

صلى الله عليه وسلم



دانشگاه شیخ بهائی  
دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی ارشد رشته مهندسی کامپیوتر - نرم افزار

استخراج دانش از ماشین بردار پشتیبان برای درمان نفريت لوپوسی

پژوهشگر

رسول بصیرت

استاد راهنما

دکتر ناصر قاسم آقائی

استاد مشاور

دکتر شهرزاد شهیدی

شهریور ۱۳۹۳

## سپاسگذاری

برخود لازم می‌دانم از اساتید بزرگوار جناب آقای دکتر ناصر قاسم آقای و خانم دکتر شهرزاد شهیدی به

خاطر سعی صدر و رهنمودهای دلسوزانه، صمیمانه تشکر کنم.

با نهایت اخلاص و عشق تقدیم به:

**پدرم به استواری کوه**

و

**مادرم به زلالی چشمه**

و

**همسرم به صمیمیت باران**

## چکیده

کسب دانش از داده‌های مربوط به بیماران نفریت لوپوسی می‌تواند کمک زیادی به ایجاد یک طرح درمانی مناسب باشد. در سه دهه اخیر، تکنیک‌های داده‌کاوی در استخراج دانش و الگوهای مخفی موجود در پایگاه داده‌ها به‌طور قابل توجهی موفق بوده‌اند. ماشین بردار پشتیبان یکی از موفق‌ترین تکنیک‌های داده‌کاوی در این زمینه به‌شمار می‌رود که در کاربردهای متنوعی به کار گرفته شده است. با وجود برتری کارایی در ماشین بردار پشتیبان، یک عیب بزرگ در این تکنیک وجود دارد و روشی که این تکنیک برای یادگیری الگوها از داده‌ها به کار می‌برد برای کاربر قابل درک نیست در نتیجه این تکنیک یک توضیح قابل درک برای دانش‌های یاد گرفته شده فراهم نمی‌کند. این مشکل باعث شده است در برخی از کاربردهای عملی به‌خصوص در حوزه‌هایی مثل تشخیص و پیش‌بینی پزشکی از این روش استفاده نشود. استخراج قواعد از ماشین بردار پشتیبان، ایده به دست آوردن قواعد از مدل به‌جای استخراج مستقیم آن‌ها از داده‌ها است. بنابراین از این طریق یک توضیح از الگوهای یاد گرفته‌شده و تعبیه‌شده در ساختار ماشین بردار پشتیبان آشکارشده و به کاربران در یک شکل قابل درک ارائه می‌شود.

در این پایان‌نامه یک روش پیشنهادی برای استخراج قواعد فازی از ماشین بردار پشتیبان برای کمک به درمان بیماری نفریت لوپوسی ارائه می‌شود. در روش پیشنهادی ابتدا ویژگی‌های تأثیرگذار با معیار نفع اطلاعات از مجموعه داده بیماران انتخاب می‌شوند. پس از این ماشین بردار پشتیبان آموزش می‌بیند و استخراج قواعد از بردارهای پشتیبان صورت می‌گیرد. در مرحله بعد قواعد به دست آمده به قواعد فازی تبدیل می‌شوند که این مرحله با استفاده از الگوریتم رقابت استعماری انجام می‌شود تا مجموعه‌های فازی به‌دقت تنظیم شوند.

ارزیابی روش پیشنهادی روی مجموعه داده بیماران نفریت لوپوسی که در طی این پژوهش جمع‌آوری شده است و مجموعه داده گل زنبق که یک مجموعه داده استاندارد برای ارزیابی است، انجام می‌شود. معیار ارزیابی نرخ خطای دسته‌بندی توسط قواعد فازی تولید شده است. مقایسه روش پیشنهادی با سه روش دسته‌بندی دیگر و بر اساس پارامترهای متفاوت ماشین بردار پشتیبان نشان‌دهنده برتری روش پیشنهادی در دسته‌بندی داده‌ها می‌باشد.

واژه‌های کلیدی: نفریت لوپوسی (Lupus Nephritis)، استخراج قواعد (Rule Extraction)، ماشین بردار

پشتیبان (Support Vector Machine)، الگوریتم رقابت استعماری (Imperialist Competitive

Algorithm)، تنظیم توابع عضویت (Membership Function Adjustment).

## فهرست مطالب

صفحه	عنوان
۱	فصل اول: مقدمه.....
۲	۱-۱ انگیزه تحقیق.....
۳	۲-۱ طرح مسئله.....
۴	۳-۱ هدف.....
۵	۴-۱ ساختار پایان نامه.....
۶	فصل دوم: ادبیات موضوع.....
۷	۱-۲ مقدمه.....
۸	۲-۲ نفریت لوپوسی.....
۱۰	۳-۲ پیش آگهی نفریت لوپوسی.....
۱۳	۴-۲ درمان های دارویی.....
۱۳	۵-۲ ماشین بردار پشتیبان.....
۱۴	۱-۵-۲ الگوریتم دسته بندی ماشین بردار پشتیبان.....
۱۹	۲-۵-۲ دسته بندی غیرخطی.....
۲۲	۶-۲ استخراج قواعد.....
۲۲	۱-۶-۲ انگیزه های استخراج قواعد.....
۲۲	۱-۱-۶-۲ قابلیت توضیح.....
۲۳	۲-۱-۶-۲ شفافیت.....
۲۳	۳-۱-۶-۲ تأیید نرم افزاری.....
۲۳	۴-۱-۶-۲ بهبود تعمیم.....
۲۴	۵-۱-۶-۲ استخراج داده و القاء تئوری های علمی.....
۲۴	۷-۲ جمع بندی.....
۲۵	فصل سوم : پیشینه تحقیق.....
۲۶	۱-۳ مقدمه.....

۲۶	۲-۳ روش‌های استخراج قواعد.....
۲۶	۱-۲-۳ روش‌هایی که از مدل ماشین بردار پشتیبان به صورت جعبه بسته استفاده می‌کنند.....
۲۷	۲-۲-۳ روش‌هایی که فقط از بردارهای پشتیبان استفاده می‌کنند.....
۲۸	۳-۲-۳ روش‌هایی که از بردارهای پشتیبان و ابر صفحه‌های جداکننده استفاده می‌کنند.....
۲۹	۴-۲-۳ روش‌هایی که از بردارهای پشتیبان، داده‌های آموزشی و ابر صفحه‌های جداکننده استفاده می‌کنند.....
۳۲	۳-۳ استخراج قواعد فازی.....
۳۴	۴-۳ جمع‌بندی:.....
۳۵	فصل چهارم: روش پیشنهادی.....
۳۶	۱-۴ مقدمه.....
۳۶	۲-۴ روش پیشنهادی.....
۳۸	۱-۲-۴ انتخاب ویژگی‌های تأثیرگذار.....
۳۸	۲-۲-۴ آموزش ماشین بردار پشتیبان.....
۳۹	۳-۲-۴ استخراج قواعد.....
۴۰	۴-۲-۴ استخراج قواعد فازی.....
۴۱	۱-۴-۲-۴ توابع عضویت.....
۴۲	۲-۴-۲-۴ شکل‌دهی به امپراطوری‌های اولیه.....
۴۷	۳-۴-۲-۴ سیاست جذب.....
۵۰	۴-۴-۲-۴ جابجایی موقعیت مستعمره و استعمارگر.....
۵۱	۵-۴-۲-۴ قدرت کل استعمارگر.....
۵۱	۶-۴-۲-۴ رقابت استعماری.....
۵۳	۷-۴-۲-۴ سقوط استعمارگران ضعیف.....
۵۴	۸-۴-۲-۴ انقلاب در کشور.....
۵۴	۹-۴-۲-۴ همگرایی.....
۵۵	۵-۴ جمع‌بندی.....
۵۶	فصل پنجم: پیاده‌سازی و ارزیابی.....

۵۷	۱-۵ مقدمه
۵۷	۲-۵ پیاده‌سازی
۶۳	۳-۵ ارزیابی مجموعه داده بیماران نفریت لوپوسی
۶۴	۱-۳-۵ فراوانی داده‌ها
۷۰	۲-۳-۵ انتخاب ویژگی‌های تأثیرگذار
۷۲	۳-۳-۵ مجموعه‌های فازی بهینه
۷۴	۴-۳-۵ قواعد فازی
۷۶	۵-۳-۵ ارزیابی نرخ خطای دسته‌بندی
۷۹	۶-۳-۵ ارزیابی همگرایی
۸۲	۴-۵ مقایسه با روش‌های دیگر
۸۳	۱-۴-۵ مجموعه‌های فازی بهینه
۸۵	۲-۴-۵ ارزیابی نرخ خطای دسته‌بندی
۸۶	۳-۴-۵ ارزیابی همگرایی
۸۷	۵-۵ جمع بندی
۸۸	فصل ششم: نتیجه گیری و کارهای آینده
۸۹	۱-۶ نتیجه‌گیری
۸۹	۲-۶ یافته‌های پژوهش
۹۰	۳-۶ کارهای آینده
۹۱	منابع
۹۶	واژه نامه فارسی به انگلیسی
۱۰۰	واژه نامه انگلیسی به فارسی
۱۰۴	پیوست
۱۰۵	مقدمه
۱۰۵	نحوه کار با نرم‌افزار



## فهرست شکل‌ها

صفحه	عنوان
۷.....	شکل ۱-۲ مراحل ایجاد لوپوس .....
۹.....	شکل ۲-۲ مراحل ایجاد نفریت لوپوسی .....
۹.....	شکل ۳-۲ رسوب سلول‌های سیستم ایمنی در ساختار گلومرول کلیه .....
۱۲.....	شکل ۴-۲ نمونه بیوپسی کلاس چهار نفریت لوپوسی .....
۱۴.....	شکل ۵-۲ حالت‌های متفاوت برای جداسازی کلاسه‌ها .....
۱۵.....	شکل ۶-۲ دو حاشیه متفاوت برای جداسازی کلاسه‌ها .....
۱۷.....	شکل ۷-۲ نمایش بردارهای پشتیبان .....
۱۹.....	شکل ۸-۲ داده‌های جداناپذیر خطی .....
۲۹.....	شکل ۱-۳ استفاده از بردارهای پشتیبان و مرز تصمیم برای استخراج قواعد .....
۳۰.....	شکل ۲-۳ استفاده از خوشه‌بندی برای تعیین بردارهای نمونه .....
۳۱.....	شکل ۳-۳ استفاده از روش پوسته محدب برای استخراج قواعد .....
۳۶.....	شکل ۱-۴ مراحل کلی روش پیشنهادی .....
۳۷.....	شکل ۲-۴ نمودار گردش الگوریتم پیشنهادی .....
۳۹.....	شکل ۳-۴ تعدادی از بردارهای پشتیبان به دست آمده از مدل .....
۳۹.....	شکل ۴-۴ قواعد تولید شده از بردارهای پشتیبان .....
۴۰.....	شکل ۵-۴ مجموعه‌های فازی برای ویژگی سن .....
۴۳.....	شکل ۶-۴ مجموعه‌های فازی اولیه .....
۴۴.....	شکل ۷-۴ مقدار دهی اولیه به متغیرهای مسئله .....
۴۵.....	شکل ۸-۴ طبقه‌بندی داده‌ها در ۵ طبقه در روش اعتبارسنجی ضربدری .....
۴۷.....	شکل ۹-۴ شکلدهی به امپراطوری‌های اولیه .....

- شکل ۴-۱۰ حرکت مستعمره به سمت استعمارگر ..... ۴۷
- شکل ۴-۱۱ حرکت مستعمره به سمت استعمارگر با در نظر گرفتن زاویه ..... ۴۸
- شکل ۴-۱۲ شرایط در نظر گرفته شده برای جلوگیری از نقض محدودیت ها ..... ۴۹
- شکل ۴-۱۳ مجموعه‌های فازی نامعتبر ..... ۵۰
- شکل ۴-۱۴ جابجایی موقعیت مستعمره و استعمارگر ..... ۵۰
- شکل ۴-۱۵ رقابت استعماری ..... ۵۲
- شکل ۴-۱۶ حذف استعمارگر بدون مستعمره ..... ۵۳
- شکل ۴-۱۷ حرکت تصادفی مستعمره جهت انقلاب ..... ۵۴
- شکل ۵-۱ شکل (الف) کلاس Imperialist، شکل (ب) کلاس Colony و شکل (ج) کلاس MyProblem را نشان می‌دهد. ..... ۵۹
- شکل ۵-۲ شکل (الف) کلاس SupportVectorMachine، شکل (ب) کلاس JCA، شکل (ج) کلاس Program و شکل (د) کلاس Randoms را نشان می‌دهد. ..... ۶۰
- شکل ۵-۳ گراف وابستگی کلاس‌های مورد استفاده ..... ۶۲
- شکل ۵-۴ نمودار جعبه‌های (الف) ویژگی Lupusnephritisage و (ب) ویژگی Lupusage را نشان می‌دهد. ۶۵
- شکل ۵-۵ قسمت (الف) ویژگی Sex و قسمت (ب) ویژگی Outcome را نشان می‌دهد. ..... ۶۵
- شکل ۵-۶ نمودار جعبه‌های (الف) ویژگی ProLast و (ب) ویژگی ProDX را نشان می‌دهد. ..... ۶۶
- شکل ۵-۷ قسمت (الف) ویژگی HematuriaLast و قسمت (ب) ویژگی HematuriaDX را نشان می‌دهد.

۶۶

- شکل ۵-۸ نمودار جعبه‌های (الف) ویژگی CrLast و (ب) ویژگی CrDX را نشان می‌دهد. ..... ۶۷
- شکل ۵-۹ قسمت (الف) ویژگی AcLIgM و قسمت (ب) ویژگی AcLIgG را نشان می‌دهد. ..... ۶۷
- شکل ۵-۱۰ نمودار جعبه‌های (الف) ویژگی Durationlupus و (ب) ویژگی Durationnephritis را نشان می‌دهد. ..... ۶۸

- شکل ۵-۱۱ قسمت (الف) ویژگی C4low و قسمت (ب) ویژگی C3low را نشان می‌دهد. ..... ۶۸

شکل ۱۲-۵ قسمت (الف) ویژگی ANALast و قسمت (ب) ویژگی ANADX را نشان می دهد. ۶۹.....

شکل ۱۳-۵ قسمت (الف) ویژگی Antimalarialast و قسمت (ب) ویژگی AntimalariaDX را نشان می دهد.

۶۹

شکل ۱۴-۵ قسمت (الف) ویژگی Cyclophosphamid و قسمت (ب) ویژگی Predisolon را نشان می دهد.

۶۹

شکل ۱۵-۵ قسمت (الف) ویژگی RenalBX و قسمت (ب) ویژگی Renal را نشان می دهد. ۶۹.....

شکل ۱۶-۵ قسمت (الف) ویژگی Cellcept و قسمت (ب) ویژگی AZA را نشان می دهد. ۷۰.....

شکل ۱۷-۵ قسمت (الف) ویژگی Tacrolimus و قسمت (ب) ویژگی Cyclospoin را نشان می دهد. ۷۰.....

شکل ۱۸-۵ ابزارهای مورد استفاده در نرم افزار Orange برای محاسبه معیار نفع اطلاعات. ۷۰.....

شکل ۱۹-۵ رتبه بندی ویژگی ها براساس معیار نفع اطلاعات ..... ۷۱

شکل ۲۰-۵ مجموعه های فازی ویژگی ProDX ..... ۷۲

شکل ۲۱-۵ مجموعه های فازی ویژگی CrDX ..... ۷۳

شکل ۲۲-۵ مجموعه های فازی ویژگی ProLast ..... ۷۳

شکل ۲۳-۵ مجموعه های فازی ویژگی CrLast ..... ۷۴

شکل ۲۴-۵ تعدادی از قواعد فازی به دست آمده ..... ۷۶

شکل ۲۵-۵ نمودار همگرایی روش پیشنهادی روی مجموعه داده لوپوس (هشت ویژگی) با انتخاب هسته Sigmoid

و پارامترهای متفاوت ..... ۸۰

شکل ۲۶-۵ نمودار همگرایی روش پیشنهادی روی مجموعه داده لوپوس (هشت ویژگی) با انتخاب هسته RBF و

پارامترهای متفاوت ..... ۸۰

شکل ۲۷-۵ نمودار همگرایی روش پیشنهادی روی مجموعه داده لوپوس (چهار ویژگی) با انتخاب هسته RBF و

پارامترهای متفاوت ..... ۸۱

شکل ۲۸-۵ نمودار همگرایی روش پیشنهادی روی مجموعه داده لوپوس (چهار ویژگی) با انتخاب هسته Sigmoid

و پارامترهای متفاوت ..... ۸۱

شکل ۲۹-۵ مجموعه های فازی ویژگی Sepallenght ..... ۸۳

- شکل ۳۰-۵ مجموعه‌های فازی ویژگی Sepalwidth ..... ۸۳
- شکل ۳۱-۵ مجموعه‌های فازی ویژگی Petallenght ..... ۸۴
- شکل ۳۲-۵ مجموعه‌های فازی ویژگی Petalwidth ..... ۸۴
- شکل ۳۳-۵ نمودار همگرایی روش پیشنهادی روی مجموعه داده گل زنبق با هسته خطی ..... ۸۶
- شکل ۳۴-۵ نمودار همگرایی روش پیشنهادی روی مجموعه داده گل زنبق با هسته خطی ..... ۸۶
- شکل ۱-۸ رابط کاربری نرم‌افزار Orange ..... ۱۰۵
- شکل ۲-۸ نحوه انتخاب مجموعه داده در نرم‌افزار Orange ..... ۱۰۶
- شکل ۳-۸ تنظیمات مؤلفه Rank ..... ۱۰۷

## فهرست جدول‌ها

عنوان	صفحه
جدول ۱-۲ میزان تصفیه خون توسط کلیه‌ها	۱۱
جدول ۱-۵ توضیح برخی از توابع مهم پیاده‌سازی شده	۶۱
جدول ۲-۵ توصیف ویژگی‌های مجموعه داده بیماران نفریت لوپوسی	۶۳
جدول ۳-۵ رابطه کلاس واقعی و کلاس پیش‌بینی شده	۷۷
جدول ۴-۵ نرخ خطای دسته‌بندی با هسته RBF و انتخاب چهار ویژگی	۷۷
جدول ۵-۵ نرخ خطای دسته‌بندی با هسته Sigmoid و انتخاب چهار ویژگی	۷۸
جدول ۶-۵ نرخ خطای دسته‌بندی با هسته RBF و انتخاب هشت ویژگی	۷۸
جدول ۷-۵ نرخ خطای دسته‌بندی با هسته Sigmoid و انتخاب هشت ویژگی	۷۹
جدول ۸-۵ مشخصات مجموعه داده گل زنبق	۸۲
جدول ۹-۵ مقایسه نرخ خطای دسته‌بندی روش پیشنهادی با سه روش دیگر روی داده‌های آموزشی	۸۵
جدول ۱۰-۵ مقایسه نرخ خطای دسته‌بندی روش پیشنهادی با سه روش دیگر روی داده‌های آزمایشی	۸۵

فصل اول:

# مقدمه

## ۱-۱ انگیزه تحقیق

لوپوس منتشر یا لوپوس اریتماتوس سیستمیک<sup>۱</sup> یک نوع بیماری خود ایمنی چند ارگانی هست که هنوز به عنوان یک بیماری با علت نامعلوم شناخته می‌شود. بیش از ۵۰ درصد از بیماران لوپوس به بیماری کلیوی دچار می‌شوند که عمدتاً به واسطه رسوب کمپلکس‌های ایمنی در کلیه‌ها می‌باشد. به این نوع لوپوس که کلیه‌ها را درگیر بیماری می‌کند نفریت لوپوسی<sup>۲</sup> گفته می‌شود. نفریت لوپوسی اصلی‌ترین عامل مرگ و میر در بیماران لوپوس می‌باشد [۱].

با وجود پیشرفت‌های حاصل شده در مدیریت بالینی<sup>۳</sup> نفریت لوپوس در سال‌های اخیر، با تشخیص زودهنگام بیماری و بهینه‌سازی رژیم‌های سرکوب کننده سیستم ایمنی، هنوز هم در حدود ده تا پانزده درصد از بیماران به بیماری مرحله نهایی<sup>۴</sup> کلیه مبتلا می‌شوند. بعد از قرار گرفتن عملکرد کلیه‌ها در مرحله نهایی، کم کم علائم از کار افتادن کلیه مثل مشکلات خواب، سردرد، خستگی، حالت تهوع و ورم دست و پا مشاهده می‌شود و نقص کامل کلیه<sup>۵</sup> نزدیک می‌شود. در صورت نقص کامل کلیه بیمار نیاز دارد با عمل پیوند کلیه<sup>۶</sup> یا فرآیند دیالیز<sup>۷</sup> این نقص را جبران کند [۲].

پیش‌آگهی یکی از واژگان پزشکی است که برای بیان پیش‌بینی آینده یک بیماری به کار می‌رود. در حقیقت پیش‌آگهی پیش‌بینی‌ای است که پزشک با توجه به وضعیت بالینی بیمار و سطح امکانات و پیشرفت‌های دانش پزشکی در آن زمینه برای بیمار خود دارد. دانستن پیش‌آگهی بیماری‌ها در اتخاذ تصمیمات درمانی تاثیر می‌گذارد. چرا که در صورتی که به طور مثال پیش‌آگهی یک بیماری خاص بسیار بد و ناامید کننده است در این صورت شاید بهتر باشد که به جای تاکید بر طول عمر بیمار بر افزایش کیفیت زندگی بیمار تاکید شود.

یک بررسی سی ساله گذشته نگر بین سال‌های ۱۹۷۵ تا ۲۰۰۵ که روی بیماران نفریت لوپوسی انجام شده است نشان می‌دهد که نرخ آنالیز زنده ماندن پنج ساله در دهه‌های اول و دوم کاهش ۶۰ درصدی داشته اما در دهه سوم

<sup>1</sup> Systemic lupus erythematosus

<sup>2</sup> Lupus nephritis

<sup>3</sup> Clinical management

<sup>4</sup> End stage kidney disease

<sup>5</sup> Kidney failure

<sup>6</sup> Kidney transplantation

<sup>7</sup> Dialysis

ثابت مانده است و میزان پیشرفت به بیماری مرحله نهایی کلیه نیز در دهه سوم افزایش داشته است. این نتایج نشان می‌دهد که مزایای درمان‌های متعارف سرکوب‌کننده سامانه ایمنی افزایش داشته است و پیشنهاد می‌دهد با توجه به عوارض جانبی این نوع درمان‌ها، باید روش‌های درمانی جدیدتری توسعه پیدا کند. بنابراین پیش‌بینی به موقع سرانجام نفرت لوپوسی موجب طرح درمانی بهتر و جلوگیری از نارسایی کلیه می‌شود [۲].

## ۲-۱ طرح مسئله

در روش‌های آنالیز بقا از ابزارهای استاندارد آماری مثل رگرسیون لجیستیک، مدل کاکس و موارد مشابهی استفاده می‌شوند. این روش‌های مدل‌های مبتنی بر جمعیت محسوب می‌شوند و پیش‌بینی در این روش‌ها از احتمال یا فاصله از برآورد جمعیت محاسبه می‌شود [۳].

از روش‌های داده‌کاوی می‌توان برای پیش‌بینی استفاده کرد. در این حالت دیدگاه متفاوت است و تاکید روی کشف دانش‌های مفید، معتبر و بدیع از داده‌ها است. پژوهش‌های بسیاری در زمینه پیش‌بینی بیماری نفرت لوپوسی انجام شده است که از روش‌های آماری استفاده می‌کنند [۴-۷]، همچنین رجیمهر و همکاران در سال ۲۰۰۲ برای پیش‌بینی نفرت لوپوسی از روش‌های داده‌کاوی استفاده کردند [۸].

رجیمهر و همکاران روش شبکه عصبی مصنوعی را برای پیش‌بینی نفرت لوپوسی در بیماران مبتلا به لوپوس ارائه کردند [۸]. در این پژوهش از شبکه عصبی برای پیش‌بینی استفاده شده است که این مدل با صحت ۶۶/۶۹ درصد پیش‌بینی را انجام می‌دهد. این مدل در مقایسه با مدل رگرسیون لجیستیک و پیش‌بینی پزشکان عملکرد بهتری را نشان می‌دهد. با این حال چندین فضای خالی تحقیقاتی در این پژوهش وجود دارد از جمله: عدم مشخص کردن ویژگی‌های تاثیرگذار، درصد صحت پائین در دسته‌بندی و مهمتر از همه این مدل به جز نتیجه پیش‌بینی هیچ دانش دیگری در اختیار کاربر قرار نمی‌دهد.



## ۱-۳ هدف

هدف اصلی در این پژوهش ارائه یک مدل مناسب برای کمک به درمان بیماران نفریت لوپوسی است. در این مدل برخلاف روش قبلی از الگوریتم‌های انتخاب ویژگی برای تعیین اهمیت ویژگی‌ها استفاده می‌شود. به علاوه برای بهبود صحت دسته‌بندی از ماشین بردار پشتیبان استفاده شده است. با این حال هر دو روش شبکه عصبی و ماشین بردار پشتیبان دارای یک عیب بزرگ هستند. به این صورت که روشی که این تکنیک‌ها برای یادگیری الگوها از داده‌ها به کار می‌برند برای کاربر قابل درک نیست، در نتیجه این تکنیک یک توضیح قابل درک برای دانش‌های یادگرفته شده فراهم نمی‌کنند. این مشکل یکی از موانع اصلی است که باعث شده است در برخی از کاربردهای عملی از این روش‌ها استفاده نشود [۹]، بنابراین توضیح یکی از مهمترین معیارهایی است که پذیرش این تکنیک‌ها توسط کاربران را تحت تاثیر قرار می‌دهد [۱۰-۱۳]. به طور مشابه، نشان داده شده است که توضیح یک تصمیم دسته‌بندی<sup>۱</sup>، یک نیازمندی حیاتی برای پذیرش این مدل جعبه سیاه<sup>۲</sup> توسط کاربران است، به خصوص در حوزه‌هایی مثل تشخیص و پیش‌بینی پزشکی [۹ و ۱۴-۱۶].

استخراج قواعد<sup>۳</sup> از ماشین بردار پشتیبان، ایده به دست آوردن قواعد از مدل به جای استخراج مستقیم آن‌ها از داده‌ها است. بنابراین از این طریق یک توضیح از الگوهای یاد گرفته شده و تعبیه شده در ساختار ماشین بردار پشتیبان یعنی بردارهای پشتیبان و پارامترهای آن‌ها و مؤلفه‌های دیگر، آشکار شده و به کاربران در یک شکل قابل درک ارائه می‌شود [۱۷].

در این پژوهش علاوه بر شناسایی ویژگی‌های تأثیرگذار از الگوریتم رقابت استعماری برای بهبود استخراج قواعد از ماشین بردار پشتیبان استفاده می‌شود.

---

<sup>1</sup> Classification

<sup>2</sup> Black box

<sup>3</sup> Rule extraction

## ۱-۴ ساختار پایان نامه

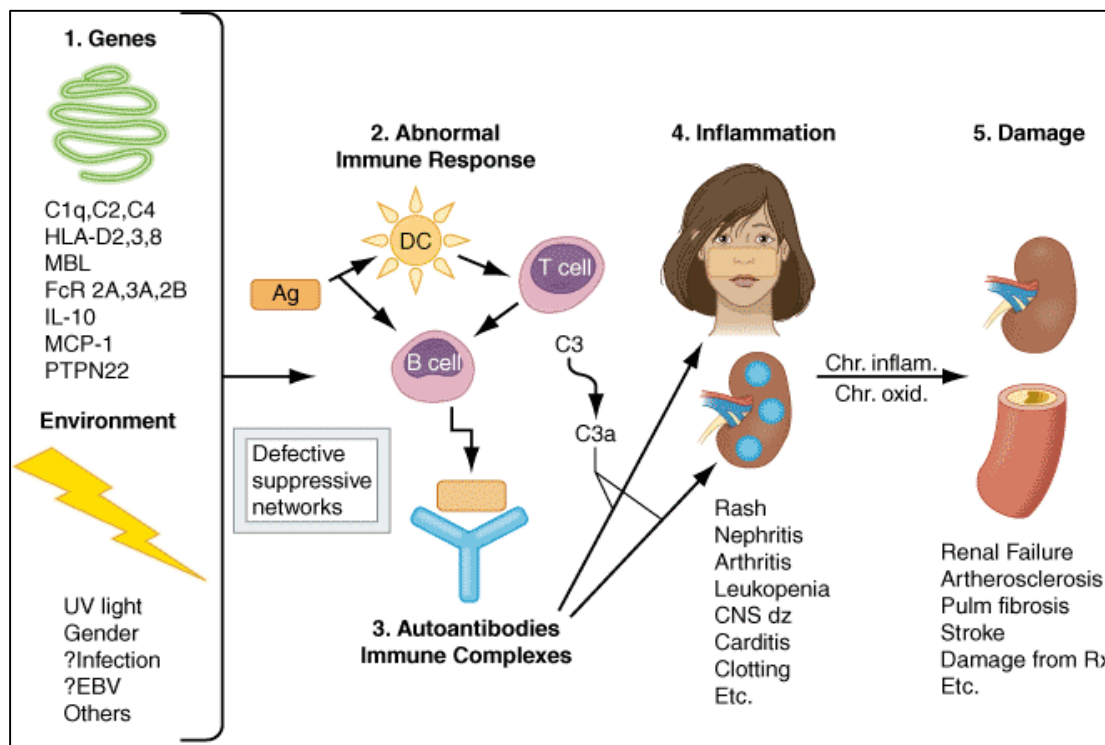
در ادامه در فصل دوم به ارائه اطلاعات پیش‌زمینه‌ای لازم در مورد بیماری نفریت لوپوسی پرداخته می‌شود. در فصل سوم، ادبیات مربوط به ماشین بردار پشتیبان و روش‌های استخراج قواعد مرور شده است به این صورت که ابتدا در مورد مفاهیم اولیه ماشین بردار پشتیبان مثل، ساختار ماشین، نحوه دسته‌بندی خطی و غیرخطی و توابع هسته بحث می‌شود. در ادامه فصل سوم، انگیزه‌های استخراج قواعد از ماشین بردار پشتیبان بیان می‌شوند و روش‌های به‌کاررفته در این زمینه مورد بررسی قرار می‌گیرند و ویژگی‌های مهم هر روش بیان می‌شود. در فصل چهارم یک روش پیشنهادی برای استخراج قواعد فازی از ماشین بردار پشتیبان ارائه می‌شود و در ادامه این فصل جزئیات مربوط به روش پیشنهادی ارائه می‌شود. در فصل پنجم ارزیابی روش پیشنهادی انجام می‌شود و فصل ششم به نتیجه‌گیری و کارهای آینده اختصاص داده شده است.

فصل دوم:

# ادبیات موضوع

در بدن انسان‌ها سامانه ایمنی نقش مهمی در نحوه بروز بیماری‌ها و بهبودی پس از آن دارد اما گاهی همین سامانه ایمنی و اختلال در کارکرد آن می‌تواند منجر به بروز بیماری‌های موسوم به بیماری‌های خود ایمنی<sup>۱</sup> گردد. لوپوس منتشر یا لوپوس اریتماتوس سیستمیک<sup>۲</sup> یک نوع بیماری خود ایمنی چند ارگانی است که هنوز به عنوان یک بیماری با علت نامعلوم شناخته می‌شود. لوپوس اریتماتوس سیستمیک یا به اختصار لوپوس، نام لاتین و به معنی گرگ، یک بیماری پیچیده و منحصر به فرد است که همراه با چالش‌های خاص خود است [۱].

این بیماری غالباً زنان را و به‌خصوص در دوران بارداری تحت تأثیر قرار می‌دهد و ابتلا به این بیماری در زنان هشت برابر مردان است. عوامل محیطی و ژنتیکی در ایجاد این بیماری مؤثر هستند. به‌علاوه افرادی که استعداد ژنتیکی دارند و محرک‌های زیست‌محیطی را که منجر به از دست دادن غیرقابل برگشت ایمنی است را تجربه کرده‌اند، در معرض خطر ابتلا به این بیماری قرار می‌گیرند [۱]. مراحل ایجاد لوپوس در شکل شماره ۱-۲ نشان داده شده است.



شکل ۱-۲ مراحل ایجاد لوپوس [۱۹]

<sup>1</sup> Auto immune

<sup>2</sup> Systemic lupus erythematosus