

## یک مدل بازی مشارکتی برای تحلیل داده‌های بیان ژنی آزمایش ریزآرایه

یونس کردوندی<sup>۱\*</sup>، حمیدرضا نویدی<sup>۲</sup>، مجید حسن پور عزتی<sup>۳</sup>

۱- کارشناس ارشد ریاضی کاربردی، گرایش تحقیق در عملیات، دانشگاه شاهد، تهران، ایران

۲- دانشیار دانشگاه شاهد، گروه ریاضی کاربردی، تهران، ایران

۳- استادیار دانشگاه شاهد، گروه زیست‌شناسی، تهران، ایران

رسید مقاله: ۶ خرداد ۱۳۹۲

پذیرش مقاله: ۲۲ مهر ۱۳۹۲

### چکیده

فناوری ریزآرایه یک ابزار تحلیلی کوچک است که به کمک آن می‌توان بیان ده‌ها هزار ژن را، به طور همزمان مورد کاوش قرار داد. سوال عمده مورد بحث در ریزآرایه‌ها، چگونگی انتخاب مارکرهای ژنی می‌باشد، تشخیص مارکرهای ژنی می‌تواند تشخیص زودتر، درمان و پیش‌بینی دقیق‌تر نتیجه بیماری و متعاقب آن استفاده از درمان‌های مناسب و کاهش هزینه‌ها را در پی داشته باشد. با استفاده از نظریه‌های آماری برای پردازش داده‌های ریزآرایه، تنها قادر به شناسایی تعداد محدودی از ژن‌های با نظم متفاوت در نمونه‌های تحت بررسی خواهیم بود. به منظور تقویت سیگنال‌های حاوی اطلاعات مفید بیولوژیکی، اخیراً محققان نظریه بازی توانسته‌اند روشی را ابداع نمایند که با به‌کارگیری آن روی داده‌های بیان ژنی آزمایش ریزآرایه، قادر خواهیم بود گروهی از ژن‌های با تاثیرگذاری مضاعف در بروز یک اختلال را شناسایی کنیم. در این روش، با استفاده از نظریه بازی‌های همکارانه، به مدل‌سازی سهم مشارکت هر یک از ژن‌های دخیل در بروز یک اختلال که بیان‌شان دچار تغییر شده است، پرداخته می‌شود. در این تحقیق، مدل مذکور با ترکیب حلی مفهوم ارزش شیپلی از بازی‌های همکارانه با قابلیت انتقال سودمندی به همراه مفهوم بوت‌استرپ آماری، گزینشی از ژن‌های با تاثیرگذاری مضاعف در بروز یک اختلال مورد بررسی قرار گرفته است. در یک مطالعه موردی، این مدل‌سازی روی داده‌های به‌دست آمده از افراد مبتلا به بیماری اوتیسم اجرا شده است. نتایج به‌دست آمده از این پژوهش همخوانی بالایی با نتایج تحقیقات بیولوژیک از خود نشان می‌دهد.

**کلمات کلیدی:** نظریه بازی‌های همکارانه، بازی ریزآرایه، تحلیل مقایسه‌ای ارزش شیپلی، بیان ژن

### ۱ مقدمه

فناوری ریزآرایه یک ابزار تحلیلی کوچک است که اجازه می‌دهد تجزیه و تحلیل ژنومی با سرعت و دقت انجام شود؛ به کمک این فناوری می‌توان بیان ده‌ها هزار ژن را به‌طور همزمان مورد بررسی قرار داد [۱]. این فناوری

\*عنده دار مکاتبات

آدرس الکترونیکی: younes.kord@yahoo.com

به طور گسترده در تحقیقات پزشکی برای شناسایی بیان ژنی خاص که عامل ایجاد بیماری است، مورد استفاده قرار می‌گیرد. از آزمون‌های آماری متفاوتی برای اندازه‌گیری تفاوت سطح بیان ژن‌ها، در آزمایشات ریزآرایه استفاده می‌شود، در تعدادی از این آزمون‌ها (از جمله آزمون  $t$ )، ژن‌ها براساس مقدار احتمال  $p$ -value که به دست می‌آورد رتبه‌بندی می‌شوند، اما  $p$ -value ها لزوماً اشاره به یک سیگنال بیولوژیکی ندارند، از طرف دیگر پردازش‌های آماری برای تحلیل  $p$ -value ها، اغلب برای تصحیح یک کشف نادرست به کار می‌روند و در بسیاری از موارد، این تصحیح‌ها، برای شناسایی سیگنال‌های حاوی اطلاعات مفید بیولوژیکی، محافظه کارانه عمل می‌کنند. از طرف دیگر، اکثر روش‌های کلاسیک آماری برای شناخت مارکرهای ژنی یک بیماری قادر به شناسایی تعداد معدودی از ژن‌های موثر در بروز یک اختلال بیولوژیکی هستند. با انگیزه نیاز به یک نظریه قوی‌تر جهت شناسایی سیگنال‌های مفید بیولوژیکی، اخیراً یک نظریه جدید بر پایه مفاهیمی از تئوری بازی‌ها ارائه شده است. امتیاز مهم این روش، محاسبه یک شاخص عددی برای ژن‌ها می‌باشد، که ارزش شیپلی نام دارد، این شاخص رابطه هر ژن را تحت یک حالت بیولوژیکی خاص مشخص می‌کند، حال آنکه به طور همزمان، رفتارهای بیان بقیه ژن‌ها را نیز، تحت همان حالت یکسان در نظر می‌گیرد [۲]. در این پژوهش، یک روش جدید برای تحلیل داده‌های بیان ژنی، بر پایه مقایسه ارزش‌های شیپلی هر ژن در دو حالت مختلف، توصیف می‌شود. این روش با ترکیب مفهوم ارزش شیپلی از بازی‌های همکارانه با مفهوم بوت‌استرپ آماری، گزینشی از ژن‌های با تاثیرگذاری مضاعف در بروز یک اختلال را نتیجه می‌دهد [۳]. آزمون  $t$  برای تحلیل داده‌های ریزآرایه، ژن‌ها را براساس تفاوت بیان فردی‌شان میان دو حالت مورد مطالعه انتخاب می‌کند، در روش نظریه بازی برای تحلیل داده‌های بیان ژنی آزمایش ریزآرایه نیز، بیان از هر ژن تحت دو حالت مختلف در نظر گرفته می‌شود، اما ارزش افزوده مقدار شیپلی، قادر به برجسته کردن نقش آن دسته از ژن‌هایی است که به طور پیوسته‌ای، با ژن‌های دیگر در تعامل اند. این روش، رابطه از ژن‌ها را به عنوان سهم حاشیه‌ای میانگین از بین تمام جایگشت‌های ممکن از ژن‌ها محاسبه می‌کند، از این رو ژن‌های دارای رابطه بالا، آنهایی هستند که با احتمال زیاد، تبیین کننده تفاوت میان دو حالت می‌باشند، زیرا این ژن‌ها، نقش مهمی را (روی میانگین) در میان تمام جایگشت‌های ممکن از ژن‌ها، نه صرفاً با در نظر گرفتن تفاوت بیان فردی‌شان، بازی می‌کنند [۲]. در این مقاله، در جهت مشخص نمودن میزان کارایی، این روش نوین روی داده‌های بیان ژنی افراد مبتلا به بیماری اوتیسم به کار گرفته می‌شود، نتایج حاصل از به کارگیری این روش روی داده‌های بیان ژنی به دست آمده از افراد دچار اختلالات طیف اوتیسم، مزیت‌های مهم استفاده از روش نظریه بازی را آشکار می‌سازد.

## ۲ نظریه بازی‌های همکارانه با قابلیت انتقال سودمندی

نظریه بازی شاخه‌ای از تحقیق در عملیات است که در بستر علم اقتصاد توسعه یافته و به بررسی رفتار استراتژیک بین عوامل عقلانی بازی می‌پردازد. رفتار استراتژیک زمانی بروز می‌کند که مطلوبیت هر عامل، نه فقط به استراتژی انتخاب شده توسط وی بلکه به استراتژی انتخاب شده توسط بازیگران دیگر هم بستگی داشته باشد [۴].

بازی توصیفی از تعاملات استراتژیکی است که بین افراد تصمیم گیرنده رخ می دهد، بنابراین هر الگوی بازی متشکل از مجموعه بازیگران، استراتژی های ممکن برای هر بازیگر و توابع بازدهی (تابعی از کنش فرد و قیمت ها و کنش دیگران) می باشد. بازی ها دارای ابعاد مختلفی می باشند و به همین خاطر طبقه بندی های مختلفی از آنها ارائه شده است، یکی از مهمترین این طبقه بندی ها، طبقه بندی براساس همکاری و عدم همکاری میان بازیکنان است، اگر بازیکنان پیرامون انتخاب یک استراتژی با هم متحد شوند و این استراتژی بین آنها قابل اجرا و عملی باشد، بازی را همکاریانه و در غیر اینصورت بازی را غیر همکاریانه می نامند [۵].

بازی های با مطلوبیت قابل انتقال نوع خاصی از بازی های همکاریانه است که در آن درآمد ائتلاف، می تواند به شکل دلخواهی بین اعضایش تقسیم شود. یک بازی با مطلوبیت قابل انتقال یا به طور خلاصه یک بازی TU، به صورت دوتایی  $(N, v)$ ، تعریف می شود که  $N$  مجموعه متناهی از بازیکنان را نمایش می دهد و  $v: 2^N \rightarrow \mathbb{R}$  تابع مشخصه با  $v(\emptyset) = 0$  است. گروه بازیکنان  $T \subseteq N$  یک ائتلاف نامیده می شود. تخصیص یا بردار بازدهی  $(x_i)_{i \in N}$  در بازی، برداری است که بازدهی بازیکنان را توصیف می کند. یک جواب برای بازی TU تابعی است که به هر بازیکن تخصیص نسبت می دهد. معروف ترین جواب در خانواده بازی های همکاریانه با مطلوبیت قابل انتقال، ارزش شیلی است که به هر بازیکن، سهم حاشیه ای میانگین اش را از میان تمام جایگشت های ممکن بازیکنان اختصاص می دهد [۶].

فرض کنید،  $(N, v)$  یک بازی با مطلوبیت قابل انتقال بوده، و  $\sigma$  جایگشتی از مجموعه بازیکنان باشد، تصور کنید بازیکنان یک به یک، به ترتیب  $\sigma(1)$ ،  $\sigma(2)$  و ...  $\sigma(n)$  وارد یک اتاق شده و به هر کدام سهم حاشیه اش را که در بازی ایجاد کرده پرداخت می شود، مجموعه بازیکنان ماقبل  $i$  در  $\sigma$  را به صورت ائتلاف  $p_\sigma(i)$  تعریف می کنیم، بردار حاشیه ای  $m^\sigma$  را به صورت زیر تعریف می کنیم:

$$m_i^\sigma = v(p_\sigma(i) \cup \{i\}) - v(p_\sigma(i)) \quad (1)$$

بدین ترتیب بردار حاشیه ای  $m^\sigma$ ، سهم حاشیه ای هر بازیکن را در ائتلافی که با ورود او شکل گرفته براساس ترتیب  $\sigma$  پرداخت می کند.

بازی مشارکتی  $(N, v)$  را در نظر بگیرید، برای هر بازیکن  $i \in N$  ارزش شیلی  $\phi_i(v)$  به صورت میانگین بردارهای حاشیه ای بازی تعریف می شود:

$$\phi_i(v) = \frac{1}{n!} \sum_{\sigma \in \Pi(N)} m_i^\sigma \quad (2)$$

(در این تعریف  $\Pi(N)$  نمایانگر مجموعه جایگشت های  $N$  است).

### ۳ مدل بازی ریزآرایه

ایده اساسی مدل بازی ریزآرایه، از نظریه بازی‌های همکارانه با مطلوبیت قابل انتقال نشأت می‌گیرد، به طور خاص از چارچوب بازی‌های ساده که نوعی بازی TU است استفاده می‌کنیم که به طور گسترده‌ای در تجزیه و تحلیل توان بازیکنان در موقعیت‌های تعاملی مانند شورا، مجلس و ... کاربرد دارد.

این مدل بازی از مقایسه دو ماتریس بیان ژن که یکی نمونه‌های توموری و دیگری نمونه‌های سالم است، حاصل می‌شود. این مقایسه در واقع به کارگیری یک روش تفکیک‌کننده، روی هر نمونه برای تقسیم مجموعه تمام ژن‌ها به دو نمونه است؛ یعنی آن ژن‌هایی که سطح بیان متفاوتی از نمونه‌های سالم دارد و آن ژن‌هایی که سطح بیان شان با نمونه‌های سالم متناظر است؛ سپس با استفاده از یک رابطه علت و معلولی تابع مشخصه بازی را می‌یابیم [۷].

فرض کنید  $G = \{1, 2, \dots, n\}$  مجموعه  $n$  ژن و  $S_R = \{1, 2, \dots, r\}$  مجموعه  $r$  نمونه سالم و  $S_D = \{1, 2, \dots, d\}$  مجموعه  $d$  نمونه بیمار باشد؛ هدف آزمایش ریزآرایه، نسبت دادن یک پروفایل بیان  $(a_{ij})_{i \in G}$  به هر نمونه  $j \in S_R \cup S_D$  است؛ یعنی  $a_{ij} \in \mathbb{R}$  مقدار بیان مربوطه هر ژن  $i$  در نمونه  $j$  است. مجموعه داده‌های آزمایش می‌تواند به شکل دو ماتریس بیان  $A^{S_R} = (A^j)_{j \in S_R}$  و  $A^{S_D} = (A^j)_{j \in S_D}$  نشان داده شود که در آن اندیس نمایانگر یک ستون است. به طور خلاصه چندتایی  $E = \langle G, S_R, S_D, A^{S_R}, A^{S_D} \rangle$  را به عنوان یک وضعیت آزمایشگاهی ریزآرایه (MES) می‌شناسیم.

وضعیت آزمایشگاهی ریزآرایه  $E$  را در نظر بگیرید. برای پی بردن به این مساله که چه ژن‌هایی در نمونه‌های بیمار به طور غیر نرمال بیان شده‌اند از یک روش تفکیک‌کننده استفاده می‌کنیم و به هر ژن  $i \in G$  که در نمونه  $j \in S_D$  بیان غیر نرمال دارد، مقدار یک و به ژنی که بیان نرمال دارد، مقدار صفر نسبت می‌دهیم  $(B^{E,m} \in \{0,1\}^{G \times S_D})$ . در ادامه میانگین تعداد نمونه‌های توموری که به علت بیان غیرنرمال گروهی از ژن‌ها ایجاد شده‌اند، را تعیین کرده و از این مفهوم به عنوان تابع مشخصه بازی استفاده می‌کنیم.

فرض کنید  $a \in \{0,1\}^n$  که  $n \in \{1, 2, \dots\}$ ، در این صورت محمل  $a$  که با  $\text{sp}(a)$  نمایش داده می‌شود به صورت مجموعه  $\text{sp}(a) = \{i \in \{1, \dots, n\} \mid a_i = 1\}$  تعریف می‌شود. بازی ریزآرایه به صورت بازی  $(N, v)$  تعریف می‌شود که مجموعه بازیکنان را مجموعه ژن‌ها در نظر گرفته و تابع مشخصه را برای هر ائتلاف  $T \in 2^N \setminus \{\emptyset\}$  به صورت زیر تعیین می‌کنیم:

$$v(T) = \frac{|\Theta(T)|}{|S_D|} \quad (3)$$

که در آن  $|\Theta(T)|$ ، کاردینالیته مجموعه  $\Theta(T) = \{k \in S_D \mid \text{sp}(B^{E,m}(k)) \subseteq T, \text{sp}(B^{E,m}(k)) \neq \emptyset\}$  است و هم‌چنین  $|\Theta(\emptyset)| = 0$  [۸].

اکنون بازی TU متناظر با وضعیت ریزآرایه MES را می‌توان به صورت زیر تعریف کرد:  
مجموعه بازیکنان، مجموعه ژن‌های  $G$  هستند، و تابع مشخصه بازی، در معادله (۳) آمده است.

ارزش شیپلی یک بازی ریز آرایه از رابطه زیر به دست می آید:

$$\phi_i(V) = \frac{1}{|S_D|} \sum_{S \subseteq N: i \in S} \frac{\bar{\lambda}_S}{|S|} \quad (4)$$

که در آن:  $\bar{\lambda}_S = |\{k \in S_D : sp(B^{E,m}(k)) = S\}|$ .

ارزش شیپلی یک بازی ریز آرایه، امکان محاسبه یک شاخص عددی برای هر ژن را در اختیار ما قرار می دهد، این شاخص، ارتباط هر ژن را تحت یک حالت بیولوژیکی خاص، با در نظر گرفتن رفتارهای بیان ژن های دیگر تحت همان حالت، به دست می آورد. ارزش شیپلی بالای برای هر ژن، در یک بازی ریز آرایه، در واقع نشان دهنده ژن های با تاثیر گذاری مضاعف در بروز بیماری مورد مطالعه می باشد [۷].

#### ۴ تحلیل مقایسه ای ارزش شیپلی

در این قسمت یک روش نوین برای تحلیل رابطه از ژن ها شده تحت دو حالت مختلف، معرفی می شود، در این روش از یک الگوریتم بر مبنای مفهوم بوت استرپ آماری، برای یک تحلیل مقایسه ای از رفتار ژن ها تحت حالت های مختلف، به منظور شناسایی آن دسته از ژن هایی استفاده می شود که نقش مهمی را در ارتباط با بروز یک اختلال بیولوژیکی خاص ایفا می کنند.

بوت استرپ، روشی ساده، اما در عین حال قوی از روش های باز نمونه گیری مونت کارلو است که برای تعیین دقت آماری یا بر آورد کردن توزیع از روی آماره های نمونه می باشد.

بوت استرپ، نمونه گیری ای است که با تبدیل از یک نمونه به دست می آید. یا به عبارت دیگر بوت استرپ نمونه گیری ای است درون یک نمونه. روش بوت استرپ با تکیه بر نمونه شخصی انجام می شود که اغلب، آن نمونه شخصی تنها منبعی است که یک محقق برای تحقیق دارد و این بر اهمیت روش بوت استرپ می افزاید [۹].

این روش که از مدل بازی ریز آرایه نشأت می گیرد، امکان مقایسه شاخص های ارتباط هر ژن (ارزش های شیپلی برای هر ژن در بازی ریز آرایه) میان دو حالت مختلف را فراهم می سازد، هدف از این مقایسه مشخص کردن گروهی از ژن های با تاثیر گذاری مضاعف در بروز یک اختلال بیولوژیکی، و هم چنین شناسایی آن دسته از ژن هایی می باشد که دارای بیشترین ارزش شیپلی می باشند، این ژن ها می توانند به عنوان ژن هایی مطرح شوند که با بیشترین احتمال، ژن های با تاثیر گذاری مضاعف در بروز یک اختلال بیولوژیکی می باشند.

همان طور که عنوان شد، در این روش، از یک الگوریتم اصلاح شده مبتنی بر بوت استرپ، به کار گرفته شده روی آماره های نمونه ارزش شیپلی، به منظور یک تحلیل مقایسه ای از ارزش شیپلی استفاده می شود [۳].

دو گروه از داده های آزمایش ریز آرایه روی مجموعه یکسان از ژن های  $G = \{1, 2, \dots, n\}$ ، که تحت دو حالت مختلف، (به عنوان مثال، سالم و بیمار) جمع آوری شده اند را در نظر بگیرید؛ فرض کنید  $B^1 \in \{0, 1\}^{n \times k}$  و  $B^2 \in \{0, 1\}^{n \times h}$ ، دو ماتریس بولی باشند، به طوری که  $B^1$  توسط یک روش جداسازی از یک مجموعه داده بیان با  $k$  نمونه بیولوژیکی تحت حالت ۱ و  $B^2$  توسط یک روش جداسازی از مجموعه داده بیان از  $h$  نمونه

بیولوژیکی تحت شرایط ۲ به دست آمده باشد؛ فرض  $\bar{v}_1$  و  $\bar{v}_2$  بازی‌های ریزآرایه مطابق با ماتریس‌های بولی  $B^1$  و  $B^2$  باشند  $\phi(\bar{v}^1)$  و ارزش شیپلی روی بازی  $\bar{v}_1$  و  $\phi(\bar{v}^2)$  ارزش شیپلی روی بازی  $\bar{v}_2$  باشد؛ تفاضل مطلق ارزش شیپلی زیر را در نظر بگیرید:

$$\delta_i(\phi(\bar{v}^1), \phi(\bar{v}^2)) = |\phi_i(\bar{v}^1) - \phi_i(\bar{v}^2)| \quad (5)$$

جایی که  $\phi_i(\bar{v}^1)$  ارزش شیپلی ژن  $i$  در بازی ریزآرایه متناظر با ماتریس بولی  $B^1$  و  $\phi_i(\bar{v}^2)$  ارزش شیپلی ژن  $i$  در بازی ریزآرایه متناظر با ماتریس بولی  $B^2$  می‌باشد.

حال الگوریتم ۱ را برای تست این فرض اولیه که آیا هر ژن  $i \in N$ ، ارزش شیپلی یکسانی میان دو حالت بیولوژیکی ۱ و ۲ دارد یا خیر، به کار می‌بریم، در حقیقت ما می‌خواهیم فرض صفر  $\delta_i(\phi(\bar{v}^1), \phi(\bar{v}^2)) = 0$  را در برابر فرض مقابل  $\delta_i(\phi(\bar{v}^1), \phi(\bar{v}^2)) \neq 0$  مورد آزمون قرار دهیم. به طور دقیق‌تر یک روش آزمون، بر پایه روش بوت استرپ غیر پارامتری با جای‌گذاری را ارایه می‌دهیم که قادر می‌باشد فرض اولیه نبود تفاوت میان ارزش شیپلی دو نمونه تصادفی را مورد آزمایش قرار دهد.

## الگوریتم ۱

### ورودی الگوریتم:

- دو ماتریس بولی  $B^1 \in \{0, 1\}^{n \times k}$  و  $B^2 \in \{0, 1\}^{n \times h}$  با  $n, k, h \in \{1, 2, \dots, n\}$
- عدد صحیح  $b$  از باز نمونه‌گیری بوت استرپ با جای‌گذاری.

### خروجی الگوریتم:

- برآورد بوت استرپ از توزیع صفر تفاضل ارزش شیپلی روی  $n$  ژن.
- بردار  $n$  تایی از سطح معناداری برای هر ژن از به دست آمده الگوریتم ۱.

### شروع الگوریتم:

- تفاوت مطلق ارزش شیپلی برای هر ژن  $i$  میان دو حالت ۱ و ۲ را محاسبه کنید:

$$\delta_i(\phi(\bar{v}^1), \phi(\bar{v}^2)) = |\phi_i(\bar{v}^1) - \phi_i(\bar{v}^2)| \quad (6)$$

- برای  $r = 1, 2, \dots, b$  مراحل زیر را انجام دهید:
  - $r$  - امین باز نمونه‌گیری بوت استرپ (با جای‌گذاری) را روی اندیس‌های ستونی  $B^1 \in \{1, 2, \dots, k\}$  محاسبه کنید.

○  $r^2$  - امین باز نمونه گیری بوت استرپ (با جای گذاری) را روی اندیس های ستونی  $\{1, 2, \dots, h\} \in B^r$  را محاسبه کنید.

○ ماتریس های بولی  $B^{s^r, 1} \in \{0, 1\}^{n \times k}$  متناظر با  $r^1$  - امین باز نمونه گیری بوت استرپ و  $B^{s^r, 2} \in \{0, 1\}^{n \times h}$  متناظر با  $r^2$  - امین باز نمونه گیری را تعریف کنید.

○ برای هر  $i \in N$ ، تفاضل مطلق ارزش شیپلی بوت استرپ را از طریق رابطه زیر محاسبه کنید:

$$\beta_i^r(\phi(\bar{v}_r^1), \phi(\bar{v}_r^2)) = |(\phi_i(\bar{v}_r^1) - \phi_i(\bar{v}^1)) - (\phi_i(\bar{v}_r^2) - \phi_i(\bar{v}^2))| \quad (7)$$

به طوری که  $\bar{v}_r^1$  و  $\bar{v}_r^2$ ، به ترتیب بازی های ریز آرایه متناظر با ماتریس های بولی  $B^{s^r, 1}$  و  $B^{s^r, 2}$  می باشند.

پایان حلقه تکرار

• برای هر ژن  $i \in N$ ، سطح معناداری به دست آمده یا مقدار احتمال p-value را از رابطه زیر به دست آورید:

$$p_i = \frac{\text{card}(\{r : \beta_i^r(\phi(\bar{v}_r^1), \phi(\bar{v}_r^2)) \geq \delta_i(\phi(\bar{v}^1), \phi(\bar{v}^2))\}}{b} \quad (8)$$

### پایان الگوریتم

با به کار گیری الگوریتم ۱ روی داده های بیان ژنی، برای هر کدام از ژن ها، یک سطح معناداری آماری یا p-value

به دست می آید؛ در مرحله نهایی، ژن های واجد شرایط:

• P-value تولید شده توسط الگوریتم ۱، کمتر از یک جداسازی از پیش تعریف شده (معمولاً ۰/۰۵)،

• تفاضلات مطلق ارزش شیپلی بزرگ تر از انحراف معیار به علاوه میانگین برای هر گروه؛

برای تحلیل های نهایی انتخاب می شوند.

### مثال ۱:

در این قسمت، یک فهم شهودی از مفهوم ارزش شیپلی در یک بازی ریز آرایه، در قالب یک مثال ارائه می شود،

هم چنین این مثال در فهمیدن این مساله که چگونه روش توصیفی در الگوریتم ۱، برای تحلیل تفاوت های بیان ژن

میان دو حالت عمل می کند، حایز اهمیت است.

ماتریس بولی  $B^1$  از جدول ۱ و بازی ریز آرایه  $(N, \bar{v})$  متناظر با آن را در نظر بگیرید، تابع مشخصه  $\bar{v}$ ،

ارزشی را به هر کدام از زیر مجموعه های غیر تهی  $N$  نسبت می دهد.

جدول ۱. ماتریس بولی  $\bar{B}^1$

	نمونه ۱	نمونه ۲	نمونه ۳	نمونه ۴	نمونه ۵	نمونه ۶	نمونه ۷
ژن ۱	۱	۰	۰	۰	۰	۰	۱
ژن ۲	۰	۰	۰	۱	۰	۱	۰
ژن ۳	۰	۰	۰	۱	۱	۱	۱
ژن ۴	۰	۰	۰	۰	۱	۱	۰
ژن ۵	۰	۰	۱	۱	۰	۰	۰
ژن ۶	۰	۰	۰	۰	۱	۱	۰
ژن ۷	۰	۱	۰	۱	۰	۰	۰
ژن ۸	۰	۰	۰	۰	۰	۱	۱

به عنوان مثال، ائتلاف  $\{1, 3, 8\}$  را در نظر بگیرید، در این صورت داریم:  $\bar{v}(\{1, 3, 8\}) = \frac{2}{7}$ ، چون  $\{1, 3, 8\}$

دو بار به عنوان ائتلاف برنده از میان هفت ستون ماتریس  $\bar{B}^1$  می‌باشد (ستون‌های متناظر با نمونه‌های ۱ و ۷).

ارزش شیپلی بازی ریزآرایه  $(N, \bar{v})$  در ستون دوم جدول ۳ محاسبه شده است (محاسبات مربوط به

ارزش شیپلی از یک بازی ریزآرایه و هم‌چنین محاسبات الگوریتم، توسط کدهای نوشته شده در محیط نرم

افزارهای R و MATLAB، انجام شده است).

حال ماتریس بولی  $\bar{B}^2$  از جدول ۲ را در نظر بگیرید، به‌طوری‌که سطر  $i$  - ام آن، یک جایگشت از عناصر -

$i$ امین سطر جدول ۱ می‌باشد.

جدول ۲. ماتریس بولی  $\bar{B}^2$

	نمونه ۱	نمونه ۲	نمونه ۳	نمونه ۴	نمونه ۵	نمونه ۶	نمونه ۷
ژن ۱	۱	۰	۱	۰	۰	۰	۰
ژن ۲	۱	۱	۰	۰	۰	۰	۰
ژن ۳	۰	۰	۰	۱	۱	۱	۱
ژن ۴	۰	۱	۱	۰	۰	۰	۰
ژن ۵	۱	۰	۱	۰	۰	۰	۰
ژن ۶	۱	۰	۱	۰	۰	۰	۰
ژن ۷	۰	۱	۱	۰	۰	۰	۰
ژن ۸	۱	۰	۱	۰	۰	۰	۰

نکته اینکه تنها عناصر بولی متناظر با سطر سوم بدون هیچ تغییری باقی می ماند، بنابراین هیچ آنالیز آماری ژن - ژنی، قادر به تمایز رفتار ژن ۳ میان دو حالت ۱ و ۲ نمی باشد، اما ما در اینجا به کمک روش نظریه بازی، ثابت می کنیم نقش این ژن در تبیین تفاوت میان دو حالت ۱ و ۲، بسیار تعیین کننده است. با توجه به ماتریس بیان  $\bar{B}^2$ ، می توان به اهمیت ژن ۳ پی برد، این ژن نقش مهمی را در نمونه های ۴، ۵، ۶، و ۷ بازی می کند، زیرا در این نمونه ها، این ژن، تنها ژن با بیان غیرنرمال می باشد.

با توجه به مطالبی که درباره ارزش شیپلی یک بازی ریز آرایه عنوان شد، این ژن باید دارای بیشترین ارزش شیپلی در حالت ۲ باشد، با توجه به محاسبات انجام شده از بازی  $\bar{V}^2$  در ستون دوم جدول ۳، نقش تعیین کننده این ژن در ارتباط با حالت ۱، با ارزش شیپلی بسیار بالای این ژن، در مقایسه با بقیه ژن ها، نشان داده شده است.

**جدول ۳.** ستون های  $\phi_i(\bar{V}^1)$  و  $\phi_i(\bar{V}^2)$  ارزش های شیپلی بازی های ریز آرایه، به ترتیب متناظر با ماتریس های بولی  $\bar{B}^1$  و  $\bar{B}^2$ ، ستون  $\delta_i(\phi(\bar{V}^1), \phi(\bar{V}^2))$  تفاضل مطلق ارزش شیپلی میان دو حالت ۱ و ۲، و ستون پنجم، p-value های به دست آمده از الگوریتم ۱ برای هر کدام از ژن ها را نمایش می دهد.

	$\phi_i(\bar{V}^1)$	$\phi_i(\bar{V}^2)$	$\delta_i(\phi(\bar{V}^1), \phi(\bar{V}^2))$	p-value
ژن ۱	۰/۱۹۰۴۷۶۱۹	۰/۰۵۲۳۸۰۹۵	۰/۱۳۸۰۹۵۲۴	۰/۳۳۶
ژن ۲	۰/۰۶۴۲۸۵۷۱	۰/۰۷۶۱۹۰۴۸	۰/۰۱۱۹۰۴۴۷	۰/۸۵۷
ژن ۳	۰/۱۵۹۵۲۳۸۱	۰/۵۷۱۴۲۸۵۷	۰/۴۱۱۹۰۴۷۶	۰/۰۳۹
ژن ۴	۰/۰۷۶۱۹۰۴۸	۰/۰۷۱۴۲۸۵۷	۰/۰۰۴۷۶۱۹۱	۰/۹۳۳
ژن ۵	۰/۱۷۸۵۷۱۴۳	۰/۰۵۲۳۸۰۹۵	۰/۱۲۶۱۹۰۴۸	۰/۴۰۷
ژن ۶	۰/۰۷۶۱۹۰۴۸	۰/۰۵۲۳۸۰۹۵	۰/۰۲۳۸۰۹۵۳	۰/۶۹۱
ژن ۷	۰/۱۷۸۵۷۱۴۳	۰/۰۷۱۴۲۸۵۷	۰/۱۰۷۱۴۲۸۶	۰/۴۵۸
ژن ۸	۰/۰۷۶۱۹۰۴۸	۰/۰۵۲۳۸۰۹۵	۰/۰۲۳۸۰۹۵۳	۰/۷۰۶

از طرف دیگر، با مقایسه ارزش شیپلی ژن ها تحت دو حالت ۱ و ۲، ملاحظه می شود که ارزش شیپلی ژن ۳ حدود ۴ برابر شده است، این افزایش قابل توجه، با در نظر گرفتن ماتریس بولی  $\bar{B}^2$  و نقش ژن ۳ در این ماتریس، و مقایسه آن با نقش این ژن در ارتباط با حالت ۱ و ماتریس بولی  $\bar{B}^1$ ، قابل توجیه است.

حال الگوریتم ۱ را روی ماتریس های  $\bar{B}^1$  و  $\bar{B}^2$  با تعداد تکرار بازنمونه گیری بوت استرپ ۱۰۰۰ به کار می گیریم، در ستون ۴ از جدول ۳، p-value های به دست آمده از الگوریتم ۱ نشان داده است، در سطح معناداری ۰/۰۵، فرض صفر نبود تفاوت میان ارزش شیپلی بین حات های ۱ و ۲، تنها برای ژن ۳ می تواند رد شود. نکته جالب توجه اینجاست که ژن ۳ دارای رفتار بیان یکسانی میان دو حالت ۱ و ۲ دارد، با این وجود در حالت ۲، این ژن تنها ژن با بیان غیرنرمال برای نمونه های ۴، ۵، ۶، و ۷ می باشد.

در نتیجه نقش ژن ۳ در مشخص کردن تفاوت میان دو حالت، بسیار تعیین کننده است، این واقعیت توسط تفاوت ارزش شیبیلی این ژن، بین دو حالت ۱ و ۲ ارایه شده است، یک تفاوت که منجر به نتایج قابل توجهی در سطح معناداری ۰/۰۵، با توجه به روش آزمون توصیف شده در الگوریتم ۱، شده است. مثال ارایه شده بر این مساله تاکید دارد که به کمک الگوریتم ۱، ما می‌توانیم ژن‌های دارای بیشترین ارتباط ممکن با یک حالت بیولوژیکی خاص را شناسایی کنیم.

## ۵ مطالعه موردی

در این قسمت قصد داریم، در جهت مشخص کردن کارآیی روش نظریه بازی در شناسایی ژن‌های موثر در بروز یک اختلال، روش توصیف شده در الگوریتم ۱ را روی داده‌های آزمایش ریزآرایه، به دست آمده از افراد مبتلا به طیف اختلالات اوتیسم و افراد سالم، به کار ببریم؛ داده‌های این پردازش از سایت جامع بیان ژن (GEO) گرفته شده است، این زیر مجموعه از داده‌های بیان ژن در ۱۷ نمونه بیمار اوتیسمی و ۱۲ نمونه از کودکان سالم تشکیل شده است، تمامی محاسبات مربوط به این پردازش و تحلیل روابط بین ژن‌ها، در محیط نرم افزاری R اجرا و به دست آمده است.

همان‌طور که در مقدمه عنوان شد، از آزمون آماری  $t$  به منظور شناسایی ژن‌هایی استفاده می‌شود که سطح بیان متفاوتی از بقیه ژن‌ها دارند، ژن‌های به دست آمده از این پردازش آماری، ژن‌هایی هستند که با بیشترین احتمال ممکن، مارکرهای ژنی یک اختلال خاص به حساب می‌آیند.

ما در اینجا به منظور دستیابی به نتایج دقیق‌تر، داده‌های اولیه حاصل از آزمایش ریزآرایه را به کمک آزمون  $t$  مورد پردازش قرار داده شده است، بدین ترتیب روش توصیف شده در الگوریتم ۱، روی ژن‌هایی اعمال می‌شود که با بیشترین احتمال ممکن، حاوی سیگنال‌های معنادار بیولوژیکی می‌باشند، پس از اعمال آزمون  $t$ ، ماتریس نهایی شامل سطح بیان از ۴۹۶ ژنی بودند که با استفاده از آزمون  $t$ ،  $p$ -value اولیه کمتر از ۰/۰۱ را احراز نمودند. این ماتریس شامل سطح بیان از ۴۹۶ ژن، در ۲۹ نمونه از افراد بود که از این تعداد ۱۷ نمونه بیمار اوتیسمی (AnR) و ۱۲ نمونه از کودکان سالم (C) می‌باشند. در ادامه در جهت رسیدن به نتایج دقیق‌تر، تغییرات اندکی را مراحل روش اعمال می‌کنیم، به این صورت که ژن‌های با بیان غیر نرمال را به دو دسته ژن‌های با بیان بیش از حد نرمال و ژن‌های با بیان کمتر از حد نرمال تقسیم می‌کنیم، بدین منظور، با توجه به بیان‌های ژن اندازه‌گیری شده در کودکان سالم، روش تفکیک کننده  $m$  را روی بردارهای سطری  $X_i = \{x_{i_1}, x_{i_2}, \dots, x_{i_{29}}\}$  از ماتریس  $B$  به کار می‌بریم (mean مخفف میانگین و  $stdv$  مخفف انحراف معیار می‌باشد):

$$(m(X_i, B))_i = \begin{cases} 1 & X_{ij} \geq \text{mean}[x_i^c] + \text{stdv}[x_i^c] \\ 0 & O.W. \end{cases} \quad (8)$$

بنابراین ماتریس بولی  $B^+$  با ۴۹۶ سطر و ۲۹ ستون با مقادیر  $\{0,1\}$  به دست می آید، هم چنین به طور جداگانه جداسازی زیر را جهت تفکیک ژن های با بیان کمتر از حد نرمال، روی بردارهای سطری  $X_i = \{x_{i1}, x_{i2}, \dots, x_{i29}\}$  از ماتریس  $B$  به کار می بریم:

$$(m(X_i, B))_i = \begin{cases} 1 & X_{ij} \leq \text{mean}[x_i^c] - \text{stdv}[x_i^c] \\ 0 & O.W. \end{cases} \quad (9)$$

بدین ترتیب ماتریس باینری  $B^-$  با ۴۹۶ سطر و ۲۹ ستون با مقادیر  $\{0,1\}$  از  $X$  به دست می آید. با توجه به تمایز نمونه های سالم و بیمار، ماتریس بولی  $B^+$  به دو ماتریس بولی  $B^{c+}$  و  $B^{AnR+}$ ، و ماتریس بولی  $B^-$  به دو ماتریس بولی  $B^{c-}$  و  $B^{AnR-}$  تقسیم می شود. با توجه به رابطه ۳ در بخش ۳، روی ماتریس بولی  $B^{AnR+}$  بازی ریز آرایه  $W^{AnR+}$  و روی ماتریس بولی  $B^{AnR-}$  بازی ریز آرایه  $W^{AnR-}$  را تعریف می کنیم، در واقع یک بازی انتلافی خاص با ۴۹۶ ژن به عنوان بازیکن ها و تابع مشخصه رابطه ۳ به عنوان سهم مشارکت هر زیر مجموعه از ژن ها، روی داده های ماتریس  $B^{AnR-}$  تعریف می کنیم، سپس ارزش های شیپلی متناظر با هر ژن در هر دو حالت را نیز محاسبه می کنیم. در بخش قبل شرح دادیم که هدف از ارایه الگوریتم ۱، مقایسه میان ارزش های شیپلی هر ژن، تحت دو حالت متفاوت، به منظور پیدا کردن ژن هایی می باشد که دارای بیشترین شاخص ارتباط (ارزش شیپلی) میان دو حالت متفاوت می باشند.

بدین منظور ابتدا برای هر ژن، سطح تفاوت مطلق ارزش شیپلی میان دو حالت را محاسبه می کنیم، مقادیر به دست آمده از این مرحله را برای حالت های بیان بیش از حد نرمال و کمتر از حد نرمال، با مقادیر زیر نمایش می دهیم:

$$\begin{cases} \delta_i(\phi(w^{c+}), \phi(w^{AnR+})) = |\phi_i(w^{c+}) - \phi_i(w^{AnR+})| & \text{حالت بیش از حد نرمال} \\ \delta_i(\phi(w^{c-}), \phi(w^{AnR-})) = |\phi_i(w^{c-}) - \phi_i(w^{AnR-})| & \text{حالت کمتر از حد نرمال} \end{cases} \quad (10)$$

در این مرحله ۱۰۹ و ۷۲ ژن دارای تفاضل مطلق ارزش شیپلی، بیش از میانگین به علاوه انحراف معیار، به ترتیب برای بازی های میکروارای  $W^{AnR-}$  و  $W^{AnR+}$  بودند.

حال الگوریتم ۱ را با تعداد تکرار بوت استرپ ۱۰۰۰، جهت مقایسه ارزش های شیپلی هر ژن میان دو حالت سالم و بیمار، در دو وضعیت جداگانه حالت بیش از حد نرمال و حالت کمتر از حد نرمال به کار می گیریم، در این مرحله برای هر ژن یک مقدار احتمال p-value به دست می آید.

با توجه به محاسبات نرم افزاری انجام شده در این مرحله ۸۵ ژن در گروه ژن‌های با بیان بیش از حد منظم و ۸۴ ژن در گروه ژن‌های با بیان کمتر از حد منظم، مقدار احتمال  $P\text{-value} \leq 0.05$  را کسب کردند. به عنوان گام نهایی از مراحل فیلترینگ ژن، با توجه به توضیحاتی که در الگوریتم ۱ عنوان شد، ژن‌های نشان دهنده دو ویژگی:

- $P\text{-value} \leq 0.05$  به دست آمده از الگوریتم ۱،
  - تفاضل مطلق ارزش شیپلی (روابط ۱۰) بزرگتر از میانگین به علاوه انحراف معیار برای هر گروه،
- برای تحلیل‌های نهایی انتخاب شدند (جدول ۴).

**جدول ۴.** ژن‌های مرتبط با بیماری اوتیسم که به کارگیری از الگوریتم ۱ به دست آمده‌اند

---

ADCY7, AK3L1, APEH, ARIH2, ATP2A2, ATP5A1, ATP6AP2, BYSL, C21orf23, C9orf114, CD3Z, CHAF1B, CHST2, CHTF18, CKLFSF3, CLIC3, CREB1, CX3CR1, CXorf12, DCTN5, DDX56, DST, ECRG4, FLJ23441, FLJ38984, FXN, GFM1, GLE1L, GPSN2, GTF3C4, GUSB, GZMB, HERPUD1, HNRPA1, IL18R1, IL2RB, ITGB2, KBTBD4, KCNIP4, KIR3DL1, KIR3DL3, KSP37, LARS2, LECT2, LHFP, LOC129138, LOC143381, LOC284409, LOC285749, LOC339448, LOC440350, LRP1B, MCM3, MORF4L1, NCALD, NKG7, NMT2, NONO, PCDH17, PDPR, PMM1, PRF1, RFC2, RNASEH2A, SAMD3, SH2D1B, SMYD3, SPON2, TRA@, UNC5A, UTP11L, VMD2L2, WDR70, YIPF5, YWHAE, ZNF330, ZNF618, ZW10.

---

نکته جالب توجه اینجاست که ژن c21orf23، که یکی از دو ژنی بود که در آزمون تصحیح فرضیات چندگانه روی ژن‌های به دست آمده از آزمون t، به عنوان ژن متفاوت و عامل بیماری به دست آمده بود، در میان ژن‌های نهایی به دست آمده توسط راه حل نظریه بازی، قرار دارد.

اهمیت بیولوژیکی کاندیداهای نهایی به دست آمده از الگوریتم ۱، و هم‌چنین سادگی و مناسب بودن نظریه بازی ریزآرایه برای تحلیل داده‌های بیان ژنی، می‌تواند با توجه به نتایج اساسی که در تحلیل‌های عملکردی از ژن‌ها به دست آمده، محرز شود.

با توجه به تحقیقاتی که وال و همکارانش [۱۰] در سال ۲۰۰۸، روی داده‌های بیان ژنی کودکان مبتلا به اوتیسم انجام دادند تعداد ۱۹ ژن به طور چشم‌گیری با بیماری‌های عصبی رابطه داشتند که این ۱۹ ژن مشمول در ۷۸ ژن نهایی به دست آمده در الگوریتم ۱ می‌باشند (جدول ۵).

**جدول ۵.** ۱۹ ژن مرتبط با بیماری‌های عصبی در افراد مبتلا به اوتیسم در جدول زیر آورده شده است که تمامی این ۱۹ ژن مشمول در ۷۸ ژن نهایی به دست آمده در الگوریتم ۱ می‌باشند

---

ADCY7, ARIH2, ATP5A1, ATP6AP2, CHAF1B, CLIC3, CREB1, DST, FXN, GFM1, GLE1, IL18R1, IL2RB, ITGB2, NKG7, PRF1, RNASEH2A, TRA@, YWHAE

---

علاوه بر این، براساس مطالعاتی که روی مجموعه داده یکسان، اما توسط یک روش پردازش متفاوت توسط گریگ و همکارانشان [۱۱] انجام پذیرفته است، ۲۰ ژن مشمول در این ۷۸ ژن نیز، به عنوان ژن‌های عامل بیماری اوتیسم شناسایی شدند (جدول ۶).

**جدول ۶.** ژن‌های به دست آمده از تحقیقات آقای گریگ که تمامی آن‌ها، در بین ۷۸ ژن نهایی به دست آمده به کمک روش نظریه بازی قرار دارند

---

CLIC3, CX3CR1, DCTN5, GFM1, GZMB, IL18R1, IL2RB, ITGB2, KIR3DL3, KSP37, NCALD, NKG7, NMT2, PRF1, SAMD3, SH2D1B, SPON2, WDR70, YIPF5, ZNF330

---

### ۶ جمع‌بندی و نتیجه‌گیری

فناوری ریزآرایه، یک ابزار تحلیلی کوچک است که به کمک آن می‌توان بیان ده‌ها هزار ژن را به طور همزمان، میان دو حالت بیولوژیکی مختلف، مورد کاوش قرار داد، این فناوری یک ماتریس از داده‌های بیان ژن را تولید می‌کند که سطرها نشان‌دهنده ژن‌ها و ستون‌ها، نمونه‌های مورد مطالعه را نشان می‌دهند، مقادیر این ماتریس ارزش‌های بیان ژن در نمونه‌های مختلف را نمایش می‌دهند؛ از آزمون‌های آماری متفاوتی برای تجزیه و تحلیل داده‌های بیان ژنی به دست آمده از آزمایشات ریزآرایه استفاده می‌شود، متأسفانه مشکل مهم در انتخاب اینکه کدام نظریه آماری برای تحلیل داده‌های ریزآرایه استفاده شود، این است که اغلب روش‌های آماری استفاده شده، رابطه مناسبی با یک تفسیر بیولوژیکی درست ندارند، در این روش‌ها ژن‌ها معمولاً براساس مقادیر احتمال  $p$ -value شان، رتبه‌بندی می‌شوند؛ از آنجا که هیچ معنای بیولوژیکی‌ای، لزوماً مرتبط با مفهوم  $p$ -value ها نمی‌باشد، از این روش تفسیر از  $p$ -value های جداگانه، درون سیستم‌های بیولوژیکی که در آن‌ها چندین ژن در تعامل متقابل با یکدیگر هستند، کار دشواریست، هم‌چنین پردازش‌های آماری از  $p$ -value ها، نظیر روش‌های تصحیح آزمون چندگانگی، اغلب برای تصحیح کشف‌های نادرست به کار می‌روند، در بسیاری از موارد، این تصحیح‌ها، برای شناسایی سیگنال‌های حاوی اطلاعات مفید بیولوژیکی، محافظه‌کارانه عمل می‌کنند. از طرف دیگر، اکثر روش‌های کلاسیک آماری برای شناخت مارکرهای ژنی یک بیماری، قادر به شناسایی تعداد معدودی از ژن‌های موثر در بروز یک اختلال بیولوژیکی هستند.

با انگیزه نیاز به یک نظریه قوی‌تر جهت شناسایی سیگنال‌ها، یک روش جدید برای آنالیز بیان‌های ژن، بر پایه مفاهیمی از نظریه بازی ارائه شده است، امتیاز مهم از این روش، محاسبه یک شاخص عددی برای ژن‌ها می‌باشد که ارزش شیپلی نام دارد، این شاخص، رابطه هر ژن را تحت یک حالت بیولوژیکی خاص مشخص می‌کند، حال آنکه به طور همزمان، رفتارهای بیان بقیه ژن‌ها را نیز، تحت همان حالت، در نظر می‌گیرد. مقدار بالای ارزش شیپلی، در یک بازی ریزآرایه، نشان‌دهنده ژن‌های دارای تاثیر بیشتر در بروز یک اختلال مورد نظر می‌باشند در این روش، از یک الگوریتم بکاربرده شده روی آماره نمونه ارزش شیپلی، به منظور یک تجزیه

و تحلیل مقایسه‌ای ارزش شیپلی میان دو حالت مختلف، استفاده شد. از نتایج بکارگیری این روش، شناسایی آن‌دسته از ژن‌هایی است که نتایج شاخص ارتباط پایداری در برابر اختلالات بیان ژن از خود نشان می‌دهند. روش‌های نوین برای تحلیل داده‌های آزمایش ریزآرایه، بر پایه مفاهیم نظریه بازی، راه را برای دسترسی محققان بیولوژیست به نتایج دقیق‌تر هموار ساخته است. با توجه به حجم زیاد داده‌های آزمایش، محاسبات روش‌های ذکر شده تنها به کمک برنامه‌های کامپیوتری خاص امکان‌پذیر است. در اکثر مقالات معتبر در این زمینه، محققان با ارایه صورت کلی از روش، تنها به ذکر نتایج به دست آمده از تحقیقات خود اکتفا کرده و از ذکر کدهای کامپیوتری استفاده شده در طی مراحل روش، که هسته اصلی محاسبات را تشکیل می‌دهند، خودداری می‌کنند. در این صورت عملاً امکان استفاده مستقیم محققان دیگر (از جمله محققان داخلی)، جهت به کارگیری این روش‌ها در موارد مشابه به حداقل می‌رسد.

در این مقاله، ما ضمن توصیف کامل یک روش نوین برای تحلیل داده‌های آزمایش ریزآرایه، برای تمامی مراحل روش ذکر شده، از جمله محاسبات آزمون t، روش‌های تفکیک‌سازی اولیه داده‌ها و همچنین محاسبات مربوط به الگوریتم روش، با نوشتن کدهای کامپیوتری، امکان استفاده عملی از روش را روی داده‌های آزمایش ریزآرایه فراهم کرده‌ایم.

این روش با ترکیب مفهوم شاخص ارتباط از نظریه بازی و مفاهیمی از معناداری آماری، گزینشی از ژن‌های با تاثیرگذاری مضاعف در بروز یک اختلال را نتیجه می‌دهد.

در یک مطالعه موردی، روی داده‌های بیان ژنی از افراد دچار طیف اختلالات اوتیسم، نتایج جالب توجهی از کاربرد این روش در شناسایی مارکرهای ژنی به دست آمد. اهمیت بیولوژیکی از کاندیداهای نهائی به دست آمده از روش نظریه بازی و همچنین کارآئی این روش در تحلیل داده‌های بیان ژنی، می‌تواند به توجه به نتایج اساسی به دست آمده محرز شود، در تجزیه و تحلیل‌های عملکردی روی ژن‌های نهائی به دست آمده از بکارگیری الگوریتم ۱ روی داده‌های بیان ژنی کودکان مبتلا به اوتیسم، مشاهده می‌شود که چندین گروه از ژن‌های مرتبط با جنبه‌های مختلف بیماری اوتیسم، و همچنین ژن‌هایی که در نتایج پیشین دانشمندان به عنوان ژن‌های عامل اختلالات طیف اوتیسم شناخته شده است، مشمول در این تعداد ژن به دست آمده به کمک روش نظریه بازی می‌باشند.

## ۷ سپاسگزاری

در پایان وظیفه خود می‌دانم مراتب تشکر و قدردانی خود را از کلیه عزیزانی که در طول مدت آماده‌سازی مقاله با نظرات، پیشنهادات و حمایت‌های بی‌دریغشان، اینجانب را در جهت رسیدن به اهداف تحقیق یاری نمودند، ابراز نمایم. از زحمات استاد گرانقدر جناب آقای دکتر حمیدرضا نویدی که آماده‌سازی این مقاله بدون یاری و راهنمایی‌های ایشان ممکن نبود تشکر و قدردانی می‌نمایم. از جناب آقای دکتر مجید حسن پورعزتی، به خاطر زحمات و مشاوره‌های ارزشمندشان کمال تشکر را دارم، هم‌چنین شایسته است از دوست عزیزم آقای محمد

جواد جعفری که در تحریر و اصلاح کدهای کامپیوتری مقاله با زبان برنامه‌نویسی R، کمک شایانی به بنده کردند صمیمانه تشکر کنم.

## منابع

- [۱] نودسن، س.، نقوی، م.، رامشینی، ح.، (۱۳۸۵). تجزیه داده‌های ریزآرایه DNA، تهران، نقش مهر.
- [۴] نویدی، ح. و همکاران، (۱۳۹۰). مدخلی بر نظریه‌ی بازی‌ها، انتشارات دانشگاه شاهد.
- [۵] عبدلی، ق.، (۱۳۸۶). نظریه بازی‌ها و کاربردهای آن، انتشارات جهاد دانشگاهی واحد تهران.
- [۱۲] آزادی، ن.، (۱۳۸۲). کاربرد مدل‌های آماری در تحلیل داده‌های بیان ژنی به دست آمده از تکنولوژی DNA میکروواری، پایان نامه کارشناسی ارشد، دانشکده بهداشت، دانشگاه علوم پزشکی تهران.
- [2] Moretti, S., Vasilakos, A.V., (2010). An overview of recent applications of game theory to bioinformatics, *Info Sci*, 180(22), 4312-4322.
- [3] Moreeti, S., Van leeuwen, D., et al., (2008). Combining Shapley value and statistics to the analysis of gene expression data in children exposed to air pollution, *BMC Bioinformatics*, 9:361.
- [6] Gilles, P.R., (2009). *The Cooperative Game Theory of Networks and Hierarchies*. Springer.
- [7] Moretti, S., Patrone, F., Bonassi, S., (2007). The clases of microarray games and the relevance index for genes; *Top*, 15:256–80.
- [8] Moretti, S., (2006). *Game Theory applied to gene expression analysis*, PhD Dissertation, University of Genoa, Italy.
- [9] Efron, E., Tibshirani, R. J., (1993). *An Introduction to the Bootstrap*, New York, Chapman & Hall/CRC.
- [10] Wall, D.P., et al., (2009). Comparative analysis of neurological disorders focuses genome-wide search for autism genes, *Genomics*, 93,120-129.
- [11] Gregg, J.P., et al., (2008). Gene expression changes in children with autism, *Genomics* 91, 22-29.
- [13] Fragnelli, V., Moretti, S., (2008). A game theoretical approach to the classification problem in gene expression data analysis. *Computers and Mathematics with Applications* 55.
- [14] Moretti, S.,(2009). *Statistical analysis of the Shapley value for microarray games*, Genoa, Italy, unit of Molecular Epidemiology, National center research institute.
- [15] Peleg, H., Sudhölter, P., (2007). *Introduction to the Theory of Cooperative Games*. Second Edition, Springer.