نماذج القرار بالذكاء الاصطناعي تدخل أدوات مراكز العمليات الأمنية: ما الذي ينبغي أن يدرسه قادة الأمن



ظهرت فئة جديدة من نماذج «القرار» بالذكاء الاصطناعي خلال أسابيع، مع إصدارات من عدة جهات، وهو ما يؤثر في طريقة بناء منصات مراكز العمليات الأمنية وتقييمها.

- أصدرت TypeSafe AI نموذج Jev في 15 سبتمبر ووصفته بأنه نموذج «النظام الأول»، في إشارة إلى التفكير السريع والحدسي.
- أعلنت OpenAI عن واجهة Decisions API في DevDay وفتحت نسخة تجريبية عامة في 6 أكتوبر.
- نشرت Cribl النموذج cribl-decision، وهو محوّل LoRA مفتوح يمكن تشغيله محليًا.
- يذكر المصدر أن هذه الفئة انتقلت من العدم إلى الانتشار خلال نحو شهر.
- المقتطف المتاح جزئي ولا يتناول الأداء أو التسعير أو التقييمات الأمنية.

ما الذي حدث

وفق موقع The Hacker News، انتقلت فئة جديدة من نماذج الذكاء الاصطناعي المعنية باتخاذ القرار من العدم إلى حضور واسع خلال نحو شهر. فقد أصدرت TypeSafe AI نموذج Jev في 15 سبتمبر ووصفته بأنه نموذج «النظام الأول»، في إشارة إلى فكرة كانمان عن التفكير السريع والحدسي. وأعلنت OpenAI عن واجهة Decisions API في DevDay وفتحت نسخة تجريبية عامة في 6 أكتوبر. ونشرت Cribl النموذج cribl-decision، وهو محوّل LoRA مفتوح يمكن للمؤسسات تشغيله محليًا. ويشير المصدر إلى دخول جهات أخرى في هذا المجال.

لماذا يهم قادة مراكز العمليات الأمنية

يربط عنوان المصدر بين هذه النماذج وتصميم منصات مراكز العمليات الأمنية المدعومة بالذكاء الاصطناعي. ولا يتضمن المقتطف المتاح بيانات عن الأداء أو الدقة أو التكلفة أو نتائج أمنية محددة، لذا ينبغي اعتبار أي ادعاءات حول القدرات غير مثبتة إلى حين اختبارها. والواضح أن الخيارات تمتد اليوم من واجهات مستضافة إلى نماذج يمكن تشغيلها محليًا، وهو ما يؤثر في معالجة البيانات ومخاطر الأطراف الثالثة والتحكم في كيفية اتخاذ قرارات الفرز الآلي.

التوجه العملي

لا حاجة إلى الاستعجال. يمكن للجهات التي تدرس الفرز المدعوم بالذكاء الاصطناعي أن تستثمر هذه المرحلة لوضع معايير التقييم قبل أن يطرق الموردون أبوابها، بدلاً من اعتماد أداة بدافع زخم السوق. فنماذج النشر المستضافة والمحلية تثير أسئلة مختلفة عن مكان معالجة بيانات القياس الأمني ومن يستطيع مراجعة القرارات الناتجة.

ما يجب فعله الآن

1. حدّدوا معايير تقييم للفرز المدعوم بالذكاء الاصطناعي، تشمل اختبار الدقة على تنبيهاتكم التاريخية، قبل أي تعاقد مع مورد.
2. قارنوا بين الواجهات المستضافة والنماذج المشغّلة محليًا من حيث مكان معالجة بيانات القياس الأمني وتخزينها.
3. تحققوا من اتساق استخدام خدمات ذكاء اصطناعي خارجية مع بيانات المركز لمتطلبات نظام حماية البيانات الشخصية (PDPL) وقواعد تصنيف البيانات لديكم.
4. أبقوا مراجعة بشرية للقرارات الآلية المؤثرة في الاحتواء أو التصعيد إلى أن يثبت الأداء.
5. أدرجوا مزوّدي نماذج الذكاء الاصطناعي ضمن تقييم مخاطر الأطراف الثالثة.