

## روش هوشمند پیش‌بینی توالی زمانی بیان ژن‌ها مبتنی بر شبکه‌های فازی-عصبی پویا

صدف ایرانپور طاری<sup>۱</sup>، علی احمدی<sup>۲</sup>

<sup>۱</sup> دانشکده برق و کامپیوتر دانشگاه خواجه نصیر طوسی  
sadafricanpour@ee.kntu.ac.ir

<sup>۲</sup> دانشکده برق و کامپیوتر دانشگاه خواجه نصیر طوسی  
ahmadi@eed.kntu.ac.ir

### چکیده

از آن‌جا که فرآیند تنظیم ژن‌ها در سطح سلولی یک روند دینامیک و پویا است، شبکه‌های عصبی با قابلیت آموزش و تنظیم بر اساس داده‌های آموزشی و سیستم‌های فازی با قابلیت تفسیرپذیری، الگوریتم‌های مناسبی جهت این نوع محاسبات هستند. در پژوهش پیش‌روی، روش‌هایی بر مبنای شبکه‌های عصبی فازی جهت پیش‌بینی ارتباطات پیچیده میان ژن‌ها انجام شده است. روش پیشنهادی، ژن‌هایی که بیشترین تاثیر بر هم دارند را به عنوان ژن‌های تنظیم‌کننده می‌یابد و نوع روابط آن‌ها شامل اثر سدکنندگی، فعال‌کنندگی و یا خنثی را مشخص می‌کند و در نهایت شبکه تنظیم ژنی ترسیم می‌شود. برای آموزش، از مجموعه داده‌های استاندارد ریزآرایه مربوط به ۱۲ ژن شاخص موثر در مدت سیکل جوانه زدن نوعی مخمر (*Saccharomyces Cerevisiae*) استفاده شد. فعل و انفعالات حاصل شده با نتایج آزمایشگاهی بیولوژیکی پیشین، تحت معبرسازی قرار گرفته‌اند و نتایج نشان‌دهنده آن است که با اجرای روش پیشنهادی، ۱۵٪ از تعداد قوانین استخراج شده جهت بخش‌بندی فضای ورودی- خروجی کاسته شده و این منجر به کاهش شدید محاسبات می‌شود، در حالی که میزان مجموع مربعات خطای الگوریتم نیز در مقایسه با نزدیک‌ترین روش از لحاظ الگوریتم، ۰٫۳- کاهش یافته است.

### کلمات کلیدی

بیوانفورماتیک، بیان ژن، شبکه تنظیم ژن، شبکه‌های عصبی- فازی.

### ۱- مقدمه

توانند به صورت مستقل یا به صورت ترکیبی، رونویسی از روی یک ژن را فعال یا متوقف کنند. شبکه‌های تنظیم ژنی این ارتباطات را در قالب یک شبکه مدل می‌کنند. در حال حاضر درک شبکه‌های تنظیم‌کننده ژنی و فهم فرایندهای تنظیمی در یک سلول در سطح ژن، یک هدف مهم در بیولوژی محاسباتی است.

در سال‌های اخیر، تکنولوژی جدیدی به نام ریزآرایه‌های DNA، امکان مطالعه هم‌زمان بیان هزاران ژن را در یک آزمایش برای محققان فراهم می‌کند. مدل‌سازی شبکه‌های ژنی از داده‌های تجربی ریزآرایه، تشریح اعمال سلولی در سطح مولکولی را ساده‌تر می‌سازد. داده‌های سری زمانی ریزآرایه به صورت یک ماتریس عددی با هزاران سطر (ژن) و چند ده ستون (نقطه زمانی) هستند. این داده‌ها مشکلات خاصی دارند که تحلیل آن‌ها را مشکل می‌کند. برخی از این مشکلات شامل وجود نویز و تعداد نمونه‌های کم در مقایسه با تعداد پارامترهای

بیوانفورماتیک دانش استفاده از علوم کامپیوتر و آمار و احتمالات در شاخه زیست‌شناسی مولکولی است. در چند دهه اخیر، با پیشرفت در زیست‌شناسی مولکولی و با بهره از تجهیزات مورد نیاز تحقیق، با استفاده از کامپیوتر، نرم‌افزارهای متعدد و بانک‌های اطلاعاتی بیولوژیکی، تحقیقات وسیعی در خصوص پروتئین‌ها و ژن‌ها به عمل آمده است [۱]. تلاش‌های پژوهشی اصلی در این رشته عبارتند از: تنظیم توالی، کشف ژن، گردآوری ژنوم، تنظیم ساختار پروتئینی، پیش‌گویی ساختار پروتئینی، پیش‌بینی بیان ژن و تعاملات پروتئین- پروتئین.

تنظیم بیان ژن یا تنظیم ژن، به کنترل مقدار و زمان تغییرات در تولید محصول یک ژن اشاره می‌کند. محصولات ژن‌های مختلف می-