

Data Mining: Applying and comparing the performance of classification algorithms on a clinical dataset.

Our project had 5 main achievements:

1. Applying different well-known prediction algorithms on a clinical dataset obtained from Razan Center for Infertility which are:
 - Naïve Bayes Classifier
 - Random Forest
 - Logistic Regression
 - K-NN algorithm
2. Applying the RIMARC algorithm, which is a supervised (prediction) algorithm that learns a scoring function to rank instances, developed and published as a scientific research and an executable jar file.
3. Compare the performance of RIMARC and the other 4 prediction algorithm regarding the same clinical data using 2 main indicators:
 - Area Under the Roc Curve indicator (AUC).
 - Execution time indicator.
4. Document the results of the previous mentioned experiment in a scientific research under the title of: Applying Prediction Algorithms on ICSI Treatment Related Data for Performance Comparison.
5. Develop a code from the logic of SERA, which is based on the ranking algorithm RIMARC, to estimate the success rate for the couples going through ICSI treatment to help both Razan Center for Infertility and patients with more informed decisions on whether to go through with the treatment or not.

تم تحقيق 5 انجازات رئيسية في المشروع:

1. تطبيق خوارزميات تنبؤ مختلفة معروفة على مجموعة بيانات طبية تم الحصول عليها من مركز رزان للعقم وهي:
 - Naïve Bayes Classifier
 - Random Forest
 - Logistic Regression
 - K-NN algorithm
2. تطبيق خوارزمية (رايمارك) وهي خوارزمية (التعلم التنبؤي) حيث تتعلم توظيف النقاط لترتيب الحالات او المثيلات, تم تطويرها ونشرها كبحث علمي.
3. مقارنة اداء الرايمارك مع خوارزميات التنبؤ الاربعة الاخرى بنفس البيانات الطبية بناءا على استخدام مؤشرين رئيسيين وهما: المساحة تحت المنحنى و وقت التنفيذ.
4. تم توثيق نتائج التجربة المذكورة سابقا في بحث علمي تحت عنوان: تطبيق خوارزميات التنبؤ على البيانات المتعلقة بمعالجة الحقن المجهري (الزراعة) لمقارنة اداءهم.
5. تم تطوير كود باستخدام لغة البايثون بناءا على منطق ال (sera) والذي يعتمد على خوارزمية الترتيب (رايمارك) , وذلك لتقدير معدل نجاح عملية الحقن المجهري (الزراعة) للازواج وذلك يساعد كل من مركز رزان للعقم وايضا المرضى في اتخاذ قرارات افضل فيما اذا كان عليهم المضي قدما في العلاج ام لا.