وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات

استخدام الطريقة الهجينة ARIMA-RNN للتكهن ببيانات التلوث الجوي

آلاء زهير مصطفى الخشاب

رسالة دبلوم عالٍ
الإحصاء والمعلوماتية

بإشراف
د.أسامة بشير شكر الحنون

١٤٣٩ هـ ٢٠١٨ م

المستخلص

ان للتلوث الجوي دوراً سلبياً كبيراً يؤدي الى مشكلات كثيرة على البيئة وعلى صحة الإنسان لذا فان من الضروري دراسته والتكهن له للحد او للتقليل من اضراره او السيطرة عليه، تستخدم قياسات جسيمات المادة (~~PM~~$_{10}$PM$_{10}$) Particular Matter في الهواء لبيان تلوث الهواء او جودته. ان المشكلة الرئيسة تكمن في غير خطية هذا النوع من البيانات وفوضويتها. ومن اشهر نماذج السلاسل الزمنية استخداما هو نموذج الانحدار الذاتي والمتوسطات المتحركة Autoregressive and Moving Average model Integrated (ARIMA) كنموذج تقليدي شائع ولكن هذا النموذج لا يمكن التعامل مع البيانات غير الخطية فيُظهر تكهنات غير دقيقة. في هذه البحث تم استخدام بيانات ~~PM~~$_{10}$PM$_{10}$ والتي استحصلت للفترة من 1 كانون الثاني لسنة ٢٠١٣ الى ٣١ تشرين الأول لسنة ٢٠١٥ في ماليزيا. يهدف هذه البحث الى ايجاد انموذج يتعامل مع البيانات غير الخطية ويعطي بنسبة اقل خطا مقارنة بالنماذج التقليدية و اقتراح نموذج يوافق بين الشبكات العصبية المعاودة Recurrent Neural Network (RNN) و نموذج ARIMA عن طريق الاعتماد على هيكلية مدخلاته في حالتين مختلفتين اولهما يراعي الاخذ بقيم معلمات النموذج والاخرى بدونها. باستخدام الطرائق المقترحة تم الحصول على خطا اقل ومن ثم نتائج اكثر دقة وكما اعطت مدخلات ARIMA مع قيم المعلمات نتائج افضل وكذلك فان خوارزمية الضبط البيزي للتدريب(BR) Bayesian regularization اعطت نتائج ادق من خوارزمية Levenberg-Marquardt (LM) في مرحلة التدريب. من ذلك نستنتج بأن الطريقة المقترحة RNN بالاعتماد على ARIMA هي طريقة جيدة وتعطي نتائج اكثر دقة من نتائج النماذج التقليدية.

نسّقه:الخط: 9 نقطة، خط اللغة العربية وغيرها: 9 نقطة

University of Mosul

COLLEGE OF COMPUTERS SCIENCES

AND MATHEMATICS

Using Hybrid method ARIMA-RNN for Air Pollution Forecasting

Alaa' Zuhair Mostafa Alkhashaab

Higher Diploma Thesis

Statistics

Supervised By

Dr. Osamah Basheer Shukur

2018 A.D. 1439 A.H

Abstract

Air pollution has negative role leads to many problems for the environment and human health, therefore the study and forecasting of this subject is necessary to control and reduce the damages. The measurements of the particular matter (~~PM_{10}~~PM_{10}) can be used to present the air pollution and its quality. The main problem in this type of dataset is the nonlinearity pattern and noise. Autoregressive and Moving Average model Integrated (ARIMA) is the most common time series model, but it can't deal with a problem of nonlinearity and its result will be in accurate. In this study, PM_{10} data from Malaysia (1^{st}/Jan/2013– 31^{st}/Nov/2015) has been used and studied. The aim of study is to find the suitable model can deal with the nonlinear dataset and result more minimum forecasting error comparing to other classical models. ARIMA and recurrent neural network (RNN) will be combined to compose the hybrid method by depending on the structure of the ARIMA inputs (with and without the parameters). Using this proposed method gives more minimum forecasting error and more accurate results. Depending on ARIMA inputs with parameters gives more accurate results and Bayesian regularization (BR) training algorithm gives more accurate results than Levenberg-Marquardt (LM) training algorithm within training process. As conclusion, the proposed method is better than the classical methods and will be gives more accurate results.

نسّقه:الخط: 9 نقطة، خط اللغة العربية وغيرها: 9 نقطة

نسّقه:الخط: 9 نقطة، خط اللغة العربية وغيرها: 9 نقطة