

مجله علوم پزشکی مدرس: آسیب‌شناسی زیستی
دوره ۱۳، شماره ۱: از ۱-۱۶
بهار ۱۳۸۹

ارایه روشی به منظور طبقه‌بندی انواع لوسمی براساس تئوری بازی‌های مبتنی بر همکاری و مقدار شیپلی

عاطفه ترکمن^۱، نصراله مقدم‌چرکری^{۲*}، مهناز آقایی‌پور^۳، کامبیز بدیع^۴

- ۱- کارشناس ارشد، گروه مهندسی صنایع (گرایش فناوری اطلاعات)، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، تهران، ایران
- ۲- استادیار، گروه مهندسی کامپیوتر، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، ایران
- ۳- دانشیار، مرکز تحقیقات انتقال خون، سازمان انتقال خون ایران، تهران، ایران
- ۴- دانشیار، گروه مهندسی دانش و سامانه‌های هوشمند، پژوهشکده IT، مرکز تحقیقات مخابرات ایران، تهران، ایران

پذیرش مقاله: ۸۸/۱۱/۰۷

دریافت مقاله: ۸۸/۰۶/۲۵

چکیده

هدف: طبقه‌بندی انواع مختلف لوسمی حاد براساس تئوری بازی‌های مبتنی بر همکاری و مقدار شیپلی
مواد و روش‌ها: در این تحقیق از نمونه داده‌های بیمار، جمع‌آوری شده از آزمایش‌های فلوسایتومتری در سازمان انتقال خون ایران، استفاده شد. ۳۰۴ نمونه بیماری مختلف در ۸ طبقه لوسمی حاد بررسی شد. در ابتدا این نتایج به صورت عددی بودند. در مرحله بعد براساس حد آستانه تعریف شده، داده‌ها به شکل بولین تبدیل شد. سپس براساس تئوری بازی مبتنی بر همکاری و مقدار شیپلی، وزن‌های خاصی به این داده‌ها تخصیص داده شد. بدین ترتیب، نمونه‌های مختلف بیماری از یکدیگر جدا و امکان طبقه‌بندی انواع لوسمی حاد حاصل شد (مرحله یادگیری).
در مرحله تشخیص، با استفاده از معیارهای شباهت، شباهت نمونه جدید مورد مطالعه با نمونه‌های آموزشی ارزیابی شد و نوع لوسمی نمونه مورد آزمایش، آشکار شد.
نتایج: دقت طبقه‌بندی لوسمی براساس تئوری بازی مبتنی بر همکاری، ۹۶/۳ درصد بود که نشان می‌دهد روش پیشنهادی از دقت قابل توجهی در تفکیک درست طبقه‌ها برخوردار است. به منظور سنجش کارایی روش پیشنهادی، نتیجه با شبکه عصبی که از جمله الگوریتم‌های یادگیری کاراست، مقایسه شد. نتایج حاصل از دقت طبقه‌بندی با شبکه عصبی پایه شعاعی، ۹۱/۸۰ درصد بود.
نتیجه‌گیری: با توجه به نتایج حاصل، روش پیشنهادی در طبقه‌بندی لوسمی حاد، امیدوارکننده است. از این رو می‌تواند به هماتولوژیست‌ها و پزشکان در تشخیص دقیق تر انواع لوسمی کمک کند تا تصمیمات مناسب‌تری در مورد نحوه تشخیص و درمان بگیرند.

کلیدواژه‌گان: تئوری بازی، بازی مبتنی بر همکاری، مقدار شیپلی، طبقه‌بندی، لوسمی، فلوسایتومتری

۱- مقدمه

در بیماری سرطان یا چنگار (واژه فارسی سرطان)، موارد به سایر سلول‌ها هجوم می‌برند. کنترل سلول‌ها توسط سلول‌های بدن به صورت کنترل نشده‌ای تکثیر یافته و در اکثر عوامل رشد، پیام‌های محیطی و برخی پروتئین‌ها و

*نشانی مکاتبه: تهران، دانشگاه تربیت مدرس، دانشکده مهندسی برق و کامپیوتر، گروه مهندسی کامپیوتر، کدپستی: ۱۴۱۱۷۱۳۱۱۶
Email: charkari@modares.ac.ir

این امر در تصمیم‌گیری دقیق‌تر، یاری می‌رساند و خطاهای احتمالی را کاهش می‌دهد.

اغلب در سیستم‌های تصمیم‌ساز، الگوریتم‌های مختلف یادگیری ماشین (Machine learning) به کار گرفته می‌شود. یادگیری ماشین، شاخه‌ای از هوش مصنوعی (Artificial intelligence) است که با طرح و به کارگیری الگوریتم‌ها به کامپیوترها این امکان را می‌دهد که کارایی خود را براساس یادگیری، بهینه نمایند. یادگیری ماشین با مباحث داده کاوی (Data Mining)، آمار (Statistics)، استدلال استنتاجی (Induction reasoning) و تشخیص الگو (Pattern recognition) مرتبط است. کاربردهای یادگیری ماشین شامل پردازش زبان طبیعی (Natural language processing)، تشخیص الگوی نحوی (Syntactic pattern recognition)، موتورهای جستجو (Search engines)، تشخیص پزشکی (Medical diagnosis)، بیوانفورماتیک (Bioinformatics)، طبقه‌بندی داده‌ها (Data classification) و بازی‌ها (Games) است. الگوریتم‌های متداول در این زمینه شامل یادگیری با نظارت (Supervised learning) و بدون نظارت (Unsupervised)، یادگیری نیمه‌نظارتی (Semi-supervised learning)، یادگیری تقویتی (Reinforcement learning) و یادگیری برای یادگیری (Learning to learn) است. در حوزه پزشکی، یادگیری ماشین به منظور استنتاج قوانین تشخیص و پیش‌بینی بیماری‌ها و حل مشکلات آن‌ها کاربردهای زیادی دارد.

طبقه‌بندی داده‌ها که یکی از مباحث مهم علوم تصمیم‌گیری و یادگیری ماشین است، در اغلب سیستم‌های پشتیبانی تصمیم، استفاده می‌شود. به طور کلی، طبقه‌بندی کارا سبب جستجوی سریع‌تر و راحت‌تر داده‌ها می‌شود. بسیاری از مسایل نظیر تشخیص پزشکی، شناسایی تصویر (Image recognition) و ارزیابی اعتبار (Credit evaluation) از روش‌های مختلف طبقه‌بندی استفاده می‌کند [5]. تحقیق حاضر نیز با هدف طبقه‌بندی خودکار بدخیمی‌های لوسمی با به کارگیری تئوری بازی مبتنی بر همکاری (Cooperative Game Theory) و

پیام‌رسان‌های سلولی صورت می‌پذیرد. جهش‌هایی که منجر به تغییر هریک از عوامل مؤثر در سرنوشت سلول، می‌شود باعث به هم خوردن نظم دقیقی که در تنظیم رشد و تکثیر و تمایز سلول‌ها وجود دارد می‌شود و می‌تواند منجر به بروز سرطان شود [1]. امروزه، بیش از ۱۰۰ نوع مختلف از سرطان‌ها در دنیا شناخته شده است که از میان آن‌ها، لوسمی (Leukemia) یکی از انواع شایع و مهلک سرطان‌ها است. سرطان خون یا لوسمی، بیماری پیشرونده و بدخیم اعضای خون‌ساز بدن به‌ویژه مغز استخوان است که با تکثیر و تکامل ناقص سلول‌های خون و پیش‌سازهای آن در خون و مغز استخوان ایجاد می‌شود. ابزارهای مختلف به‌منظور شناسایی این نوع بیماری وجود دارد. آزمایش خون و شمارش گلبول‌های سفید خون، آزمایش‌های فلوسایتومتری (Flowcytometry) روی خون و مغز استخوان، نمونه‌برداری از گره‌های لنفوی و نمونه‌برداری از مایعات بدن از جمله موارد تشخیص این نوع بیماری‌ها است. در این میان، فلوسایتومتری یک روش توانا، سریع و دقیق برای تشخیص و پیگیری بدخیمی‌های خون است. این روش با خصوصیات منحصر به فردی که دارد، سلول‌های مختلف را براساس آنالیز چند پارامتری در مدت زمان کوتاه تشخیص داده و این مسئله باعث برتری آن نسبت به سایر روش‌ها است [۲-۴]. در کنار داده‌های حاصل از آزمایش‌های فلوسایتومتری، همواره وجود یک متخصص برای ارزیابی، تفسیر داده‌ها و تشخیص تعلق آن‌ها به طبقه بیماری مربوط به خود، ضروری است. ارزیابی درست و دقیق داده‌ها و گرفتن تصمیم مناسب از مهم‌ترین فعالیت‌ها در تشخیص سرطان‌ها است.

امروزه، توسعه سیستم‌های تشخیصی که به پزشکان در تشخیص انواع سرطان‌ها و تصمیم‌گیری دقیق‌تر در درمان این نوع بیماری‌ها کمک نمایند، متداول شده است. همین امر باعث کاهش مرگ و میر ناشی از این نوع بیماری‌ها است. در این جا، اهمیت سیستم‌های پشتیبان-تصمیم‌ساز پزشکی (Medical Decision Support System: MDSS) (تصمیم‌یار پزشک)، تبیین می‌شود. این سیستم‌ها به پزشکان و متخصصین

در ادامه نمونه‌هایی از کاربرد این نوع سیستم‌ها را در تشخیص پزشکی، بیان خواهیم نمود.

گارسیا-گومز (Garcia-Gomez) و همکاران یک MDSS به‌منظور کمک به متخصصین رادیولوژی در شناسایی تومورهای بافت نرم طراحی نمودند که براساس معماری توزیع شده با سه گره [نود (Node)] مشخص کار می‌کند؛ رابط بصری رادیولوژیست (Radiologist Visual Interface)، سیستم اطلاعات و سرویس‌های پشتیبان-تصمیم مبتنی بر وب MDSS [۷]. دیگری نیز به‌منظور کمک به بیماران آسمی توسط سفیون (Sefion) و همکاران ارائه شد که به پزشکان در کنترل این بیماری مزمن کمک می‌کند [۸] باکار (Bakar) و همکاران، یک MDSS به‌منظور شناسایی انواع مشکلات شنوایی طراحی نمودند که به متخصصین در کنترل و حل مشکلات شنوایی کمک می‌کند [۹]. سیستم تصمیم یار پزشکی دیگری نیز به‌منظور مدیریت سرطان‌های خون توسط چائو (Chae) و همکاران پیاده‌سازی شده است [۱۰]. در این سیستم، به‌منظور تشخیص و پیش‌بینی سرطان‌های خون، چهار روش مختلف طبقه‌بندی به‌کار گرفته شده است؛ روش مبتنی بر قواعد (Rule-Based Reasoning: RBR)، روش مبتنی بر نمونه‌های ضبط شده (Case-Based Reasoning: CBR)، شبکه عصبی و آنالیز تمایز (Discriminant Analysis: DA).

۲-۲ مروری اجمالی بر روش‌های طبقه‌بندی

انواع سرطان

تحقیقات زیادی در زمینه تشخیص و طبقه‌بندی انواع سرطان‌ها صورت گرفته است. بسیاری از آن‌ها از الگوریتم‌های یادگیری به‌منظور طبقه‌بندی داده‌ها بهره می‌جویند. به‌عنوان مثال در مراجع [۱۱-۱۴] هریک به نوعی به‌منظور طبقه‌بندی داده‌ها از الگوریتم‌های یادگیری استفاده نموده‌اند. در این مراجع، به‌منظور کاهش اندازه داده‌های آموزشی و تعریف زیر مجموعه‌های مناسب از متغیرهای ورودی، از رویکرد استخراج ویژگی (Feature selection) استفاده شده است. استخراج

مقدار شیبیلی (Shapley Value) انجام شده است. این امر می‌تواند به هماتولوژیست‌ها در تشخیص دقیق‌تر و سریع‌تر این نوع بیماری‌ها کمک نماید.

طبق مطالعات انجام شده، استفاده از روش‌هایی نظیر شبکه عصبی و ماشین بردار-پشتیبان (Support vector machine: SVM) به‌منظور طبقه‌بندی داده‌ها در حوزه پزشکی، بسیار متداول است. همان‌طور که می‌دانیم، از مزایای عمده این روش‌ها، توانایی انجام محاسبات غیرخطی، توانایی شناسایی و دسته‌بندی الگوها و موارد دیگر است. اما از سوی دیگر؛ این نوع روش‌ها قابلیت توسعه نداشته و بنابراین در مواجهه با یک طبقه جدید، آموزش دوباره سیستم ضروری است. از این رو، روش ارائه شده در این مقاله با تکیه بر تئوری بازی‌های مبتنی بر همکاری قابلیت توسعه را به سیستم طبقه‌بند (Classifier)، افزوده است.

در این مقاله ابتدا کارهای انجام شده به‌طور اجمالی مرور خواهد شد. سپس به‌صورت مختصر به تئوری بازی مبتنی بر همکاری اشاره می‌شود. پس از آن مدل پیشنهادی براساس تئوری بازی مبتنی بر همکاری بررسی خواهد شد.

۲-۲ مروری اجمالی بر کارهای انجام شده

۲-۱-۱ کارهای انجام شده در زمینه استفاده از

MDSSها

همان‌گونه که بیان شد، MDSSها زیرمجموعه سیستم‌های پشتیبان تصمیم‌ساز است که به پزشکان در راستای تصمیم‌گیری و تشخیص دقیق بیماری‌ها کمک می‌نمایند. ایده سیستم‌های پشتیبان تصمیم‌ساز برای اولین بار توسط وُن‌نیومان (VonNeuman) و مورگن‌استرن (Morgenstern) در سال ۱۹۴۴ مطرح شد. در سال ۱۹۸۰ بحث MDSSها عنوان شد و در نتیجه جدیدی را در مبحث سیستم‌های پشتیبان-تصمیم‌ساز گشود [۶]. این سیستم‌ها در تشخیص و پیگیری بیماری‌ها به پزشکان کمک می‌نمایند و خطاهای احتمالی را کاهش می‌دهد.

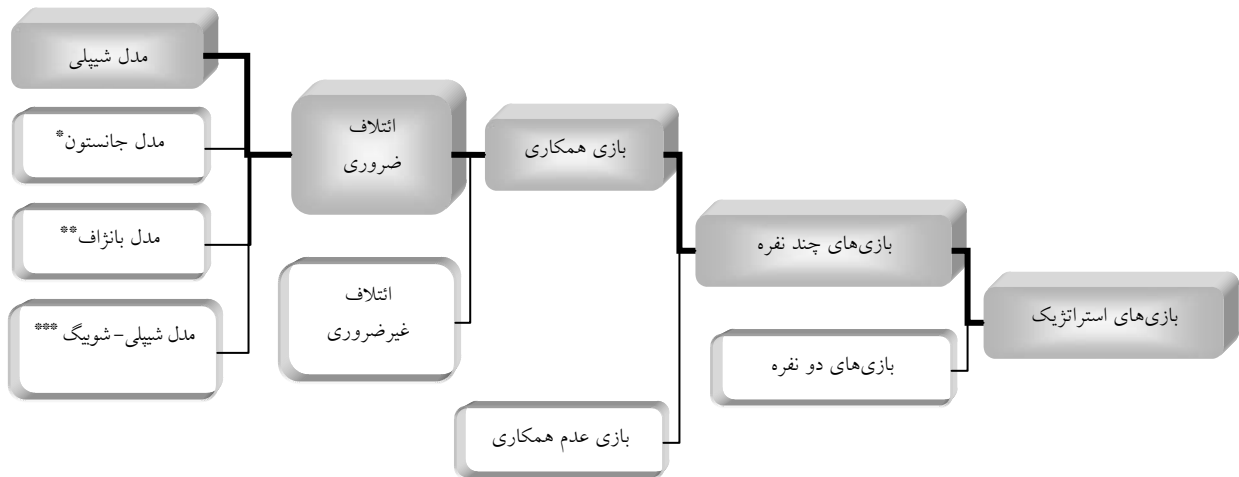
طرح‌ریزی کرده‌اند [۱۳]. قوانین انجمنی به‌منظور کاهش ابعاد پایگاه داده (Database) سرطان سینه و شبکه عصبی برای طبقه‌بندی هوشمند، استفاده شده‌اند. فراگنلی (Fagnelli) و همکاران با اعمال سه روش درخت C4.5، پرسپترون چند لایه (Multi Layer Perceptron: MLP) و Naïve Bayes روش‌های نشانگرهای (Markers) خاص تومورهای سینه، انواع مختلف سرطان سینه را طبقه‌بندی کردند [۱۴]. به‌منظور طبقه‌بندی انواع لوسمی براساس بیان آنالیز داده‌ای ژن‌ها میچی (Michi) و همکاران از روش یادگیری CBR استفاده کردند [۱۵]. در این روش، نمونه‌های قبلی به خاطر سپرده شده و در مواجهه با نمونه جدید از تجربیات گذشته استفاده می‌شود. تشخیص مناساز گره‌های لنفاوی (Lymph Nodes Metastases: LNM) قبل از عمل جراحی بسیار مشکل و پیچیده است به همین دلیل، براندنبورگر (Brandenburger) و همکاران از روش شبکه عصبی مصنوعی برای تشخیص این موضوع در بیماران مبتلا به سرطان معده، استفاده کردند [۱۶]. برانزی (Branzei) و همکاران نیز به‌منظور طبقه‌بندی سرطان سینه از روش SVM استفاده نمودند [۱۷]. از آنجایی‌که تمایز بین تومورهای خوش‌خیم و بدخیم، بسیار مشکل و زمان‌بر است، به‌کارگیری یک شیوه خودکار برای تشخیص این موارد، ضروری به‌نظر می‌رسد. در این مقاله مجموعه‌ای از تصاویر سینه با استفاده از روش نمونه برداری آسپیراسیون با سوزن باریک (Fine Needle Aspiration: FNA) در نظر گرفته شده است. برای تشخیص تفاوت بین طبقه‌های سالم و بیمار از روش ماشین بردار-پشتیبان استفاده شده است.

همچنین در تحقیقات دیگری نیز از توری بازی‌ها به‌منظور تعامل بین ژن‌ها، سلول‌ها و محیط پیرامون آن‌ها، استفاده شده است. به‌عنوان مثال، کوهن (Cohen) و همکاران برای مدل‌سازی جهش، هجوم (Onset)، پیشرفت (Progression) و رقابت بین سلول‌های سرطانی (Immune competition cancer cells)، از تئوری بازی اتفاقی (Stochastic game theory) استفاده کرده‌اند [۱۸]. بازی‌های تکاملی (Evolutionary game) به‌منظور ارزیابی جمعیت سلول‌های تومور با تعامل بین سلولی، توسط محققان دیگر مطرح شده است [۱۹]. به‌کارگیری روش

ویژگی، گزینش زیر مجموعه‌ای از متغیرهای ورودی است که کارایی طبقه‌بند را افزایش می‌دهد. بلومو (Bellomo) و همکاران به‌منظور طبقه‌بندی الگوی سرطان سینه (Breast Cancer)، ابزار شناسایی بهتری به‌نام سیتولوژی آسپیراسیون با سوزن باریک (Fine Needle Aspiration Cytology: FNAC) - یکی از روش‌های تشخیص سرطان سینه - را به‌کار گرفتند [۱۱]. در این راستا، شیوه رگرسیون وقتی چند متغیره (Multivariate Adaptive Regression Splines: MARS) برای تشخیص زیر مجموعه مناسب از متغیرهای ورودی از یک مدل شبکه عصبی، به‌کار گرفته شده است. بنابراین، دقت طبقه‌بندی در روش شناسایی (Methodology) ترکیبی ارایه شده، افزایش می‌یابد. باج (Bach) و همکاران از سه مدل شبکه عصبی تئوری تشدید وقتی (Adaptive Resonance Theory Based Neural Network: ART) شبکه خود سازمانده (Self Organizing Map Based Neural Network: SOM) و شبکه پس انتشار خطا (Back Propagation Neural Network: BPN) به‌منظور طبقه‌بندی انواع سرطان سینه استفاده کردند [۱۲]. در این میان روش BPN نسبت به دو روش دیگر عملکرد بهتری دارد. BPN یکی از انواع شبکه‌های عصبی است که براساس یادگیری سعی و خطا (Try and error) عمل نموده و سعی دارد با حداقل کردن مقدار یک تابع خطا، ورودی‌های داده شده را بر خروجی‌ها منطبق نماید. یادگیری در این دسته از شبکه‌ها به‌صورت با نظارت است.

ART یکی از انواع شبکه‌های بدون نظارت است و به‌گونه‌ای طراحی شده است که به کاربر این امکان را می‌دهد که درجه شباهت الگوهایی که در یک خوشه قرار می‌گیرند را با تنظیم پارامترهای هوشیاری کنترل نماید. یادگیری SOM نیز بدون نظارت است و از روش یادگیری رقابتی برای آموزش استفاده می‌کند. واحدهای پردازشگر آن در یک فرآیند یادگیری رقابتی نسبت به الگوهای ورودی منظم می‌شود. منصور (Mansury) و همکاران سیستم خبره‌ای (Expert system) برای آشکارسازی (Detection) سرطان سینه براساس قوانین انجمنی (Association Rules: AR) و شبکه عصبی،

انجام حرکت در یک بازی برای برآوردن سود بیشتر است [۲۲، ۲۳]. بازی استراتژیک یا بازی به شکل طبیعی، یکی از انواع بازی‌ها است که در آن هر بازیکن حرکتی را برای بازی خود انتخاب می‌کند و تصمیمات به‌صورت همزمان گرفته می‌شود. بازیکنان علاوه بر حرکت‌های خود به تصمیمات دیگر بازیکنان نیز توجه دارند. بنابراین در این نوع بازی‌ها، تلاش تمام بازیکنان در راستای انتخاب بهترین استراتژی است. این بازی‌ها به دو دسته، بازی‌های دو نفره و چند نفره تقسیم می‌شوند. شکل ۱، طبقه‌بندی این نوع بازی‌ها را نشان می‌دهد. تمایز این طبقه‌ها در چگونگی شکل‌دهی ارتباط بین بازیکنان است [۲۴، ۲۵].



* Johnston
 ** Banzhaf
 *** Shapley- Shubik

شکل ۱ طبقه‌بندی بازی‌ها

تعیین پاداش هر یک از اعضای ائتلاف مطرح می‌شود. خطوط مشکی پر رنگ در شکل ۱، مسیر هدف در تحقیق حاضر را مشخص می‌سازد. در ادامه، بازی‌های مبتنی بر همکاری و مدل شیلی به‌طور مفصل بررسی می‌شود. به‌طور کلی، بازی مبتنی بر همکاری شامل مجموعه‌ای از بازیکنان و تابع مشخصه (Characteristic function) است.

بازی‌های تکاملی در یک مدل مبتنی بر عامل (Agent-based) تومور مغز تحقیق دیگری در زمینه طبقه‌بندی سرطان‌ها است [۲۰]. در تحقیق دیگری نیز روشی مبتنی بر تئوری بازی برای حل مشکل طبقه‌بندی در بیان آنالیز داده‌های ژن‌ها ارائه شده است [۲۱].

۳- مروری اجمالی بر تئوری بازی‌های مبتنی بر همکاری

همان‌طور که می‌دانیم، تئوری بازی، شاخه‌ای از ریاضیات کاربردی است که به مطالعه موقعیت‌های استراتژیک بازیکنان یک بازی می‌پردازد. هدف این تئوری، تخمین راه‌های مختلف

در بازی‌های چند نفره، ممکن است هر یک از بازیکنان با یکدیگر به رقابت بپردازند، این نوع بازی‌ها یا مبتنی بر عدم همکاری است یا این‌که با یکدیگر در حل یک مشکل مشارکت می‌کنند (بازی‌های مبتنی بر همکاری). بازی‌های مبتنی بر همکاری به دو دسته ائتلاف ضروری و غیرضروری تقسیم می‌شوند. در بازی‌های ائتلاف ضروری، مدل‌هایی به‌منظور

شیپلی اختصاص داده شده به هریک از بازیکنان به صورت زیر است [۲۶، ۲۵]:

(۱)

$$\phi_i(v) = \frac{1}{n!} \sum_{\pi} (v(P(\pi; i) \cup \{i\}) - v(P(\pi; i)))$$

که π جایگشت بازیکنان و $P(\pi; i)$ مجموعه‌ای از بازیکنان است که مقدم بر بازیکن i در جایگشت π هستند. اگر دو بازیکن i و j متقارن باشند، یعنی رابطه

$$v(S \cup \{i\}) = v(S \cup \{j\}), \forall S \subseteq N \setminus \{i, j\}$$

آن‌گاه $\phi_i(v) = \phi_j(v)$ خواهد بود.

اگر بازیکن i سود حاشیه‌ای نداشته باشد

$$\phi_i(v) = 0, \forall S \subseteq N \setminus \{i\}, v(S \cup \{i\}) = v(S),$$

خواهد بود.

بنابراین با توجه به مطالب مذکور، به منظور حل مسایل و مشکلات فعالیت‌های مشارکتی می‌توان از تئوری بازی و مزایای مهم مقدار شیپلی در تقسیم عادلانه پاداش بازیکنان، بهره جست.

۴- روش ارایه شده براساس تئوری بازی

مبتنی بر همکاری

۴-۱- چارچوب مفهومی روش پیشنهادی

همان‌گونه که بیان شد، هدف، ارایه سیستمی است که بتواند در تشخیص و طبقه‌بندی انواع سرطان خون به هماتولوژیست‌ها و متخصصین امر یاری رسانده و خطاهای احتمالی در این زمینه را کاهش دهد. در این سیستم به منظور طبقه‌بندی این نوع بیماری‌ها، از تئوری بازی‌های مبتنی بر همکاری استفاده شد.

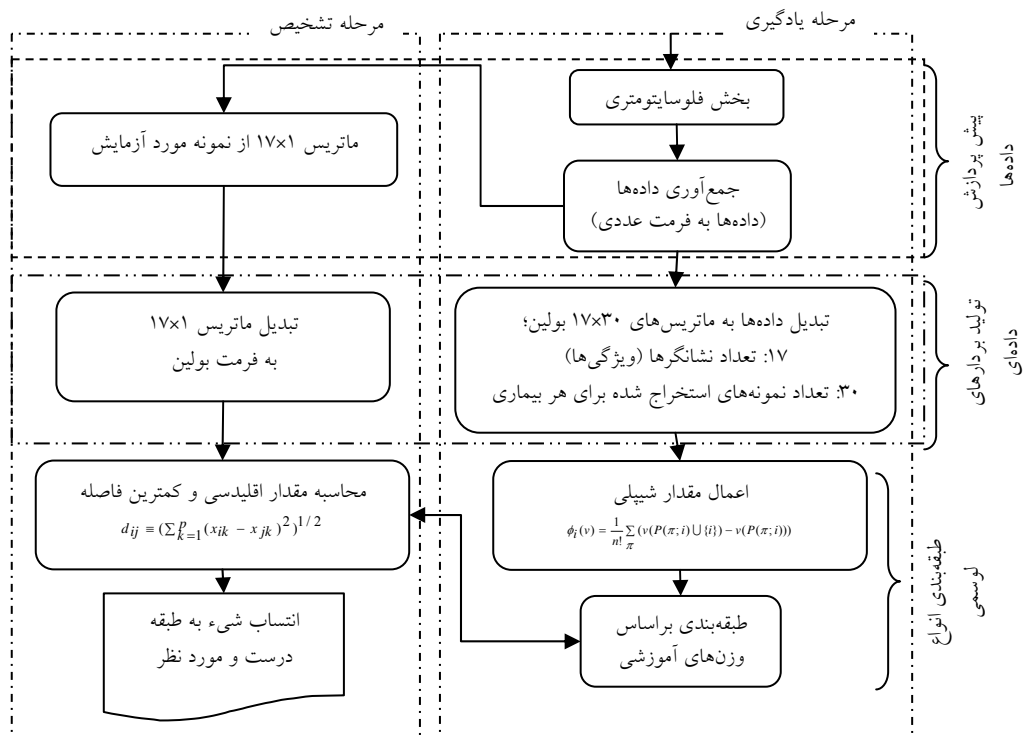
برای این منظور، روش پیشنهادی در دو مرحله مجزای (۱) یادگیری (Training phase) و (۲) مرحله تشخیص نمونه‌های جدید (Diagnosis phase or Test phase)، بررسی شد (شکل ۲).

تابع مشخصه، مقدار ایجاد شده توسط زیر مجموعه‌های بازیکنان مختلف در بازی را مشخص می‌کند. فرض کنید $N = \{1, 2, \dots, n\}$ مجموعه (متناهی) از بازیکنان و $v(S)$ تابع مشخصه این بازی است که بیان می‌کند v با هر زیر مجموعه S از N ارتباط دارد. می‌توان $v(S)$ را به مقداری تفسیر نمود که زمانی ایجاد می‌شود که اعضای S با یکدیگر متحد شده و با هم تعامل دارند. مقادیر، بسته به شرکت بازیکنان در به دست آوردن سود بیشتر، تعیین می‌شوند.

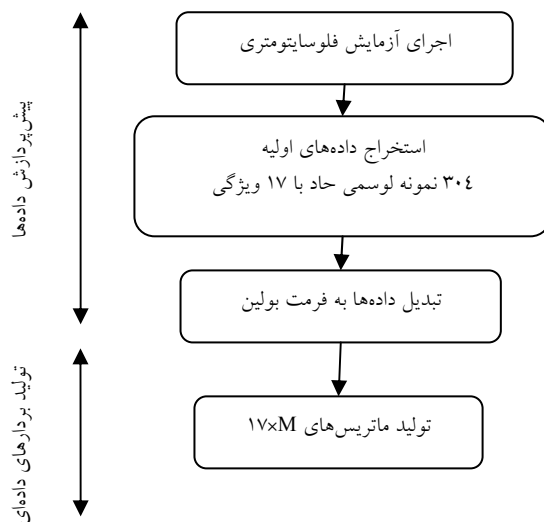
یکی از مفاهیم مهم در بازی‌های مبتنی بر همکاری، مفهوم سود حاشیه‌ای (Marginal contribution) است. در واقع، در یک بازی مبتنی بر همکاری، همواره این سؤال مطرح می‌شود که چگونه سود حاصل از بازی بین تمامی بازیکنان مختلف تقسیم شود. در واقع، سود حاشیه‌ای عبارتست از مقداری که در صورت ترک یک بازیکن خاص از صحنه بازی از کل مقدار ایجاد شده، کسر می‌شود.

بازی مبتنی بر همکاری با سودمندی انتقال‌پذیر (Transferable Utility game: TU-game)، با زوج مرتب (v, N) بیان می‌شود. N مجموعه بازیکن و v تابع مشخصه $(v : 2^N \rightarrow \mathbb{R})$ است. به طوری که $v(\emptyset) = 0$. $v(S)$ به عنوان بالاترین ارزش یا پس‌انداز هزینه (Cost savings)، برای ائتلاف S در نظر گرفته می‌شود. راه حلی برای بازی‌های $(TU\text{-game}(N, v))$ تابع v است که به هر بازیکن، یک بردار سود $v(v)$ را تخصیص می‌دهد. مقدار شیپلی که اولین بار توسط شیپلی مطرح شد [۲۶]، یکی از بهترین و متداول‌ترین راه حل‌ها در بازی‌های TU-game است. این راه حل، به هر بازیکن، در همه حالت‌های ممکن (جاب‌جایی بازیکنان) متوسط سود حاشیه‌ای مختص خودش را عطا می‌کند [۲۱].

فرض کنید $G = (N, v)$ یک بازی است. در این بازی، هریک از بازیکنان به صورت $i \in N$ فرض می‌شوند. مقدار



شکل ۲. نمای انتزاعی مدل ارائه شده به منظور طبقه‌بندی لوسمی‌ها



شکل ۳. مراحل پیش‌پردازش و تولید بردار داده‌ای (M بیانگر تعداد نمونه‌های هر بیماری است).

۴-۲- مرحله یادگیری

در مرحله یادگیری مراحل مختلف پیش‌پردازش داده‌ها، تولید بردارهای داده‌ای از بیماری‌های مورد نظر و طبقه‌بندی انواع لوسمی مد نظر است. شکل ۳ مراحل مختلف پیش‌پردازش و تولید بردار را نشان می‌دهد. در ادامه توضیحات بیشتری در مورد این مراحل ارائه می‌شود.

۴-۲-۱- پیش‌پردازش داده‌ها

در این بخش مجموعه داده‌های اولیه تهیه می‌شود. این مجموعه داده باید به طریقی انتخاب شوند که جامعیت لازم را در نشان دادن انواع بیماری‌های مورد نظر در حوزه لوسمی داشته باشد. برای تأمین این هدف بایستی از مجموعه داده‌های استاندارد استفاده شود. بنابراین، در این تحقیق از نمونه داده‌های بخش فلوسایتومتری سازمان انتقال خون ایران استفاده شد. در ابتدا توضیح مختصری راجع به نحوه انجام آزمایش فلوسایتومتری که در راستای تولید داده‌های اولیه این تحقیق استفاده شد، بیان می‌شود.

در این آزمایش، ابتدا نمونه خونی مورد نظر وارد دستگاه فلوسایتومتری شد. سلول‌های موجود در سوسپانسیون از مرکز یک جریان پوششی (Sheath) محصور و به‌صورت یک ردیف مرکزی با سرعت ثابت از لوله‌ای شیشه‌ای عبور کرد و در این مسیر تحت تابش یک یا چند منبع نوری که غالباً لیزر است قرار گرفت. پس از برخورد نور لیزر با سلول‌ها، بازتاب نور لیزر در جهات مختلف پراکنده شده و توسط آشکارکننده‌هایی (Detectors) که در زوایای مختلف قرار داشت، برای کسب اطلاعات جمع‌آوری شد [۲]. به این ترتیب تقریباً همه مولکول‌های موجود در قسمت‌های مختلف سلول قابل ردیابی خواهند شد. مولکول‌های سطحی سلول‌ها تحت نام عمومی CD نشانگر (Cluster Differentiation Marker) شناخته می‌شود. CD نشانگرهای هر سلول بیانگر وضعیت آن سلول است. هر نوع بدخیمی لوسمی با حضور نشانگرهای خاصی در کنار هم ایجاد می‌شود. تمرکز محققان حاضر در این بررسی روی ۸ نوع لوسمی حاد در ایران است.

با بررسی نتایج حاصل از فلوسایتومتری بدخیمی‌های ALL (Acute Lymphoblastic Leukemia) و AML (Acute Myeloid Leukemia)، ۱۷ نشانگر به‌عنوان داده ورودی (ویژگی) در نظر گرفته شد (جدول ۱).

جدول ۱ CD۱۷۱ نشانگر که به‌عنوان ویژگی‌های اولیه در نظر گرفته می‌شود.

ویژگی	
HLA-DL	۱
CD3	۲
CD20	۳
CD5	۴
CD10	۵
CD13	۶
CD33	۷
CD19	۸
CD34	۹
CD14	۱۰
CD45	۱۱
CD117	۱۲
CD64	۱۳
CD71	۱۴
CD41	۱۵
CD61	۱۶
گلیکوفورین A	۱۷

در مجموع، ۳۰۴ نمونه از انواع مختلف لوسمی حاد با ۱۷ ویژگی از سازمان انتقال خون ایران جمع‌آوری شد که ۸۰ درصد داده‌ها در مرحله یادگیری (به‌طور متوسط در هر نوع بیماری ۳۰ نمونه) و مابقی در مرحله تشخیص استفاده خواهد شد.

۴-۲-۲- تولید بردارهای داده‌ای

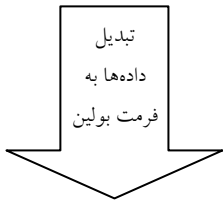
داده‌های حاصل از آزمایش فلوسایتومتری به فرمت عددی است. در این بخش داده‌ها به فرمت بولین (Boolean format) تبدیل شد و با به‌کارگیری نظر هماتولوژیست، مقدار A به‌عنوان حد آستانه انتخاب شد. مقادیر بیشتر از A در تبدیل بولین، یک و در غیر این صورت ۰ در نظر گرفته شد. در بررسی حاضر مقدار A برابر با ۲۰ در نظر گرفته شده است. تعیین میزان CD

به این ترتیب داده‌ها در قالب ماتریس‌های بولین $17 \times M$ قرار داده شد. M در مرحله یادگیری معادل ۳۰ در نظر گرفته شد (نمونه‌های در نظر گرفته شده برای هر نوع بیماری).

نشانه‌ها، کار بسیار مهمی است و مقدار آن‌ها بر طبقه‌های داده تأثیر می‌گذارد. جدول ۲ تبدیل نشانه‌های عددی به فرمت بولین را نشان می‌دهد.

جدول ۲ تبدیل داده‌ها به فرمت بولین

CD33	CD13	CD19	CD10	CD5	CD20	CD3	HLA/DL	CD نشانه‌ها	
۵۴/۳	۳۳/۹	۲/۱	۳	۷/۴	۵/۱	۱۰/۱	۶۳/۲		بدخیمی‌ها
									AML M4M5
۴۴/۷	۳۴/۶	۱/۲	۰/۵	۱	۱/۵	۲/۱	۳۲/۱		AML M7



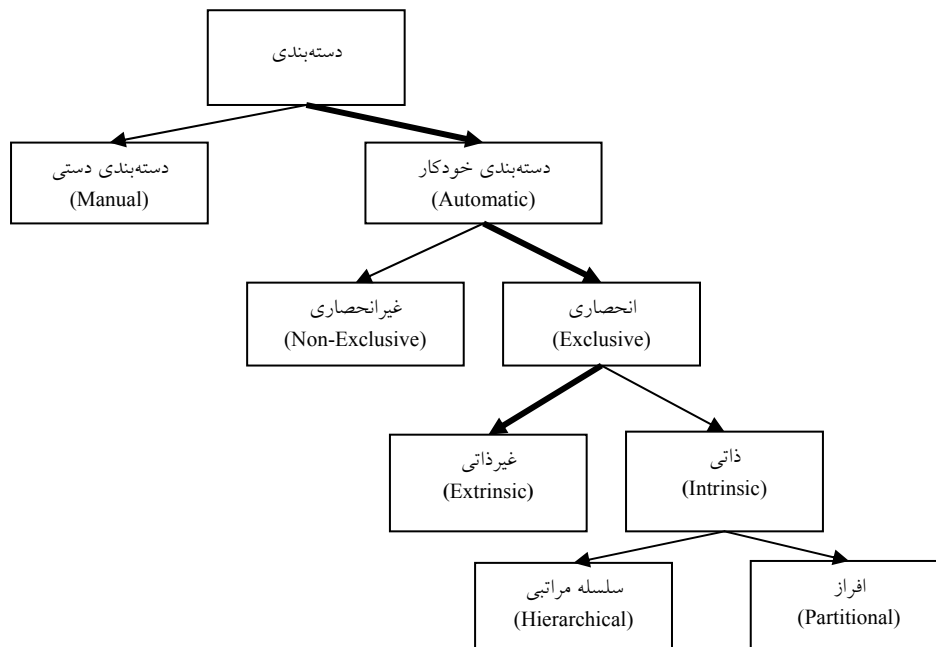
CD33	CD13	CD19	CD10	CD5	CD20	CD3	HLA/DL	CD نشانه‌ها	
۱	۱	۰	۰	۰	۰	۰	۱		بدخیمی‌ها
									AML M4M5
۱	۱	۰	۰	۰	۰	۰	۱		AML M7

فقط به یک زیر مجموعه یا خوشه تعلق دارد؛ در حالی که در طبقه‌بندی غیرانحصاری یک شیء می‌تواند به طبقه‌های مختلف از اشیا نسبت داده شود. در طبقه‌بندی ذاتی از ماتریس همسایگی، برای اجرای طبقه‌بندی استفاده می‌شود. به این نوع طبقه‌بندی، «یادگیری بدون نظارت» نیز گفته می‌شود. در طبقه‌بندی ذاتی، اغلب طبقه‌بندی براساس اطلاعات کیفی که در ذات شیء وجود دارد، انجام می‌شود. در مقابل، طبقه‌بندی غیرذاتی با قراردادن برجسب‌هایی روی اشیا، آن‌ها را طبقه‌بندی می‌نماید. مسیرهای پر رنگ در شکل ۵ بیان‌کننده نحوه طبقه‌بندی این تحقیق است.

۴-۲-۳- طبقه‌بندی انواع لوسمی

هدف نهایی مرحله یادگیری طبقه‌بندی و مجزا نمودن بدخیمی‌های مذکور است.

همان‌طور که می‌دانیم، طبقه‌بندی به معنای اختصاص یک شیء در رده یا به عبارتی طبقه مناسب خود است [۳]. در حالت کلی می‌توان طبقه‌بندی را به صورت شکل ۴ مطرح نمود. طبقه‌بندی دستی توسط یک شخص متخصص صورت می‌گیرد و بسیار کند و پر هزینه است. طبقه‌بندی خودکار توسط کامپیوتر و الگوریتم‌های طبقه‌بندی انجام می‌شود و سرعت بالاتری دارد. در طبقه‌بندی انحصاری، هر شیء فقط و



شکل ۴ دیدگاه کلی طبقه‌بندی

بازیکنان در یک بازی، از اصول اولیه بازی‌های مبتنی بر همکاری است.

در این بازی، نشانگرها به‌عنوان بازیکن، در نظر گرفته می‌شود. همان‌طور که می‌دانیم، یکی از مزایای مهم مقدار شیپلی، توانایی آن در فراهم آوردن راه‌حل‌های عادلانه و منحصر به فرد (Fair and Unique) است. منظور از عادلانه بودن یعنی سهم واقعی هر یک از بازیکنان در بازی را به‌صورت دقیق تعیین شود. بنابراین به‌منظور وزن‌دهی به بازیکنان (نشانگرها) از مقدار شیپلی (فرمول ۱ که در بخش ۳ مطرح شد) استفاده شد. در این زمینه باید به اصول زیر توجه نمود:

اصل ۱- عدم اهمیت جایگشت‌ها یا تقارن: برای هر بازی (N, v) ، جایگشت π روی N به‌صورت $\phi_i(v) = \phi_{\pi(i)}(\pi v)$ تعریف می‌شود.

این اصل بیان می‌کند که تغییر نام یا ترتیب نشانگرها در

یک طبقه‌بند $h: X \rightarrow Y$ شیء $x \in X$ را با برچسب طبقه‌بندی $y \in Y$ مشخص می‌کند. می‌توان براساس اطلاعات هر x_i ، آن‌ها را با یکی از برچسب‌های فوق مشخص نمود یا به‌عبارتی آن‌ها را طبقه‌بندی نمود.

به این ترتیب، در تحقیق حاضر نیز به‌منظور طبقه‌بندی انواع لوسمی، با استفاده از تئوری بازی‌های مبتنی بر همکاری به هر یک از CD نشانگرها یا در واقع داده‌های اولیه، وزن خاصی اختصاص داده شد. از آن‌جایی که در ایجاد هر نوع بدخیمی، نشانگرهای خاصی در کنار هم حضور پیدا می‌کند، این مسئله را می‌توان به ائتلاف نشانگرها تشبیه نمود. بنابراین می‌توان فرض کرد که بیماری‌ها در یک بازی شرکت کرده و CD نشانگرهای متعلق به هر بیماری به‌عنوان بازیکنان این بازی به‌منظور تشکیل یک بیماری خاص با یکدیگر تشکیل ائتلاف داده‌اند و همکاری می‌نمایند. همکاری و مشارکت

مقدار نهایی، بی‌اثر است. تعیین نمود. جدول ۳ نمونه‌ای از وزن‌های اختصاص داده شده به نشانگرها را نشان می‌دهد. به این ترتیب در بخش یادگیری می‌توان وزن دقیق هر یک از نشانگرها را براساس سهمی که در یک بیماری دارد، با استفاده از مقدار شیپلی تعیین کرد. وزن‌های به‌دست آمده به‌عنوان دانش اولیه برای سیستم خبره مورد نظر، به‌منظور طبقه‌بندی داده‌ها، در نظر گرفته می‌شود. بنابراین با توجه به مطالب فوق و مباحث طبقه‌بندی، می‌توان براساس وزن‌های به‌دست آمده برای هریک از نشانگرها، طبقه‌های مختلف بیماری را از یکدیگر جدا نمود.

اصل ۲- خصوصیت گنگی (dummy): برای هر بازی (N, ν) اگر رابطه $\nu(S \cup \{i\}) = \nu(S), \forall S \subseteq N \setminus \{i\}$ برقرار باشد، آن‌گاه $\phi_i(\nu) = 0$ خواهد بود. اصل ۲ بیان می‌کند که نشانگری (بازیکنی) که در یک بیماری تأثیر ندارد، مطمئناً مقدار آن صفر خواهد بود.

فرض کنید $B \in \{0,1\}^{n \times k}$ یک ماتریس بولین است که n تعداد نشانگرها و k تعداد نمونه‌های هر بیماری را نشان می‌دهد. با محاسبه مقدار شیپلی (فرمول ۱ در بخش ۳) برای ماتریس B به راحتی می‌توان وزن هریک از نشانگرها را

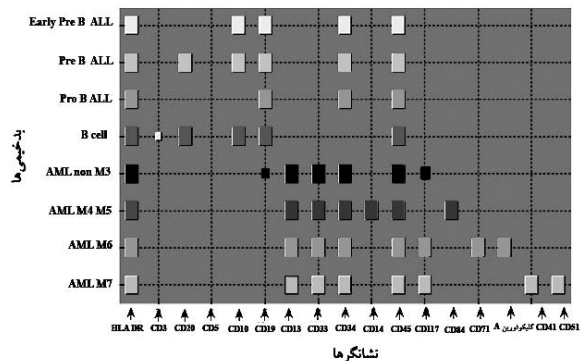
جدول ۳ وزن نشانگرهای مختلف بدخیمی‌ها

AML-Non-M3	AML-M7	AML-M6	AML-M4-M5	B-Cell-ALL	Pro-B-ALL	Pre-B-ALL	Early-Pre-B-ALL	طبقه نشانگر (ویژگی)
۰/۳۹۱۲	۰/۲۲۰۶	۰/۱۵۲۳	۰/۲۹۳۳	۰/۳۶۴۷	۰/۵۰۰۰	۰/۳۲۹۴	۰/۳۷۹۴	HLA-DL
.	.	.	.	۰/۰۲۹۴	.	.	.	CD3
.	.	.	.	۰/۳۶۴۷	.	۰/۳۲۹۴	.	CD20
.	CD5
.	.	.	.	۰/۲۷۶۵	.	۰/۳۲۹۴	.	CD10
۰/۰۲۹۴	.	.	.	۰/۳۶۴۷	۰/۵۰۰۰	۰/۳۲۹۴	۰/۳۷۹۴	CD19
۰/۲۸۸۲	۰/۲۲۰۶	۰/۱۹۴۳	۰/۲۶۳۹	.	.	.	۰/۳۷۹۴	CD13
۰/۳۳۲۴	۰/۲۲۰۶	۰/۲۳۶۳	۰/۲۹۳۳	CD33
۰/۲۷۳۵	۰/۲۲۰۶	۰/۲۳۶۳	۰/۱۳۴۵	.	۰/۳۲۳۵	۰/۱۱۷۶	۰/۲۴۷۱	CD34
.	.	.	۰/۲۲۲۷	CD14
۰/۳۹۱۲	۰/۲۲۰۶	۰/۲۳۶۳	۰/۲۹۳۳	۰/۳۶۴۷	۰/۴۴۱۲	۰/۳۲۹۴	۰/۳۷۹۴	CD45
۰/۰۵۸۸	۰/۲۲۰۶	۰/۲۳۶۳	CD117
.	.	.	۰/۲۶۳۹	CD64
.	.	۰/۲۳۶۳	CD71
.	.	۰/۲۳۶۳	گلیکوفورین A
.	۰/۲۲۰۶	CD41
.	۰/۲۲۰۶	CD61

می‌رود [۲۸]. مقادیر وزن‌ها در دیاگرام با مربع نشان داده می‌شوند که اندازه این مربع‌ها با توجه به مقدار آن‌ها متفاوت است و رنگ آن‌ها، وجود نشانگرهای مختلف را در بدخیمی، نشان می‌دهد.

به‌منظور بیان و مطالعه دقیق‌تر وزن‌ها، می‌توان از نمایش گرافیکی دیاگرام هنتون (Hinton diagram) (شکل ۵) استفاده نمود. دیاگرام هنتون به‌منظور نمایش کیفی مقادیر یک ماتریس داده‌ای (به‌ویژه یک ماتریس وزن‌دار) به‌کار

استفاده شد. به طوری که X_{ik} بیانگر نمونه آزمایشی و X_{jk} طبقه آموزشی را مشخص می‌کند. بنابراین با استفاده از فرمول ۲ می‌توان شباهت نمونه آزمایشی را با هر یک از طبقه‌های آموزشی، تعیین نمود. در نهایت با احتساب کمترین فاصله (Minimum distance) می‌توان دسته واقعی هر بیماری را معین نمود.



شکل ۵ دیاگرام هیتون به منظور نمایش وزن نشانگرها

۵- نتایج شبیه‌سازی

۱-۵- پایگاه داده بیماری‌های لوسمی

پایگاه داده مورد بحث در مطالعه حاضر، شامل ۳۰۴ نمونه از مغز استخوان و خون محیطی لوسمی انسان است که شامل ۱۷ ویژگی است. این ویژگی‌ها، CD نشانگرهای مختلف بدخیمی‌های مذکور را مشخص می‌کنند. داده‌های حاصل از آزمایش‌های مختلف، به صورت ماتریس در نظر گرفته می‌شود. به طوری که سطرها، CD نشانگرها (ویژگی‌ها) را بیان می‌کند و ستون‌ها، نمونه‌های مختلف بیماری را مشخص می‌کند. همه نمونه‌ها از طبقه بدخیم (Malignant class) است. در این فعالیت، هشت طبقه لوسمی به صورت زیر در نظر گرفته شده است (جدول ۴).

جدول ۴ هشت طبقه لوسمی حاد مورد استفاده شده در تحقیق حاضر

Early Pre B Acute Lymphoid Leukemia	Early Pre B ALL	
pre B Acute Lymphoid Leukemia	Pre B ALL	
Pro B Acute Lymphoid Leukemia	Pro B ALL	ALL
B cell Acute Lymphoid Leukemia	B Cell ALL	
Acute Myelomonoid Leukemia Non M3	AML Non M3	
Acute Myelomonoid Leukemia M4 M5	AML M4 M5	
Acute Myelomonoid Leukemia M6 (Erythroleukemia)	AML M6	AML
Acute Myelomonoid Leukemia M7 (Acute Megakaryocytic Leukemia)	AML M7	

به وسیله دیاگرام هیتون می‌توان به سرعت دریافت که چه نشانگرهایی در یک نمونه بدخیمی فعال هستند. تمایز طبقه‌های مختلف بیماری در این دیاگرام به خوبی نشان داده شده است.

۳-۴- مرحله تشخیص

هدف بررسی حاضر در این مرحله، تشخیص تعلق نمونه جدید به هر یک از طبقه‌های مختلف آموزشی است. برای این منظور، مراحل پیش‌پردازش (بخش ۴-۲-۱) و تولید بردارهای داده‌ای (بخش ۴-۲-۲) مانند مرحله یادگیری انجام می‌شود. مقدار M (در بخش ۴-۲-۱ در مرحله یادگیری بیان شد) یک در نظر گرفته می‌شود. در واقع می‌خواهیم ببینیم یک نمونه مورد آزمایش به کدام یک از طبقه‌های آموزشی متعلق است یا به عبارت دیگر؛ به کدام یک از طبقه‌های موجود، شباهت بیشتری دارد. به منظور تعیین شباهت بین یک نمونه و طبقه‌های موجود، می‌توان از معیارهای شباهت استفاده نمود. برخی از معیارهای شباهت کمی عبارتند از معیار مینکوفسکی (Minkowski metric)، معیار سیتی بلاک (City block)، فاصله اقلیدسی (Euclidean Distance)، معیار کنبرا (Canberra)، معیار تفکیک زاویه‌ای (Angular separation) و ضریب همبستگی (Correlation coefficient) [۲۸]. در پژوهش حاضر، از معیار فاصله اقلیدسی

$$d_{ij} \equiv (\sum_{k=1}^p (x_{ik} - x_{jk})^2)^{1/2}. \quad (2)$$

جدول ۶ دقت طبقه‌بندی، میزان حساسیت و میزان پیشگویی مثبت در مدل

مورد نظر	
میزان حساسیت	۹۳/۷ درصد
میزان پیشگویی مثبت	۹۳/۷ درصد
دقت طبقه‌بندی	۹۶/۳ درصد

همان‌گونه که در جدول ۵ مشاهده می‌شود، دقت طبقه‌بندی ۹۶/۳ در نظر گرفته شده است. این میزان بیان می‌کند که روش پیشنهادی از دقت قابل توجهی برخوردار است و توانایی تفکیک درست طبقه‌ها را دارد.

میزان حساسیت، ۹۳/۷ است. هر چقدر میزان این معیار بالاتر باشد، طبقه‌بندی کامل‌تر بوده و بیماری‌ها به‌درستی از یکدیگر تفکیک و تشخیص داده شده است؛ بنابراین روش درمانی مناسب برای هر بیمار استفاده می‌شود. میزان پیشگویی مثبت، احتمال این‌که در مرحله تشخیص، یک فرد بیمار تشخیص داده شود و در حالت واقعی نیز این فرد بیمار باشد را بیان می‌کند. در تشخیص پزشکی این میزان از اهمیت بالایی برخوردار است. به عبارت دیگر؛ بیان می‌کند که سیستم طبقه‌بندی تا چه اندازه مفید است. تشخیص درست بیماری یک فرد بیمار به پزشک معالج در گرفتن تصمیم مناسب در شیوه درمان و تجویز داروی مناسب کمک می‌نماید.

۶- مقایسه روش ارایه شده با شبکه عصبی پایه

شعاعی (Radial Basis Function: RBF)

اگر چه نتایج به‌دست آمده از به‌کارگیری تئوری بازی‌های مبتنی بر همکاری و مقدار شیپلی در طبقه‌بندی انواع لوسمی بسیار امیدوارکننده است، اما در این مرحله سعی شده است تا کارایی این روش با یکی از روش‌های پایه و دقیق در حوزه یادگیری مقایسه شود. در این خصوص با انتخاب شبکه RBF که از جمله کاراترین شبکه‌های عصبی است، این مقایسه انجام شده است. ساختمان این شبکه در شکل ۶ ارایه شده است.

۸۰ درصد داده‌ها در مرحله یادگیری و بقیه در مرحله تشخیص، استفاده شد. جدول ۵، مجموع مقادیر انطباق نمونه‌های مورد آزمایش با طبقه مربوط به خود را ۹۳/۷۵ نشان می‌دهد. با توجه به این مقدار، می‌توان گفت وزن‌های تعیین شده به‌وسیله مقدار شیپلی، بهترین کارایی را در طبقه‌بندی داده‌ها دارد. خطای قابل مشاهده در نمونه‌های Early pre B، Pre B All، All به‌علت همپوشانی این طبقه‌ها در نمونه‌های آموزشی است.

جدول ۵ تعداد نمونه‌های درست از مدل ارایه شده

نوع بدخیمی	انطباق نمونه‌های مورد آزمایش با طبقه مربوط به خود
Early pre B ALL	۹۲/۵ درصد
Pre B ALL	۸۷/۵ درصد
Pro B ALL	۱۰۰ درصد
B Cell ALL	۱۰۰ درصد
AML M4 M5	۱۰۰ درصد
AML M6	۱۰۰ درصد
AML m7	۱۰۰ درصد
AML Non M3	۱۰۰ درصد
مجموع	۹۳/۷۵ درصد

به‌منظور ارزیابی بهتر کارایی مدل ارایه شده، از معیارهای اندازه‌گیری مختلفی نظیر دقت طبقه‌بندی (Classification accuracy)، میزان حساسیت (Sensitivity)، میزان پیشگویی مثبت (Positive predictive value) استفاده شد که عبارتند از:

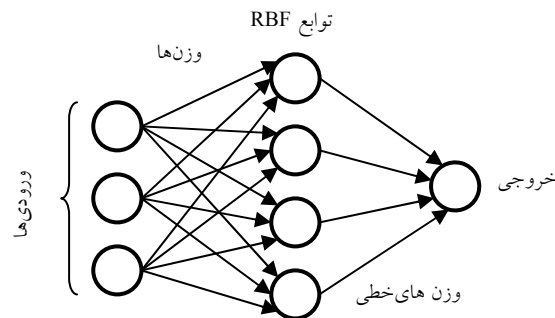
$$\text{دقت طبقه‌بندی (درصد)} = \frac{TP + TN}{TP + FP + FN + TN} \quad (۳)$$

TP: True Positive
TN: True Negative
FN: False Negative
FP: False Positive

$$\text{میزان حساسیت (درصد)} = \frac{TP}{TP + FN} \times 100 \quad (۴)$$

$$\text{میزان پیشگویی مثبت} = \frac{TP}{TP + FP} \times 100 \quad (۵)$$

مقادیر فوق در جدول ۶ نمایش داده شده است.



شکل ۶ شبکه عصبی RBF

همان‌طور که ملاحظه می‌شود، دقت طبقه‌بندی بیماری‌های لوسمی در روش پیشنهادی از شبکه عصبی RBF بهتر است.

۷- نتیجه‌گیری و پیشنهادات آتی

در این مقاله، روشی براساس تئوری بازی مبتنی بر همکاری به‌منظور طبقه‌بندی انواع لوسمی‌های نوع حاد، ارائه شد. تحقیق مورد نظر، روی داده‌های واقعی با روش فلوسیتومتری از نمونه‌های مختلف لوسمی - برگرفته از سازمان انتقال خون ایران - صورت پذیرفت. مجموعه داده، شامل ۳۰۴ نمونه از لوسمی انسانی است. ۸۰ درصد داده‌ها در مرحله یادگیری و مابقی آن‌ها در مرحله تشخیص استفاده شد. دقت طبقه‌بندی روش پیشنهادی ۹۶/۳ درصد است که در مقایسه با شبکه RBF از دقت بالاتری برخوردار است. به‌منظور ارزیابی کارایی روش پیشنهادی از معیارهای ارزیابی نظیر دقت طبقه‌بندی، میزان حساسیت و میزان پیش‌گویی مثبت استفاده شد. با توجه به نتایج حاصل، می‌توان گفت روش پیشنهادی در طبقه‌بندی بدخیمی‌های مذکور، موفقیت‌آمیز بوده است. بنابراین می‌تواند به هماتولوژیست‌ها در تشخیص دقیق‌تر انواع لوسمی کمک نماید تا تصمیمات بهتری در مورد نحوه درمان این بیماران بگیرند.

از سوی دیگر تئوری بازی‌های مبتنی بر همکاری می‌تواند در راستای طبقه‌بندی دقیق داده‌ها می‌تواند مفید واقع شود. هر چند که در این زمینه محدودیت‌هایی نیز وجود دارد. از

این شبکه‌ها از جمله شبکه‌های پیش‌خور (Feed forward) بوده که از سه لایه ثابت تشکیل شده است. لایه ورودی که محل تزریق علائم ورودی به شبکه است؛ لایه میانی یا طبقه RBF که شامل توابع RBF است (در اکثر موارد از توابع گوسی (Gaussian function) در لایه استفاده می‌شود) و در نهایت لایه خروجی که ترکیبی خطی از کلیه خروجی‌های طبقه RBF را می‌سازد. وزن‌های اتصالات بین لایه RBF و لایه خروجی در هنگام یادگیری و با استفاده از الگوریتم یادگیری، مشخص می‌شود. این نوع شبکه‌ها می‌تواند به‌صورت کارا کلیه توابع عمومی را تخمین بزند. سرعت الگوریتم آموزشی آن‌ها سریع است؛ بنابراین می‌توان گفت، شبکه عصبی RBF از جمله قدرتمندترین نوع شبکه‌های عصبی در حل مسایل پیچیده مهندسی است [۲۹].

به‌منظور مقایسه کارایی روش پیشنهادی، یک‌بار دیگر مسئله مورد نظر با شبکه RBF پیاده‌سازی شد. در این جا نیز ۸۰ درصد داده‌ها در مرحله یادگیری و باقیمانده آن‌ها در مرحله تشخیص استفاده شد. نتایج حاصل از این مقایسه در جدول ۷ نشان داده شده است.

جدول ۷ مقایسه دقت طبقه‌بندی روش ارائه شده با شبکه عصبی RBF

دقت طبقه‌بندی	روش طبقه‌بندی
۹۶/۳	روش ارائه شده براساس تئوری بازی مبتنی بر همکاری
۹۱/۸۰	شبکه عصبی پایه شعاعی (RBF)

۷- تشکر و قدردانی

از کلیه کارکنان بخش فلوسایتومتری سازمان انتقال خون ایران که در زمینه گردآوری داده‌ها و ارایه مطالب به محققان حاضر یاری رساندند، صمیمانه تشکر می‌نمایم.

آنجایی که که تابع شیپلی، به صورت نمایی است، بنابراین در حجم بیشتر داده‌ها، پیچیدگی سیستم بیشتر شده و زمان پردازش در بخش آموزش با مرتبه نمایی افزایش می‌یابد. در نهایت توسعه سیستم در راستای MDSS می‌تواند به عنوان چشم‌انداز جدیدی در عرصه سیستم‌های پزشکی مطرح شود.

۸- منابع

- [1] Gordon AD. Classification. 2nd edition, New York: Chapman & Hall/CRC Press, 1999; 29-33.
- [2] Mahmoud Aadi A. Leukemia in Simple Language. Tehran: Kerdegari Press, 2007; p: 1-30. (Persian)
- [3] Nikoogoftar M. Hemathoogy and Blood Transfer. Tehran: Andishe Rafie, 2007; p: 251-74. (Persian)
- [4] Wood B, Borowitz M, Abraham N, Massey H, Bluth M, Miller, Threatte G, .Hutchison R, Unger E, Lifshitz M. Henry's Clinical Diagnosis AND Management BY Laboratory Methods Twenty First Edition. New York: Sounders Elsevier Press, 2007; 599-600.
- [5] Michie D, Spiegelhalter DJ, Taylor CC. Machine learning, neural and statistical classification. New York: Ellis Horwood Press, 1994; p: 49-60.
- [6] Miller RA. Medical diagnostic decision support systems--past, present, and future: a threaded bibliography and brief commentary. J Am Med Inform Assoc 1994; 1(1): 8-27
- [7] Garcia-Gomez JM, Vidal C, Vicente J, Marti-Bonmati L, Robles M. Medical decision support system for diagnosis of soft tissue tumors based on distributed architecture. Conf Proc IEEE Eng Med Biol Soc 2004; 5: 3225-8.
- [8] Sefion I, Gaiihardou M, Ennaji A. A medical decision support system for asthmatic patient health care. IASTED International Conference on Artificial Intelligence and Applications (AIA'02) Spain 2002; 83-6
- [9] Bakar AA, Othman Z, Ismail R, Zakari Z. Using rough set theory for mining the level of hearing loss diagnosis knowledge. International Conference on Electrical Engineering and Informatics (ICEEI '09) 2009; 1: 7-11.
- [10] Chae YM, Park KS, Park Q, Bae MY. Development of medical decision support system for leukemia management. Stud Health Technol Inform 1998; 52 Pt 1: 449-52.
- [11] Chou S.M, Lee T.S, Shao Y. E, Chen I. Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines. Expert Systems with Applications Elsevier 2004; 27, 133-42.
- [12] Mumtaz K, Sheriff S.A, Duraiswamy K. Evaluation of three neural network models using Wisconsin breast cancer database. International Conference on Control, Automation, Communication and Energy Conservation (INCACEC 09) 2009; 1-7.
- [13] Karabatak M, Cevdet M. An expert system for detection of breast cancer based on association

- rules and neural network. *Expert Systems with Applications Elsevier* 2009; 36, 3465–9.
- [14] Soria D, Garibaldi JM, Biganzol E, Ellis IO. A Comparison of Three Different Methods for Classification of Breast Cancer Data. *Seventh International Conference on Machine Learning and Applications (ICMLA '08) San Diego, US* 2008; 619–24.
- [15] De Paz JF, Rodriguez S, Bajo J, Corchado JM. CBR System for Diagnosis of Patients. *Eighth International Conference on Hybrid Intelligent Systems (HIS '08)* 2008; 807–12.
- [16] Bollschweiler EH, Mönig SP, Hensler K, Baldus SE, Maruyama K, Hölscher AH. Artificial neural network for prediction of lymph node metastases in gastric cancer: a phase II diagnostic study. *Ann Surg Oncol* 2004; 11(5): 506–11.
- [17] Sewak M, Vaida P, Chan C, Duan Z. SVM Approach to Breast Cancer Classification. *Second International Multi symposium on Computer and Computational Sciences* 2007; 32-7.
- [18] Bellomo N, Delitala M. From the mathematical kinetic, and stochastic game theory to modelling mutations, onset, progression and immune competition of cancer cells. *JPL Rev* 2008; 5(4): 183–206.
- [19] Bach LA, Bentzen SM, Alsner J, Christiansen FB. An evolutionary-game model of tumour-cell interactions: possible relevance to gene therapy. *Eur J Cancer* 2001; 37(16): 2116–20.
- [20] Mansury Y, Diggory M, Deisboeck TS. Evolutionary game theory in an agent-based brain tumor model: exploring the 'Genotype-Phenotype' link. *J Theor Biol* 2006; 238(1): 146–56.
- [21] Fragnelli V, Moretti S. A game theoretical approach to the classification problem in gene expression data analysis. *J.CAMWA* 2007; 55: 950-9.
- [22] Myerson RB. *Game theory: analysis of conflict*. New York: Harvard University Press, 1997; p: 12-45.
- [23] Gibbons R. *A Primer in Game Theory*. New York: Prentice Hall Press, 2000; 1-5.
- [24] Brandenburger A. *Cooperative Game Theory: Characteristic Functions, Allocations, Marginal Contribution*. 2007; <http://pages.stern.nyu.edu/~abranden/coop-01-04-07.pdf>
- [25] Branzei R, Dimitrov D, Tijs S. *Models in Cooperative Game Theory*. 2nd edition, Berlin: Springer Press, 2008; 26-30.
- [26] Shapley L. A value for n-person games. In: Kuhn HW, Tucker AW (Eds.). *Contributions to the Theory of Games II Annals of Mathematics Studies*. Princeton University Press, 1953; 28: 307–17.
- [27] Cohen S, Ruppin E. *Feature Selection Based on the Shapley Value*. Technical Report 2005; 1-6. <http://www.ijcai.org/papers/0763.pdf>
- [28] Skapura David M, *Building neural networks*. New York: Addison Wesley Press, 1996; 88-93.
- [29] Karayiannis NB, Randolph-Gips MM. On the construction and training of reformulated radial basis function neural network. *IEEE Transactions on Neural Networks* 2003; 14(4): 835–46.