



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

تصنيف السلاسل الزمنية بالطرائق الإحصائية والشبكات العصبية العميقة

رسالة مقدمة

الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة ماجستير علوم في الإحصاء

من قبل

عمر أكرم محمد سعيد بكر

بإشراف

أ.م.د. أسامة بشير شكر الحنون

المستخلص

في هذه الرسالة تم تصنيف السلاسل الزمنية باستخدام الانحدار اللوجستي Logistic Regression (LR) وشبكة المدرك العصبية Perceptron Neural Network (PNN) وآلة المتجه الداعم Support Vector Machine (SVM) و K الجار الأقرب K-nearest neighbor (KNN) وطريقة النواة الكاوسية Gaussian Kernel Method (GK). وتمت الاستفادة من متغيرات الانحدار الذاتي Auto Regressive (AR) وذلك من خلال رتبة أنموذج الانحدار الذاتي $AR(p)$ في تحديد متغيرات الإدخال للطرائق المستخدمة. ولغرض المقارنة وزيادة الدقة في التصنيف تم استخدام الشبكة العصبية الالتفافية Convolutional Neural Network (CNN) بوصفها إحدى تقانات التعلم العميق الخاصة في التصنيف الصوري، ولغرض انشاء متغير الإدخال لطريقة CNN حُوِّلت بيانات السلاسل الزمنية المستخدمة من صيغتها الرقمية وشُكِّلت صوراً ثنائية البعد. وقد استخدمت في هذه الدراسة بيانات السلسلة الزمنية لدودة الربداء الرشيقية Caenorhabditis Elegans (CE) بزوايا حركتها الموجية التي ستمثل حالة الدراسة. إن البيانات المستخدمة في هذه الرسالة هي من السلاسل الزمنية الطويلة وبعدها مشاهدات كبير جدا فضلا عن تعدد الخصائص والصفات التي تمر بها السلسلة الزمنية قد يكون سببا في عدم تجانس البيانات الذي قد يؤدي الى تصنيفات لا ترقى الى ان تكون دقيقة. ان استخدام طريقة CNN ستؤدي الى نتائج دقيقة؛ وذلك لانها ملائمة للتعامل مع بيانات الدراسة لحل مشكلات عدم التجانس من خلال التمثيل الصوري للبيانات الرقمية. ومن خلال مقارنة النتائج بين الطرائق المستخدمة تبين ان طريقة CNN يتفوق على بقية الطرائق الأخرى المستخدمة. ولذلك فمن الممكن استنتاج ان استخدام طريقة CNN التي تعتمد التصنيف الصوري الثنائي تؤدي الى نتائج تصنيفية دقيقة مقارنة بالطرائق الأخرى التي تعتمد التصنيف الرقمي الثنائي.

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



Time Series Classification based on Statistical Methods and Deep Neural Network

**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
Omar Akram Mohammed Saeed**

**Supervised by
Assist. Prof. Dr. Osamah Basheer Hannon**

2024 A.D.

1445 A.H.

Abstract

In this thesis, time series were classified using Logistic Regression (LR), Perceptron Neural Network (PNN), Support Vector Machine (SVM), K-nearest neighbor (KNN), and Gaussian Kernel Method (GK). Autoregressive (AR) variables were utilized through the rank of the AR (p) model in determining the input variables for the methods used. For the purpose of comparison and increasing accuracy in classification, the Convolutional Neural Network (CNN) was used as one of the special deep learning techniques in image classification, and for the purpose of creating the input variable for the CNN method, the time series data used were converted from their digital form and formed two-dimensional images. This study used time series data for the worm *Caenorhabditis elegans* (CE) with its wave movement angles, which will represent the case of the study. The data used in this thesis is from long time series with a very large number of observations, in addition to the multiplicity of characteristics and characteristics that the time series experiences, which may be a reason for the heterogeneity of the data, which may lead to classifications that are not accurate. Using the CNN method will lead to accurate results; This is because it is suitable for dealing with study data to solve heterogeneity problems through the visual representation of digital data. By comparing the results between the methods used, it was found that the CNN method outperforms the rest of the other methods used. Therefore, it is possible to conclude that using the CNN method that relies on binary image classification leads to accurate classification results compared to other methods that rely on binary digital classification.