عرض ملخص كتب صادرة حديثاً

أ.د. محمد محمد الهادي

عمليات تشغيل تعلم الآلة لتفعيل علم البيانات

عمليات تشغيل تعلم الألة لتفعيل علم البيانات: أربـع خطوات لتحقيق قيمة علم البيانات خلال عمليات النموذج.

تأليف ديفيد سوينر وآخرون

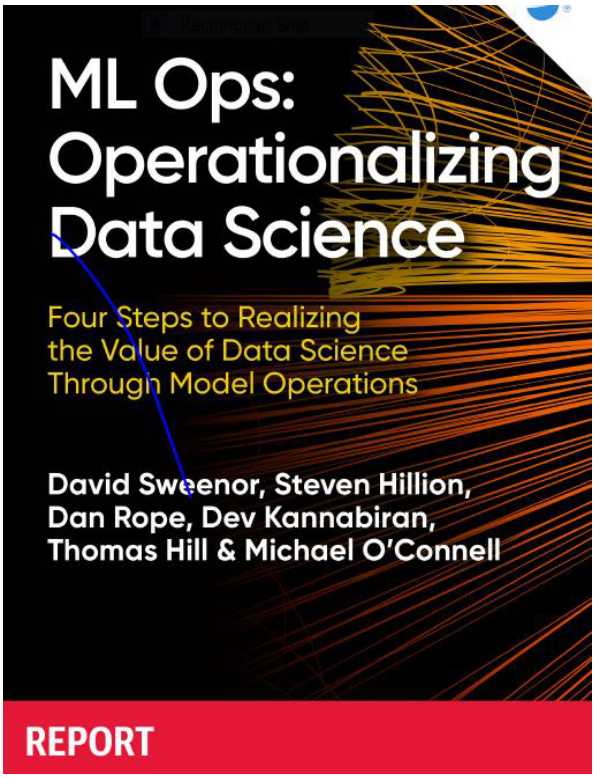
الناشر: أو رايلي

تاريخ النشر: ٢٠٢٠

عدد الصفحات: ٣٨ صفحة

تضمن هذا الكتيب الذي يتعلق بعمليات تشغيل تعلم الآلة لتفعيل وتحقيق قيمة علم البيانات ، أهمية تواجد أربع خطوات أساسية ترتبط بتحقيق ذلك التي تتمثل في بناء النموذج المحتاج إليه؛ إدارة هذا النموذج بطريقة سليم؛ نشر وتكامل النموذج المصمم؛ ثم أخيرا مراقبة تنفيذ النموذج واختباره.

كما استعرض هذا العمل دراستي حالة: الأولي ترتبط بتفعيل علم البيانات في الصناعة التحويلة عبر نماذج التؤم الرقمية؛ والحالة الثانية تستعرض تفعيل علم البيانات في صناعة التأمين التي تتعلق بنماذج التسعير الديناميكية.



استعرض هـذا الكتيب أهمية نشر النماذج وخاصة النماذج التنبؤية ، حيث يكم نضكين عدد من التحويلات وقواعد الأعمال الممكن الحصول عليها فيما يتعلق بطريقة تفعيل علم البيانات باستخدام عمليات تشغيل تعلم الآلة. وبذلك يتضمن النموذج المصمم لعمل التنبؤات عدة عناصر أو متغيرات أساسية ترتبط بالمجال المحدد المطلوب تطويره، وتتواجد هذه المتغيرات في نطاق البيانات الخام المدخلة بدون أي حسابات إضافية أو نقص فيها، إلا أنها مرتبطة بإمكانية تفعيل هذا النموذج المطور من خلال مجموع عناصره ، مكوناته، متغيراته بالإضافة للتحويلات المتطلبة له. التي تدار كلها كحقيقة اصطناعية من خلال خط تدفق محدد أثناء دورة حياة النموذج.

وقد استعرض هذا الكتيب أربع خطوات اسايية يجب القيام بها لتحقيق قيمة علم البيانات خلال عمليات النموذج. وهذه الخطوات تتمثل في التالي:

- بناء النموذج: وفي هذه الخطوة تستخدم لغات برمجة مثل لغتي Python, R بالإضافة ألي التطبيقات التجارية المتاحة بالفعل من أجلإنشاء خط تدفق التحليلات، كما يستخدم معيار خوارزميات تعلم الآلة لبناء النماذج التنبؤية وأداء هندسة الأوجه المميزة المؤدية لإنشاء التحليلات الأكثر احتماليا في تقرير نموذج تعلم الآلة القوي.
- إدارة النموذج المصمم: من الملاحظ أن النماذج تشتمل علي دورة حياة الموذج المدارة أحسن من مستودع مـركـزي تتواجد فيه الأصول، تعيين الإصدار، التطابق مع القواعد والبيانات المنظمة لذلك، التوافق مع الإنتاج المستهدف، الاختبار، النشر والاستبدال النهائيالتي تتابع ، إضافة إلي واصفات البيانات (ما وراء البيانات) المرتبطة بحقائق النموذج ألالاصطناعي Artifacts، إدارة المنصة والمستودع التي يجب ان تتبع مقاييس الدقة إلي جانب الوصلات بين النماذج ومجموعات البيانات.
- النشر والتكامل: في هذه الخطوة يؤخذ خط تدفق علم البيانات من بيئة التطوير الأصلية ويعبر عنه في تطبيقات الأعمال. علي سبيل المثال، النموذج المطور المستخدم لغة برمجة Python قد يحتاج لأن يحول كود لغة التساؤل المهيكلة SQL للتنفيذ في قاعدة بيانات المؤسسة، أو قد ينشئ آلية لتصدير تفقات عمل علم البيانات كقود لغة جافا Java مثلا الذي عندئذ أن يصبح ضمنيا في كود التطبيق المعين. كما أن معماريات عمليات تشغيل تعلم الآلة قد تساند أيضا نشر حقائق النموذج الاصطناعية لبيئات السحابة العامة أو الخاصة مما يمكن من وقت أسرع للإنتاج وتسجيا الطلبات . وعلي ذلك، توجد حاجة من أجل التمكن في نشر خط التدفق في الشكل واللغة الملائمة لهدف بيئة التشغيل المتوافق مع متطلبات أعمال المنظمة.
- المراقبة: بعد نشر النموذج يجب مراقبته من أجل دقة تنبؤاته بالإضافة لسلامة النموذج ومعايير الأعمال الهامة الأخري بغية تقديم رؤية كاملة في تأثيره العام علي الأعمال. ويمكن عمل المراقبة عبر أدواتذكاء الأعمال BI المرتبطة بالتنبؤات ومداخل النموذج، أو أن تأخذ شكل تقارير جدولة نظامية علي مؤشرات أداء رئيسية (KPIs)، وعلي ذلك في الإمكان تحسين الدقة ومعايير الأداء الرئيسية إما يدويا خلال التكرار العرضي بإدخال من الخبير المسؤل البشري، او آليا خلال إعادة التدريب الجاري وحلقات التحدي بواسطة الإنسان علي سبيل المثال.

ولتحقيق هذه الخطوات الأربعة تم استعراض الموضوعات التاية: اعتبارات البيانات من حيث الهياكل والـوصـول؛ تحديد أبعاد هندسة الأوجه المتطلبة؛ إختبار النموذج المصمم؛ الخروج في مقدمة انتشار النموذج؛ تحرير وموافقات ورقابة الإصـدارة؛ إعـادة استخدام وإعـادة تحديد غرض النموذج من مستودع البيانات المدار مركزيا؛ النشر والتكامل وأين يتقابلا؛ تطوير تطبيق الأعمال . وفي إطار المراقبة تحديد القياسات الاحصائية؛ قياسات الأداء؛ قياسات الأعمال والعائد من الاستثمار؛ متي تنحرف النماذج عن مسارها؛ مع إعادة التدريب وإعادة النمذجة؛ زمراقبة آلية اللقاءات.

كما تم استعراض دراستي حالة: الحاة الأولي اختصت بتفعيل علم البانات في الصناعة التحويلية من خلال نماذج التؤم الرقمية؛ أما دراسة الحالة الثانية فقد تعلقت بتفعيل علم البيانات في صناعة التأمين من خلال نماذج التسعير الديناميكية.