



بايثون Python لتحليل البيانات وأساسيات تعلم الآلة باستخدام pandas وlearn-scikit

12 - 16 Jul 2027

دبي - فندق خمس نجوم في وسط النعزال



بايثون Python لتحليل البيانات وأساسيات تعلم الآلة باستخدام pandas وscikit-learn

المرجع: 1248_7402 التاريخ: 12 - 16 Jul 2027 دبي - فندق خمس نجوم في وسط الأعمال الرسوم: SAR 19500

العقود التقني: مهارس - أسلوب التطبيق: مختبر تطبيقي

مقدمة

تمتلك مؤسسات كثيرة بيانات تشغيلية تفوق قدرة تحليلها على معالجتها في جداول البيانات، فتبقى الأسئلة المتعلقة بالعملاء والتكاليف والطلب على الخدمات بانتظار الفرق المتخصصة، أو تعالج بحلول يدوية لا يمكن تكرارها. تبني هذه الدورة من كور كونسبت قدرة عملية على استخدام لغة بايثون Python لدى المحللين، بدءاً من تنظيف البيانات واستكشافها بمكتبة pandas وحتى تدريب نماذج تعلم الآلة الأولى والتحقق منها بمكتبة learn-scikit. وتنفذ جميع الجلسات في بيئة عملهم واقع من مشكلة على مبني أساسي ونموذج التنفيذ لإعادة قابل تحليل بدفتر المشاركون ويخرج، واقعية بيانات على Jupyter

أهداف الدورة

- إعداد بيئة تحليل بلغة بايثون قابلة لإعادة التنفيذ باستخدام Jupyter والبيانات الافتراضية وإدارة الإصدارات
- تنظيف البيانات الجدولية وإعادة تشكيلها ودمجها بمكتبة pandas، وتقييم صلاحيتها للاستخدام وفق معايير جودة البيانات المعتمدة
- استكشاف البيانات ووثيقها بصرياً باستخدام Matplotlib وseaborn للاختبار فرضيات العمل وعرض النتائج
- صياغة سؤال العمل كمشكلة تعلم خاضع للإشراف أو غير خاضع للإشراف وفق منهجية DM-CRISP
- تدريب نماذج الانحدار والتصنيف والتجميع بمكتبة learn-scikit والتحقق منها وضبطها واختيار مقاييس التقييم المناسبة
- توثيق التحليل في دفتر وبطاقة نموذج يستطيع الزلاء إعادة تنفيذهما ومراجعتها

الفئة المستهدفة

- محللو البيانات والأعمال المنتقلون من جداول البيانات إلى التحليل البرمجي
- محللو المالية والموارد البشرية والعمليات المسؤولون عن التقارير الدورية والتحليلات الطارئة
- مطورو ذكاء الأعمال الذين يضيفون التحليل الإحصائي والتنبؤي إلى لوحات المعلومات
- الباحثون والإحصائيون في وحدات الخدمات العامة والسياسات الذين يعملون على البيانات الإدارية
- المهندسون وأخصائيو الجودة الذين يحلون بيانات العمليات والحساسات والفحص
- علماء البيانات المبتدئون الذين يرسخون أساسياتهم في بايثون وتعلم الآلة

محاور الدورة

اليوم الأول: أساسيات بايثون ومسار العمل التحليلي

- مراحل منهجية DM-CRISP لتتقيب البيانات وربطها بالعمل التحليلي اليومي
- إعداد دفتر Jupyter والبيئة الافتراضية وإدارة الحزم
- أنواع البيانات وجمال التحكم والدوال في بايثون لغراض التحليل
- مصفوفات NumPy والحساب المتجهي
- استيراد بيانات CSV وExcel وقواعد SQL ووواجهات API إلى pandas

اليوم الثاني: هياكل البيانات ومعالجتها ومعايير جودتها

- الفهرسة والاختيار في هياكل Series و DataFrame بـ مكتبة pandas
- معالجة القيم المفقودة والتكرارات وتحويل الأنواع في pandas
- إعادة التشكيل والدمج باستخدام groupby و merge و table_pivot و melt
- فحص جودة البيانات وفق معايير المواصفة IEC/ISO 5259-2 لجودة بيانات التحليلات وتعلم الآلة
- مكونات أنظمة تعلم الآلة ومصطلحاتها وفق المواصفتين IEC/ISO 23053 و IEC/ISO 22989

اليوم الثالث: التحليل الاستكشافي والنماذج النولى

- الإحصاء الوصفي وتحليل التوزيعات في pandas
- اختيار مخططات Matplotlib و seaborn المناسبة للنسلة التحليلية
- تحليل الارتباط واختبار الفرضيات بـ مكتبة SciPy
- الانحدار والتصنيف والتجميع بخوارزمية Means-k عبر مسارات learn-scikit
- معايير التقييم: RMSE والدقة والاستدعاء و AUC ROC ومعامل Silhouette

اليوم الرابع: التحقق من النماذج ومزالتها وتحسينها

- التحقق من التقاطع Fold-k ومنع تسرب البيانات
- تشخيص فرط التخصيص بمنحنيات التعلم وأساليب التنظيم
- معالجة عدم توازن الفئات بالتقسيم الطبقي وأوزان الفئات
- ضبط المعاملات الفائقة باستخدام GridSearchCV
- أهمية المتغيرات بطريقة التبدل Permutation لتفسير النموذج

اليوم الخامس: التطبيق العملي ودفتر التحليل القابل لإعادة التنفيذ

- مختبر تسرب العملاء: بناء نموذج تصنيف وتفسير نتائج
- مختبر الطلاب على الخدمات العامة: التجميع وتوصيف الشرائح
- إعداد بطاقة النموذج Card Model للنموذج الأساسي
- بناء دفتر التحليل القابل لإعادة التنفيذ والنموذج الأساسي
- مراجعة الزملاء للشيفرة وفق أعراف PEP 8 وتسليم الدفتر

المهارات المكتسبة

- البرمجة بايثون لغراض التحليل
- معالجة البيانات وتجهيزها
- تقييم جودة البيانات
- التحليل الاستكشافي للبيانات
- اختبار الفرضيات الإحصائية
- التعلم الخاضع للإشراف
- التحقق من النماذج
- التحليل القابل لإعادة التنفيذ

لهذا تحضر هذه الدورة

- تعود إلى عملك بدفتر تحليل قابل لإعادة التنفيذ ونموذج أساسي مبني على مشكلة بيانات من واقع عملك وقد راجعه الزملاء
- تستبدل الخطوات المتكررة في جداول البيانات بمسارات برمجية يعاد تشغيلها كلها وصلت بيانات جديدة
- تكتشف النتائج التي تبدو أفضل من أن تكون صحيحة، من تسرب البيانات إلى فرط التخصيص، قبل وصولها إلى صناع القرار
- تعمل على بيانات المختبر نفسها مع تحليلين من قطاعات أخرى وتقارن أساليب معالجة مشكلات البيانات الشائعة

الخاتمة

أصبحت بايثون لغة العمل الأساسية في التحليلات لأنها تتيح للهلال الانتقال من البيانات الخام إلى نهج مختبرة ضمن مسار عمل واحد قابل للتكرار. تبني هذه الدورة ذلك المسار خطوة بخطوة: إعداد البيئة ومعالجة البيانات، وفحص الجودة وفق مقاييس معتمدة، والتحليل الاستكشافي، والنهائج النولي، وانضباط التحقق الذي يحويها من النتائج المضللة. وبحول اليوم الأخير هذه المعارف إلى دفتر تحليل قابل لإعادة التنفيذ ونموذج أساسي يعود بهما المشاركون إلى عملهم ويطورونها مع وصول بيانات وأسئلة جديدة.