

تحلیل آماری و داده‌کاوی بر روی عوامل بیماری‌های قلبی

سید شهاب‌الدین شاه ابراهیمی^۱، روزبه قوسی^۲

^۱دانشجوی کارشناسی ارشد، دانشگاه علم و صنعت ایران؛ shahebrahimi@ind.iust.ac.ir

^۲عضو هیئت‌علمی دانشکده صنایع، دانشگاه علم و صنعت ایران؛ ghousi@iust.ac.ir

چکیده

داده‌کاوی فنی برای کشف دانش جدید از پایگاه‌های داده است. تشخیص بیماری‌های مختلف در علم پزشکی یکی از زمینه‌های پرکاربرد داده‌کاوی محسوب می‌شود. بیماری‌های قلبی که یکی از علل اصلی مرگ‌ومیر در جهان است. در این تحقیق به تحلیل و شناسایی عوامل تأثیرگذار بر بیماری‌های قلبی پرداخته، سپس مؤثرترین فیلدهای پیشگو در بیماری‌های قلبی بررسی می‌شود و در نهایت به ارزیابی مدل‌ها پرداخته می‌شود و مجموعه قوانینی در تشخیص زودهنگام بیماری قلبی و جلوگیری از مرگ بیماران ارائه می‌شود. در این تحقیق از فن‌های درخت تصمیم استفاده شده که توسط قوانین انجمنی استخراج شده که درخت تصمیم ایجاد شده را پشتیبانی می‌نمایند. در ادامه هم به تحلیل آماری برخی عوامل تأثیرگذار در بیماری قلبی پرداخته شده است.

کلمات کلیدی

داده‌کاوی، عوامل بیماری‌های قلبی، درخت تصمیم C5.0، تحلیل آماری بیماران قلبی.

Statistical Analysis and Data mining Approach for Diagnosis of Heart Diseases

Seyed Shahabedin Shahebrahimi, MS,¹ Dr.Rouzbeh Ghousi,PHD,²

¹Socio-Economical Engineering, School of Progress Engineering, Iran University of Science and Technology, Tehran, Iran

²Proffessor of Industrial Engineering, School of industrial Engineering, Iran University of Science and Technology,Iran

ABSTRACT

Data mining is a new technique for extracting new knowledge from data bases. Diagnosis of different diseases in health care is one the most popular aspects of data mining. Heart diseases are one of the most important death factors in the world. In this research, analysis and identification of effective factors on heart diseases are aimed, then the most effective field will be identified and at the end there will be an assessment and judgment between data mining models and a set of rules for forecasting of heart diseases will be presented. Decision tree rules are used in this research and they have been supported by association rules.in the continuous of the research a set of statistical analysis is presented about effective fields in heart disease.

KEYWORDS

Data mining, Heart diseases factors, Decision tree c5.0, Heart diseases statistical analysis.

^۱ سید شهاب‌الدین شاه ابراهیمی تلفن همراه : ۰۹۱۸۷۲۳۵۵۹۷