## ملخص البحــث رقـــم ( 7)

**السيد الأستاذ الدكتور/ مقرر اللجنة العلمية الدائمة لترقية الأساتذة والأساتذة المساعدين للحاسبات والمعلومات**

**تحية طيبة وبعد –احيط سيادتكم علما بان البحث رقم 7 بياناتة كالتالي**

**عنوان البـــــحث باللغة الانجليزية:**

**Self-Tuning Parameters for Decision Tree Algorithm Based on Big Data Analytics**

**عنوان البـــــحث باللغة العربية**

**معلماتالضبط الذاتي لخوارزمية شجرة القراربناء ً على تحليلات البيانات الضخمة**

**اسماء المؤلفين:**

**Manar Mohamed Hafez1, *, Essam Eldin F. Elfakharany1, Amr A. Abohany2 and Mostafa Thabet3**

**مكان النشر و تاريخه:**

**Information Sciences Letters An International Journal , November 2022**

**ملخص البـــــحث باللغة العربية :**

يهتم العديد من الباحثين والممارسين في هندسة النظم الصناعية وعلم التحكم الآلي "بالبيانات الضخمة". تختلف البيانات الضخمة عن البيانات العادية. البيانات الضخمة ، على سبيل المثال ، غالبًا ما تكون غير منظمة ويجب تحليلها في الوقت الفعلي. حظيت تحليلات البيانات الضخمة باهتمام كبير من رجال الأعمال والأوساط الأكاديمية بسبب مزاياها الهائلة في خفض التكاليف وصنع القرار. على سبيل المثال ، يمكن أن يوفر معلومات أساسية لمجموعة واسعة من الشركات. في هذا البحث ، قمنا بتطوير خوارزمية شجرة القرار (DT) ، وتحديدًا أشجار التصنيف والانحدار ، للحصول على نتائج أفضل. قمنا بتصميم نسخة معدلة من خوارزمية DT. تهدف الشجرة إلى تحقيق العمق الأمثل عن طريق الضبط الذاتي لمعلمات التشغيل وتحسين الأداء. علاوة على ذلك ، تحققنا من كفاءة وفعالية DT المعدل باستخدام مجموعتي بيانات: مجموعات بيانات المطار والحريق. تحتوي مجموعات بيانات المطار والحريق على 500000 و 600000 مثيل على التوالي. قمنا أيضًا بمقارنة خوارزميات DT المعدلة والقياسية.

أظهرت النتائج أن خوارزمية DT المعدلة تعمل بشكل أفضل من خوارزمية DT القياسية. تم إجراء هذه المقارنة على عقدة متعددة على أداة Apache Spark باستخدام خدمات الويب من Amazon. علاوة على ذلك ، حققنا دقة 6.85٪ و 8.85٪ لمجموعات بيانات الحريق والمطار على التوالي.