

عنوان مقاله:

کاربرد یادگیری ماشین در تحلیل داده های متیلاسیون DNA با هدف تشخیص بیماریهای ژنتیکی

محل انتشار:

پنجمین همایش بین المللی علوم و تکنولوژی با رویکرد توسعه پایدار (سال: 1398)

تعداد صفحات اصل مقاله: 17

نویسنده:

بهروز دیندار - گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه اسلامی واحد قزوین، قزوین، ایران

خلاصه مقاله:

متیلاسیون دی ان ای پدیده ای است که وابستگی قدرتمندی با ابتلای اشخاص به سرطان دارد. بروز این پدیده در بدو تولد دلیلی برای ابتلای شخص به سرطان است. سرطان سینه بر اثر متیلاسیون، یکی از نمونه های بروز متیلاسیون در دی ان ای اشخاص است. مهمترین دغدغه پزشکان برای تشخیص و پیش گیری از بروز سرطان، تحلیل داد هایی است که این داده ها با استفاده از نمونه برداری جمع آوری گردیده است. از آنجایی که حجم این داده ها زیاد است و دارای ابعاد و پیچیدگی بسیاری هستند، بنابراین پزشکان قادر به تحلیل چنین داده هایی نخواهند بود. چراکه تحلیل این نوع داده ها برای پزشکان و متخصصین، فرایندی طاقت فرسا و غیر قابل انجام است و برای تشخیص صحیح نمونه ها، دقت و سرعت لازم است. در این تحقیق سعی داریم تا با استفاده از یادگیری ماشین و داده کاوی به تحلیل داده های متیلاسیون دی ان ای برای تشخیص صحیح سرطان سینه و تومور های زیرمجموعه آن، پزشکان را یاری کنیم. مدل ت4ن4ک و تجزیه اندازه تکین روش هایی هستند که در این تحقیق برای کاهش بعد مورد استفاده قرار می گیرند. پس از آنکه کاهش بعد انجام شد با دو رویکرد به تحلیل داده های متیلاسیون می پردازیم. رویکرد اول مبتنی بر یادگیری بدون نظارت است و رویکرد دوم مبتنی بر یادگیری با نظارت می باشد. هدف ما در این مقاله این است که بررسی کنیم کدام روش تحلیل داده برای افزایش دقت مناسب است در رویکرد اول از خوشه بندی کننده K-میانگین و شبکه عصبی نگاشت خود سازمان دهنده استفاده می کنیم. این رویکرد مبتنی بر یادگیری بدون نظارت است. در رویکرد دوم از ماشین بردار پشتیبان و نزدیکترین همسایه وزن دار استفاده می کنیم. رویکرد دوم مبتنی بر یادگیری با نظارت است. پس از انجام آزمایش الگوریتم مورد استفاده رویکردها روی داده های متیلاسیون سرطان سینه، این دو رویکرد و الگوریتم های آنها از لحاظ دقت، نرخ خطا، حساسیت، صحت و دیگر موارد باهم مقایسه شده و الگوریتمی که بالاترین دقت را دارد معرفی می گردد.

کلمات کلیدی:

متیلاسیون دی ان ای، یادگیری ماشین، کاهش بعد، خوشه بندی، طبقه بندی، دقت

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/967394>

