

طراحی سیستم پیش بینی بیماری قلبی-عروقی با استفاده از یادگیری ماشین

محمد کاظمی^۱، امیرعباس کاظمی^۲، امیرحسین باب الحکمی^۳، نگار عباسی^۴

^۱ دانشجوی کارشناسی ارشد، گروه کامپیوتر، دانشگاه آزاد اسلامی مشهد، LPLN@Live.com

^۲ دانشجوی کارشناسی ارشد، گروه کامپیوتر، دانشگاه فردوسی مشهد، a.kazemi@mail.um.ac.ir

^۳ دانشجوی دکتری، گروه کامپیوتر دانشگاه فردوسی مشهد، Amir.Babolhakam@mail.um.ac.ir

^۴ کارشناسی ارشد، گروه آمار دانشگاه پیام نور مشهد، R.Abbaasi@gmail.com

چکیده: بیماری‌های قلبی یکی از شایع‌ترین بیماری‌هاست که در حال حاضر تعداد افراد مبتلا به این نوع بیماری‌ها در حال افزایش می‌باشد. این درحالی است در صورتی که مراقبت‌های لازم برای بیمار در زمان مناسب صورت نگیرد، می‌تواند باعث مرگ بیمار شود. در این مقاله با استفاده از رویکردی مبتنی بر انتخاب ویژگی به عنوان یک گام پیش‌پردازش، مدلی باهدف تشخیص بیماری قلبی ارائه گردد. راهکار پیشنهادی دارای ۳ گام اصلی می‌باشد که گام ۱) پیش پردازش داده‌ها با پاکسازی داده‌ها، گام ۲) انتخاب ویژگی‌های موثر براساس ۲ روش پیرسون و PCA و گام ۳) طبقه‌بندی داده‌ها برای پیش‌بینی بیماری قلبی ساخته می‌شود. نتایج بدست آمده نشان می‌دهد الگوریتم J48 با دقت ۰.۸۹، دارای بالاترین دقت است. کلید واژه- تشخیص بیماری قلبی، انتخاب ویژگی، ضریب همبستگی پیرسون، تحلیل مولفه‌های اساسی، طبقه بندی

۱- مقدمه

آمده است تا با جمع آوری اطلاعاتی مرتبط، به پیش بینی بیماری قلبی بپردازیم. یکی از عملکردهای پیش گوینه در داده کاوی، دسته بندی است. دسته بندی، فرآیند یافتن مدلی است که با تشخیص دسته ها و یا مفاهیم داده می تواند دسته ناشناخته اشیا دیگر را پیش گویی کند [1]. هدف از پیش بینی، کمک به روند تشخیص و بهبود سلامت بیماران است. تاکنون مطالعات زیادی پیرامون تکنیک های داده کاوی بر روی انباره بیماران قلبی و عروقی انجام شده است، که تکنیک های مختلف داده کاوی مانند درخت تصمیم و شبکه های عصبی را به کار برده اند که (a) پیش بینی حمله قلبی و سندرم کرونری حاد با استفاده از شبکه عصبی داده کاوی [4]-[2] (b) ارزیابی معیارهای مهم برای پیش بینی بقا بیمار با استفاده از شبکه های بیزین [5] (c) تشخیص ایسکمی قلبی در رکوردهای طولانی مدت الکتروکاردیوگرام [6] (d) ارزیابی فاکتورهای ریسکی مرتبط با بیماری قلبی با استفاده از درخت تصمیم و قوانین تلازمی [7]، [8] و (e) کشف روابط بین عوامل خطرهای قلبی با استفاده از درخت تصمیم [9] از جمله پژوهش های انجام شده در این زمینه است.

بیماری‌های قلبی یکی از شایع‌ترین بیماری‌ها می‌باشند که در حال حاضر تعداد افراد مبتلا به بیماران قلبی در حال افزایش است. به طور کلی این بیماری دارای علائم مختلفی می باشد اما از مهمترین آن می توان به درد در قفسه سینه، ناحیه فک، گردن، نفس تنگی و ... اشاره کرد. همچنین مشکلات قلبی می تواند باعث بروز بیماری عروق کرونر قلب، نارسایی قلبی و سکتة مغزی شود. با این وجود، اگرچه عامل اصلی مرگ و میر در دهه اخیر در دنیا شناخته شده است، به عنوان کنترل پذیرترین و قابل پیشگیری ترین بیماری نیز شناخته می شود. نتایج بدست آمده نشان می دهد، شناسایی و تشخیص به موقع بیماری قلبی علاوه بر امکان جلوگیری از پی شرف و بروز عوارض ناشی از آن، می تواند در درمان صحیح و کامل بیماری موثر باشد. برای تشخیص به موقع بیماری قلبی تست های مختلفی مانند نوار قلب، تست استرس و آنژیو قلب وجود دارد ولی این نوع آزمایش ها هزینه ی زیادی در بر دارند و نمی توان آنها را به طور گسترده استفاده کرد. تشخیص در مراحل ابتدایی کمک شایانی به روند درمان افراد میکند و نسبت به روش های متداول پزشکی مقرون به صرفه تر است. امروزه داده کاوی و کشف دانش به کمک ما