



علی اکبر زحمتکش، ۱۴۰۴

ارائه روشی به منظور تشخیص و بهینه سازی بیماری دیابت با استفاده از روش های داده کاوی و الگوریتم جست و جوی ممنوعه

این پژوهش با هدف ارائه روشی نوین برای تشخیص و بهینه سازی بیماری دیابت با بهره گیری از روش های داده کاوی و الگوریتم جستجوی ممنوعه انجام شده است. دیابت به عنوان یکی از بیماری های مزمن و شایع جهان، اهمیت فراوانی در سلامت عمومی دارد؛ تشخیص به موقع و دقیق آن نقش اساسی در پیشگیری از عوارض و کاهش هزینه های درمانی ایفا می کند. به همین منظور، ارائه مدل های پیشبینی کارآمد و بهینه برای شناسایی بیماران دیابتی ضروری است.

روش پژوهش مبتنی بر استفاده از داده های واقعی مراجعه کنندگان مراجعه کنندگان در سال 1402 الی 1403 به مراکز درمانی تحت پوشش علوم پزشکی شهرستان سبزوار به همراه طیف وسیعی از ویژگی های بالینی، رفتاری و زیستی (شامل 62 ویژگی) مختلف بوده است که با استفاده از تکنیک های پیش پردازش و خوشه بندی مبتنی بر درخت تصمیم بهبود یافته، کیفیت داده ها ارتقا یافته اند. در نهایت، الگوریتم جستجوی ممنوعه به عنوان یک روش بهینه سازی فراابتکاری برای تنظیم پارامترهای (C و گاما) ماشین بردار پشتیبان به عنوان یک مدل استفاده شده است تا با یافتن بهترین پارامترها، مرز تصمیم بهینه برای تشخیص بیماران دیابتی و غیردیابتی ایجا کند. ارزیابی ها نشان داده اند که مدل پیشنهادی با دقت 85/61 درصد، بازیابی 73/24، صحت 82/17 و امتیاز 84/31 توانسته است تعادل نسبی مناسبی میان دقت و بازیابی برقرار کند و در شناسایی نمونه های مثبت نسبت به چند مرجع معتبر بهبود قابل توجهی داشته باشد اما در مقایسه با روش های فراابتکاری سایر مراجع ضعیف عملکرد داشته است.

کلیدواژه‌ها: دیابت، داده کاوی، درخت تصمیم بهبود یافته، ماشین بردار پشتیبان، الگوریتم جست و جوی ممنوعه

شماره‌ی پایان‌نامه: ۱۲۷۴۲۹۲۰۷۹۰۴۲۶۳۹۰۴۰۳۱۱۶۲۹۷۳۶۸۵

تاریخ دفاع: ۱۴۰۴/۰۶/۲۵

رشته‌ی تحصیلی:

دانشکده:

استاد راهنما: دکتر علی اکبر نقابی

Thesis:

Presenting a method to diagnose and optimize diabetes by using data mining methods and Tabu search algorithm



This study aims to propose a novel method for the diagnosis and optimization of diabetes by leveraging data mining techniques and the tabu search algorithm. Diabetes, as one of the most common chronic diseases worldwide, holds significant importance in public health; timely and accurate diagnosis plays a crucial role in preventing complications and reducing treatment costs. Therefore, developing efficient and optimized predictive models to identify diabetic patients is essential. The research methodology is based on real data from patients visiting healthcare centers affiliated with the Sabzevar University of Medical Sciences during 20232024, encompassing a wide range of clinical, behavioral, and biological features (62 features in total). Data quality was enhanced through preprocessing techniques and clustering based on an improved decision tree. Ultimately, the tabu search algorithm, as a metaheuristic optimization method, was employed to tune the parameters (C and γ) of the support vector machine model to find the best parameters and establish an optimal decision boundary for distinguishing diabetic from non-diabetic patients. Evaluations showed that the proposed model achieved an accuracy of 85.61%, recall of 73.24%, precision of 82.17%, and an F1-score of 84.31%, demonstrating a reasonable balance between precision and recall and significant improvement in identifying positive cases compared to several reputable references. However, it performed weaker relative to other metaheuristic methods reported in the literature.