

مقایسه دو ابزار داده کاوی در تشخیص بیماری دیابت نوع ۲

محمد حسین زاهدی^۱، فاطمه سادات کیا^۲، راضیه عسگر نژاد^۳

۱- دانشجوی کارشناسی ارشد کامپیوتر نرم افزار، دانشگاه غیر انتفاعی سینا، کاشان، ایران

۲ و ۳- دکتری کامپیوتر نرم افزار، دانشگاه آزاد اسلامی واحد اصفهان (خوراسگان)، اصفهان، ایران و

گروه کارشناسی ارشد کامپیوتر نرم افزار، دانشگاه غیر انتفاعی سینا، کاشان، ایران

چکیده

امروزه با پیشرفت فناوری ها، به ویژه فناوری های مرتبط با اطلاعات و ارتباطات، حجم عظیمی از داده ها در حال تولید در شبکه های ارتباطی و اطلاعاتی هستند و یکی از ضرورت های موفقیت های کسب و کارها، حتی در مقیاس کوچک، امکان بهره گیری از این اطلاعات و داده هاست. مدیران کسب و کارهای مختلف، می دانند که جمع آوری اطلاعات و داده ها از مشتریان و مخاطبان یک شرکت، یکی از عوامل مورد نیاز برای رشد و توسعه این شرکت هاست. اما این فقط نصف مسیر است؛ اگر داده ها فقط جمع شوند و بلا استفاده بمانند، عملاً هدف اصلی از جمع آوری این اطلاعات، برآورده نشده است. گام مهمی که پس از جمع آوری داده ها و اطلاعات می بایست برداشته شود، استخراج دانش از اطلاعات و ملموس تر کردن داده های جمع آوری شده، برای استنتاج قواعد و نتایج کاربردی است. تا به امروز نرم افزار های تجاری و آموزشی فراوانی برای داده کاوی در حوزه های مختلف داده ها به دنیای علم و فناوری عرضه شده اند. هریک از آنها با توجه به نوع اصلی داده هایی که مورد کاوش قرار می دهند، روی الگوریتمهای خاصی متمرکز شده اند.

در این مقاله سعی شده است با مقایسه دو نرم افزار پرکاربرد در زمینه داده کاوی با نام Rapidminer و HeuristicLab نتایج بدست بیاوریم و نتیجه گیری کنیم کدام نرم افزار نتایج و عملکرد بهتری می توانند در داده کاوی داشته باشند.

کلمات کلیدی: دیتاست Pima، مدل های داده کاوی (Support vector - (Linear Discriminant Analysis machine classification), (Neural Network),