

عنوان مقاله:

بررسی تاثیر پیش پردازش بر روی کیفیت داده و مدل های پیش بینی عود مجدد سرطان پستان

محل انتشار:

چهارمین کنفرانس بین المللی مطالعات نوین در علوم کامپیوتر و فناوری اطلاعات (سال: 1396)

تعداد صفحات اصل مقاله: 12

نویسندگان:

زینب سجادینا - کارشناسی ارشد مهندسی کامپیوتر- نرم افزار، دانشگاه صنعتی شیراز، شیراز

سیدریوف خیامی - استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات دانشگاه صنعتی شیراز، شیراز

سیدمحمدرضا موسوی - استادیار، دانشکده مهندسی برق و کامپیوتر- بخش مهندسی و علوم کامپیوتر و فناوری اطلاعات دانشگاه شیراز، شیراز

مهديه ديانى - متخصص رادیوتراپی آنکولوژی، مدیر کمیته آموزش و پژوهش مرکز تخصصی رادیوتراپی آنکولوژی رضا(ع)، مشهد

خلاصه مقاله:

امروزه به دلیل شیوع انواع سرطان ها، محققان زیادی علاقه مند به کشف دانش مفید موجود در داده های سرطانی شده اند. داده کاوی یکی از ابزارهای قدرتمندی است که بدین منظور استفاده می شود اما نتایج آن زمانی قابل اعتماد و مفید است که داده های با کیفیتی در اختیار باشد. با توجه به اینکه اکثر داده های دنیای واقعی دارای نواقصی همچون ناسازگاری، خطا، نویز، مقادیر گمشده و ... میباشد، پیش پردازش داده ها امری جداناپذیر از داده کاوی است. در این مقاله بر روی یکمجموعه داده واقعی سرطان پستان کار شده که هدف بررسی تاثیر پیش پردازش بر روی کیفیت داده و نتایج مدلهایپیش بینی عود سرطان پستان می باشد. در ابتدا با استفاده از سه الگوریتم طبقه بندی نزدیکترین همسایه، بیز ساده و بهینه سازیحداقل ترتیبی، مدلهای پیش بینی ساخته شده اند. سپس پیش پردازش هایی از قبیل رفع خطا و ناسازگاری، پرکردن مقادیربدون محتوی و کاهش داده (انتخاب ویژگی) بر روی مجموعه داده اعمال و مجددا مدل های پیش بینی ساخته شده اند. درپایان نتایج مدل های پیش بینی با استفاده از دو معیار صحت و حساسیت با یکدیگر مقایسه گردیده اند. نتایج ارزیابی نشانمیدهد که مدل های پیش بینی بهینه سازی حداقل ترتیبی از نظر صحت و حساسیت به ترتیب 3.56 و 73.59، بیز ساده 4.86 و 66.16 و نزدیکترین همسایه 5.4 و 60.9 درصد نسبت به قبل از پیش پردازش دادهها رشد داشته است. بنابراین پیشپردازش موجب بهبود نتایج طبقه بندی و کیفیت داده می شود.

کلمات کلیدی:

پیش پردازش داده ها، سرطان سینه، داده کاوی، کیفیت داده، طبقه بندی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/779209>

