

ارائه مدل مناسب برای انتخاب ویژگی در بیان ژن در سیستم تشخیص سرطان مبتنی بر الگوریتم گرگ خاکستری

حمید رحیمی^۱*

۱- باشگاه پژوهشگران جوان، واحد مشهد، مشهد، ایران Rahimihmd@gmail.com



چکیده

به یکی از انواع داده‌ها که دارای تعداد ویژگی‌های زیادی هستند و اصطلاحاً دارای ابعاد بالایی است، داده‌های ژنی در علم پزشکی می‌باشد. این داده‌ها جهت ارزیابی ژن‌ها برای تشخیص بیماری‌ها به خصوص بیماری سرطان به کار می‌روند. بیماری سرطان سینه یکی از شایع‌ترین بیماری‌ها در طی سالیان اخیر بوده است که برای درمان آن باید بیماری در مراحل اولیه شناسایی گردد و این شناسایی در مراحل اولیه اهمیت ویژه‌ای دارد. به علت زیادی حجم داده‌ها، بایستی از روش‌های مبتنی بر یادگیری ماشین جهت کاهش داده‌ها و ساخت مدل تصمیم‌گیرنده استفاده کرد زیرا یکی از مشکلات داده‌های با ابعاد زیاد اینست که در بیشتر مواقع تمام ویژگی‌های داده‌ها برای یافتن دانشی که در داده‌ها نهفته است مهم و حیاتی نیستند. از این رو در این پژوهش سعی در ارائه روش جدید برای بدست آوردن زیرمجموعه ویژگی‌های مناسب مجموعه داده بیان ژن بیماری سرطان سینه با استفاده از روش الگوریتم گرگ خاکستری شده و برای یادگیری سیستم تشخیص سرطان از شبکه عصبی پرسپترون چندلایه استفاده شده است.

واژه‌های کلیدی: الگوریتم گرگ خاکستری، بیان ژن سرطان سینه، تشخیص بیماری، شبکه‌های عصبی مصنوعی

۱- مقدمه

امروزه بیماری سرطان به یکی از خطرناک‌ترین بیماری‌ها تبدیل شده است که علم پزشکی را درگیر خود کرده است. یکی از روش‌های بررسی بیماری سرطان، بررسی داده‌های ژنی است. با توجه به اینکه این داده‌ها ذاتاً حجم بالایی از نظر تعداد ویژگی‌ها دارند و همه‌ی این ویژگی‌ها دارای اطلاعات مفیدی نمی‌باشند، بنابراین باید با استفاده از روش‌هایی ویژگی‌های مهم شناسایی گردند. از طرفی حجم داده‌ها به صورت روز افزون در حال افزایش است و بررسی ویژگی‌ها به صورت دستی امکان پذیر نخواهد بود. زیرا بررسی دستی توسط انسان با توجه به خطای بالای عامل انسانی سبب می‌شود نتایج خوبی حاصل نگردد. لذا از روش‌های یادگیری ماشین می‌توان برای تحلیل این داده‌ها بهره برد. هنگامی که حجم داده‌ها زیاد باشد و هدف تحلیل داده‌ها باشد روش‌های یادگیری ماشین می‌توانند نتایج خوبی ارائه دهند. با استفاده از روش‌های یادگیری ماشین می‌توان ویژگی‌های مناسب را انتخاب کرد [1]. در این پژوهش برای انتخاب ویژگی از الگوریتم‌های فراابتکاری استفاده می‌شود. الگوریتمی که در این پژوهش مورد استفاده قرار گرفتند، الگوریتم بهینه‌سازی گرگ خاکستری است که با تعریف مناسب در این پژوهش به کار برده شده است. از طرفی برای دسته‌بندی نمونه‌های ژنی از ابزار طبقه‌بند شبکه عصبی استفاده کردیم.

^۱ کارشناسی ارشد کامپیوترگرایش هوش مصنوعی - مدرس دانشگاه آزاد اسلامی واحد مشهد