علم البيانات

نص الحلقات

INJAZ CAMPUS

علم البيانات
نص الحقات

STEM و التكنولوجيا | تحليل الأعمال | الجامعة | ٧ جلسات

مقدمة

أهمية علم البيانات للتوظيف

يعد فهم أهمية علم البيانات للتوظيف في منطقة الشرق الأوسط وشمال أفريقيا أمرًا بالغ الأهمية نظرًا للتحول الرقمي الكبير. تمر المنطقة بنقطة محورية في رحلتها الرقمية، حيث أصبحت معظم البلدان أكثر اتصالاً من أي وقت مضى وتشهد ارتفاعًا في حركة البيانات والوصول إلى النطاق العريض. ومع ذلك، فإن الاستفادة من إمكانات التطوير الكاملة لنمو البيانات تتطلب معالجة الثغرات في الأطر القانونية والتنظيمية لإنشاء البيانات واستخدامها وإعادة استخدامها.

في منطقة الشرق الأوسط وشمال أفريقيا، هناك اعتراف متزايد بأن علم البيانات والتحليلات هي المحركات الرئيسية لاتخاذ قرارات عمل سريعة ومناسبة. لقد استفادت الشركات الأكثر نجاحًا على مستوى العالم من قدرتها على جمع البيانات وتحليلها والتصرف بناءً عليها على نطاق واسع. ومع ذلك، هناك نقص ملحوظ في المهارات في المنطقة، ليس فقط فيما يتعلق بخبراء علم البيانات، ولكن أيضًا في معرفة البيانات بشكل عام بين القوى العاملة، ولمعالجة هذه المشكلة، يتم تشجيع الشركات على التركيز على تحسين مهارات موظفيها الحاليين لإنشاء قوة عاملة أكثر معرفة بالبيانات. يتضمن هذا النهج تحديد وتمكين الخبراء في القسم، وإنشاء استراتيجيات تتضمن هؤلاء الخبراء المخفيين، وتطوير دورة مستمرة من التعلم والتحسين.

علاوة على ذلك، فإن ظهور التقنيات الرقمية يعمل على تحويل سوق العمل، مما يجعل من الضروري للعاملين في جميع القطاعات اكتساب مهاراتهم الرقمية وتعزيزها. وفي هذا السياق، تعاون البنك الدولي مع LinkedIn لتحليل أسواق العمل في منطقة الشرق الأوسط وشمال أفريقيا، وتحديد المهارات الأكثر صلة ومميزة وفقًا لصناعات ومهن مختارة في قطاع التكنولوجيا والقطاع الرقمي. ويهدف هذا التحليل إلى إعلام صناع السياسات وأصحاب العمل ومقدمي التدريب لتطوير برامج تدريب مخصصة تمكن الناس من النجاح في العصر الرقمي.

بشكل عام، تقف منطقة الشرق الأوسط وشمال أفريقيا على مفترق طرق حيث يمكن للاستفادة الفعالة من البيانات أن تعزز التنمية والتعافي من جائحة كوفيد-19. ويتطلب هذا جهودًا متضافرة من الشركات وصناع السياسات ومقدمي التدريب لسد فجوة المهارات، وبناء ثقافة مؤسسية قائمة على فهم واستخدام البيانات، مما يضمن قدرة المنطقة على الاستفادة الكاملة من الفرص التي يوفرها الاقتصاد الرقمي.

أهمية علم البيانات في ريادة الأعمال

في مجال ريادة الأعمال، وخاصة في منطقة الشرق الأوسط وشمال أفريقيا، أصبحت أهمية علم البيانات واضحة بشكل متزايد. يعد علم البيانات بمثابة حجر الأساس لاتخاذ قرارات مستنيرة، مما يسمح لرواد الأعمال باستخلاص الاستراتيجيات من البيانات التجريبية. يعد هذا النهج التحليلي أمرًا بالغ الأهمية لتصميم المنتجات والخدمات وفقًا لاحتياجات العملاء وتفضيلاتهم. إن تحسين العمليات التجارية من خلال علم البيانات لا يؤدي فقط إلى تعزيز الكفاءة وخفض التكاليف ولكنه يوفر أيضًا للشركات الناشئة والشركات الصغيرة ميزة تنافسية في إدارة الموارد.

تسلط جهود التحول الرقمي والتنويع الاقتصادي في منطقة الشرق الأوسط وشمال أفريقيا الضوء على الدور المحوري لعلم البيانات بالنسبة لرواد الأعمال. تلعب مجموعة المهارات هذه دورًا أساسيًا في تعزيز الابتكار عبر مختلف القطاعات، بما في ذلك التكنولوجيا المالية والتجارة الإلكترونية والتكنولوجيا الصحية وغيرها، ومعالجة التحديات الإقليمية الفريدة مثل المناظر الطبيعية الثقافية والبيئات التنظيمية المتنوعة. بالإضافة إلى ذلك، ومع ارتفاع معدل البطالة بين الشباب في منطقة الشرق الأوسط وشمال أفريقيا، فإن ريادة الأعمال المدعومة بمهارات علم البيانات تمثل طريقًا واعدًا لخلق فرص العمل. يمكن لرواد الأعمال الشباب المجهزين بمعرفة علوم البيانات إطلاق شركات ناشئة مرنة وتنافسية.

يوفر الارتفاع الكبير في الاستهلاك الرقمي والخدمات عبر الإنترنت في المنطقة أرضًا خصبة لرواد الأعمال المتمرسين في مجال البيانات لتطوير المنتجات والخدمات للسكان المتعلمين رقميًا. علاوة على ذلك، تساعد مهارات علوم البيانات رواد الأعمال في منطقة الشرق الأوسط وشمال أفريقيا على توسيع نطاق أعمالهم على مستوى العالم، مما يوفر رؤى حول الأسواق الدولية واتجاهات المستهلكين. تتيح التحليلات التنبؤية، وهي جانب رئيسي من علوم البيانات، التنبؤ باتجاهات السوق

المستقبلية، وبالتالي تزويد الشركات بأدوات التخطيط الاستراتيجي لمواجهة التحديات والفرص المستقبلية.

في الختام، بالنسبة لرواد الأعمال في منطقة الشرق الأوسط وشمال أفريقيا، فإن علم البيانات ليس مجرد أداة عمل ولكنه عنصر حاسم في التخطيط الاستراتيجي والابتكار. فهو يمكّنهم من التنقل بفعالية في ديناميكيات السوق الفريدة في المنطقة ويضع أعمالهم في مكانة لتحقيق النجاح محليًا وعلى المسرح العالمي.

أهمية معرفة علم البيانات عندما يتعلق الأمر بالتوظيف وريادة الأعمال
تشهد منطقة الشرق الأوسط وشمال أفريقيا حاليًا تحولًا رقميًا كبيرًا، ويلعب علم البيانات دورًا محوريًا في كل من التوظيف وريادة الأعمال ضمن هذا المشهد المتطور.

على صعيد القدرة على التوظيف، يُعترف بعلم البيانات كمجال حيوي لاتخاذ قرارات مستنيرة. تشهد المنطقة اليوم اتصالًا أكبر من أي وقت مضى، مع تزايد حركة البيانات ووصول النطاق العريض للإنترنت. ومع ذلك، للاستفادة الكاملة من نمو البيانات هذا، من الضروري معالجة الثغرات في الأطر القانونية والتنظيمية. في حين استفادت الشركات العالمية الناجحة من قوة تحليل البيانات، تواجه منطقة الشرق الأوسط وشمال إفريقيا نقصًا في المهارات، بما في ذلك نقص علماء البيانات ومستوى عام من معرفة البيانات بين القوى العاملة. لمعالجة هذه الفجوة، يتم تشجيع الشركات على رفع مهارات موظفيها، وخلق قوة عاملة ملمة بالبيانات من خلال تحديد وتمكين خبراء الأقسام الداخلية. ويشمل هذا النهج تطوير استراتيجيات مستمرة للتعلم والتحسين. بالإضافة إلى ذلك، تعمل التقنيات الرقمية على إعادة تشكيل سوق العمل، يستلزم تنمية المهارات الرقمية عبر القطاعات. وتهدف عمليات التعاون مثل تلك القائمة بين البنك الدولي ولينكد إن إلى تحديد المهارات ذات الصلة لمساعدة الناس على النجاح في العصر الرقمي. بشكل عام، يمكن أن يؤدي الاستخدام الفعال للبيانات إلى تعزيز التنمية ومرحلة التعافي ما بعد الوباء في منطقة الشرق الأوسط وشمال أفريقيا الأمر الذي يتطلب جهودًا متضافرة من الشركات وصانعي السياسات ومقدمي التدريب.

في سياق ريادة الأعمال، أصبح علم البيانات أكثر أهمية لاتخاذ قرارات مستنيرة. فهو يسمح للرواد باستخلاص استراتيجيات من البيانات التجريبية، وتكييف المنتجات والخدمات لتلبية احتياجات العملاء، وتعزيز الكفاءة، وتقليل التكاليف، والحصول على ميزة تنافسية. يؤكد التحول الرقمي الذي تشهده منطقة الشرق الأوسط وشمال إفريقيا، بالإضافة إلى جهود تنويع الاقتصاد، على الدور المحوري لعلم البيانات في تعزيز الابتكار، ومعالجة التحديات الإقليمية، وخلق فرص العمل، وخاصةً للشباب. عَلاوةً على ذلك، يمكن لرواد الأعمال من ذوي المهارات في مجال البيانات الاستفادة من الزيادة في الاستهلاك الرقمي والخدمات عبر الإنترنت في المنطقة لتطوير منتجات وخدمات تتناسب مع جيل متمكن رقمياً. كما تمكن هذه المهارات رواد الأعمال من توسيع نطاق أعمالهم على الصعيد العالمي، وتوفير رؤى حول الأسواق الدولية واتجاهات المستهلكين من خلال التحليلات التنبؤية.

وباختصار، يمثل علم البيانات أداةً لا غنى عنها في مجاليّ التوظيف وريادة الأعمال في منطقة الشرق الأوسط وشمال إفريقيا. فهو يسهل اتخاذ القرارات المستنيرة والابتكار وخلق فرص العمل والتوسع العالمي، مما يمهد الطريق لنجاح الأفراد والشركات على حد سواء في ظل اقتصاد رقمي متطور باستمرار. يُعد سد الفجوات في المهارات ومعالجة التحديات التنظيمية خطواتٍ حاسمة لإطلاق العَنان الكامل لإمكانيات علم البيانات في المنطقة.

١ مقدمة في علم البيانات

الأهداف التعليمية:

- تعريف البيانات كمعلومات أولية.
- وصف علم البيانات كمجال متعدد التخصصات لاستخراج المعرفة.
- التأكيد على دور فهم الأعمال في علم البيانات.
- استكشاف تأثير تصور البيانات والتنبؤات النموذجية على عملية صنع القرار.
- تعريف النمذجة التنبؤية على أنها استخدام الخوارزميات للتنبؤات.

النص:

ما هي **البيانات**؟

البيانات هي المعلومات الأولية وغير المعالجة التي تعمل كأساس للمعرفة وصنع القرار. ويمكن أن تتخذ أشكالًا مختلفة، بما في ذلك الأرقام والنصوص والصور، وأكثر من ذلك.

في جوهرها، البيانات هي أساس بناء المعلومات، وتكمن أهميتها في القدرة على استخلاص رؤى ذات معنى عند تحليلها وتفسيرها بشكل صحيح.

ما هو **علم البيانات**؟

علم البيانات هو مجال متعدد التخصصات يستخدم الأساليب العلمية والعمليات والخوارزميات والأنظمة لاستخراج المعرفة والرؤى من البيانات المنظمة وغير المنظمة.

فهو يجمع بين عناصر الإحصاء والرياضيات وعلوم الكمبيوتر والخبرة الخاصة بالمجال لتحليل مجموعات البيانات المعقدة وكشف الأنماط والاتجاهات والعلاقات المتبادلة.

يُمكِّن علم البيانات المؤسسات من اتخاذ قرارات مستنيرة، والتنبؤ بالاتجاهات المستقبلية، واستخلاص رؤى قابلة للتنفيذ من الكميات الهائلة من البيانات التي تم إنشاؤها في العصر الرقمي الحالي.

ومثل أي كَيان آخر، يتبع مجال علوم البيانات دورة حياة خاصة به.

دورة حياة علم البيانات هي عملية منظمة توجه الممارسين خلال مراحل مختلفة للحصول على رؤى فعالة للبيانات. تشمل المراحل الرئيسية ما يلي:

- **فهم الأعمال:** تحديد أهداف المشروع مع أصحاب المصلحة لمواءمة الحلول مع الأهداف.
- **التنقيب عن البيانات (تعدين البيانات):** جمع البيانات ذات الصلة، مع الأخذ بعين الاعتبار الجودة والملاءمة وسهولة الوصول إليها.
- **تنظيف البيانات:** التعامل مع التناقضات والأخطاء والقيم المفقودة للتحليل الدقيق.
- **استكشاف البيانات:** تحليل البيانات المُنَظَّفة للاتجاهات والرؤى باستخدام الإحصائيات والمرئيات.
- **هندسة الخصائص:** تحسين أداء النموذج عن طريق تحديد الميزات أو تحويلها.
- **النمذجة التنبؤية:** تطوير نماذج التعلم الآلي وضبطها للحصول على تنبؤات دقيقة.

- **التصور البصري للبيانات:** استخدم العناصر المرئية لتوصيل النتائج المعقدة إلى أصحاب المصلحة، مما يساعد في اتخاذ القرار.

ماذا عن علماء البيانات؟ هل تساءلت يومًا عن أدوارهم؟
يُمكن وصف دور عالم البيانات بأنه "متعدد المهارات"، وذلك بسبب مجموعة المهارات المتنوعة المطلوبة للتميُّز في هذا المجال.

يرتدي عالم البيانات قبعات متعددة، حيث يجمع بين الخبرة من مختلف المجالات لاستخراج رؤى ذات معنى من البيانات.

فيما يلي استكشاف للمهارات المختلفة التي تجعل من عالم البيانات عالمًا موسوعيًا حقيقيًا:

- **عالِم رِياضيّات:**
- الدور: يستفيد علماء البيانات من المفاهيم والتقنيات الرياضية لبناء نماذج وخوارزميات قوية. من الجبر الخطي لعمليات المصفوفات إلى حساب التفاضل والتكامل للتحسين، يعد الأساس الرياضي المتين أمرًا بالغ الأهمية لفهم نماذج التعلم الآلي المعقدة وتطويرها.

- **عالم إحصاء:**
- الدور: تعتبر المعرفة الإحصائية أمرًا أساسيًا لعلماء البيانات فيما يتعلق باختبار الفرضيات ونظرية الاحتمالات والإحصاءات الاستدلالية. يستخدم علماء البيانات الأساليب الإحصائية لاستخلاص استنتاجات موثوقة من البيانات، وإجراء التنبؤات، وتقييم أهمية النتائج التي توصلوا إليها.

- **عالم:**
- الدور: يتبع علماء البيانات منهجًا علميًا في عملهم. إنهم يشكلون الفرضيات، ويصممون التجارب (أو التحليلات)، ويقيمون النتائج بشكل منهجي. تضمن هذه العقلية العلمية موثوقية النتائج وإمكانية تكرارها، مما يساهم في الدقة الشاملة لعملية علم البيانات.

- **مبرمج:**
- الدور: تعد الكفاءة في البرمجة سمة أساسية لعالم البيانات. فهم يستخدمون لغات البرمجة مثل Python أو R لمعالجة البيانات وبناء النماذج وإنشاء حلول قابلة للتطوير. تعد مهارات البرمجة ضرورية لتنفيذ الخوارزميات وأتمتة العمليات ودمج الحلول التحليلية في بيئات الإنتاج.

في مجالات العلوم والتكنولوجيا والهندسة والرياضيات (STEM)، تعتبر خبرة عالم البيانات الماهر لا تقدر بثمن لتسخير الرؤى المستندة إلى البيانات لتعزيز البحث وتحسين العمليات واتخاذ قرارات مستنيرة، والمساهمة في نهاية المطاف في التقدم العلمي والابتكار.

ما هي فوائد استخدام الأدوات؟

- **سير عمل أفضل:**

تعمل Jupyter Notebooks على تعزيز سير العمل المتكامل، والجمع بين التعليمات البرمجية والمرئيات والنص لعملية تحليل متماسكة ومتكررة.

- **توفير الوقت**:

تعمل أدوات مثل Apache Spark على تمكين المعالجة المتوازية لمجموعات البيانات الكبيرة، مما يقلل من أوقات الحساب. تتميز الأتمتة والتحسين بميزة إعفاء علماء البيانات من المهام اليدوية، مما يسمح لهم بالتركيز على العمل عالي المستوى.

- **التناسق**:

يعد الاتساق في التحليل أمرًا حيويًا للموثوقية، ويعمل Power BI على تمكين علماء البيانات من تصور البيانات، والاتصال بمصادر البيانات المتنوعة للتقارير ولوحات المعلومات التفاعلية. يساعد تكامله مع منتجات Microsoft وواجهته سهلة الاستخدام في سرد البيانات واتخاذ القرارات بشكل تعاوني.

فيما يلي بعض أدوات علم البيانات الرئيسية الأخرى:

Excel:

يعتبر برنامج Excel، وهو برنامج جداول بيانات يستخدم على نطاق واسع، ضروريًا لمعالجة البيانات الأساسية وتحليلها. تعمل واجهته البديهية على تسهيل استكشاف مجموعة البيانات، وإجراء العمليات الحسابية، والتصورات البسيطة، مما يجعلها أحد الأصول القيمة في مجموعة أدوات عالم البيانات.

Tableau:

Tableau هي أداة بديهية لتصور البيانات تمكن علماء البيانات من إنشاء لوحات معلومات تفاعلية ومفيدة مع واجهة السحب والإفلات، مما يجعل البيانات المعقدة سهلة الفهم لجميع المستويات التقنية.

عبر مختلف المجالات، بما في ذلك الأعمال التجارية والرعاية الصحية والتمويل، يسهل علم البيانات اتخاذ قرارات أكثر استنارة واستراتيجية من خلال عملية منهجية من خمس خطوات:

- **حدد أهدافك:**

حدد بوضوح أهداف عملية اتخاذ القرار المستندة إلى البيانات، وحدد التحديات أو الفرص بما يتماشى مع الأهداف التنظيمية.

- **حدد مصادر البيانات:**

قم بتقييم وتحديد مصادر البيانات، مع الأخذ في الاعتبار مجموعات البيانات الداخلية أو المصادر الخارجية أو المزيج. قييم جودة البيانات وإمكانية الوصول إليها وأهميتها.

- **بناء/ شراء حل لمعالجة البيانات:**

تطوير أو تحديد حل لمعالجة البيانات، مما يضمن الكفاءة وقابلية التوسع وأمن البيانات والتكامل مع المصادر المتنوعة.

- **استخراج القيمة من البيانات:**

قم بإجراء التحليل باستخدام الأساليب الإحصائية والتعلم الآلي لاستخلاص رؤى قابلة للتنفيذ، بما يتماشى مع الأهداف المحددة.

- **تعزيز الثقافة القائمة على البيانات في مؤسستك:**

زراعة ثقافة تنظيمية تقدر عملية صنع القرار المبنية على البيانات. رفع مستوى الوعي وتشجيع التعاون وتعزيز المعرفة بالبيانات.

دعونا نستخدم علم البيانات!

يمكن أن يتضمن العرض العملي لعلم البيانات تطبيقًا عمليًا، مثل التنبؤ بسلوك العملاء، أو تحسين عمليات سلسلة التوريد، أو تشخيص الحالات الطبية.

تخيل السيناريو التالي:
أنت عالم بيانات في شركة تبيع 3 منتجات مختلفة، وهناك خطأ ما في الأرباح.
تم تكليفك بمعرفة ما يحدث باستخدام البيانات المقدمة من الشركة
كيف ستحل هذه المشكلة وتمنع حدوث هذه المشكلة مرة أخرى؟

أولا، جَمِّعْ البيانات!
تمتلك شركتك بيانات حاسمة لحل هذه المشكلة، وتحديداً بيانات المبيعات من العام الماضي.

سترى أن البيانات منظمة بالفعل ويبدو أنها جاهزة للاستخدام. ومع ذلك، فإنه غالبًا ما تتطلب إعدادًا إضافيًا قبل أن يتم استخدامها بشكل فعال.

لتبسيط العملية، لنفترض أن البيانات منظمة مسبقًا، فلنبدأ مباشرة!

فكر في **استكشاف البيانات.**

وهو بمثابة الخطوة الأولى **بعد جمع البيانات وقبل النمذجة المتقدمة**، حيث يوفر نظرة عامة شاملة على بُنية البيانات وتوزيعها وعلاقاتها المحتملة.

تتمثل الأهداف الأساسية لاستكشاف البيانات في تحديد الأنماط والقيم المتطرفة والاتجاهات والحصول على رؤى قيمة تفيد القرارات التحليلية اللاحقة.

تأتي "**هندسة الخصائص**" كخطوة رئيسية في مجال التعلم الآلي بعد عملية استكشاف البيانات وتصورها. تركز هذه العملية على إنشاء أو تحويل المتغيرات الموجودة في مجموعة البيانات بهدف تحسين أداء النماذج التنبؤية. إنها تجمع المعرفة المتخصصة، والإبداع، وفهم مجموعات البيانات لاستخراج معلومات قيمة لتحسين التنبؤات.

والآن، حان دور **التصور البصري للبيانات**, الذي يتجاوز الأرقام والجداول الخامة ليقدم طريقة سهلة وبسيطة لنقل المعلومات المعقدة. من خلال العروض المرئية الجذابة، يستطيع علماء البيانات والمحللين إيصال الأفكار والرؤى إلى مختلف الجماهير، سواء الفنية أو غير الفنية.

أخيرًا، **النمذجة التنبؤية**
هي فرع من علم البيانات يستخدم خوارزميات إحصائية وتقنيات التعلم الآلي للتنبؤ بنتائج مستقبلية بناءً على بيانات تاريخية.
تمكن هذه المنهجية القوية المؤسسات من اكتساب رؤى حول الاتجاهات والأنماط والعلاقات داخل مجموعات البيانات الخاصة بها، مما يسمح لها باتخاذ قرارات مستنيرة وتوقع الأحداث المستقبلية.

٢ البيانات هي النفط الجديد

الأهداف التعليمية:

- التعرف على تواجد البيانات في كل مكان في الحياة اليومية، وتشابكها مع التكنولوجيا والتفاعلات.
- فهم أنواع البيانات المتنوعة وتطبيقاتها، بدءًا من قواعد البيانات المنظمة وحتى الإشارات الصوتية والمحتوى المرئي.
- اكتشف كيفية مساهمة مصادر البيانات المختلفة، بما في ذلك قواعد البيانات والسجلات المالية والبيانات التشغيلية ومعلومات البحث والتطوير وتعليقات الموظفين، في فهم العمليات التنظيمية.
- تقدير دور البيانات مفتوحة المصدر في تعزيز التعاون والابتكار واكتشاف المعرفة في المجالات المختلفة.

النص:

لا تقتصر البيانات على الخوادم وجداول البيانات؛ إنها منسوجة في نسيج وجودنا اليومي.
منذ اللحظة التي نستيقظ فيها، نتفاعل مع التكنولوجيا التي تولد البيانات وتستهلكها.

سواء أكان ذلك التحقق من الطقس على هواتفنا الذكية، أو تتبع خطواتنا باستخدام تطبيقات اللياقة البدنية، أو حتى تصفح الإنترنت - فالبيانات موجودة في كل مكان.

في عالم البيانات، هناك مجموعة متنوعة من الأنواع تحمل رؤى قيمة.

1. **البيانات الجدولية،** الشكل الأساسي للبيانات، منظمة في الصفوف والأعمدة، وهو نموذج بيانات أساسي حيث يحتوي كل صف على خلايا تمثل الخصائص الموضحة بواسطة الأعمدة. ويدعم قواعد البيانات وجداول البيانات، مما يتيح التحليل المنهجي.
2. **بيانات السلاسل الزمنية،** يلتقط المعلومات التي تتطور مع مرور الوقت، مما يسمح لنا بتمييز الأنماط والاتجاهات، وهو أمر بالغ الأهمية للتنبؤ وفهم الديناميكيات الزمنية. تستخدم بشكل رئيسي لمعالجة الإشارات وتخزين بيانات الإشارة.
3. **البيانات الصوتية**، يتم استخدام الإشارات الصوتية في تطبيقات مختلفة مثل التعرف على الكلام وتحليل الموسيقى.
4. **بيانات الفيديو،** ديناميكية تجميع الصور بمرور الوقت، توفر مصدرًا غنيًا للفهم البصري، مما يتيح تطبيقات مثل المراقبة وتحليلات الفيديو.
5. **بيانات الصور،** تمثل هذه البيانات المحتوى البصري، وتلعب دورًا رئيسيًا في مجالات مثل رؤية الكمبيوتر، حيث تساعد في التعرف على الأشياء ومعالجة الصور.

الصور بالأبيض والأسود

- نبدأ بالنوع الأساسي من الصور وهو الأبيض والأسود
- قيمة البكسل في مثل هذه الصور هي إما 0 (أسود) أو 1(أبيض).

صور التدرج الرمادي

- اللونان المتناقضان في هذا المقياس هما الأسود، ويمثله 0، والأبيض، ويرمز له بـ 255.
- تمثل القيم التي تقع بين هذه الأرقام درجات مختلفة من اللون الرمادي
- ضع شدة البكسل في مصفوفة، وهذا هو تمثيل صورتك!

صور RGB

- يتم تحديد لون كل بكسل من خلال مزيج من شدة الأحمر والأخضر والأزرق، والتي يتم تخزينها في مستويات الألوان الخاصة بها في موضع البكسل.

6. **بيانات نصية،** منتشر على الانترنت، أساسي لمعالجة اللغة الطبيعية، واستخلاص الأفكار من المصادر النصية. شكله الأولي هو ثنائي (الآحاد والأصفار)، مجمعة في بايت باستخدام ASCII (الرمز القياسي الأمريكي لتبادل المعلومات).

يساهم كل نوع من أنواع البيانات بشكل فريد في مشهد البيانات، كما تعمل خصائصها المتنوعة على تمكين الأساليب التحليلية المتنوعة عبر المجالات.

تعد الاستفادة من بيانات الشركة الداخلية أمرًا أساسيًا لاتخاذ قرارات مستنيرة.

تضم قواعد بيانات الشركات معلومات منظمّة حول العمليات، وتقدم السجلات المالية رؤى حول الربحية، بينما توفر البيانات التشغيلية نظرة شاملة عن العمليات، وتوجه بيانات البحث والتطوير الابتكار، وتضيف ملاحظات الموظفين بُعدًا نوعيًا، حيث تكشف عن رؤى حول المناخ التنظيمي.

دعونا نتحدث عن **"البيانات مفتوحة المصدر".** تُشير عادةً إلى مجموعات البيانات المتاحة مجانًا للاستخدام العام والتعديل والتوزيع.

غالبًا ما يتم إتاحة الوصول إلى مجموعات البيانات هذه بأقل قدر من القيود، مما يعزز التعاون والابتكار.
فيما يلي بعض المنصات والمستودعات الشائعة للوصول إلى البيانات مفتوحة المصدر:

- **مجموعات بيانات Kaggle:**

Kaggle عبارة عن منصة تستضيف مجموعات بيانات لمسابقات التعلم الآلي وعلوم البيانات. ويقدم مجموعة واسعة من مجموعات البيانات في مختلف المجالات التي ساهم بها المجتمع.

- **GitHub:**

يستضيف GitHub العديد من المستودعات التي تحتوي على مجموعات بيانات مفتوحة المصدر. غالبًا ما يشارك المستخدمون مجموعات البيانات المتعلقة بمجالات أو مشاريع بحثية محددة.

- **تحليلات Vidhya:**

إنها منصة لعلوم البيانات عبر الإنترنت تقدم مقالات وبرامج تعليمية ودراسات حالة ودورات وشهادات ومسابقات. وتستضيف مدونة ومنتدى لتعزيز التعلم العملي والتعاون في مجتمع علوم البيانات.

- **KDD :**

يستخرج KDD (اكتشاف المعرفة في قواعد البيانات) الرؤى من مجموعات البيانات الكبيرة عبر مراحل مثل الاختيار والمعالجة المسبقة والتحويل واستخراج البيانات والتقييم. تعمل هذه العملية التكرارية على تحويل البيانات الأولية إلى معرفة قابلة للتنفيذ لاتخاذ قرارات مستنيرة.

في هذه المرحلة، قد تكون مهتمًا بكيفية جمع هذه البيانات.

قد يكون تحديد البيانات ذات الصلة بالتحديات التي تواجهك أمرًا صعبًا، وغالبًا ما يتطلب جمع بيانات مستقلة.

فيما يلي بعض الأفكار للمساعدة:

- **استخراج البيانات من الانترنت Web scraping:** حدد موقع لاستخراج المحتوى، واستخرج محتوى HTML من صفحة الويب، واحفظ البيانات بالتنسيق المفضل لديك.

- **الاستطلاعات Surveys:** الاستطلاعات عبر الإنترنت، والاستطلاعات الهاتفية، والمقابلات الشخصية.

- **التوظيف الجماعي Crowdsourcing:** يستفيد من الأفراد الذين يشتركون في مصالح مشتركة لجمع البيانات، حيث يجمع بين أصحاب العمل الحر والمتطوعين. هذه الطريقة فعّالة من حيث التكلفة تيسر العمليات، مما يوفر الوقت والنفقات للشركات.

- **زيادة البيانات Data augmentation:** زيادة تنوع البيانات المتاحة للتدريب بشكل كبير دون جمع بيانات جديدة بشكل فعلي. تنطبق زيادة البيانات أيضًا على أنواع أخرى من البيانات

- **يتم إنشاء البيانات الاصطناعية Synthetic data:** المرئية في المقام الأول، برمجيًا باستخدام محركات العرض التي تولد الصور والتعليقات التوضيحية. تعتبر هذه البيانات المرنة والقابلة للتطوير ذات قيمة لتدريب نماذج التعلم الآلي ومحاكاة السيناريوهات المختلفة.

- **دولاب موازنة البيانات Data flywheel:** مفهوم الدائرة البيانية مثير للاهتمام، حيث يشدد على عملية دورية:

اعرض نموذجك أمام المستخدمين -> اجمع المزيد من البيانات -> قم بتحسين النموذج -> تقديم منتج أفضل للمستخدمين -> الحصول على المزيد من المستخدمين

تتيح هذه الحلقة التكرارية التحسين السريع، مما يعزز الدورة المستمرة التي تعمل على تسريع تطوير المنتج وتسهيل التحول المُبكِّر للمنتج.

في عصر اليوم الغني بالبيانات، يعد إعطاء الأولوية لخصوصية البيانات لكل من المؤسسات والأفراد أمرًا ضروريًا.

وهذا يعني التعامل مع البيانات الشخصية ومعالجتها وحمايتها بشكل مناسب لاحترام الحقوق الفردية عبر الإنترنت.

من الضروري أيضًا أن تظل يقظًا ضد التهديدات المتطورة مثل الوصول غير المصرح به، والهجمات الإلكترونية، والانتهاكات، ورسائل التصيد الاحتيالي، والبرامج الضارة، التي تعرض سرية البيانات وأمانها للخطر.

يعد تنفيذ أفضل الممارسات لخصوصية البيانات أمرًا ضروريًا لحماية المعلومات الحساسة والحفاظ على الثقة مع أصحاب المصلحة.

تشمل الممارسات الرئيسية ما يلي:

- **كلمات المرور القوية:**

تعزيز استخدام كلمات مرور قوية وفريدة مكونة من أحرف وأرقام ورموز، وتحديثها بانتظام لتحسين الأمان.

- **التحكم في مشاركة البيانات:**

تنفيذ ضوابط وصول صارمة لتقييد مشاركة البيانات على الموظفين المصرح لهم، بناءً على الأدوار والمسؤوليات الوظيفية المحددة.

- **المصادقة الثنائية (2FA):**

تمكين (2FA) لطبقة أمان إضافية، تتطلب من المستخدمين التحقق من هويتهم بعامل ثانٍ، مثل رمز من تطبيق أو رسالة نصية، إلى جانب كلمة المرور الخاصة بهم.

- **تحديثات البرنامج:**

قم بتحديث جميع البرامج بانتظام، بما في ذلك أنظمة التشغيل وبرامج مكافحة الفيروسات والتطبيقات، لتصحيح نقاط الضعف والحماية من التهديدات الأمنية المحتملة.

- **احتياطات محاولات التصيد الاحتيالي:**

قم بتثقيف المستخدمين حول مخاطر التصيد الاحتيالي وتشجيع الشك تجاه رسائل البريد الإلكتروني أو الرسائل غير المرغوب فيها. تنفيذ حلول تصفية البريد الإلكتروني لاكتشاف محاولات التصيد الاحتيالي وحظرها.

- **الشبكات الآمنة فقط:**

اتصل بشبكات آمنة وموثوقة، وتجنب شبكات Wi-Fi العامة للأنشطة الحساسة. استخدم الشبكات الخاصة الافتراضية (VPN) عند الوصول إلى البيانات عن بعد لتشفير الاتصالات وتعزيز الخصوصية.

من خلال دمج هذه الاحتياطات في ممارسات الأمن السيبراني الخاصة بك، يمكنك تحسين الوضع الأمني لكل من المستخدمين الفرديين والأنظمة التنظيمية بشكل كبير.

٣ تنظيف البيانات ومعالجتها

الأهداف التعليمية:

- فهم EDA باعتباره أمرًا ضروريًا في علم البيانات، ومزج الإحصائيات والتكنولوجيا للتحليل.
- النظر إلى EDA كحلقة وصل بين العلم والنظرية والتطبيق العملي.
- تقدير اكتشاف الحالات الشاذة واختبار الفرضيات والتحقيق في البيانات واكتشاف الأنماط.
- الإحصاء الوصفي لتوزيع البيانات والاتجاهات المركزية.
- التعرف على خطوات تشخيص البيانات: أنواع الأعمدة وأبعاد البيانات وأنواع البيانات والقيم الخالية والخصائص الإحصائية.
- التعرف على أهمية تنظيف البيانات لإعداد مجموعة البيانات.

النص:

تعرف على فاطمة، وهي محللة بيانات متحمسة، شرعت في رحلة في تحليل البيانات الاستكشافي (EDA) باستخدام مجموعة بيانات معقدة تهدف إلى توجيه استراتيجيات شركتها.

نظرًا لفهمها لـ EDA كعنصر حاسم في علوم البيانات المتأصلة في العلوم والتكنولوجيا والهندسة والرياضيات (STEM)، استخدمت التقنيات الإحصائية وتكنولوجيا الحوسبة والنهج الهندسي لحل المشكلات بشكل منهجي، ومزج النظرية العلمية مع التطبيق العملي لتحقيق الأهداف المتنوعة لـ EDA.

وبينما تعمقت فاطمة في فهم الأهداف متعددة الأوجه لتحليل البيانات الاستكشافي EDA.

كان كل هدف بمثابة ضوء إرشادي في استكشافها:

- **اكتشاف الحالات الشاذة:** تضمنت مهمتها الأولية تحديد القيم المتطرفة أو الأنماط غير العادية في البيانات، والتي غالبًا ما تكشف عن رؤى أو مجالات رئيسية تحتاج إلى مزيد من الفحص.

- **اختبار الفرضيات** في تحليل البيانات الاستكشافي سمح لفاطمة باختبار افتراضاتها حول البيانات، وهو خطوة حيوية لتأكيد أو نفي نظرياتها وفهم خصائص مجموعة البيانات.

- **التحقيق في البيانات:** كرست فاطمة الكثير من الوقت للتحقيق العميق في البيانات، وفحص المتغيرات وعلاقاتها المتبادلة، كما يفعل المحقق الذي يجمع الأدلة معًا لتكوين صورة أوسع.

- **اكتشاف الأنماط:** وجدت فاطمة الإثارة في الكشف عن أنماط مثل الاتجاهات أو الارتباطات أو التجمعات، والتي قدمت رؤى قيمة وغالباً ما وجهت المزيد من التحليل.

بالنسبة لفاطمة، عندما شرعت في رحلة تحليل البيانات، كانت الإحصائيات الوصفية هي الخطوة الأولى في فهم مجموعة البيانات المعقدة التي أمامها.

ومن خلال تطبيق هذه الإحصائيات، حصلت على نظرة عامة أولية لبياناتها:

- لقد حسبت **المتوسط والوسيط** للعثور على الاتجاه المركزي، مما منحها فهمًا سريعًا للقيم المتوسطة في مجموعة البيانات الخاصة بها.
- كشف **النطاق والانحراف المعياري** عن مدى انتشار نقاط بياناتها، مما يشير إلى مستوى الاتساق في البيانات.
-

- من خلال **فحص الانحراف** يمكن لفاطمة أن تستنتج ما إذا كانت بياناتها تحمل أي انحراف أو ارتفاع غير عادي.
- ومن خلال **تحليل الأرباع والنسب المئوية**، تمكنت من تحديد القيم المتطرفة وفهم توزيع البيانات عبر مستويات مختلفة.
- ساعدتها **توزيعات الترددات** في تصور البيانات، مما يسهل اكتشاف الأنماط والشذوذات.

أجرت فاطمة تشخيصًا للبيانات، لتحديد أي مشكلات في مجموعة البيانات اللتي قد تؤثر على دقة التحليل أو موثوقيته.

ما هي الخطوات الأساسية للتشخيص؟

- بدأت فاطمة بتحديد نوع البيانات لكل عمود في مجموعة البيانات الخاصة بها، مما ساعد في فهم ما إذا كانت البيانات هي عددية، أو تصنيفية، أو نصية، أو تاريخية/زمنية.

- ثم قامت بتصنيف أبعاد مجموعة البيانات عن طريق عد الصفوف والأعمدة، مما قدم فهمًا واضحًا لحجمها ونطاقها لتخطيط تحليلها.

- تحديد نوع البيانات في العمود: بعد ذلك، تحدد فاطمة أنواع البيانات الدقيقة لعمودين محددين، مثل عدد صحيح، أو عدد عشري، أو سلسلة نصية، أو القيمة المنطقية.

- تتحقق فاطمة من وجود قيم فارغة أو مفقودة في كل عمود، وهي خطوة حاسمة لتقييم اكتمال البيانات وسلامتها.

- تقييم الخصائص الإحصائية: في النهاية، قامت فاطمة بتقييم مقاييس إحصائية رئيسية مثل المتوسط والوسيط والانحراف المعياري لكل عمود لفهم توزيع البيانات والاتجاهات المركزية، مما يُعد المرحلة الأولى لتحليل أعمق.

تتقدم فاطمة إلى تنظيف البيانات بعد تشخيص البيانات، بهدف تحسين جودة مجموعة البيانات الخاصة بها. تبدأ بالتعامل مع القيم الخالية، وتقرر ما إذا كانت تريد إزالتها أو إدراجها. ثم تقوم بعد ذلك بتقييم مدى ملاءمة كل عمود، وإزالة الأعمدة غير المساهمة.

وبعد ذلك، تقوم بتحديد السجلات المكررة وإزالتها.
إذا واجهت فاطمة تحديات مع بنية أنواع البيانات في مجموعة البيانات الخاصة بها، وتحديداً التفريق بين الأنواع العددية والتصنيفية، فيجب عليها تصحيح هذه التصنيفات.

تلعب البيانات العددية دورًا رئيسيًا في العمليات الحسابية وتحسين استخدام الذاكرة في **إطارات البيانات (data frames)**، وهو أمر مهم للمكتبات المستخدمة في الحسابات الإحصائية وتعلم الآلة.

لكن نطاق قيمتها يمكن أن يؤثر على أداء النموذج، لذلك يتم استخدام **التوحيد Normalization أو التقليص Scaling** لتوحيد النطاقات، وتحقيق توازن في تأثير الميزات على التحليل وبناء النموذج.
يعمل التوحيد على ضبط القيم ضمن نطاق [0، 1]، مما يفيد الخوارزميات الحساسة للقياس مثل K-Nearest Neighbors.

X Normalized= Xinitial/max(X)

أما **التقييس (Standardization)**، من ناحية أخرى، فيتم إلى متوسط يساوي صفر وانحراف معياري يساوي واحد، وهو مثالي للطرق مثل آلات المتجهات الداعمة (Support Vector Machines) التي تفترض توزيعًا طبيعيًا للبيانات.

X standardized = X initial - mean(X)/std (X)

الاختيار بينهما يعتمد على الخوارزمية الخاصة وخصائص البيانات، حيث يكون التوحيد حساسًا للقيم الطرفية، في حين أن التقييس أقل حساسية لها.

تعكس مجموعة فاطمة من البيانات المتحولة، التي أصبحت الآن سردًا يتجاوز الأرقام والفئات، رحلتها الماهرة في تحليل البيانات الاستكشافي EDA. تسلط هذه العملية الضوء على قوة النهج المنهجي والثاقب في الكشف عن قصص البيانات المخفية.

٤ تَصَوُّرُ البيانات

الأهداف التعليمية:

- فهم دور وأهمية تصور البيانات.
- تعلم كيفية تحديد واستخدام أنواع مختلفة من المخططات والرسوم البيانية.
- تطوير مهارات تفسير وتحليل البيانات المرئية.
- اكتساب القدرة على اختيار التصورات المناسبة لمجموعات البيانات المختلفة.
- تطبيق تصور البيانات في سيناريوهات عملية.

النص:

في العصر الرقمي الحالي، يعد تصور البيانات جزءًا مهمًا من العملية التحليلية، وهو ضروري لتحويل البيانات المعالجة إلى استنتاجات ذات معنى.

أكثر من مجرد تقديم البيانات، فهو يحول المعلومات المعقدة إلى معلومات مفهومة وجذابة.

هذه المرحلة الحاسمة في تحليل البيانات تعمل على تحسين الدقة وتسريع عملية اتخاذ القرار من خلال توضيح مجموعات البيانات المعقدة. يعد استخدام أنواع المخططات المتنوعة لتصور البيانات بشكل فعال أمرًا حيويًا للكشف عن الرؤى المخفية في السياقات المستندة إلى البيانات.

- **مخطط الخطوط (Line Chart):** مثالي لتتبع الاتجاهات الزمنية والتقدم، مثل المبيعات السنوية أو تعداد الحيوانات.

- **مخطط الأعمدة (Bar Chart):** يظهر التغييرات عبر المتغيرات أو يقارن بين متغيرات متعددة، مفيد للبيانات المتعلقة بالوقت أو البيانات التفضيلية.

- **مخطط الدائرة (Pie Chart):** فعّال لعرض النسب والتوزيعات، يصلح بشكل خاص لتوضيح النسب المئوية.

- **الهيستوغرام (Histogram):** يلخص توزيع البيانات، ويكشف عن الأنماط من خلال تحليل التردد (المحور الأفقي: الحاويات؛ المحور الرأسي: الترددات).

- **مخطط الكثافة ثنائي الأبعاد (2D Density Plot):** هو هيستوغرام متقدم لتوزيع القيم عبر متغيرين.

- **مخطط النقاط المتناثرة (Scatter Plot):** يُظهر العلاقات بين متغيرين، محددًا الترابطات والاتجاهات.
>مخطط (x، y) لمتغيرين: يُستخدم للعثور على ترابطات وخطوط اتجاه.

- **مخطط الفقاعات (Bubble Plot):** نسخة من مخطط النقاط المتناثرة تستخدم فقاعات لتمثيل ثلاثة متغيرات رقمية وتحليل علاقاتها.

- **مخطط الصندوق والشارب (Box and Whisker):** يشمل خمس قيم محددة:
>بترتيب تصاعدي:

أدنى قيمة (أي الحد الأدنى).
الربع الأول (25%) - Q1.
الوسيط (50%).
الربع الثالث (75%) - Q3.
أعلى قيمة (أي الحد الأقصى).

يُستخدم للعثور على القيم الطرفية وعرض انحراف البيانات.

مخطط تقدير كثافة النواة (Kernel Density Estimate (KDE) Plot): يظهر وظيفة كثافة الاحتمال لمتغير مستمر.

- **مخطط الكمان (Violin Plot):** مفيد عندما يكون عرض الملاحظات الفردية غير عملي، مثالي لمجموعات بيانات كبيرة، ويصور كثافة احتمال البيانات عبر القيم.

مخطط الكمان (Violin Plot) يجمع بين جانبين من توزيع البيانات في تصوير واحد:
1. ميزات رسم الصندوق (Box Plot): الوسيط، المسافة بين الربعين.
2. وظيفة كثافة الاحتمال (PDF)، أو دلالة الكثافة للاحتمال.

- **أخيرًا، مخططات ثلاثية الأبعاد (3D Plots):** رسم نقاط ثلاثي الأبعاد، رسم أعمدة ثلاثي الأبعاد، ورسم سطح ثلاثي الأبعاد.

هناك أدوار مشتركة لتصور البيانات وتشمل:

-
- عرض التغيير مع مرور الوقت: استخدم مخططات الأعمدة، ومخططات الخطوط، ومخططات الصندوق والشارب.
- توضيح التركيب من الجزء إلى الكل: اختر المخططات الدائرية، والمخططات المساحية المكدسة، ومخططات الأعمدة المكدسة.
- مقارنة قيم المجموعات: استخدم مخططات الأعمدة، والهيستوغرامات، ومنحنيات الكثافة، ومخططات الكمان، ومخططات الصندوق والشارب.
- تحليل العلاقات بين المتغيرات: استخدم مخططات النقاط المتناثرة، ومخططات الفقاعات، والخرائط الحرارية.
- فحص البيانات الجغرافية.

اختيار نوع المخطط الصحيح أمر حاسم لتصور البيانات بشكل فعّال، اعتمادًا على نوع البيانات والسرد المقصود. من المهم أن يكون التصور المرئي متناسبًا مع طبيعة البيانات ومتطلبات الجمهور.

٥ البيانات في العالم الحقيقي

الأهداف التعليمية:

- فهم تحويل البيانات في الصناعات الحديثة.
- استكشاف الشركات في الشرق الأوسط باستخدام البيانات في الوقت الفعلي للتحليلات والقرارات المستندة إلى الذكاء الاصطناعي.
- دراسة تأثير علم البيانات في الرعاية الصحية والتحليلات الرياضية.
- فهم أخلاقيات علم البيانات: الخصوصية والأمن والتحيز والإنصاف والشفافية والموافقة.

النص:

تعمل البيانات، مثلها مثل النفط في الثورة الصناعية، على تغذية العصر الرقمي اليوم مع ارتفاع الطلب عبر الصناعات، من الشركات الناشئة إلى الشركات الكبيرة.

على عكس البيانات المحدودة أو التناظرية أو الرقمية المبكرة في الماضي، تعمل التقنيات المتقدمة اليوم على تعزيز توليد البيانات بشكل كبير.

أدى ظهور التفاعلات الرقمية وأجهزة إنترنت الأشياء إلى زيادة كبيرة في حجم البيانات، متجاوزًا الحد الأدنى لجمع البيانات يدويًا في الأوقات السابقة.

تعمل الشركات على تحويل عملية صنع القرار باستخدام البيانات في الوقت الفعلي والذكاء الاصطناعي، والانتقال من الحدس إلى الاستراتيجيات القائمة على البيانات في قطاعات مختلفة مثل الرعاية الصحية.

هذا النهج يولد أفكارًا ورؤى جديدة، خاصة في صناعات مثل ألعاب الفيديو لتعزيز التفاعل.

كما أنه يدفع الابتكار والتنبؤ باتجاهات السوق والاستراتيجيات المستهدفة، مما يعزز الكفاءة والاستدامة.

في هذا العصر، تعمل الشركات التي تركز على البيانات على تحسين الأداء والمساهمة المجتمعية والأثر البيئي، مما يؤدي إلى رفع سمعتها ومعايير أعمالها.

تعمل التحسينات المستندة إلى البيانات على تحويل مختلف الصناعات، وتخصيص الاستراتيجيات والعمليات لتكون أكثر كفاءة وفعالية وتركز على العملاء.

- **استراتيجية العمل**

تستخدم شركات الشرق الأوسط بيانات العملاء والسوق لتعزيز الاستراتيجيات، مع التركيز على التسويق الشخصي، وتحسين التوظيف في المتجر، والمخزون المبسط لتحسين تجربة العملاء والكفاءة التشغيلية.

- **الرعاية الصحية**

في مجال الرعاية الصحية في الشرق الأوسط، مثل النظام الصحي الإقليمي في منطقة الشرق الأوسط وشمال أفريقيا، يعد علم البيانات أمرًا أساسيًا لتعزيز التشخيص وخطط العلاج ورعاية المرضى، وذلك باستخدام تحليلات البيانات لتحسين نتائج المرضى وتقديم الرعاية الصحية بكفاءة.

- تستفيد المؤسسات المالية في الشرق الأوسط من البيانات لإدارة المخاطر، والخدمات الشخصية، واستراتيجية الاستثمار، وتحسين عملية صنع القرار وتقديم المشورة المخصصة للعملاء.

- **إدارة الإمدادات**

تستخدم الشركات في الشرق الأوسط علم البيانات لتحسين الخدمات اللوجستية، وإدارة المخزون بكفاءة، والتنبؤ الدقيق بالطلب، مما يؤدي إلى قدر أكبر من الشفافية، وأوقات عبور أقصر، وانخفاض النفقات التشغيلية.

- **الموارد البشرية**

تستخدم أقسام الموارد البشرية، بما في ذلك تلك الموجودة في شركات مثل Akhtaboot، البيانات لفهم اتجاهات المواهب، والتعرف على أنماط التوظيف، وتعزيز استراتيجيات التوظيف، مما يؤدي إلى توظيف أفضل الكفاءات، وتحسين مشاركة الموظفين، وارتفاع معدلات الاحتفاظ.

- **التحليلات الرياضية**

في الرياضات في العالم العربي، يستخدم الفرق والمدربون البيانات لتحسين أداء اللاعبين، وتحسين الاستراتيجيات، والتوظيف المستنير، وإحداث ثورة في كيفية الاستعداد والتنافس.

- **التعليم**

تستخدم المؤسسات والمنصات التعليمية البيانات لتخصيص التعلم وتعزيز النتائج من خلال الجمع بين التدريس البشري وأدوات الذكاء الاصطناعي للحصول على تجربة تعليمية شاملة وفعالة.

دعونا نتفحص دراسة حالة فعلية في علم البيانات التطبيقية.

قم بإلقاء نظرة على متاجر باندا للبيع بالتجزئة التابعة لمجموعة سافولا:
التنبؤ بالمبيعات وإدارة المخزون:

- تعريف المشكلة: من المرجح أن تركز متاجر باندا للبيع بالتجزئة، التابعة لمجموعة سافولا على تحسين إدارة المخزون لتحقيق التوازن بين العرض والطلب بشكل فعال.
- جمع البيانات: يتضمن ذلك جمع البيانات حول اتجاهات المبيعات وتفضيلات العملاء والخدمات اللوجستية لسلسلة التوريد وديناميكيات السوق.
- تحليل البيانات: من المحتمل أن تستخدم متاجر باندا هذه البيانات لتحليل الأنماط والتنبؤ باتجاهات المبيعات المستقبلية، مما يساعد في إدارة المخزون بكفاءة.
- النتيجة والتطبيق: يمكن أن تكون النتائج المتوقعة إدارة أكثر فعالية للمخزون، وتقليل النفايات، وتحسين رضا العملاء، وزيادة كفاءة المبيعات.

ماذا عن كشف الاحتيال؟
بالتركيز بشكل خاص على دراسة جرائم الاحتيال والاقتصاد PwC الشرق الأوسط لتحليل كشف الاحتيال:

- تعريف المشكلة: يسلط استطلاع PwC الضوء على زيادة في أنواع مختلفة من الاحتيال، بما في ذلك المشتريات والعملاء والرشوة والفساد والجرائم الإلكترونية، مما يشكل تحديات كبيرة للشركات الإقليمية في مشهد تكنولوجي دائم التغير.

- جمع البيانات: تقوم شركة PwC بجمع البيانات من خلال دراسات استقصائية واسعة النطاق تشمل المديرين التنفيذيين وكبار موظفي مكافحة الاحتيال، والتي تغطي جوانب مثل أنواع الاحتيال وتكراره وتأثيره التنظيمي.

- تحليل البيانات: في استطلاع PwC تتضمن مرحلة التحليل فحص البيانات المجمعة بحثًا عن الاتجاهات والقواسم المشتركة وأنماط الاحتيال الناشئة. والهدف هو فهم طبيعة التهديد، وتقييم تدابير مكافحة الاحتيال، وتحديد مجالات التحسين، وخاصة في مجال التكنولوجيا والامتثال.

- النتيجة والتطبيق: يكشف استطلاع PwC عن وجود انفصال بين نوايا المنظمات في الشرق الأوسط لمكافحة الاحتيال وأدائها الفعلي. ولسد هذه الفجوة، توصي PwC بتعزيز جهود مكافحة الاحتيال، وتعزيز الضوابط الداخلية، وتحسين السياسات، والاستثمار في التدريب، وإعطاء الأولوية لاعتماد التكنولوجيا للامتثال وأنظمة مكافحة الاحتيال.

تعتبر الاعتبارات الأخلاقية في علم البيانات بمثابة ألغاز معقدة تتطلب معالجة دقيقة. وتشمل الجوانب الرئيسية خصوصية البيانات والأمن، والتي تعتبر حيوية لحماية المعلومات الشخصية من الوصول غير المصرح به. التحيز والعدالة تعتبر هذه المخاوف بالغة الأهمية، لأنها يمكن أن تؤدي إلى قرارات منحرفة في مجالات مثل التوظيف، وتطبيق القانون، والتمويل.الشفافية وقابلية التفسير تعد نماذج البيانات المعقدة، وخاصة في التعلم الآلي، ضرورية للمساءلة وفهم القرار. وأخيرا،الموافقة وملكية البيانات تتعلق بحقوق الأفراد في بياناتهم، مع التركيز على أهمية الموافقة المستنيرة والتحكم في المعلومات الشخصية.

وفي منطقة الشرق الأوسط وشمال أفريقيا، قامت دول مثل مصر بسن قوانين لحماية البيانات مستوحاة من اللوائح الدولية مثل اللائحة العامة لحماية البيانات. يتوافق **قانون حماية البيانات المصري (القانون رقم 151 لعام 2020)**، الذي دخل حيز التنفيذ منذ أكتوبر 2020، مع مبادئ اللائحة العامة لحماية البيانات مثل تقليل البيانات والدقة والأمن والمشروعية.

كما ينص القانون على الإبلاغ عن الانتهاكات في غضون 72 ساعة ويفرض عقوبات كبيرة، بما في ذلك الغرامات وأحكام السجن المحتملة، على الانتهاكات المختلفة.

٦ المسارات الوظيفية في الذكاء الاصطناعي

الأهداف التعليمية:

- تتبع تطور علم البيانات: من التحليل الأساسي إلى الذكاء الاصطناعي المتقدم، مع ملاحظة التعقيد المتزايد.
- استكشاف المسارات الوظيفية المتنوعة: من مهندسي التعلم الآلي إلى محللي البيانات، واستيعاب المهارات المطلوبة.
- التعرف على تأثير علم البيانات: تمكين الابتكار وتجهيز الأفراد لأسواق العمل المستقبلية عبر القطاعات المختلفة.

النص:

في المشهد التكنولوجي المتطور باستمرار، يبرز علم البيانات (Data Science) كمجال ديناميكي سريع النمو.

فهو يمزج الخبرة الإحصائية والمعرفة الخاصة بالمجال ومهارات البرمجة لاستخراج رؤى ذات معنى من البيانات.

علم البيانات، الذي يدمج مجالات مثل الرياضيات والإحصاء وعلوم الحاسوب، يقدم مجموعة متنوعة من فرص الحياة المهنية، تلبي احتياجات الأدوار التقنية وغير التقنية على حد سواء.

يؤثر هذا المجال سريع النمو على قطاعات مختلفة، مما يسمح بحل المشكلات بشكل مبتكر واتخاذ القرارات القائمة على البيانات. إن نطاقها الواسع من الأدوار، بدءًا من متخصصي التعلم الآلي إلى محللي البيانات الإستراتيجية، يسلط الضوء على طبيعتها المتنوعة في عالم اليوم.

إن العمل في مجال علم البيانات يضعك في طليعة صياغة المستقبل، ويقدم لك فرصًا فريدة من أجل:

- **تشكيل تقنيات المستقبل:** يساهم علماء البيانات في تطوير تقنيات جديدة، ليكونوا بمثابة الأساس للاكتشافات المستقبلية.

- **حل المشكلات العالمية:** يعالج علم البيانات القضايا العالمية الملحة، مثل تغير المناخ وأزمات الرعاية الصحية، مع إمكانية إحداث تأثير عميق على المجتمع.

- **تحفيز الابتكار:** يدفع علماء البيانات الحدود ويخلقون الإمكانيات في مجالات مثل الذكاء الاصطناعي والتحليلات التنبؤية، مما يدفع الابتكار.

- **تعزيز الاستعداد للعمل:** يزودك الانخراط في علم البيانات بمهارات متعددة الاستخدامات، مما يعزز استعدادك لاقتصاد عالمي يعتمد على البيانات.

لنذهب مع سامي في رحلة عبر تطور مسارات مهنة علم البيانات!

المرحلة الأولى: الظهور والنمو المبكر

بدأ سامي رحلته في علم البيانات خلال الأيام الأولى لهذا المجال، حيث قام بدمج الإحصاء وتكنولوجيا المعلومات.
بعد تخرجه حديثًا من الكلية، بدأ باستخدام أدوات مثل SAS وSPSS، وهي جزء من مجموعة رائدة تكشف عن رؤى في البيانات والأنماط، وتضع الأساس لعلم البيانات الحديث.

المرحلة الثانية: التوسع والتبني الرئيسي
في بدايات العقد الأول من القرن الحادي والعشرين، شهد سامي ولعب دورًا في النمو السريع لهذا المجال. أحدثت البيانات الضخمة وتعلم الآلة ثورة في معالجة البيانات وتحليلها. قام سامي بالتكيُّف عن طريق تعلم لغات البرمجة Python و R واستكشاف نماذج بيانات معقدة. شهد سامي كيف توسع مجال علم البيانات ليشمل العديد من الصناعات، وكيف ظهرت برامج تعليمية لتلبية الطلب على علماء البيانات المهرة. كان هذا العقد بمثابة تحول عميق في مسيرة سامي المهنية وفي مشهد علم البيانات ككل.

المرحلة الثالثة: النضج والتخصص
في المرحلة الحالية من علم البيانات، يعمل سامي في مجال متخصص وناضج حيث تكون تقنيات الذكاء الاصطناعي المتقدمة قياسية، وتتنوع الأدوار.
بصفته عالم بيانات متمرس، فهو متخصص في التكامل الأخلاقي للذكاء الاصطناعي في استراتيجيات الأعمال ويقوم بتوجيه عشاق البيانات الطموحين، مع التركيز على الأخلاقيات والتخصص. لقد تطور علم البيانات بشكل كبير، ليصبح عنصرًا محوريًا في استراتيجية الأعمال الحديثة، حيث يمزج بين البراعة التقنية والمسؤولية الأخلاقية.

يتوقع سامي التطور السريع لهذا المجال، مع التركيز على التعلم الآلي المتقدم ودعم الاستخدام الأخلاقي للبيانات. كما يتصور تطبيقات خاصة بمجالات محددة في الوقت الفعلي تطمس الخطوط الفاصلة بين علم البيانات والحياة اليومية، مما يجعلها جزءًا لا يتجزأ من جوانب مختلفة من العمل والحياة، وذلك بفضل التقنيات الجديدة والتعاون متعدد التخصصات. لقد أتاحت مسيرة سامي المهنية له أن يشهد ظهور أدوار متنوعة في علم البيانات.

شهد سامي التطور من علماء البيانات العامة إلى أدوار متخصصة مختلفة:

- مسجلو البيانات ومدخلي البيانات يضمنون دقة البيانات المستخدمة في التحليل.
- مهندسو البيانات ومهندسو تصميم البيانات يبنون البنية التحتية الأساسية للبيانات.
- مديرو منتجات الذكاء الاصطناعي ومحللو الأعمال يربطون بين الأهداف التقنية والأهداف التجارية.
- باحثو الذكاء الاصطناعي ومهندسو التعلم الآلي ومهندسو عمليات التعلم الآلي يبتكرون باستخدام تقنيات التعلم الآلي.
- مهندسو الرؤية الحاسوبية ومهندسو معالجة اللغة الطبيعية يتخصصون في مجالات متقدمة.
- علماء البيانات ومحللو البيانات يفسرون البيانات الرقمية لاتخاذ قرارات مدروسة.

تدفع هذه الأدوار النمو الديناميكي وتأثير مجال علم البيانات.
في الساحة التكنولوجية المعتمدة على البيانات، يظهر دورين رئيسيين هما عالم البيانات ومهندس التعلم الآلي. لكل منهما مسؤوليات ومهارات فريدة:

عالم البيانات:

- يحدد ويحقق المشاكل التجارية المناسبة لحلول التعلم الآلي.
- يحلل ويصور البيانات عبر دورة حياة التعلم الآلي باستخدام أدوات مثل Tableau وPower BI.
- يتفوق في التحليل الإحصائي وتقديم الأفكار لأصحاب المصلحة.

مهندس التعلم الآلي:

- يركز على الجوانب التقنية للتعلم الآلي، بما في ذلك نشر النماذج وشرح البيانات.

- يحدد مجموعات البيانات الإضافية ويحسن نماذج الأداء.
- يراقب أداء النموذج والصيانة وتصحيح الأخطاء، غالبًا باستخدام الأنظمة الأساسية السحابية مثل AWS، Azure أو برنامج Google Cloud Platform.
- إجراء اختبار الاستدلال على الأجهزة المختلفة، من خلال نهج عملي.

تساهم هذه الأدوار في الاستخدام الفعال للبيانات في التكنولوجيا.

يعد فهم المسميات الوظيفية لعلم البيانات أمرًا بالغ الأهمية لاختيار المسار الوظيفي المناسب، حيث غالبًا ما تتداخل مسؤولياتها.

المسميات الوظيفية القادمة في علم البيانات هي:

- أخصائي حوكمة البيانات
- مسؤول خصوصية البيانات
- مهندس البيانات السحابية
- خبير تكامل البيانات
- مدير جودة البيانات
- مطور البيانات الضخمة
- محلل امتثال البيانات
- مطور تصور البيانات
- أخصائي التحليلات في الوقت الحقيقي
- مدرب ثقافة البيانات
- محلل بيانات إنترنت الأشياء
- خبير التحليلات التنبؤية
- راوي البيانات
- مستشار أخلاقيات البيانات

لماذا تعتقد أن هناك حاجة متزايدة لمهارات علوم البيانات؟

- يتم إنتاج كمية هائلة من البيانات من مصادر متنوعة، مثل:

- الأجهزة المتصلة بإنترنت الأشياء (IoT).
- منصات وسائل التواصل الاجتماعي
- المعاملات التي تتم عبر التجارة الإلكترونية
- استخدام الهاتف المحمول
- أنواع مختلفة من أجهزة الاستشعار
- الكاميرات وأجهزة التصوير

- تحظى مهارات علم البيانات بتقدير كبير في سوق العمل اليوم، حيث تشمل كل شيء بدءًا من التحليل الإحصائي وحتى الرؤية الإستراتيجية. تفتح هذه الكفاءات العديد من المسارات الوظيفية المتطورة في مختلف الصناعات
- يحدث علم البيانات ثورة في مجموعة واسعة من الصناعات، بما في ذلك تكنولوجيا المعلومات والتعليم والرعاية الصحية وتجارة التجزئة، مما يؤدي إلى زيادة الطلب على متخصصي علم البيانات.

- من المتوقع أن يرتفع الطلب على المتخصصين في علم البيانات بشكل كبير، مدفوعًا بالتقدم التكنولوجي المستمر في تحليل البيانات وتطبيقها المتوسع عبر الصناعات المتنوعة.

لبدء حياتك المهنية في مجال علم البيانات:

- أنشئ محفظة قوية لعلم البيانات تعرض مهاراتك وإنجازاتك، على غرار المحفظة الفنية للفنان.

- استخدم منصات مثل Kaggle، المعروفة بدورها في مجال علم البيانات، لتعلم مهارات ومشاريع البيانات وتطويرها ومشاركتها.

- ابقى على اطلاع على أحدث الاتجاهات والتقنيات من خلال التسجيل في الدورات التدريبية عبر الإنترنت والحصول على شهادات من منصات مثل Coursera أو Udacity أو edX.

- انضم إلى مجتمعات علم البيانات، وساهم في المشاريع مفتوحة المصدر، وتعاون في تعلم مفاهيم علم البيانات، واقرأ الأوراق البحثية بانتظام للتقدم في هذا المجال.

والأهم من ذلك هو أنه يمكننا الاستفادة من إمكانات علم البيانات في عالم ريادة الأعمال وتحديد الفرص التجارية الفريدة من خلال استيعاب:

متطلبات السوق + كيف يمكن للرؤى تقديم حلول استثنائية + قدرات تقنيات البيانات

تتبع شركة Code Three Fourteen Technologies، وهي شركة إماراتية متخصصة في الحلول التقنية، خطوات التطوير الرئيسية التالية:

- تكوين الأفكار: يقومون بالعصف الذهني وإنشاء أفكار مبتكرة لمواجهة التحديات التنظيمية.
- تجميع الفِرَق: يقومون بتجميع فريق ماهر يتمتع بخبرات متنوعة.
- التطوير: يقوم الفريق ببناء الحلول التكنولوجية.
- الاختبار: يضمن الاختبار الصارم الأداء الوظيفي والموثوقية.
- الإطلاق: يقدمون حلولاً للسوق أو ينفذونها في المنظمات.

تعتبر هذه الخطوات حيوية لتقديم حلول تقنية فعالة ومبتكرة.

تواجه كل شركة تقنية، بما في ذلك الشركات العاملة في مجال البيانات والذكاء الاصطناعي، تحديات مثل المعضلات الأخلاقية، وتكاليف البيانات، وموارد الحوسبة، وخصوصية البيانات، وتوافر البيانات، وجدوى المشروع. هذه هي العقبات النموذجية في الصناعة.

٧ الدور الحاسم لعلم البيانات لرواد الأعمال وفي مكان العمل

الأهداف التعليمية:

- تحليل البيانات الضخمة للحصول على رؤى السوق وسلوك العملاء وفرص الأعمال، مما يؤدي إلى نجاح ريادة الأعمال في سوق تعتمد على البيانات.
- تطبيق علم البيانات لتحسين العمليات من خلال التحليلات التنبؤية والأتمتة والتسويق المبني على البيانات، مما يعزز كفاءة الأعمال والربحية.

النص:

أصبح علم البيانات، وهو مجال ديناميكي ومتطور، أداة لا غنى عنها في مجموعة أدوات ريادة الأعمال، ولكن هل تساءلت يومًا عن السبب؟ تخيل أنك رائد أعمال في عالم اليوم الذي يعتمد على البيانات. هل ستبقى في صدارة المنافسة، وتفهم أعمق رغبات عميلك، وتتخذ القرارات التي تدفع عملك إلى الأمام؟ تكمن الإجابة في قوة علم البيانات.

تخيل أنك تستطيع التنبؤ بالاتجاه الكبير التالي للسوق! يُمكِّن علم البيانات رواد الأعمال من تحقيق ذلك. من خلال **تحليل مجموعات البيانات الضخمة**، يمكن لرواد الأعمال اكتشاف أنماط ورؤى خفية تكشف سلوكيات المستهلكين وتحولات السوق. هذه النظرة الثاقبة لا تقدر بثمن في وضع استراتيجيات تبقي الشركات في طليعة التقدم.

لا يقتصر دور علم البيانات فقط على النظر للخارج، بل يهتم أيضًا بالنظر إلى الداخل. **هل تُدار عملياتك بكفاءة؟ هل تُستَغَل مواردك بأقصى إمكانياتها؟** من خلال التحليلات التنبؤية والأتمتة، يعمل علم البيانات على تبسيط العمليات التجارية، مما يعزز الإنتاجية ويقلل من النفقات غير الضرورية.

يمتد تأثير علم البيانات إلى عملية صنع القرار أيضًا. كم مرة تعثرت الشركات بسبب القرارات المبنية على المشاعر الغريزية بدلاً من الحقائق الثابتة؟ تتيح الرؤى المستندة إلى البيانات اتخاذ قرارات محسوبة ومستنيرة بشكل أكبر، مما يقلل من المخاطر المرتبطة بخيارات العمل ويزيد من احتمالية النجاح.

وإلى جانب ذلك، فكر في مجال التسويق وعلاقات العملاء. كيف تضمن أن رسائل التسويق الخاصة بك تحقق أثرًا قويًا؟ يمكن لعلم البيانات إجراء تجزئة دقيقة للعملاء، وذلك يسمح لرواد الأعمال بتخصيص جهودهم التسويقية لجمهور محدد، مما يؤدي إلى حملات أكثر فعالية وعائد أفضل على الاستثمار.

لذا، كرائد أعمال أو محترف في مجال الأعمال، ألم يحن الوقت لاعتناق قوة علم البيانات وإطلاق الإمكانات الكاملة لعملك؟