

بسم الله الرحمن الرحيم



مؤسسه آموزش عالی کوشیار (غیردولتی - غیرانتفاعی)

گروه : مهندسی کامپیوتر

گرایش : نرم افزار

بهینه سازی شبکه های عصبی فازی با استفاده از الگوریتم های تکاملی برای  
مدلسازی ریسک سرمایه گذاری

از

مهدی نظری ویشکائی

استاد راهنما

دکتر سودابه پورذاکر عربانی

استاد مشاور

دکتر علی اکبر نجفی

شهریور ۱۴۰۰

## تقدیم :

سپاس و ستایش مرخدای را جل و جلاله که آثار قدرت او بر چهره روز روشن، تابان است و انوار حکمت او در دل شب تار، درفشان. آفریدگاری که خویشتن را به ما شناساند و درهای علم را بر ما گشود و عمری و فرصتی عطا فرمود تا بدان، بنده ضعیف خویش را در طریق علم و معرفت بیازماید.

ماحصل آموخته هایم را تقدیم می کنم به مادرم، آنکه آفتاب مهرش در آستانه قلبم، همچنان پابرجاست و هرگز غروب نخواهد کرد و به استوارترین تکیه گاهم ، به سبزترین نگاه زندگیم ، خانواده ی عزیزم که هرچه آموختم در مکتب عشق شما آموختم و هرچه بکوشم قطره ای از دریای بی کران مهربانیتان را سپاس نتوانم بگویم .امروز هستی ام به امید شماست و فردا کلید باغ بهشتم رضای شما بوسه بر دستان پرمهرتان .

## تقدیر و تشکر

اکنون که به یاری پروردگار و یاری و راهنمایی اساتید بزرگ موفق به پایان این رساله شده‌ام وظیفه خود دانسته که نهایت سپاسگزاری را از تمامی عزیزانی که در این راه به من کمک کرده‌اند را به عمل آورم:

در آغاز از استاد بزرگ سرکار خانم دکتر پورذاکر عربانی که راهنمایی این پایانامه را به عهده داشته‌اند کمال تشکر را دارم.

از جناب آقایان دکتر نجفی که استاد مشاور این پایانامه بوده‌اند نیز قدردانی می‌نمایم.

از داوران گرامی که زحمت داوری و تصحیح این پایانامه را به عهده داشتند کمال سپاس را دارم.

خالصانه از تمامی اساتید و معلمان و مدرسانی که در مقاطع مختلف تحصیلی به من علم آموخته و مرا از سرچشمه دانایی سیراب کرده‌اند متشکرم.

و در پایان این پایان نامه را تقدیم می‌کنم به مادرم که مهر آسمانی شان آرام بخش آلام زمینی ام است و نفس خیرشان و دعای روح پرورشان بدرقه ی راه و مرا در این راه استوار و ثابت قدم نموده است.

## فهرست مطالب

عنوان	شماره
<u>صفحه</u>	
فصل اول.....	۲
۱-۱ مقدمه.....	۲
۱-۲ بیان مسئله.....	۲
۱-۳ ضرورت مسئله.....	۳
۱-۴ اهداف پژوهش.....	۳
۱-۵ پرسشهای پژوهش.....	۳
۱-۶ فرضیه‌های پژوهش.....	۳
۱-۷ ساختار پایان‌نامه.....	۴
فصل دوم.....	۵
۲-۱ مقدمه.....	۶
۲-۲ یادگیری ماشین.....	۶
۲-۲-۱ خوشه‌بندی.....	۷
۲-۲-۲ درخت تصمیم.....	۷
۲-۲-۲ ماشین پشتیبانی تصمیم (SVM).....	۸
۲-۲-۱ یادگیری ژرف.....	۹
۲-۲-۱ یادگیری تقویتی.....	۱۴
۲-۳ الگوریتم‌های تکاملی.....	۱۴
۲-۳-۱ الگوریتم ژنتیک.....	۱۶
۲-۳-۲ تکامل تفاضلی.....	۱۶
۲-۳-۳ الگوریتم بهینه‌سازی کلونی مورچگان.....	۱۶
۲-۳-۳ الگوریتم بهینه‌سازی ازدحام ذرات.....	۱۷
۲-۳-۴ الگوریتم کلونی زنبور عسل.....	۱۸
۲-۴ پیشینه تحقیق.....	۱۹
فصل سوم.....	۲۱
۳-۱ مقدمه.....	۲۶
۳-۲ شبکه‌های عصبی فازی.....	۲۶
۳-۳ الگوریتم تکامل تفاضلی تطبیقی.....	۲۷

۳-۴	متدولوژی پیشنهادی.....	۲۸
<b>فصل چهارم.....</b>		
۴-۱	مقدمه.....	۳۹
۴-۲	تنظیمات آزمایشها.....	۳۱
۴-۳	مجموعه داده‌ها.....	۳۱
۴-۴	معیارهای ارزیابی.....	۳۲
۴-۵	نتایج و تحلیل.....	۳۴
<b>فصل پنجم.....</b>		
۵-۱	نتیجه‌گیری.....	۴۱
۵-۲	پیشنهادهای.....	۴۳
<b>منابع و مآخذ.....</b>		
<b>فهرست منابع فارسی.....</b>		
<b>فهرست منابع انگلیسی.....</b>		

## فهرست جدول‌ها

جدول ۱-۲: تحلیلی بر روش‌های مطرح‌شده برای ریسک سرمایه‌گذاری..... ۱۹

جدول ۱-۴: ویژگی‌های داده‌ای موردبررسی ..... ۳۳

جدول ۲-۴: عملکرد مدل‌ها برای پیش‌بینی مقدار ریسک بازگشت اعتبار..... ۳۵

## فهرست شکل‌ها

- شکل ۲-۱: طبقه‌بندی یادگیری ماشین..... ۶
- شکل ۲-۲: نمای کلی ساخت مدل با استفاده از الگوریتم‌های یادگیری ماشین..... ۷
- شکل ۲-۳: نمایش گرافیکی یک درخت تصمیم..... ۸
- شکل ۲-۴: دسته‌بندی داده‌ها با استفاده از SVM..... ۹
- شکل ۲-۵: یک شبکه عصبی ساده با سه لایه..... ۹
- شکل ۲-۶: الگوریتم یادگیری بازگشت به عقب..... ۱۰
- شکل ۲-۷: معماری اصلی مدل شبکه عصبی پیچشی برای طبقه‌بندی چند کلاسه..... ۱۱
- شکل ۲-۸: نمودار کلی از رمزگذار خودکار..... ۱۲
- شکل ۲-۹: نمونه از شبکه SAE..... ۱۲
- شکل ۲-۱۰: معماری DBN که توسط RBM ساخته شده‌اند..... ۱۳
- شکل ۲-۱۱: شبکه عصبی بازگشتی..... ۱۳
- شکل ۲-۱۲: یک فرایند کلی از یادگیری تقویتی..... ۱۴
- شکل ۲-۱۳: طبقه‌بندی الگوریتم‌های تکاملی..... ۱۵
- شکل ۲-۱۴: انتقال یک ذره..... ۱۸
- شکل ۳-۱: ساختار شبکه‌های عصبی فازی..... ۲۶
- شکل ۳-۲: چارچوب پیشنهادی..... ۲۹
- شکل ۳-۳: ساختار یک کروموزوم..... ۲۹
- شکل ۴-۱: معماری ANN طراحی شده برای این پژوهش..... ۳۱
- شکل ۴-۲: ساختار شبکه عصبی فازی طراحی شده برای پیش‌بینی ریسک..... ۳۲
- شکل ۴-۳: متوسط خطای مدل‌های پیش‌بینی از نظر تعداد تکرار..... ۳۵
- شکل ۴-۴: عملکرد مدل‌های هوش مصنوعی در پیش‌بینی مقدار ریسک بازگشت اعتبار..... ۳۶
- شکل ۴-۵: عملکرد مدل‌های هوش مصنوعی در پیش‌بینی مقدار ریسک سیستماتیک..... ۳۷
- شکل ۴-۶: تأثیر نرخ یادگیری بر روی عملکرد مدل‌های پیش‌بینی..... ۳۸
- شکل ۴-۷: تأثیر اندازه جمعیت بر روی عملکرد مدل FNNDE..... ۳۹
- شکل ۴-۸: عملکرد مدل‌های هوش مصنوعی در پیش‌بینی مقدار ریسک کلی..... ۴۰
- شکل ۴-۹: تأثیر نرخ یادگیری بر روی عملکرد مدل‌های پیش‌بینی..... ۴۱
- شکل ۴-۱۰: تأثیر اندازه جمعیت بر روی عملکرد مدل FNNDE..... ۴۱



بهینه‌سازی شبکه‌های عصبی فازی با استفاده از الگوریتم‌های تکاملی برای مدل‌سازی ریسک سرمایه‌گذاری

مهدی نظری

در دنیای مالی، مدیریت ریسک فرآیند شناسایی، تجزیه و تحلیل و پذیرش یا کاهش عدم اطمینان در تصمیمات سرمایه‌گذاری است. مدیریت ریسک زمانی اتفاق می‌افتد که یک سرمایه‌گذار یا مدیر صندوق پتانسیل زیان در یک سرمایه‌گذاری را مورد تجزیه و تحلیل و سپس تلاش کند با توجه به اهداف سرمایه‌گذاری و تحمل ریسک، اقدام مناسب (یا عدم اقدام) را انجام دهد. مدیریت ریسک در همه حوزه‌های مالی اتفاق می‌افتد. در این پژوهش، یک مدل پیش‌بینی ریسک سرمایه‌گذاری هوشمند با استفاده از تکنیک‌های هوش مصنوعی توسعه داده می‌شود که ترکیبی از شبکه‌های عصبی فازی (FNN) و رویکردهای تکاملی است. در روش پیشنهادی، الگوریتم تکامل تفاضلی تطبیقی ( $ADE^1$ ) به عنوان بهینه‌ساز برای الگوریتم FNN استفاده می‌شود. پارامترهای لایه‌های فازی سازی و عدم فازی سازی توسط ADE به صورت خودکار تنظیم می‌شوند. مدل پیشنهادی با استفاده از یک مجموعه داده استاندارد با مدل‌های شبکه‌های عصبی مصنوعی (ANN) و FNN مقایسه می‌شود. سه معیار ارزیابی RMSE، MAE و  $R^2$  برای این مقایسه‌ها استفاده می‌شود.

**کلید واژه :** شبکه عصبی فازی، بستر مالی، اندازه‌گیری ریسک، الگوریتم‌های تکاملی

<sup>1</sup> Adaptive Differential Evolution (ADE)

# فصل اول

## کلیات پژوهش

## ۱-۱ مقدمه

در سال‌های اخیر، دولت‌ها در تلاش برای ارتقا ساخت زیرساخت‌های شهری، توسعه اقتصادی محلی و توسعه شهرنشینی، به همراه استقرار تعداد زیادی از پلتفرم‌های سرمایه‌گذاری و تأمین مالی، وام‌های بانکی را برای تأمین مالی گرفته‌اند. از زمان اصلاح سیستم مدیریت مالیات، اندازه دولت مرکزی به تدریج افزایش یافته است [۱]. با این حال، اصلاحات فقط تقسیم مالکیت املاک بود، نه اختیارات اداری. بنابراین، در اصلاحات، مالکیت املاک در تمام سطوح به سمت مقامات بالاتر و اختیارات اداری به سمت مقامات پایین‌تر می‌رود. در این حالت، دولت‌های محلی تحت فشار مالی قرار دارند و با افزایش بدهی‌های ضمنی حقوق بازنشستگی کاهش می‌یابد [۲]. ماده ۲۸ قانون بودجه سال ۱۹۹۴ جمهوری خلق چین پیش‌بینی کرده است که بودجه‌های محلی در همه سطوح بر اساس اصل تأمین هزینه تأمین می‌شوند. دولت‌های محلی نباید اوراق قرضه دولت محلی را منتشر کنند مگر اینکه توسط قانون و شورای دولتی مقرر شده باشد [۳، ۴]. در این زمینه، در سال‌های اخیر، به منظور ارتقا توسعه اقتصادی محلی و ارتقا شهرنشینی، برخی از دولت‌های محلی نوآوری‌هایی را در سیستم‌های سرمایه‌گذاری و تأمین مالی انجام داده‌اند و به‌طور متوالی برخی از پلتفرم‌های سرمایه‌گذاری و مالی وابسته به دولت‌های محلی را ایجاد کرده‌اند، در نتیجه وام‌های بانکی را با پروژه‌های دولتی و کاهش فشار مالی اخذ کرده‌اند.

## ۱-۲ بیان مسئله

ریسک اعتبار اصلی‌ترین ریسکی است که مؤسسات مالی با آن روبرو هستند. ارزیابی ریسک اعتباری یک بانک به تلاش برای جستجوی گام‌هایی برای مقابله با ریسک بامطالعه ماهیت و تأثیر آن بر اساس برآورد کمی ریسک اعتباری اشاره دارد. بنابراین، برآورد کمی مقدمه‌ای برای ارزیابی ریسک است. بر اساس توافق بازل، بانک‌های تجاری می‌توانند ارزیابی ریسک را به ارزیابی داخلی و خارجی با توجه به موضوعات مختلف تقسیم کنند و هسته اصلی ارزیابی ایجاد یک مدل رتبه‌بندی اعتبار است [۵]. مدل‌های رتبه‌بندی را می‌توان به‌طور کلی به سه دسته تقسیم کرد: مدل قضاوت ذهنی، مدل آماری و مدل علی<sup>۱</sup>. مدل قضاوت ذهنی اغلب به قضاوت ذهنی کارشناسان اعتبار متکی است. مدل آماری به دو نوع تقسیم می‌شود: مدل پارامتریک و مدل غیرپارامتریک. مدل پارامتریک شامل تجزیه و تحلیل‌های تک متغیره و چند متغیره است، در حالی که مدل غیر پارامتریک، همانند هوش مصنوعی، از فناوری پیشرفته رایانه‌ای برای شبیه‌سازی تصمیم‌گیری انسان استفاده می‌کند. مدل علی به دلایل پیش فرض شرکت بر اساس ساختار سرمایه شرکت اشاره می‌کند. مدل قضاوت ذهنی اغلب از پرسشنامه‌های معمولی، نظرات کارشناسان، سیستم‌های خبره و سیستم‌های منطق فازی استفاده می‌کند. مدل‌های آماری پارامتری اغلب از تجزیه و تحلیل رگرسیون، تحلیل تفکیک، مدل‌های لجستیکی، و غیره استفاده می‌کنند. مدل‌های آماری غیر پارامتریک اغلب از درختان تصمیم، شبکه عصبی،  $KNN^2$  و غیره استفاده می‌کنند.

در عمل، متداول‌ترین ساخت مازول رتبه‌بندی اعتبار داخلی بانک، ترکیبی از مدل قضاوت ذهنی و دو نوع مدل دیگر است. شیوه‌ها و تحقیقات نشان می‌دهد که تجزیه و تحلیل تفکیک چند متغیره و مدل لجستیکی به‌طور گسترده در تکنیک‌های آماری استفاده می‌شود، در حالی که مدل‌های غیر پارامتریک همانند هوش مصنوعی جدید هستند.

<sup>1</sup> Causal Model<sup>2</sup> K-Nearest Neighbor (KNN)

## ۳-۱ ضرورت مسئله

مدل‌سازی ریسک سرمایه‌گذاری شامل فرآیند شناسایی ریسک‌های احتمالی در سرمایه‌گذاری و تجزیه و تحلیل آن‌ها و انجام اقدامات لازم برای جلوگیری از آن‌ها است. در مورد کسب و کارها، هنگامی که سرمایه‌گذاری مالی انجام می‌شود، مدیریت ریسک باید چنان کارآمد انجام شود که بتواند خطرات احتمالی اقتصادی، تأثیرات آن‌ها و راه‌های غلبه بر آن‌ها را شناسایی کند. مدیریت ریسک زمانی صورت می‌گیرد که یک سرمایه‌گذار یا مدیر صندوق کمی از ضررهای احتمالی را کمی کرده و اقدامات لازم را برای مقابله با ریسک موجود در سرمایه‌گذاری انجام دهد. سرمایه‌گذاری هنگامی که بازدهی آن تغییرات زیادی را نشان دهد، یک سرمایه‌گذاری خطرناک تلقی می‌شود. بازدهی مورد انتظار با توجه به نوع اوراق بهاداری که در آن سرمایه‌گذاری می‌شود متفاوت است. هرچه تغییرات بین بازدهی مورد انتظار و واقعی بیشتر باشد، خطرپذیری سرمایه‌گذاری بیشتر است. هنگام تصمیم‌گیری در مورد سرمایه‌گذاری، اندازه‌گیری سطح صحیح ریسک مرتبط با انواع مختلف اوراق بهادار موجود برای سرمایه‌گذاری ضروری است. درک بهتر ریسک در اشکال مختلف آن به سرمایه‌گذاران کمک می‌کند تا از فرصت‌های مختلف، معامله و هزینه‌های مختلف سرمایه‌گذاری مطلع شوند. مدل‌سازی ریسک در زندگی همه اجتناب‌ناپذیر است. بنابراین فرد باید توانایی مقابله با خطر و غلبه بر آن را داشته باشد. مزیت اصلی مدل‌سازی ریسک سرمایه‌گذاری این است که خطر ضرر را به حداقل می‌رساند. در صورت عملکرد ضعیف یک سرمایه‌گذاری، ممکن است سایر سرمایه‌گذاری‌ها عملکرد بهتری داشته باشند و در نتیجه از دست دادن سرمایه‌گذاری کاسته شود. مشکل اصلی در مدل‌سازی ریسک، دشواری در اجرای آن است زیرا جمع‌آوری اطلاعات مربوط به برنامه‌های استراتژیک زمان زیادی می‌خواهد. یک مدل ریسک به پیش‌بینی خطر کمک می‌کند.

## ۴-۱ اهداف پژوهش

- ارائه یک الگوریتم شبکه عصبی فازی جدید برای بهبود عملکرد مدل ریسک سرمایه‌گذاری.
- استفاده از الگوریتم‌های تکاملی به عنوان بهینه‌ساز الگوریتم شبکه‌های عصبی فازی.
- فرموله کردن یک تابع تناسب کارآمد مرتبط با مسئله پیش‌بینی با استفاده از معیارهایی همانند  $RMSE^1$ .

## ۵-۱ پرسش‌های پژوهش

- چگونه می‌توان یک مدل هوشمند برای مدل‌سازی ریسک سرمایه‌گذاری طراحی کرد؟
- آیا می‌توان عملکرد مدل ریسک سرمایه‌گذاری را بهبود داد؟
- آیا می‌توان چالش استفاده از مدل‌های هوش مصنوعی همانند شبکه عصبی کاهش داد؟

## ۶-۱ فرضیه‌های پژوهش

- ترکیب الگوریتم تکاملی و شبکه‌های عصبی فازی می‌تواند مدل‌سازی برای پیش‌بینی ریسک سرمایه‌گذاری را بهینه کند.
- به دلیل قابلیت یادگیری سیستم هوش مصنوعی بر اساس پارامترهای مختلف بهینه‌سازی الگوریتم‌ها، عملکرد مدل‌های پیش‌بینی کننده ریسک بالا می‌رود.
- استفاده از الگوریتم‌های تکاملی همانند ژنتیک، کلونی مورچگان، ازدحام ذرات به افزایش کارایی الگوریتم‌های مبتنی به شبکه‌های عصبی کمک می‌کند و باعث افزایش همگرایی و فرار از بهینه محلی می‌شود و در نتیجه مدل‌های بهتری آموزش داده می‌شود.

<sup>1</sup> Root Mean Square Error (RSME)

## ۷-۱ ساختار پایان‌نامه

رئوس مطالب پایان‌نامه به شرح ذیل است. در فصل دوم، ابتدا تکنیک‌های یادگیری ماشین تشریح می‌شوند که به سه دسته یادگیری ماشین مرسوم (همانند خوشه‌بندی و درخت تصمیم)، یادگیری ژرف و یادگیری تقویتی تقسیم می‌شوند. سپس الگوریتم‌های تکاملی مشهور (همانند ژنتیک، تفاضل تکاملی، و بهینه کلونی مورچگان) ذکر می‌شوند. در نهایت، روش‌های پیشین ارائه‌شده برای پیش‌بینی ریسک سرمایه‌گذاری که از هوش مصنوعی و یادگیری ماشین استفاده می‌کنند بیان می‌شوند. در فصل سوم یک روش جدید برای پیش‌بینی ریسک سرمایه‌گذاری پیشنهادی می‌شود که ترکیبی از شبکه‌های عصبی فازی و الگوریتم تکامل تفاضلی است. در فصل چهارم مدل پیشنهادی پیاده‌سازی و با استفاده از داده‌های استاندارد با روش‌های مشابه مقایسه و نتایج ارزیابی می‌شود و سرانجام در فصل پنجم یک نتیجه‌گیری کلی و مختصر از مطالب مطرح‌شده در این پایان‌نامه ارائه می‌شود.

# فصل دوم

## بررسی مبانی نظری و

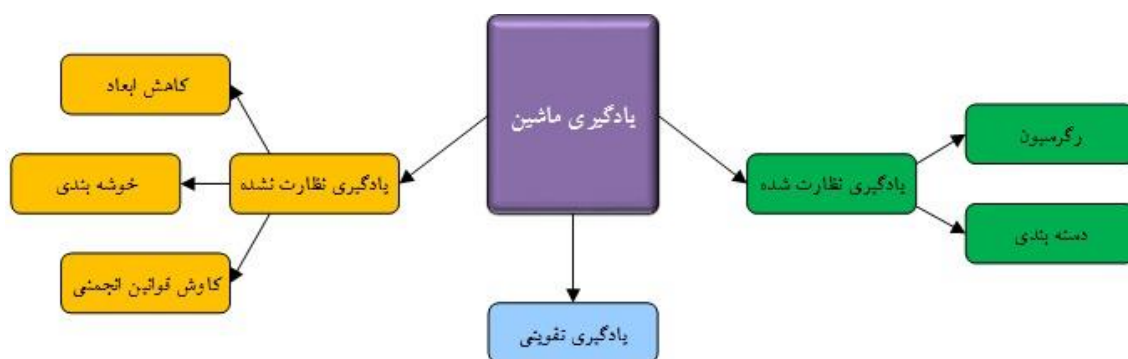
## پیشینه پژوهش

## ۲-۱ مقدمه

برآورد ریسک یکی از جنبه‌های مهم تصمیم‌گیری سرمایه‌گذاری است. از طریق پیش‌بینی ریسک، ریسک‌هایی که سرمایه‌گذاران با آن روبرو هستند شناخته می‌شوند، زیرا ریسک معیاری برای سرمایه‌گذاری است. علاوه بر بازدهی، سرمایه‌گذاران همیشه ریسک سرمایه‌گذاری را در نظر می‌گیرند، زیرا سرمایه‌گذاران افراد منطقی هستند، یعنی افرادی که همیشه معامله بین بازدهی و ریسک را در نظر می‌گیرند. در یک سطح مشخص از بازدهی، سرمایه‌گذاران تمایل به انتخاب سرمایه‌گذاری با کمترین سطح ریسک دارند. برعکس، در سطح معینی از ریسک، سرمایه‌گذاران تمایل به انتخاب سرمایه‌گذاری با بالاترین نرخ بازده دارند.

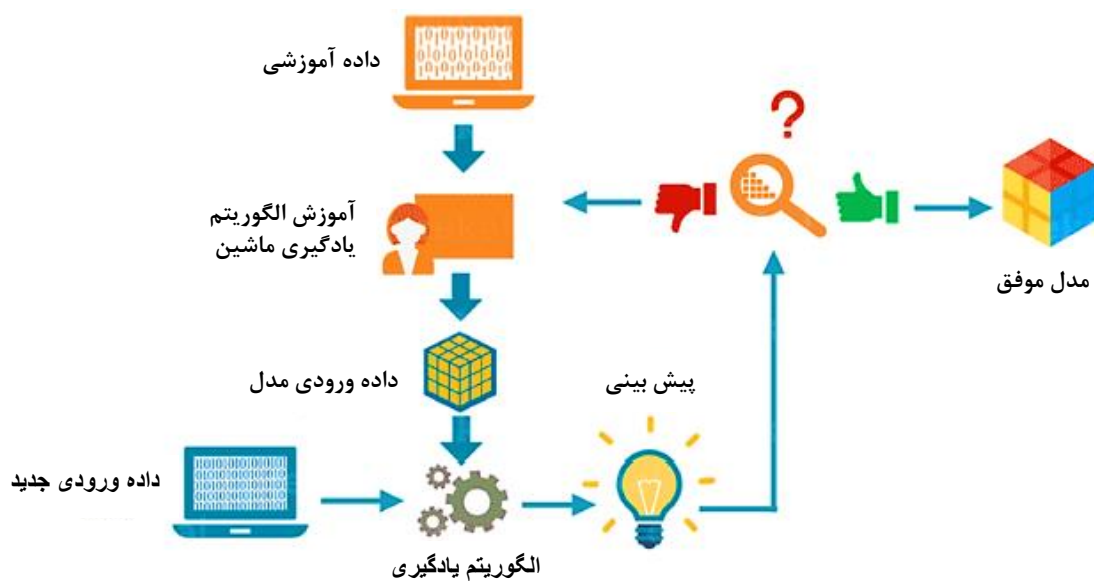
## ۲-۲ یادگیری ماشین

یادگیری ماشین، زیرمجموعه‌ای از هوش مصنوعی، الگوریتم‌های ریاضی را برای استخراج الگوها با استفاده از بهینه‌سازی، آمار و روش‌های کشف دانش به کار می‌برد. الگوریتم‌های یادگیری ماشین از سه دسته اصلی تشکیل شده است (شکل ۲-۱): (۱) یادگیری تحت نظارت، (۲) یادگیری بدون نظارت و (۳) یادگیری تقویتی.



شکل ۲-۱: طبقه‌بندی یادگیری ماشین [۶]

یادگیری تحت نظارت، یکی از روش‌های اصلی پردازش داده‌های یادگیری ماشین است که سعی در یافتن رابطه بین مجموعه‌ای از ورودی‌ها و خروجی‌هایی دارد که هنگام آموزش سیستم ارائه می‌شود. الگوریتم‌های یادگیری تحت نظارت مدلی را ایجاد می‌کنند که روابط بین ویژگی‌های ورودی و خروجی‌های هدف پیش‌بینی را نشان می‌دهد. این الگوریتم‌ها شامل دودسته اصلی هستند: (۱) طبقه‌بندی و (۲) پیش‌بینی. هر دو برای پیش‌بینی هستند. تنها تفاوت متغیرهای هدف (پاسخ) است. در طبقه‌بندی، متغیر هدف، دسته‌ها (برچسب‌های کلاس) است، مانند مسئله کلاس باینری و چندطبقه. باینری، در پیش‌بینی، متغیر هدف پیوسته است. شکل ۲-۲ نمای کلی ساخت مدل نظارت‌شده با استفاده از الگوریتم‌های یادگیری ماشین را نشان می‌دهد. طبقه‌بندی در کاربردهای مختلف همانند تشخیص تقلب، بازاریابی هدف، پیش‌بینی عملکرد، صنعت و تشخیص پزشکی استفاده می‌شود. برای مثال، یک مدل طبقه‌بندی می‌تواند برای طبقه‌بندی وام‌های بانکی به عنوان امن یا خطرناک ایجاد شود. افسر وام بانکی نیاز به تجزیه و تحلیل داده‌های خود دارد تا بتواند بداند که کدام وام‌گیرندگان "امن" هستند و کدام یک برای بانک خطرناک هستند.



شکل ۲-۲: نمای کلی ساخت مدل با استفاده از الگوریتم‌های یادگیری ماشین [۶]

### ۲-۲-۱ خوشه‌بندی

خوشه‌بندی یک روش یادگیری بدون نظارت است که اشیاء داده‌ای بدون برچسب را با توجه به شباهت‌های موجود در اشیاء گروه‌بندی می‌کند بطوریکه که اشیاء درون یک خوشه بسیار شبیه یکدیگر هستند اما با اشیاء موجود در خوشه‌های دیگر متفاوت هستند. مشخصه اصلی خوشه‌بندی این است که هیچ دانش قبلی از توزیع داده‌ها وجود ندارد. خوشه‌بندی مقطعی<sup>۱</sup> و خوشه‌بندی سلسله‌مراتبی<sup>۲</sup> دو گروه اصلی از الگوریتم‌های خوشه‌بندی هستند. روش‌های خوشه‌بندی مقطعی مجموعه داده‌ها را بر اساس معیارهای تناسب در تعداد تکرارهای از پیش تعریف‌شده به گروه‌های خاصی تقسیم می‌کنند. سادگی و هزینه محاسباتی پایین دو مزیت اصلی الگوریتم‌های خوشه‌بندی مقطعی هستند، همانند *k-means*. باین‌حال، دو چالش اصلی برای این الگوریتم‌ها وجود دارد. اول، آن‌ها به مقداردی اولیه بسیار حساس هستند و احتمال گرفتار شدن در بهینه‌های محلی را افزایش می‌دهند. دوم، لازم است قبل از شروع الگوریتم خوشه‌بندی تعداد خوشه‌ها مشخص شود. تعداد خوشه‌های کمتر منجر به از دست رفتن اطلاعات اصلی پنهان می‌شوند. در مقابل، تعداد خوشه زیاد می‌تواند منجر به همگنی زیاد مجموعه خوشه شود [۷].

در روش سلسله‌مراتبی از توپولوژی درخت برای نشان دادن روابط بین مجموعه‌های خوشه‌ای استفاده می‌شود. روش‌های سلسله‌مراتبی می‌توانند داده‌ها را با استفاده از یک رویکرد تقسیم‌کننده یا یک رویکرد تجمیعی خوشه‌بندی کنند. روش اول خوشه‌های کوچک‌تر را به بزرگ‌ترها ادغام می‌کند، درحالی‌که روش دوم خوشه‌های بزرگ را به کوچک‌ترها تقسیم می‌کند. خوشه‌بندی سلسله‌مراتبی یک مزیت نسبت به خوشه‌بندی مقطعی دارد و آن این است که مشخص کردن تعداد خوشه‌ها نیازی به تنظیم قبلی ندارد. باین‌حال، مشکل خوشه‌بندی سلسله‌مراتبی این است که هر عنصر می‌تواند فقط به یک خوشه اختصاص یابد. بنابراین، هنگام جداسازی خوشه‌های همپوشان، عملکرد این روش ضعیف است [۸].

### ۲-۲-۲ درخت تصمیم

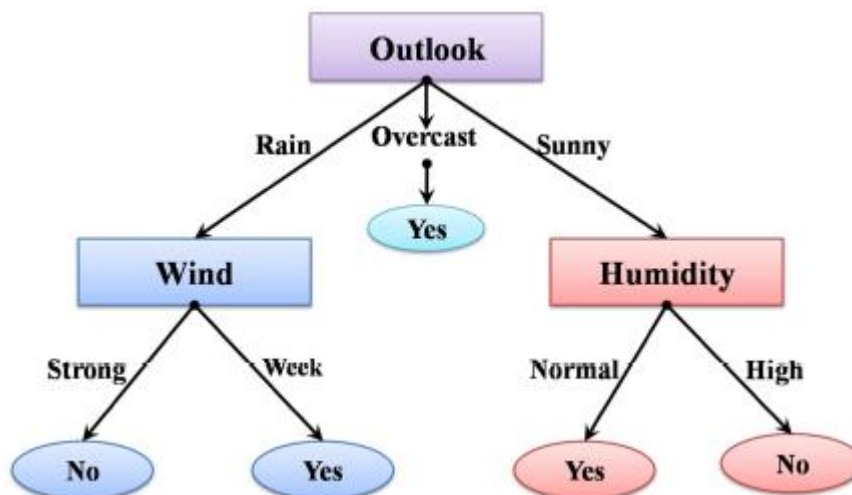
درختان تصمیم یک روش یادگیری ماشین نظارت‌شده برای طبقه‌بندی بر اساس مجموعه‌ای از قوانین IF-THEN برای افزایش خوانایی است. یک درخت تصمیم شامل دو نوع گره است: گره‌های برگ و گره‌های تصمیم‌گیری است. درخت تصمیم با استفاده از ایجاد یک مدل آموزش بر اساس قوانین تصمیم‌گیری استنباط شده از داده‌های آموزشی، از یک کلاس یا هدف استفاده می‌کند. یک نمایش گرافیکی از یک درخت تصمیم در شکل ۲-۳ نشان داده شده است. مزایای

<sup>۱</sup> Partitional clustering

<sup>۲</sup> hierarchical clustering



اصلی درخت تصمیم شفافیت است که ابهام در تصمیم‌گیری را کاهش می‌دهد و امکان تجزیه و تحلیل جامع را فراهم می‌کند [۹].



شکل ۲-۳: نمایش گرافیکی یک درخت تصمیم [۹]

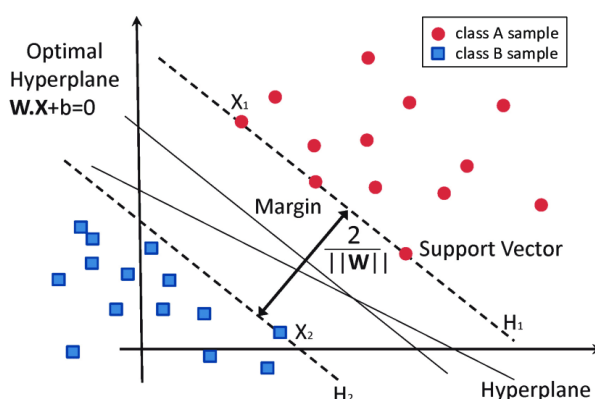
### ۲-۲-۲ ماشین پشتیبانی تصمیم (SVM)

ماشین بردار پشتیبانی ( $SVM^1$ ) یک طبقه‌بندی کننده نظارت شده است که یک ابرصفحه<sup>۲</sup> مطلوب برای دسته‌بندی داده‌ها می‌یابد. فضای الگوی اصلی در SVM ابتدا با استفاده از توابع غیرخطی در یک فضای ویژگی با ابعاد بالا ترسیم می‌شود. سپس، یک ابرصفحه جداکننده بهینه از فضای ویژگی ایجاد می‌شود. ماشین بردار پشتیبانی بهترین طبقه‌بندی را با استفاده از ابرصفحه انجام می‌دهد. بسیاری از داده‌های آموزشی اضافی پس از ایجاد یک مرز شناسایی می‌شوند. برای یافتن یک مرز به مجموعه‌ای از نقاط نیاز است و نقاطی که برای یافتن مرز استفاده می‌شوند به عنوان بردارهای پشتیبانی نامیده می‌شوند. ماشین بردار پشتیبانی بهترین طبقه‌بندی را از یک مجموعه داده شده ارائه می‌دهد. بنابراین، پیچیدگی مدل ماشین بردار پشتیبانی تحت تأثیر تعداد ویژگی‌های موجود در داده‌های آموزشی قرار نمی‌گیرد. به همین دلیل، ماشین بردار پشتیبانی برای انجام کارهای یادگیری که تعداد ویژگی‌ها با توجه به تعداد موارد آموزشی زیاد است، مناسب است.

ماشین بردار پشتیبانی در ابتدا با موفقیت برای مسائل طبقه‌بندی دودویی اعمال شد و سپس برای پیش‌بینی رگرسیون و پیش‌بینی سری زمانی مورد استفاده قرار گرفت. هنگام استفاده از SVM برای طبقه‌بندی چند کلاسه، مسئله از طریق روش تجزیه به چندین زیر مسئله دودویی تقسیم می‌شود. هر زیرمسئله توسط SVM حل می‌شود و خروجی پیش‌بینی‌ها باهم ترکیب می‌شوند. ماشین بردار پشتیبانی نه تنها برای طبقه‌بندی و مسائل رگرسیونی بلکه برای یادگیری بدون نظارت نیز استفاده می‌شود. شکل ۲-۴ نمای کلی از دسته‌بندی دودویی با استفاده از الگوریتم SVM را نشان می‌دهد.

<sup>1</sup> Support Vector Machine (SVM)

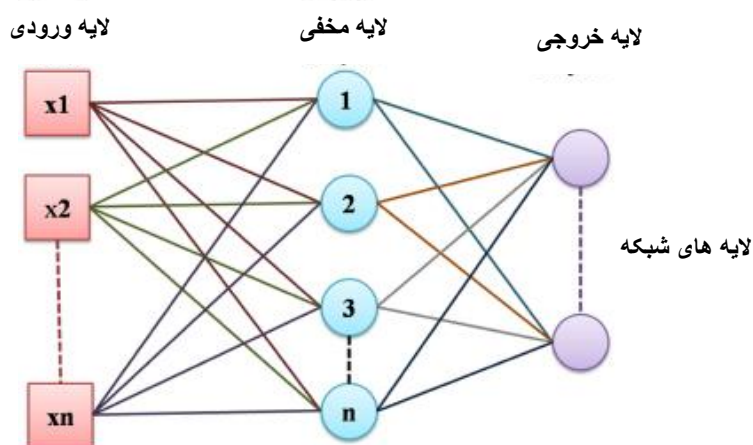
<sup>2</sup> Hyperplane



شکل ۲-۴: دسته‌بندی داده‌ها با استفاده از SVM [۱۰]

### ۲-۲-۱ یادگیری ژرف

یک شبکه عصبی<sup>۱</sup> استاندارد شامل بسیاری از پردازنده‌های متصل به نام نورون است. نورون‌های ورودی مقادیر را از محیط دریافت می‌کنند، درحالی‌که نورون‌های دیگر مقادیر را از طریق اتصالات وزنی از سلول‌های عصبی قبلی دریافت می‌کنند. شبکه عصبی اغلب روی لایه‌ها کار می‌کند، این لایه‌ها با گره‌ها به یکدیگر متصل می‌شوند. شکل ۲-۵ ساختار لایه اصلی یک شبکه عصبی را نشان می‌دهد. هر شبکه عصبی شامل سه نوع لایه به نام لایه ورودی، یک یا چند لایه مخفی و لایه‌های خروجی است. یک شبکه عصبی به‌راحتی مجموعه داده‌های پیچیده و غیرخطی را طبقه‌بندی می‌کند و همانند سایر روش‌های طبقه‌بندی محدودیتی برای ورودی‌ها وجود ندارد.



شکل ۲-۵: یک شبکه عصبی ساده با سه لایه [۹]

تمرکز اصلی در فرآیند یادگیری وزن‌ها، یافتن یک مجموعه مطلوب یا کاملاً مناسب از وزن اتصال برای یک شبکه با استفاده از انتشار به عقب است. انتشار به عقب، یک الگوریتم کارآمد برای محاسبه شیب تابع است و روش غالب برای آموزش شبکه‌های عصبی در زمینه شبکه‌های عصبی است. الگوریتم شکل ۲-۶ یک الگوریتم انتشار به عقب را نشان می‌دهد.

<sup>۱</sup> Neural Network

**ALGORITHM 2: Back Propagation**

Input: Data  $D = \{(x_k, y_k)\}_{k=1}^n$ , learning rate  $\eta$ , cost function  $E$

Output: optimal weights  $w$

1. Randomly initialize all weights and threshold
2. while *Stopping criteria is not met* do
3. for all  $(x^{(i)}, y^{(i)}) \in D$  do
4. Compute  $w_{ij} = w_{ij} - \eta \frac{\partial E}{\partial w_{ij}}$
5. End
6. End

**شکل ۲-۶: الگوریتم یادگیری بازگشت به عقب**

یادگیری ژرف زیرمجموعه‌ای از شبکه‌های عصبی است. رویکردهای یادگیری ژرف شامل روش‌های بازنمایی یادگیری داده با چندین لایه (بین لایه ورودی و لایه خروجی) است. ترکیب لایه‌های شبکه با ماژول‌های غیرخطی تبدیل نمایش از لایه پایین به لایه بالاتر را ساده می‌کند. این روش از الگوهای ارتباطی و پردازش اطلاعات در سیستم‌های عصبی انسان الهام گرفته شده است. مزایای اصلی یادگیری ژرف استخراج ویژگی‌های سطح بالا از داده‌ها، کار با داده‌های برچسب دار یا بدون برچسب است. این می‌تواند در حوزه‌های مختلف مانند بیوانفورماتیک، تجزیه و تحلیل شبکه‌های اجتماعی، هوش تجاری، پردازش تصویر پزشکی و تشخیص گفتار مفید باشد. [۵].

از زمان یادگیری ژرف در مجله Science در سال ۲۰۰۶، این موضوع به یک موضوع تحقیقاتی بسیار داغ در جامعه یادگیری ماشین تبدیل شده است. در چند سال گذشته مدل‌های مختلف یادگیری ژرف ایجاد شده است. متداول‌ترین مدل‌های یادگیری ژرف شامل خودرنگذار انباشته ( $SAE^1$ )، شبکه باور عمیق ( $DBN^2$ )، شبکه عصبی پیچشی ( $CNN^3$ ) و شبکه عصبی بازگشتی ( $RNN^4$ ) است که از مدل‌های پرکاربرد هستند. بیشتر مدل‌های یادگیری ژرف دیگر می‌توانند از این چهار معماری الگو بگیرند. در این بخش، چهار مدل یادگیری ژرف را به‌طور خلاصه مرور می‌کنیم.

**۱-۴-۲-۲ شبکه‌های عصبی پیچشی**

شبکه عصبی پیچشی (CNN) یکی از تکنیک‌های یادگیری ژرف است که به‌طور گسترده برای طبقه‌بندی و تشخیص تصویر استفاده می‌شود زیرا یادگیری ژرف برای داده‌های دوبعدی مناسب‌تر است. یادگیری ژرف از سه لایه اصلی تشکیل شده است: لایه‌های کانولوشن، لایه‌های نمونه‌برداری (ادغام<sup>۵</sup>) و لایه اتصال کامل. لایه کانولوشن نمایش ویژگی ورودی را می‌آموزد. هر لایه کانولوشن از چندین هسته کانولوشنی برای نگاشت ویژگی‌ها تشکیل شده است. هر واحد لایه پیچشی بر اساس وزن به نگاشت‌های محلی در نگاشت ویژگی‌های لایه قبلی متصل می‌شود. اغلب یک تابع غیرخطی همانند واحد خطی اصلاح شده ( $ReLU^6$ )،  $f(z) = \max(z, 0)$ ، پس از لایه کانولوشن استفاده می‌شود تا مقادیر منفی حاصل از مجموع وزنی محاسبه شده از لایه قبلی را به صفر تبدیل کند. تابع ReLU همگرایی در شبکه‌های دارای لایه‌های زیاد را سریع‌تر می‌کند. در یادگیری ژرف، یک لایه ادغام بعد از یک یا چند لایه کانولوشن اضافه می‌شود تا ابعاد نگاشت ویژگی‌ها

<sup>1</sup> Stacked Auto-Encoder (SAE)

<sup>2</sup> Deep Belief Network (DBN)

<sup>3</sup> convolutional Neural Network (CNN)

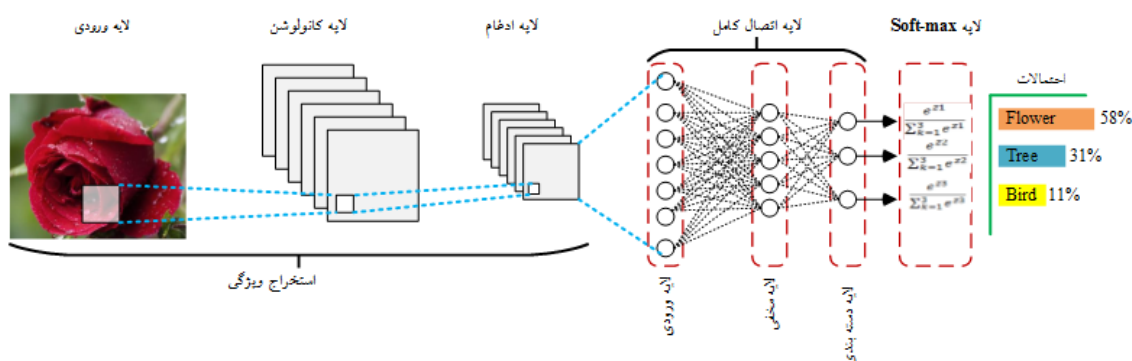
<sup>4</sup> Recurrent Neural Network (RNN)

<sup>5</sup> Pooling

<sup>6</sup> Rectified Linear Unit (ReLU)

کاهش یابد و ویژگی‌های معنایی مشابه را در یک مقدار ادغام کند. دو تکنیک ادغام میانگین<sup>۱</sup> و ادغام حداکثری<sup>۲</sup> اغلب استفاده می‌شوند.

پس از لایه‌های کانولوشن و ادغام، یک لایه اتصال کامل و یک لایه softmax برای طبقه‌بندی و تشخیص در نظر گرفته می‌شود. تابع softmax بردار وزن‌های تجمیع شده را می‌گیرد و به تعداد کلاس‌ها مقادیر احتمالی متناسب را برمی‌گرداند. کلاس با بالاترین احتمال به عنوان یک کلاس پیش‌بینی شده در نظر گرفته می‌شود. شکل ۲-۷ مراحل طبقه‌بندی چند کلاسه مبتنی بر CNN را نشان می‌دهد.



شکل ۲-۷: معماری اصلی مدل شبکه عصبی پیچشی برای طبقه‌بندی چند کلاسه

## ۲-۲-۴-۲ خودرمزگذار

خودرمزگذار<sup>۳</sup> یک شبکه عصبی سه لایه است که از الگوریتم بازگشت به عقب بدون نظارت استفاده می‌کند زیرا نیازی به برچسب‌گذاری یا پردازش قبلی ندارد (شکل ۲-۸). هدف از خودرمزگذار به حداقل رساندن خطا بین داده‌های ورودی و بازسازی شده (خروجی) است. اندازه لایه ورودی و لایه خروجی (تعداد نورون‌ها) برابر است. در خودرمزگذار، داده‌های ورودی به نمایشی با ابعاد پایین‌تر فشرده می‌شوند و سپس خروجی از بازنمایی بازسازی شده تولید می‌شود. هنگامی که تعداد سلول‌های عصبی در لایه کدگذاری کمتر از ورودی باشد، می‌توان بازنمایی با ابعاد پایین‌تر را به دست آورد. تابع رمزگذاری (معادله ۲-۱) بردار ورودی  $x \in [0,1]^d$  را به یک بازنمایی پنهان شده  $h$  تبدیل می‌کند.

$$h = f(W_1 x + b_1) \quad (2-1)$$

در این معادله، اصطلاح  $f$  یک تابع فعال‌ساز غیرخطی است،  $W_1 \in \mathbb{R}^{d \times z}$  و  $b_1 \in \mathbb{R}^z$  به ترتیب ماتریس وزن و بردار بایاس لایه پنهان هستند. عبارات  $d$  و  $z$  به ترتیب تعداد نورون‌های موجود در لایه پنهان و لایه خروجی هستند. یک تابع logistic sigmoid (رابطه ۲-۲) اغلب به عنوان تابع فعال‌ساز استفاده می‌شود.

$$f(x) = 1/(1 + \exp(-x)) \quad (2-2)$$

تابع رمزگشایی<sup>۴</sup> (معادله ۲-۳)  $h$  را برای تولید بردار خروجی  $x' \in [0,1]^d$  بازسازی می‌کند.

$$x' = g(W_2 h + b_2) \quad (2-3)$$

جاییکه  $g(\cdot)$  یک تابع فعال‌ساز رمزگشا غیرخطی است. اصطلاحات  $W_2 \in \mathbb{R}^{z \times d}$  و  $b_2 \in \mathbb{R}^d$  به ترتیب ماتریس وزن و بردار بایاس لایه خروجی هستند.

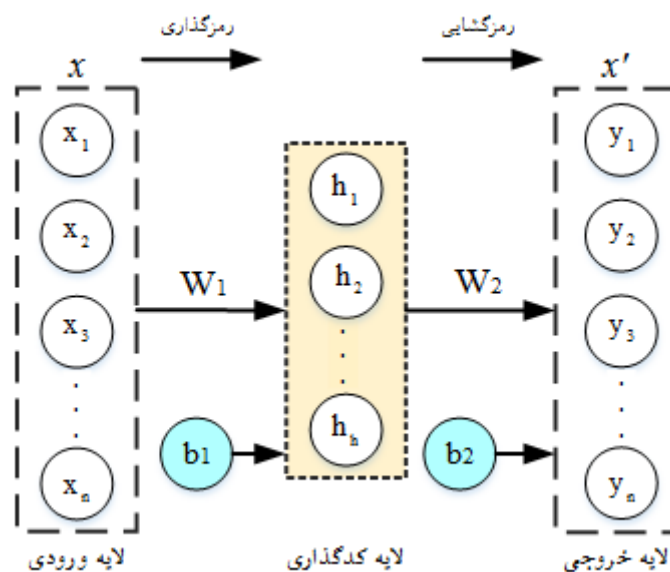
<sup>1</sup> Average pooling

<sup>2</sup> Max pooling

<sup>3</sup> Auto-encoder

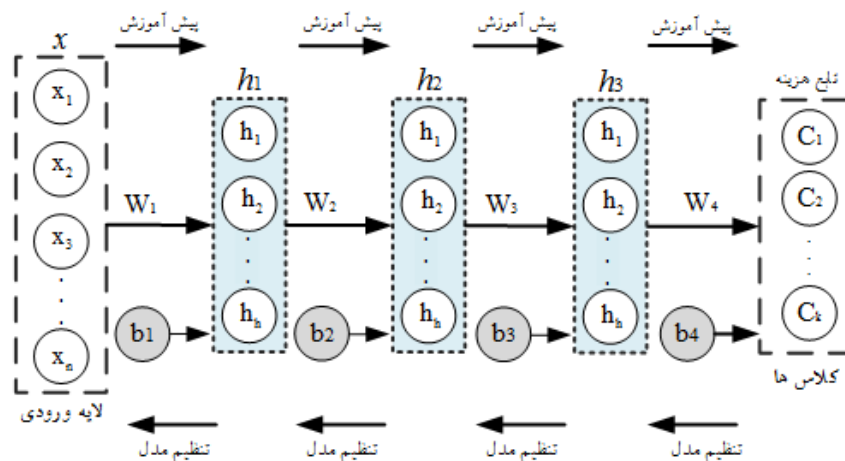
<sup>4</sup> Decoding Function

پارامترهای شبکه  $\theta = \{W_1, b_1, W_2, b_2\}$  از طریق الگوریتم نزول گرادینت<sup>۱</sup> و الگوریتم انتشار به عقب بهینه‌سازی می‌شوند تا میانگین خطای بازسازی به حداقل برسد.



شکل ۲-۸: نمودار کلی از رمزگذار خودکار

یک خودرمزگذار تکی می‌تواند ویژگی‌های سطح پایین را استخراج کند. پشته کردن چندین خودرمزگذار، خودرمزگذار پشته‌ای (SAE<sup>۲</sup>) نامیده می‌شود و می‌تواند با ویژگی‌های سطح عمیق‌تر و سطح بالاتر عملکرد بهتری داشته باشد. برای آموزش SAE از یک فرآیند دومارحله‌ای استفاده می‌شود. در مرحله اول، به نام پیش آموزش<sup>۳</sup>، هر مدل خودرمزگذار از طریق روش بدون نظارت آموزش داده می‌شود، در این حالت هر خودرمزگذار ورودی به خروجی را بازسازی می‌کند تا پارامترهای شبکه آموزش داده شوند. سپس، خروجی خودرمزگذار به آخرین خودرمزگذار تغذیه می‌شود. پس از انجام مرحله پیش آموزش، SAE با استفاده از پارامترهای آموزش دیده آغاز می‌شود. در مرحله دوم (Finetuning) پارامترهای شبکه بر اساس خطای دسته‌بندی بهینه می‌شوند. شکل ۲-۹ نمونه‌ای با سه خودرمزگذار پشته شده را نشان می‌دهد.



شکل ۲-۹: نمونه از شبکه SAE

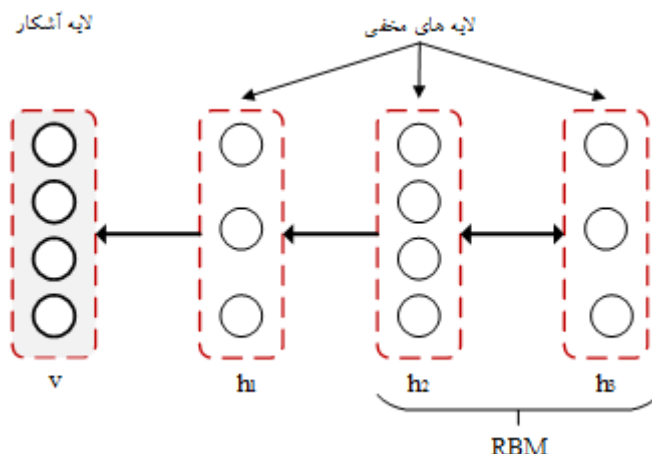
<sup>۱</sup> Gradient Descent

<sup>۲</sup> Stacked Auto-Encoder

<sup>۳</sup> Pre-training

## ۳-۲-۴-۲ شبکه باور عمیق

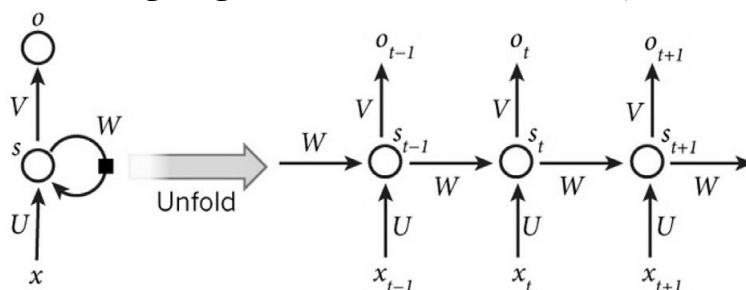
شبکه باور عمیق (DBN) شبیه SAE است، جایی که چندین ماشین Boltzmann محدود ( $RBM^1$ ) باهم پشته می‌شوند و DBN با استفاده از یک استراتژی دومرحله‌ای همانند SAE آموزش می‌بیند. پارامترهای اولیه در مرحله پیش‌آموزش به روش نظارت‌شده و حریصانه آموزش می‌بینند درحالی‌که از استراتژی نظارت‌شده مبتنی بر softmax برای تنظیم دقیق پارامترها استفاده می‌شود.  $RBM$  از دولا به تشکیل شده است: لایه قابل مشاهده و لایه پنهان، همان‌طور که در شکل ۲-۱۰ نشان داده شده است. از نمونه‌گیری Gibbs در  $RBM$  برای آموزش پارامترها استفاده می‌شود. در  $RBM$ ، مقدار هر واحد در هر دولا به پنهان و قابل مشاهده با استفاده از احتمال شرطی  $P(h|v)$  محاسبه می‌شود. در DBN، لایه قابل مشاهده  $RBM$  لایه پنهان برای  $RBM$  قبلی است.



شکل ۲-۱۰: معماری DBN که توسط  $RBM$  ساخته شده‌اند

## ۴-۲-۴-۲ شبکه عصبی بازگشتی

مدل‌های سنتی یادگیری عمیق همانند SAE، DBN و CNN، سری‌های زمانی را در نظر نمی‌گیرند، بنابراین برای یادگیری ویژگی‌های مربوط به داده‌های سری زمانی مناسب نیستند. یک جمله زبان طبیعی که نوعی داده سری زمانی است را به عنوان مثال در نظر بگیرید. از آنجاکه هر کلمه با کلمات دیگر در یک جمله ارتباط نزدیک دارد، هنگام استفاده از کلمه فعلی برای پیش‌بینی کلمه بعدی باید از یک یا چند کلمه قبلی استفاده شود. بدیهی است که مدل‌های یادگیری عمیق پیشرو نمی‌توانند برای این کار به خوبی کار کنند، زیرا اطلاعات ورودی‌های قبلی را ذخیره نمی‌کنند. شبکه عصبی بازگشتی (RNN) یک مدل یادگیری متوالی معمولی است. این شبکه ویژگی‌ها را برای ذخیره ورودی‌های قبلی که در حالت داخلی شبکه عصبی ذخیره شده‌اند، برای داده‌های سری می‌آموزد. همان‌طور که در شکل ۲-۱۱ ارائه شده است، یک چرخه مستقیم برای ایجاد ارتباطات بین سلول‌های عصبی معرفی شده است.



شکل ۲-۱۱: شبکه عصبی بازگشتی [۹]

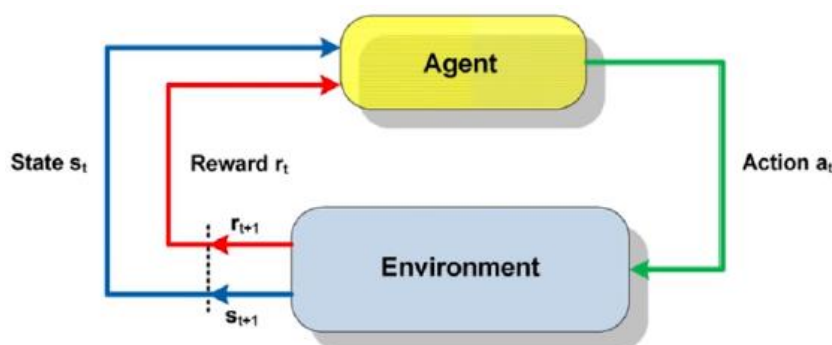
شبکه عصبی بازگشتی وابستگی بین نمونه فعلی  $x_t$  با نمونه قبلی  $x_{t-1}$  را با تلفیق بازنمایی پنهان قبلی  $s_{t-1}$  به دست می‌آورد. شبکه عصبی بازگشتی می‌تواند وابستگی‌هایی به طول دلخواه را به دست آورد. باین حال، به دلیل از بین رفتن

<sup>1</sup> Restricted Boltzmann Machines (RBMs)

شیب با استراتژی بازگشت به عقب برای آموزش پارامترها، برای یک RNN دشوار است که یک وابستگی طولانی مدت را به دست آورد. برای مقابله با این مشکل، برخی از مدل‌ها همانند حافظه کوتاه مدت طولانی ( $LSTM^1$ )، با جلوگیری از بین رفتن شیب یا انفجار شیب ارائه شده‌اند.

### ۲-۲-۱ یادگیری تقویتی

یادگیری تقویتی یک تکنیک مهم در یادگیری ماشین است که برای مدیریت فرآیند تصمیم‌گیری مارکوف استفاده شده است. در یادگیری تقویتی، یک عامل<sup>۲</sup> سیاست بهینه خود را از طریق تعامل با محیط خود می‌آموزد. یادگیری تقویتی بر روی به حداکثر رساندن پاداش فعالیت‌های عوامل نرم‌افزاری در یک محیط متمرکز است. ماهیت یادگیری تقویتی شامل یک عامل خودمختار همانند یک شخص، حیوان و یا ربات است، همان‌طور که شکل ۲-۱۲ نشان می‌دهد که با هدف به حداکثر رساندن پاداش عددی، در یک محیط نامشخص حرکت می‌کنند. ورزش نمونه خوبی از یادگیری تقویتی است. عامل باید با استراتژی و اقدامات مستمر که در یک رویداد ورزشی همانند مسابقه تنیس رخ می‌دهد، مقابله کند. عامل باید در مسابقه تنیس اعمالی مانند سرویس زنی، برگشت و والی را در نظر بگیرد. این اقدامات وضعیت بازی یا به عبارت بهتر مجموعه فعلی، بازیکن فعلی پیش رو و موارد مشابه را که بخشی از تنیس هستند تغییر می‌دهد. هر عملی برای دریافت پاداش انجام می‌شود، همانند کسب یک امتیاز که منجر به برنده شدن در بازی، ست و یا مسابقه می‌شود. عامل برای به حداکثر رساندن امتیاز نهایی ملزم به دنبال کردن خط‌مشی یا مجموعه‌ای از معیارها، قوانین و استراتژی‌ها است [۱۲].



شکل ۲-۱۲: یک فرایند کلی از یادگیری تقویتی [۶]

### ۲-۳ الگوریتم‌های تکاملی

رویکردهای محاسبات تکاملی از اصول تکامل طبیعی الهام گرفته شده است. یک رویکرد محاسبات تکاملی یک مسئله را به صورت راه‌حلی رمزگذاری می‌کند تا با هدف بهبود کیفیت راه‌حل‌ها تکامل یابد. به عنوان مثال، عملگرهای ژنتیکی، از جمله تقاطع، جهش و انتخاب، برای تولید نسل جدید استفاده می‌شود. بر اساس مکانیسم بقاء نسل تفاضل، فقط بهترین راه‌حل‌ها باقی می‌مانند. الگوریتم‌های محاسبات تکاملی با استفاده از یک روش اکتشافی و تکراری فضای جستجو را جستجو می‌کنند تا به تدریج راه‌حل‌های بهتری به دست آورند. دو نوع الگوریتم ابتکاری وجود دارد: جمعیت محور<sup>۳</sup> و مبتنی بر راه‌حل تکی<sup>۴</sup>. رویکرد جمعیت محور فرآیند تکامل را با استفاده از مجموعه‌ای از راه‌حل‌های تصادفی اولیه آغاز می‌کند. برخی از نمونه‌ها شامل الگوریتم ژنتیک ( $GA^5$ )، بهینه‌سازی کلونی مورچه‌ها ( $ACO^6$ ) و بهینه‌سازی جمعیت ذرات ( $PSO^7$ ) نمونه‌هایی از الگوریتم‌های جمعیت محور هستند.

<sup>1</sup> Long Short Term Memory (LSTM)

<sup>2</sup> Agent

<sup>3</sup> Population-based

<sup>4</sup> Single solution-based

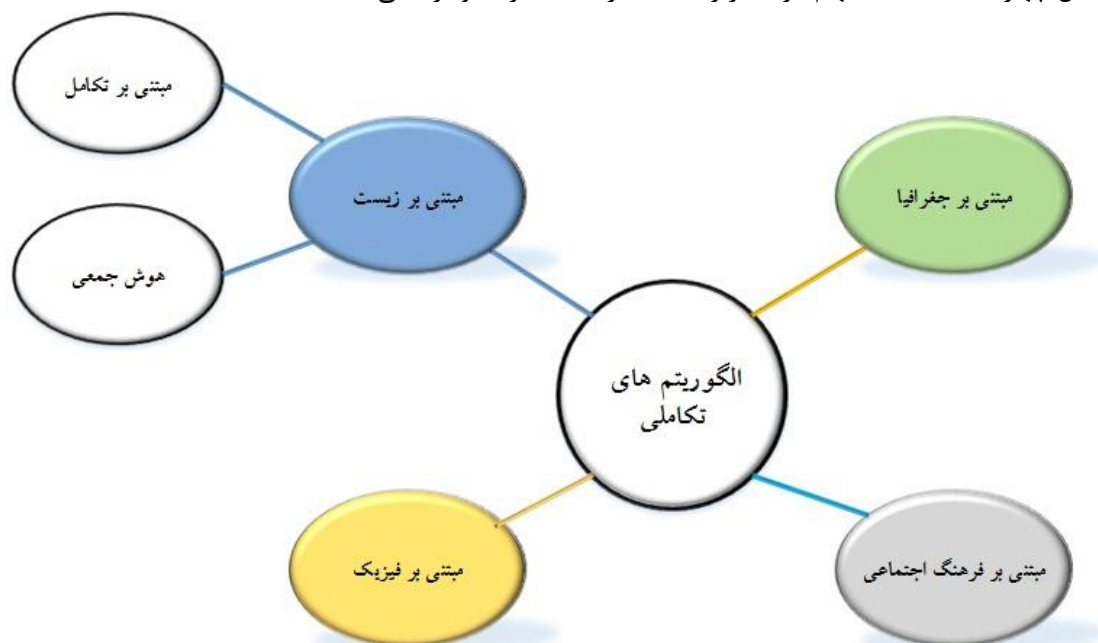
<sup>5</sup> Genetic Algorithm (GA)

<sup>6</sup> Ant Colony Optimization (ACO)

<sup>7</sup> Particle Swarm Optimization (PSO)



روش مبتنی بر راه حل تکی که بهینه سازی مسیر نیز نامیده می شود، از یک راه حل تصادفی اولیه آغاز می شود. جستجوی Tabu نمونه ای از این الگوریتم ها است. شکل ۲-۱۳ دسته بندی دیگری از الگوریتم های محاسبات تکاملی را نشان می دهد که شامل چهار دسته هستند: الهام گرفته از زیست<sup>۱</sup>، فیزیک<sup>۲</sup>، جغرافیا و فرهنگی.



شکل ۲-۱۳: طبقه بندی الگوریتم های تکاملی [۶]

(۱) **مبتنی بر زیست:** این دسته شامل رویکردهای مبتنی بر هوش تجمعی<sup>۳</sup> و الگوریتم هایی با الهام از تکامل است که از رفتار طبیعی موجودات نشئت می گیرند. هوش تجمعی در زندگی گروهی خود چگونه رفتار در یک کلونی (به عنوان مثال، پرندگان، ماهی ها و حشرات) را شبیه سازی می کند. هر یک از اعضای کلونی می توانند به طور مشترک بسیاری از کارهای پیچیده مورد نیاز برای بقای خود را انجام دهند. خودسازمانی و کنترل غیرمتمرکز دو ویژگی اصلی سیستم های مبتنی بر ازدحام هستند که به علت فعل و انفعالات موضعی بین عوامل ازدحام، منجر به رفتارهای آشکار می شوند. PSO و ACO دو الگوریتم اصلی از هوش تجمعی هستند. ریشه الگوریتم های مبتنی بر تکامل<sup>۴</sup> در اصول داروین مرتبط با تکامل طبیعی است که به موجودات زنده کمک می کند تا به خوبی با محیط خود سازگار شوند. خودسازمان دهی و سازگاری قوی دو ویژگی اصلی این رویکردها است. در این الگوریتم ها می توان کل جمعیت ها را از یک نسل به نسل دیگر با عملگرهایی مانند تقاطع و جهش جایگزین کرد. الگوریتم ژنتیک، برنامه نویسی تکاملی<sup>۵</sup> و برنامه نویسی ژنتیکی (GP)<sup>۶</sup> سه نوع اصلی از مکانیسم های الهام گرفته از تکامل هستند.

(۲) **الهام گرفته از فیزیک:** منشأ الگوریتم های الهام گرفته از فیزیک از قوانین فیزیکی / شیمیایی ناشی می شود. الگوریتم جستجوی گرانشی<sup>۷</sup> از این دسته است.

(۳) **الهام گرفته از جغرافیا:** این الگوریتم ها راه حل های تصادفی در فضای جستجوی جغرافیایی ایجاد می کنند. Tabu Search در این گروه قرار می گیرد.

<sup>1</sup> Bio-inspired

<sup>2</sup> Physics-inspired

<sup>3</sup> Swarm intelligence

<sup>4</sup> Evolution-inspired

<sup>5</sup> Evolutionary programming

<sup>6</sup> Genetic Programming (GP)

<sup>7</sup> Gravitational Search Algorithm



(۴) الهام گرفته از فرهنگ: این الگوریتم‌ها از رفتار انسانی که در جریان تعاملات فرهنگی با دیگران دیده می‌شود الهام گرفته شده‌اند. مشاهده رفتارهای طبیعی و ذاتی سایر افراد به اشخاص کمک می‌کند تا دانش جدیدی بیاموزند و رفتار خود را بهبود بخشند. الگوریتم Memetic را می‌توان یکی از این رویکردها دانست که از طریق یک روش ابتکاری محلی از روند جهش تقلید می‌کند.

### ۲-۳-۱ الگوریتم ژنتیک

الگوریتم ژنتیک [۱۱] یک تکنیک محاسبات تکاملی محبوب با استفاده گسترده است. یک الگوریتم ژنتیک پایه‌ای شامل پنج مرحله است: شروع، ارزیابی، تولیدمثل، تقاطع و جهش. نمایش متداول یک کروموزوم یک رشته باینری با طول ثابت است. جمعیت اولیه به‌طور تصادفی در آغاز الگوریتم GA تولید می‌شود. هر کروموزوم با استفاده از یک تابع تناسب ارزیابی می‌شود تا با استفاده از مکانیسم انتخاب، راه‌حل‌ها را برای تولید فرزندان انتخاب کند. عملگرهای تقاطع و جهش، جمعیت را به یک جمعیت جدید با مقادیر تناسب بالاتر تبدیل می‌کنند [۱۲].

### ۲-۳-۲ تکامل تفاضلی

تکامل تفاضلی (DE) [۱۳] یک الگوریتم تکاملی قوی و قابل اعتماد است. عملگرهای ژنتیکی همانند تقاطع، جهش و انتخاب در DE نیز استفاده می‌شود. برخلاف الگوریتم‌های ژنتیکی که به عملکرد تقاطع متکی هستند، DE بر روی عملگر جهش متمرکز است. در DE، تفاوت بین کروموزوم‌ها برای مکانیسم جستجو استفاده می‌شود. عملیات جهش، تقاطع و انتخاب اعمال شده در DE به ترتیب در رابطه‌های ۲-۴، ۲-۵ و ۲-۶ ارائه شده است.

$$v_i^G = x_{r1}^G + F \cdot (x_{r2}^G - x_{r3}^G) \quad (2-4)$$

$$u_{i,j}^G = \begin{cases} x_{i,j}^G & \text{if } rand_{i,j}[0,1] \leq CR \text{ or } i = i_{rand} \\ v_{i,j}^G & \text{otherwise} \end{cases} \quad (2-5)$$

$$u_i^{G+1} = \begin{cases} u_i^G & \text{if } f(u_i^G) < f(x_i^G) \\ x_i^G & \text{otherwise} \end{cases} \quad (2-6)$$

در این رابطه‌ها،  $G$  حداکثر تعداد نسل را نشان می‌دهد و  $F$  پارامتری است که طول اکتشاف را تعیین می‌کند ( $x_{r2} - x_{r3}$ ). دو نوع عملیات تقاطع در DE استفاده می‌شود: دودویی و نمایی. رابطه ۲-۶ تابع متقاطع دودویی را نشان می‌دهد، جایی که تابع  $rand_{i,j}[0,1]$  مقداری بین ۰ تا ۱ را به‌طور تصادفی ایجاد می‌کند و  $CR$  یک پارامتر کنترل به‌عنوان نرخ تقاطع است. هنگامی که یک کروموزوم جدید توسط عملگر جهش (رابطه ۲-۵) تولید می‌شود، مقدار تناسب آن با والدین مقایسه می‌شود تا کروموزومی که برای نسل بعدی زنده خواهد ماند تعیین شود.

### ۲-۳-۳ الگوریتم بهینه‌سازی کلونی مورچگان

بهینه‌سازی کلونی مورچگان (ACO) برای حل مشکلات محاسباتی از رفتار جستجوی مورچه‌های واقعی الهام گرفته شده است. مورچه‌ها به‌طور تصادفی شروع به کشف منطقه اطراف لانه می‌کنند [۱۴]. مورچه‌ها پس از یافتن منابع غذایی، کمیت و کیفیت غذا را ارزیابی کرده و مقداری از غذا را به لانه حمل می‌کنند. مورچه‌ها با سفر بر روی نمودار راه‌حلی را می‌سازند. در طول سفر بازگشت، آن‌ها فرمون‌های<sup>۱</sup> شیمیایی را بر اساس کمیت و کیفیت غذا در مسیر خود برای سایر مورچه‌ها می‌گذارند. این فرایند به مورچه‌ها کمک می‌کند تا کوتاه‌ترین مسیر بین لانه و منبع غذایی مناسبی را که در

<sup>۱</sup> Pheromone

قبل توسط مورچه‌های دیگر کشف شده بود را پیدا کنند. سیستم کلونی مورچگان (ACS<sup>۱</sup>) [۱۵] محبوب‌ترین نوع ACO است. در ACS، قوانین وضعیت و به‌روزرسانی فرمون برای بهبود اثربخشی ACO اصلاح شده است. جزئیات ACS با استفاده از یک مسئله فروشنده دوره‌گرد توصیف می‌شود. اگر مورچه  $k$  در حال حاضر در شهر  $j$  است، شهر بعدی مورچه برای بازدید با استفاده از رابطه ۷-۲ تعیین می‌شود.

$$s = \begin{cases} \arg n \in R_k(j) \max\{\tau(j, n)^\alpha \times [\eta(j, n)]^\beta\} & \text{if } q \leq q_0 \\ i \text{ with a probability } P_k(j, i) & \text{if } q > q_0 \end{cases} \quad (۷-۲)$$

در این رابطه،  $R_k(j)$  نشان‌دهنده شهرهایی است که توسط مورچه  $k$  ملاقات نشده‌اند،  $\tau(j, n)$  فرمون تجمع شده بین شهرهای  $j$  و  $n$  است. همچنین  $\alpha$ ،  $\beta$  تأثیر نسبی فرمون در مقابل فاصله بین دو شهر را نشان می‌دهند. علاوه بر این،  $q$  یک عدد تصادفی بین ۰ و ۱ است و  $P_k(j, i)$  احتمال از شهر  $j$  به شهر  $i$  است که با استفاده از رابطه ۸-۲ محاسبه می‌شود:

$$P_k(j, i) = \begin{cases} \frac{[\tau(j, n)]^\alpha \times [\eta(j, n)]^\beta}{\sum_{n \in R_k(j)} [\tau(j, n)]^\alpha \times [\eta(j, n)]^\beta}, & \text{if } i \in R_k(j) \\ 0, & \text{otherwise} \end{cases} \quad (۸-۲)$$

هنگامی که همه مورچه‌ها مأموریت‌های خود را به پایان رساندند، تراکم فرمونی بهترین مأموریت اصلاح می‌شود. بهترین راه‌حل تکراری و بهترین راه‌حل عمومی دو نوع از به‌روزرسانی هستند. اولی بهترین قاعده را از بین تمام تکرارهای اجرا شده انتخاب می‌کند، درحالی که دومی تکرار فعلی را انتخاب می‌کند. قانون به‌روزرسانی عمومی توسط رابطه‌های ۹-۲ و ۱۰-۲ نشان داده شده است.

$$\tau_{t+1}(j, s) = (1 - \alpha) \times \tau_t(j, s) + \alpha \times \Delta\tau(j, s) \quad (۹-۲)$$

$$\Delta\tau(j, s) = \begin{cases} \frac{1}{L_{best}}, & \text{if } (j, s) \in \text{best tour} \\ 0, & \text{otherwise} \end{cases} \quad (۱۰-۲)$$

### ۳-۳-۲ الگوریتم بهینه‌سازی ازدحام ذرات

بهینه‌سازی ازدحام ذرات (PSO) [۱۶] یک الگوریتم بهینه‌سازی جمعیتی برای مسائل عددی غیرخطی است که از رفتار حیوانات اجتماعی همانند پرندگان و ماهی الهام گرفته شده است. کارآمدی، سادگی و سرعت همگرایی سریع از مزایای اصلی PSO است. هر ذره نشان‌دهنده راه‌حل‌های مسئله است و دارای سرعتی است که جهت پرواز به سمت سایر راه‌حل‌ها را نشان می‌دهد. در PSO چهار مفهوم وجود دارد:

$$(1) X_i(t) = (X_{i1}, X_{i2}, \dots, X_{iD}) \text{ subject to: } X_{i,n}(t) \in [l_n, u_n], 1 \leq n \leq N$$

$$(2) pbest_i(t) = (p_{i1}, p_{i2}, \dots, p_{iD})$$

$$(3) gbest_i(t) = (g_{i1}, g_{i2}, \dots, g_{iD})$$

$$(4) v_i(t) = (v_{i1}, v_{i2}, \dots, v_{iD})$$

<sup>۱</sup> Ant Colony System (ACS)

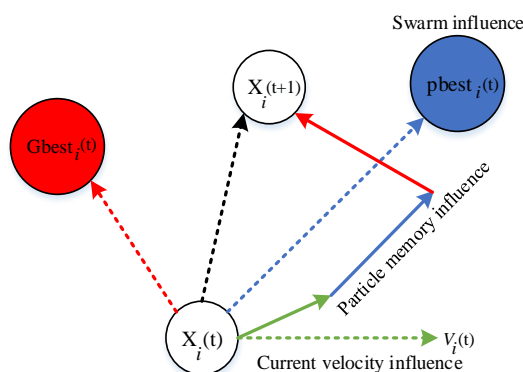
اصطلاح  $X_i(t)$  موقعیت ذره  $i$  در تکرار  $t$  است،  $pbest_i(t)$  و  $gbest_i(t)$  به ترتیب بهترین موقعیت‌های محلی و عمومی هستند و  $v_i(t)$  سرعت در تکرار  $i$  است. در طی هر تکرار، هر ذره بر اساس  $pbest$  و  $gbest$  به‌روز می‌شود. در PSO برای ارزیابی کیفیت ذره از موقعیت ذره  $i$  در تکرار  $t$  استفاده می‌شود. معادلات (۲-۱۱) و (۲-۱۲) نحوه محاسبه سرعت و موقعیت ذره را در تکرار بعدی نشان می‌دهد:

$$v_{i,j}(t+1) = w * v_{i,j}(t) + 2 * rand(0,1) * (pbest_{i,j}(t) - X_{i,j}(t)) + \quad (2-11)$$

$$2 * rand(0,1) * (gbest_{i,j}(t) - X_{i,j}(t))$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (2-12)$$

در این رابطه‌ها،  $w$  وزن سکون<sup>۱</sup> بین ۰.۸ و ۱.۲ است. در PSO پایه، روی ۱ تنظیم شده بود. مقدار  $w$  کوچک‌تر باعث اکتشاف محلی می‌شود، درحالی‌که  $w$  بزرگ‌تر، منجر به اکتشاف عمومی می‌شود. شکل ۲-۱۴ حرکت ذره را در دو بعد با توجه به معادلات (۲-۱۱) و (۲-۱۲) نشان می‌دهد.



شکل ۲-۱۴: انتقال یک ذره [۱۷]

#### ۲-۳-۴ الگوریتم کلونی زنبور عسل

رفتارهای مختلف زنبورهای عسل همانند جفت‌گیری، پرورش و جستجوی غذا توسط چندین الگوریتم بهینه‌سازی مبتنی بر زنبور عسل تقلید شده است. بیشتر الگوریتم‌های الهام گرفته از زنبور عسل از رفتار تغذیه زنبورهای عسل تقلید می‌کنند. در این الگوریتم‌ها، از یک جستجوی اکتشافی تصادفی برای یافتن مکان‌های امیدوارکننده استفاده می‌شود. سپس این الگوریتم‌ها برای دستیابی به یک راه‌حل بهینه، جستجوی استثماری را در مکان‌ها اعمال می‌کنند. الگوریتم زنبور مصنوعی ( $ABC^2$ ) یک الگوریتم معروف است که از رفتار تغذیه زنبورهای عسل برای کشف غنی‌ترین و در دسترس‌ترین منابع غذایی استفاده می‌کند. این الگوریتم شامل سه نوع زنبور عسل با الگوهای مختلف پرواز است: زنبورهای کارگر، ناظر و زنبورهای پیشاهنگی. هر زنبور عسل با موقعیتی همراه است که نشان‌دهنده یک راه‌حل عملی برای یک مسئله بهینه است. در هر تکرار، زنبورهای دارای بهترین و بدترین تناسب به ترتیب به‌عنوان زنبورهای کارگر جستجوگر و پیشاهنگی انتخاب می‌شوند. زنبورهای ناظر اطلاعات شهد ارائه‌شده توسط زنبورهای کارگر را ارزیابی می‌کنند تا مسیر حرکت آن را در دفعات بعدی تنظیم کنند. ناهمگنی در الگوهای پرواز زنبورها منجر به تعادل بین بهره‌برداری و اکتشاف می‌شود. زنبورهای پیشاهنگی برای کنترل تنوع درحالی‌که زنبورهای باتجربه جستجوگر به بهره‌برداری از فضای جستجو کمک می‌کنند [۱۸].

<sup>1</sup> Inertia

<sup>2</sup> Artificial Bee Colony (ABC)

هر کدام از زنبورهای کارگر و ناظر در الگوریتم ABC نیمی از کندو را اشغال می کنند. برای هر منبع غذایی یک زنبور عسل کارگر اختصاص داده می شود. وقتی زنبور عسل کارگر منبع غذایی خود را رها کند، تبدیل به یک زنبور عسل پیشاهنگی می شود. مراحل اصلی ABC به شرح زیر است. جمعیت اولیه به صورت تصادفی به تعداد SN منبع غذایی با استفاده از رابطه ۲-۱۳ تولید می شود، جایی که SN توسط کاربر تعریف می شود و نشان دهنده تعداد منابع غذایی است.

$$X_{i,j} = X_{min,j} + rand(0, 1) (X_{max,j} - X_{min,j}) \quad (2-13)$$

هر منبع غذایی  $X_i$  یک بردار به اندازه  $D$  که تعداد پارامترهای مسئله بهینه سازی را نشان می دهد. هر زنبور عسل کارگر ابتدا منبع غذایی را با استفاده از رابطه ۲-۱۴ اصلاح می کند و سپس مقدار شهد منبع غذایی جدید را محاسبه می کند. اگر مقدار تناسب راه حل جدید بیشتر از راه حل قبلی نباشد، فرآیند با راه حل قدیمی ادامه می یابد. در غیر این صورت، زنبور عسل کارگر به سمت راه حل جدید حرکت می کند. در واقع، یک مکانیسم انتخاب حریصانه برای انتخاب موقعیت جدید استفاده شده است.

$$V_{i,j} = X_{i,j} + \theta_{i,j} (X_{i,j} - X_{i,k}) \quad (2-14)$$

زنبورهای ناظر بر اساس اطلاعات منبع غذایی یافت شده توسط زنبورهای کارگر در مورد انتخاب منبع غذایی تصمیم می گیرند. یک زنبور عسل با توجه به مقدار شهد با استفاده از یک احتمال  $(P_i)$ ، منبع غذایی را انتخاب می کند (رابطه ۲-۱۵).

$$P_i = \frac{f_i}{\sum_{n=1}^{SN} f_n} \quad (2-15)$$

#### ۲-۴ پیشینه تحقیق

سرمایه گذاران در بازارهای مالی می خواهند بازده را بر اساس ریسک قابل قبول به حداکثر برسانند. یک سیستم خبره مبتنی بر قانون که نگرش سرمایه گذار به ریسک و بازده را در نظر می گیرد، اغلب می تواند بازدهی خوبی داشته باشد.

#### جدول ۲-۱: تحلیلی بر روش های مطرح شده برای ریسک سرمایه گذاری

ردیف	منبع / سال انتشار	الگوریتم	مجموعه داده
۱	Hu and Ansell [۱۹]/(۲۰۰۷)	بیزین ساده شبکه عصبی رگرسیون	مطالعه موردی از ۱۹۵ مشتری در آمریکا
۲	Ince and Aktan [۲۰]/(۲۰۰۹)	CART رگرسیون شبکه عصبی	مطالعه موردی
۳	Finlay et al. [۲۱]/(۲۰۱۱)	ANN	داده مصنوعی
۴	Karakaya et al. [۲۲]/(۲۰۱۳)	رگرسیون	داده های وام
۵	Baklouti [۲۳]/(۲۰۱۴)	CART	تونس
۶	Cao [۲۴]/(۲۰۱۸)	رگرسیون	مطالعه موردی از ۱۰۰۰ مشتری در لهستان
۷	Lv [۲۵]/(۲۰۱۸)	CVaR	چین
۸	Huang et al. [۲۶]/(۲۰۱۸)	PNN	چین
۹	Tang [۳۸]/(۲۰۱۸)	یادگیری تقویتی	داده های سهام شامل ۵۰۰۰۰ نمونه
۱۰	Tang [۳۹]/(۲۰۱۸)	یادگیری تقویتی	داده های سهام چین

سوئد	SVM	Zhu et al. [۲۷]/(۲۰۱۹)	۱۱
مطالعه موردی از ۱۱۰ مشتری	FNN	Zhang [۲۸]/(۲۰۲۰)	۱۲
آمریکا	NLP	Antoncic [۲۹]/(۲۰۲۰)	۱۳
ارمنستان	MLP	Baghdasaryan et al. [۳۰]/(۲۰۲۱)	۱۴

در حوزه بانکی، هوش مصنوعی به یک راه حل قوی برای بانک‌ها تبدیل شده است. همچنین در سرمایه‌گذاری بانکی و خدمات مرتبط همانند پیش‌بینی سهام و رتبه‌بندی اعتباری بدون تماس با مشتری اعمال می‌شود. حضور هوش مصنوعی در بانکداری تجاری در عملیات اصلی که اغلب بر وام، پردازش پرداخت‌ها و مدیریت سپرده‌ها متمرکز است، هنوز محدود است. هوش مصنوعی می‌تواند با کاهش خسارات وام، افزایش امنیت در پردازش پرداخت‌ها، خودکار کردن کارهای مربوط به انطباق و بهبود هدف‌گیری مشتری، مدیریت سپرده‌ها و خدمات مربوطه مفید باشد.

هوش مصنوعی می‌تواند میزان استفاده از نیروی کار در بانک‌ها را تغییر دهد [۳۱]. هوش مصنوعی مورد استفاده همراه با تجزیه و تحلیل داده‌ها می‌تواند با کارآمدتر کردن عملکردهای معمول همانند گزارش انطباق، منجر به صرفه‌جویی در هزینه کار شود که ۵۷ درصد هزینه‌های بانکی را نشان می‌دهد. هوش مصنوعی می‌تواند از طریق ردیابی، ریسک را کاهش داده و افراد را قادر می‌سازد تا روی موارد ضروری تمرکز کنند. علاوه بر این، با افزایش تعداد مقررات جدید، زمان صرف شده به دنبال همه هشدارهایی که سیستم ایجاد می‌کند باید کاهش یابد زیرا این هشدارها اغلب مثبت کاذب هستند و مدیران باید به‌طور مداوم سیستم را به‌روز کنند تا تعداد مطابقت‌ها متناسب شود.

بانک‌ها از روش‌های آماری برای درک علت بازپرداخت شدن برخی از وام‌ها و عدم پرداخت برخی از وام‌های دیگر استفاده می‌کنند. به عنوان مثال، یک مطالعه از ثبت اعتبار در ایتالیا نشان داد که وام‌های تضمینی تمایل بیشتری به عدم بازپرداخت دارند [۳۲]. بنابراین، بانک‌ها باید مدل‌هایی را ایجاد کنند که به آن‌ها اجازه دهد درک کنند که آیا هنوز هم دادن وام با ضمانت بهتر است یا خیر زیرا زیان‌ها به لطف ضمانت‌نامه‌ها قابل بازیابی است.

در [۲۱] از الگوریتم انتشار عقب<sup>۱</sup> برای ساخت مدل ارزیابی اعتبار شبکه عصبی استفاده شد که از مجموعه داده‌های وام بنگاه‌های کوچک و متوسط برای تجزیه و تحلیل استفاده شد. نتایج مدل و مدل رگرسیون لجستیک نشان داد که مدل شبکه عصبی می‌تواند به دقت پیش‌بینی بهتری در ارزیابی ریسک اعتبار دست یابد. ناراکایا و همکاران یک مدل شبکه عصبی چند لایه را با مدل ماشین بردار پشتیبانی (SVM<sup>۲</sup>) مقایسه کردند. داده‌های وام یک بانک تجاری به عنوان نمونه‌ای برای تحقیق در نظر گرفته شد. نتایج نشان داد که دقت پیش‌بینی مدل شبکه عصبی چندلایه بهتر از مدل SVM است که می‌تواند نتایج پیش‌بینی بهتری را به دست آورد [۲۲].

در مطالعه دیگر مدل رگرسیون خطی برای مطالعه ریسک اعتباری وام‌گیرنده استفاده شد [۲۷]. در صورت عدم وجود داده‌های کمی بانکی، ارزیابی ریسک پیش‌فرض اعتبار وام‌گیرنده از طریق داده‌های قبلی مربوط به اعتبار قابل اعتمادتر است، که همچنین می‌تواند درآمد سرمایه‌گذاران را به‌طور قابل توجهی بهبود بخشد.

کائو<sup>۳</sup> از نسبت‌های مالی و شاخص‌های مالی برای ارزیابی ریسک مالی شرکت‌ها استفاده کرد و یک مدل احتمال ارزیابی ریسک مالی لجستیک ارائه داد و سپس شرکت‌های نمونه را با استفاده از مدل محاسبه کرد تا وضعیت ریسک مالی فعلی کوچک و متوسط را ارزیابی کند [۲۴]. همبستگی و ابعاد بالای داده‌های اعتباری شرکت‌ها بر روند و نتایج تجزیه و تحلیل در فرایند تجزیه و تحلیل رگرسیون لجستیک برای پیش‌بینی و مطالعه ریسک پیش‌فرض شرکت تأثیر می‌گذارد. نویسندگان تا حد امکان اطلاعات ارائه شده توسط همه شاخص‌ها به منظور بهبود دقت شناسایی و پیش‌بینی وضعیت

<sup>۱</sup> Back propagation<sup>۲</sup> Support Vector Machine (SVM)<sup>۳</sup> Cao

عملکرد شرکت در نظر گرفتند. در پژوهش موردنظر، تجزیه و تحلیل مؤلفه‌های اصلی ( $PCA^1$ ) به منظور کاهش ابعاد و به حداقل رساندن اتلاف اطلاعات موجود در داده‌های اصلی استفاده شد. در روش پیشنهادی، شش مؤلفه اصلی از ۱۴ نسبت مالی انتخاب و نرخ مشارکت تجمعی به حدود ۸۵٪ رسید. در همین حال، با توجه به ارتباط بین اجزای اصلی و داده‌های اصلی، پیامدهای اقتصادی متفاوتی به آن‌ها داده شد. تجزیه و تحلیل مدل رگرسیون لجستیک برای ارزیابی ریسک اعتباری بر اساس داده‌های نسبت‌های مالی شرکت‌های فهرست شده در چین نشان داد برای سرمایه‌گذاران تعیین ریسک سرمایه‌گذاری و کاهش ریسک مناسب است.

به منظور کاهش خطر آشفتگی بازار مالی، روش VaR به عنوان ابزاری برای اندازه‌گیری ریسک ارائه شد و تحقیقات تأیید کرد که روش VaR می‌تواند دقت پیش‌بینی بازار مالی در چین را که از اهمیت بسیاری برخوردار است، بهبود بخشد [۲۵]. شش روش مختلف VaR مورد آزمایش قرار گرفت. برای اعتبارسنجی نتایج آماری، از آزمون کوپیک برای ارزیابی نسبت تخلفات و از آزمون کریستوفرسن برای بررسی سرعت تعدیل مدل در برابر نوسانات بازار استفاده شد. دو تحلیل انجام شد؛ تحلیل اول یک بازه زمانی ۱۰۰۰ روزه و دومی ۲۵۲ روزه را در نظر گرفتند. برای هر دو تحلیل روش پیشنهادی دارای بالاترین عملکرد بود. مدل‌ها دارای پیش‌بینی دقیق برای شاخص‌های اوراق قرضه خصوصی و کالاها بودند که نشان‌دهنده روش‌های متمرکز بر ریسک بازار برای این دسته از دارایی‌ها مناسب نیست. نتایج همچنین نشان داد که یک بازه زمانی کوچک‌تر برای برآورد خسارت برای دارایی‌های با نوسان بالا مناسب‌تر است.

با هدف حل مسئله نحسی ابعاد بالا، در [۳۸] از برنامه‌نویسی پویای تقریبی برای ایجاد یک مدل تصمیم‌گیری مارکوف استفاده شد. یک الگوریتم بازیگر-منتقد<sup>۲</sup> مبتنی بر مدل در محیط نامشخص پیشنهاد شد، که در آن تابع ارزش مطلوب بر اساس محدوده ریسک و تعداد محدودی از بودجه به دست می‌آید و سرمایه‌گذاری بهینه هر دوره با استفاده از برنامه‌ریزی پویا از تعداد محدودی تعیین می‌شود. مسئله موردنظر به صورت ماتریس ریسک بر اساس داده‌های قبلی فرض می‌شود و از قانون ریسک درآمد سهام در جهت مثبت پیروی می‌کند. از سه اصل برای راهنمایی سرمایه‌گذار در انتخاب سهام استفاده شد. اول، ماتریس درآمد خالص دارای سه عنصر مهم است: نوع سهام، درآمد خالص و مقدار پول سرمایه‌گذاری شده. دوم، انتخاب بین سهام مهم نیست و این بر عهده سرمایه‌گذار است که نوع و تعداد سهام را تعیین کند. سوم، کل سود شامل سود بانکی شامل نرخ سود از قبل تعیین‌شده، هزینه‌های معامله و سود است. علاوه بر صندوق‌های سرمایه‌گذاری و سود از پیش تعیین‌شده، هزینه‌های معامله نیز مشخص است، درحالی‌که سایر موارد نامشخص است.

در [۳۹]، سه الگوریتم یادگیری تقویت مداوم پیشرفته برای مدیریت ریسک سرمایه‌گذاری توسعه داده شد. عملکرد الگوریتم‌ها با تنظیمات مختلف، از جمله نرخ‌های یادگیری مختلف، توابع هدف، و ترکیب ویژگی‌ها به منظور ارائه بینشی برای تنظیم پارامترها، انتخاب ویژگی‌ها و آماده‌سازی داده‌ها ارائه شد. همچنین آزمایش‌ها بر روی بازار سهام چین انجام شد. علاوه بر این، یک روش آموزشی به اصطلاح آموزش خصمانه<sup>۳</sup> پیشنهاد شد و نشان داده شد که می‌تواند کارایی آموزش را تا حد زیادی بهبود بخشد و به طور قابل توجهی میانه‌گین بازدهی روزانه را ارتقا دهد. ابتدا، تفاوت مرتبه دوم<sup>۴</sup> برای پارامترهای شبکه عصبی استراتژی خروجی و انتظار در Q-value برای همگرایی الگوریتم ضروری است. با توجه به الگوریتم، فقط سیاست بهینه در مجموعه تابع مرتبه دوم، به جای مجموعه تابع سیاست جستجو می‌شود. سپس، الگوریتم نیاز به انتقال ثابت دارد. با این حال، به دلیل بی‌نظمی در بازار و مداخله دولت، تغییرات دولت در بازار سهام ممکن است در زمان متغیر باشد.

ژانگ و همکاران از مدل CoVaR برای اندازه‌گیری نفوذ ریسک سیستماتیک استفاده کردند و مدل جدیدی برای تخصیص دارایی بر اساس Mean-CoVaR ایجاد کردند تا بتواند این مشکل را حل کند [۳۳]. این مدل می‌تواند مقدار

<sup>1</sup> Principal Component Analysis (PCA)

<sup>2</sup> Actor-Critic

<sup>3</sup> Adversarial Training

<sup>4</sup> Second-Order

خسارت سبد دارایی در هنگام مواجهه با ریسک‌های مالی، به‌ویژه ریسک‌های شدید، را تخمین بزند. نتایج نشان داد که مدل Mean-CoVaR در هنگام بررسی تأثیرات ریسک سیستماتیک برای تخصیص نمونه کارها کارآمدتر است. همچنین به طور تجربی عملکرد مجموعه‌های ارائه شده با ۱۰ استراتژی در پنج مجموعه داده مقایسه شدند. نتایج بررسی‌ها نشان داد که سبدهای اوراق بهادار محدود اغلب نسبت ریسک بیشتری نسبت به استراتژی‌های سبد دارایی دارند.

یک تجزیه و تحلیل بر روی تأثیر اطلاعات مالی بر روی ریسک شرکت‌های تولیدکننده سهام در بورس سهام اندونزی طی یک دوره پنج‌ساله از ژانویه ۲۰۱۱ تا دسامبر ۲۰۱۵ در [۹] ارائه شد. اطلاعات مالی با استفاده از چهار متغیر حسابداری اندازه‌گیری شد، یعنی قدرت مالی، نقدینگی، سودآوری و اندازه شرکت. نتایج تجزیه و تحلیل داده‌ها با استفاده از روش رگرسیون خطی چندگانه اثبات کرد که در سطح ۰.۰۵ تنها شرکت‌ها با اندازه متغیر هستند که به‌طور قابل توجهی بر روی ریسک تأثیر می‌گذارند. در همین حال، قدرت مالی، نقدینگی و سودآوری بر روی ریسک تأثیر نمی‌گذارد. این نتایج با نتایج چندین مطالعه قبلی مغایرت داشت. این تناقض ممکن است به دلیل مشکلات اندازه‌گیری حسابداری متغیر، دوره اجرای مطالعه و استفاده از نمونه‌های تحقیق مختلف باشد.

کوالنکو<sup>۱</sup> و همکاران [۱۴] مسئله بدهی اعتباری معوق و ایجاد روش‌های مؤثر برای مدیریت بدهی‌های مسئله‌دار در بانک‌ها و مؤسسات مالی را مورد بررسی قرار دادند. بر اساس ترکیبی از ابزارهای تئوری منطق فازی و شبکه‌های عصبی مصنوعی، یک مدل ریاضی برای امتیازدهی ساخته شد. برای تنظیم پارامترهای توابع عضویت در فرآیند فازی سازی متغیرهای کمی از نگاشت‌های خودتنظیمی کوهن<sup>۲</sup> استفاده شد. داده‌ها از وبسایت‌های رسمی چهار بانک بلغارستان برای ۲۰۱۵-۲۰۱۹ گرفته شده است. حجم نمونه تهیه‌شده شامل ۱۰۰۰ قرارداد اعتباری با پرداخت‌های معوقه فعال بود.

یک مطالعه نشان می‌دهد که شبکه‌های عصبی تکاملی با مدل پیش‌بینی خروجی خطی می‌تواند با در نظر گرفتن ریسک سرمایه‌گذاری برای تجارت مورد استفاده قرار گیرد [۳۴]. نمایه‌های مختلف سرمایه‌گذاری می‌تواند ایجاد شود و نمونه کارها برای هر نمایه ریسک قابل تهیه باشد. هدف ارائه مدل پیش‌بینی شبکه عصبی Evolino، هماهنگی و متناسب‌سازی با تجارت سودآور در بازار ارز با در نظر گرفتن ریسک سرمایه‌گذاری است. روش پیشنهادی بیشتر بر مدل و قابلیت اطمینان و دقت تمرکز می‌کند. همچنین هدف آن نشان دادن دامنه وسیع فرصت‌های سرمایه‌گذاری است. مدل توسعه‌یافته بر اساس شبکه عصبی مکرر Evolino و بر اساس روش‌های متخصص، استفاده از آن ساده است و ابزار مناسبی برای سرمایه‌گذاران است. قابلیت اطمینان این مدل که با ضریب همبستگی اندازه‌گیری می‌شود، برای تجارت سودآور در بازار مالی به‌اندازه کافی بالا است. اعتبار پیش‌بینی مدل سودآوری سرمایه‌گذاری را افزایش می‌دهد. این مدل به سرمایه‌گذار اجازه می‌دهد تا مجموعه‌های سرمایه‌گذاری متفاوتی را بر اساس انتخاب استراتژی‌های مختلف سرمایه‌گذاری با سطوح مختلف ریسک ایجاد کند.

یک مطالعه به بررسی چگونگی استفاده از هوش مصنوعی برای کنترل تصمیمات مدیریت پول مبتنی بر ریسک پرداخته است [۳۵]. یک روش منطقی فازی برای شناسایی و طبقه‌بندی قوانین فنی پیشنهاد شد. این مدل به‌صورت پویا مناطق با عملکرد بالاتر را اولویت‌بندی می‌کند و سیاست‌های مدیریت پول را برای به حداکثر رساندن عملکرد تنظیم‌شده توسط ریسک، منطبق می‌کند. این مدل عملکرد تقریبی تعدیل‌شده توسط ریسک را محاسبه می‌کند و به‌طور خودکار تخصیص سرمایه را با توجه به سناریوهای ترجیحی روند و نوسانات طی یک روز معاملات متعادل می‌کند. این مطالعه نشان داد که یک روش ترکیبی با استفاده از قوانین فازی به همراه یک مدل پیش‌بینی شبکه‌های عصبی بهتر از شبکه‌های عصبی استاندارد است.

<sup>1</sup> Kovalenko

<sup>2</sup> Kohonen self-organizing maps

همچنین مطالعه دیگری نشان داد که شبکه‌های عصبی در ترکیب با منطق فازی بهتر از شبکه‌های عصبی پایه‌ای هستند. تفاوت بین این دو نوع شبکه‌های عصبی این است که کسب دانش از طریق منطق فازی است اما در مورد دوم از طریق الگوریتم‌ها تحقق می‌یابد. قوانین استدلال فازی قابل مشاهده است، درحالی که شبکه‌های عصبی انتشار به عقب یک جعبه سیاه است و روند آموزش آن نامرئی است. شبکه‌های عصبی فازی دارای لایه‌های اضافی است، یک لایه برای فازی سازی و یک لایه برای خارج کردن مقادیر از حالت فازی هستند. استفاده از شبکه‌های عصبی فازی در سیستم ارزیابی ریسک سرمایه‌گذاری سه مزیت دارد: اول، FNN دارای مکانیسم‌های پردازش موازی قوی است. دوم، با تغییر در محیط یاد می‌گیرند و می‌توانند قوانین خاص خود را ایجاد کنند. و سوم، از آنجاکه یک فرآیند مدل سازی غیرخطی است، می‌تواند به سرعت مدل‌های خود را بسازد. ژانگ [۲۸] از شبکه‌های عصبی مصنوعی برای پیش‌بینی ریسک اعتباری استفاده کرد. داده‌های پروژه سرمایه‌گذاری جمع‌آوری شده برای متناسب سازی رابطه بین ریسک سرمایه‌گذاری و بازدهی مورد استفاده قرار گرفت. علاوه بر این، مدل تجزیه و تحلیل حساسیت و معیار مدیریت کیفیت برای به دست آوردن شاخص ریسک معرفی شد. همچنین، ۱۵ شاخص ریسک با توجه به درجه حساسیت در ۶ دسته طبقه‌بندی شدند که کنترل ریسک برای سرمایه‌گذاران را فراهم می‌کند تا به آن‌ها در کنترل ریسک‌های سرمایه‌گذاری کمک کند.

بانک‌ها و وام‌دهندگان از امتیازدهی اعتبار برای ارزیابی توانایی بازپرداخت وام‌گیرندگان استفاده می‌کنند. بانک‌ها از دهه ۱۹۴۰ از اعتبارسنجی استفاده کرده‌اند. ریسک اعتباری بر اساس داده‌های آموزش تعیین می‌شوند. این روش با رتبه‌بندی اعتباری که بر اساس تجربه ذهنی ارزیابی کنندگان تعیین می‌شود متفاوت است. یک الگوریتم امتیازدهی اعتباری شامل فاکتورهای ریسک و وزن آن‌ها است. مؤسسات مالی کوچکی که از امتیازدهی اعتباری استفاده می‌کنند می‌توانند عملکرد خود را از نظر رشد و کارایی بهبود بخشند اما هزینه آن‌ها هم برای توسعه ابزار امتیازدهی اعتبار و هم برای غربالگری وام‌گیرندگان خوب زیاد است.

از آنجاکه الگوریتم‌های امتیازدهی اعتبار با استفاده از تکنیک‌های آماری سنتی نیاز به داده‌های ساختاری دارند، پژوهشگران علاقه‌مند بودند تا ببینند آیا هوش مصنوعی با استفاده از داده‌های غیر ساختاری می‌تواند بهتر کار کند یا خیر. این کار می‌تواند منجر به کاهش هزینه‌های توسعه مدل شود. یک مطالعه اولیه پنج روش امتیازدهی اعتبار را برای ریسک صنعت خرده‌فروشی مقایسه کرد. این روش‌ها شامل بیزین ساده، رگرسیون لجستیک، تقسیم‌بندی بازگشتی، ANN و بهینه‌سازی حداقل متوالی ( $SMO^1$ ) بودند [۱۹]. با این حال، یک مطالعه متفاوت نشان داد که شبکه‌های عصبی به خوبی برای مسئله امتیازدهی اعتبار کار می‌کنند. این مطالعه روش سنتی و هوش مصنوعی را مقایسه می‌کند: تجزیه و تحلیل تفکیکی<sup>۲</sup>، رگرسیون لجستیک، ANN، و درخت طبقه‌بندی و رگرسیون ( $CART^3$ ). این تحقیق نشان داد که CART و NN از لحاظ دقت پیش‌بینی و خطا از مدل‌های سنتی امتیازدهی اعتباری بهتر عمل می‌کنند [۲۰].

در اعطای وام به مشتریان بی‌بضاعت، مؤسسات اعتباری کوچک از قضاوت ذهنی مبتنی بر روانشناسی مشتریان برای تصمیم‌گیری استفاده می‌کنند. در یک مطالعه در بانک تونس، محققان سعی کردند ساختار مشتری و متغیرهای روان‌شناختی را در یک مدل CART قرار دهند. محققان دریافتند که این مدل کمک می‌کند تا نسبت وام‌های بد طبقه‌بندی شده به عنوان وام‌های خوب ۳.۱۲۵٪ کاهش می‌یابد، که منجر به کاهش زیان MFI تا ۴.۸٪ می‌شود. این مدل در مقایسه با مدل‌های تحلیل تفکیکی و رگرسیون لجستیک، به ترتیب ۶.۸ و ۱۳.۵ درصد عملکرد بهتری داشت [۲۳].

مزیت مدل‌های هوش مصنوعی نسبت به تکنیک‌های آماری سنتی این است که به توزیع نرمال نیازی ندارند. آن‌ها می‌توانند به صورت مستقیم مدل‌ها را از مجموعه داده‌های آموزشی تولید کنند و برای روابط غیرخطی بهتر عمل می‌کنند. با این حال، همه مدل‌های هوش مصنوعی در طبقه‌بندی شرکت‌ها برای رتبه‌بندی اعتباری به یک اندازه مؤثر نیستند.

<sup>1</sup> Sequential Minimal Optimization (SMO)

<sup>2</sup> Discriminant Analysis

<sup>3</sup> Classification and Regression Trees (CART)



مطالعه‌ای بر روی ۴۶ بنگاه اقتصادی در چین نشان داد که یک شبکه عصبی احتمالی بهتر از شبکه عصبی انتشار به عقب و یا شبکه عصبی رگرسیونی عمل می‌کند [۲۶].

رتبه‌بندی پایدار برای سرمایه‌گذاران یک مسئله چالش‌برانگیز است. مؤسسات رتبه‌بندی به داده‌های خود اظهاری شرکت‌ها وابسته هستند. هر موسسه معیارهای مختلفی را در نظر می‌گیرد و عملکرد ذهنی است. داده‌های حجیم، هوش مصنوعی (AI) و پردازش زبان طبیعی (NLP) می‌توانند راه‌حلی برای این مسئله باشند. برای سنجش عملکرد شرکت‌ها و مؤسسات اغلب از داده‌های موجود در رسانه‌های اجتماعی و اطلاعات موجود در دسترس عموم استفاده می‌شود. علاوه بر از بین بردن جهت‌گیری‌های گزارش‌شده، داده‌های پنهان می‌توانند کشف شوند که منجر به امتیازات منفی و مثبت بودن می‌شود، درحالی‌که داده‌های خود اظهاری تنها امتیازات مثبت ایجاد می‌شود [۲۹].

در یک پژوهش دیگر [۳۰]، در یک مطالعه موردی بر روی داده‌های اعتباری ارمنستان نشان داده شد که مدل‌های یادگیری عمیق نسبت به مدل‌های ساختاریافته مرسوم بهتر عمل می‌کنند. داده‌ها شامل متغیرهای مالی و غیرمالی بیش از ۹۰۰۰ وام‌گیرنده وام کشاورزی از سال ۲۰۱۲ تا ۲۰۱۷ است. عملکرد شبکه‌های عصبی عمیق در مقایسه با روشهای متداول و گسترده مقایسه شدند. تأخیر در وام‌های گذشته، همراه با میزان وام و ارز، عوامل اصلی موثر در پیش‌بینی احتمال وام هستند. مجموعه‌ای از متغیرهای آماری مهم بین مدل‌های یادگیری عمیق متفاوت است و ثابت می‌کند که می‌تواند روابط غیر خطی را به تصویر بکشد.

# فصل سوم

## روش شناسی پژوهش

## ۳-۱ مقدمه

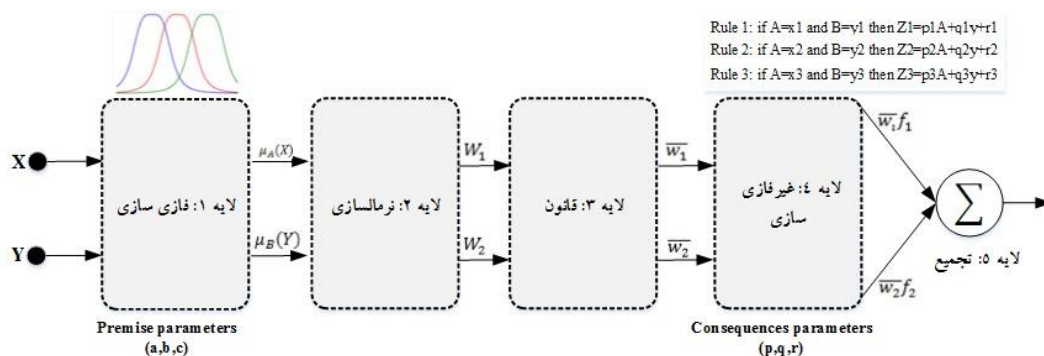
در سال‌های اخیر، هوش مصنوعی و یادگیری ماشین راه‌حل‌های زیادی برای مدیریت ریسک سرمایه‌گذاری فراهم کرده است. روش‌های بررسی‌شده نشان داده است که تکنیک‌های هوش مصنوعی می‌توانند قدرتمندتر و بیشتر شوند، اما محدود هستند و به‌کارگیری آن‌ها مشکل و چالش‌برانگیز است. در این فصل، یک روش جدید برای پیش‌بینی ریسک سرمایه‌گذاری با استفاده از شبکه عصبی فازی ارائه می‌شود. برای بهینه‌سازی FNN، از الگوریتم تکاملی استفاده می‌شود تا پارامترها را به‌صورت خودکار در طول آموزش بهینه کند.

## ۳-۲ شبکه‌های عصبی فازی

این مدل ترکیب تئوری فازی و شبکه عصبی بازگشت به عقب است. درواقع، مزایای شبکه عصبی و تئوری فازی است و یادگیری، شناسایی و پردازش اطلاعات را با هم ادغام می‌کند. طبق نظریه فازی، ویژگی‌های اساسی اشیاء در مرحله انتقالی نامشخص است و جنس آن‌ها مشخص نیست. انواع عدم اطمینان منطق فازی اغلب به عدم اطمینان تصادفی و عدم اطمینان واژگان تقسیم می‌شوند. عدم اطمینان تصادفی اغلب به احتمال تصادفی بودن اتفاقات اشاره دارد. بزرگ‌ترین مزیت منطق فازی این است که به کاربران اجازه می‌دهد رفتار سیستم مورد انتظار را با یک رابطه ساده توصیف کنند. با ترکیب منطق فازی و شبکه عصبی بازگشت به عقب، می‌توان نوعی شبکه عصبی ایجاد کرد که می‌تواند به‌صورت خودکار اطلاعات فازی را پردازش کند.

تفاوت بین شبکه عصبی<sup>۱</sup> FNN و انتشار به عقب<sup>۲</sup> (BP) این است که: (۱) برای منطق فازی، نیازی به دانش متخصصان است، درحالی‌که شبکه عصبی BP به کسب دانش از طریق الگوریتم پی می‌برد. (۲) منطق فازی از تابع عضویت برای نشان دادن دانش استفاده می‌کند، درحالی‌که شبکه عصبی BP از نمایش توزیع‌شده استفاده می‌کند. (۳) قوانین استدلال فازی قابل‌مشاهده هستند، درحالی‌که شبکه عصبی BP جعبه سیاه است و روند آموزش آن نامرئی است. (۴) ساختار شبکه عصبی BP اغلب شامل لایه ورودی، لایه پنهان و لایه خروجی است، درحالی‌که ساختار FNN اغلب شامل لایه ورودی، لایه فازی، لایه قانون، لایه ضد فازی و لایه خروجی است.

ساختار شبکه عصبی فازی شامل پنج لایه است، همان‌طور که در شکل ۳-۱ نشان داده شده است. لایه فازی سازی<sup>۳</sup>، لایه قانون<sup>۴</sup>، لایه نرمال‌سازی<sup>۵</sup>، لایه غیرفازی سازی<sup>۶</sup> و لایه تجمیع<sup>۷</sup> است.



شکل ۳-۱: ساختار شبکه‌های عصبی فازی [۳۶]

<sup>1</sup> Fuzzy Neural Network (FNN)

<sup>2</sup> Back Propagation (BP)

<sup>3</sup> Fuzzification Layer

<sup>4</sup> Rule Layer

<sup>5</sup> Normalization Layer

<sup>6</sup> Defuzzification Layer

<sup>7</sup> Summation Layer

۱. **لایه فازی سازی:** در این لایه، خوشه‌های فازی از مقادیر ورودی با استفاده از توابع عضویت، همانند تابع  $gbell$  و تابع مثلثی  $trimf$  به دست می‌آیند. لایه فازی سازی وظیفه تشکیل توابع عضویت با استفاده از پارامترهای موجود در قسمت پیشین<sup>۱</sup> (شامل  $a, b, c$ )، را بر عهده دارد. درواقع، این پارامترها درجه هر تابع عضویت را تعیین می‌کنند و با استفاده از رابطه‌های ۱-۳ و ۲-۳ بیان می‌شوند.

$$\mu_{A_i} = gbellmf(x; a, b, c) = \frac{1}{1 + \left| \frac{x-c}{a} \right|^{2b}} \quad (3-1)$$

$$O_i^1 = \mu_{A_i}(x) \quad (3-2)$$

۲. **لایه قانون:** در این لایه، وزن ( $w_i$ ) برای قوانین از طریق مقادیر عضویت محاسبه شده در لایه اول ایجاد می‌شود. وزن‌ها با ضرب مقادیر عضویت تولید می‌شوند (رابطه ۳-۳).

$$O_i^2 = w_i = \mu_{A_i}(x) \cdot \mu_{B_i}(y) \quad i=1,2 \quad (3-3)$$

۳. **لایه نرمال سازی:** در این لایه، وزنه‌ای نرمال شده هر قانون محاسبه می‌شود. وزن نرمال شده درصد قدرت تأثیرگذاری یک قانون نسبت به کل قدرت تأثیرگذاری است (رابطه ۳-۴):

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i=1,2 \quad (3-4)$$

۴. **لایه غیرفازی سازی:** در این لایه، وزن قوانین در هر گره با استفاده از چندجمله‌ای مرتبه اول محاسبه می‌شود (رابطه ۳-۵).

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (3-5)$$

در این معادله،  $\bar{w}_i$  خروجی لایه سوم است و سه پارامتر  $p_i, q_i$  و  $r_i$  پارامترهای نتیجه‌گیری<sup>۲</sup> هستند (یعنی پارامترهایی که در قسمت نتیجه هستند). به عنوان یک قاعده، تعداد پارامترهای قسمت نتیجه برای هر قانون یک بیشتر از تعداد ورودی است.

۵. **لایه تجمیع:** مقدار پیش‌بینی شده در آخرین لایه با تجمیع تمام خروجی‌های به دست آمده برای هر قانون در لایه چهارم به دست می‌آید (رابطه ۳-۶).

$$O_i^5 = \text{overall output} = \sum_i \bar{w}_i f_i = \frac{\sum_i \bar{w}_i f_i}{\sum_i \bar{w}_i} \quad (3-6)$$

### ۳-۳ الگوریتم تکامل تفاضلی تطبیقی

الگوریتم تکامل تفاضلی تطبیقی JADE [۳۷] یک نسخه خودسازگار از DE است که بجای عملگر جهش “DE/rand/1” (رابطه ۳-۷) از عملگر “DE/current-to-pbest/1” (رابطه ۳-۸) برای جهش استفاده می‌کند. عملگر “DE/rand/1” تمام راه‌حل‌ها برای ترکیب را به صورت تصادفی انتخاب می‌کند، درحالی‌که، عملگر “DE/current-to-pbest/1” بهترین راه‌حل در نسل جاری را با راه‌حل‌های تصادفی ترکیب می‌کند.

$$v_i^G = x_i^G + F \cdot (x_{best}^G - x_i^G) + F \cdot (x_{r1}^G - x_{r2}^G) \quad (3-7)$$

<sup>1</sup> Antecedent /premise

<sup>2</sup> Consequence/Conclusion

در روش پیشنهادی، عملگر تقاطع دوجمله‌ای غیرمتوالی<sup>۱</sup> برای تولید فرزندان از دو والد اعمال می‌شود. به‌جای مقدار ثابت برای همه کروموزوم‌ها، برای هر کروموزوم یک احتمال تقاطع اختصاصی وجود دارد. در مرحله انتخاب، بهترین کروموزوم از بین والدین و فرزندان بر اساس تناسب آن‌ها انتخاب می‌شود. در مرحله تطبیق پارامتر، از توزیع نرمال با میانگین  $\mu_{Cr}$  و انحراف استاندارد  $\sigma = 0.1$  برای تولید احتمال تقاطع هر کروموزوم به‌طور مستقل استفاده می‌شود. توزیع کوشی<sup>۲</sup> با پارامتر مکانی  $\mu_F$  و پارامتر مقیاس  $0.1$  برای تولید عامل جهش هر کروموزوم به‌طور مستقل استفاده شده است. سازگاری پارامتر برای هر دو  $\mu_{Cr}$  و  $\mu_F$  در پایان هر نسل انجام می‌شود.

### ۴-۳ متدولوژی پیشنهادی

#### ۴-۳-۱ چارچوب کنترل ریسک

در مرحله اول، بر اساس شاخص‌های ریسکی که بر درآمد سرمایه‌گذاری پروژه تأثیر می‌گذارند، روش FNN برای تطبیق رابطه منطقی بین ارزش شاخص ریسک و نرخ بازده معرفی می‌شود. پس‌از آن، مقدار ریسک هر شاخص از پروژه سرمایه‌گذاری به‌صورت پیوسته به‌عنوان مقدار ورودی تغییر می‌کند. علاوه بر این، از روش تجزیه و تحلیل حساسیت برای شناسایی حساسیت هر ریسک به درآمد استفاده می‌شود. شبکه عصبی مصنوعی یک سیستم محاسباتی چندلایه با تکنیک پیشخور<sup>۳</sup> است که قادر به یادگیری مستقل است، به همراه مزایایی همانند کاربرد بالا و عملی بودن. بنابراین، این پایان‌نامه مدل FNN را متناسب با رابطه بین ریسک سرمایه‌گذاری تأمین مالی و نرخ بازدهی معرفی می‌کند.

#### ۴-۳-۲ طراحی پیش‌بینی کننده FNN

سه مزیت استفاده از FNN در سیستم ارزیابی ریسک سرمایه‌گذاری وجود دارد: اول، FNN دارای مکانیسم پردازش موازی قوی، قابلیت سازگاری و استدلال قوی و توانایی خوب در پردازش اطلاعات فازی است، بنابراین FNN برای ارزیابی ریسک اعتباری سرمایه‌گذاران مناسب است. دوم، FNN توانایی زیادی در یادگیری با تغییر محیط دارد. این مسئله می‌تواند از طریق تعداد زیادی از داده‌ها قوانینی پیدا کند و استدلال صحیحی را به دست آورد، که برای در حالت اطلاعات کم مرتبط با سرمایه‌گذاران مناسب است. سوم، FNN یک فرایند مدل‌سازی غیرخطی طبیعی است، که می‌تواند به‌طور مؤثر بر مشکل انتخاب توابع مناسب در روند مدل‌سازی و تجزیه و تحلیل سنتی غلبه کند، بنابراین به‌سرعت مدل‌ها را می‌سازد.

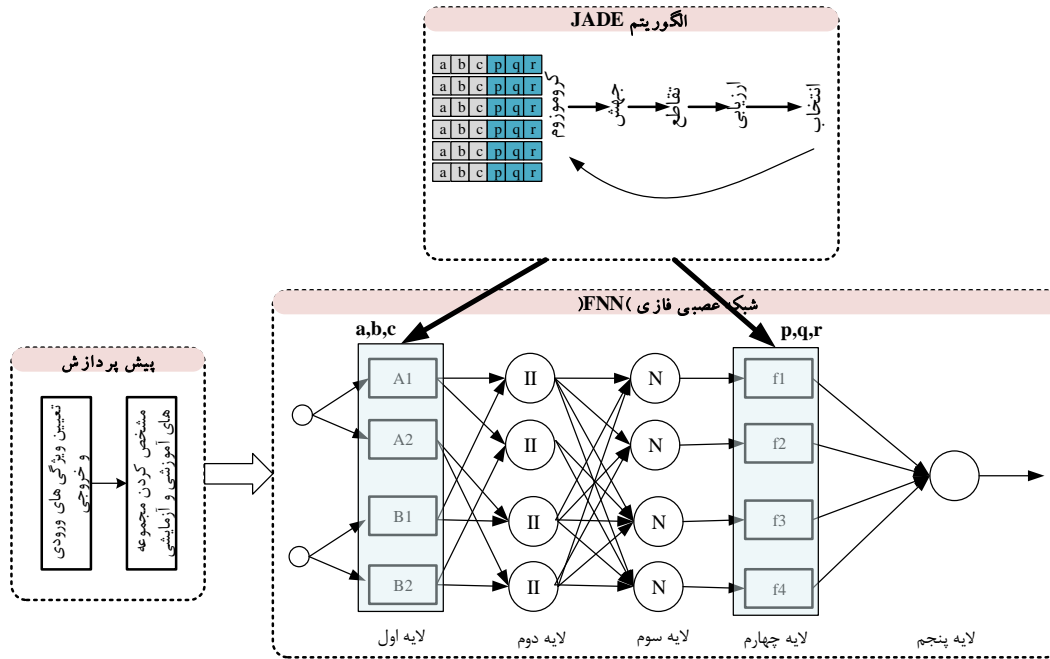
الگوریتم‌های بهینه‌سازی می‌توانند به دو روش برای بهینه‌سازی پارامترهای پیشین و نتیجه‌گیری در مدل FNN استفاده شوند. روش اول تنظیم همه پارامترهای دو قسمت تنها از طریق الگوریتم‌های بهینه‌سازی است. راه دوم این است که می‌توان در هر یک از لایه‌های FNN از یک الگوریتم بهینه‌سازی مستقل استفاده کرد. در متدولوژی پیشنهادی از روش اول استفاده می‌شود و الگوریتم JADE در هر دو قسمت پیشین و نتیجه‌گیری استفاده می‌شود.

هنگام استفاده از FNN در سیستم‌های ریسک سرمایه‌گذاری، عملکرد مدل FNN تحت تأثیر پارامترهای وابسته است. تنظیم مناسب پارامترها به‌طور مستقیم بر مدل پیش‌بینی نهایی تأثیر می‌گذارد. به‌طور کلی، پارامترهای FNN چالش‌های مختلفی را برای فرآیند جستجو از جمله اکتشاف و اکتشاف نامناسب، همگرایی کند و به دام انداختن در حداقل‌های محلی ایجاد می‌کنند. بنابراین تنظیم دقیق مقادیر پارامترهای FNN چالش اساسی است. شکل ۳-۲ چارچوب پیشنهادی برای بهینه‌سازی FNN با استفاده از آن را نشان می‌دهد.

<sup>1</sup> Non-consecutive binomial

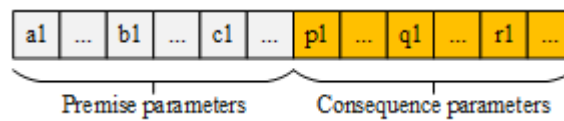
<sup>2</sup> Cauchy

<sup>3</sup> Feedforward



شکل ۳-۲: چارچوب پیشنهادی

در الگوریتم پیشنهادی، تمام پارامترهای پیشین و نتیجه به عنوان کروموزوم فرموله سازی می شوند. بنابراین، رویکرد JADE برای یافتن بهترین پارامترهای پیشین و نتیجه در فضای جستجو بکار گرفته می شود. نمایشی از یک کروموزوم در شکل ۳-۳ آورده شده است.



شکل ۳-۳: ساختار یک کروموزوم

برای محاسبه کیفیت کروموزومها، خطای میانگین مربع ریشه ( $RMSE^1$ ) به عنوان معیار تناسب در نظر گرفته شده است. برای محاسبه  $RMSE$ ، از مقدار پیش بینی شده و مقدار واقعی استفاده می شود (رابطه ۳-۸):

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \bar{y}_i)^2}{N}} \quad (3-8)$$

در این معادله،  $y_i$  به مقدار پیش بینی شده اشاره دارد،  $\bar{y}_i$  مقدار واقعی است و  $N$  اندازه مجموعه داده است. روش JADE سعی می کند مقادیر  $RMSE$  را به حداقل برساند.

<sup>1</sup> Root Mean Square Error (RMSE)

# فصل چهارم

## پیاده سازی و تحلیل یافته ها

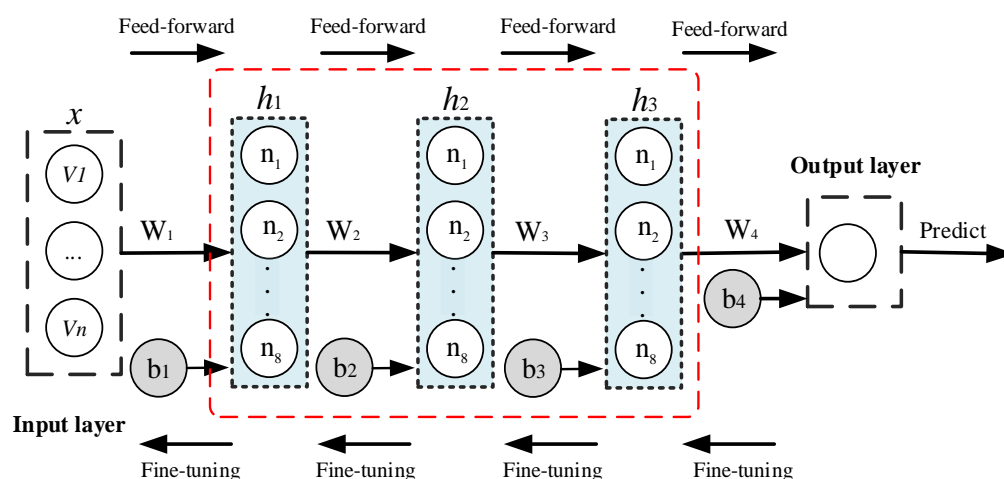
## ۴-۱ مقدمه

به منظور بررسی اثربخشی الگوریتم پیشنهادی در سیستم‌های پیش‌بینی ریسک سرمایه‌گذاری، آزمایش‌های شبیه‌سازی برای الگوریتم پیشنهادی در مقایسه با شبکه‌های عصبی فازی (FNN) و شبکه‌های عصبی پایه‌ای (ANN) انجام می‌شود. پیاده‌سازی‌ها با استفاده از زبان برنامه‌نویسی پایتون در محیط‌های Spyder و Jupyter انجام شد.

در زمان مدل‌سازی و برای جلوگیری از مشکل بیش برآزش، از دو استراتژی استفاده می‌شود: اعتبار سنجی ۵ تایی<sup>۱</sup> و توقف زودهنگام<sup>۲</sup>. در استراتژی اعتبارسنجی، مجموعه داده به پنج بخش تقسیم می‌شود و مدل با استفاده از چهار مجموعه آموزش داده می‌شود در حالی که با استفاده از یک مجموعه دیگر اعتبارسنجی می‌شود. این استراتژی با یک استراتژی توقف اولیه دنبال می‌شود، جایی که اگر مقدار زیان مربوط به مجموعه داده اعتبارسنجی برای چندین تکرار بدون تغییر باقی بماند، مرحله آموزش به پایان می‌رسد.

## ۴-۲ تنظیمات آزمایش‌ها

یک معماری ANN طراحی شده است که شامل سه لایه پنهان است که در هر لایه هشت نورون وجود دارد. اندازه لایه ورودی برابر با تعداد متغیرهای وابسته در مجموعه داده است، در حالی که مدل فقط یک نورون خروجی دارد که مقدار ریسک سرمایه‌گذاری است. در این آزمایش‌ها، مقدار پارامتر برای میزبان یادگیری روی ۰.۱ تنظیم شده است. این پارامتر اندازه به‌روزرسانی وزن را تنظیم می‌کند تا مقدار تابع هزینه را به حداقل برساند. معماری مدل ANN در شکل ۴-۱ به تصویر کشیده شده است. در این پژوهش، از الگوریتم آموزش انتشار به عقب برای بهینه‌سازی وزن‌ها و بایاس استفاده می‌شود. از تابع Sigmoid غیرخطی در لایه پنهان و لایه خروجی استفاده می‌شود. تعداد تکرارها ۱۰۰ تعیین می‌شود. حداقل خطای به‌دست‌آمده در فرآیند آموزش به عنوان یک معیار توقف تعریف می‌شود که در آن وقتی که مقدار خطا به ۰.۰۰۱ می‌رسد، آموزش متوقف می‌شود.



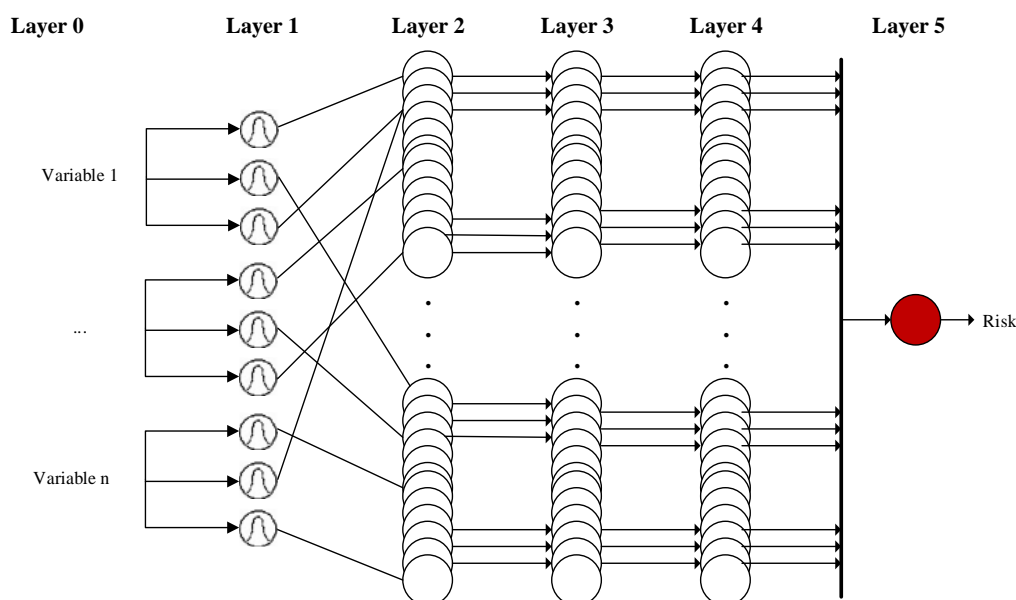
شکل ۴-۱: معماری ANN طراحی شده برای این پژوهش

<sup>1</sup> 5-Fold Cross Validation

<sup>2</sup> Early Stopping



برای ایجاد مدل FNN، از تابع فازی گاوسی<sup>۱</sup> برای فازی سازی هر متغیر استفاده شد (شکل ۴-۲). یک تابع فازی گاوسی با دو پارامتر نشان داده می‌شود: میانگین و واریانس. تعداد پارامترهای قابل اصلاح، که تحت تأثیر نوع و تعداد تابع فازی و همچنین سیستم است، زمان محاسبه و بار مورد نیاز برای تبدیل را تعیین می‌کند. در مدل طراحی شده، از روش prod برای رفع فازی استفاده می‌شود.



شکل ۴-۲: ساختار شبکه عصبی فازی طراحی شده برای پیش‌بینی ریسک

### ۴-۳ مجموعه داده‌ها

در این پایان‌نامه از یک مجموعه داده ریسک سرمایه‌گذاری استفاده شده است که انواع ریسک‌ها همانند ریسک بازگشت سرمایه و ریسک سیستماتیک را پوشش می‌دهد. این مجموعه داده از مخزن داده‌ای یادگیری ماشین<sup>۲</sup> گردآوری شده است. این مجموعه داده بر اساس امتیازدهی به معیارهای مختلف با استفاده از داده‌های بازار سهام ایالات متحده جمع‌آوری شده است. این مجموعه داده در چهار بازه زمانی گردآوری شدند و دارای ۲۴۸ نمونه است.

این مجموعه داده شامل ۶ ویژگی وابسته و ۶ ویژگی نهایی است که نشان‌دهنده انواع ریسک‌های سرمایه‌گذاری است. جدول ۴-۱ ویژگی‌های موجود در مجموعه داده مورد بررسی را نشان می‌دهد.

<sup>۱</sup> Gaussian

<sup>۲</sup> <https://archive.ics.uci.edu/ml/datasets/Stock+portfolio+performance>

جدول ۴-۱: ویژگی‌های داده‌ای موردبررسی

نام متغیر	نوع متغیر
Large B/P	متغیر ورودی
Large ROE	
Large S/P	
Large Return Rate in the last quarter	
Large Market Value	
Small systematic Risk	
Annual Return	متغیر خروجی
Excess Return	
Systematic Risk	
Total Risk	
Abs. Win Rate	
Rel. Win Rate	

دو نوع شاخص بازگشت و ریسک اغلب برای ارزیابی عملکرد ریسک سرمایه‌گذاری استفاده می‌شوند:

- نرخ بازگشت سالانه<sup>۱</sup>:

$$\text{Annual Return} = \left( (1 + R)^{\frac{1}{t}} \right) - 1$$

جایی که  $R$  نرخ بازگشت انباشته است؛  $t$  بازه زمان به تعداد سال است.

- نرخ بازگشت مازاد<sup>۲</sup>: می‌تواند از مقدار ضریب رگرسیون در معادله رگرسیون محاسبه شود. اگر نرخ بازدهی مازاد مثبت باشد، نشان می‌دهد که بازدهی سرمایه‌گذاری بهتر از بازدهی بازار است.

$$R_i - R_f = \alpha_i + \beta_i(R_m - R_f)$$

جایی که  $R_f$  نرخ بازگشت بدون ریسک؛  $R_m$  نرخ بازدهی بازار؛  $R_i$  نرخ بازدهی ریسک سرمایه‌گذاری هستند.

- Absolute winning rate:

$$Win_{abs} = \frac{n_1}{N}$$

جایی که  $n_1$  تعداد دوره‌های سرمایه‌گذاری است وقتی که نرخ بازدهی سرمایه‌گذاری بیش از ۰ است؛  $N$  تعداد کل دوره‌های سرمایه‌گذاری است.

- Relative winning rate:

$$Win_{Rel} = \frac{n_2}{N}$$

<sup>۱</sup> Annual Return

<sup>۲</sup> Excess return

جایی که  $n_1$  تعداد دوره‌های سرمایه‌گذاری است وقتی که نرخ بازدهی سرمایه‌گذاری نرخ بازدهی بازار را بیشتر می‌کند و  $N$  تعداد کل دوره‌های سرمایه‌گذاری است.

- **ریسک سیستماتیک:** می‌تواند از مقدار ضریب رگرسیون تخمین زده شود. هرچه مقدار ضریب بیشتر باشد، ریسک سیستماتیک ریسک سرمایه‌گذاری بیشتر است.
- **ریسک کلی:** می‌تواند از انحراف استاندارد نرخ بازگشت پورتفولیو برآورد شود. این معیار نشان‌دهنده نوسانات نرخ بازگشت در یک دوره خاص است. نوسانات بیشتر برای نرخ بازگشت نشان‌دهنده خطر پورتفولیو بالاتر است.

#### ۴-۴ معیارهای ارزیابی

پس از توسعه مدل‌های پیش‌بینی برای پیش‌بینی ریسک سرمایه‌گذاری، نیاز به ارزیابی مدل‌های توسعه‌یافته وجود دارد. روند ارزیابی در این مطالعه با استفاده از برخی از شاخص‌های ارزیابی همانند RMSE، MAE و  $R^2$  برای هر دو مرحله آموزش و آزمایش انجام شد. رابطه‌های ۴-۱، ۴-۲ و ۴-۳ به ترتیب نحوه محاسبه این معیارها را نشان می‌دهند.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4-1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4-2)$$

$$R^2 = \frac{[\sum_{i=1}^n (y_i - y_{mean})^2] - [\sum_{i=1}^n (y_i - \hat{y}_i)^2]}{[\sum_{i=1}^n (y_i - y_{mean})^2]} \quad (4-3)$$

$\hat{y}_i$ : مقدار پیش‌بینی شده برای نمونه  $i$

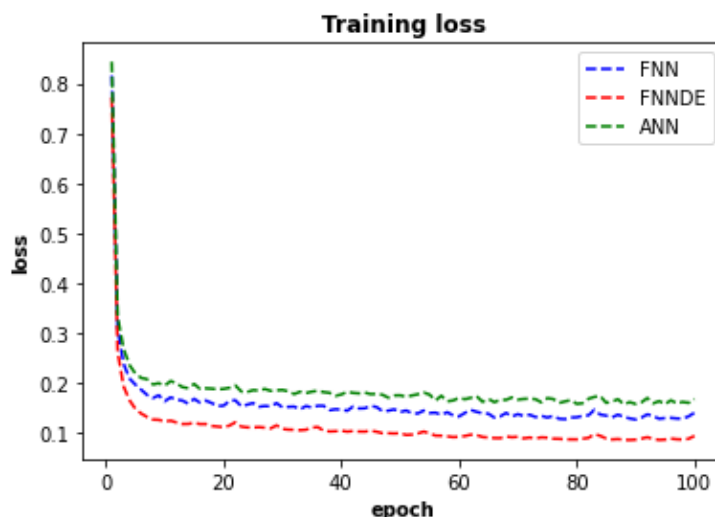
$y_i$ : مقدار واقعی برای نمونه  $i$

$n$ : تعداد نمونه‌های داده‌ای

$y_{mean}$ : مقدار میانگین برای متغیر هدف

#### ۴-۵ نتایج و تحلیل

شکل ۴-۳ میانگین خطای مدل‌های پیش‌بینی ریسک سرمایه‌گذاری در طول آموزش برای تکرارهای مختلف را نشان می‌دهد. این آزمایش با ۱۰۰ تکرار روند مدل‌سازی انجام شد. همان‌طور که مشاهده می‌شود، الگوریتم FNNDE توانست نسبت به الگوریتم‌های FNN و ANN در پیش‌بینی ریسک سرمایه‌گذاری خطای کمتری داشته باشد. میانگین خطای FNNDE حدود ۱٪ کمتر از FNN و حدود ۱.۵٪ کمتر از ANN بود. این به دلیل استفاده از الگوریتم DE تطبیقی به‌عنوان یک الگوریتم قدرتمند بهینه‌سازی برای بهینه‌سازی توابع فازی سازی در مدل FNN است.



شکل ۴-۳: متوسط خطای مدل‌های پیش‌بینی از نظر تعداد تکرار

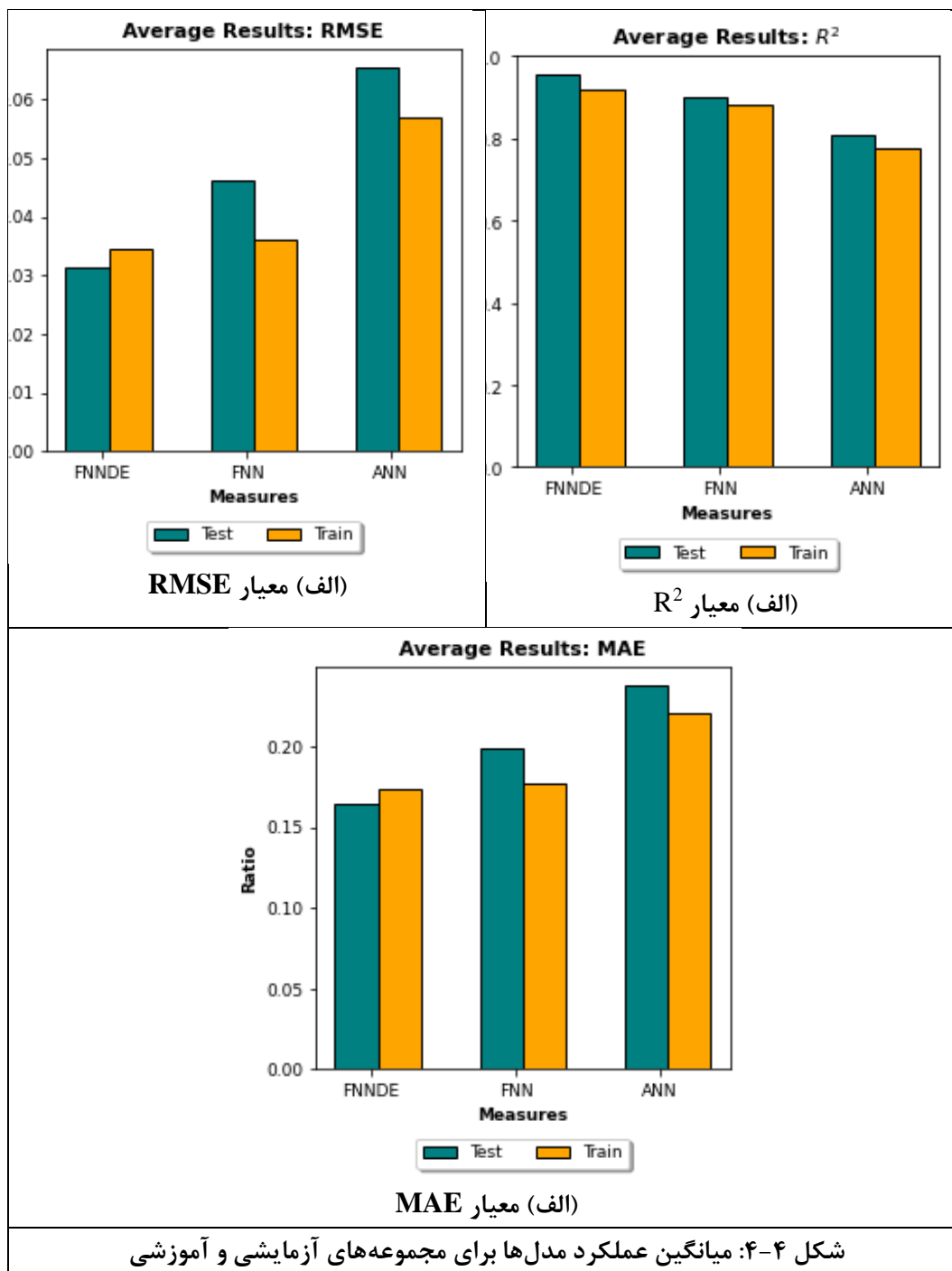
جدول ۴-۲ عملکرد مدل‌های پیش‌بینی فاکتورهای نهایی را از نظر معیارهای مختلف، از جمله RMSE، MAE،  $R^2$  و نشان می‌دهد. مطابق این جدول، کمترین اختلاف بین مقادیر واقعی و مقادیر پیش‌بینی‌شده توسط الگوریتم FNNDE فراهم می‌شود.

جدول ۴-۲: مقدار عملکرد مدل‌ها بر حسب معیارهای مختلف برای پیش‌بینی اهداف مختلف

$R^2$		MAE		RMSE		مدل	هدف
آزمایشی	آموزشی	آزمایشی	آموزشی	آزمایشی	آموزشی		
0.915	0.957	0.1946	0.158	0.0422	0.028	FNNDE	بازگشت اعتبار
0.898	0.909	0.177	0.1945	0.0354	0.0436	FNN	
0.869	0.791	0.21	0.2182	0.05	0.0537	ANN	
0.946	0.952	0.154	0.169	0.02874	0.0328	FNNDE	ریسک سیستماتیک
0.91	0.931	0.1788	0.1796	0.037	0.0383	FNN	
0.763	0.771	0.21	0.268	0.0515	0.0832	ANN	
0.923	0.95	0.187	0.164	0.0398	0.0316	FNNDE	بازگشت مازاد
0.887	0.9	0.173	0.2	0.0352	0.046	FNN	
0.741	0.848	0.247	0.225	0.073	0.057	ANN	
0.897	0.955	0.1565	0.1674	0.0276	0.0327	FNNDE	ریسک کلی
0.837	0.865	0.181	0.221	0.0371	0.0574	FNN	
0.733	0.817	0.213	0.238	0.0531	0.0672	ANN	

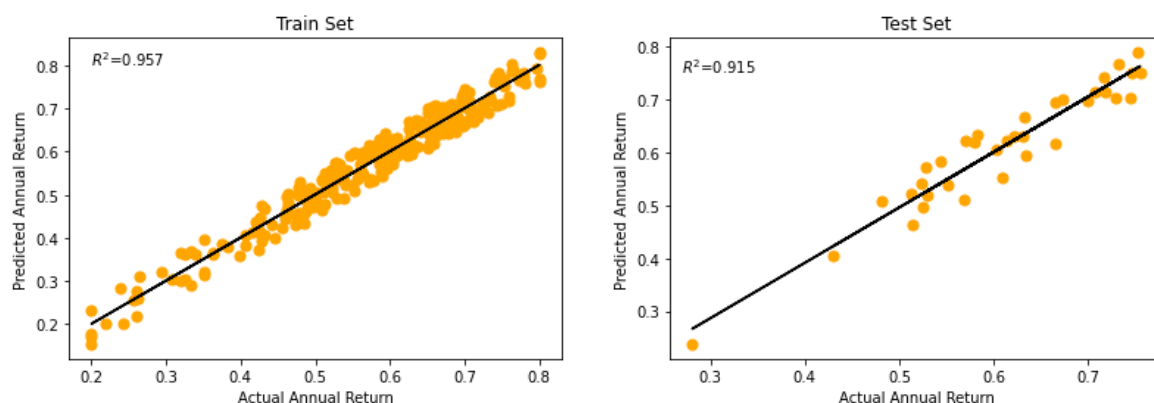
با توجه به جدول ۴-۲، شکل ۴-۴ میانگین عملکرد مدل‌های پیش‌بینی کننده از نظر معیارهای مختلف برای مجموعه‌های آزمایشی و آموزشی را نشان می‌دهد. همانگونه که مشاهده می‌شود، مدل FNNDE پیشنهادی برای تمام معیارها بالاترین عملکرد را بدست آورد. برای معیار RMSE، مدل پیشنهادی دارای میانگین نرخ تقریبی ۳٪ است که حدود ۱٪ کمتر از مدل FNN و ۳٪ کمتر از ANN است. میانگین MAE برای مدل FNNDE حدود ۱۶.۸٪ است، درحالیکه این مقدار برای مدل‌های FNN و ANN به ترتیب حدود ۱۸.۸٪ و

۲۲.۸٪ هستند. نتایج برای معیار  $R^2$  نشان می‌دهد که روش پیشنهادی توانسته است مقدار عملکرد ۹۳.۶٪ را بدست آورد، در مقابل ۸۹.۲٪ و ۷۹.۱٪ به ترتیب برای مدل‌های FNN و ANN.

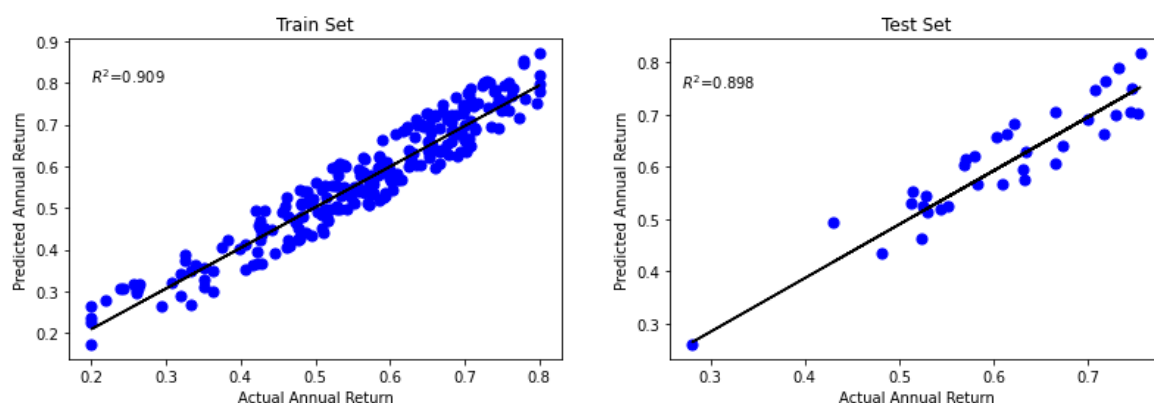


شکل‌های ۴-۵، ۴-۶، ۴-۷ و ۴-۸ عملکرد مدل‌های پیش‌بینی به ترتیب برای مقدار ریسک بازگشت اعتبار، ریسک سیستماتیک، بازگشت مازاد و ریسک گلی را نشان می‌دهند. با توجه به نمودارهای شکل ۴-۴، مشخص شد که مدل FNNDE با مقادیر  $R^2$  برابر با ۰.۹۵۷ و ۰.۹۱۵ به ترتیب برای مراحل آموزش و آزمایش

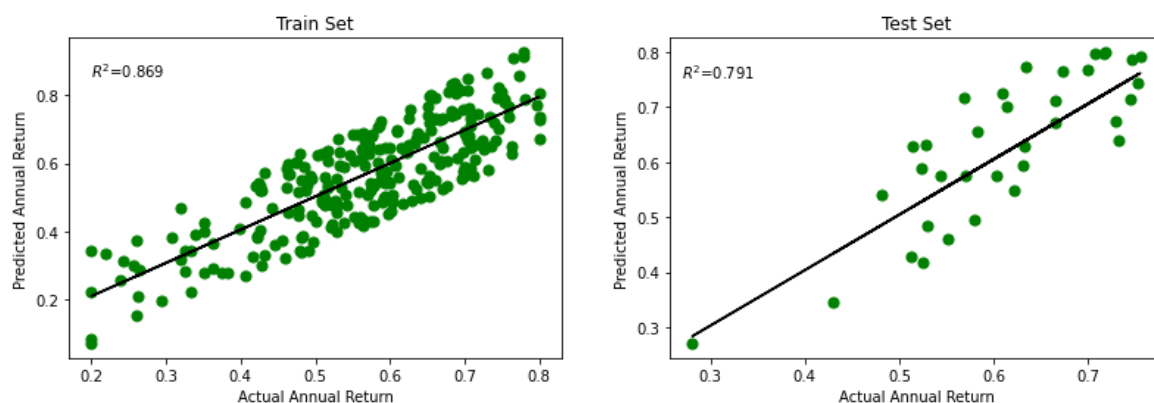
مطمئن‌ترین روش پیش‌بینی برای ریسک بازگشت اعتبار است و به دنبال آن مدل‌های FNN و ANN با مقادیر (۰.۸۹۸ و ۰.۹۰۹) و (۰.۸۶۹، ۰.۷۹۱).



(الف) الگوریتم FNNDE



(ب) الگوریتم FNN



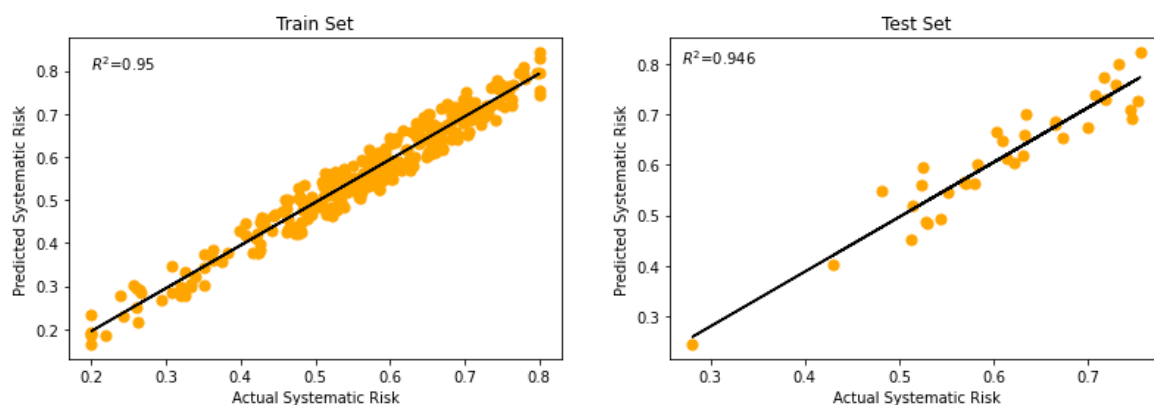
(ج) الگوریتم ANN

شکل ۴-۵: عملکرد مدل‌های هوش مصنوعی در پیش‌بینی مقدار ریسک بازگشت اعتبار

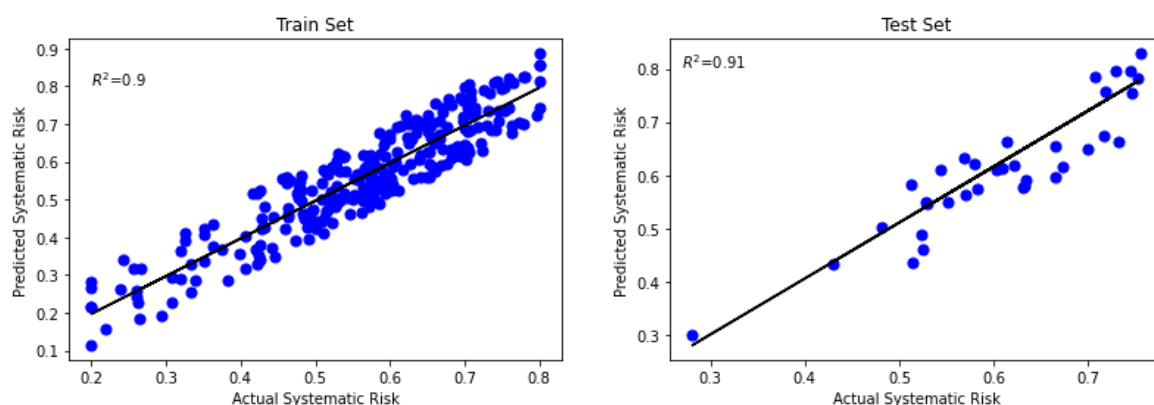
ریسک سیستماتیک به ریسک ذاتی کل بازار اشاره دارد. ریسک سیستماتیک همچنین به عنوان "ریسک بازار" شناخته می‌شود. این ریسک نه تنها بر روی یک سهام یا صنعت خاص، بلکه بر روی کل بازار تأثیر می‌گذارد. ریسک سیستماتیک زمینه‌ساز سایر ریسک‌های سرمایه‌گذاری مانند ریسک صنعت است. برای مثال، اگر سرمایه‌گذار بیش از حد بر سهام امنیت سایبری تأکید کرده باشد، می‌توان با سرمایه‌گذاری در طیف وسیعی از سهام در بخش‌های دیگر، مانند بهداشت و درمان و زیرساخت‌ها، تنوع ایجاد کرد. با این حال، ریسک سیستماتیک

شامل تغییرات نرخ بهره، تورم، رکود و جنگ و سایر تغییرات عمده است. جابجایی در این دامنه‌ها می‌تواند کل بازار را تحت تأثیر قرار دهد و با تغییر موقعیت در سبد، سهام عمومی نمی‌تواند کاهش یابد.

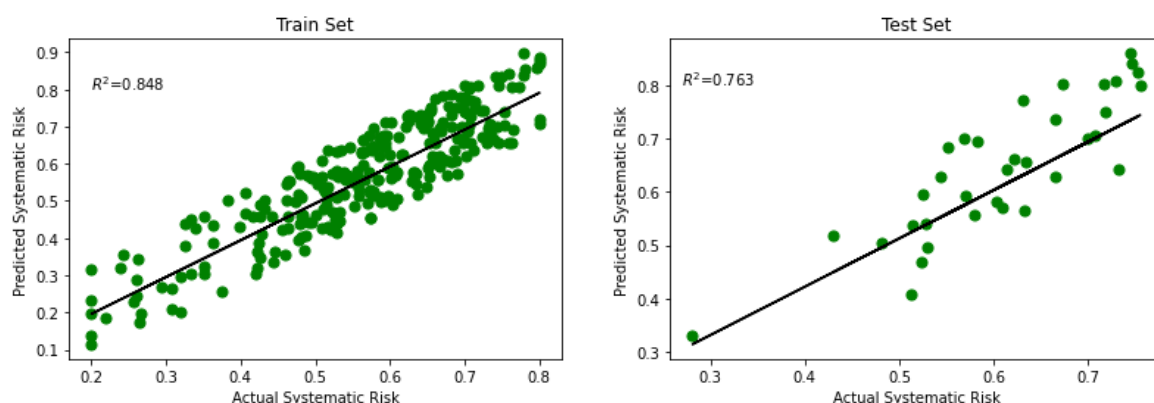
شکل ۴-۶ مقدار پیش‌بینی‌شده در مقابل مقدار واقعی برای ریسک سیستماتیک به همراه مقدار  $R^2$  را نشان می‌دهد. مدل FNNDE توانست مقدار  $R^2$  برابر با ۰.۹۵۲ و ۰.۹۴۶ به ترتیب برای نمونه‌های آموزشی و آزمایشی فراهم کند. مقادیر RMSE و MAE این الگوریتم به ترتیب (۰.۰۳۲۸، ۰.۰۲۸۷) و (۰.۱۶۹، ۰.۱۵۴) برای مجموعه‌های آموزشی و آزمایشی است.



(الف) الگوریتم FNNDE

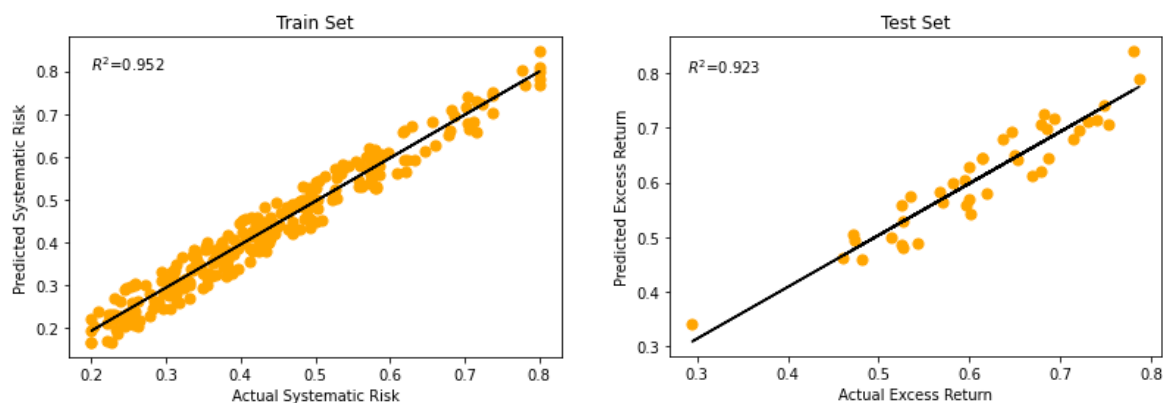


(ب) الگوریتم FNN

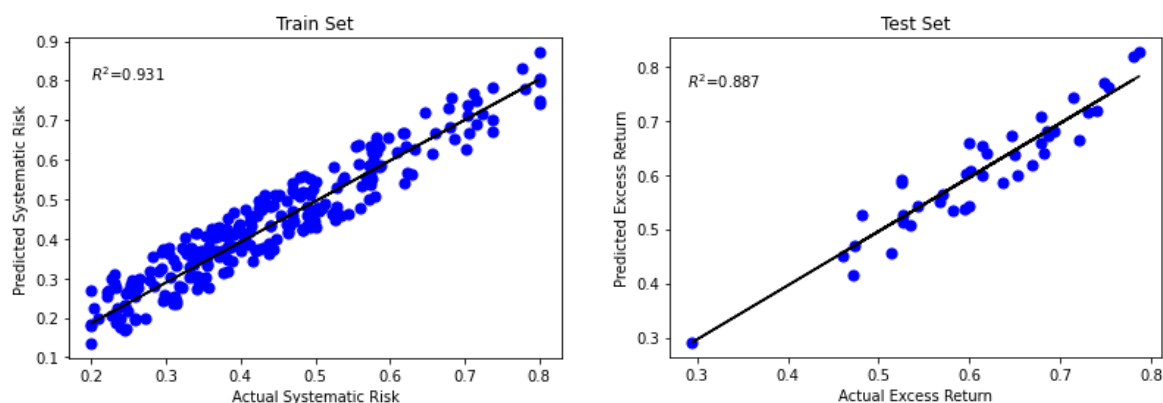


(ج) الگوریتم ANN

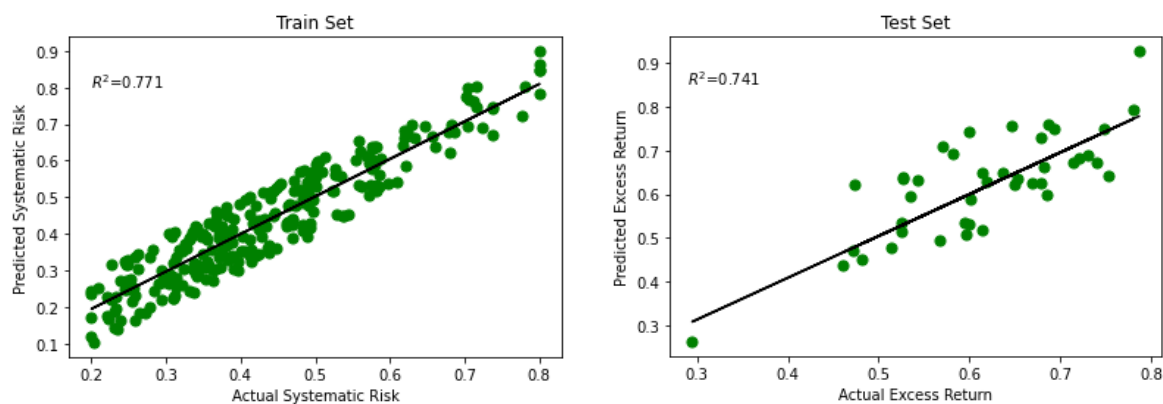
شکل ۴-۶: عملکرد مدل‌های هوش مصنوعی در پیش‌بینی مقدار ریسک سیستماتیک



(الف) الگوریتم FNND



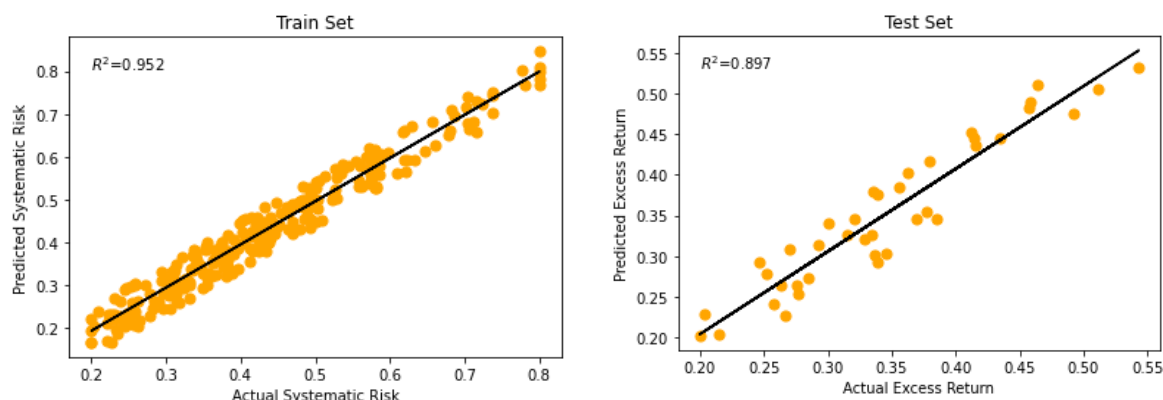
(ب) الگوریتم FNN



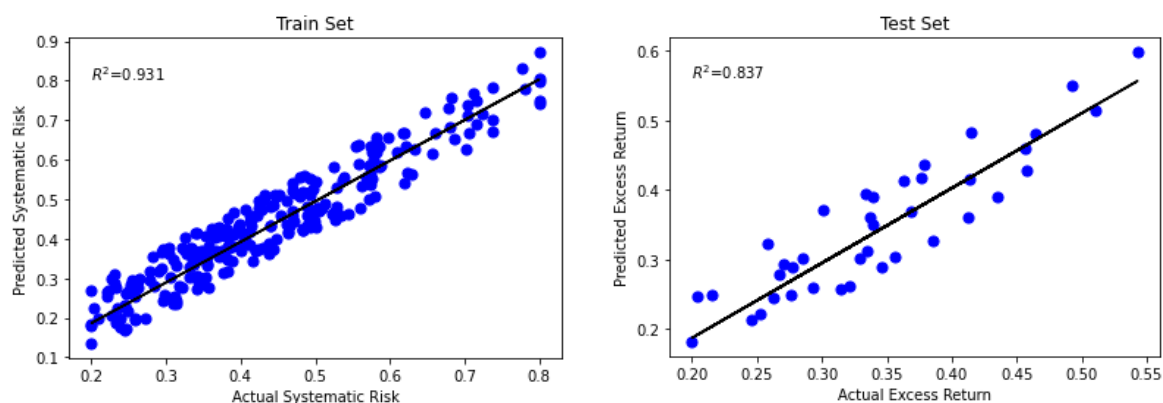
(ج) الگوریتم ANN

شکل ۴-۷: عملکرد مدل‌های هوش مصنوعی در پیش‌بینی مقدار ریسک سیستماتیک

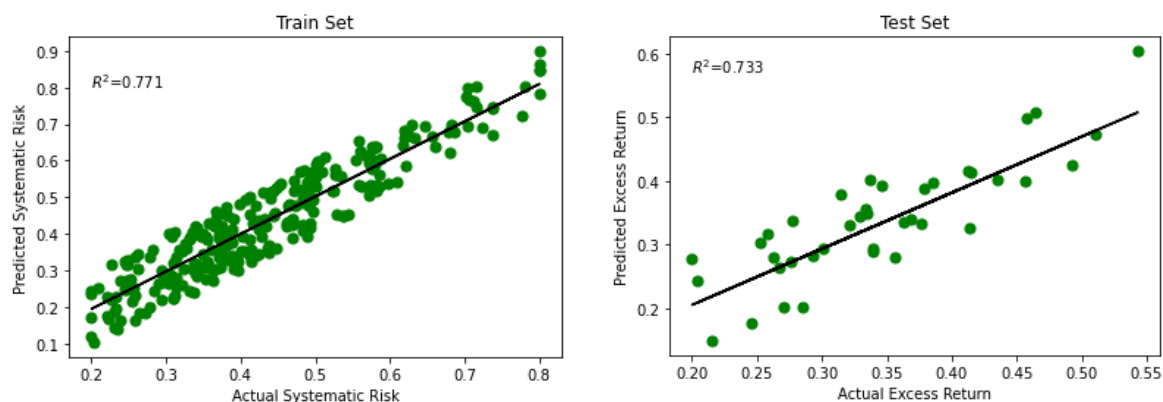




(الف) الگوریتم FNND



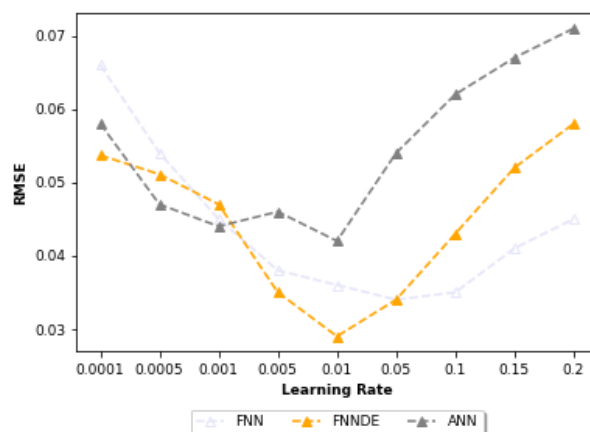
(ب) الگوریتم FNN



(ج) الگوریتم ANN

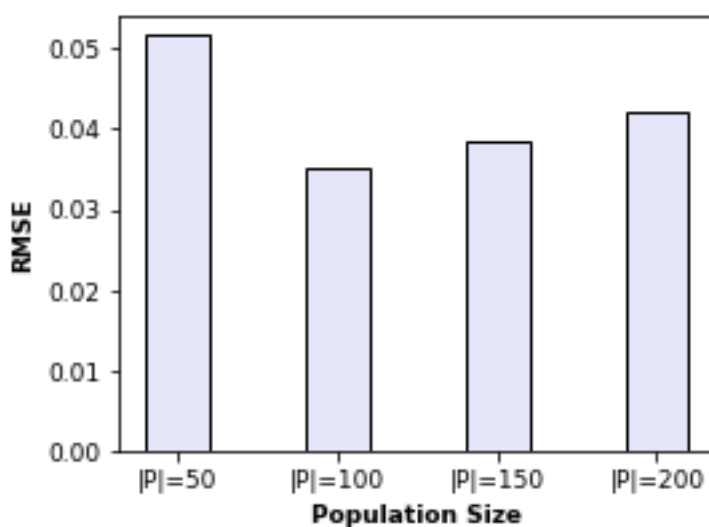
شکل ۴-۸: عملکرد مدل‌های هوش مصنوعی در پیش‌بینی مقدار ریسک کلی

شکل ۴-۹ تأثیر نرخ یادگیری را نشان می‌دهد که مقدار نرخ یادگیری از ۰.۰۰۰۱ تا ۰.۲ متغیر است. با تغییر مقدار نرخ یادگیری، مقدار RMSE تغییر می‌کند بطوریکه با افزایش نرخ یادگیری مقدار RMSE کاهش می‌یابد و بهترین عملکرد برای نرخ یادگیری ۰.۰۱ به‌دست‌آمده است.



شکل ۴-۹: تأثیر نرخ یادگیری بر روی عملکرد مدل‌های پیش‌بینی

اندازه جمعیت یکی از مهم‌ترین پارامترهای تأثیرگذار در الگوریتم‌های تکاملی است. در این آزمایش، چهار اندازه جمعیت ۵۰، ۱۰۰، ۱۵۰ و ۲۰۰ در نظر گرفته شده است. شکل ۴-۱۰ مقدار RMSE مدل FNNDE را هنگام تغییر تعداد کروموزوم نشان می‌دهد. نشان داده شده است که بهترین مقادیر زمانی به دست می‌آیند که اندازه جمعیت ۱۰۰ و ۱۵۰ باشد.



شکل ۴-۱۰: تأثیر اندازه جمعیت بر روی عملکرد مدل FNNDE

# فصل پنجم

## نتیجه گیری و پیشنهادها

## ۵-۱ نتیجه‌گیری

پیش‌بینی ریسک سرمایه‌گذاری نقش اساسی در مدیریت منابع مالی دارد که باعث بیشینه کردن مقدار سود سرمایه‌گذاری می‌شود. در این پژوهش، یک مدل هوشمند ترکیبی جدید با استفاده از الگوریتم شبکه‌های مصنوعی فازی (FNN) و رویکردهای تکاملی توسعه داده شد. تعیین مقدار پارامترهای مربوط به توابع عضویت در لایه‌های فازی سازی یک چالش اساسی در الگوریتم FNN است. تعیین نامناسب این مقادیر منجر به بیش برآزش و گیرکردن الگوریتم در بهینه محلی می‌شود که در نتیجه باعث پیش‌بینی نادرست ریسک سرمایه‌گذاری می‌شود. در روش پیشنهادی، از الگوریتم JADE (که یک نسخه جدید از الگوریتم تکامل تفاضلی با تطبیق خودکار مقادیر پارامترهای DE است) برای بهینه‌سازی مقدار پارامترهای توابع عضویت در لایه‌های فازی استفاده می‌شود. پارامترهای توابع عضویت به صورت راه‌حل‌های کروموزومی کدگذاری می‌شوند که شامل یک بردار با مقادیر پیوسته بین ۰ و ۱ است.

روش پیشنهادی با الگوریتم شبکه عصبی مصنوعی (ANN) و FNN پایه‌ای با استفاده از معیارهای مختلف همانند RMSE، MAE و  $R^2$  مقایسه شد. از یک مجموعه داده که شامل ویژگی‌های مربوط به وزن‌دهی امتیازی است استفاده شده است. این ویژگی‌های به‌عنوان ورودی مدل استفاده می‌شود و انواع مختلف ریسک همانند ریسک برگشت سرمایه و ریسک سیستماتیک به‌عنوان خروجی‌های مدل در نظر گرفته می‌شوند. برای هر یک از پیش‌بینی ریسک‌ها معیارهای ارزیابی محاسبه و با خروجی مدل‌های دیگر مقایسه شدند.

نتایج شبیه‌سازی‌ها نشان دادند که روش پیشنهادی توانست پیش‌بینی دقیق‌تری نسبت به مدل‌های ANN و FNN فراهم کند. میانگین مقدار  $R^2$  برای روش پیشنهادی برابر با ۹۳٫۶٪ در مقابل ۹۰٪ و ۸۳٪ به ترتیب برای FNN و ANN است. هرچقدر مقدار  $R^2$  برای یک مدل پیش‌بینی کننده به ۱ نزدیک‌تر باشد نشان‌دهنده بهتر بودن مدل است. برای معیار RMSE، مدل‌های FNN، FNNDE و ANN به ترتیب ۰٫۳۵۱، ۰٫۳۹۵ و ۰٫۵۱۸ است. برخلاف  $R^2$ ، هرچقدر مقدار RMSE برای یک مدل پیش‌بینی کننده ریسک به ۰ نزدیک‌تر باشد به معنی بهتر بودن مدل پیش‌بینی است. مدل FNNDE برای تمام معیارهای موردبررسی بهتر عمل می‌کند.

## ۲-۵ پیشنهادها

یادگیری فدرالی<sup>۱</sup> برای جلوگیری از نشت داده‌ها، روش‌های آموزش توزیع‌شده را اتخاذ می‌کند. هدف از یادگیری فدرالی حفظ حریم خصوصی و امنیت داده‌ها با آموزش یک مدل مشترک بدون دسترسی به داده‌های نگهداری شده توسط شخص سوم است. برای پرداختن به چالش‌های حریم خصوصی داده‌ها در امور مالی، می‌توان از تکنیک یادگیری فدرالی استفاده کرد تا آموزش مدل چند طرفه بدون مبادله داده‌های سرمایه‌گذاری خصوصی مشتری یا حساس انجام شود. ترکیب یادگیری فدرالی و حریم خصوصی تفاضلی<sup>۲</sup> یکی از تکنیک‌های پر استفاده است که نویز بدون تأثیرگذاری بر روی عملکرد مدل نهایی به داده‌ها اضافه می‌شود.

برای مسئله پیش‌بینی ریسک سرمایه‌گذاری که اغلب دچار ویژگی‌های داده‌ای پویا و نویزی هستند، روش‌های یادگیری قوی<sup>۳</sup> یکی از راه‌حل‌های محتمل است. از آنجاکه این ویژگی‌های داده‌ای فقط مختص حوزه مالی نیستند، یک جریان تحقیقاتی در حوزه‌های مختلف وجود دارد تا از داده‌های نویزی یاد بگیریم. روش‌های یادگیری قوی ممکن است هم در یک زمینه طبقه‌بندی یا در زمینه سری زمانی اجرا شوند.

استفاده از الگوریتم‌های یادگیری ژرف همانند شبکه‌های عصبی کانولوشن (CNN<sup>۴</sup>) و خودرمزگذارها<sup>۵</sup> می‌توانند برای پیش‌بینی ریسک سرمایه‌گذاری استفاده شوند. یادگیری ژرف قادر به یادگیری سلسله مراتبی در معماری‌های عمیق هستند. الگوریتم‌های یادگیری ژرف نیاز به حجم داده‌های زیادی دارند، درحالی‌که مجموعه داده‌های مرتبط با ریسک سرمایه‌گذاری داده‌های دارای تعداد نمونه‌های کمی هستند. بنابراین استفاده از رویکردهایی همانند نمونه‌برداری تقویتی (Oversampling) در ترکیب با یادگیری ژرف می‌تواند مفید باشد.

به‌عنوان فرصت تحقیقاتی بعدی، از رویکرد تجمیعی می‌توان برای بهبود دقت پیش‌بینی ریسک استفاده کرد. در این استراتژی الگوریتم‌های پیش‌بینی کننده می‌توانند با یکدیگر تجمیع شوند و نزدیک‌ترین پیش‌بینی و یا میانگین پیش‌بینی‌ها به‌عنوان پیش‌بینی نهایی در نظر گرفته شود. چندین الگوریتم یکسان و یا چندین الگوریتم متفاوت می‌توانند در مرحله تجمیع استفاده شوند.

<sup>1</sup> Federated Learning

<sup>2</sup> Differential Privacy

<sup>3</sup> Robust Learning

<sup>4</sup> Convolutional Neural Network (CNN)

<sup>5</sup> autoencoders

# منابع و مآخذ

- [1] R. Zhang, Research on brand trust and financing risk preference of E-commerce based on neuroeconomic experiment, *NeuroQuantology* 16 (4) (2018) 12–15, <http://dx.doi.org/10.14704/nq.2018.16.4.1215>.
- [2] M. Gong, Research and application of credit rating model in small and micro enterprises based on fuzzy neural network, *JDMSC* 20 (4) (2017) 817–834, <http://dx.doi.org/10.1080/09720529.2017.1364042>.
- [3] W. Zhou, Research and application of data mining algorithm based on fuzzy neural network for nonlinear problems in large data environment, *J. Comput. Theor. Nanosci.* 13 (7) (2016) 4735–4738, <http://dx.doi.org/10.1166/jctn.2016.5345>.
- [4] G. Kang, N. Zhuang, Research on early risk predictive model and discriminative feature selection of cancer based on real-world routine physical examination data, *BIBM* 173 (2017) 1512–1519, <http://dx.doi.org/10.1109/bibm.2016.7822746>.
- [5] Z.O. Xu, J.C. Li, S. Xiao, Y.Z. Yuan, Study on security risk assessment of power system based on BP neural network, *J. Comput. Theor. Nanosci.* 13 (8) (2016) 5277–5280, <http://dx.doi.org/10.1166/jctn.2016.5414>.
- [6] Telikani, A., Tahmassebi, A., Wolfgang, B., and Gandomi, A., 2021. Evolutionary Machine Learning: A Survey. *ACM Computing Surveys*, 54(8), pp.11-50.
- [7] Joost Verbraeken, Matthijs Wolting, Jonathan Katzy, Jeroen Kloppenburg, Tim Verbelen, and Jan S Rellermeyer. 2020. A Survey on Distributed Machine Learning. *ACM Computing Surveys (CSUR)* 53, 2 (2020), 1–33.
- [8] Ibrahim Aljarah, Majdi Mafarja, Ali Asghar Heidari, Hossam Faris, and Seyedali Mirjalili. 2019. Clustering analysis using a novel locality-informed grey wolf-inspired clustering approach. *Knowledge and Information Systems* (2019), 1–33.
- [9] Puspitaningtyas, Z. (2017). Estimating systematic risk for the best investment decisions on manufacturing company in Indonesia. *Investment management and financial innovations*, (14, № 1), 46-54.
- [10] Härdle, W. K., Prastyo, D., & Hafner, C. (2012). Support vector machines with evolutionary feature selection for default prediction.
- [11] J.H. Holland, *Adaptation in natural and artificial systems*, University of Michigan Press, Michigan, 1975.
- [12] I. Boussaïd, J. Lepagnot, P. Siarry, A survey on optimization metaheuristics, *Information Sciences* 237 (2013) 82-117.
- [13] R. Storn, K. Price, Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces, *Journal of global optimization* 11(4) (1997) 341-359.
- [14] M. Dorigo, T. Stutzle, *Ant Colony Optimization*, MIT Press, 2004.
- [15] M. Dorigo, L.M. Gambardella, Ant colony system: a cooperative learning approach to the traveling salesman problem, *IEEE Trans. Evol. Comput.* 1 (1) (1997) 53–66.

- [16] J. Kennedy, R.C. Eberhart, Particle swarm optimization, in: IEEE International Conference on Neural Networks, 1995, pp. 1942–1948.
- [17] Telikani, A., Gandomi, A. H., & Shahbahrani, A. (2020). A survey of evolutionary computation for association rule mining. *Information Sciences*, 524, 318-352.
- [18] B. Yuce, M. Packianather, E. Mastrocinque, D. Pham, A. Lambiase, Honey bees inspired optimization method: the bees algorithm, *Insects* 4 (4) (2013) 1577 646–662.
- [19] Hu Y-C & Ansell J. (2007). Measuring retail company performance using credit scoring techniques. *European Journal of Operational Research* 183(3):1595-1606.
- [20] Ince H & Aktan B. (2009). A Comparison of Data Mining Techniques for Credit Scoring In Banking: A Managerial Perspective. *Journal of Business Economics & Management* 10(3):233-240.
- [21] J.P. Finlay, X.X. Su, Y.D. Gu, I.D. Jenkinson, X.J. Ren, Inverse analysis of material parameters of multiple foam layers based on artificial neural network, *Adv. Mater. Res.* 189–193 (2011) 3313–3316, <http://dx.doi.org/10.4028/www.scientific.net/amr.189-193.3313>.
- [22] N. Karakaya, F. Evrendilek, K. Gungor, D. Onal, Predicting diel, diurnal and nocturnal dynamics of dissolved oxygen and chlorophyll- a using regression models and neural networks, *Acta Hydrochem. Hydrobiol.* 41 (9) (2013) 872–877, <http://dx.doi.org/10.1002/clen.201200683>.
- [23] Baklouti I. (2014). A Psychological Approach To Microfinance Credit Scoring Via A Classification And Regression Tree. *Intelligent Systems in Accounting, Finance & Management* 21(4):193-208.
- [24] J.F. Cao, Application of logistic regression model in financial investment risk assessment, *J. Netw. Syst. Manage.* 254 (09) (2018) 24–26, <http://dx.doi.org/10.2991/jrarc.2018.8.3.5>.
- [25] D.J. Lv, Theoretical analysis and empirical research on risk measurement of investment funds based on garch-cvar model, *Mod. Econ. Inf.* 18 (2018) 299–300, <http://dx.doi.org/10.2495/mesi140801>.
- [26] Huang X, Liu X & Ren Y. (2018). Enterprise credit risk evaluation based on neural network algorithm. *Cognitive Systems Research* 52:317-324.
- [27] J. Zhu, M. Khurshid, Z. Barsoum, Assessment of computational weld mechanics concepts for estimation of residual stresses in welded box structures, *Procedia Struct. Integr.* 17 (2019) 704–711, <http://dx.doi.org/10.1016/j.prostr.2019.08.094>.
- [28] Zhang, J. (2020). Investment risk model based on intelligent fuzzy neural network and VAR. *Journal of Computational and Applied Mathematics*, 371, 112707.
- [29] Antoncic, M. (2020). Uncovering hidden signals for sustainable investing using Big Data: Artificial intelligence, machine learning and natural language processing. *Journal of Risk Management in Financial Institutions*, 13(2), 106-113.
- [30] Baghdasaryan, V., Davtyan, H., Grigoryan, A., & Khachatryan, K. (2021). Comparison of econometric and deep learning approaches for credit default classification. *Strategic Change*, 30(3), 257-268.
- [31] Kerkez N. (2020). Artificial Intelligence and Machine Learning Can Repurpose Humans, Not Replace Them. American Bankers Association. ABA Banking Journal 112(6):30-32.



- [32] Gai L & Ielasi F. (2017). Credit quality and guarantees: How to interpret the Central Credit Register for a better access to credit. *Strategic Change* 26(4):291-299.
- [33] B.S. Zhang, T. Jiang, X.H. Zhou, et al., A new approach to portfolio optimization: mean Covar model, *Bmc Med. Inform. Decis.* 11, 38. <http://dx.doi.org/10.2495/mesi140801>.
- [34] Maknickienė N & Maknickas A. (2013). Financial Market Prediction System with Evolino Neural Network and Delphi Method. *Journal of Business Economics & Management* 14(2):403-413.
- [35] Vella V & Ng WL. (2015). A Dynamic Fuzzy Money Management Approach for Controlling the Intraday Risk-Adjusted Performance of AI Trading Algorithms. *Intelligent Systems in Accounting, Finance & Management* 22(2):153-178.
- [36] Karaboga, D., & Kaya, E. (2019). Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey. *Artificial Intelligence Review*, 52(4), 2263-2293.
- [37] Zhang, J., & Sanderson, A. C. (2009). JADE: adaptive differential evolution with optional external archive. *IEEE Transactions on evolutionary computation*, 13(5), 945-958.
- [38] Tang, L. (2018). An actor-critic-based portfolio investment method inspired by benefit-risk optimization. *Journal of Algorithms & Computational Technology*, 12(4), 351-360.
- [39] Liang, Z., Chen, H., Zhu, J., Jiang, K., & Li, Y. (2018-). Adversarial deep reinforcement learning in portfolio management. *arXiv preprint arXiv:1808.09940*.

## Abstract

Improving intelligent fuzzy neural network using evolutionary algorithms for investment risk model

Mahdi Nazari

In the financial world, risk management is the process of identification, analysis, and acceptance or mitigation of uncertainty in investment decisions. Essentially, risk management occurs when an investor or fund manager analyzes and attempts to quantify the potential for losses in an investment, such as a moral hazard, and then takes the appropriate action (or inaction) given the fund's investment objectives and risk tolerance. Risk management occurs everywhere in the realm of finance. In this research, an intelligent investment risk forecasting model is developed using artificial intelligence techniques that is a combination of Fuzzy Neural Networks (FNN) and evolutionary approaches. In the proposed method, the Adaptive Differential Evolution (ADE) algorithm is used as an optimizer for the FNN algorithm. The parameters of fuzzification and defuzzification layers are adjusted automatically by JADE. The proposed model is compared with the artificial neural network (ANN) and FNN models using a standard dataset. Three evaluation criteria of RMSE, MAE and  $R^2$  are used for these comparisons.

**Keywords:** Fuzzy neural network, Financing platform, Risk measurement, Evolutionary algorithms



Kooshyar Institute Of Higher Education

Department Of Computer

Improving intelligent fuzzy neural network using evolutionary algorithms for  
investment risk model

Mahdi Nazari

Supervisor

Dr. Soodabeh Arabani

Advisor

Dr. Ali Akbar Najafi

Sep 2021



Ministry Of Science, Research And Technology  
Kooshyar Institute Of Higher Education  
Masters Dissertation

Improving intelligent fuzzy neural network using evolutionary algorithms for  
investment risk model

Mahdi Nazari

Supervisor

Dr. Soodabeh Arabani

Advisor

Dr. Ali Akbar Najafi

Sep 2021