

## آنالیز بیان ژنهای سلول با استفاده از Support Vector Machines

امیر حسین اسدی<sup>۱</sup>; حسام الدین ترابی دشتی<sup>۲</sup>; نسیم جمالی<sup>۳</sup>.

آنالیز رفتار سلولهای موجودات زنده به منظور ادراک فرایندهای حیاتی سلول یکی از مهمترین مسائل مطرح در زمینه تحقیقاتی بیوانفورماتیک بشمار می‌رود. آنالیز رفتار سلول از طریق بررسی تعداد دفعات کپی‌برداری - که به آن بیان ژن می‌گوییم - از ژنهایش صورت می‌گیرد. رفتاری که سلول در شرایط مختلف باید انجام دهد از طریق تغییر مقدار بیان برخی از ژنها تعریف می‌گردد. در مورد تغییر مقدار بیان ژن، نکته حائز اهمیت، تاثیر تغییر بیان یک ژن بر مقدار بیان سایر ژنهاست. همانگونه که اشاره شد، تعیین تاثیر بیان ژنها بر یکدیگر مسئله حائز اهمیتی در آنالیز رفتار سلول می‌باشد. تحقیقات زیست‌شناسان در این زمینه نشان داده است که این تاثیر به صورت گروهی صورت می‌پذیرد. این بدان معنی است که تغییر بیان دسته‌ای از ژنها باعث تغییر بیان ژن دسته یا دسته‌های دیگری از ژنها می‌گردد. این تاثیرات به صورت سلسله‌مراتبی به سایر دسته‌ها<sup>۴</sup> نیز منتقل می‌شود؛ بعنوان مثال افزایش/کاهش مقدار بیان ژنهای دسته A باعث افزایش/کاهش مقدار بیان ژنهای دسته B می‌گردد، که این تغییر خود بر مقدار بیان ژنهای دسته C می‌تواند تاثیر گذار باشد و این سلسله تغییرات ادامه پیدا می‌کند. این مثال بیانگر درجه اهمیت تشخیص گروه ژنهای تاثیر گذار بر یکدیگر در آنالیز رفتار سلول می‌باشد. گروه بندی<sup>۵</sup> ژنهایی که تاثیر یکسان بر سایر ژنها دارند از طریق بررسی مکاشفای مقدار تغییر بیان ژنها صورت می‌گیرد. توجه به این نکته ضروری است که تاثیر بیان ژنها بر یکدیگر نیازمند گذر زمان است و لذا برای بررسی تغییر بیان ژنها بر یکدیگر باید یک سری زمانی از مقدار بیان ژنها را در اختیار داشته باشیم. پایگاه داده‌ای که ما که برای گروه بندی ژنهای سلولهای مغز استفاده کرده‌ایم مقدار بیان ژن را برای ۴ زمان مختلف در اختیار ما قرار داده است.

متد محاسباتی ارائه شده بر پایه Support Vector Machines به گروه بندی این داده‌ها می‌پردازد. نتایج بدست آمده از این تحقیق نشان می‌دهد که الگوریتم طراحی شده، ۹۸٪ گروه بندی زیست‌شناسان - که بر پایه اطلاعات آزمایشگاهی می‌باشد - را به درستی تشخیص داده و همچنین گروه بندی‌های جدید و سلسله مراتب‌های جدیدی را نیز استخراج نموده است.

Keywords: Clustering, Gene expressions, Gene pathways, Support Vector Machine.

<sup>1</sup> دکتری بیوماتمیک، دانشگاه ویسکانسین، آمریکا.

<sup>2</sup> دانشجوی مقطع کارشناسی ارشد علوم کامپیوتر، دانشگاه تهران . HTDashti@ut.ac.ir

<sup>3</sup> مهندسی نرم‌افزار، مرکز عالی بیوماتمیک، دانشگاه تهران. Nasim.Jamali@gmail.com

<sup>4</sup> به این تاثیرات سلسله‌مراتبی Gene Pathway می‌گویند.

<sup>5</sup> Clustering