جامعة الموصل
كلية علوم الحاسوب والرياضيات

اختيار الميزة باستخدام طريقة مقترحة بين امثلة اسراب الجسيمات الثنائية وتقنية المعلومات المتبادلة الضبابية

مصطفى ايهم عبد الحافظ الزرري

رسالة ماجستير

الرياضيات/الحاسوبية

بإشراف

الأستاذ المساعد

د. عمر صابر قاسم

1441هـ 2019م

المستخلص

إن الكم الهائل من المعلومات المستخلصة من التطبيقية المختلفة (خصوصا التطبيقات الطبية الجزيئية) يحتاج الى عمليات حسابية معقدة بسبب حجم البيانات، لذا فكّر العلماء بطريقة استخلاص البيانات المهمة والاقتصار عليها في المعالجة من خلال مفاهيم اختيار الميزة Feature selection. إذ تعُّد مسألة اختيار الميزات من أهم المسائل المستخدمة في تحسين عملية تصنيف البيانات، واذ يوجد العديد من الخوارزميات التطورية Evolutionary algorithms التي تستخدم لهذا الغرض، مثل خوارزمية تحسين أسراب الجسيمات Particle swarm optimization (PSO)، بعد تحويلها من الفضاء المستمر Continuous إلى الفضاء المتقطع Discrete من خلال مفهوم Binary PSO. إذ تستمد خوارزمية تحسين سرب الجسيمات الثنائية BPSO آلية عملها من خوارزمية PSO الاعتيادية ولكن في مساحة بحث متقطعة ثنائية. في هذه الدراسة، تم اقتراح طريقة مهجنة بين خوارزمية BPSO مع تقنية المعلومات المتبادلة الضبابية Fuzzy mutual information (FMI) للحصول على مجموعات فرعية من الميزات من خلال مرحلتين اساسيتين: تتمثل الأولى باستخدام تقنية المعلومات المتبادلة الضبابية FMI والتي يتم من خلالها ترتيب الميزات بشكل تنازلي وفق لأهميتها وتأثيرها في التصنيف، أمَّا المرحلة الثانية، فيتم من خلالها تحديد آخر للميزات التي تم اختيارها من المرحلة الأولى وذلك من خلال استخدام خوارزمية BPSOالتي تعتمد على دالة لياقة Fitness function مقترحة. وقد اثبتت نتائج الخوارزمية المقترحة FMI-BPSO الكفاءة والفعالية من خلال الاقتصار على ميزات محددة والحصول على دقة تصنيف أعلى مقارنة بالطريقة الاعتيادية BPSO.

UNIVERSITY OF MOSUL
COLLEGE OF COMPUTER SCIENCES
AND MATHEMATICS

Feature Selection Using a Proposed Method between Binary Particle Swarm Optimization and Fuzzy Information Techniques

Mustafa Ayham Abdulhafedh

M.Sc./Thesis

Mathematics/Computational

Supervised by

Assistant Professor

Dr. Omar Saber Qasim

2019 A.D. **١٤٤٠ A.H.**

Abstract

The vast amount of information derived from the various applications (especially molecular medical applications) needs complex calculations because of the size of the data, so scientists have thought about how to extract important data and concluded to processing through the concepts of feature selection. The selection of features is one of the most important issues in improving data classification. There are many evolutionary algorithms used for this purpose, such as the Particle Swarm Optimization (PSO) algorithm after conversion from continuous space to discrete space through the PSO Binary concept. The BPSO optimization algorithm derives its mechanism from the default PSO algorithm but in a binary discrete search area. In this study, a hybrid approach was proposed between the BPSO algorithm and the Fuzzy Mutual Information (FMI) algorithm to obtain subsets of features through two basic phases: the first is the use of FMI, in which the features are arranged in descending order according to their importance and the second phase, in which the other features selected from the first stage are determined by using the BPSO algorithm, which is based on a suggested fitness function. The results of the proposed FMI-BPSO algorithm have proven efficiency and effectiveness by limiting specific features and obtaining higher classification accuracy than the default BPSO method.