تقرير بناء وتشغيل نظام كاشف للاحتيال في النصوص العربية

مقدمة وهدف المشروع

تم في هذا المشروع إنجاز دورة حياة كاملة لتطوير نظام ذكاء اصطناعي، بهدف بناء وتشغيل مصنف نصوص قادر على التمييز بدقة وموثوقية بين الرسائل النصية الاحتيالية والآمنة باللغة العربية واللهجة السعودية. تجاوز الهدف مجرد تدريب نموذج أولي، ليشمل بناء نظام متكامل يبدأ من هندسة البيانات المتقدمة، مرورًا بتطبيق تقنيات تدريب احترافية، وصولًا إلى نشر النموذج كخدمة (API) جاهزة للتكامل مع تطبيقات الويب.

المنهجية المتبعة والإنجازات التقنية

اعتمد أساس المشروع على بناء مجموعة بيانات محصّنة وذكية. بدلاً من الاعتماد على البيانات الأولية فقط، تم إثراؤها بشكل استراتيجي بأمثلة "صعبة" لرفع قدرة النموذج على التعميم، مثل إضافة رسائل آمنة تحتوي على روابط وعروض حقيقية (Hard Hams)، ورسائل احتيالية خفية تعتمد على الهندسة الاجتماعية بدون روابط (Subtle Spam).

تم بعد ذلك تدريب وتقييم نموذج لغوي متقدم من عائلة BERT، وتحديدًا asafaya/bert-mini-arabic، لقدرته على فهم السياق بفعالية وكفاءة. لمكافحة مشكلة "فرط التخصيص" (Overfitting)، تم تطبيق آلية **"الإيقاف المبكر" (Early Stopping)** التي تقوم بمراقبة أداء النموذج على مجموعة تحقق (Validation Set) وحفظ أفضل نسخة منه فقط، مما يضمن أن النموذج تعلم "التعميم" بدلاً من "الحفظ". ولضمان المصداقية العلمية للنتائج، تم إجراء فحص برمجي للتحقق من **عدم وجود أي تسرب للبيانات (Data Leakage)** بين مجموعات التدريب والاختبار، لتأكيد أن النتائج النهائية التي تم تحقيقها حقيقية وغير مضللة.

التحديات الهندسية والحلول المطبقة

التحدي الأبرز الذي واجه المشروع كان **"هشاشة" النموذج الأولية (Model Brittleness)**. على الرغم من دقته العالية في بيئة الاختبار، فقد فشل النموذج في التعرف على أساليب احتيال جديدة ومختلفة عند اختباره عمليًا (Adversarial Testing). على سبيل المثال، فشل في تصنيف رسالة واضحة عن "اختراق جهاز" لأن هذا السياق كان غائبًا عن بيانات التدريب الأولية.

تم حل هذه المشكلة الجذرية عبر اعتماد **"دورة التحسين المتكررة" (Iterative Improvement Cycle)**. تتلخص هذه المنهجية في اختبار النموذج بشكل خصومي لتحديد نقاط ضعفه، ثم جمع هذه الحالات الصعبة وإضافتها إلى مجموعة التدريب، وإعادة تدريب النموذج. هذه الدورة حولت النموذج من كونه "هشًا" إلى كونه "صلبًا" (Robust)، وقادرًا على مواكبة الأساليب المتغيرة.

بالإضافة إلى ذلك، تم حل سلسلة من التحديات المتعلقة بإعداد البيئة المحلية على نظام ويندوز، شملت مشاكل توافق إصدارات بايثون، وتثبيت المكتبات التي تتطلب بناءً برمجيًا، وإعداد وتشغيل الخوادم الخلفية والأمامية بشكل متزامن.

الخلاصة والنتيجة النهائية

المنتج النهائي لهذا المشروع هو نظام متكامل وموثوق، يبدأ من مجموعة بيانات قوية، مرورًا بخط أنابيب تدريب احترافي، وصولًا إلى نموذج مُدرَّب بدقة تم نشره كخدمة (API) باستخدام إطار العمل **FastAPI**. تم إنجاز عملية نشر محلية كاملة، حيث تم ربط الخادم الخلفي بنجاح مع واجهة أمامية تفاعلية مبنية بـ **React**، مما يثبت جاهزية المشروع للتطبيق العملي. ويوصى كخطوات مستقبلية بالاستمرار في إثراء البيانات بشكل دوري، وتجربة نماذج أخرى للمقارنة، وإضافة ميزات تحليلية متقدمة للواجهة.