وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم الإحصاء والمعلوماتية

تحسين تنبؤ بيانات السلاسل الزمنية الضخمة لإسعار المعادن

رسالة مقدمة
إلى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة ماجستير علوم في الإحصاء

من قبل

ابتهال منيب يحيى ذنون

بإشراف

الأستاذ الدكتور

زكريا يحيى نوري يحيى

1446هـ 2024م

المستخلص

شهدت العقود الأخيرة تزايدا في كمية ونوع البيانات وأصبحت العديد من التطبيقات مصدراً لتدفق هذه البيانات، إذ إن الزيادة والتراكم في حجم البيانات يحتاج إلى ابتكار طرائق وأساليب لتلخيص ودراسة وتنقيب هذه البيانات من اجل فهمها والاستفادة منها، إذ إن علم تنقيب البيانات والبحث عن المعرفة من العلوم الحديثة التي مازالت في حالة تطوير مستمر، خصوصا بعد ظهور العديد من الخوارزميات الذكائية (Intelligent Algorithms) التي أثرت بشكل ملحوظ في علم مجال التنقيب عن المعرفة ومعالجة البيانات. يعد موضوع التنبؤ ببيانات السلاسل الزمنية من أكثر المواضيع تطرقا واستخداما في العديد من التطبيقات العلمية. وتتميز أية سلسلة زمنية بأن بياناتها مرتبة بالنسبة للزمن، وأن المشاهدات المتتالية غالباً ما تكون غير مستقلة أي تعتمد على بعضها البعض. فدراسة ظاهرة ما على مدة زمنية معينة يمكننا من التعرف على نمطها واستخدامها للتنبؤ بالقيم المستقبلية للظاهرة ، والذي يساعد في وضع الخطط السليمة. تعّد آلة متجه الانحدار الداعم من الطرائق المهمة وواسعة الاستخدام في التنبؤ، ومع ذلك فإن هذه الطريقة تفقد دقتها عندما تكون البيانات ضخمة وتحتاج إلى اختيار مسبق لمعلمات الدوال اللبية (دالة النواة) المستخدمة. ولغرض تحقيق الهدف وُظفت إحدى الخوارزميات المستوحاة من الطبيعة وهي خوارزمية ابن آوى الذهبي في تقدير قيم معلمات الدالة اللبية المستخدمة في آلة متجه الانحدار الداعم ، ومن ثمّ الحصول على دقة تنبؤ أعلى. طُبّق توظيف الطريقة المقترحة على بيانات حقيقية عالمية في مجال الاقتصاد ، إذ شملت التنبؤ بالأسعار اليومية لبعض المعادن ومنها معدن الفضة والذهب والنحاس ومعدن النفط برنت. شملت الدراسة الفترة الزمنية الممتدة من 2000\1\1 ولغاية 2023\11\9. وبالاعتماد على عدد من معايير المقارنة ، فقد بينت النتائج تفوق الطريقة المقترحة على باقي الطرائق الأخرى التقليدية لحصولها على أقل القيم لمعايير خطأ التنبؤ وأعلى القيم لمعايير دقة التنبؤ ولكافة أسعار المعادن المستخدمة.

Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Statistics and Informatics

Improving Big Time Series Data Forecasting for Metal Prices

**A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of Master of Science**

**in
Statistics**

By

Ibtihal Muneeb Yahya

Supervised by

Professor

DR. Zakariya Yahya Algamal

2024 A.D. **1446 A.H.**

ABSTRAC

The last decades have witnessed an increase in the quantity and type of data، and many applications have become a source of the flow of this data، as the increase and accumulation in the volume of data requires the creation of methods to summarize، study and mine this data in order to understand and benefit from it، as the science of data mining and the search for knowledge is one of the Modern science is still in a state of continuous development، especially after the emergence of many intelligent algorithms that have significantly influenced the science of knowledge mining and data processing. The topic of forecasting time series data is one of the most discussed and used topics in many scientific applications. Any time series is characterized by the fact that its data are arranged in relation to time and that successive observations are often not independent، that is، dependent on each other. By studying a phenomenon over a specific period of time، we can identify its pattern and use it to predict the future values of the phenomenon، which helps in making sound plans. Support vector regression is an important and widely used method in forecasting. However، this method loses its accuracy when the data is large and requires prior selection of the parameters of the kernel functions used. In order to achieve the goal، one of the algorithms inspired by nature، which is the Golden Jackal algorithm، was used to estimate the values of the parameters of the pulp function used in the support vector regression machine، thus obtaining a higher prediction accuracy. The proposed method was applied to real global data in the field of economics، as it included forecasting the daily prices of some metals، including silver، gold، copper، and Brent crude oil. The study included the time period extending from January 1، 2000 to November 9، 2023. Relying on a number of comparison criteria، the results showed that the proposed

method is superior to other traditional methods because it obtained the lowest values for the prediction error criteria and the highest values for the prediction accuracy criteria and for all metal prices used.