مقارنة بين الطرق المعلمية واللامعلمية لتأثير العلاج بالحجامة الرطبة

اعداد

هاجر صالح الجهني

باشراف

د.لمياء باحارث

المستخلص

تعتبر عمليه Data miningمن اهم أدوات هذا العصر، وذلك لأنها تهدف الى تحليل البيانات الضخمة وتلخيصها الى معلومات مفيدة. يعتمد نظام Data mining على النماذج الإحصائية في الحصول على المعلومات المفيدة، حيث ان من احدى مهام هذا النظام هو ملاحظة المعلومات الغير مرئية والغير معروفة في البيانات الضخمة لإيجاد علاقة منطقيه، وبالتالي تلخيصها الى معلومات مفيدة ومفهومة يمكن ان تستخدم كقاعدة بيانات منظمة لإجراء العديد من الأبحاث بفضلها. أيضا يعد التصنيف باستخدام النماذج الإحصائية من الأساليب التي تفسر البيانات بالإضافة الى قدرتها على التنبؤ. في الأبحاث الطبية، يعد تفسير البيانات واجراء التحليل الدقيق لها بالإضافة الى التنبؤ بالتشخيص من اهم الإجراءات للدراسة. قيمت هذه الأطروحة إمكانية استخدام كل من النماذج المعلمية واللامعلمية في التصنيف والتنبؤ، حيث ان الهدف الرئيسي من هذه الدراسة هو التحقق من تأثير العلاج بالحجامة الرطبة على المرضى باستخدام الشبكات العصبية الاصطناعية وشجرة القرار والتحليل التمييزي ونماذج الانحدار اللوجستي. بالإضافة الى دراسة مستوى التحسن في الصحة، وعلى وجه الخصوص في المرضى الذين تلقوا علاج الحجامة الرطبة في عيادات الطب النبوي بمستشفى الملك عبد العزيز الجامعي بين سبتمبر ٢٠١٣ وأغسطس ٢٠١٧. توفر هذه النماذج المقترحة نظام يعطي تفسيرًا دقيقًا وغير مكلف ومناسب للتنبؤ بتأثير الحجامة الرطبة على المريض بالإضافة الى التنبؤ بمستويات التحسن. ستساعد هذه النماذج الممارس الصحي على التنبؤ بتحسن المريض واتخاذ القرار الصحيح باستجابة المريض للعلاج بالحجامة قبل اجرائها. تم بناء هذه النماذج بتقسيم البيانات إلى ٧٠٪ للتدريب و٣٠٪ لاختبار أداء النموذج. تم تقييم أداء كل نموذج باستخدام العديد من اختبارات تقييم دقت التصنيف مثل (Accuracy, Error Rate, Sensitivity, Specificity, False positive rate, False positive rate, Balanced accuracy rate, Discriminant pawer, Youden's index, area under the Receiver Operating Characteristic carve, F-Measure, Matthews Correlation Coefficient test). علاوة على ذلك، تم إجراء

المقارنات بين النماذج لتحديد ايهم أفضل واقل تكلفة على المريض بناءً على حالته الصحية وتشخيص الطبيب.

A Comparison of Parametric and Non Parametric Methods for the Effectiveness of Wet Cupping Therapy

Thesis by
Hajar Saleh Ahmed Aljohani
Supervised by
Lamya Abdulbaset Baharith
Associate Professor

Abstract

Data mining is the process of noticing unknown and unseen useful information, it is aimed to gain beneficial and understandable information from the huge data sets. The classification is a supervised method used to classify the data by learning a part of it as sample then predicts other data to test the performance. The interpretation of data and the analysis are the most important parts in medical researches. This thesis evaluates the possibility of using both parametric and nonparametric methods in classification and prediction. This retrospective study applied on patients who did wet cupping, and objects to investigate the effect of wet cupping therapy (WCT) on some diagnosis using an artificial neural network, decision tree, discriminant analysis and logistic regression model. The proposed models determined the important variables for each data set and employed as a systems that give an accurate, inexpensive and suitable explanation to predict the effect of WCT on the patient additionally to forecast level of improvement. These models will help the practitioner to take the right decision of predicting the patients' response for WCT. The investigated data set that have major impact on predicting the effect of WCT are liver function, the problems of cells that circulate in blood, thyroid function, kidney function, lipid, bone and calcium problems. The models constructed by dividing the data into 70% to train the model and 30% to test their performance using different classification assessments. The results of these models present the importance of WCT on patient's health depending on the diagnosis, where there are significant difference on most of blood tests. The comparisons between models were made to identify a comprehensive evaluation about the models. The study conclude that WCT is an effective treatment in improving the quality of life and reducing pain. The proposed models identify the important features that have a major influence in each data set. These models will help the clinicians to predict the

level of patients' improvement. Any model which proposed could be applied as a supportive tool to make an early decision regarding the response of patient to the WCT. It would be helpful to apply a nonparametric method with a parametric method and identify the performance of each methods to select the optimal one for classification. The results would provide guidelines for the improvement levels by targeting some types of blood tests and select better method in prediction power.