



دورة تحليل البيانات الضخمة وذكاء القرار



دورة تحليل البيانات الضخمة وذكاء القرار

المرجع: 159761_786 التاريخ: 23 - 27 مايو 2027 المكان: جوهانسبرغ الرسوم: 4500 Euro

نظرة عامة

دورة تحليل البيانات الضخمة وذكاء القرار دورة متوسطة لمدة 5 أيام لوظائف تحليل البيانات وذكاء الأعمال وهندسة البيانات والتحول الرقمي ودعم القرار، ويعود المشاركون منها بحزمة تحليل البيانات الضخمة وأدلة القرار. تربط الدورة خصائص البيانات الضخمة والمعالجة الموزعة والتحليل الدفعي والتحليل المتدفق وخطوط البيانات والاستعلامات القابلة للتوسع وجودة البيانات والتصور وحوكمة البيانات الضخمة. تطور أجائل للتدريب قدرات تحليل البيانات الضخمة التطبيقية.

الفئة المستهدفة

- وظائف تحليل البيانات المسؤولة عن التهيئة واسعة النطاق والأدلة التحليلية
- وظائف ذكاء الأعمال المسؤولة عن الاستعلامات القابلة للتوسع وتقارير القرار
- وظائف هندسة البيانات المسؤولة عن خطوط البيانات وأنماط المعالجة الموزعة
- وظائف التحول الرقمي المسؤولة عن حالات الاستخدام وخيارات المنصات
- وظائف دعم القرار المسؤولة عن تفسير الرؤى وعرضها

تفترض الدورة أن المشاركين يستطيعون الاستعلام عن البيانات المهيكلية وتفسير التحليلات الوصفية وتتبع تدفق البيانات، وتستبعد البرمجة التمهيدية وإدارة البنية التحتية وشهادات المنصات ونشر أنظمة الإنتاج.

الإدارات والقطاعات المستهدفة

تدعم الدورة الإدارات والقطاعات التي تحلل بيانات مرتفعة الحجم أو السرعة أو التنوع.

- بيانات الخدمات المالية وتحليلات المخاطر
- التجزئة وتحليلات العملاء الرقمية
- التصنيع وعمليات بيانات المستشعرات
- تحليلات شبكات الاتصالات
- تحليل بيانات خدمات الرعاية الصحية
- التحول الرقمي في الخدمات العامة



أهداف التعلم

بنهاية هذه الدورة، سيتمكن المشاركون من:

- تحليل خصائص البيانات الضخمة وحالات استخدامها
- مقارنة أنماط التخزين والمعالجة الموزعة
- بناء تصاميم خطوط المعالجة الدفعية والمتدفقة
- تطبيق أساليب التهيئة والاستعلام القابلة للتوسع
- تقييم جودة البيانات والحوكمة ومخاطر المعالجة
- بناء أدلة قرار مرئية من مجموعات بيانات كبيرة

جدول الدورة

اليوم 1: تأطير تحليل البيانات الضخمة

- لوحة خصائص البيانات الضخمة وحالة الاستخدام
- خريطة أصحاب الأدوار في بنية NIST المرجعية للبيانات الضخمة
- مصفوفة أنماط التخزين والحوسبة الموزعة
- ملف عبء التحليل وزمن الاستجابة
- ورقة تعريف سؤال القرار والأدلة

اليوم 2: تصميم خطوط البيانات

- خريطة تدفق البيانات من المصدر إلى الاستخدام
- خط المعالجة الدفعية للاستيعاب والتحويل
- خط الأحداث والمعالجة المتدفقة
- سجل تعريف المخططات والبيانات الوصفية
- نموذج مناطق بحيرة البيانات ودورة حياتها



اليوم 3: تنفيذ التحليل القابل للتوسع

- خطة التقسيم الموزع والمعالجة المتوازية
- نمط الاستعلام القابل للتوسع باستخدام Spark SQL
- نمط التحليل المتدفق باستخدام Spark Structured Streaming
- نمط التحويل والتجميع باستخدام Kafka Streams
- قائمة التحقق من صحة النتائج التحليلية

اليوم 4: حوكمة الأدلة وعرضها

- بطاقة أبعاد جودة البيانات الضخمة
- خريطة تتبع نسب البيانات والمعالجة
- سجل مخاطر الأمن والخصوصية
- مصفوفة اختيار التصور واسع النطاق
- موجز رؤية القرار وحدودها

اليوم 5: ممارسة قرارات البيانات الضخمة

- تمرين: تصنيف حالة استخدام وبنية للبيانات الضخمة
- تمرين: تصميم خطوط المعالجة الدفعية والمتدفقة
- تمرين: اختيار أنماط الاستعلام والمعالجة القابلة للتوسع
- تمرين: التحقق من الجودة والحوكمة والأدلة المرئية
- المشروع الختامي: حزمة تحليل البيانات الضخمة وأدلة القرار

التمارين التطبيقية

تستخدم الدورة أنشطة مقترحة تستند إلى المعاملات المالية ورحلات التجزئة والمستشعرات الصناعية وعمليات الخدمات.

- نشاط مقترح: تصنيف متطلبات الحجم والسرعة والتنوع والمعالجة والقرار لحالة استخدام مختارة.
- نشاط مقترح: رسم بيانات المعالجة الدفعية والمتدفقة من المصدر حتى التحويل والتحليل والاستخدام.
- نشاط مقترح: مقارنة أنماط الاستعلام الموزع ومعالجة التدفق وفق زمن الاستجابة والحالة والتحقق.
- نشاط مقترح: جمع الجودة والنسب والمخاطر والتصوير وحدود القرار في حزمة الأدلة النهائية.



الأسئلة الشائعة

لمن تناسب دورة تحليل البيانات الضخمة، وماذا تفترض؟

تناسب وظائف التحليل وذكاء الأعمال وهندسة البيانات والتحول ودعم القرار، وتفترض خبرة في الاستعلامات المهيكلية والتحليلات الوصفية ومنطق تدفق البيانات.

كيف يختلف تحليل البيانات الضخمة عن تدريب تحليل البيانات العام؟

يركز تحليل البيانات الضخمة على التخزين الموزع والمعالجة المتوازية وخطوط المعالجة الدفعية والمتدفقة والاستعلامات القابلة للتوسع وتحديات الحوكمة عندما لا تكفي الأنظمة المنفردة.

متى يستخدم تحليل البيانات الضخمة المعالجة الدفعية أو المتدفقة؟

تناسب المعالجة الدفعية البيانات المحدودة والتحليل المجدول، بينما تناسب المعالجة المتدفقة الأحداث المستمرة والقرارات التي تتطلب زمنا أقصر أو حالة متواصلة.

كيف يدعم Spark وKafka تحليل البيانات الضخمة؟

يدعم Spark الاستعلامات والتحويلات الموزعة والتحليل المتدفق المهيكل، بينما يدعم Kafka نقل الأحداث وأنماط معالجة التدفق؛ وتستخدمهما الدورة كأمانة لا كموضوعات لشهادة منتج.

كيف يجري التحقق من أدلة قرار البيانات الضخمة؟

يجري التحقق عبر ضوابط المخططات وفحوص الجودة ونسب البيانات واختبارات المعالجة والمطابقة التحليلية والمراجعة المرئية وتوثيق المخاطر وحدود التفسير.

الخاتمة

يعود المشاركون بحزمة تحليل البيانات الضخمة وأدلة القرار التي تربط حالات الاستخدام والبنية والخطوط والاستعلامات القابلة للتوسع وأنماط التدفق والجودة والحوكمة والتصور والتفسير. تحول الحزمة أنشطة البيانات الكبيرة المنفصلة إلى مسار تحليل قابل للتتبع. وتدعم قرارات منسقة بين التحليل والهندسة وذكاء الأعمال والتحول والعمليات.