€TRAINING

تحليل البيانات الضخمة



28 إبريل - 2 مايو 2024

كوالالمبور (ماليزيا)

تحليل البيانات الضخمة

رمز الدورة: D525 تاريخ الإنعقاد: 28 إبريل - 2 مايو 2024 دولة الإنعقاد: كوالالمبور (ماليزيا) - التكلفة: 5850 يورو

مقدمة عن البرنامج التدريبي:

البيانات الضخمة هي عامل تغيير يتحدى الطرق التقليدية التي اتخذها قادة المنظمات في اتخاذ القرارات، في هذا البرنامج سيتم بناء للمشاركين الثقة لتصميم معماريات البيانات الكبيرة لدعم الحلول القائمة على التحليلات داخل مؤسساتهم، وتوفرأيضاً خبرة في استخدام تقنيات البيانات الهامة الرئيسية المستخدمة في نشر تطبيقات البيانات المكثفة، وسيكتسب المشاركون المعرفة والمهارات التي يحتاجونها لتجميع وإدارة مشروع تحليلي ضخم للبيانات الضخمة، وأخيراً سيحصل المشاركون على مقدمة مفاهيمية لهياكل البيانات التي تدعم خوارزميات التعلم الآلي وحالات استخدام الذكاء الاصطناعي. وسيعمل المشاركون على تحديد المجالات داخل مؤسستهم التي يمكن تحسينها من خلال التطبيقات الكبيرة القائمة على البيانات، وأنواع التحسينات التي يمكن إجراؤها من خلال العمليات التحليلية، وسيقود المشاركون من خلال سلسلة من التدريبات العملية وورش العمل، حيث ستتاح لهم الفرصة لتطبيق أساليب الاختبار والنهج العملية التي يتعلمونها خلال الدورة، وفي نهاية الدورة سينتج المشاركون خطة بيانات كبيرة ومخططاً معمارياً كبيراً لاستخدامها كمقترح مخطط داخل منظماتهم الخاصة.

أهداف البرنامج التدريبي:

في نهاية البرنامج سيكون المشاركون قادرين على:

- تصميم خطط تنفيذ البيانات الضخمة وإنشاء استراتيجيات للحلول التي تعتمد على البيانات.
- شرح تحديات البيانات الضخمة والتقنيات التقليدية مثل "Excel"
- مناقشة التحديات والمزايا الرئيسية للنظام الإيكولوجي "Hadoop" والبنى الموزعة للبيانات الضخمة الأخرى.
- عرض ومناقشة التقنيات الرئيسية لتخزين البيانات الكبيرة والحوسبة ، مثل "PostgreSQL" و "MongoDB".
- مناقشة خوارزميات تعلم الآلة الشائعة وأهمية الأخلاقيات في تحليلات البيانات والذكاء الاصطناعي.
- تقديم مخطط معماري لحالات الاستخدام التي تركز على التحليلات.

الفئات المستهدفة:

- مدراء قواعد البيانات ومسؤولو النظام.
- محللو الأعمال ومتخصصو استخبارات الأعمال
- مهنيو البيانات
- موظفو البنوك والمؤسسات المالية.
- المهنيون الإداريون والذين يسعون إلى فهم استراتيجيات وتقنيات البيانات الضخمة.

محاور البرنامج التدريبي:

الوحدة الأولى:

تخزين البيانات الكبيرة:

- ما هي البيانات الكبيرة؟
- .الكبيرة البيانات من "V" 5
- كيف ترتبط البيانات الكبيرة بتحليلات البيانات.
- تأثير البيانات الضخمة على التقنيات.
- مفاهيم البيانات الكبيرة الرئيسية وأنواع البيانات.
- النص والصوت والصور.

- أدوار مهنية كبيرة البيانات.
- معماريات البيانات الكبيرة والنماذج.
- نظام "Hadoop" البيئي.
- نظرة عامة على "Hadoop".
- نظام الملفات الموزعة "{Hadoop {HDFS".
- معالجة متوازية للغاية "MPP" مقابل التطبيقات الموزعة في الذاكرة.
- NoSQL DBs. مقابل RDBMSs
- .كاساندرا ، MongoDB ، PostgreSQL
- تدفق البيانات.
- تخزين البيانات مقابل "Mart Data".
- .العمارة "Kappa" مقابل العمارة "Lambda"

الوحدة الثانية:

حوسبة البيانات الكبيرة:

- كيفية الوصول إلى البيانات الكبيرة.
- دور الحوسبة السحابية.
- خطر حركة البيانات.
- التشبيك والمشاركة في الموقع.
- استخراج البيانات الكبيرة، التحويل، التحميل {ETL}.
- تقنيات حساب البيانات الكبيرة.
- الحوسبة الموزعة.
- مجموعات عالية الأداء.
- بث: "Storm" ، "Spark"جريان منظم.
- تقنيات البيانات الكبيرة الأخرى: كافكا ، إلخ.
- تطبيقات السحاب للبيانات الكبيرة.

الوحدة الثالثة:

تقديم تحليلات البيانات الضخمة والذكاء الاصطناعي {AI}:

- أساسيات تحليلات البيانات.
- الأدوار والأهداف.
- المفاهيم الأساسية في الرياضيات والإحصاء.
- التقنيات والتطبيقات الرئيسية.
- بنية "Analytics".
- سحابة مقابل على فرضية.
- مخزن البيانات.
- أدوات التحليلات.
- حلول "ML Cloud" & "AI".
- مقدمة في الذكاء الاصطناعي.
- الجبر الخطي 101.
- تصنيف الصور.
- أهمية الأخلاق.

الوحدة الرابعة:

تخطيط مشروع البيانات الضخمة للتحليلات:

- كيف تلبي مشاريع البيانات الكبيرة الاحتياجات التنظيمية.
- دراسة حالة البيانات الكبيرة:
  - Netflix.
  - LinkedIn.

- Facebook.
- Google.
- Orbitz.
- Dell.
- وأخرين.

- أفضل الممارسات في تصميم المشروع.
- تقييم الحالة الحالية لمنظمتك.
- فرق البيانات الرأسية والمناقشات.
- اعتبارات لخطط مشروع البيانات الكبيرة.
- عصف ذهني لاستراتيجية تستند إلى البيانات.
- ممارسة تصميم المخططات المعمارية.

الوحدة الخامسة:

تصميم حلول البيانات الكبيرة:

- تحديد الفرص التحليلية.
- تحديد وتقييم المشكلة.
- وصف تأثير واستخدام البيانات لمعالجة المشكلة.
- تحديد مصادر البيانات المحتملة.
- عصف ذهني لاستراتيجية تحليلية للتنفيذ.
- التخزين والحوسبة.
- تحديد استراتيجية بيئة سحابة.
- تدوير أنظمة التخزين الرئيسية وحوسبة البيانات.