

کاربرد شبکه ایمنی مصنوعی در داده کاوی و یادگیری ماشین

کاظم نیک فرجام

دانشگاه آزاد اسلامی - واحد بیرجند

گروه کامپیوتر

Farjam@zohatel.com

چکیده: سیستم ایمنی بدن انسان، در بسیاری از زمینه ها مانند بهینه سازی جستجو، امنیت کامپیوتری، قابلیت یادگیری و تشخیص الگو و ... کاربردهای گسترده ای پیدا کرده است. از سوی دیگر یکی از تکنولوژی های مهم تحلیل اطلاعات، فرآیند داده کاوی و کشف دانش است که امروزه در حوزه های مختلف تصمیم گیری کاربرد وسیعی پیدا کرده است. در این مقاله ضمن معرفی مبانی و برخی مفاهیم اساسی شبکه ایمنی مصنوعی و فرآیند داده کاوی، یک الگوریتم یادگیری ماشین جهت ساخت شبکه ایمنی مصنوعی ارائه شده که در فرآیند داده کاوی می تواند مورد استفاده قرار گیرد.

کلمات کلیدی: شبکه ایمنی مصنوعی - داده کاوی و کشف دانش - یادگیری ماشین

۱- مقدمه

توانایی سیستم ایمنی بدن در شناسایی و یادگیری آنتی ژن های جدید، باعث شده که از آن بعنوان الگویی جهت یادگیری ماشین استفاده شود. تکنیک های یادگیری ماشین روشهای محاسباتی هستند که به منظور یادگیری و کشف نکات جدید در مورد داده ها، روی مجموعه ای از داده ها اعمال می شوند. تکنیک های یادگیری ماشین را میتوان به دو دسته تقسیم کرد: الف) تکنیک های یادگیری با سرپرست که در این تکنیک از مجموعه ای داده آموزشی برای کلاس بندی انواع داده های مشابه، به کلاس های از قبل تعریف شده استفاده می گردد و ب) تکنیک های یادگیری بدون سرپرست که بدون استفاده از داده های آموزشی به کشف و شناسایی کلاس های داده کمک می کند. [۱]

اولین تلاش ها در زمینه کاربرد سیستم ایمنی بدن به منظور یادگیری ماشین در سال ۱۹۹۵ انجام گرفت [۵] که نتایج این تلاش، در مقایسه با روش های دیگر یادگیری ماشین بسیار امیدوار کننده بود. در سیستم ایمنی با الهام از مفاهیمی چون سلول B، سطح برانگیختی سلول B، تنوری شبکه ایمنی، همونوع سازی و تکثیر سلولی، فرا جهش جسمی و تولید آنتی بادی جهت شناسایی آنتی ژن جدید، می توان یک مدل جدید جهت یادگیری ماشین ارائه داد. در تنوری شبکه ایمنی، هر سلول B به عنوان یک گره در شبکه عمل می کند که با سلول های دیگر با درجات متفاوتی اتصال دارد. در هر سلول B اطلاعاتی همچون، چگونگی انجام عمل انطباق (اتصال)، کتابخانه ای از ژن ها جهت تولید آنتی بادی جدید، سطح برانگیزش سلول B و ... نگهداری می شود. همچنین با مدل سازی آنتی ژن ها به عنوان داده های جدید که باید توسط شبکه ایمنی یادگیری شوند، آنتی ژن با مولفه انطباق سلول های B (پاراتوپ ها) مقایسه می شود و در صورتیکه انطباق و اتصال بین آنتی ژن و سلول B صورت پذیرد، سلول B تحریک می شود و عمل همونوع سازی و تکثیر سلول و جهش صورت می پذیرد. این فرایند اساس مدل یادگیری ارائه شده در این مقاله می باشد. در ادامه این مقاله ابتدا مروری بر سیستم ایمنی بدن و فرآیند داده کاوی خواهیم داشت. در بخش دوم پارامترهای مهم مورد نیاز برای مدلسازی تکنیک یادگیری معرفی و با الهام از سیستم ایمنی یک الگوریتم یادگیری ماشین ارائه می شود. در انتها نتایج حاصل از اجرای این الگوریتم و تأثیر پارامترهای مهم مدل یادگیری به همراه پیشنهاداتی برای ادامه کار ارائه خواهد شد. [۶] [۷]

۲- سیستم ایمنی بدن و تئوری شبکه ایمنی

مهم ترین جزء در سیستم دفاعی بدن انسان ، گلبول های سفید هستند که نوع خاصی از این گلبول ها بنام لنفوسیت های B و T ، بطور اختصاصی کار اصلی دفاع را انجام می دهند . پس از مشاهده یک عامل غیر خودی (آنتی ژن) توسط لنفوسیت های B پروتئینی به نام آنتی بادی تولید می شود . عمل شناسایی آنتی ژن ، توسط گیرنده های آنتی ژنی که روی سطح لنفوسیت B قرار دارد ، انجام می گیرد . این گیرنده های آنتی ژنی فقط و فقط قادر به شناسایی آنتی ژن های خاصی هستند که از نظر شکل مکمل آن باشد که در این صورت عمل اتصال^۱ صورت می پذیرد و سلول B فعال شده و بصورت کلونی رشد کرده و تکثیر می شود و سلول های B جدیدی بنام سلول B حافظه ای ، تولید می کند که در خون مجلول شده و باعث نابودی آنتی ژن ها می گردند .

به طور خلاصه خصوصیات اساسی سیستم ایمنی عبارتند از : ۱- قابلیت مقیاس پذیری ۲- تنوع ۳- یادگیری و حافظه ۴- خود تنظیمی ۵- فرا دینامیکی و فرا جهشی

آنتی ژن هر چه بیگانه تر باشد (به این معنی که شباهت کمتری با آنتی بادی های سلول B داشته باشد) بیشتر می تواند پاسخ ایمنی سیستم را برانگیزد و سطح برانگیختگی سلول B نیز افزایش می یابد . یک آنتی بادی زمانی فعال می شود که سطح انگیزش آن از یک سطح آستانه ای بیشتر باشد . سیستم ایمنی باید مقدار زیادی از آنتی ژن هایی که قبلاً ندیده است را شناسایی کند و تصمیم گیری کند که چگونه به آنها پاسخ دهد . سیستم ایمنی آنتی ژن ها را بخاطر می سیارد و اگر آنها را دوباره شناسایی کند سریع تر دفع می کند . همچنین در حول هر سلول B یکسری گوی تشخیص بنام ARB وجود دارد و مقدار تطابق و اتصال هر سلول B با آنتی ژن با فاصله d نشان داده می شود (فاصله d مقدار تفاوت سلول B با آنتی ژن است) و بعنوان مثال در نمایش باینری آنتی ژن ها و سلول B بر اساس تفاوت تعداد بیت های آنها می توان فاصله d را تعیین نمود به شکل زیر دقت کنید :

آنتی ژن : ۱۰۱۱۱۰۰۱	
آنتی بادی : ۱۰۱۰۱۱۰۰	

XOR : ۰۰۰۱۰۱۰۱	
میزان شباهت = ۳ (برابر با تعداد یک های جواب)	

محتوای داخلی هر ARB شامل تعدادی آنتی بادی ، تعدادی آنتی ژن ، سطح برانگیختگی و نوع انطباق و اتصال می باشد . یک شبکه از ARB ها تشکیل یک شبکه ایمنی (شبکه ایدیوتایپ) می دهد که بین ARB ها در شبکه نینک های با درجات متفاوت وجود دارد . سطح برانگیختگی هر ARB نسبت معکوس با فاصله میانگین از مجموعه آنتی ژن ها دارد . بر اساس تئوری شبکه ایمنی و با توجه به اینکه پاسخ اولیه سیستم ایمنی در مقابل آنتی ژن با پاسخ های ثانویه سیستم ایمنی متفاوت است می توان یک الگوریتم یادگیری ماشین با سرپرست در این فرآیند پیدا و مدل سازی کرد . بدین منظور در ابتدا مجموعه ای از داده ها جهت آموزش شبکه ایمنی به شبکه وارد می شود همچنین مجموعه ای از داده های آموزشی بعنوان آنتی ژن در نظر گرفته می شود که باید توسط شبکه ایمنی یاد گرفته شوند .

یک آنتی ژن بطور تصادفی از مجموعه آنتی ژن های آموزشی به شبکه وارد شده و با بخشی از ARB های شبکه جهت ایجاد اتصال و انطباق مقایسه می شود . بر اساس میزان شباهت (فاصله d) بین آنتی ژن و سلول B بعضی از سلول B ها تحریک شده و شروع به همونوع سازی و تکثیر سلولی می کنند (میزان شباهت باید از مقدار خاصی (آستانه) بیشتر باشد تا سلول B تحریک شود . همچنین سلول B باید به اندازه کافی تحریک شود (با یک مقدار آستانه ای مقایسه شود) تا بتواند عملیات کلونی (همونوع سازی و تکثیر و جهش) را انجام دهد .

۱- binding

۲- Artificial Recognition Ball

سلول های B تولید شده نیز باید بتوانند با شبکه ARB ها یکپارچه شوند که برای یکپارچه سازی سلول های B جدید به شبکه ، عمل ترکیب بین ۲ تا ARB نزدیک بهم صورت می گیرد . حال چنانچه هیچ ARB ای در شبکه نتواند آنتی ژن را شناسایی کند (عمل انطباق و اتصال جهت شناسایی آنتی ژن و میزان شباهت پایین باشد) از آن آنتی ژن بعنوان یک الگوی جدید در شبکه یاد خواهد شد یعنی به عنوان یک ARB جدید به شبکه متصل می شود . در این روش جمعیت ARB ها ممکن است بصورت نمایی رشد کند و مشکل انفجار جمعیتی رخ دهد بنابراین در هر مرحله ، درصدی از ضعیف ترین ARB را (با سطح انگیزش پایین) از جمعیت حذف خواهیم کرد برای محاسبه سطح برانگیختگی ARB و سلول B از هر یک از معادلات زیر ممکن است استفاده گردد :

$$\delta t = c \left[\sum_{j=1}^N m(ARB, x e_j) - K_1 \sum_{j=1}^N m(ARB, x p_j) + K_2 \sum_{j=1}^N m(ARB, y_i) \right] - K_3$$

در این معادله ARB جاری ، m تابع انطباق (شباهت) و N تعداد آنتی بادی ها و n تعداد آنتی ژن ها C, K_1, K_2, K_3 نرخ ثابت (که بر اساس تعداد مقایسات در هر بار و نرخ تولید سلول B های تحریک شده محاسبه می شود)

$x e_j =$ زمین اپیتوپ سلول B و $x p_j =$ زمین پاراتوپ سول B و $y_i =$ زمین آنتی ژن
روش دوم بر اساس میزان فاصله اقلیدسی بین ARB ها داریم :

$$W_{ij} = \exp\left(-\frac{d'_{ij}}{2\delta_i}\right)$$

سطح بر انگیزش (i) ARB برابر است با : (رابطه ی شماره دو) $\delta t_i = \frac{\sum_{j=1}^n W_{ij}}{\delta_i}$ و n تعداد آنتی ژنها و δt شعاع گوی ARB

$$\delta_i^* = \frac{\sum W_{ij} d_i^*}{\sum W_{ij}}$$

با ماکزیمم کردن سطح انگیزش در هر ARB و مشتق گیری از δt داریم :
dij فاصله بین آنتی ژن z تا مرکز (i) ARB (که با XOR کردن و شمارش تعداد یک ها می تواند انجام پذیرد)
روش دوم با استفاده از فاصله اقلیدسی نرمال که فاصله ها و وزن ها بین ۰ تا ۱ نگاشت شده است سطح انگیزش برابر است با :

$$\delta t = \sum_{x=0}^n (1 - pd_x) + \sum_{x=0}^n (1 - dis_x) - \sum_{x=0}^n dis_x$$

$Pd =$ فاصله نرمال شده بین آنتی ژن و ARB است $0 \leq pd \leq 1$ و $dis_x =$ فاصله نرمال شده بین X مین همسایه با ARB و n تعداد آنتی ژن ها

۳- داده کاوی و کشف دانش در بانک های اطلاعاتی

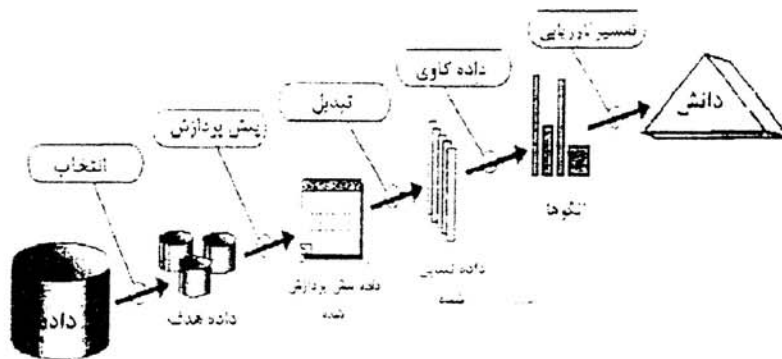
کشف دانش و داده کاوی ، هنر و علم کشف و استخراج روابط جدید ، مفید و سودمند در داده هاست . پروسه تحلیل داده هاست از منظرهای متفاوت و خلاصه کردن آنها به اطلاعات مفید ، اطلاعاتی که می توانند برای افزایش سودآوری یا کاهش هزینه ها مفید واقع شوند . بصورت خیلی خاصتر ، میتوان کشف دانش را پروسه کشف خودکار الگوها ، قوانین و سایر محتویات با قاعده ای دانست که بصورت تولیدی در حجم انبوهی از داده ها حضور دارند . داده کاوی نیز به کشف الگوها در یک مجموعه داده گفته می شود . اغلب کشف دانش و داده- کاوی را بصورت مشابه در نظر می گیرند ، اما در واقع داده کاوی ، هسته ی مرکزی پروسه ی کشف دانش و قوانین است . نمونه ای از کاربردهای داده کاوی شامل کشف کلاهبرداریها در معاملات با کارت اعتباری ، تشخیص کاراکتر در خواندن خودکار کدهای فشرده و پیش بینی فعالیتهای ترکیبی در زمینه کشف داروهاست .

داده های خام ذخیره شده در بانک های اطلاعاتی ، به ندرت مستقیماً مورد استفاده قرار می گیرند . در کاربردهای عملی ، داده ها در فرم های تغییر یافته متناسب با نیاز کاربر به او ارائه می شوند . در چنین حالتی نیز ، باز کاربر باید داده ها را کم و بیش بصورت دستی ، و با استفاده از پردازشگرهای پرس و جو ، تحلیل کند . این کار شاید برای داده های دارای حجم کم امکان پذیر باشد ولی بر ای حجم عظیمی از اطلاعات قطعاً ، نه ، شدنی است و نه پذیرفته شده . به چیزی مانند یک تحلیل گر خودکار داده نیاز خواهد بود و این دقیقاً همان چیزی است که فعالیتهای کشف دانش و داده کاوی در اختیار می گذارند . این ابزارها علاوه بر نیازمندیهای مشخص تعریف شده برای کاربر ، می توانند حقایق و روابطی درون مجموعه داده ها برای او روشن کند که نه به ذهن او خطور کرده اند و نه به دنبال آنها بوده است و در عین حال جالب و کاربردی اند . با استفاده از این روابط و حقایق انسان می تواند عوامل موثر در کار و تجارت را به بهترین نحو برنامه ریزی و مدیریت کند .

نوفاً سیستم های کشف دانش و داده کاوی بصورت همه منظوره طراحی نمی شوند و اغلب نحوه طراحی و کارشان وابسته به کاربرد است .

۱-۳ : ساختار پروسه ای سیستم های کشف دانش

کشف دانش بصورت یک پروسه انجام می گیرد و یک فعالیت یک مرحله ای نیست . به همین دلیل همانند همه ی پروسه ها محیط خاص خود را دارد و در فازهای متعدد با محدودیت ها و مفروضات مشخص انجام می شود . شکل ۱ مراحل پروسه مزبور را نشان می دهد .



(شکل ۱- فازهای مختلف پروسه کشف دانش)

همانگونه که در شکل دیده می شود ، منبع اصلی اطلاعات یک بانک اطلاعاتی است که حاوی حجم عظیمی از داده هاست . هدف یافتن الگوهایی در این بانک است . این داده ها در طی فاز انتخاب فیلتر شده ، فقط داده های خاص و مدنظر برای کشف الگو انتخاب می شوند . نتیجه حاصل از این فاز ، داده های هدف هستند . در فاز پیش پردازش ، نویزهای موجود از داده های هدف حذف می شوند و تا آنجا که ممکن است شامل موارد نادرست و غلط ، بی دقت ، مبهم ، ناسازگار ، استثنایی یا مفقود باشد .

در فاز بعد داده های پیش پردازش شده بگونه ای تغییر می یابند که متناسب با وظایف داده کاوی در فاز بعد باشند . چگونگی تغییرات روی داده ها در این فاز کاملاً وابسته به کاربرد است و متناسب با نوع عملکرد فاز داده کاوی می باشند . از آنجایی که فاز داده کاوی فقط روی زیر مجموعه ای از فیلدهای رکوردها متمرکز می شود ، اغلب در این فاز شاهد کاهش تعداد فیلدها در رکوردهای داده هستیم . حتی تغییرات و ترکیباتی نیز در فیلدهای باقیمانده ممکن است صورت گیرد تا بهترین فرم برای عرضه به فاز داده کاوی را تولید کرده باشیم . در فاز داده کاوی ، از ابزارها و وظایف خاصی برای بیرون کشیدن الگوهای موجود در داده ها استفاده می شود . نتیجه ی این فاز همان گونه که در شکل مشخص است ، الگوهای ممکن خواهند بود . الزاماً همه ی این الگوها ، کاربردی و مفید نیستند . لذا در فاز بعد که فاز تفسیر و ارزیابی است ، فقط الگوهایی پذیرفته خواهند شد که کارایی قابل ملاحظه

ای داشته باشند و بقیه حذف خواهند شد. نتیجه‌ی کار دانش مکشوف خواهند بود. به منظور تولید الگوهای یخته تر و دارای ارزش بیشتر، ممکن است در پروسه‌ی مزبور، حلقه‌هایی ایجاد گردید و یک یا چند فاز به صورت تکراری انجام شوند. حتی می‌توان روی چند مجموعه‌ی داده این پروسه را تکرار کرد.

۲-۳: نمایش دانش مکشوف

دانش مکشوف به روش فوق را با تکنیک‌های متفاوتی می‌توان نشان داد. از جمله قوانین استنتاج (If-then)، درختهای تصمیم، جداول، دیاگرام، تصاویر، عبارتهای تحلیلی و مانند اینها.

درخت‌های تصمیم، یک روش کاملاً مناسب برای این کار است. تبدیل نحوه‌ی بازنمایی از درخت‌های تصمیم به قوانین استنتاج نیز کار بسیار ساده‌ای است. یکی از محاسن این روش، امکان استفاده از تکنیک‌های یادگیری ماشین برای تولید خودکار آن است. در ادامه به رویکردی از تکنیک‌های درخت‌های تصمیم که از شبکه ایمنی استفاده می‌کند، اشاره خواهد شد.

دانش مکشوف توسط سیستم‌های کشف دانش، اغلب با شاخص‌های احتمالاتی خاصی همراه هستند، مانند شاخص قطعیت یا پوشش. این شاخص‌ها بیشتر روی میزان اطمینان و قطعیت الگوها نظر می‌دهند.

دانش مکشوف، علاوه بر نمایش برای کاربر می‌تواند به عنوان یک منبع اطلاعاتی برای سیستم‌های هوشمند دیگری نظیر سیستم‌های خبره نیز بکار رود. برای این منظور، باید به فرم مورد نیاز برای آن سیستم تبدیل شود. کاربرد دیگر این دانش، اضافه شدن آن به پایگاه دانش سیستم اصلی است، که می‌تواند در پروسه‌های بعدی کشف دانش برای اخذ الگوهای دیگر مفید واقع شود.

وظایف مرتبط با بحث داده کاوی، کاملاً وابسته به کاربرد است و به نوع دانشی که قرار است اخذ شود و نوع منبع اطلاعاتی ربط دارد. ولی در حالت کلی می‌توان موارد زیر را نام برد.

۱- تمایز: تشخیص ویژگی‌های جداگانه و منحصر به فردی است که می‌توانند دسته بندی داده‌ها را به کلاسهای متفاوت امکان پذیر سازند.

۲- دسته بندی اطلاعات: تشخیص تعلق یک عنصر به یکی از کلاسهای از پیش تعریف شده تعریف کلاسهای متصور نیز معمولاً از روی مقادیر یک یا چند فیلد خاص انجام می‌شود.

۳- تعیین کلاسترها: توصیفی برای کلاسهای مفروض تولید می‌شود و اینکه چگونه و چه داده‌هایی به هر کلاستر تعلق دارند. مانند فواصل اقلیدسی مختلف که برای تعیین تعلق عناصر به کلاس خاصی استفاده می‌شود.

۴- خلاصه سازی: یافتن خصوصیات مشترک و لازم برای گروه‌هایی از داده‌ها و قراردادنشان در یک فرم فشرده را خلاصه سازی می‌گوییم.

۵- مدلسازی وابستگی: توصیف وابستگی‌های مهم بین داده‌ها در یک مجموعه داده. وابستگی به معنای آن است که مقدار یک داده را می‌توان با یک ضریب اطمینان مشخص از مقدار داده دیگری نتیجه گرفت. مثلاً $A \rightarrow B$ با $CF = 0.91$. مجموعه‌ای از وابستگی‌های بین داده‌ها را اغلب با گراف وابستگی نشان می‌دهند.

۶- تشخیص انحراف یا تغییر: گونه‌ای از دانش قابل استخراج از مجموعه‌ی داده‌ها، تغییرات و میزان انحرافی است که در شاخص‌های آماری داده‌ها، نسبت به تغییرات داده‌ها ایجاد می‌شود. مانند تغییر در میانگین، انحراف معیار، واریانس و سایر شاخص‌های آماری.

اغلب در مباحث مربوط به داده کاوی، دسته بندی از کلیدی ترین فعالیتهای متصور است. از آنجایی که در بیشتر بانکهای اطلاعاتی و منابع دانش بشری، اجزای اطلاعاتی به فرم متن هستند، دسته بندی در این مقوله نیز بیشتر از جنس دسته بندی متنی خواهد بود. دسته بندی متنی، علاوه بر کاربرد در تحلیل بانکهای اطلاعاتی، می‌تواند در تحلیل و داده کاوی‌های مبتنی بر وب نیز مورد استفاده قرار گیرد. بارشد سریع اطلاعات بر خط، دسته بندی متنی یکی از کلیدی ترین تکنیک‌ها برای کار با داده‌های متنی است. از دسته بندی متنی می‌توان برای یافتن اطلاعات مورد علاقه در صفحات وب یا راهنمایی کاربران برای داشتن جستجوی بهتر استفاده کرد. دسته بندی

متون با استفاده از دست ، کاری بسیار مشکل و در اغلب موارد نشدنی است و آموزش مدل‌های دسته بندی یک ضرورت انکار ناپذیر بنظر می رسد .

۴- معرفی الگوریتم یادگیری :

- ۱- جمعیت آنتی ژن ها را بارگذاری کن
- ۲- جمعیتی از سلول های b را بطور تصادفی مقدار دهی اولیه کن و بعنوان شبکه سلول های B در نظر بگیر
- ۳- تا زمانیکه شرایط خاتمه رخ نداده مراحل زیر را تکرار کن
 - ۱-۳) از جمعیت آنتی ژن ها ، بطور تصادفی یک آنتی ژن انتخاب کن
 - ۲-۳) یک نقطه در شبکه سلول B بطور تصادفی بمنظور درج آنتی ژن انتخاب کن
 - ۳-۳) درصدی از سلول های B محلی (اطراف) نقطه درج را انتخاب کن
 - ۴-۳) برای هر سلول B محلی عملیات زیر را تکرار کن
 - ۱-۳-۴) آنتی ژن را با سلول B محلی مقایسه و میزان انطباق (شباهت) را تعیین کن
 - ۲-۳-۴) اگر میزان انطباق از میزان خاصی (مقدار آستانه) بیشتر بود سطح تحریک و برانگیختگی سلول B را تعیین کن (طبق معادله)
 - ۳-۳-۴) اگر سطح برانگیختگی کافی بود (از مقدار آستانه بیشتر بود) عملیات همونوع سازی و تکثیر و رشد سلولی را انجام بده
 - ۳-۵) اگر هیچ سلول B ای نتوانست به آنتی ژن انطباق پیدا کند یک سلول B جدید تولید کن بطوریکه بتواند آنتی ژن به آن متصل شود (الگوی جدید)
- ۳-۶) سلول های B را بر اساس سطح انگیزش مرتب کن
- ۳-۷) ۵٪ از ضعیف ترین سلول های B را از جمعیت حذف کن (مکانیزم کنترل جمعیت)
- ۳-۸) به اندازه ۲۵٪ جمعیت اولیه سلول B جدید تولید کن
- ۳-۹) از جمعیت جدید تولید شده ، ۵٪ سلول B جهت الحاق و یکپارچه شدن به شبکه بطور تصادفی انتخاب کن .

۳-۱۰) پایان حلقه

۴- پایان الگوریتم

۴-۱ : الگوریتم یادگیری دوم :

- ۱- شبکه ایمنی مصنوعی را مقدار دهی اولیه کن
 - ۲- برای هر آنتی ژن عملیات زیر را تا رسیدن به شرایط پایانی تکرار کن
 - ۱-۲) به هر ARB آنتی ژن را ارائه کن
 - ۲-۲) سطح انگیزش ARB را محاسبه کن
 - ۳-۲) تعدادی سلول B به ARB تخصیص بده (بر اساس میزان برانگیزش آن)
 - ۴-۲) ضعیف ترین ARB ها (که هیچ سلول B ای ندارند) را حذف کن
 - ۵-۲) بین مابقی ARB ها عمل همونوع سازی و تکثیر سلولی و جهش را انجام بده
 - ۶-۲) ARB های جدید را با شبکه یکپارچه کن
- ۳- پایان
- عملیات یک پارچه سازی بدین شکل است که اگر فاصله بین دو ARB کمتر از یک ϵ بود آن دو ARB را بهم پیوند بزن (محتوای داده های آنتی بادی دو ARB را از یک نقطه تقاطع (crossover) ، ترکیب کن) .

۵- نتیجه گیری و پیشنهادات

بعضی از نکات مهم در این الگوریتم عبارتند از :

- سطح برانگیزش سلول B ، زنده بودن یا نبودن آن سلول را در شبکه تعیین می کند .
- تعیین مقادیر آستانه ای مناسب برای انطباق و برانگیختی از موارد اساسی در الگوریتم می باشد .
- با توجه به خواص فراجهشی * و فرادینامیکی بودن سیستم ایمنی که بر علیه هجوم آنتی ژن ها عمل می کنند ممکن است مکانیزم کنترل جمعیت را با مشکل مواجه کنند . از این رو یک الگوریتم بهبود یافته دیگر در این بخش که از مفاهیم گوی تشخیص مصنوعی ARB ها (هر ARB شامل مجموعه ای از سلول B ها ، آنتی بادی ها و آنتی ژن ها است) و در آن کنترل جمعیت و کنترل اندازه شبکه به شکل بهتری است ارائه می شود .
- در شبکه ایمنی مصنوعی با ایجاد شبکه ای از سلول های B و آنتی بادی ها ، یک آنتی ژن جدید بعد از ورود به شبکه با آنتی بادی های شبکه مقایسه می شود . چنانچه از قبل وجود داشته باشد (بتواند به یک آنتی بادی بچسبد) شناسایی شده و چنانچه موجود نباشد بعنوان یک الگوی جدید (یک گره) به شبکه ایمنی مصنوعی اضافه می شود . عمل انطباق ، تولید شبکه و رشد آن و یادگیری همه و همه مدلی است از آنچه در واقعیت رخ می دهد . از این مدل یادگیری می توان در فرآیند داده کاوی بخصوص در کاربردهای مبتنی بر کلاس بندی متن استفاده کرد . پیشنهاد می شود با توجه به خاصیت فراجهشی و فرادینامیکی بودن شبکه از آن برای کاوش وب و دسته بندی سایت ها بر اساس محتویاتشان استفاده شود .

مراجع :

- [۱] Hossein Ali Abbas, Ruhul Amin Saker, Charles S,(۲۰۰۴), "Data Mining A Hurestic Approach" , University of New South Wales, Australia . IGP Press.
- [۲] D.E Goldberg, " Genetic and Evolutionary Algorithms Come of Age ",
- [۳] Dasgupta, D(۱۹۹۹). Infprmation processing in immune system. Springer-Verlag.
- [۴] Goldberg, D.(۱۹۹۰)." Genetic algorithm in search, optimization and machine learning". Addison Wesley.
- [۵] Holland,j. (۱۹۹۸)." Adaptation in natural and artifical systems|". MIT Press.
- [۶] Cooke, D. and Hunt, j.(۱۹۹۵)." Recognising promoter sequence using an Artificial Immune System". AAAI press.
- [۷] Baoding liu (۲۰۰۲), " Theory and practice of uncertain programming",A Springer-Verlag Company.
- [۸] Stephanie Forrest, (۲۰۰۲), "Information Immune System" and Genetic Algorithms", LLC Press. .
- [۹] Ansrzej Osyczka , (۲۰۰۲) " Evolutionary Algorithms for single and multicriteria Design Optimization " , A Springer – Verlag Company .

- [۱۰] اکبرزاده توتونچی محمد رضا ، یغمایی حسین ، " استفاده از سیستم ایمنی مصنوعی فازی برای امنیت شبکه های کامپیوتری " ، پنجمین کنفرانس سراسری سیستم های هوشمند مهر ماه ۱۳۸۲ مشهد تهران .
- [۱۱] علی یاری زنوز نوید ، " چکیده ایمونولوژی " ، سال ۲۰۰۰ ویرایش چهارم انتشارات تیمور زاده