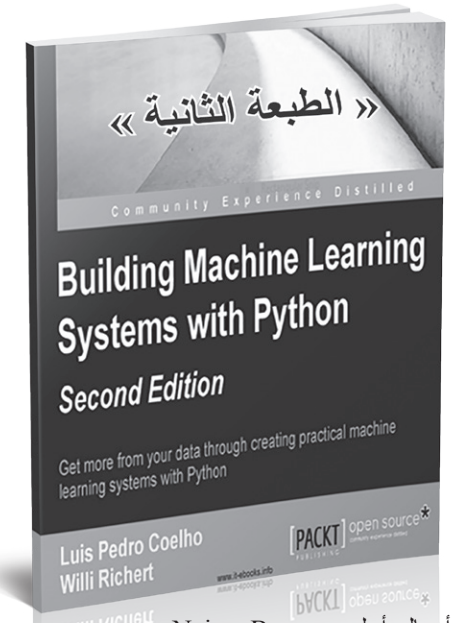
بناء نظم تعلم الآلة باستخدام لغة «Python»

أ.د. محمد محمد الهادي



هذا الكتاب من تأليف كل من Luis Pedro Coelho and Willi Richert ونشرته دار نشر PACKT Open Source وصدرت طبعته الثانية في عام ٢٠١٥. تحت ترقيم دولي ISBN ٩٧٨-١-٧٨٤٣٩-٢٧٧-٢ ويقع في حوالي ٣٢٦ صفحة. ويشتمل على اثني عشر فصلا وملحق.

وهذا الكتاب موجه في الأساس لمبرمجي لغة Python الذين يبغون تعلم كيفية أداء أسلوب تعلم الآلة باستخدام المصدر المفتوح. وبذلك تم استعراض نماذج تعلم الآلة الأساسية المبنية علي مثال فعلي. وفي نفس الوقت، فإن هذا العمل موجه أيضا لمتعلمي الآلة الذين يريدون البدء في استخدام لغة Python لبناء أنظمتهم. وتعتبر لغة Python ذات طبيعة مرنة للنمذجة التمهيدية، بينما تكتب اللوغاريتمات المحددة باستخدام لغة C أو C++. ويعتبر الكود الناتج سريع الاستخدام في الإنتاج أيضا. وتدرس لغة تعلم الآلة الآلات للتعلم في أداء المهام ذاتيا. وهي لغة سهلة الاستخدام بصفة عامة.

ويوضح هذا العمل كيف يمكن إيجاد الأنماط المختلفة المتواجدة في البيانات الخام باستخدام تعلم الآلة لاكتساب بصائر أعمق من البيانات المتاحة، مما يمثل مهارة متطلبة من قبل مطوري ومحللي التطبيقات المعاصرة وذلك باستخدام لغة Python لتطوير تطبيقات تعلم الآلة.

وفيما يتعلق بفصول هذا العمل الاثني عشر، فإنها تبدأ بفصل أول عن البدء باستخدام لغة Python الخاصة بتعلم الآلة. الذي يقدم فكرة عامة عن لغة الآلة الأساسية مع مثال سهل فعلي عنها. وعلي الرغم من سهولة هذه اللغة، فقد تتحدي القارئ مع تضمين مخاطر الملاءمة الفائقة أي الزائدة. ويختص الفصل الثاني بعرض موضوع التصنيف Classification مع أمثلة عن العالم الحقيقي لها، حيث تستخدم بيانات حقيقية عن التصنيف، بينما يدرب الحاسب الآلي على القدرة في تمييز القسام المختلفة المتضمنة في التصنيف. وقد ارتبط كل من الفصل الخامس والسادس والتاسع بموضوع التصنيف أيضا، حيث تم استعراض التصنيف المرتبط باكتشاف الإجابات الضعيفة الذي يوضح كيفية استخدام تباين التمييز فيما يتصل بنماذج تعلم الآلة، أي الارتباط المباشر باستخدام أسلوب الانحدار المنطقي لإيجاد ما إن كانت إجابة المستخدم على السؤال المطروح جيدة أم لا. واستطرد في الفصل السادس عن التصنيف أيضا تحليل الرأي العاطفي Sentiment من خلال شرح أعمال أسلوب Naive Bayes وكيفية استخدامه، إلى جانب استعراض أساليب الانحدار مثل أسلوب Lasso وأسلوبElasticNets. أما الفصل التاسع المرتبط بالتصنيف أيضا فقد ارتبط بتصنيف الموسيقى.

بينما تعرض الفصل الثالث من هذا الكتاب لموضوع العنقدة Clustering المتصل بإيجاد المواقف المرتبطة ويوضح قوة مدخل الكلمات عند تطبيقها لإيجاد مواقف شبيهة بدون تفهم الواقع؛ في حين يبين الفصل الرابع مجال النمذجة Modeling أي التحرك فيما وراء كليف كل موقف بعنقود مفرد وتحديد ذلك لموضوعات عديدة ترتبط بالنص الحقيقي مع موضوعات عديدة. ويتطرق الفصل الثامن لمجال التوصيات Recommendations حيث أن بناء نظم التوصية يتم بناء علي تقديرات منتج المستهلك، إلي جانب كيفية بناء التوصيات من بيانات التسويق دون الحاجة لبيانات التقديرات التي لا يوفرها المستخدمون.

وفي الفصل العاشر من هذا العمل، استعرض مجال الرؤية الكمبيوترية Computer Vision الذي يطبق التصنيف في سياق تداول الأشكال المعينة بواسطة استخلاص الأوجه المميزة لها من البيانات؛ بينما استعرض الفصل الحادي عشر مدي تقليل البعدية Dimensionality Reduction التي تتعلق بالطرق الأخرى المتواجدة وممكنها أن تساعد في تقليل حجم البيانات للتعامل معها بواسطة لوغاريتمات تعلم الآلة؛ وارتبط الفصل الثاني عشر الأخير ببيانات أكبر Bigger Data واكتشاف بعض المداخل التي تتعامل مع بيانات أكبر من خلال ميزة الأجزاء المركزية المتعددة أو العناقيد الكمبيوترية، إلي جان الاشتمال علي مقدمة لاستخدام الحوسبة السحابية Cloud Computing التي تستخدم خدمات شركة Amazon المتاحة علي الويب وتعتبر مقدم خدمة الحوسبة.

أما ملحق الكتاب النهائي فيرتبط بالإجابة عن التساؤل أين يمكن للمستخدم اتقان تعليم «تعلم الآلة» أكثر؟ حيث حدد لذلك قائمتين، إحداهما عن المقررات أو الدورات الدراسية المتاحة على الخط، والكتب المنشورة عن الموضوع.