في عنا، وال اللي هي الموجوده بالنص وال. يعني مثلا إذا بدنا نطبق هذا النموذج التعلم الآلي اللي شايفينه هون على. تبع تصنيف صور الحيوانات من صورة هروة صورة كلب بها الحالة تبعنا حتكون مألفة من عقد. لهذا اللاير، حتعملها نوع من يعني عنا بس احتمالات بهالحالة تورك معقد أكثر،

Summarized by © lakhasly.com

وهون بيجي تبع أو التعلم العميق ليش ديب ليه العمق من وين جاي، هيدا الكلمة جاية من فكرة إنه عنا عدة مش هدول الوحيدة، من أربع ميزات ممكن تكون مثل ا أربعة أو أربع معلومات مختلفة، نحنا عم Fفالعمق هو العمق بالهذن مثل ما شايفين هون عنا. نستخدمها لنعمل عملية تصنيف عنا هون، مثل ا تقريب ا خمسة هدن لير والأوتلير هون عنا ثلاث فئات مثل ا هيدا الشي ممكن كيف بنعرف إذا لازم ،layerايطبق لتصنيف لحيوانات بثلاث فئات باندا وكلب وهر، ففينا نعرف كم فئة عنا بعملية التصنيف، نستخدم ديب نورناتورك ولا؟ عادي الوحيدة؟ وكيف ممكن نعرف كم هذا الليبار؟ نحنا لازم نحط بالنموذج تبعنا أو كم نود لازم نحط المختلفين، كلن هودي نحنا لازم نعطيهم للموديل تبعنا، لازم نقول نحنا هيدا النمبر تبع اللي بدنا إياه، وهيدا تبع العقد اللي بدنا ياه بكل اللياقة، فكيف بنحدد هيدا الشي بصراحة؟ بس فينا نجرب يعني نعمل. فينا نبلش بهذا الوحيدة بسيط ونشوف قدي دقة النموذج اللي نحنا بنيناه، ممكن نجرب. ممكن نزيد هذا الير؟ ممكن نزيد عدد الموجودين بهذا ال معينة ونجرب مجددا إذا لقينا إنو هاي الزياده أو للعقد عم تؤدي لدقة أفضل، معناتها أفضل إنو نروح وأفضل نروح لمرور، إذا مثلا لاحظنا إنو إذا انتقلنا من زيادة بالدقة عم تكون قليلة، يعني ما لازم نعمل هيدا الانتقال، لازم نضلنا مع تو هيدا الإيرز، لأنو الشي اللي عم نعملوا هون عم نزيد التعقيد تبع النموذج تبعنا، وعم نستعمل موارد حاسوبية أكثر بكثير من قبل من غير ما نزيد الدقة كثير،

Summarized by © lakhasly.com