



دورة تحليل البيانات الضخمة

دورة شاملة لتعلم تحليل البيانات الضخمة باستخدام Hadoop و Spark و Hive وتقنيات التعلم الآلي والمعالجة في الوقت الحقيقي وقواعد بيانات NoSQL.

المدينة :	باريس	الفندق :	لو موريس
تاريخ البداية :	2026-01-26	تاريخ النهاية :	2026-01-30
الفترة :	Week 1	السعر :	\$ 5950

فكرة الدورة التدريبية

في العصر الرقمي الحالي، تقوم المؤسسات بتوليد كميات هائلة من البيانات بمعدل غير مسبوق. ومع ذلك، فإن القيمة الحقيقية تكمن في القدرة على استخلاص رؤى ذات معنى من هذه البيانات لدفع عملية صنع القرار المستنيرة واكتساب ميزة تنافسية. وبذلك فإن تحليل البيانات الضخمة هي المفتاح لإطلاق هذه الإمكانيات. تم تصميم هذه الدورة التدريبية حول تحليل البيانات الضخمة لتزويد المشاركين بالمعرفة والمهارات اللازمة لتسخير قوة البيانات الضخمة للتحليل ودعم القرار. من خلال مزيج من المفاهيم النظرية والتدريبات العملية ودراسات الحالة الواقعية، سيكتسب المشاركون فهماً عميقاً للمبادئ والأدوات والتقنيات المستخدمة في تحليلات البيانات الضخمة.

أهداف الدورة التدريبية

في نهاية هذه الدورة التدريبية، ستتعلم ما يلي:

- فهم أساسيات البيانات الضخمة: سيكتسب المشاركون فهماً قوياً لخصائص البيانات الضخمة، بما في ذلك الحجم والسرعة والتنوع والصدق، وكيف تشكل تحديات فريدة لأساليب معالجة البيانات التقليدية وتحليلها.
- استكشاف تقنيات البيانات الضخمة: سيتم تعريف المشاركين بتقنيات البيانات الضخمة الشائعة مثل قواعد بيانات Hadoop و Apache Spark و Apache Kafka و NoSQL. سوف يتعلمون كيف تتيح هذه التقنيات معالجة قابلة للتطوير وموزعة لمجموعات البيانات الكبيرة.
- معالجة وتحليل البيانات الرئيسية: سيتعلم المشاركون كيفية معالجة وتحليل البيانات الضخمة باستخدام أطر الحوسبة الموزعة مثل Apache Spark و MapReduce. سوف يكتسبون خبرة عملية في استخدام أدوات استيعاب البيانات وتحويلها والاستعلام عنها.
- اكتشاف التعلم الآلي باستخدام البيانات الضخمة: سوف يتعمق المشاركون في تقنيات التعلم الآلي المصممة لبيئات البيانات الضخمة. سوف يتعلمون كيفية بناء وتقييم النماذج التنبؤية باستخدام Apache Spark MLlib واستكشاف موضوعات متقدمة مثل التعلم العميق.
- فهم تحليلات البيانات في الوقت الفعلي: سوف يستكشف المشاركون معالجة البيانات في الوقت الفعلي وتحليلات التدفق باستخدام تقنيات مثل Apache Kafka و Apache Flink. سوف يتعلمون كيفية استيعاب ومعالجة وتحليل البيانات المتدفقة للحصول على رؤى واتخاذ القرار في الوقت المناسب.
- تطبيق تحليلات البيانات الضخمة في سيناريوهات عملية: من خلال دراسات الحالة والتمارين العملية، سيطبق المشاركون معرفتهم بتحليلات البيانات الضخمة على سيناريوهات العالم الحقيقي عبر مختلف الصناعات. سوف يتعلمون أفضل الممارسات لتصميم وتنفيذ حلول تحليلات البيانات الضخمة.
- تطوير مهارات التفكير النقدي وحل المشكلات: سيواجه المشاركون طوال الدورة تحدياً للتفكير النقدي وحل المشكلات المعقدة.

باستخدام تقنيات تحليل البيانات الضخمة. سيقومون بتطوير مهارات التفكير التحليلي الأساسية لاستخلاص رؤى ذات معنى من البيانات.

الفئات المستهدفة

هذه الدورة التدريبية موجهة لـ:

- محللو البيانات
- محترفي ذكاء الأعمال
- متخصصو تكنولوجيا المعلومات
- المديرين والتنفيذيين
- محترفي التسويق والمبيعات
- المتخصصين في الشؤون المالية والعمليات
- رجال الأعمال وأصحاب الأعمال
- المهنيين الأكاديميين والباحثين
- أي شخص مهتم بتحليلات البيانات الضخمة

منهجية الدورة

مع تدريب عملي على نموذج Hadoop وHDFS تبدأ الدورة بتقديم أساسيات البيانات الضخمة وخصائصها، ثم التعرف على نظام مع تنفيذ تمارين عملية على Hive وApache Spark البيئية مثل Hadoop يلي ذلك تعلم معالجة البيانات باستخدام أدوات MapReduce. وتقنيات التقييم والتحسين، مع Spark MLlib معالجة البيانات وتحليلها. بعد ذلك، يتم الانتقال إلى التعلم الآلي للبيانات الضخمة باستخدام كما تغطي الدورة تحليلات البيانات الضخمة في الوقت الحقيقي باستخدام TensorFlow وPyTorch مقدمة لأطر التعلم العميق مثل وقواعد بيانات الرسم NoSQL مع تطبيقات عملية على تدفق البيانات ومعالجتها. أخيرًا، يتم استعراض قواعد بيانات Apache Kafka وFlink البياني، دراسة حالات تطبيقية، ومناقشة الاتجاهات الناشئة في تحليل البيانات الضخمة، مع تلخيص شامل لمفاهيم ومهارات الدورة.

محاور الدورة

اليوم الأول: مقدمة في تحليل البيانات الضخمة

- مقدمة عن البيانات الضخمة وخصائصها
- نظرة عامة على تحليل البيانات الضخمة
- فهم أهمية وتطبيقات تحليل البيانات الضخمة
- مقدمة لنظام Hadoop البيئي
- أساسيات نظام الملفات الموزعة (HDFS) Hadoop
- مقدمة إلى نموذج MapReduce
- تدريبات عملية على Hadoop و MapReduce باستخدام عينات من مجموعات البيانات

اليوم الثاني: معالجة البيانات باستخدام أدوات النظام البيئي Hadoop

- مقدمة إلى أباتشي سبارك
- فهم Spark RDDs (مجموعات البيانات الموزعة المرنة)
- نظرة عامة على Spark SQL و DataFrames
- تمارين عملية مع Spark لمعالجة البيانات وتحليلها
- مقدمة إلى Apache Hive لتخزين البيانات والاستعلام عنها
- تمارين عملية مع Hive للاستعلام عن البيانات وتحليلها

اليوم الثالث: التعلم الآلي مع البيانات الضخمة

- مقدمة لمفاهيم التعلم الآلي
- التحديات والاعتبارات المتعلقة بالتعلم الآلي باستخدام البيانات الضخمة
- نظرة عامة على خوارزميات التعلم الآلي للبيانات الضخمة
- تمارين عملية مع التعلم الآلي باستخدام Apache Spark MLlib
- تقنيات التقييم والتحسين النموذجي
- مقدمة لأطر التعلم العميق للبيانات الضخمة (مثل TensorFlow و Py Torch)

اليوم الرابع: تحليلات البيانات الضخمة في الوقت الحقيقي

- مقدمة لمعالجة البيانات في الوقت الحقيقي
- نظرة عامة على Apache Kafka لتدفق البيانات في الوقت الفعلي
- مقدمة إلى Apache Flink لمعالجة التدفق
- تمارين عملية مع Apache Kafka لاستيعاب البيانات ومعالجتها في الوقت الفعلي
- تمارين عملية مع Apache Flink لمعالجة البث في الوقت الفعلي

اليوم الخامس: موضوعات وتطبيقات متقدمة

- مقدمة إلى قواعد بيانات NoSQL (مثل Apache Cassandra و MongoDB) للبيانات الضخمة
- نظرة عامة على قواعد بيانات الرسم البياني (على سبيل المثال، Neo4j) وتطبيقاتها

- دراسات الحالة وأفضل الممارسات في تحليلات البيانات الضخمة
- مناقشة حول الاتجاهات الناشئة والاتجاهات المستقبلية في تحليل البيانات الضخمة
- تحليل البيانات الضخمة نظرة عامة كاملة

الشهادات المُعتمدة

عند إتمام هذا البرنامج التدريبي بنجاح، سيتم منح المشاركين شهادة هاي بوينت رسمياً، اعترافاً بمعارفهم وكفاءاتهم المثبتة في الموضوع. تُعد هذه الشهادة دليلاً رسمياً على كفاءتهم والتزامهم بالتطوير المهني.