



دانشگاه تهران  
دانشکده علوم و فنون نوین

## استفاده از یادگیری عمیق به منظور پیش‌بینی در بازارهای سرمایه

پایان‌نامه برای دریافت درجه کارشناسی ارشد در رشته علوم تصمیم و مهندسی دانش

احسان حسین‌زاده

استاد راهنما

دکتر سامان هراتی‌زاده

تابستان ۱۳۹۸



الحمد لله رب العالمين





دانشگاه تهران  
دانشکده علوم و فنون نوین

## استفاده از یادگیری عمیق به منظور پیش‌بینی در بازارهای سرمایه

پایان‌نامه برای دریافت درجه کارشناسی ارشد در رشته علوم تصمیم و مهندسی دانش

احسان حسین‌زاده

استاد راهنما

دکتر سامان هراتی‌زاده

تابستان ۱۳۹۸



دانشگاه تهران

دانشکده علوم و فنون نوین

### گواهی دفاع از پایان نامه کارشناسی ارشد

هیأت داوران پایان نامه کارشناسی ارشد آقای / خانم احسان حسین زاده به شماره دانشجویی ۸۳۰۵۹۵۰۰۶ در رشته علوم تصمیم و مهندسی دانش را در تاریخ ..... با عنوان « استفاده از یادگیری عمیق به منظور پیش‌بینی در بازارهای سرمایه »

به حروف	به عدد	با نمره نهایی
		ارزیابی کرد.

و درجه

ردیف داوران	مشخصات هیئت	نام و نام خانوادگی	مرتبه دانشگاهی	دانشگاه یا مؤسسه	امضـ
۱	استاد راهنمای اول	دکتر سامان هراتی زاده	استادیار	دانشگاه تهران	
۲	استاد داور داخلی	دکتر هادی ویسی	استادیار	دانشگاه تهران	
۳	استاد مدعو	دکتر باقر بابا علی	استادیار	دانشگاه تهران	
۴	نماینده تحصیلات تکمیلی دانشکده			دانشگاه تهران	

نام و نام خانوادگی معاون تحصیلات تکمیلی و نام و نام خانوادگی معاون آموزشی و تحصیلات

پژوهشی دانشکده / گروه: تکمیلی پرdis دانشکده‌های فنی:

تاریخ و امضاء: تاریخ و امضاء:

## تعهدنامه اصالت اثر

باسم‌هه تعالی

این جانب احسان حسین‌زاده تائید می‌کنم که مطالب مندرج در این پایان‌نامه حاصل کار پژوهشی این جانب است و به دستاوردهای پژوهشی دیگران که در این نوشه از آنها استفاده شده است مطابق مقررات ارجاع گردیده است.  
این پایان‌نامه قبل‌اً برای احراز هیچ مدرک هم‌سطح یا بالاتر ارائه نشده است.

نام و نام خانوادگی دانشجو: احسان حسین‌زاده

تاریخ و امضای دانشجو:

کلیه حقوق مادی و معنوی این اثر  
متعلق به دانشگاه تهران می باشد.

تقدیم به:

پدر، مادر و برادر مهربانم که پشتیبان همیشگی من بودند.

## قدردانی

از زحمات استاد گران قدرم دکتر سامان هراتیزاده صمیمانه سپاسگزارم که کمک‌ها و راهنمایی‌های او راهگشای من در حل مسائل و گذر از بنیست‌ها شد. آموخته‌های من در این دوره از استاد بزرگوارم فراتر از درس و بحث علمی است و انسانیت، صبوری و محبت ایشان کمک بزرگی به من برای طی کردن این مسیر دشوار بود.

احسان حسین‌زاده

تابستان ۱۳۹۸

## چکیده

پیش‌بینی بازارهای مالی همواره یکی از موضوعات جذاب برای دانشمندان در حوزه‌های مختلف علم از جمله آماردانان، متخصصین مالی، ریاضیدانان و افراد فعال در حوزه علوم کامپیوتر بوده است. هدف این پایان‌نامه یافتن چارچوب‌هایی برای پیش‌بینی جهت حرکت بازارهای مالی مختلف به کمک الگوریتم‌های مبتنی بر یادگیری عمیق است. در میان این دسته از الگوریتم‌ها، شبکه‌های پیچشی جدیدتر از سایر الگوریتم‌ها هستند. در این پایان‌نامه سه چارچوب بر پایه شبکه‌های پیچشی توسعه داده شده است. اولین چارچوب که CNNpred نام دارد، می‌تواند از طیف وسیع و غنی‌ای از ویژگی‌های مالی نظیر تکنیکال اندیکاتورها، نرخ تبدیل ارز، قراردادهای آتی، وضعیت شاخص‌های مهم جهان و غیره برای پیش‌بینی بازارهای مختلف مالی استفاده کند. برخورداری از طراحی منحصر به فرد و متناسب با ویژگی‌های مالی استفاده شده و توجه به تاثیرات بازارهای مختلف بر یکدیگر از دیگر نقاط قوت این چارچوب است. از این چارچوب برای پیش‌بینی ۵ شاخص اصلی بازار آمریکا استفاده شده است که نتایج حاصل برتری، معنادار آماری را نسبت به سایر الگوریتم‌های پایه نشان می‌دهد. در چارچوب دوم که HT-CNNpred نام دارد، به بهبود کیفیت شبکه‌های پیچشی از طریق انتخاب هوشمندانه تر وزن‌های ابتدایی این شبکه‌ها با استفاده از داده‌های برچسب‌دار، پرداخته‌ایم. ارائه روشی سلسه‌مراتبی برای آموزش شبکه‌های پیچشی از مزایای این چارچوب است که باعث بهبود کیفیت در پیش‌بینی ۴۵۸ شرکت حاضر در بازار بورس آمریکا و غلبه آماری بر الگوریتم‌های پایه شده است. استفاده از یادگیری انتقالی به HT-CNNpred امکان پیش‌بینی بازارهای جدید را می‌بخشد و پیش‌بینی موفق ۱۴ شاخص معروف در سرتاسر جهان، گواهی بر موفقیت این چارچوب است. در آخرین چارچوب که C-CNNpred نامیده شده است، استفاده از خوشه‌بندی برای قرار دادن بازارهای دارای رفتار مشابه و انتخاب نماینده هر خوش‌بینی نهایی در دستور کار قرار دارد. سپس، اطلاعات نماینده هر خوش‌بینی نهایی در کنار بازار اصلی که قصد پیش‌بینی آن را داریم قرار می‌گیرد تا به بهبود پیش‌بینی کمک کند. استفاده از این چارچوب در پیش‌بینی شاخص کل بازار بورس تهران، منجر به برتری آماری بر سایر الگوریتم‌های پایه شده است.

**واژگان کلیدی:** پیش‌بینی بازارهای مالی، یادگیری عمیق، شبکه‌های پیچشی، یادگیری انتقالی، خوشه‌بندی

# فهرست مطالب

ج	فهرست شکل‌ها
خ	فهرست جداول‌ها
د	فهرست الگوریتم‌ها
۱	فصل ۱: مقدمه
۱	۱-۱ پیش‌بینی بازار و اهمیت آن
۲	۲-۱ تاریخچه‌ی مختصری از تحقیق
۳	۳-۱ اهداف تحقیق
۴	۴-۱ نوآوری تحقیق
۴	۴-۵ ساختار پایان‌نامه
۵	فصل ۲: مروری بر مفاهیم پایه
۵	۱-۲ مقدمه
۶	۲-۲ شبکه‌های عصبی
۶	۱-۲-۲ شبکه پیشرو
۷	۲-۲-۲ شبکه عصبی پیچشی
۸	۱-۲-۲-۲ لایه پیچشی
۹	۲-۲-۲-۲ لایه ادغام
۹	۳-۲-۲-۲ لایه اتصال کامل

۱۰ . . . . .	آموزش شبکه های پیچشی . . . . .	۳-۲-۲
۱۱ . . . . .	انتشار به جلو . . . . .	۱-۳-۲-۲
۱۱ . . . . .	انتشار به عقب . . . . .	۲-۳-۲-۲
۱۲ . . . . .	بروزرسانی وزن ها . . . . .	۳-۳-۲-۲
۱۲ . . . . .	شبکه های بازگشتی . . . . .	۴-۲-۲
۱۳ . . . . .	فراموشی در دنباله های طولانی . . . . .	۱-۴-۲-۲
۱۴ . . . . .	شبکه حافظه کوتاه-مدت ماندگار . . . . .	۲-۴-۲-۲
۱۴ . . . . .	بلوک حافظه کوتاه-مدت ماندگار . . . . .	۳-۴-۲-۲
۱۷ . . . . .	حذف تصادفی . . . . .	۵-۲-۲
۱۷ . . . . .	یادگیری انتقالی . . . . .	۶-۲-۲
۱۷ . . . . .	الگوریتم ژنتیک . . . . .	۳-۲
۱۹ . . . . .	مدل سازی مسئله . . . . .	۱-۳-۲
۱۹ . . . . .	تشکیل جمعیت اولیه . . . . .	۲-۳-۲
۱۹ . . . . .	ارزیابی جمعیت . . . . .	۳-۳-۲
۱۹ . . . . .	انتخاب والدین . . . . .	۴-۳-۲
۲۰ . . . . .	بازترکیب . . . . .	۵-۳-۲
۲۰ . . . . .	جهش . . . . .	۶-۳-۲
۲۰ . . . . .	انتخاب فرزندان . . . . .	۷-۳-۲
۲۱ . . . . .	شرط خاتمه الگوریتم . . . . .	۸-۳-۲
۲۱ . . . . .	کاهش بعد و استخراج ویژگی . . . . .	۴-۲
۲۲ . . . . .	تحلیل مؤلفه های اساسی . . . . .	۱-۴-۲
۲۴ . . . . .	خودرمندگار . . . . .	۲-۴-۲
۲۶ . . . . .	ماشین بولتزمن محدود . . . . .	۳-۴-۲
۲۸ . . . . .	خوشبندی . . . . .	۵-۲
۲۹ . . . . .	خوشبندی k میانگین . . . . .	۱-۵-۲
۲۹ . . . . .	جمع بندی . . . . .	۶-۲

۳۱	<b>فصل ۳: مروری بر ادبیات تحقیق</b>
۳۱	۱-۳ مقدمه . . . . .
۳۱	۲-۳ شبکه‌های کم‌عمق . . . . .
۳۲	۱-۲-۳ شبکه پرسپترون چندلایه . . . . .
۳۳	۲-۲-۳ ترکیب شبکه عصبی و الگوریتم ژنتیک . . . . .
۳۸	۳-۲-۳ ترکیب شبکه عصبی و تحلیل مؤلفه‌های اساسی . . . . .
۴۰	۳-۳ شبکه‌های عمیق . . . . .
۴۰	۱-۳-۳ شبکه پرسپترون چندلایه عمیق . . . . .
۴۶	۲-۳-۳ یادگیری عمیق و الگوریتم‌های نمایش داده . . . . .
۴۹	۳-۳-۳ شبکه‌های بازگشتی . . . . .
۵۳	۴-۳-۳ شبکه‌های پیچشی . . . . .
۵۵	۴-۳ جمع بندی کارهای انجام شده . . . . .
۵۷	<b>فصل ۴: چارچوب‌های پیشنهادی</b>
۵۷	۱-۴ مقدمه . . . . .
۵۸	۲-۴ چارچوب CNNpred . . . . .
۶۰	۱-۲-۴ 2D-CNNpred . . . . .
۶۳	۲-۲-۴ 3D-CNNpred . . . . .
۶۵	۳-۲-۴ ویژگی‌های استفاده شده برای هر شاخص . . . . .
۶۷	۳-۴ چارچوب HT-CNNpred . . . . .
۷۰	۱-۳-۴ شبکه پیچشی ابتدایی . . . . .
۷۰	۲-۳-۴ نمایش داده‌ها . . . . .
۷۰	۳-۳-۴ آموزش سلسله‌مراتبی . . . . .
۷۲	۴-۳-۴ یادگیری انتقالی . . . . .
۷۲	۴-۴ چارچوب C-CNNpred . . . . .
۷۲	۱-۴-۴ خوشبندی بازارها . . . . .

۷۳ . . . . .	نمايش دادهها ۲-۴-۴
۷۳ . . . . .	استخراج ویژگی های روزانه ۳-۴-۴
۷۳ . . . . .	استخراج ویژگی های بلند مدت ۴-۴-۴
۷۴ . . . . .	پيش‌بياني نهايی ۵-۴-۴
۷۴ . . . . .	جمع‌بندی ۵-۴
۷۵ . . . . .	<b>فصل ۵: نتایج، بحث و ارزیابی</b>
۷۵ . . . . .	۱-۵ مقدمه
۷۵ . . . . .	۲-۵ جمع‌آوري و پيش‌پردازش دادهها
۷۶ . . . . .	CNNpred ۱-۰-۲-۵
۷۶ . . . . .	HT-CNNpred ۲-۰-۲-۵
۷۷ . . . . .	C-CNNpred ۳-۰-۲-۵
۷۸ . . . . .	۱-۲-۵ نرمال‌سازی دادهها
۷۸ . . . . .	۳-۵ معیار ارزیابی
۷۹ . . . . .	۴-۵ پارامترهای شبکه
۸۰ . . . . .	۵-۵ الگوريتم‌های پایه
۸۱ . . . . .	۶-۵ نتایج
۸۱ . . . . .	CNNpred ۱-۶-۵
۸۲ . . . . .	HT-CNNpred ۲-۶-۵
۸۶ . . . . .	C-CNNpred ۳-۶-۵
۸۸ . . . . .	۷-۵ جمع‌بندی
۹۰ . . . . .	<b>فصل ۶: جمع‌بندی، نتیجه‌گیری و پيشنهاهاي برای تحقیقات آينده</b>
۹۱ . . . . .	۱-۶ پيشنهاهاي برای پژوهش‌های آينده
۹۲ . . . . .	<b>مراجع</b>
۱۰۳ . . . . .	<b>پيوست آ: شناخت دادهها</b>

اول

واژه نامه فارسی به انگلیسی

دوم

واژه نامه انگلیسی به فارسی

چهارم

نمایه

# فهرست شکل‌ها

۱-۲	شبکه عصبی پیشرو	۷
۲-۲	شبکه عصبی پیچشی [۱]	۷
۳-۲	لایه پیچشی [۲]	۹
۴-۲	لایه ادغام [۳]	۱۰
۵-۲	ساختار شبکه بازگشته [۴]	۱۳
۶-۲	فراموشی در دنباله‌های طولانی.	۱۴
۷-۲	ساختار شبکه عصبی حافظه کوتاه-مدت ماندگار	۱۵
۸-۲	شمای کلی یک بلوک حافظه کوتاه-مدت ماندگار	۱۶
۹-۲	مراحل مختلف الگوریتم ژنتیک [۵]	۱۸
۱۰-۲	ساختار یک خودرمزنگار [۶]	۲۵
۱۱-۲	ساختار یک ماشین بولتزمن محدود	۲۶
۱-۳	میزان دقیق شبکه عصبی در پیش‌بینی سال‌های مختلف [۷]	۲۳
۲-۳	کاربرد الگوریتم ژنتیک در شبکه عصبی [۸]	۲۴
۳-۳	مقایسه نتایج [۸]	۲۵
۴-۳	پارامترهای بررسی شده [۹]	۲۵
۵-۳	مقایسه نتایج [۹]	۳۶
۶-۳	نحوه ترکیب شبکه عصبی و الگوریتم ژنتیک [۱۰]	۳۷
۷-۳	پارامترهای استفاده شده [۱۰]	۳۸
۸-۳	مقایسه نتایج [۱۰]	۳۹

۹-۳ میزان واریانس داده‌ها با درنظر گرفتن مؤلفه‌های مختلف [۱۱]	۳۹
۱۰-۳ میزان دقت پیش‌بینی با درنظر گرفتن مؤلفه‌های مختلف [۱۱]	۴۰
۱۱-۳ نتایج پیش‌بینی با استفاده از داده‌های ۴ روز قبل [۱۲]	۴۲
۱۲-۳ نتایج پیش‌بینی با استفاده از داده‌های ۹ روز قبل [۱۲]	۴۲
۱۳-۳ نحوه تقسیم داده‌ها به داده آموزشی و داده آزمون	۴۴
۱۴-۳ نتایج حاصل با درنظر گرفتن پارامترهای مختلف [۱۳]	۴۴
۱۵-۳ مقایسه قیمت پیش‌بینی شده و قیمت واقعی در روز $n$ با استفاده از داده‌های تاریخ $1 - n$	۴۵
۱۶-۳ مراحلی طی شده برای پیش‌بینی با استفاده از شبکه حافظه کوتاه-مدت ماندگار [۱۵]	۴۹
۱۷-۳ شرکت‌های انتخاب شده و رقبای آنها [۱۵]	۵۰
۱۸-۳ بررسی تعداد داده‌های مختلف در یک دنباله [۱۵]	۵۱
۱۹-۳ نتایج نهایی حاصل از استفاده شبکه حافظه کوتاه-مدت ماندگار در پیش‌بینی ۶ شرکت [۱۵]	۵۱
۲۰-۳ نتایج نهایی حاصل از استفاده شبکه حافظه کوتاه-مدت ماندگار در مقایسه با سایر الگوریتم‌ها	۵۲
۲۱-۳ نمای گرافیکی از شبکه پیچشی استفاده شده [۱۷]	۵۴
۲۲-۳ نمای گرافیکی از شبکه پیچشی استفاده شده [۱۸]	۵۵
۱-۴ نمای گرافیکی از نحوه نمایش و استفاده از داده در CNNpred	۵۹
۲-۴ نمای گرافیکی از ماتریس ورودی در 2D-CNNpred	۶۱
۳-۴ اعمال فیلتری با ابعاد $\times 1$ تعداد ویژگی‌های روزانه، بر ماتریس ورودی	۶۲
۴-۴ نمای گرافیکی از 2D-CNNpred	۶۳
۵-۴ نمای گرافیکی از تنسور ورودی در 3D-CNNpred	۶۴
۶-۴ اعمال فیلتری $1 \times 1$ بر تنسور ورودی	۶۵
۷-۴ نمای گرافیکی از 3D-CNNpred	۶۶
۸-۵ مقایسه الگوریتم‌های مختلف در داشتن بهترین عملکرد از نظر F-measure	۸۴
۹-۵ مجموع فاصله درون خوشه‌ای نسبت به تعداد خوشه‌ها	۸۷

### ۳-۵ بهترین نتایج الگوریتم‌های مختلف . . . . .

۸۸

# فهرست جدول‌ها

۱-۳ مقایسه خطای قیمت پیش‌بینی شده [۱۴]	۴۶
۲-۳ میانگین دقت پیش‌بینی ۳۸ شرکت با استفاده از الگوریتم‌های مختلف [۱۹]	۴۷
۳-۳ میانگین مجدد میانگین مربعات خطای ۳۸ شرکت [۱۹]	۴۸
۴-۳ میانگین میانگین مطلق خطای ۳۸ شرکت [۱۹]	۴۸
۵-۳ مقایسه مقالات توضیح‌داده شده	۵۶
۱-۵ توضیحات نمادها	۷۷
۲-۵ ویژگی‌های استفاده شده برای نمایش شاخص‌های بازار تهران	۷۸
۳-۵ میانگین F-measure الگوریتم‌های مختلف	۸۱
۴-۵ بهترین F-measure الگوریتم‌های مختلف	۸۱
۵-۵ میانیگن F-measure الگوریتم‌های مختلف در پیش‌بینی ۴۵۸ شرکت	۸۳
۶-۵ عملکرد زیرشبکه‌های تشکیل شده در پیش‌بینی وضعیت شرکت‌ها	۸۴
۷-۵ F-measure الگوریتم‌های مختلف در پیش‌بینی شاخص‌های آمریکایی	۸۵
۸-۵ F-measure الگوریتم‌های مختلف در پیش‌بینی شاخص‌های مهم جهان	۸۵
۹-۵ میانگین F-measure الگوریتم‌های مختلف	۸۷
۱۰-۱ توضیحات ویژگی‌های استفاده شده	۱۰۵
۱۰-۲ لیست شاخص‌های حاضر در بازار ایران	۱۰۵

# فهرست الگوریتم‌ها

۶۹	.....	HT-CNN-pred	۱-۴
۷۱	.....	Training Hierarchical	۲-۴

# فصل ۱

## مقدمه

### ۱-۱ پیش‌بینی بازار و اهمیت آن

پیش‌بینی بازارهای مالی همواره یکی از موضوعات جذاب برای دانشمندان در حوزه‌های مختلف علم از جمله آماردانان، متخصصین مالی، ریاضیدانان و افراد فعال در حوزه علوم کامپیوتر بوده است. این بازارها تأثیر زیادی از تحولات سیاسی و اقتصادی می‌گیرند و بدلیل اینکه مدل کردن بازار با درنظر گرفتن تمامی جوانب کاری بسیار دشوار است، عمدتاً با نویز زیادی برخورد می‌کنیم. بعضًا این نظر دیده شده است که بخارطه موارد گفته شده پیش‌بینی این بازارها امکان‌پذیر نیست، اما نتایج بدست آمده و سرمایه‌گذاری‌های انجام شده توسط کارگزاری‌ها و سایر نهادهای مالی‌ای که در این امر دخیل هستند، حکایت از این دارد که این کار شدنی و امکان‌پذیر است. از دیگر عواملی که باعث علاقه افراد به این حوزه می‌شود، می‌توان به سوددهی بالا در صورت گرفتن نتیجه مطلوب اشاره کرد. علاوه بر این، بهبود تدریجی الگوریتم‌ها و دستاوردهای افراد نیز نقش مهمی ایفا می‌کنند. این بهبود تدریجی را بوضوح می‌توان در نرم‌افزارهایی که برای بازی شطرنج طراحی شده‌اند مشاهده کرد که چطور از باخت در ابتدایی ترین مراحل در سال‌های گذشته، به برد قهرمانان جهان در حال حاضر رسیده‌اند. بنابراین می‌توان انتظار داشت که در حوزه پیش‌بینی بازارهای مالی نیز شاهد رشدی تدریجی باشیم.

هر چند بخش اعظم فعالیت‌های انجام شده مربوط به بازار بورس است، اما پیش‌بینی بازارهای مالی تنها به

بورس محدود نمی‌شود و شامل پیش‌بینی قیمت بیت‌کوین<sup>۱</sup>، نفت، طلا و فارکس<sup>۲</sup> نیز می‌شود [۲۰، ۲۱، ۲۲، ۲۳].

باید یادآور شویم که هدف از این پایان‌نامه پیش‌بینی جهت حرکت بازارهای مالی است و این مسئله ذاتاً یک مسئله دسته‌بندی است. اهمیت پیش‌بینی جهت در این است که از نتایج حاصل از پیش‌بینی می‌توان در یک سیستم خرید و فروش استفاده کرد. هرچه دقیق‌تر پیش‌بینی انجام شده در مورد وضعیت آینده بازارهای مالی بیشتر باشد، می‌توان انتظار سود بیشتری داشت. یائو و دیگران در مورد اهمیت پیش‌بینی تشخیص جهت حرکت بازار صحبت کرده‌اند [۲۴]. تحقیقات بسیاری در حوزه پیش‌بینی جهت حرکت بازار تا کنون انجام گرفته است [۲۵، ۲۶، ۲۷، ۲۸، ۲۹، ۷، ۲۵].

## ۲-۱ تاریخچه‌ی مختصری از تحقیق

از گذشته تا به امروز روش‌های متعدد و متنوعی برای پیش‌بینی آینده بازار ارائه شده است. تحلیل این بازارها عمده‌تاً به دو صورت تکنیکال<sup>۳</sup> و بنیادی<sup>۴</sup> انجام می‌پذیرد. در تحلیل تکنیکال، از طریق پیدا کردن الگوهای معنادار در داده‌های گذشته، پیش‌بینی رفتار آینده بورس انجام می‌پذیرد. تحلیل بنیادی نیز بر پایه ویژگی‌های درونی یک شرکت استوار است. ظهور تکنیک‌های داده‌کاوی، محققین و سرمایه‌گذاران را بر آن داشته است تا نهایت استفاده را از این تکنیک‌ها در مسئله پیش‌بینی بورس ببرند. شبکه‌های عصبی<sup>۵</sup>، ماشین بردار پشتیبان<sup>۶</sup> از جمله الگوریتم‌های پرکاربرد در این حوزه هستند [۳۱، ۳۲، ۳۳، ۳۴]. الگوریتم‌های تکاملی و بخصوص الگوریتم ژنتیک<sup>۷</sup>، از جمله مواردی هستند که بعضاً به صورت مکمل شبکه‌های عصبی و بعضاً به صورت مستقل در پیش‌بینی بازار بورس استفاده شده‌اند [۳۵، ۳۶، ۳۷، ۳۸، ۳۹].

نسل جدید از الگوریتم‌های مبتنی بر شبکه‌های عصبی که یادگیری عمیق<sup>۸</sup> نامیده می‌شوند، در سال‌های گذشته و با افزایش داده‌ها و همین‌طور توان پردازشی بسیار محبوب شده‌اند [۴۰]. این الگوریتم‌ها در واقع همان

<sup>1</sup>Bitcoin

<sup>2</sup>Forex

<sup>3</sup>Technical

<sup>4</sup>Fundamental

<sup>5</sup>Neural networks

<sup>6</sup>Support vector machine

<sup>7</sup>Genetic algorithm

<sup>8</sup>Deep learning

شبکه‌های عصبی‌ای هستند که تعداد لایه‌های مخفی آن‌ها افزایش یافته است. از موهبت‌های یادگیری عمیق می‌توان به استخراج ویژگی‌های بسیار پیچیده‌تر نسبت به شبکه‌های عصبی کم‌عمق اشاره کرد. این برتری باعث می‌شود تا استخراج ویژگی‌های مناسب که در گذشته بر عهده کارشناسان بوده است، تا حد زیادی به شبکه‌های عمیق واگذار شود و فرآیند مهندسی ویژگی‌ها نقش کمزنگ‌تری در کارایی الگوریتم‌ها ایفا کند. اگرچه یادگیری عمیق نتایج درخشانی در سایر حوزه‌ها نظیر پردازش تصویر<sup>۹</sup>، بینایی ماشین<sup>۱۰</sup> و پردازش زبان طبیعی<sup>۱۱</sup> از خود نشان داده است [۴۱، ۴۲]، اما کاربرد آن در پیش‌بینی بازارهای مالی محدود به حدود ۳ سال اخیر است. از جمله این الگوریتم‌ها می‌توان به شبکه پرسپترون چندلایه<sup>۱۲</sup> عمیق، شبکه‌های بازگشتی<sup>۱۳</sup>، شبکه‌های پیچشی<sup>۱۴</sup>، ماشین بولتزمن محدود<sup>۱۵</sup> و خودرمزگذارها<sup>۱۶</sup> اشاره کرد [۴۳، ۴۴، ۴۵، ۱۳]. حال اگر به مقالات ارائه شده در مورد بازار تهران بنگریم، متأسفانه پژوهشی در این حوزه انجام نشده است و نیاز به آن کاملاً حس می‌شود.

## ۱-۳ اهداف تحقیق

اهداف ما از انجام این تحقیق ساخت مدل یا مدل‌هایی برای پیش‌بینی بازارهای مالی است که از نظر دقیقت نسبت به سایر مدل‌ها بهتر باشند. اهداف این تحقیق به شرح ذیل می‌باشند:

۱. بررسی تعاملات بازارها بر یکدیگر و استفاده از این تعاملات برای پیش‌بینی بازارهای مالی.
۲. ارائه چارچوبی برای افزایش دقیقت پیش‌بینی بازارهای مالی
۳. بررسی میزان کارایی الگوریتم‌های مبتنی بر یادگیری عمیق در پیش‌بینی بازارهای مالی
۴. بررسی تاثیر نحوه آموزش شبکه بر کیفیت پیش‌بینی
۵. بررسی میزان شباهت رفتاری بازارهای مالی مختلف

<sup>9</sup>Image processing

<sup>10</sup>Computer vision

<sup>11</sup>Natural language processing

<sup>12</sup>Multi layer perceptron

<sup>13</sup>Recurrent neural networks

<sup>14</sup>Convolutional neural networks

<sup>15</sup>Restricted Boltzmann machine

<sup>16</sup>Autoencoders

## ۱-۴ نوآوری تحقیق

نوآوری‌های تحقیق ما به شرح زیر است:

۱. ارائه چارچوبی که برای اولین بار از یک تنسور سه-بعدی برای ترکیب بازارهای مختلف استفاده می‌کند.
۲. استفاده از طیف وسعی از ویژگی‌های مالی برای پیش‌بینی بازارهای مالی
۳. طراحی فیلترهای شبکه پیچشی بصورت اختصاصی برای کار با داده‌های مالی برای اولین بار
۴. پیشنهاد آموزش سلسله‌مراتبی برای آموزش شبکه‌های پیچشی
۵. استفاده از یادگیری انتقالی برای پیش‌بینی بازارهایی که در آموزش شبکه عصبی مورد استفاده قرار نگرفته‌اند
۶. استفاده از خوشبندی برای قرار دادن بازارهای مشابه در یک خوش و استفاده از نماینده آن خوش برای پیش‌بینی سایر بازارها

## ۱-۵ ساختار پایان‌نامه

در فصل بعدی، الگوریتم‌های استفاده شده در مقالات مطالعه شده و همینطور پژوهش گزارش شده در این پایان‌نامه مورد بررسی قرار می‌گیرد. این الگوریتم‌ها شامل شبکه عصبی پیشرو، شبکه بازگشتی، شبکه حافظه کوتاه-مدت مانندگار، شبکه خودرنگذار، ماشین بولتزمن محدود، الگوریتم خوشبندی k میانگین، الگوریتم رنتیک و الگوریتم تحلیل مؤلفه‌های اصلی است.

در فصل سوم گزارشی از کاربرد روش‌های مبتنی بر یادگیری ماشین و بخصوص شاخه شبکه‌های عصبی و یادگیری عمیق در پیش‌بینی و تحلیل بازارهای مالی بیان می‌شود.

فصل چهارم به گزارش راهکارهای ارائه شده توسط ما و توضیح کامل آن‌ها اختصاص پیدا کرده است.

فصل پنجم نتایج حاصل از الگوریتم‌های پیشنهادی و همچنین تحلیل آن‌ها را شامل می‌شود.

در فصل آخر جمع‌بندی و همچنین پیشنهاداتی برای ادامه پژوهش ذکر شده است.

پیوست آ ویژگی‌های داده‌های استفاده شده در این پژوهش را تشریح می‌کند.

پیوست ب نیز مقالات پژوهشی استخراج شده از این پایان‌نامه را معرفی می‌کند.

## فصل ۲

### مروری بر مفاهیم پایه

#### ۱-۲ مقدمه

این فصل به مرور مفاهیم و الگوریتم‌های استفاده شده در قسمت مرور ادبیات تحقیق و همچنین پژوهش انجام شده در این پایان‌نامه اختصاص دارد تا پیش از رسیدن به این بخش‌ها، مقدمات لازم برای درک کامل مطالب، فراهم شده باشد.

در ابتدا، مروری بر شبکه‌های عصبی و انواع مختلف آن از جمله شبکه‌های عصبی پیش‌رو، شبکه‌های پیچشی و همینطور شبکه‌های بازگشتی، حذف تصادفی<sup>۱</sup> و یادگیری انتقالی<sup>۲</sup> خواهیم داشت. سپس به معرفی الگوریتم رئیتیک می‌پردازیم که از آن می‌توان برای آموزش شبکه‌های عصبی استفاده کرد. پس از آن مفاهیم مربوط به کاهش بعد و استخراج ویژگی با مرور الگوریتم‌هایی نظیر ماشین بولتزمن محدود، خودر مزگذارها و همینطور الگوریتم تحلیل مؤلفه‌های اساسی<sup>۳</sup> مرور خواهد شد. در انتها نیز اشاره‌ای به خوشبندی<sup>۴</sup> خواهیم داشت.

---

<sup>1</sup>Dropout

<sup>2</sup>Transfer learning

<sup>3</sup>Principal component analysis

<sup>4</sup>Clustering

## ۲-۲ شبکه‌های عصبی

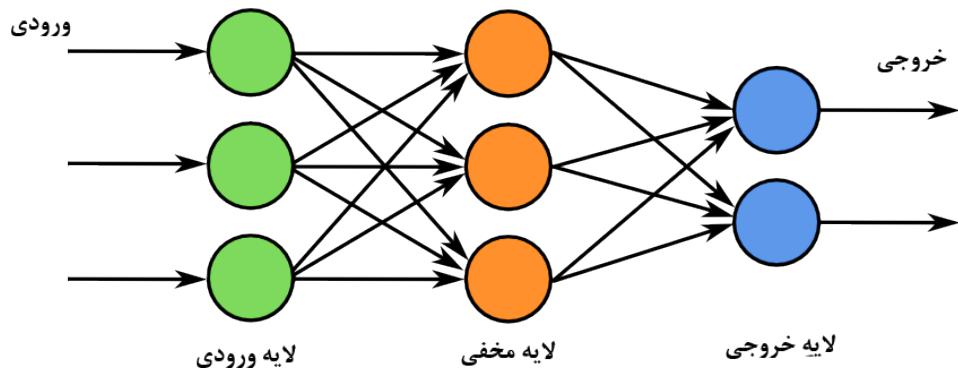
شبکه‌های عصبی مصنوعی بر اساس مدل‌سازی ریاضی از ساختار شبکه عصبی و مغز انسان طراحی شده‌اند. ساختار شبکه‌های عصبی بر پایه شبکه‌ای از واحدهای پردازنده کوچک و یال‌های وزن دار بین آن‌ها شکل گرفته است. پردازش اطلاعات در واحدهای پردازنده انجام می‌شود، سیگنال‌ها بین واحدهای پردازنده از طریق یال‌ها (بیوند) منتقل می‌شوند که هر یک از این یال‌ها با توجه به اهمیت سیگنال خروجی واحد پردازنده مبدأ، وزن مربوط به خود را دارند. هر یک از واحدهای پردازنده یک تابع فعال‌سازی<sup>۵</sup> دارد که ورودی این تابع جمع وزن دار مقادیر ورودی به این واحد و خروجی آن مقدار فعال‌سازی واحد مورد نظر است. یادگیری عمیق شاخه‌ای از یادگیری ماشین است که بر اساس مجموعه‌ای از الگوریتم‌ها، تلاش می‌کند تا مفاهیم انتزاعی سطح بالا را با استفاده از شبکه عمیق با چندین لایه خطی و غیر خطی مدل کند. به زبان ساده‌تر، هدف یادگیری عمیق، استخراج خودکار ویژگی‌ها به صورت هوشمندانه است که نگاهی نو به شبکه‌های عصبی است [۴۶، ۴۷]. در واقع شبکه عمیق، شبکه عصبی چند لایه‌ای است که با روی هم قرار گرفتن بیش از دو لایه مخفی تشکیل می‌شوند.

[۴۶، ۴۸]

## ۱-۲-۲ شبکه پیشرو

شبکه‌ایی که در آن‌ها یال‌ها تشکیل دور نمی‌دهند، شبکه‌های عصبی پیش‌رو نامیده می‌شوند. در شبکه عصبی پیشرو اطلاعات تنها در جهت رو به جلو از لایه ورودی به سمت لایه خروجی حرکت می‌کند. از زمان معرفی شبکه‌های عصبی مدل‌های متعدد با ویژگی‌های مختلفی از شبکه‌های عصبی معرفی شده و مورد استفاده قرار گرفته‌اند. شبکه پرسپترون در سال ۱۹۵۸ توسط فرانک روزنبلت ابداع شد [۴۹]. این شبکه یکی از ساده‌ترین انواع شبکه‌های عصبی پیش‌رو است که توانایی تفکیک خطی مسائل را دارد. از مدل‌های دیگر شبکه‌های عصبی پیش‌رو می‌توان به پرسپترون چندلایه (MLP) اشاره کرد. در شکل (۱-۲) می‌توان نمونه‌ای از ساختار شبکه عصبی پیشرو مشاهده کرد.

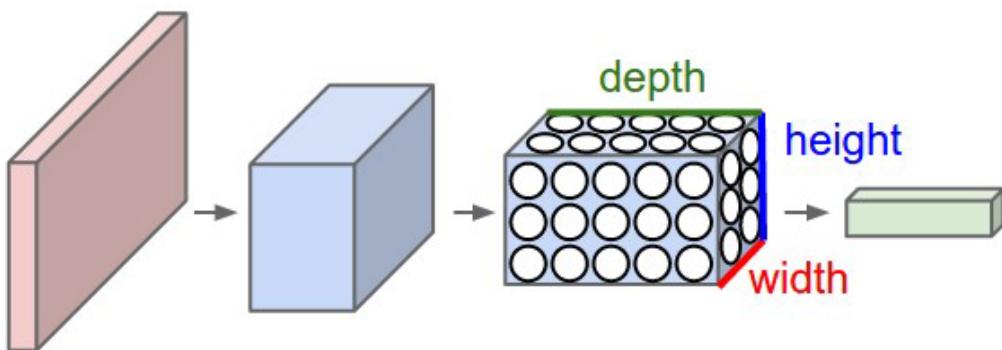
<sup>۵</sup>Activation function



شکل ۲: شبکه عصبی پیشرو

## ۲-۲-۲ شبکه عصبی پیچشی

این شبکه از سیستم بینایی انسان الهام گرفته شده است و برای اولین بار توسط لی کان<sup>۶</sup> و هکارانش مدل سازی شده است [۵۰]. شبکه های عصبی پیچشی<sup>۷</sup> یک دسته از شبکه های عصبی<sup>۸</sup> هستند که در ابتدا برای پردازش تصاویر ساخته شده اند و معمولاً<sup>۹</sup> از سه لایه پیچشی<sup>۹</sup>، لایه ادغام<sup>۱۰</sup> و لایه اتصال کامل<sup>۱۱</sup> تشکیل می شوند.



شکل ۲: شبکه عصبی پیچشی [۱]

<sup>6</sup>LeCun

<sup>7</sup>Convolution Neural Network(CNN)

<sup>8</sup>Neural Network

<sup>9</sup>Convolutional layer

<sup>10</sup>Pooling layer

<sup>11</sup>Fully connected layer

### ۱-۲-۲-۲ لایه پیچشی

اولین لایه‌ای که ورودی را دریافت می‌کند لایه پیچشی<sup>۱۲</sup> نام دارد. این لایه یک عمل پیچش را بر ورودی انجام می‌دهد. برای این کار، فیلتری با سایر معین تعیین می‌شود. اندازه این فیلتر نشان‌دهنده حوزه اعمال تغییر توسط آن فیلتر است. هر فیلتر با تعدادی وزن مشخص می‌شود و وزن‌های هر فیلتر برای تمامی ویژگی‌های ورودی به آن لایه یکسان است. سپس به منظور غیرخطی‌سازی خروجی این لایه، از یک تابع فعال‌سازی استفاده می‌شود. گزینه‌های متفاوتی برای تابع فعال‌سازی وجود دارد اما تابع یکسوساز<sup>۱۳</sup> به دلیل غیرخطی بودن و محاسبه سریع، گزینه مناسب و محبوبی در پژوهش‌های یادگیری عمیق است. این تابع را در رابطه ۱-۲ می‌توان یافت. در طول فرآیند یادگیری شبکه، وزن‌های هر فیلتر به نحوی تغییر می‌کنند تا ویژگی‌هایی را استخراج کنند که در پیدا کردن خروجی نهایی شبکه، تأثیر بیشتری داشته باشند. نحوه عملکرد این لایه در رابطه‌های ۲-۲، ۳-۲ نشان داده شده است که در آن اندازه فیلتر  $m \times m$ ، وزن‌های فیلتر با ماتریس  $w$ ، تابع فعال‌سازی با  $\delta$  و مقدار سطر  $\gamma$  و ستون  $\zeta$  از لایه  $l$  با  $v_{i,j}^l$  نمایش داده می‌شود. شکل ۳-۲ نیز نمای گرافیکی از نحوه عملکرد این لایه را نشان می‌دهد.

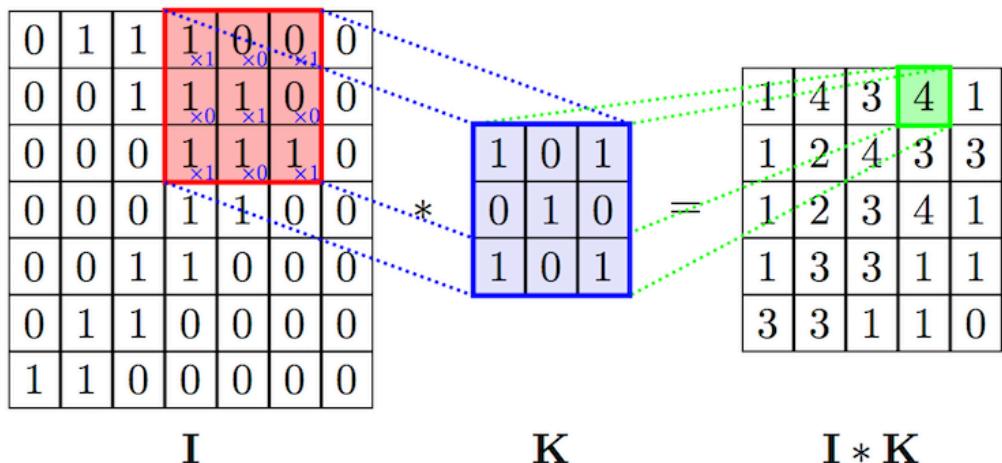
$$f(x) = \max(0, x) \quad (1-2)$$

$$x_{i,j}^l = \sum_{t=0}^{m-1} \sum_{y=0}^{m-1} w_{t,y} v_{i+t,j+y}^{l-1} \quad (2-2)$$

$$v_{i,j}^l = \sigma(x_{i,j}^l) \quad (3-2)$$

<sup>12</sup>Convolution Layer

<sup>13</sup>Relu



شکل ۲-۳: لایه پیچشی [۲]

### ۲-۲-۲-۲ لایه ادغام

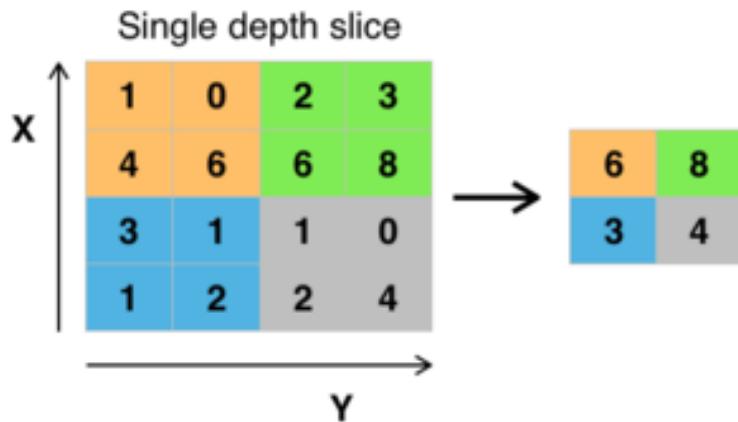
مسئولیت اصلی این لایه، ادغام ویژگی‌های استخراج شده در لایه پیچشی است. این فرآیند علاوه بر کاهش ویژگی‌ها که منجر به کاهش هزینه محاسباتی می‌شود، ابزاری مناسب برای حذف نویز و اطلاعات اضافی در شبکه‌های پیچشی است. از دیگر فواید این لایه می‌توان به کاهش تعداد پارامترهایی که باید توسط شبکه یاد گرفته شود و درنتیجه کاهش احتمال بیش‌برازش<sup>۱۴</sup> اشاره کرد. ادغام بیشینه<sup>۱۵</sup> معروف‌ترین نوع ادغام است که در آن ویژگی با بزرگ‌ترین مقدار در هر پنجمراه  $m \times m$  انتخاب می‌شود. شکل ۴-۲ نشان دهنده نحوه عملکرد ادغام بیشینه است.

### ۳-۲-۲-۲ لایه اتصال کامل

این لایه در انتهای شبکه پیچشی قرار دارد و ویژگی‌های استخراج شده در لایه‌های قبلی را به پاسخ نهایی تبدیل می‌کند. نرون‌های حاضر در این لایه، به تمامی نرون‌های لایه‌های قبلی و بعدی خود متصل هستند و از این منظر، این لایه یک شبکه پرسپترون چندلایه است. نحوه عملکرد این لایه در دورابطه ۴-۲ و ۵-۲ نشان داده شده است که در آن  $w$  ماتریس وزن بین دو لایه  $1 - l$  و  $l - 1$  مقدار نuron  $k$  در لایه  $1 - l$  است.

<sup>14</sup>Overfitting

<sup>15</sup>Max pooling



شکل ۴-۲: لایه ادغام [۳]

$$x_i^l = \sum_k v_k^{l-1} w_{k,i}^{l-1} \quad (4-2)$$

$$v_i^l = \sigma(x_i^l) \quad (5-2)$$

### ۳-۲-۲ آموزش شبکه‌های پیچشی

آموزش شبکه‌های پیچشی همانند سایر شبکه‌های عصبی از سه مرحله انتشار به جلو<sup>۱۶</sup>، انتشار به عقب<sup>۱۷</sup> و بروزرسانی وزن‌ها تشکیل می‌شود. در مرحله انتشار به جلو، ورودی به شبکه داده می‌شود تا خروجی آن ورودی محاسبه شود. این مرحله در قسمت‌های قبلی توضیح داده شده است. در مرحله انتشار به عقب، با مقایسه برچسب اصلی ورودی داده شده به شبکه و همینطور خروجی تولید شده توسط شبکه، خطای محاسبه می‌شود. سپس این خطای سرتاسر شبکه انتشار داده می‌شود تا میزان خطای برای لایه‌های مختلف محاسبه شود. با توجه به خطای محاسبه شده در این مرحله، وزن‌های شبکه عصبی توسط یک الگوریتم بهینه‌سازی بروزرسانی می‌شود تا عملکرد

<sup>16</sup>Forward propagation

<sup>17</sup>Backward propagation

شبکه پیچشی بهبود یابد. به عبارتی دیگر، از گرادیان محاسبه شده در مرحله انتشار به عقب برای بروز رسانی وزن‌های شبکه استفاده می‌شود. سه مرحله ذکر شده، در زیر با جزئیات بیشتری توضیح داده شده‌اند.

### ۱-۳-۲-۲ انتشار به جلو

این مرحله با دادن ورودی به شبکه شروع می‌شود. پس از آن، لایه‌های پیچشی و ادغام بر ورودی اعمال می‌شوند. خروجی این لایه‌ها به لایه اتصال کامل داده می‌شود تا خروجی نهایی شبکه برای ورودی داده شده را تولید کند.

### ۲-۳-۲-۲ انتشار به عقب

بدلیل تفاوتی که میان نحوه عملکرد لایه‌های اتصال کامل و پیچشی وجود دارد، نحوه محاسبه و بازنگشت خطاب نیز در این دو لایه متفاوت است.

در اولین مرحله، خطاب در لایه خروجی و توسط رابطه  $6-2$  محاسبه می‌شود. توسط رابطه‌های  $7-2$  و  $8-2$  خطاب در هر کدام از لایه‌های اتصال کامل محاسبه و به لایه قبلی از آن انتشار داده می‌شود. در نهایت گرادیان وزن‌های مختلف لایه اتصال کامل توسط رابطه  $9-2$  محاسبه می‌شود.

$$\frac{\partial E}{\partial v_i^L} = \frac{d}{dv_i^L} E(v^L) \quad (6-2)$$

$$\frac{\partial E}{\partial x_j^l} = \sigma'(x_j^l) \frac{\partial E}{\partial v_j^L} \quad (7-2)$$

$$\frac{\partial E}{\partial v_j^l} = \sum w_{i,j}^l \frac{\partial E}{\partial x_j^{l+1}} \quad (8-2)$$

$$\frac{\partial E}{\partial w_{i,j}^l} = v_i^l \frac{\partial E}{\partial x_j^{l+1}} \quad (9-2)$$

## فصل ۲. مروری بر مفاهیم پایه

همین فرآیند در مورد لایه پیچشی نیز اتفاق می‌افتد، با این تفاوت که به دلیل نحوه عملکرد متفاوت این لایه نسبت به لایه اتصال کامل، تغییراتی در رابطه‌های بالا خواهیم داشت. به همین خاطر، از رابطه‌های [۱۰-۲](#) و [۱۱-۲](#) برای محاسبه خطای لایه پیچشی و انتشار آن به لایه‌های قبل استفاده می‌شود. رابطه [۱۲-۲](#) نیز محاسبه گرادیان وزن‌های فیلترها را انجام می‌دهد.

$$\frac{\partial E}{\partial w_{i,j}^l} = \sum_{k=0}^{N-F} \sum_{m=0}^{N-F} \frac{\partial E}{\partial x_{m,k}^l} v_{i+k,j+m}^{l-1} \quad (10-2)$$

$$\frac{\partial E}{\partial x_{i,j}^l} = \frac{\partial E}{\partial v_{i,j}^l} \sigma'(x_{i,j}^l) \quad (11-2)$$

$$\frac{\partial E}{\partial v_{i,j}^{l-1}} = \sum_{k=0}^{F-1} \sum_{m=0}^{F-1} \frac{\partial E}{\partial x_{i-k,j-m}^l} w_{k,m} \quad (12-2)$$

### ۳-۳-۲-۲ بروزرسانی وزن‌ها

آخرین قدم، بروزرسانی وزن‌های شبکه با استفاده از گرادیان‌های محاسبه شده در مرحله انتشار به عقب است. محققین، الگوریتم‌های مختلفی را برای این منظور پیشنهاد داده‌اند. کاهش گرادیان تصادفی [۱۸](#) [۵۱]، [۱۹](#) [۵۲] یکی از روش‌هایی است که محققین در گذشته از آن برای بروزرسانی وزن‌های شبکه‌های عصبی استفاده می‌کردند. به مرور زمان و با افزایش داده‌های موجود در دسترس محققین، نیاز به الگوریتم‌های سریع‌تر حس شد. نتیجه این نیاز، الگوریتم‌های جدیدی نظیر Adam [۵۳](#) [۵۴] Adadelta [۵۵](#) [۵۶] هستند.

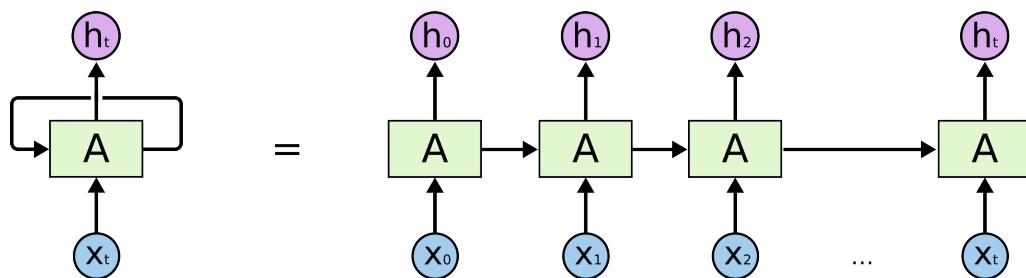
### ۴-۲-۲ شبکه‌های بازگشتی

این نوع شبکه عصبی از تعدادی واحد تشکیل شده است که با یال‌هایی جهت‌دار و وزن‌دار به یکدیگر متصل شده‌اند. ایده اصلی شبکه‌های بازگشتی استفاده از ماهیت دنباله‌ای [۱۹](#) اطلاعات است. در مدل‌های قبلی فرض

<sup>18</sup>Stochastic gradient descent

<sup>19</sup>Sequential

بر این بود که ورودی‌ها (خروجی‌ها) از یکدیگر مستقل هستند، در حالی که این فرض در برخی از مسائل درست نیست. علت نامگذاری شبکه‌های عصبی بازگشتی<sup>۲۰</sup> این است که شبکه برای هر ورودی از دنباله، ورودی‌های قبلی موجود در دنباله را نیز در نظر می‌گیرد [۵۶].



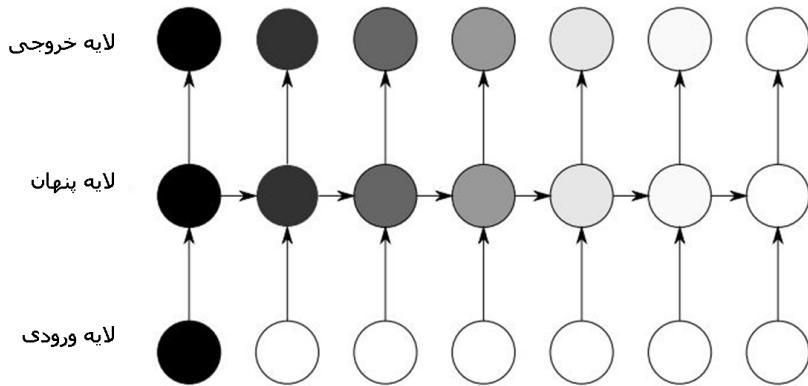
شکل ۲-۵: ساختار شبکه بازگشتی [۴]

## ۱-۴-۲-۲ فراموشی در دنباله‌های طولانی

شبکه‌های عصبی بازگشتی نسبت به شبکه‌های عصبی پیش‌رو در پیش‌بینی دنباله‌ای موفق‌تر عمل می‌کنند. علت این امر توانایی شبکه‌های بازگشتی در استفاده از اطلاعات قبلی در زمان نگاشت بین دنباله ورودی و دنباله خروجی است. ولی مشکلی که در شبکه‌های مطرح شده تاکنون همچنان وجود دارد، محدود بودن دامنه محتوای قابل دسترس است. همان‌طور که در شکل ۶-۲ مشاهده می‌کنید، در زمان  $t = 1$  اولین داده ورودی به شبکه وارد می‌شود، (تیره‌تر بودن رنگ لایه‌ها، حساسیت بیشتر هر لایه به اولین داده ورودی را در هر مرحله زمانی نشان می‌دهد)، در زمان  $t = 1$  تمام لایه‌ها بیشترین حساسیت را نسبت به اولین داده ورودی دارند، ولی در زمان‌های بعدی با توجه به وارد شدن داده‌های جدید، شبکه حساسیت خود را به ورودی اول از دست می‌دهد و عملاً بعد از گذشت چند مرحله زمانی، ورودی اول تاثیری در شبکه ندارد. از این رو شبکه‌های عصبی بازگشتی در پیش‌بینی دنباله‌های طولانی با مشکل مواجه هستند.

راه حلی که برای رفع این مشکل ارائه شد، شبکه عصبی حافظه کوتاه-مدت ماندگار بود. این شبکه با قرار دادن ۳ دروازه کنترلی روی هر یک از واحدهای لایه‌ی پنهان خود، تاثیر داده‌های مختلف را در طول زمان بر خروجی شبکه کنترل می‌کند.

<sup>20</sup>Recurrent Neural Network



شکل ۲-۶: فراموشی در دنباله های طولانی.

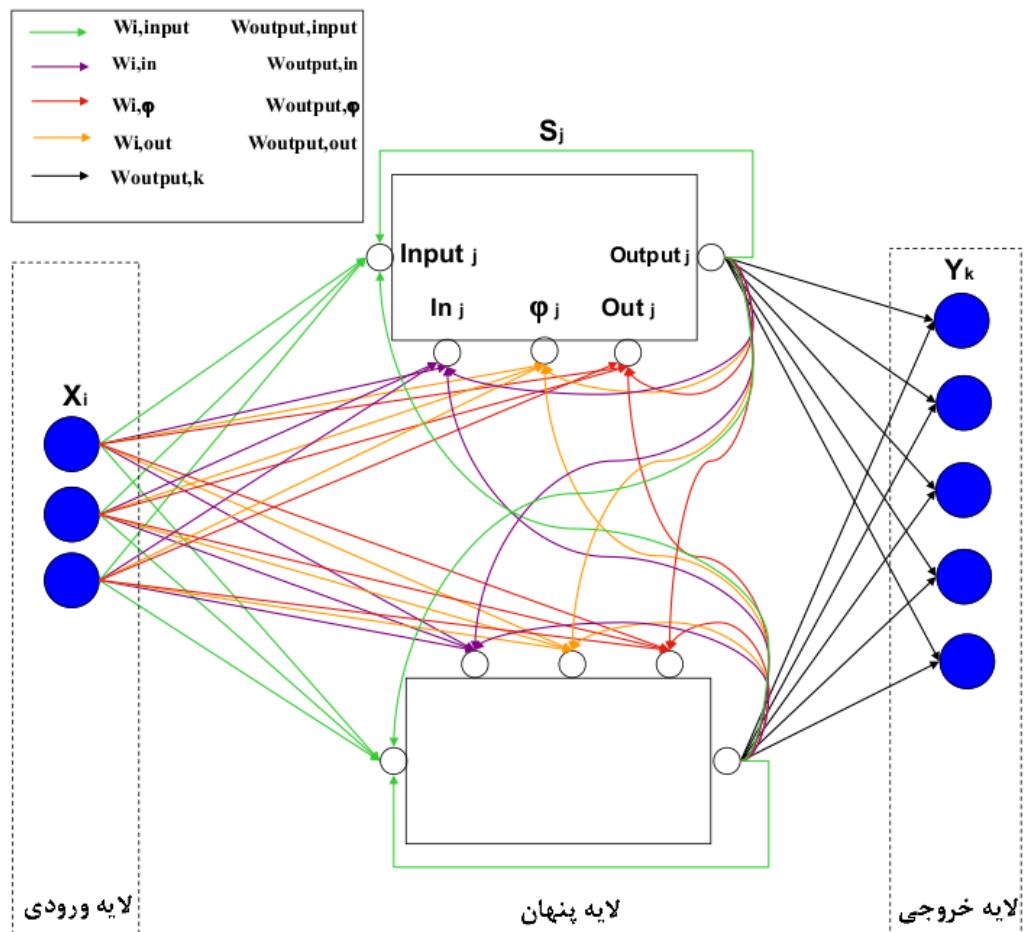
#### ۲-۴-۲-۲ شبکه حافظه کوتاه-مدت ماندگار

شبکه حافظه کوتاه-مدت ماندگار<sup>۲۱</sup> یک شبکه بازگشته است و از سه لایه تشکیل شده است که برای اولین بار در سال ۱۹۹۷ معرفی شد. [۵۷] لایه اول، ورودی است. لایه دوم این شبکه از تعدادی بلوک که به آنها بلوکهای حافظه کوتاه-مدت ماندگار گفته می‌شود، تشکیل شده است که در ادامه در مورد ساختار این بلوک توضیح داده خواهد شد. لایه سوم یا همان لایه خروجی این شبکه نیز از تعدادی واحد تشکیل شده که مقدار خروجی شبکه را مشخص می‌کنند. شکل ۷-۲ ساختار یک شبکه عصبی حافظه کوتاه-مدت ماندگار متتشکل از سه واحد ورودی، دو بلوک حافظه تک سلولی در لایه مخفی و پنج واحد خروجی را نشان می‌دهد.

#### ۳-۴-۲-۲ بلوک حافظه کوتاه-مدت ماندگار

هر بلوک شامل سه دروازه ورودی و یک دروازه خروجی است. دروازه‌های ورودی برای ورود داده‌ی جدید، فراموش کردن مقدار ذخیره شده در بلوک و اجازه‌ی ورود داده‌ی جدید به بلوک در نظر گرفته شده‌اند. دروازه خروجی برای بررسی امکان خروج داده‌ی جدید از بلوک تعریف می‌شود.

<sup>21</sup>Long short-term memory



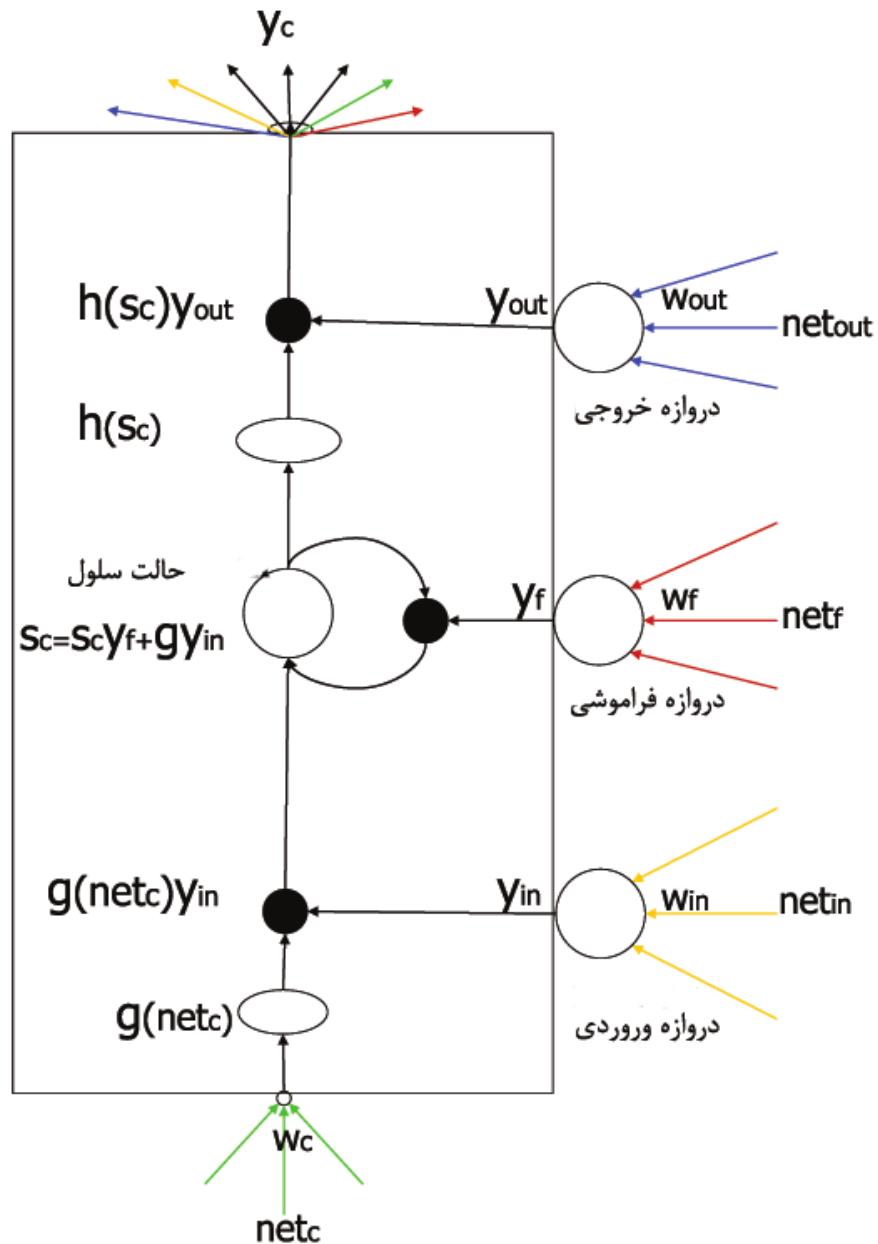
شکل ۷-۲: ساختار شبکه عصبی حافظه کوتاه-مدت ماندگار

همان طور که در شکل ۸-۲ مشاهده می‌کنید، ابتدا مقدار ورودی بلوک وارد شبکه می‌شود و پس از اعمال مقادیر دروازه‌های بلوک، مقدار خروجی بلوک مشخص می‌گردد.

مقدار خروجی دروازه ورودی مقداری بین صفر و یک است که نزدیک به صفر بودن این مقدار باعث بی‌اثر کردن مقدار وارد شده به بلوک است.

خروجی دروازه فراموشی مقداری بین صفر و یک دارد که نزدیک به صفر بودن مقدار این خروجی باعث از بین رفتن مقدار ذخیره شده در حافظه سلول می‌شود. در واقع بلوک مقداری را که تا این مرحله به خاطر سپرده بود فراموش می‌کند.

دروازه خروجی مقداری بین صفر و یک دارد که نزدیک به یک بودن آن به این معنی است که شبکه می‌تواند مقدار ذخیره شده در خود را به عنوان خروجی ارائه کند.



شکل ۲-۸: شمای کلی یک بلوک حافظه کوتاه-مدت ماندگار

## ۵-۲-۲ حذف تصادفی

حذف تصادفی<sup>۲۲</sup> یکی از راههای رایج برای مقابله با مشکل بیشبرازش در شبکه‌های عمیق است که با گسترش این نوع شبکه‌های عصبی در سال‌های اخیر، بسیار محبوب شده است. ایده اصلی پشت این تکنیک، جلوگیری از یادگیری بیش از اندازه مدل توسط داده‌های آموزشی است. بنابراین، در هر دور آموزش شبکه، هر نورون با احتمالی برابر با نرخ حذف تصادفی، ممکن است در آموزش شبکه شرکت نکند. این ایده، از انعطاف‌پذیری بیش از حد مدل در برابر داده‌های آموزشی جلوگیری می‌کند. حاصل این کار مدلی خواهد بود که از قدرت تعمیم‌پذیری بیشتری در برابر داده‌های جدید برخوردار است [۵۸، ۵۹].

## ۶-۲-۲ یادگیری انتقالی

در حوزه‌هایی از کاربرد یادگیری عمیق نظیر بینایی ماشین و پردازش زبان طبیعی، برای آموزش یک مدل کارا به حجم بسیار زیادی داده و همینطور توان پردازشی فراوان احتیاج است که همین امر آموزش یک شبکه عمیق را به کاری سخت و زمان بر تبدیل می‌کند. به همین دلیل، پژوهشگران با استفاده از روشی به نام یادگیری انتقالی<sup>۲۳</sup>، از مدل‌های از پیش آموزش دیده برای پیش‌بینی داده‌های جدید استفاده می‌کنند. در این روش، پژوهشگران شبکه‌های عمیق را یک بار با استفاده از حجم زیادی از داده‌ها آموزش می‌دهند. سپس، برای مجموعه داده‌های مشابه، از مدل آموزش داده شده استفاده می‌کنند و با اعمالی تغییرات جزئی، آن را منطبق بر مجموعه داده جدید می‌کنند [۶۰، ۶۱]. با استفاده از این روش، دانش به دست آمده در زمان آموزش شبکه با حجم بسیار زیاد داده، ذخیره‌سازی شده و مورد استفاده مجدد قرار می‌گیرد که در کاهش زمان آموزش بسیار موثر است.

## ۳-۲ الگوریتم ژنتیک

الگوریتم ژنتیک<sup>۲۴</sup> الهامی از نظریه ژنتیک و نظریه تکاملی داروین است و بر اساس بقای برترین‌ها یا انتخاب طبیعی استوار است. یک کاربرد متداول الگوریتم ژنتیک، استفاده از آن به عنوان تابع بهینه کننده است. [۵]

<sup>22</sup>Dropout

<sup>23</sup>Transfer learning

<sup>24</sup>Genetic Algorithm

## فصل ۲. مروری بر مفاهیم پایه

الگوریتم‌های ژنتیکی، نحوه تکامل ژنتیکی موجودات زنده شبیه‌سازی می‌شود. این الگوریتم‌ها با الهام از روند تکاملی طبیعت مسائل را حل می‌نمایند. یعنی مانند طبیعت یک جمعیت از موجودات را تشکیل می‌دهند و پس از طی مراحلی به یک مجموعه بهینه و یا موجود بهینه دست می‌یابند.

مراحل حل مسئله با استفاده از الگوریتم ژنتیک شامل موارد زیر است:

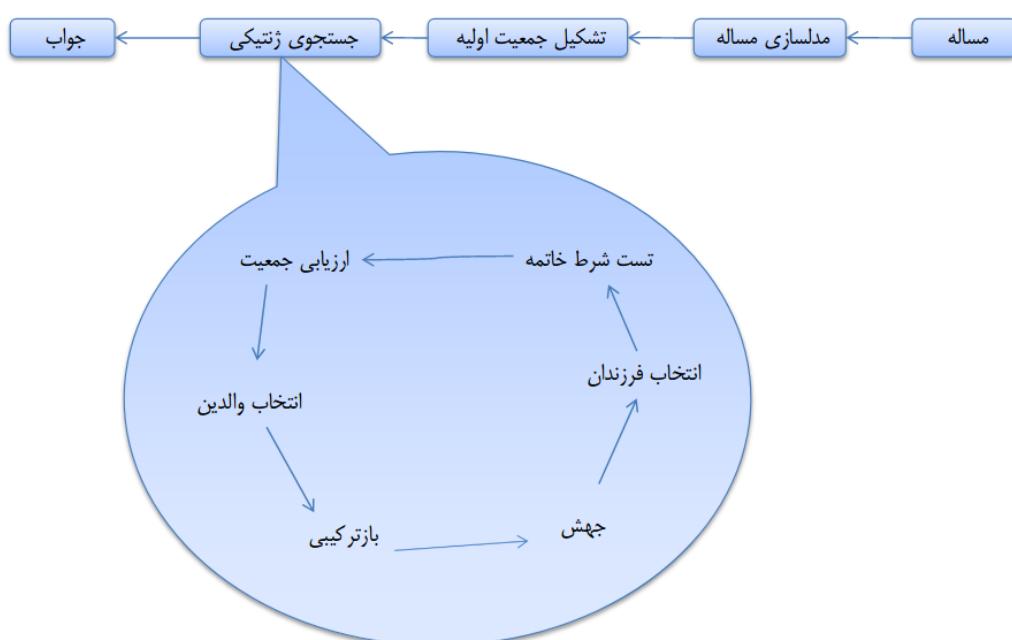
۱. مدل‌سازی مسئله

۲. تشکیل جمعیت اولیه

۳. جستجو ژنتیکی

۴. جواب

همانطور که در شکل ۹-۲ دیده می‌شود مرحله جستجوی ژنتیکی خود از چند بخش تشکیل شده است. در ادامه توضیحاتی در مورد موارد مطرح شده در شکل داده شده است.



شکل ۹-۲: مراحل مختلف الگوریتم ژنتیک [۵]

### ۱-۳-۲ مدل‌سازی مسئله

برای اینکه بتوانیم یک مسئله را با استفاده از الگوریتم ژنتیک حل کنیم، بایستی آن را به فرم مخصوص مورد نیاز این الگوریتم‌ها تبدیل کنیم. نحوه مدل‌سازی تأثیر بسیاری بر میزان پیچیدگی و کارایی الگوریتم دارد.

### ۲-۳-۲ تشکیل جمعیت اولیه

هر الگوریتم ژنتیک نیازمند یک تخمین اولیه از پاسخ است که در قالب جمعیت اولیه می‌توان به آن پرداخت. جمعیت اولیه، اولین پاسخ برای مسئله است که در خلال مرحله جستجوی ژنتیکی بهبود داده می‌شود. ممکن است جمعیت اولیه بصورت کاملاً تصادفی انتخاب شوند و یا اینکه با استفاده از روشی هوشمندانه، پاسخ‌هایی تولید شود که به پاسخ نهایی نزدیک باشند.

اندازه جمعیت اولیه معمولاً ثابت است و هرچه این میزان بیشتر باشد، میزان پویش<sup>۲۵</sup> و انتفاع<sup>۲۶</sup> بیشتر می‌شود، هرچند بار محاسباتی نیز افزایش پیدا می‌کند [۶۲].

### ۳-۳-۲ ارزیابی جمعیت

کروموزوم‌های تولید شده در هر مرحله بایستی از نظر میزان خوبی ارزیابی شوند. برای این منظور تابع برازش<sup>۲۷</sup> تعریف می‌شود تا به هر کروموزوم میزانی از خوب بودن آن کروموزوم را نسبت دهد. خروجی این تابع برای هر کروموزوم یک عدد است.

### ۴-۳-۲ انتخاب والدین

مرحله بعدی در الگوریتم ژنتیک انتخاب تعدادی از کروموزوم‌ها به عنوان والدین برای ایجاد فرزندان جدید است. در این انتخاب سعی بر این است تا والدین از کروموزوم‌هایی انتخاب شوند که تابع برازش مناسب‌تری دارند. با

<sup>25</sup>Exploration

<sup>26</sup>Exploitation

<sup>27</sup>Fitness Function

## فصل ۲. مروری بر مفاهیم پایه

این کار پاسخ‌ها را به سمت فضای جستجویی سوق می‌دهیم که امکان پیدا کردن پاسخ‌های باکیفیت در آن بیشتر است. این قسمت از الگوریتم بدنال تضمین بقای اصلاح داروین است.

### ۵-۳-۲ بازنگری

باززنگری<sup>۲۸</sup> یا هم‌برش<sup>۲۹</sup> مرحله‌ای دیگر در الگوریتم ژنتیک است که طی آن ژن‌های والدین را به امید یافتن پاسخ بهتر ترکیب می‌کنیم. همانطور که در مرحله قبل ذکر شده است، سعی بر این است که والدین از بین شایسته‌ترین کروموزوم‌ها انتخاب شوند، بنابراین می‌توان این احتمال را در نظر داشت که جستجو در فضای ترکیب ژن‌های آن‌ها، منجر به پاسخ‌های مناسبی شود. خروجی این کار فرزندانی جدید است که به میزان قابل قبولی ویژگی‌های والدین خود را حفظ کرده‌اند.

### ۶-۳-۲ جهش

اگر در تولید فرزندان جدید فقط به بازنگری اکتفا شود، این امکان وجود دارد که الگوریتم در نقاط بهینه محلی گیر بیفت. برای جلوگیری از این اتفاق و همچنین افزایش قابلیت پویش، از جهش<sup>۳۰</sup> استفاده می‌کنیم. در طی این فرآیند برخی از ژن‌های فرزندان تغییر پیدا می‌کند تا این امکان نیز در نظر گرفته شود که برخی از ژن‌های خوب ممکن است در ترکیب ژن‌های والدین حضور نداشته باشند.

### ۷-۳-۲ انتخاب فرزندان

پس از اینکه فرزندان تولید شدند، بایستی تصمیم گرفت که به چه صورتی از فرزندان و والدین آن‌ها در تولید پاسخ‌ها استفاده کرد. دوره‌یکرد کلی در این مورد وجود دارد که در زیر به آن‌ها اشاره شده است.

#### • جایگزینی حالت پایدار<sup>۳۱</sup>

<sup>28</sup> Recombination

<sup>29</sup> Crossover

<sup>30</sup> Mutation

<sup>31</sup> Steady State Replacement

در این حالت، فرض بر این است که جمعیت فعلی از وضعیت مناسبی برخوردار است. بنابراین حفظ آن‌ها در اولویت قرار می‌گیرد و بخش کوچکی از فرزندان با والدینشان جایگزین می‌شوند.

### • جایگزینی نسلی<sup>۳۲</sup>

در این حالت الگوریتم بخش اعظم والدین را با فرزندانشان جایگزین می‌کند. این روش باعث همگرایی سریع الگوریتم می‌شود اما احتمال گیر افتادن در بهینه محلی را نیز افزایش می‌دهد.

## ۸-۳-۲ شرط خاتمه الگوریتم

از آنجایی که الگوریتم ژنتیک بر مبنای تولید و تست می‌باشد، جواب بهینه را نمی‌دانیم و به همین دلیل مجبور هستیم تا شرایط دیگری را برای اتمام اجرای الگوریتم در نظر بگیریم.  
در زیر برخی از این موارد ذکر شده است:

- محدود کردن الگوریتم به تعداد مشخصی نسل
- عدم تغییر جمعیت در نسل‌های متوالی
- شمارش تعداد نسل‌هایی که بهترین پاسخ تغییر نکرده است

## ۴-۲ کاهش بعد و استخراج ویژگی

استخراج دانش از داده‌های با بعد بالا، مستلزم صرف زمان زیادی است. بنابراین منطقی به نظر می‌رسد که ما روش‌هایی را برای کاهش بعد<sup>۳۳</sup> داده‌ها به کار ببریم.  
دلایل زیادی برای کاهش بعد داده‌ها وجود دارد که از جمله آن‌ها می‌توان به موارد زیر اشاره کرد:

- سرعت الگوریتم‌ها با داده‌های با بعد کمتر بیشتر می‌شود
- فضای ذخیره سازی کمتری نیاز است

<sup>32</sup>Generational Replacement

<sup>33</sup>Dimension reduction

- احتمال بیش برازش<sup>۳۴</sup> کاهش می‌باید و بنابراین قدرت تعمیم الگوریتم‌های یادگیری بیشتر می‌شود
- برای ترسیم و بدست آوردن درکی از مجموعه داده‌ها گاهی بعد داده‌ها را به دو یا سه، تقلیل می‌دهند تا بتوانند نموداری از داده‌های با ابعاد زیاد ترسیم کنند.

در برخی موارد ممکن است داده‌هایی که در اختیار داریم، خام و ابتدایی باشد. در این شرایط می‌توان با استفاده از روش‌های استخراج ویژگی<sup>۳۵</sup>، ویژگی‌هایی را استخراج کرد که نقش مهم‌تری را در استخراج دانش از داده‌ها ایفا کنند.

روش آماری تحلیل مؤلفه‌های اساسی و روش‌های مبتنی بر شبکه عصبی ماشین بولتزمن محدود، خود را مزگذار و شبکه باور عمیق این موارد را برای ما انجام می‌دهند. [۶۳]

## ۱-۴-۲ تحلیل مؤلفه‌های اساسی

روش تحلیل مؤلفه‌های اساسی<sup>۳۶</sup> یکی از معروف‌ترین روش‌های مورد استفاده برای کاهش بعد است. این الگوریتم از جمله با ارزش‌ترین کاربردهای جبرخطی است که در علوم مختلف استفاده شده است زیرا یک روش آسان برای استخراج اطلاعات مرتبط از یک مجموعه پیچیده است.

این روش واریانس کل ویژگی‌ها را تحلیل می‌کند و مؤلفه‌ها طوری برآورده می‌شوند که پس از تبدیل به بعد کمتر، بخش اعظم واریانس حفظ شود. در واقع مؤلفه‌های اصلی تولید شده ترکیبی خطی از ویژگی‌های ابتدایی می‌باشند. [۶۴]

این روش از مقادیر ویژه<sup>۳۷</sup> و بردارهای ویژه<sup>۳۸</sup> ماتریس کوواریانس<sup>۳۹</sup> داده‌ها استفاده می‌کند. ماتریس کوواریانس یک ماتریس متقارن است و عناصر قطر اصلی آن مقادیر واریانس‌های ویژگی‌ها هستند. بنابراین با فرض اینکه بعد داده‌های ما  $n$  باشد، ما به یک ماتریس  $n \times n$  مربعی دست خواهیم یافت.

مرحله بعدی بدست آوردن بردارها و مقادیر ویژه برای این ماتریس است. از معادله زیر برای بدست آوردن

<sup>34</sup>Overfitting

<sup>35</sup>Feature extraction

<sup>36</sup>Principal component analysis (PCA)

<sup>37</sup>Eigenvalue

<sup>38</sup>Eigenvector

<sup>39</sup>Covariance Matrix

آن‌ها استفاده می‌شود.

$$C \times E_{vector} = \lambda \times E_{vector} \quad (13-2)$$

در فرمول ۱۳-۲،  $C$  ماتریس کوواریانس،  $E_{vector}$  بردار ویژه و  $\lambda$  مقدار ویژه هستند. بردارهای ویژه متعامد<sup>۴۰</sup> هستند و این نکته موضوع مهمی است زیرا می‌خواهیم داده‌ها را با این بردارها توصیف کنیم.

پس از محاسبه  $n$  مقدار ویژه، آن‌ها را به صورت نزولی مرتب می‌کنیم.

$$\lambda_1 > \lambda_2 > \dots > \lambda_n \quad (14-2)$$

در انتها با انتخاب عدد  $k$  که معمولاً کوچک‌تر از  $n$  می‌باشد، ابعاد داده را از  $n$  به  $k$  کاهش می‌دهیم. انتخاب  $k$  با کمک یک سطح آستانه که می‌تواند توسط کاربر مشخص شود، تعیین می‌شود. اگر مقدار ۹۵.۰ برای حد آستانه مشخص شود، مقدار  $k$  را می‌توان از از عبارت زیر بدست آورد.

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^n \lambda_i} \geq 0,95 \quad (15-2)$$

این بدين معنی است که پس از تبدیل به  $k$  بعد، ۹۵.۰ درصد از خصوصیات داده‌های اصلی حفظ شده است. مقدار  $k$  کوچکترین مقداری است که در نامساوی ۱۵-۲ صدق کند.

مرحله پایانی در الگوریتم بدست آوردن داده‌های جدید است. این داده‌ها طبق فرمول ۱۶-۲ بدست می‌آیند.

$$Data_{new} = Data_{original} \times Transform \quad (16-2)$$

در فرمول ۱۶-۲،  $Transform$  ماتریس بردارهای ویژه است. در ستون‌های این ماتریس  $k$  بردار ویژه متناظر با بزرگترین  $k$  مقدار ویژه قرار داده شده است.

---

<sup>40</sup>Orthogonal

## ۲-۴-۲ خودرمزگذار

خودرمزگذار<sup>۴۱</sup> یک نوع شبکه عصبی است که از آن برای بدست آوردن یک نمایش مطلوب از داده‌های ورودی استفاده می‌شود. عمدهاً از خودرمزگذار در یادگیری بدون ناظر<sup>۴۲</sup> برای کاهش بعد استفاده می‌شود. در واقع ساختار یک خودرمزگذار به گونه‌ای بنا شده است که در آن بدنال بدست آوردن وزن‌هایی باشیم که خطای بازسازی<sup>۴۳</sup> را کمینه کند.

در خودرمزگذار همانند سایر شبکه‌های عصبی لایه نام شبکه از طریق فرمول زیر محاسبه می‌شود.

$$h_i = \delta_i(W_i h_{i-1} + b_i) \quad (17-2)$$

در فرمول ۱۷-۲،  $h_i$  خروجی لایه،  $b_i$  ورودی لایه،  $h_{i-1}$  بایاس لایه و  $\delta_i$  تابع فعالسازی لایه است. اگرچه می‌توان از توابع فعالسازی مختلف برای هر لایه استفاده کرد اما مرسوم است که از تابع سیگموید<sup>۴۴</sup> برای تمامی لایه‌ها استفاده شود. فرمول ۱۸-۲ نشان دهنده تابع سیگموید است.

$$\delta(z) = \frac{1}{1 + \exp(-z)} \quad (18-2)$$

با فرض اینکه شبکه دارای  $2L$  لایه باشد،  $L$  لایه اول مربوط به ساختن نمایشی جدید از داده‌ها است و  $L$  لایه آخر مرتبط با بازسازی نمایش داده است. یعنی داریم:

$$u = \phi(x) = h_L \circ \dots \circ h_1(x) \quad (19-2)$$

$$x_{rec} = \varphi(u) = h_{2L} \circ \dots \circ h_{L+1}(u) \quad (20-2)$$

<sup>41</sup>Autoencoder

<sup>42</sup>Unsupervised Learning

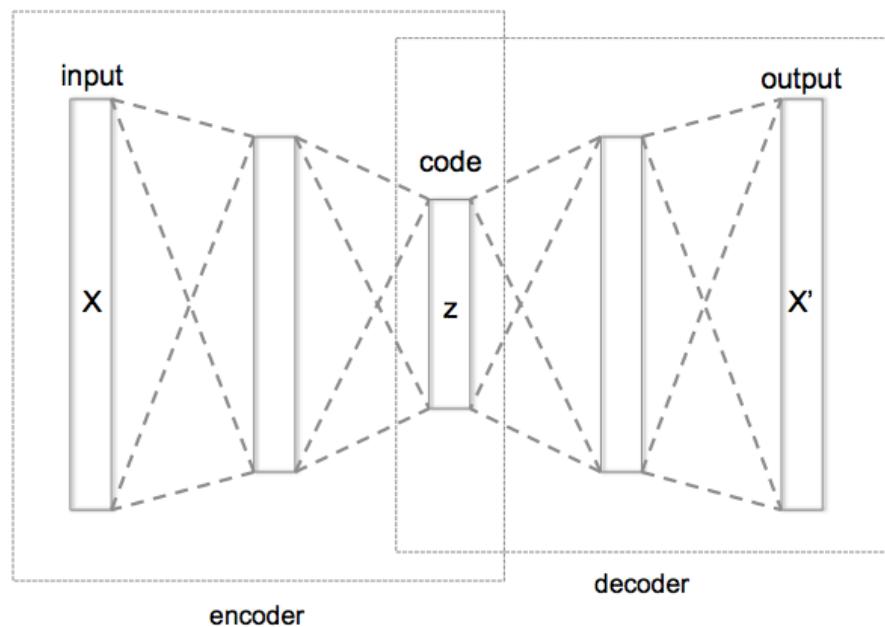
<sup>43</sup>Reconstruction Error

<sup>44</sup>Sigmoid

تابع هزینه نیز به شکل فرمول ۲۱-۲ تعریف می‌شود.<sup>[۱۹]</sup> در حین یادگیری بدنال کمینه کردن خطای بازسازی از طریق تعیین مناسب پارامترهای  $\theta = \{W_i, b_i\}, i = 1, \dots, 2L$  هستیم.

$$\min_{\theta} \frac{1}{N} \sum_{n=1}^N \|x^n - \varphi \circ \phi(x^n)\|^2 \quad (21-2)$$

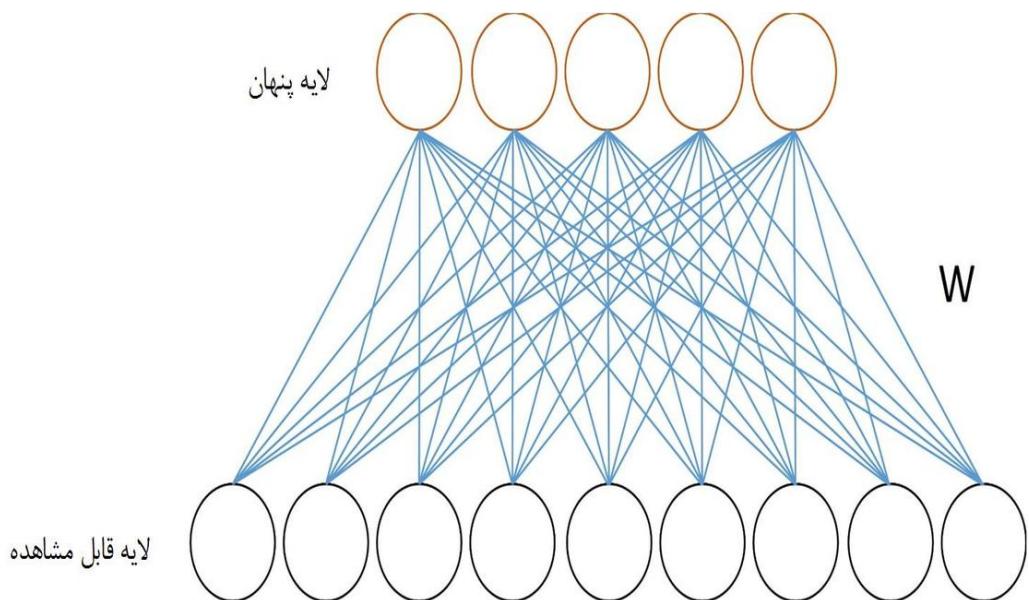
شکل ۱۰-۲ ساختار یک خودرمزنگذار را نشان می‌دهد که در مرحله اول ورودی را کد می‌کند و در مرحله دوم کدگشایی می‌کند.



شکل ۱۰-۲: ساختار یک خودرمزنگذار [۶]

### ۳-۴-۲ ماشین بولتزمن محدود

ماشین بولتزمن محدود<sup>۴۵</sup> یک نوع شبکه عصبی است که از دو مجموعه تشکیل شده است. واحدهای قابل مشاهده که در واقع همان داده آموزشی هستند و واحدهای پنهان<sup>۴۶</sup> که پس از آموزش شبکه، با استفاده از آنها می‌توان بازنمایی ورودی را بدست آورد.<sup>۴۷</sup> [۶۵] محدودیت قید شده در نام این شبکه به این دلیل است که در هر لایه بین هیچ دو نوروں داخلی اتصالی وجود ندارد. هدف ماشین بولتزمن محدود بدست آوردن بازنمایی داده‌ها است. در این شبکه یکتابع انرژی تعریف می‌شود که واحدهای شبکه و اتصالات بین آنها را مدل می‌کند. هر اتصال بین دو واحد دارای یک وزن است. شکل ۱۱-۲ نشان دهنده ساختار یک ماشین بولتزمن محدود است. ماشین بولتزمن محدود این قابلیت را دارد که تعداد ابعاد داده‌های ورودی را به اندازه قابل توجهی کاهش دهد و همچنین ویژگی‌های اساسی که در مجموعه داده وجود دارند را حفظ کند.



شکل ۱۱-۲: ساختار یک ماشین بولتزمن محدود

این ماشین یکتابع توزیع روی یک بردار به نام بردار قابل مشاهده  $v \in \{0, 1\}^n$  و یک بردار به نام بردار پنهان  $h \in \{0, 1\}^m$  دارد. به عبارت دیگر، ماشین بولتزمن محدود تلاش می‌کند که مقادیر  $h = h_1, h_2, \dots, h_m$  را طوری یاد بگیرد که یک بازنمایی از  $v_1, v_2, \dots, v_n = v$  باشد.

<sup>45</sup>Restricted Boltzman Machine

<sup>46</sup>Visible

<sup>47</sup>Hidden

هر کدام از لایه‌ها شامل نورون‌های دودویی هستند که می‌توانند فعال<sup>۴۸</sup> و یا غیرفعال<sup>۴۹</sup> باشند. به علت عدم وجود اتصال بین واحدهایی که در یک لایه هستند، احتمال شرطی<sup>۵۰</sup> لایه پنهان که از یکدیگر مستقل<sup>۵۱</sup> هستند به صورت زیر محاسبه می‌شوند:

$$P(h|v) = \prod_{i=1}^m P(h_i|v) \quad (22-2)$$

همچنین، احتمال اینکه مقدار یک واحد از لایه پنهان برابر یک قرار گیرد، برابر است با:

$$p(h_i = 1|v) = \sigma\left(\sum_{j=1}^n w_{ij}v_j + b_i\right) \quad (23-2)$$

این شبکه برای یادگیری مقادیر متغیرهای لایه پنهان، یک تابع بر مبنای لگاریتم درست‌نمایی بیشینه<sup>۵۲</sup> تعریف می‌کند و تلاش می‌کند آن را بیشینه کند. فرمول<sup>۲۴-۲</sup> نشان دهنده این تابع است.

$$l(\theta) = \frac{1}{n} \log P(v) = \frac{1}{n} \sum_i \log(P(v_i)) \quad (24-2)$$

که  $p(v)$  را می‌توان از طریق توزیع حاشیه‌ای<sup>۵۳</sup> روی مقادیر احتمال توأم  $v$  و  $h$  به صورت زیر به دست آورد:

$$p(v) = \sum_h p(v, h) \quad (25-2)$$

احتمال توأم دو بردار  $v$  و  $h$  به شکل زیر تعریف می‌شود:

$$P(v, h) = \frac{\exp(-E(v, h))}{z} \quad (26-2)$$

<sup>48</sup>Active

<sup>49</sup>Inactive

<sup>50</sup>Conditional probability

<sup>51</sup>Independent

<sup>52</sup>Log Likelihood

<sup>53</sup>Marginal distribution

که  $\hat{z}$  یک مقدار نرمال‌سازی است.

تابع انرژی نیز به صورت زیر تعریف می‌شود:

$$E = -h^T w v - b^T h - c^T v \quad (27-2)$$

مقادیر احتمال شرطی  $p(h_i = 1|v)$  و  $p(v_j = 1|h)$  به ترتیب در معادله ۲۲-۲ و معادله ۲۳-۲ آورده شده‌اند.

همچنین، مقادیر احتمال شرطی  $p(v_j|h)$  به این صورت بدست می‌آیند:

$$P(v_j|h) = \prod_{j=1}^q P(v_j|h) \quad (28-2)$$

$$p(v_j = 1|h) = \sigma\left(\sum_{i=1}^p w_{ij} h_i + c_j\right) \quad (29-2)$$

## ۵-۲ خوشبندی

به قرار دادن نمونه‌های مختلف داده در خوشه‌های مختلف بطوری که اعضای یک خوش، بیشترین شباهت را به یکدیگر داشته باشند، خوشبندی<sup>۵۴</sup> گفته می‌شود [۶۶، ۶۷]. الگوریتم‌های خوشبندی، بنابر نحوه عملکردشان به دسته‌های مختلفی تقسیم می‌شوند که در زیر به دو موردشان اشاره شده‌است.

- **مرکز محور**<sup>۵۵</sup>: در این خوشبندی، هر خوش یک مرکز دارد و الگوریتم، داده‌ها را به خوشه‌ای نسبت می‌دهد که داده به مرکز آن خوش نزدیک‌تر باشد.

- **سلسله‌مراتبی**<sup>۵۶</sup>: این نوع خوشبندی نیز داده‌ها را براساس فاصله‌شان تا سایر داده‌ها در یک خوشه قرار می‌دهد. برخلاف سایر الگوریتم‌های خوشبندی، در این نوع خوشبندی نتیجه نهایی پس از طی تعداد

<sup>54</sup>Clustering

<sup>55</sup>Center-Based

<sup>56</sup>Hierarchical

زیادی ادغام و یا تقسیم تولید می‌شود و خروجی الگوریتم وابستگی زیادی به معیار توقف الگوریتم دارد [۶۸].

## ۱-۵-۲ خوشبندی $k$ میانگین

این الگوریتم، یک الگوریتم خوشبندی مبتنی بر مرکز است که در آن هر کدام از  $k$  خوش با یک مرکز خوش مشخص می‌شوند و در هر مرحله الگوریتم، داده‌ها به خوش‌ای که کمترین فاصله را از مرکز آن دارند، اختصاص می‌دهد. سپس، مرکز هر خوش با میانگین اعضای آن خوش تغییر پیدا می‌کند و این روند تا زمانی که شرایط توقف فراهم نشده است، ادامه پیدا می‌کند [۶۹]. نحوه کار این الگوریتم را می‌توان در زیر مشاهده کرد:

۱. انتخاب تصادفی  $k$  مرکز خوش

۲. تا زمانی که مراکز خوش‌ها تغییر می‌کند مراحل زیر تکرار شوند:

(آ) انتصاب هر کدام از داده‌ها به خوش‌ای که فاصله کمتری با آن دارد (فاصله اقلیدسی یا کسینوسی

(...)

(ب) به روزرسانی مرکز خوش‌ها با استفاده از میانگین‌گیری میان تمامی داده‌های مرتبط با آن خوش

## ۶-۲ جمعبندی

در این فصل تعدادی الگوریتم مورد بررسی قرار گرفت که می‌توان آن‌ها را در چهار دسته کلی شبکه‌های عصبی و یادگیری عمیق، الگوریتم ژنتیک، کاهش بعد و استخراج ویژگی و خوشبندی قرار داد. شبکه‌های عصبی تلاشی برای حل مسئله با الگوبرداری از شبکه‌های عصبی ذهن انسان هستند. این شبکه‌ها از سه لایه ورودی، مخفی و خروجی تشکیل شده‌اند. یادگیری عمیق تلاشی جدید است که در آن با افزایش لایه‌های مخفی، سعی بر این است تا استخراج ویژگی بر عهده شبکه قرار گیرد. در شبکه عصبی پیشرو اطلاعات تنها در جهت رو به جلو از لایه ورودی به سمت لایه خروجی حرکت می‌کنند. شبکه پیچشی، با هدف استخراج خودکار ویژگی (عمدتاً در مورد تصاویر) ابداع شده است. شبکه‌های بازگشتی ابزار مناسبی برای مدل کردن

## فصل ۲. مروری بر مفاهیم پایه

داده‌های دنباله‌ای هستند و از معروف‌ترین آن‌ها می‌توان به شبکه حافظه کوتاه-مدت ماندگار اشاره کرد. حذف تصادفی، یکی از روش‌های محبوب برای حل مشکل بیش‌برازش در شبکه‌های عمیق است. یادگیری انتقالی نیز امکان ذخیره‌سازی و استفاده مجدد از دانش بدست‌آمده در هنگام آموزش یک شبکه عمیق را فراهم می‌کند. الگوریتم ژنتیک نیز سعی دارد تا با الهام از طبیعت و نظریه انتخاب اصلاح به حل مسائل پردازد. این روش مانند طبیعت، یک جمعیت از پاسخ‌های مسئله را تشکیل می‌دهد و سپس طی فرآیند جستجوی ژنتیکی، آن‌ها را بهبود می‌دهد تا به پاسخی بهینه دست یابد.

زمانی که با تعداد زیادی از ویژگی‌ها مواجه هستیم، بدلیل سختی و زمان بر بودن پردازش و یا عدم وجود ویژگی‌های مناسب، نیاز به کاهش بعد و استخراج ویژگی حس می‌شود. تحلیل مؤلفه‌های اساسی روشی آماری است که ابعاد ویژگی‌ها را با در نظر گرفتن واریانس آن‌ها کاهش می‌دهد. ماشین بولتزمن محدود نوعی شبکه عصبی است که نمایشی از داده را پیدا می‌کند که به داده اصلی نزدیک باشد. خودر مزگذار نیز نوع دیگری از شبکه عصبی است که با کمینه کردن خطای بازسازی داده‌ها، بعد آن‌ها را می‌تواند کاهش دهد.

هدف اصلی از خوشه‌بندی پیدا کردن داده‌هایی است که خصوصیات مشابه دارند. الگوریتم k میانگین را می‌توان معروف‌ترین الگوریتم خوشه‌بندی مرکز محور دانست که در آن داده‌ها به خوشه‌ای که کمترین فاصله را با آن دارند، انتصاب پیدا می‌کنند.

## فصل ۳

### مروری بر ادبیات تحقیق

#### ۱-۳ مقدمه

پس از معرفی الگوریتم‌های مورد استفاده در مقالات، نوبت به تشریح کارهای مشابه می‌رسد. با توجه به اینکه این پایان‌نامه در مورد کاربرد الگوریتم‌های یادگیری عمیق که نوع خاصی از شبکه‌های عصبی هستند در بازارهای سرمایه است، کارهای مشابه به دو دسته کلی شبکه‌های کم‌عمق و شبکه‌های عمیق تقسیم شده است. برخی از پژوهشگران در هنگام استفاده از یک شبکه کم‌عمق از الگوریتم‌هایی نظیر الگوریتم ژنتیک و همینطور تحلیل مؤلفه‌های اساسی بهره برده‌اند که در این قسمت به آن‌ها نیز اشاره خواهد شد. الگوریتم‌های مبتنی بر یادگیری عمیق نیز به دسته‌های مختلف نظیر شبکه پرسپترون چندلایه عمیق، شبکه‌های بازگشتی و شبکه پیچشی تقسیم شده‌است.

#### ۲-۳ شبکه‌های کم‌عمق

شبکه‌های کم‌عمق الگوریتم‌هایی هستند که سالیان دراز در پیش‌بینی بازارهای مالی حضور داشته‌اند. این دسته از الگوریتم‌ها عمدتاً به شبکه پرسپترون چندلایه محدود می‌شوند. داشتن تنها یک لایه مخفی ویژگی‌ای است که شبکه‌های کم‌عمق را از شبکه‌های عمیق متمایز می‌کند. بدلیل همین ویژگی، توانایی آن‌ها در استخراج ویژگی‌های

## فصل ۳. مروری بر ادبیات تحقیق

مناسب در مقایسه با شبکه‌های عمیق محدود است. پژوهش‌های مطرح شده در این بخش به سه دسته شبکه پرسپترون چندلایه، ترکیب شبکه عصبی و الگوریتم ژنتیک و همچنین ترکیب شبکه عصبی و تحلیل مؤلفه‌های اساسی دسته‌بندی می‌شوند.

### ۱-۲-۳ شبکه پرسپترون چندلایه

بنابر پژوهش‌های مختلف [۷۰]، این نوع شبکه‌ها محبوب‌ترین الگوریتم‌های سنتی یادگیری ماشین هستند که در بازارهای مالی به کار گرفته شده‌اند [۷۱، ۷۲، ۷۳]. استفاده از الگوریتم پس انتشار خطأ<sup>۱</sup> برای آموزش این نوع شبکه‌های عصبی بسیار رایج است [۷۴، ۷۵].

در [۷۶] نویسنده‌گان توانستند با موفقیت از یک شبکه پرسپترون چندلایه برای پیش‌بینی بازار بورس ترکیه استفاده کنند. شبکه استفاده شده در این پژوهش یک پرسپترون چندلایه است که دارای یک لایه مخفی می‌باشد. ورودی به این شبکه ۱۰ تکنیکال اندیکاتور است و خروجی آن نیز ۱ نورون است که جهت تغییرات قیمت را برای روز آینده پیش‌بینی می‌کند. تعداد نورون‌های لایه مخفی نیز یکی از پارامترهایی است که پژوهشگران برای یافتن بهترین نتیجه، اعداد مختلفی را امتحان کرده‌اند.

داده‌های استفاده شده در این پژوهش مربوط به تغییرات جهتی شاخص بورس استانبول در دوره زمانی ژانویه ۱۹۹۷ تا دسامبر ۲۰۰۷ است. ابتدا ۲۰٪ از داده‌های هر سال برای پیدا کردن پارامترهای بهینه مورد استفاده قرار گرفته‌اند. سپس، داده‌های هر سال به دو قسمت تقسیم شده‌اند که قسمت اول برای آموزش و قسمت دوم برای آزمودن کارایی روش پیشنهادی مورد استفاده قرار گرفته‌اند. لازم به ذکر است که بهترین نتایج زمانی حاصل شده است که تعداد نورون‌های لایه مخفی ۹۰، نرخ یادگیری ۱٪ و تعداد دوره‌های آموزش شبکه ۶۰۰۰ بوده است. شکل ۱-۳ نشان دهنده میزان دقت کسب شده توسط شبکه عصبی در پیش‌بینی تغییرات جهتی سال‌های مختلف است.

<sup>۱</sup>Error Backpropagation

Prediction performance (%) of ANN model for best parameter combinations ( $lr = 0.1$ ).

Parameter combination (ep; mc; n)						
Year	(5000; 0.7; 30)		(7000; 0.1; 90)		(6000; 0.4; 90)	
	Training	Holdout	Training	Holdout	Training	Holdout
1997	100	76.98	100	79.37	100	79.37
1998	99.19	73.29	99.19	73.39	98.39	76.61
1999	99.16	75.21	99.16	76.92	99.16	75.21
2000	100	76.42	100	79.67	100	78.86
2001	100	69.11	100	70.73	99.20	71.54
2002	98.43	76.80	99.21	79.20	100	77.60
2003	100	69.92	99.19	68.29	100	73.98
2004	100	75.00	100	72.58	100	72.58
2005	100	78.57	100	76.98	100	76.98
2006	98.41	74.19	97.62	70.97	97.62	74.19
2007	97.62	75.40	97.62	73.81	97.62	76.19
Average	99.35	74.63	99.27	74.72	99.27	75.74

شکل ۳-۱: میزان دقت شبکه عصبی در پیش‌بینی سال‌های مختلف [۵]

## ۲-۲-۳ ترکیب شبکه عصبی و الگوریتم ژنتیک

پس انتشار خطا یکی از روش‌های مرسوم برای یادگیری شبکه‌های عصبی است اما همانطور که در مقالات این بخش به آن اشاره می‌شود، بدلیل نویزی بودن داده‌های موجود در بازار بورس ممکن است به نتیجه مطلوب نرسد. لذا می‌توان در تعیین وزن‌های بین نورون‌ها از الگوریتم ژنتیک استفاده کرد تا این نقص را پوشش دهد. در این مقاله [۸] نویسنده‌گان توانسته‌اند با استفاده از ترکیب شبکه عصبی و الگوریتم ژنتیک، جهت تغییرات قیمت را برای روز آینده در بازار سهام کره پیش‌بینی کنند.

شبکه عصبی در نظر گرفته شده در این مقاله شامل ۱۲ نورون در قسمت ورودی، ۱۲ نورون در لایه مخفی و یک نورون در لایه خروجی است. هدف شبکه تعیین بالا یا پایین رفتن قیمت سهام است که یک مسئله از نوع دسته‌بندی می‌باشد.

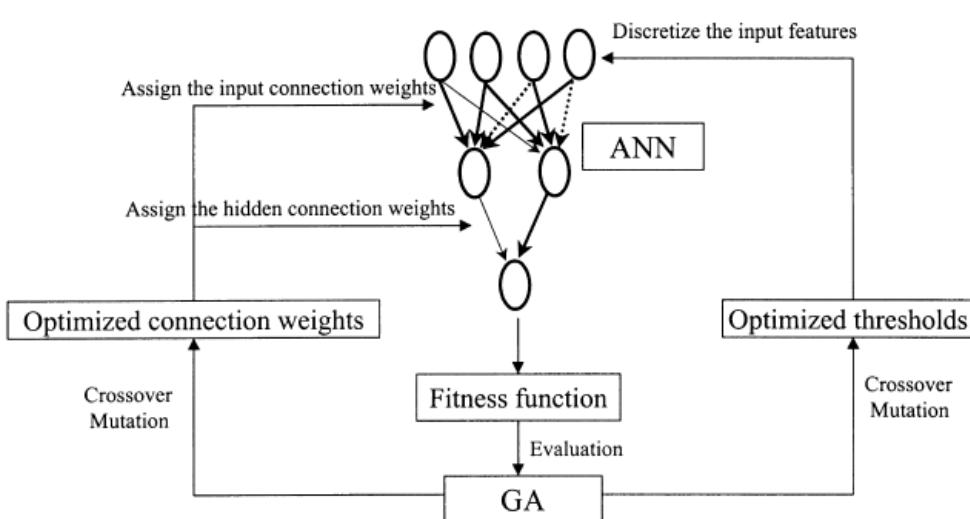
علاوه بر این برای غلبه بر نویز داده‌ها، نویسنده‌گان از تکنیک گسسته‌سازی<sup>۲</sup> مقادیر ویژگی‌ها نیز بهره برده‌اند. هر ویژگی به ۵ کلاس گسسته شده است. تعیین مقادیر آستانه برای گسسته‌سازی نیز بر عهده الگوریتم ژنتیک قرار دارد.

<sup>2</sup>Discretization

### فصل ۳. مروری بر ادبیات تحقیق

نحوه ساخت کروموزوم‌ها: لایه‌های ورودی و مخفی هرکدام ۱۲ نورون دارند، بنابراین ۱۴۴ وزن در این بین وجود دارد که ۱۴۴ ژن اول کروموزوم را تشکیل می‌دهند. همچنین بین لایه مخفی و خروجی نیز ۱۲ وزن وجود دارد که ۱۲ ژن بعدی را تشکیل می‌دهد. در انتها نیز به ۴۸ مقدار آستانه برای گسترش‌سازی احتیاج داریم که ۴۸ ژن بعدی را تشکیل می‌دهد. (۱۲ ویژگی که هرکدام به ۴ مقدار آستانه نیاز دارند)

شکل ۲-۳ نشان دهنده ساختار شبکه عصبی و کاربرد الگوریتم ژنتیک در آن است.



شکل ۲-۳: کاربرد الگوریتم ژنتیک در شبکه عصبی [۸]

داده‌های این مقاله مربوط به سال‌های ۱۹۸۹-۱۹۹۸ بازار سهام کره می‌باشد. ۸۰٪ از داده‌های هر سال برای آموزش و ۲۰٪ باقیمانده برای آزمون استفاده شده است. نویسنده‌گان روش خود را با ۲ روش دیگر مقایسه کردند. روش اول استفاده از شبکه عصبی و یادگیری با پس انتشار خط است (BPLT). روش دوم استفاده از شبکه عصبی و الگوریتم ژنتیک برای تعیین وزن‌ها است (GALT). روش نویسنده‌گان نیز استفاده از شبکه عصبی و الگوریتم ژنتیک برای گسترش‌سازی و تعیین وزن‌ها است. شکل ۳-۳ نتایج بدست آمده را بررسی و مقایسه می‌کند. بهبود نتایج هم در قسمت آموزش و هم در قسمت آزمون کاملاً مشهود است. این نتایج را می‌توان این‌گونه تحلیل کرد که گسترش‌سازی یکی از ابزارهای مهم در برابر داده‌های نویزی است.

در این مقاله [۹] نویسنده‌گان توانسته‌اند با استفاده از ترکیب شبکه عصبی و الگوریتم ژنتیک، تغییرات قیمت را برای روز آینده در بازار سهام را پیش‌بینی کنند.

شبکه عصبی در نظر گفته شده در این مقاله یک شبکه ۳ لایه شامل ۱۸ نورون در قسمت ورودی، تعداد

Year	BPLT		GALT		GAFD	
	Training	Holdout	Training	Holdout	Training	Holdout
1989	59.05	48.28	57.33	49.12	68.10	59.65
1990	62.23	49.15	59.23	56.90	66.95	60.34
1991	58.97	53.45	53.42	50.00	63.25	56.90
1992	61.02	51.72	60.17	44.83	66.95	58.62
1993	54.01	44.07	54.43	44.07	67.09	61.02
1994	62.45	64.41	61.18	59.32	63.29	62.71
1995	63.83	44.83	63.83	53.45	69.36	65.52
1996	61.28	60.35	61.70	50.00	64.26	67.24
1997	46.15	50.00	50.43	50.00	64.10	62.07
1998	55.98	51.72	56.84	48.28	64.53	62.07
Total	58.50%	51.81%	57.86%	50.60%	65.79%	61.70%

شکل ۳-۳: مقایسه نتایج [۸]

متغیری نورون در لایه مخفی و یک نورون در لایه خروجی است. هدف شبکه تعیین میزان بازده قیمت سهام است که یک مسئله از نوع رگرسیون می‌باشد.

نویسندهان برای بدست آوردن پارامترهای بهینه شبکه عصبی اقدام به بررسی ترکیب تعداد زیادی از پارامترها کرده‌اند.

شکل ۴-۲: نشان دهنده پارامترهای بررسی شده است.

Parameter	Meaning	Level(s)
<i>n</i>	The number of neurons in the hidden layer	10, 20, ... 100
<i>ep</i>	number of iterations	1000, 2000, ... 10000
<i>mc</i>	momentum constant	0.1, 0.2, ... 0.9
<i>lr</i>	value of learning rate	0.1

شکل ۴-۳: پارامترهای بررسی شده [۹]

نحوه ساخت کروموزوم‌ها: هدف استفاده از الگوریتم ژنتیک پیدا کردن وزن‌های بهینه است. درنتیجه هر کروموزوم بایستی شامل مقادیر وزن‌های بین نورون‌ها در لایه‌های مختلف باشد.

### فصل ۳. مروری بر ادبیات تحقیق

داده‌های این مقاله مربوط به سال‌های ۱۹۹۳-۲۰۱۳<sup>۳</sup> شاخص نیکی ۲۲۵ می‌باشد. این شاخص یکی از مهم‌ترین شاخص‌های بازار ژاپن است که ترکیب وزن‌داری از ۲۲۵ سهام مختلف در بازار ژاپن است. حدود ۷۰٪ از داده‌ها برای آموزش و از ۳۰٪ باقیمانده برای آزمون استفاده شده است. نویسنده‌گان برای نشان دادن کارایی روش خود، یکبار داده‌ها را فقط با استفاده از الگوریتم پس انتشار خط<sup>۴</sup> آموزش داده‌اند و بار دیگر ابتدا وزن‌ها را با استفاده از الگوریتم ژنتیک تقریب زده‌اند و سپس از الگوریتم پس انتشار خط برای آموزش استفاده کرده‌اند. برای محاسبه خط از میانگین مربعات خط<sup>۵</sup> استفاده شده است که در زیر فرمول آن را مشاهده می‌کنید.

$$\text{MSE} = \frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2 \quad (1-3)$$

در فرمول بالا  $N$  تعداد داده‌ها،  $y_t$  مقدار واقعی و  $\hat{y}_t$  مقدار خروجی الگوریتم است.

شکل ۳-۵ نتایج بدست آمده را بررسی و مقایسه می‌کند.

Models	BPNN best	BPNN average	GABPNN best	GABPNN average
MSE	0.0044	0.1077	0.0043	0.0090
CPU Time*	68		1080	1322

شکل ۳-۵: مقایسه نتایج [۹]

همانطور که مشاهده می‌شود بهترین نتیجه دوروش تغییر محسوسی را نشان نمی‌دهد اما در قسمت میانگین خط، تفاوت چشمگیر است. دلیل این امر می‌تواند این باشد که استفاده از الگوریتم ژنتیک احتمال گیر افتادن شبکه در کمینه محلی را کاهش می‌دهد، بنابراین میانگین خط تفاوت چشمگیری بین دوروش را نشان می‌دهد. مقاله [۷۶] نیز کاری شبیه به کاری است که نویسنده‌گان مقاله قبلی انجام داده‌اند. در این کار نیز از ترکیب شبکه عصبی و الگوریتم ژنتیک استفاده شده است منتها روند یادگیری قدری متفاوت است. مقاله [۱۰] نیز کاری مشابه است. مراحل یادگیری الگوریتم به شرح زیر است:

#### ۱. نرمال سازی داده‌ها

<sup>3</sup>Nikkei 225 index

<sup>4</sup>Error Backpropagation

<sup>5</sup>Mean Squared Error

## ۲. تشکیل جمعیت اولیه

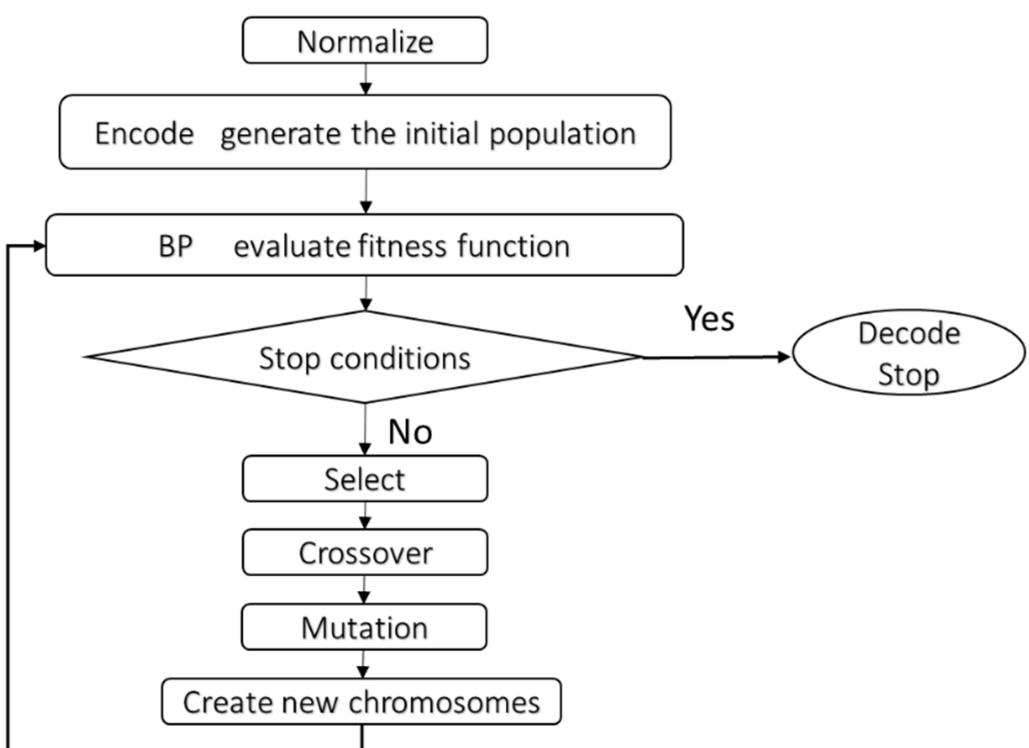
### ۳. آموزش شبکه عصبی و محاسبه تابع برازش

### ۴. انتخاب والدین و تشکیل فرزندان جدید

### ۵. بازنگری و جهش

### ۶. توقف در صورت برقراری شرط توقف

شکل ۶-۳ الگوریتم استفاده شده را به تصویر می کشد.



شکل ۶-۳: نحوه ترکیب شبکه عصبی و الگوریتم ژنتیک [۱۰]

شکل ۷-۳ پارامتر های استفاده شده را به تصویر می کشد.

تفاوت روند یادگیری این کار با کار اخیر در این است که در این مقاله از الگوریتم پس انتشار خطا در حین کار الگوریتم ژنتیک استفاده می شود در حالی که در مقاله قبلی پس از اتمام کار الگوریتم ژنتیک از آن استفاده می شد.

### فصل ۳. مروری بر ادبیات تحقیق

Variable	Value	Definition
n	10	number of neurons in the hidden layer of the ANN model
ep	3000	number of iterations for the hybrid model
mc	0.4	momentum constant of the ANN model
l	0.1	value of learning rate of the ANN model
pcre	0.7	crossover rate of the GA-ANN model
pmut	0.2	mutation rate of the GA-ANN model
popu	100	Initial population number of the GA-ANN model

شکل ۷-۳: پارامتر های استفاده شده [۱۰]

در ارزیابی نیز این دو کار شبیه به هم هستند و تفاوت شان در این است که در این کار، پس از بدست آوردن بازده بصورت رگرسیونی آن را با استفاده از فرمول زیر به درصد تبدیل می کنند:

$$\text{Hit ratio} = \frac{1}{N} \sum_{i=1}^N p_i \quad (2-3)$$

$$p_i = \begin{cases} 1 & (y_{i+1} - y_i)(\hat{y}_{i+1} - \hat{y}_i) > 0 \\ 0 & \text{else} \end{cases} \quad (3-3)$$

شکل ۸-۳ نتایج بدست آمده را با سایر کارهای مشابه بررسی و مقایسه می کند. در میان کارهای مقایسه شده بعضاً بورس سایر کشورها نیز مشاهده می شود که بدليل متفاوت بودن داده ها نتوان برای اثبات برتری الگوریتم پیشنهادی بطور کامل به آن ها استناد کرد اما تفاوت نتایج بدست آمده و دو کار انجام شده در بازار ژاپن می تواند برتری روش پیشنهادی را کاملاً نشان دهد.

### ۳-۲-۳ ترکیب شبکه عصبی و تحلیل مؤلفه های اساسی

تحلیل مؤلفه های اساسی از جمله الگوریتم های معروف در حوزه کاهش بعد و استخراج ویژگی خطی است که می تواند به همراه شبکه های عصبی استفاده شود [۳۳، ۷۷].

Studies	Methods	Stock market	Hit ratio (%)
Kim and Han [29]	GA feature discretization	Korea	61.70
Leung et al. [2]	Classification model	US, UK, Japan	68 (Nikkei 225)
Huang et al. [9]	SVM	Japan	75
Kara et al. [10]	BPNN	Istanbul	75.74
Our study	GA-ANN hybrid model	Japan	81.27

شکل ۳-۸: مقایسه نتایج [۱۰]

در این مقاله [۱۱] نویسنده‌گان توانستند با استفاده از ترکیب شبکه عصبی و تحلیل مؤلفه‌های اساسی، شاخص S&P 5000 را پیش‌بینی کنند.

داده‌های استفاده شده مربوط به ژوئن ۲۰۰۳ تا می ۲۰۱۳ است. هدف از این پیش‌بینی تعیین بالا یا پایین رفتن قیمت شاخص در روز بعد است. از ۷۰٪ ابتدایی داده‌ها به عنوان داده آموزشی استفاده شده است و از ۱۵٪ آخر به عنوان داده آزمون و ۱۵٪ باقیمانده هم به عنوان داده اعتبار سنجی<sup>۶</sup> استفاده شده است.

ویژگی‌های ابتدایی شامل ۶۰ مورد می‌شود. سپس از الگوریتم تحلیل مؤلفه‌های اساسی برای تولید ویژگی‌های مطلوب‌تر استفاده شده است. می‌دانیم که با کاهش بعد ممکن است واریانس داده‌های جدید نسبت به داده‌های ابتدایی کاهش پیدا کند. هر چه تعداد مؤلفه‌های در نظر گرفته شده بیشتر شود، واریانس داده‌ها نسبت به داده‌های اولیه نیز افزایش پیدا می‌کند. شکل ۹-۳ میزان واریانس حفظ شده با در نظر گرفتن مؤلفه‌های مختلف را نشان می‌دهد. با در نظر گرفتن ۱۱ مؤلفه حدود ۹۹٪ واریانس حفظ شده است.

PCs	Cumulative proportion						
1	0.930842	11	0.990272	21	0.998352	31	0.999981
2	0.947948	12	0.991921	22	0.998711	32	0.999987
3	0.961163	13	0.993275	23	0.999035	33	0.99999
4	0.9696	14	0.994235	24	0.999233	34	0.999993
5	0.974644	15	0.995146	25	0.999423	35	0.999996
6	0.978549	16	0.995968	26	0.999604	36	0.999999
7	0.981723	17	0.996543	27	0.999742	37	1
8	0.984181	18	0.997069	28	0.999841	38	1
9	0.986476	19	0.997575	29	0.999933	39	1
10	0.988453	20	0.99797	30	0.999962	40	1

شکل ۹-۳: میزان واریانس داده‌ها با در نظر گرفتن مؤلفه‌های مختلف [۱۱]

پس از کاهش بعد، داده‌های حاصل به یک شبکه عصبی پیشرو<sup>۳</sup> لایه با ۱۰ نورون در لایه مخفی و ۲ نورون در لایه خروجی داده می‌شود. خروجی این شبکه ۲ احتمال است. احتمال بالا رفتن قیمت و احتمال پایین رفتن قیمت که هر کدام عدد بیشتری را نشان بدهد به عنوان خروجی نهایی در نظر گرفته می‌شود.

<sup>6</sup>Validation

### فصل ۳. مروری بر ادبیات تحقیق

در نهایت میزان دقت پیش‌بینی در داده‌های آزمون، آموزشی و اعتبارسنجی استخراج شده است که در شکل ۱۰-۳ قابل ملاحظه است. بهترین نتیجه در تک تک بخش‌های آموزش، اعتبارسنجی و آزمون مربوط به زمانی است که ۳۱ مؤلفه داشته باشیم. این تعداد مؤلفه حدود ۹۹٪ از واریانس داده‌ها را حفظ می‌کند. البته زمانی که به اندازه ویژگی‌ها مؤلفه داشته باشیم (یعنی ۶۰ مؤلفه) نیز نتایج مطلوب است.

PCs	PCA			
	Training	Validation	Testing	Total
1	54.8	53.6	56.8	54.9
3	55.2	53.3	57.3	55.2
6	54.9	53.6	57.3	55
10	56.4	54.6	57.3	56.3
15	56.3	53.3	57.6	56
22	55.2	54.6	58.1	55.5
26	55.1	53.1	58.1	55.2
31	57.5	57.3	58.1	57.5
34	56.2	56	57.3	56.4
37	55	54.4	57	55.2
40	56.2	56.2	56.2	56.2
60	57.5	54.1	58.1	57.1

شکل ۱۰-۳: میزان دقت پیش‌بینی با درنظر گرفتن مؤلفه‌های مختلف [۱۱]

## ۳-۳ شبکه‌های عمیق

دسته دوم الگوریتم‌هایی که مورد بررسی قرار می‌گیرند، الگوریتم‌های مبتنی بر یادگیری عمیق هستند. این الگوریتم‌ها که در واقع شبکه‌های عصبی‌ای با بیش از یک لایه مخفی هستند، نسبت به شبکه‌های کم‌عمق توانایی بیشتری در استخراج ویژگی از داده‌های خام دارند. از جمله این الگوریتم‌ها می‌توان به شبکه پرسپترون چندلایه عمیق، شبکه‌های بازگشتی و شبکه‌های پیچشی اشاره کرد.

### ۱-۳-۳ شبکه پرسپترون چندلایه عمیق

شبکه‌های پرسپترون چندلایه عمیق به شبکه‌های پرسپترون چندلایه‌ای گفته می‌شود که بیش از یک لایه مخفی دارند. این نوع شبکه‌ها از جمله اولین شبکه‌های عمیقی بودند که برای پیش‌بینی بازارهای مالی توسط پژوهشگران

مورد استفاده قرار گرفته‌اند [۷۸].

در مقاله [۱۲] نویسنده‌اند با استفاده از شبکه عصبی پیش‌رو شاخص نزدک را پیش‌بینی کنند. هدف از این کار پیش‌بینی قیمت شاخص با در اختیار داشتن قیمت روزهای گذشته و همینطور در نظر گرفتن اینکه پیش‌بینی در چه روزی از هفته اتفاق می‌افتد است که در فرمول زیر نشان داده شده است:

$$y(k) = f(y(k-1), y(k-2), \dots, y(k-n), D(k)) \quad (4-3)$$

در فرمول بالا  $y(k)$  قیمت شاخص در روز  $k$ ,  $n$  تعداد روزهایی که برای پیش‌بینی در نظر می‌گیریم و  $D(k)$  نشان دهنده این است که در کدام روز هفته قصد پیش‌بینی داریم.

معیار اندازه‌گیری خطأ، ضریب تشخیص<sup>۴</sup> است که در فرمول زیر به آن اشاره شده است:

$$R^2 = 1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2} \quad (5-3)$$

در فرمول بالا  $R^2$  مقدار واقعی،  $\hat{y}$  مقدار خروجی الگوریتم و  $\bar{y}$  میانگین مقادیر واقعی است. داده‌های جمع‌آوری شده مربوط به ژانویه تا ژوئیه سال ۲۰۱۵ است. نویسنده‌اند پیش‌بینی خود را در دو حالت انجام داده‌اند. حالت اول استفاده از داده‌های ۴ روز گذشته است و حالت دوم استفاده از داده‌های ۹ روز گذشته است. همچنین برای پیش‌بینی از شبکه‌های عصبی با ساختار مختلف استفاده شده است. ساختار شبکه عصبی و همینطور میزان ضریب تشخیص را می‌توان در شکل‌های ۱۱-۳ و ۱۲-۳ دید. با یک تقریب کلی می‌توان گفت که پیش‌بینی با استفاده از داده‌های ۴ روز گذشته به مراتب نتایج بهتری را نسبت به ۹ روز نشان می‌دهد. پس می‌توان این نتیجه‌گیری را انجام داد که برای پیش‌بینی هر روز نبایستی خیلی نگاه به گذشته داشته باشیم. نکته بعدی که می‌توان از نتایج دریافت کرد این است که با افزایش یا کاهش لایه‌ها تغییرات منظمی را شاهد نیستیم و نمی‌توان گفت که تغییر ساختار شبکه از طریق کاهش یا افزایش لایه‌ها و همچنین تغییر در تعداد نورون‌های هر لایه منجر به بهبود نتایج خواهد شد یا خیر.

در این مقاله [۱۳] نویسنده‌اند با بهره‌گیری از یادگیری عمیق قیمت سهام شرکت اپل را در زمان بحران اقتصادی آمریکا، پیش‌بینی کنند.

داده‌های جمع‌آوری شده کوچکترین تغییرات قیمت را در بازه زمانی بین ۲ سپتامبر تا ۹ نوامبر ۲۰۰۸ را که

<sup>7</sup>Determination Coefficient

### فصل ۳. مروری بر ادبیات تحقیق

No.	Structure	R <sup>2</sup>			
		Train	Test	Validation	Total
1	2	0.8177	0.9616	0.9493	0.8692
2	5	0.9250	0.9188	0.9605	0.9264
3	5-5	0.9229	0.9724	0.8631	0.9212
4	5-10	0.2185	-0.0165	-0.2065	0.0838
5	10-10	0.9534	0.9602	0.6811	0.9344
6	10-20	0.9059	0.9758	0.9108	0.9263
7	40-40	0.9003	0.9639	0.9616	0.9264
8	50-100	0.9576	0.9324	0.9393	0.9483
9	100-200	0.9390	0.9466	0.9533	0.9393
10	200-300	0.9267	0.9642	0.8822	0.9276
11	20-40-20	0.9386	0.8917	0.9408	0.9267
12	20-50-20	0.9403	0.9417	0.9077	0.9374
13	50-100-50	0.6837	0.8108	0.7785	0.7326
14	20-40-40-20	0.8990	0.8445	0.8093	0.8739
15	10-20-20-10	0.8977	0.9602	0.9015	0.9109
16	10-20-20-20-10	0.9304	0.9341	0.9456	0.9329

شکل ۱۱-۳: نتایج پیش‌بینی با استفاده از داده‌های ۴ روز قبل [۱۲]

No.	Structure	R <sup>2</sup>			
		Train	Test	Validation	Total
1	5	0.8195	0.8480	0.8425	0.8274
2	10	0.8025	0.7113	0.8608	0.7882
3	5-5	0.8437	0.7762	0.8518	0.8280
4	5-10	0.8127	0.8554	0.8262	0.8180
5	20-20	0.8344	0.9116	0.7524	0.8389
6	20-30	0.8859	0.8690	0.8756	0.8707
7	50-100	0.8335	0.9028	0.8300	0.8292
8	200-300	0.8473	0.7899	0.8938	0.8506
9	300-400	0.8476	0.8685	0.7988	0.8435
10	20-50-20	0.8648	0.7937	0.8809	0.8372
11	20-40-20	0.9318	0.8827	0.9622	0.9262
12	50-100-50	0.8697	0.8367	0.8388	0.8552
13	20-40-40-20	0.7761	0.9205	0.8601	0.8188
14	20-50-50-20	0.8359	0.8462	0.8977	0.8443
15	10-20-20-20-10	0.8578	0.8348	0.6704	0.8107

شکل ۱۲-۳: نتایج پیش‌بینی با استفاده از داده‌های ۹ روز قبل [۱۲]

شامل ۱۴، ۸۳۹، ۳۹۴ مشاهده است را شامل می‌شود. برای ساماندهی داده‌ها آن‌ها را در هر دقیقه دسته‌بندی کرده‌اند که نتیجه آن ۱۹۱۱۰ سری زمانی می‌شود.

قبل از معرفی ویژگی‌هایی که استفاده شده‌اند نیاز است تا برخی مفاهیم تعریف شوند:

<sup>۸</sup> لگاریتم بازده

$$R = \ln \frac{p_t}{p_{t-1}} = \ln p_t - \ln p_{t-1} \quad (6-3)$$

در فرمول ۶-۳  $p_t$  نشان دهنده قیمت پایانی در دقیقه  $t$  است.

لگاریتم شبه‌بازده<sup>۹</sup> : بیانگر اختلاف لگاریتم میانگین قیمت در دقایق متوالی است در حالی که لگاریتم بازده بیانگر اختلاف لگاریتم قیمت پایانی دقایق متوالی است.

شاخص روند<sup>۱۰</sup> : برای محاسبه این شاخص بر تمامی داده‌های آن دقیقه یک مدل خطی  $b$  می‌برازیم. عدد  $a$  بیانگر همین شاخص است که نشان دهنده روند قیمت در آن دقیقه است.

ویژگی‌های استفاده شده را می‌توان در ۴ دسته جای داد: دسته اول شامل ساعت و دقیقه کنونی می‌شود. دسته بعدی شامل  $n$  لگاریتم شبه‌بازده آخر می‌شود. دسته سوم از  $n$  انحراف استاندارد<sup>۱۱</sup> آخر قیمت‌ها تشکیل شده است و در دسته آخر  $n$  شاخص روند آخر وجود دارد. در مجموع  $2 + 3n$  ویژگی وجود دارد و  $n$  یکی از پارامترهایی است که تعیین می‌شود.

داده‌های جمع‌آوری شده به ۲ قسمت مساوی تقسیم می‌شود. ۱۵٪ آخر هر قسمت به عنوان داده آزمون انتخاب می‌شود و بقیه به عنوان داده آموزشی مورد استفاده قرار می‌گیرند.

شکل ۱۳-۳ گویای این مطلب است.

از دو معیار برای گزارش نتایج استفاده شده است. مقیاس اول میانگین مربعات خطأ است که در فرمول ۱-۲ معرفی شده است. مقیاس بعدی نیز در فرمول زیر معرفی شده است

$$\text{Directional Accuracy} = 100 \times \text{Hit ratio} \quad (7-3)$$

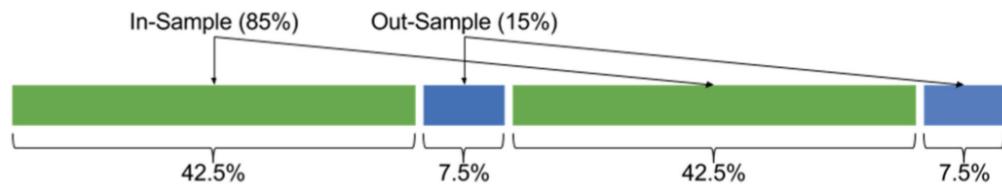
<sup>8</sup>log-return

<sup>9</sup>pseudo-log-return

<sup>10</sup>Trend Indicator

<sup>11</sup>Standard Deviation

## فصل ۳. مروری بر ادبیات تحقیق



شکل ۱۳-۳: نحوه تقسیم داده ها به داده آموزشی و داده آزمون

نحوه محاسبه Hit ratio در فرمول ۲-۳ توضیح داده شده است.

شکل ۱۴-۳ نتایج حاصل از استفاده از شبکه با پارامتر های مختلف را نشان می دهد. لازم به ذکر است که شبکه برای هر مجموعه از پارامترها چندین بار اجرا شده است و بهترین، بدترین و میانگین نتایج گزارش شده است. بهترین پنجره های زمانی ۲، ۳، ۴ دقیقه هستند و میزان دقت در این پنجره ها بین ۶۳٪ تا ۶۶٪ است. در مقاله قبلی مشاهده کردیم که در مقایسه بین پیش بینی با استفاده از داده های ۴ روز و ۹ روز قبل، نتایج پیش بینی ۴ روزه به مراتب بهتر بود. در این کار نیز مشاهده شد که اگر تعداد دقایق درنظر گرفته شده زیاد شود، دقت کاهش می یابد. در نتیجه با اطمینان بیشتری می توان گفته بازه زمانی داده هایی که برای پیش بینی درنظر گرفته می شود نباید خیلی بزرگ باشد.

Window Size	DNN Architecture	Maximum		Minimum		Mean		$\sigma$	
		MSE	DA (%)	MSE	DA (%)	MSE	DA (%)	MSE	DA (%)
2	8 8:6:4:3:1 1	0.07832	65.71328	0.06768	61.63236	0.07042	64.47506	0.00294	1.30354
3	11 11:8:6:4:2 1	0.07678	66.15492	0.06823	63.67759	0.07125	65.17794	0.00233	0.71620
4	14 14:11:8:5:2 1	0.09158	65.71328	0.07197	63.30659	0.07576	65.07847	0.00579	0.70564
5	17 17:13:10:6:3 1	0.10132	66.05024	0.07569	64.30565	0.08561	64.83949	0.00729	0.52326
6	20 20:16:12:7:3 1	0.10512	65.74816	0.07574	62.99267	0.08514	64.91105	0.00874	0.83872
7	23 23:18:13:9:4 1	0.10383	65.63154	0.08251	63.22400	0.08929	64.43475	0.00634	0.73761
8	26 26:20:15:10:5 1	0.09873	65.60865	0.07813	63.41123	0.08754	64.40879	0.00648	0.68195
9	29 29:23:17:11:5 1	0.09020	65.49197	0.07628	63.78227	0.08437	64.59874	0.00475	0.45499
10	32 32:25:19:12:6 1	0.10250	65.50401	0.07400	63.09731	0.08476	64.55877	0.00829	0.76088
11	35 35:28:21:13:6 1	0.10565	65.24773	0.07702	62.90997	0.08537	64.25680	0.00892	0.83953
12	38 38:30:22:15:7 1	0.09440	65.32961	0.07746	63.69026	0.08698	64.46110	0.00562	0.62078
13	41 41:32:24:16:8 1	0.09442	64.61967	0.07437	61.89811	0.08491	63.79972	0.00618	0.96908
14	44 44:35:26:17:8 1	0.09833	65.32961	0.07849	62.39972	0.08781	64.25531	0.00635	0.94794
15	47 47:37:28:18:9 1	0.09871	64.86392	0.08201	61.86322	0.08916	63.64270	0.00531	1.13385

شکل ۱۴-۳: نتایج حاصل با درنظر گرفتن پارامترهای مختلف [۱۲]

در این مقاله [۱۴] نویسنده کان توانستند با استفاده از یادگیری عمیق شاخص بازار سهام سنگاپور را پیش بینی

کنند.

داده‌های استفاده شده مربوط به ژانویه ۲۰۱۷ تا ژانویه ۲۰۱۰ است. هدف از این پیش‌بینی تعیین قیمت شاخص در روز بعد است. از ۷۵٪ ابتدایی داده‌ها به عنوان داده آموزشی و از ۲۵٪ آخر به عنوان داده آزمون استفاده شده است.

از ۴۰ ویژگی به عنوان ورودی شبکه عمیق استفاده شده است که شامل قیمت آغازین، پایانی، حداکثر و حداقل قیمت در ۱۰ روز آخر منتهی به روز پیش‌بینی است.

شبکه عصبی استفاده شده یک شبکه ۵ لایه است که علاوه بر ۲ لایه ورودی و خروجی از ۳ لایه مخفی نیز بهره می‌برد. هر کدام از لایه‌های مخفی از ۱۰ نورون تشکیل شده‌اند.

نویسنده‌گان با استفاده از داده‌های تا روز  $1 - n$  قیمت را در روزهای  $n + 1, n + 2, \dots, n + 4$  پیش‌بینی کرده‌اند. به عبارت دیگر، قیمت در چند روز بعد نیز پیش‌بینی شده است. شکل ۱۵-۳ قیمت واقعی و پیش‌بینی شده در روز  $n$  با استفاده از داده‌های تا روز  $1 - n$  را نشان می‌دهد.



شکل ۱۵-۳: مقایسه قیمت پیش‌بینی شده و قیمت واقعی در روز  $n$  با استفاده از داده‌های تا روز  $1 - n$  [۱۴]

نویسنده‌گان برای ارزیابی از دو معیار استفاده کرده‌اند. معیار اول خطای مجدول میانگین مربعات است که در

t(days)	RMSE	MAPE
0	32.77	0.75
1	48.40	1.20
2	51.38	1.26
3	72.00	1.83
4	74.64	1.84

جدول ۱-۳: مقایسه خطای قیمت پیش‌بینی شده [۱۴]

۱-۳ معرفی شده است. معیار دوم خطای میانگین درصد مطلق خطای<sup>۱۲</sup> است که در فرمول زیر معرفی شده است:

$$\text{MAPE} = \frac{100}{N} \sum_{t=1}^N \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (8-3)$$

در فرمول بالا  $N$  تعداد داده‌ها،  $y_t$  مقدار واقعی و  $\hat{y}_t$  مقدار خروجی الگوریتم است.

جدول ۱-۳ میزان خطای پیش‌بینی را با معیارهای گفته شده نشان می‌دهد. نکته‌ای که نتایج نشان می‌دهد این است که با در اختیار داشتن داده‌های تا امروز، هرچه آینده دورتری را بخواهیم پیش‌بینی کنیم، نتایج ضعیفتر خواهد بود.

### ۲-۳-۳ یادگیری عمیق و الگوریتم‌های نمایش داده

از شبکه‌های عصبی می‌توان برای پیدا کردن نمایش بهتری از داده‌ها پیش از استفاده از یک الگوریتم پیش‌بینی استفاده کرد. برای این منظور، ماشین بولتزمن محدود و خودمرمزگذارها مناسب به نظر می‌رسند [۴۳، ۴۵]. در این مقاله [۱۹] نویسنده‌گان توانستند با بهره گیری از روش‌های مختلف نمایش داده‌ها، بازده سهام شرکت‌های مختلف را در بازار کوه جنوبی پیش‌بینی کنند.

داده‌های این کار مربوط به ژانویه ۲۰۱۰ تا دسامبر ۲۰۱۴ است. ۳۸ شرکت بزرگ که که اطلاعات آن‌ها از ابتدا تا انتها موجود بوده است برای پیش‌بینی انتخاب شده‌اند. از ۸۰٪ داده‌های جمع‌آوری شده برای آموزش و از ۲۰٪ باقیمانده برای آزمون استفاده شده است.

از جمله اهدافی که نویسنده‌گان در نظر داشتند، این است که کارایی روش‌های مختلف نمایش داده نظیر

<sup>12</sup>Mean Absolute Percentage Error

Representation	Number of features	Training Accuracy	Test Accuracy
RawData	380	0.6361	0.6201
PCA200	200	0.5780	0.5551
PCA380	380	0.6364	0.6175
RBM400	400	0.6079	0.5843
RBM800	800	0.6169	0.5928
AE400	400	0.5712	0.5341
AE800	800	0.6060	0.5692

جدول ۲-۳: میانگین دقت پیش‌بینی ۳۸ شرکت با استفاده از الگوریتم‌های مختلف [۱۹]

خودرمزگذار، ماشین بولتزمن محدود و تحلیل مؤلفه‌های اساسی را در برابر داده خام بررسی کنند. داده‌های اولیه شامل ۳۸۰ ویژگی است اما نویسنده‌گان با استفاده از الگوریتم‌های گفته‌شده تغییراتی در ابعاد ویژگی‌ها می‌دهند تا تأثیر آن‌ها را مورد بررسی قرار دهند.

اولین بررسی به صورت دسته‌بندی و با استفاده از الگوریتم رگرسیون لجستیک<sup>۱۳</sup> انجام شده است. دو دسته مختلف در اختیار داریم و برای هر نمونه از داده، احتمال تعلق به هر کدام از دسته‌ها محاسبه می‌شود. سپس آن دسته‌ای را که محتمل‌تر است، به عنوان برچسب آن داده انتخاب می‌شود. فرمول زیر نحوه محاسبه احتمال را نشان می‌دهد:

$$P(y|u_t, w, b) = \frac{1}{1 + \exp(-(w^T u_t + b)y)} \quad (9-3)$$

در فرمول بالا  $u_t$  ورودی جدید،  $w$  ضریب ویژگی‌ها و  $y$  نشان دهنده برچسب آن نمونه از داده است. برچسب هر نمونه نیز با توجه به میزان بازده آن تعیین می‌شود. بدین ترتیب که اگر بازده بیش از ۲۰٪ باشد داریم  $y = 1$  و در غیر این صورت داریم  $y = -1$ .

جدول ۲-۴ نشان دهنده میانگین دقت پیش‌بینی بر روی سهام ۳۸ شرکت است.

در قدم بعدی نویسنده‌گان از یک شبکه عمیق و یک شبکه عصبی برای پیش‌بینی میزان بازده استفاده کرده‌اند. در شبکه عمیق از دو لایه به عنوان لایه‌های مخفی استفاده شده است که لایه اول شامل ۲۰۰ نورون و لایه دوم شامل ۱۰۰ نورون است. همچنین از دو معیار مجدد میانگین مربعات خطأ<sup>۱۴</sup> و میانگین مطلق خطأ که در زیر

<sup>13</sup>Logistic Regression

Representation	Number of features	ANN	DNN
RawData	380	0.8357	0.8216
PCA380	380	0.8379	0.8230
RBM400	400	0.8376	0.8248
AE400	400	0.8374	0.8220

جدول ۳-۳: میانگین مجدور میانگین مربعات خطای ۳۸ شرکت [۱۹]

Representation	Number of features	ANN	DNN
RawData	380	0.5690	0.5900
PCA380	380	0.5675	0.5888
RBM400	400	0.5684	0.5927
AE400	400	0.5678	0.5899

جدول ۴-۳: میانگین میانگین مطلق خطای ۳۸ شرکت [۱۹]

معرفی شده است برای ارزیابی استفاده کردند.

$$\text{MAE} = \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t| \quad (10-3)$$

در فرمول بالا  $N$  تعداد داده‌ها،  $y_t$  مقدار واقعی و  $\hat{y}_t$  مقدار خروجی الگوریتم است.

جدول ۳-۳ میانگین مجدور میانگین مربعات خطای ۳۸ شرکت و جدول ۴-۳ میانگین میانگین مطلق خطای ۳۸ شرکت را نشان می‌دهد.

هدف نویسنده‌گان از نوشتمن این مقاله بررسی روش‌های مختلف نمایش داده بوده است. در مجموع نمی‌توان یک مدل را به عنوان مدل برتر برای نمایش داده معرفی کرد. نتایج دسته‌بندی و همینطور مجدور میانگین مربعات حکایت از این دارد که داده خام برای پیش‌بینی مناسب‌تر است در حالی که میانگین مطلق خطای چیز دیگری را نشان می‌دهد. نکته دیگر این است که شبکه عمیق در مقیاس مجدور میانگین مربعات عملکرد بهتری را نشان می‌دهد اما مقیاس میانگین مطلق خطای شبکه عصبی معمولی را بهتر می‌داند. در مجموع بهترین نتایج مربوط به داده خام و یا تحلیل مؤلفه اساسی با ۳۸۰ مؤلفه است.

### ۳-۳-۳ شبکه‌های بازگشتی

از دیگر شبکه‌های عمیق که در حوزه پیش‌بینی بازارهای مالی مورد استفاده قرار گرفته‌اند، می‌توان به شبکه‌های بازگشتی اشاره کرد. این نوع از شبکه‌ها بدلیل داشتن حافظه داخلی، برای داده‌های دنباله‌ای مناسب به نظر می‌رسند [۷۹، ۸۰، ۸۱]. در میان انواع مختلف شبکه‌های بازگشتی، شبکه حافظه کوتاه-مدت ماندگار محبوب‌تر از سایرین به نظر می‌رسد [۸۲، ۸۳، ۸۴، ۸۵، ۸۶].

در [۱۵] نویسنده توانسته است با استفاده از شبکه حافظه کوتاه-مدت ماندگار برخی از شرکت‌های حاضر در بازار نزدک<sup>۱۴</sup> را پیش‌بینی کند. مراحل کار به شرح زیر است:

۱. استخراج داده‌های خام

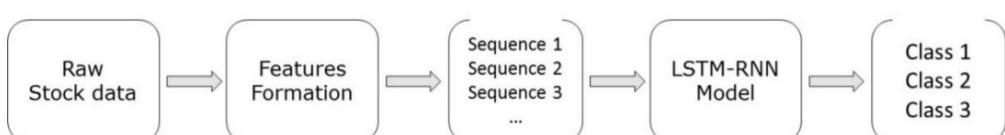
۲. ساخت ویژگی‌ها با توجه به شرکت‌های رقیب

۳. تشکیل دنباله‌ای از داده‌ها

۴. استفاده از شبکه عصبی

۵. پیش‌بینی

شکل ۱۶-۳ روند کار را نشان می‌دهد.



شکل ۱۶-۳: مراحل طی شده برای پیش‌بینی با استفاده از شبکه حافظه کوتاه-مدت ماندگار [۱۵]

داده‌های مورد استفاده جزئیات قیمت سهام را در هر ساعت کاری شامل می‌شود و هدف پیش‌بینی تغییر قیمت سهام به یکی از سه دسته کاهش قیمت، افزایش ۱٪-۵٪ و افزایش بیش از ۱٪ است.

۶ شرکت برای پیش‌بینی انتخاب شده‌اند و در کنار این شرکت‌ها به منظور افزایش دقیق پیش‌بینی برای هر کدام داده‌های مربوط به دو شرکت رقیب آن‌ها نیز بررسی شده است.

شکل ۱۷-۳ شرکت‌های انتخاب شده و رقبای آن‌ها را نشان می‌دهد.

<sup>14</sup>NASDAQ

### فصل ۳. مروری بر ادبیات تحقیق

Target Stock	Competitor A	Competitor B
BLK	State Street Corp.(STT)	Legg Mason Inc.(LM)
GOOGL	Apple Inc.(AAPL)	Facebook Inc.(FB)
QCOM	Microsoft Corp.(MSFT)	Texas Instruments Inc.(TXN)
XOM	Devon Energy Corp.(DVN)	Cabot Oil & Gas Corp.(COG)
IBM	Cisco Systems Inc. (CSCO)	HP Inc. (HPQ)
JPM	American Express (AXP)	Citigroup Inc. (C)

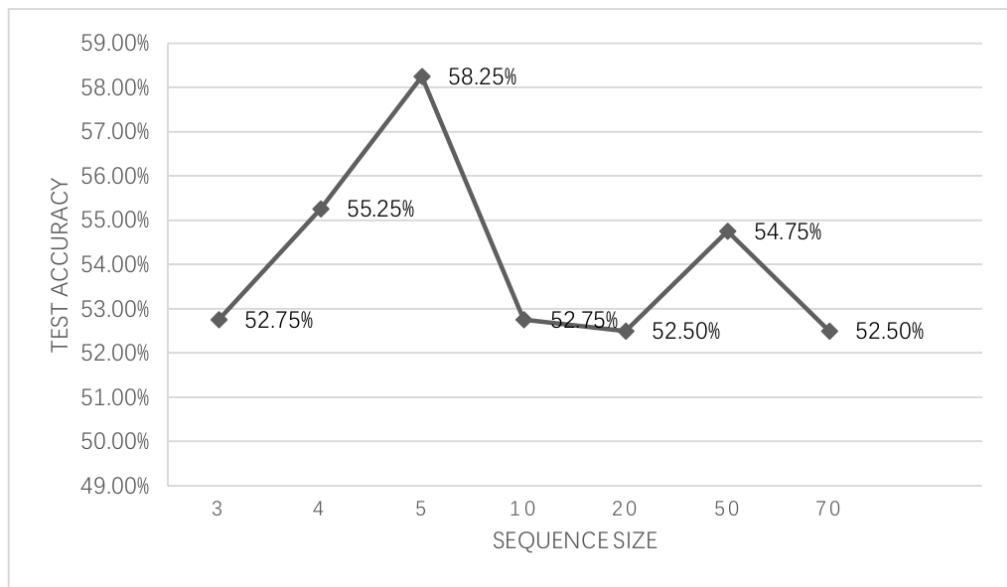
شکل ۳-۱۷: شرکت‌های انتخاب شده و رقبای آن‌ها [۱۵]

برای ساخت ویژگی‌ها نیز قیمت ۶۰ ساعت اخیر خود شرکت و دو رقیب آن، شاخص اس‌اندپی ۵۰۰<sup>۱۵</sup> شاخص نزدک و برخی موارد دیگر را در کنار هم قرار می‌دهیم تا مجموع ویژگی‌ها به ۳۶۹ برسد. در مرحله دنباله‌سازی تصمیم گرفته می‌شود که چند داده متوالی تحت عنوان یک دنباله به شبکه داده شود. نویسنده برای بدست آوردن تعداد بهینه داده‌ها در هر دنباله، اقدام به آزمایش تعداد مختلفی داده در دنباله کرده است و نتایج حاصل از آن که در شکل ۱۸-۲ نشان داده شده است عدد ۵ را بهترین گزینه می‌داند. و در نهایت خروجی کار را می‌توان در شکل ۱۹-۳ دید. میزان دقت برای یک شرکت زیر ۵٪ است که با در نظر گرفتن این نکته که برخلاف سایر موارد بررسی شده، در این کار سه دسته وجود دارد کاملاً توجیح می‌شود. از دلایلی که برای سه دسته در نظر گرفتن پیش‌بینی می‌توان بر شمرد، محاسبه کارمزد خرید و فروش سهام جدی‌تر به نظر می‌رسد. به عبارتی دیگر، بدست آوردن میزانی کمتر از ۱٪ سود برای پوشش کارمزد خرید و فروش سهام کفایت نمی‌کند.

در مقاله [۱۶] نویسنده‌گان توانستند با استفاده از شبکه حافظه کوتاه-مدت ماندگار به پیش‌بینی ۵ شرکت حاضر در بازار بزریل پردازنند.

داده‌های استفاده شده در این پژوهش مربوط به سال‌های ۲۰۰۸ تا ۲۰۱۵ بوده است. از دادگان خام که شامل بیشترین قیمت، کمترین قیمت، قیمت پایانی، قیمت آغازین و حجم معامله برای هر ۱۵ دقیقه بوده است، ۱۷۵

<sup>۱۵</sup>S&P 500



شکل ۱۸-۳: بررسی تعداد داده‌های مختلف در یک دنباله [۱۵]

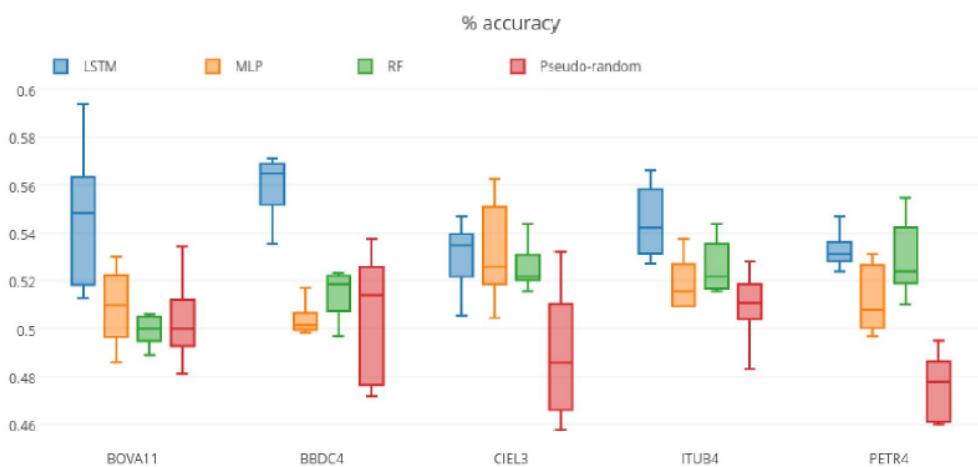
Stock Name	Test Accuracy
BLK	59.50%
GOOGL	49.75%
QCOM	58.25%
XOM	55.50%
IBM	54.50%
JPM	51.50%

شکل ۱۹-۳: نتایج نهایی حاصل از استفاده شبکه حافظه کوتاه-مدت ماندگار در پیش‌بینی ۶ شرکت [۱۵]

تکنیال اندیکاتور استخراج شده است. سپس، این تکنیکال اندیکاتورها بعلاوه داده‌های خام به شبکه مورد نظر داد شده‌اند تا از طریق لایه مخفی مت Shank از ۲۰ بلوک حافظه، تغییرات جهتی را برای ۱۵ دقیقه بعدی پیش‌بینی کنند. نویسنده‌گان برای مقایسه الگوریتم خود، نتایج حاصل را با الگوریتم‌های شبکه عصبی پرسپترون چندلایه، جنگل تصادفی و همینطور پیش‌بینی تصادفی بر حسب توزیع برچسب داده‌ها مقایسه کرده‌اند. شکل ۲۰-۳ نشان

## فصل ۳. مروری بر ادبیات تحقیق

دهنده عملکرد این الگوریتم‌ها است. همانطور که قابل مشاهده است، شبکه حافظه کوتاه-مدت ماندگار بهترین عملکرد را نسبت به سایر الگوریتم‌ها داشته است.



شکل ۳-۲۰: نتایج نهایی حاصل از استفاده شبکه حافظه کوتاه-مدت ماندگار در مقایسه با سایر الگوریتم‌ها [۱۶]

در مقاله‌ای دیگر [۴۴] نویسندهای از ظرفیت‌های شبکه حافظه کوتاه-مدت ماندگار بهره برداشت تغییرات جهتی شرکت‌های موجود در شاخص اس‌اند‌پی ۵۰۰ را برای فردا پیش‌بینی کنند.

برای پیش‌بینی وضعیت هر شرکت، دنباله‌ای نرمال شده از میزان بازده همان شرکت در ۲۴۰ روز گذشته به شبکه داده می‌شود که درواقع همان لایه ورودی شبکه را تشکیل می‌دهد. سپس، این ۲۴۰ ورودی توسط ۲۵ بلوک در لایه مخفی پردازش می‌شوند و درنهایت ۲ خروجی که نشان دهنده میزان احتمال بالا رفتن یا پایین آمدن هستند، خروجی نهایی کار را نشان می‌دهند.

داده‌های استفاده شده در این پژوهش مربوط به بازه سال‌های ۱۹۸۹ تا ۲۰۱۵ است. این بازه زمانی، خود به زیر بازه‌های ۴ ساله تقسیم می‌شود که داده‌های ۳ سال اول حدود ۵۰۰ شرکت حاضر در این شاخص برای آموختش شبکه و داده‌های یک سال آخر برای آزمون کارایی شبکه مورد استفاده قرار می‌گیرد.

پژوهشگران به منظور بررسی میزان کارایی الگوریتم پیشنهادی، آن را با جنگل تصادفی، رگرسیون لوجستیک و شبکه پرسپترون چندلایه عمیق مقایسه کرده‌اند. با دقیق حدود ۵۴٪/۳، شبکه حافظه کوتاه-مدت ماندگار بهترین عملکرد را در بین الگوریتم‌های آزمایش شده دارد درحالی‌که جنگل تصادفی، شبکه عمیق و رگرسیون لوجستیک بهترین دقیق معادل ۵۳٪/۸، ۵۳٪/۷ و ۵۲٪/۲ دارند.

### ۴-۳-۳ شبکه‌های پیچشی

شبکه‌های عصبی پیچشی، ابزارهای بسیار قدرتمندی در استخراج ویژگی‌های کارا از داده‌های خام محسوب می‌شوند. استخراج ویژگی‌های مناسب از دادگان بازار بورس، امری حیاتی برای پیش‌بینی دقیق رفتار آینده این بازار است. این نوع شبکه‌ها در ادبیات پیش‌بینی بازارهای مالی کاملاً جدید هستند [۳۰، ۸۸، ۸۹].

دی‌پرسیو و هنکار [۹۰] با استفاده از قیمت پایانی و سه شبکه عصبی پیچشی، حافظه کوتاه-مدت ماندگار و پرسپترون چندلایه، تغییرات جهت قیمت را برای بازه زمانی بعدی شاخص اس‌اندپی ۵۰۰ پیش‌بینی کردند.

در ابتدا برداری شامل قیمت پایانی ۳۰ روز متنه به روز پیش‌بینی با استفاده از تبدیل موجک پیش‌پردازش شده است و پس از آن به شبکه پیچشی داده شده است. سپس، دو لایه پیچشی که در میانشان دو لایه ادغام بیشینه قرار دارند، عمل استخراج ویژگی از سری زمانی ذکر شده را بر عهده داشته‌اند. دو لایه اتصال کامل نیز پس از لایه‌های پیچشی قرار گرفته‌اند تا ساختار شبکه تکمیل شود.

داده‌های استفاده شده در این پژوهش مربوط به سال‌های ۱۹۵۰ تا ۲۰۱۶ شاخص اس‌اندپی ۵۰۰ است. از ۸۰٪ داده‌ها برای آموزش و از ۲۰٪ باقیمانده برای آزمون الگوریتم‌های مطرح شده استفاده شد است. همچنین، نویسنده‌گان از ۱۰٪ داده‌های آموزشی برای پیدا کردن پارامترهای مناسب آموزش شبکه استفاده کردند.

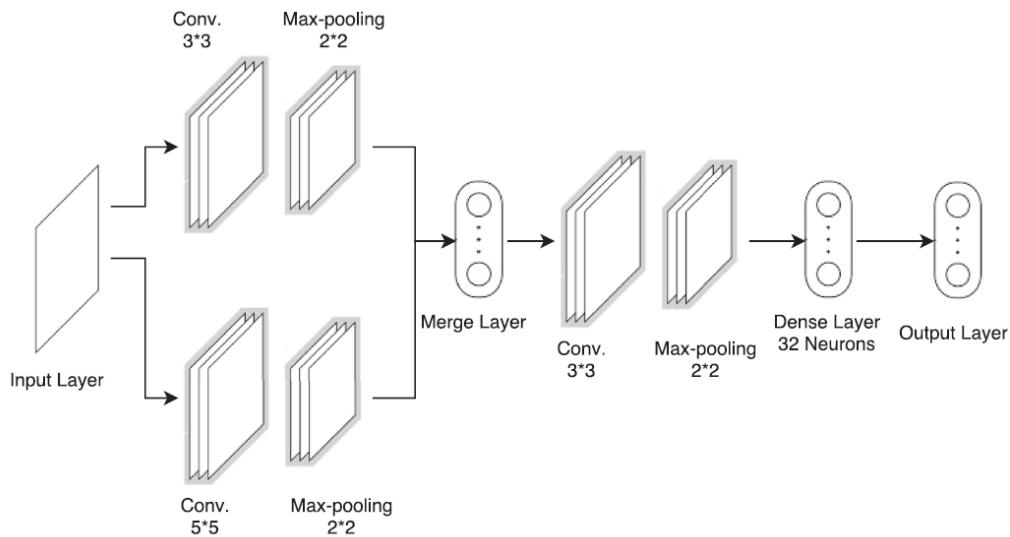
نتایج حاصل از مقایسه عملکرد الگوریتم‌های مختلف در پیش‌بینی تغییرات جهتی نشان می‌دهد که ترکیب شبکه پیچشی و تابع موجک با کسب دقیق حدود ۵۵٪ بهترین نتیجه را کسب کرده است در حالی که شبکه عصبی پیچشی، شبکه بازگشتی و شبکه پرسپترون چندلایه به ترتیب دقیقی برابر با ۶٪، ۲٪ و ۱٪ کسب کرده‌اند.

گوندوز و همکاران [۱۱] توانستند با بهره‌گیری از شبکه عصبی پیچشی به پیش‌بینی تغییرات جهتی سهام ۱۰۰ شرکت موجود در بورس استانبول برای ساعت بعد پیردازنند.

داده‌های استفاده شده در این پژوهش مربوط به ژانویه ۲۰۱۱ تا دسامبر ۲۰۱۵ می‌شود و از داده‌های تاسال ۲۰۱۴ برای آموزش و از داده‌های پس از آن برای آزمون استفاده شده است. ورودی به این شبکه ۷۵ ویژگی شامل تکنیکال اندیکاتورهای استخراج شده و سایر عملگرهای تاخیری سری زمانی است که از سهام این شرکت‌ها استخراج شده است.

شبکه پیچشی استفاده شده در این مقاله شامل ۸ لایه است: لایه ورودی، ۲ لایه موازی پیچشی به همراه ۲ لایه ادغام، یک لایه برای ترکیب ویژگی‌های استخراج شده از ۲ لایه پیچشی موازی، یک لایه پیچشی دیگر، لایه ادغام،

لایه اتصال کامل و درنهایت لایه خروجی. شکل ۲۱-۳ ساختار شبکه استفاده شده را نشان می‌دهد.

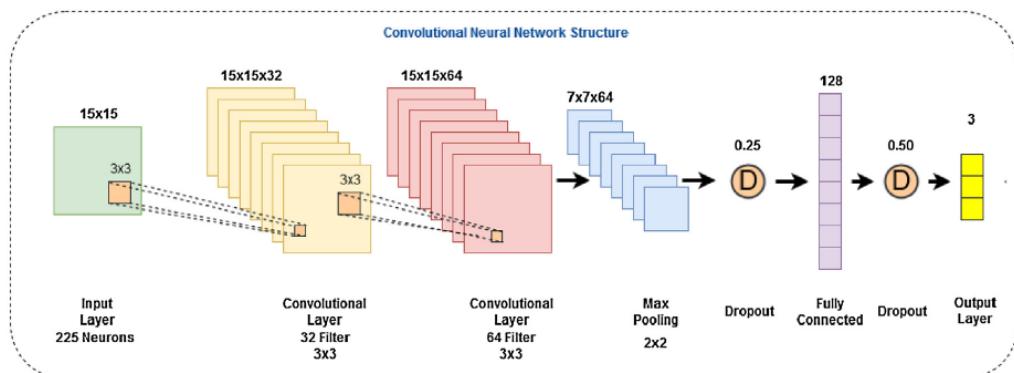


شکل ۲۱-۳: نمای گرافیکی از شبکه پیچشی استفاده شده [۱۷]

نویسندهان، خروجی شبکه ذکر شده را با الگوریتم رگرسیون لوجستیک از نظر macro-averaged F-measure مقایسه کردند و برای هر دو روش به عددی حدود ۵۴٪ رسانیدند. در تلاشی دیگر برای بهبود پیش‌بینی، ۷۵٪ ویژگی موجود از طریق روش خوشبندی سلسله‌مراتبی در خوشبهای مختلف قرار گرفتند. سپس، ویژگی‌هایی که در یک خوشه قرار داشتند، در زمان چینش ورودی شبکه پیچشی کنار یکدیگر قرار داده شدند. با این تغییر دقت پیش‌بینی به ۵۶٪ افزایش پیدا کرد. نتیجه کسب شده نشان می‌دهد که نحوه چینش ویژگی‌ها در لایه ورودی تأثیر مستقیم بر کیفیت پیش‌بینی دارد.

از شبکه‌های پیچشی می‌توان برای پیش‌بینی زمان خرید، فروش و یا نگهداری سهام استفاده کرد. سزر و ازبایقلو [۱۸] با استفاده از دادگان ۱۵ روز گذشته توانستند خرید، فروش و یا نگهداری سهام را برای روز بعد پیش‌بینی کنند.

در این تحقیق برای نمایش دادگان هر روز از ۱۵ تکنیکال اندیکاتور استفاده شده است؛ بنابراین، پس از وارد کردن ورودی که یک ماتریس  $15 \times 15$  بوده است به شبکه پیچشی با دو لایه پیچشی، یک لایه ادغام و یک لایه اتصال کامل (شکل ۲۲-۳)، شبکه نگهداری، خرید و یا فروش سهام را برای روز آینده پیش‌بینی می‌کند. داده‌های استفاده شده در این پژوهش مربوط به سال‌های ۲۰۰۲ تا ۲۰۱۷ است که به بازه‌های زمانی ۵ ساله تقسیم شده است. در هر بازه زمانی، از داده‌های ۴ سال اول برای آموزش شبکه و از داده‌های سال پنجم برای



شکل ۲۲-۳: نمای گرافیکی از شبکه پیچشی استفاده شده [۱۸]

آزمودن کارایی شبکه استفاده شده است.

استفاده از مدل پیشنهادی در این مقاله برای معامله سهام شرکت هایی که در بازار بورس آمریکا حاضر هستند، منجر به سود سالانه حدود ۱۳٪ در بازه زمانی سال های ۲۰۱۷-۲۰۰۷ شده است.

## ۴-۳ جمع بندی کارهای انجام شده

در این فصل مقالات متعددی بطور خلاصه مورد بررسی قرار گرفت. برای درک بهتر موارد گفته شده و مقایسه کارها، در جدولی تمامی موارد مطرح شده، جمع آوری شده است. در این جدول سال چاپ مقاله، بازار مورد پیش‌بینی قرار گرفته، الگوریتم‌های به کار گرفته شده، هدف پیش‌بینی، بازه زمانی جمع آوری داده و مقیاس‌های کارایی نتایج مطرح شده است. روش‌های مبتنی بر شبکه‌های عصبی و همینطور الگوریتم ژنتیک سالیان درازی است که مورد استفاده قرار می‌گیرند اما یادگیری عمیق موضوعی جدید و جذاب می‌باشد که در سال‌های اخیر و پس از موفقیت آن در سایر موضوعات نظری پردازش تصویر، صوت و متن مورد توجه افراد فعال در این حوزه قرار گرفته است. جدول ۵-۳ حاوی این جمع بندی است.

### فصل ۳. مروری بر ادبیات تحقیق

---

Year	Data Type	Target Output	Sampling Period	Method	Performance Measure
2011 [7]	Borsa Istanbul 100 Index	market direction	Jan-1997 to Dec-2007	ANN	accuracy
2000 [8]	KOSPI	market direction	Jan-1989 to Dec-1998	GA+ANN	accuracy
2016 [9]	Nikkei 225 index	stock return	Nov-1993 to Jul-2013	GA+ANN	MSE
2016 [76]	Nikkei 225 index	market direction	Jan-2007 to Dec-2013	GA+ANN	MSE Hit ratio
2016 [10]	Nikkei 225 index	market direction	Jan-2007 to Dec-2015	GA+ANN	MSE Hit ratio
2017 [11]	S&P 500 index	market direction	Jun-2003 to May-2013	PCA+ANN	accuracy
2016 [12]	NASDAQ	stock price	Jan-2015 to Jun-2015	DNN	determination coefficient
2016 [13]	AAPL Inc.	stock price	Sep-2008 to Nov-2008	DNN	MSE+ directional accuracy
2017 [14]	S&P 500 index	stock price	Jan-2010 to Jan-2017	DNN	RMSE MAPE
2017 [19]	Korea KOSPI 38 stock returns	stock returns	Jan-2010 to December-2014	data representation +DNN	RMSE MAE
2016 [15]	NASDAQ+S&P 6 stock returns	market direction	Aug-2013 to Nov-2015	LSTM	accuracy
2017 [16]	5 Stocks in Brazilian market	market direction	2008 to 2015	LSTM	accuracy
2018 [44]	Stocks in S&P 500 index	market direction	1989 to 2015	LSTM	accuracy
2016 [90]	S&P 500 index	market direction	1950 to 2016	MLP+LSTM wavelet+CNN	accuracy
2017 [17]	100 stocks in Borsa Istanbul	market direction	Jan 2011 to Dec 2015	CNN + Clustering	F-measure
2018 [18]	stocks in U.S. market	trading strategy	2012 to Dec 2017	CNN	trading simulation
2018 [18]	stocks in U.S. market	trading strategy	2012 to Dec 2017	CNN	trading simulation

جدول ۳-۵: مقایسه مقالات توضیح داده شده

## فصل ۴

### چارچوب‌های پیشنهادی

#### ۱-۴ مقدمه

این فصل به توضیح کامل و با جزئیات ۳ چارچوب پیشنهادی در این پایان‌نامه اختصاص داده شده است. هدف از ارائه چارچوب‌های پیشنهادی که CNNpred، HT-CNNpred و C-CNNpred نام دارند، پیش‌بینی تغییرات جهتی بازارها و شاخص‌ها برای روز آینده است. طراحی این چارچوب‌ها به شکلی انجام شده است که توانایی انطباق با هر بازاری را دارند.

به طور خلاصه این چارچوب‌ها عبارتند از:

- **CNNpred**

این چارچوب که خود شامل دو زیر چارچوب دو بعدی 2D-CNNpred و سه بعدی 3D-CNNpred است، قابلیت بهره‌گیری از طیف وسیعی از ویژگی‌ها و منابع نظیر تکنیکال اندیکاتورها، قراردادهای آتی، شاخص بورس سایر کشورها و تغییرات نرخ دلار به سایر ارزها را دارد. در حالی که ورودی به 2D-CNNpred یک ماتریس ۲-بعدی شامل ویژگی‌هایی نظیر تکنیکال اندیکاتورها و تغییرات آن‌ها در چند روز گذشته است، ورودی به 3D-CNNpred علاوه بر موارد مطرح شده شامل بازارهای دارای تأثیر بر بازار هدف نیز است.

- **HT-CNNpred: Hierarchical Transferable CNNpred**

#### فصل ۴. چارچوب‌های پیشنهادی

این چارچوب توسعه‌ای است بر چارچوب 2D-CNNpred. استفاده از آموزش لایه به لایه شبکه که منجر به پیدا کردن وزن‌های ابتدایی مناسب‌تر برای آموزش شبکه‌های می‌شود، مزیت HT-CNNpred نسبت به 2D-CNNpred 2 است. همچنین این چارچوب قادر است تا با استفاده از یادگیری انتقالی، به پیش‌بینی بازارهایی که در آموزش شبکه حضور نداشتند اقدام کند.

#### • **C-CNNpred: Clustering + CNNpred**

این چارچوب توسعه‌ای است بر چارچوب 3D-CNNpred. مزیت این چارچوب، استفاده کردن از روشی الگوریتمیک برای خوشبندی و انتخاب بازارهای دارای تأثیر بر بازار هدف است.

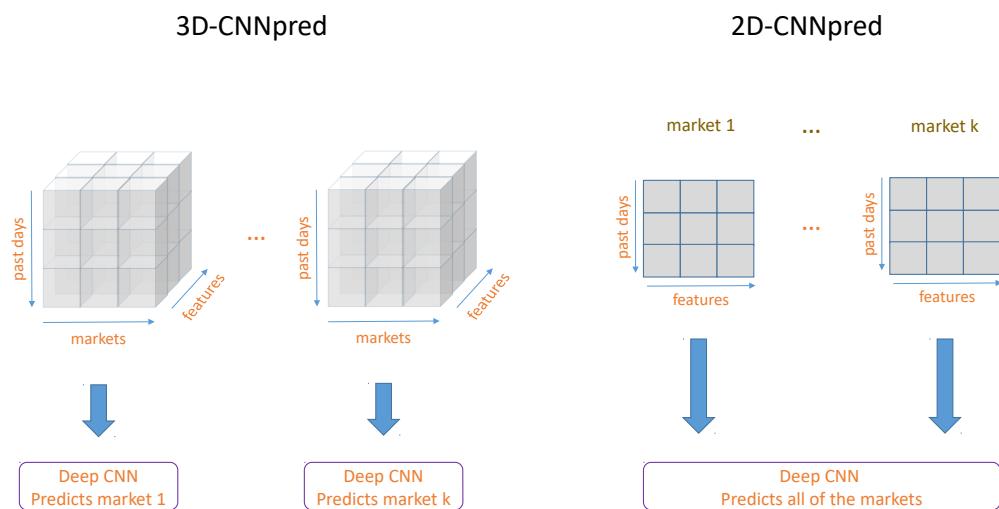
هرکدام از چارچوب‌های ذکر شده بطور مفصل در این فصل توضیح داده شده اند و نتایج این چارچوب‌ها در فصل ۵ بررسی خواهد شد.

## ٢-٤ چارچوب CNNpred

شبکه‌های پیچشی دارای پارامترهای مختلفی از جمله تعداد لایه‌ها، تعداد فیلترها، سایز فیلترها، نرخ حذف تصادفی و ... هستند که باستی هوشمندانه انتخاب شوند تا به نتیجه مطلوب برسند. اگرچه استفاده از فیلترهای پیچشی با سایز  $3 \times 3$  و یا  $5 \times 5$  بسیار رایج هستند، اما برای استفاده در پیش‌بینی بازارهای مالی، باستی سایز این فیلترها متناسب با ویژگی‌های مالی ای که قرار است پردازش بکنند باشد. بنابراین، در این قسمت توضیحات مفصلی درباره طراحی چارچوب CNNpred ارائه می‌شود که از ۴ مرحله نمایش داده، استخراج ویژگی‌های روزانه، استخراج ویژگی‌های بلند مدت و پیش‌بینی نهایی تشکیل شده است.

**نمایش داده:** CNNpred از داده‌ها و اطلاعات بازارهای مختلف برای پیش‌بینی بازار هدف استفاده می‌کند. همانطور که در مقدمه اشاره شد، این چارچوب خود به ۲ نوع 2D-CNNpred و 3D-CNNpred تشکیل می‌شود که هرکدام رویکردی متفاوت برای پیش‌بینی دارند. هدف از 2D-CNNpred ارائه یک مدل عمومی برای پیش‌بینی تمامی بازارهایی است که در آموزش شبکه مورد استفاده قرار گرفته‌اند. به بیانی دیگر، یک مدل با استفاده از داده‌های تمامی بازارهایی که قرار است پیش‌بینی در مورد آن‌ها انجام شود، آموزش می‌بیند و وظیفه پیش‌بینی تمامی این بازارها را برعهده می‌گیرد. در 2D-CNNpred تمامی اطلاعات مرتبط با یک بازار توسط یک ماتریس ۲-بعدی به شبکه پیچشی داده می‌شود و به همین دلیل 2D-CNNpred 3D-CNNpred نام دارد. بر این فرض

استوار است که برای پیش‌بینی هر بازار مدلی جداگانه نیاز است که می‌تواند از اطلاعات تمامی بازارها برای پیش‌بینی بازار هدف استفاده کند. درواقع، ایده پشت این مدل بر این مبنای است که مکانیزم رفتاری بازارها متفاوت از یکدیگر است و به همین دلیل نیاز به مدلی جداگانه برای هر بازار است. ورودی به چارچوب 3D-CNNpred یک تصور ۳-بعدی است که تمامی اطلاعات مرتبط با تمامی بازارها را ترکیب کرده و به عنوان ورودی به شبکه می‌دهد. در این مدل اگرچه ساختار مدل برای تمامی بازارها یکسان خواهد بود، اما داده‌ای که آموخته شبکه توسط آن انجام می‌شود، برای هر بازار متفاوت خواهد بود. شکل ۱-۴ نحوه نمایش داده‌های ورودی و همچنین بکارگیری آن‌ها در چارچوب CNNpred نشان می‌دهد.



شکل ۱-۴: نمای گرافیکی از نحوه نمایش و استفاده از داده در CNNpred

**استخراج ویژگی‌های روزانه:** هر روز از داده‌های موجود با استفاده از تعدادی ویژگی نظری قیمت آغازین و قیمت پایانی نمایش داده می‌شود. در روش‌های سنتی پیش‌بینی بازارهای مالی، تحلیل‌گران از این ویژگی‌های روزانه در قالب‌هایی نظری نمودار شمعی<sup>۱</sup> استفاده می‌کردند تا با پیدا کردن الگوهایی معنادار در آن‌ها به پیش‌بینی رفتارهای آینده بازارها بپردازنند. ایده اصلی طراحی اولین لایه پیچشی در CNNpred از این حقیقت ناشی می‌شود. در اولین مرحله از هر دو نوع مختلف CNNpred یک لایه پیچشی قرار دارد تا با استفاده از ویژگی‌های روزانه بازار، ویژگی‌های سطح بالاتری برای آن روز تولید کند.

**استخراج ویژگی‌های بلند مدت:** برخی دیگر از اطلاعات مفید برای پیش‌بینی بازار از طریق مطالعه رفتار

<sup>1</sup>Candlestick

#### فصل ۴. چارچوب‌های پیشنهادی

بازار در طولانی مدت بدست می‌آید. این مطالعه به یافتن الگوهایی در روند حرکتی بلند مدت بازارها کمک بسیاری می‌کند و از این جهت تاثیر بسزایی در بهبود کیفیت پیش‌بینی می‌گذارد. بنابراین، ترکیب ویژگی‌های چندین روز متوالی به منظور یافتن ویژگی‌هایی پیچیده‌تر که نشان‌دهنده روند حرکتی بازار در طولانی مدت است، امری مهم است. هر دو 2D-CNNpred و 3D-CNNpred لایه‌هایی دارند که ویژگی‌های روزانه را از لایه پیچشی اول که مسئول استخراج ویژگی‌های روزانه است تحویل می‌گیرند و ویژگی‌های سطح بالاتری که در برگیرنده چندین روز است از آن‌ها استخراج می‌کنند.

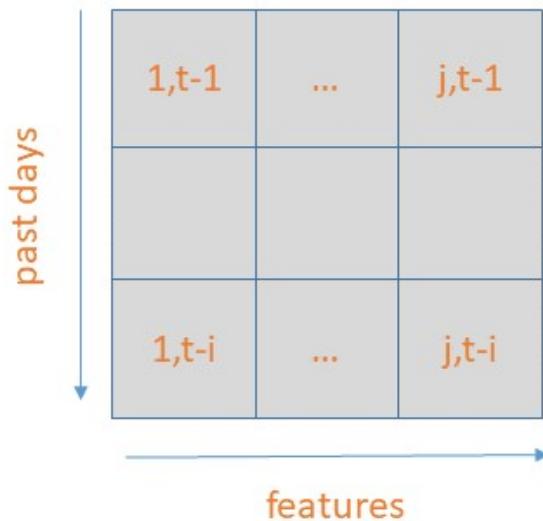
**پیش‌بینی نهایی:** در آخرین مرحله، ویژگی‌هایی که در طی مراحل قبلی استخراج شده‌اند به یک بردar-1- بعدی تبدیل می‌شوند. این بردار به یک لایه اتصال کامل داده می‌شود تا خروجی نهایی شبکه پیچشی را تولید کند.

در دو قسمت بعدی، بطور خاص نحوه طراحی هر کدام از چارچوب‌های 2D-CNNpred و 3D-CNNpred به نحوی که قابلت انطباق با داده‌های هر بازاری را پیدا کند، توضیح داده می‌شود. همچنان نحوه انطباق چارچوب‌ها با مجموعه داده‌ای که در این پایان‌نامه برای آزمودن کارایی آن‌ها مورد استفاده قرار می‌گیرد نیز ذکر می‌شود. مجموعه داده‌ای که قرار است برای آزمودن کارایی CNNpred مورد استفاده قرار بگیرد شامل ۵ شاخص اصلی بازار آمریکا می‌شود. هر شاخص ۸۲ ویژگی دارد که این بدین معنی است که هر روز از یک شاخص با ۸۲ ویژگی نمایش داده می‌شود. برای پیش‌بینی هر روز، از داده‌های ۶۰ روز گذشته استفاده می‌شود.

### **2D-CNNpred ۱-۲-۴**

**نمایش داده:** همانطور که اشاره شد، ورودی به 2D-CNNpred یک ماتریس ۲-بعدی است. ابعاد این ماتریس به تعداد ویژگی‌هایی که در هر روز برای بازار موجود است و همچنان تعداد روزهایی که قرار است از داده‌های آن‌ها برای پیش‌بینی استفاده شود وابسته است. بنابراین، اگر قرار باشد پیش‌بینی براساس داده‌های  $n$  روز گذشته انجام شود که هر کدام از این روزها با  $z$  ویژگی نمایش داده می‌شوند، ابعاد ماتریس ورودی  $z \times n$  خواهد بود. شکل ۲-۴ نمایی گرافیکی از ماتریس ورودی را نشان می‌دهد.

**استخراج ویژگی‌های روزانه:** برای استخراج ویژگی‌های روزانه در 2D-CNNpred از فیلترهای پیچشی با سایز تعداد ویژگی‌های روزانه  $\times 1$ ، استفاده می‌شود. هر کدام از این فیلترها تمامی ویژگی‌های روزانه را پوشش می‌دهند و از آن‌ها برای ساختن یک ویژگی سطح بالاتر استفاده می‌کنند. با استفاده از این لایه، 2D-CNNpred

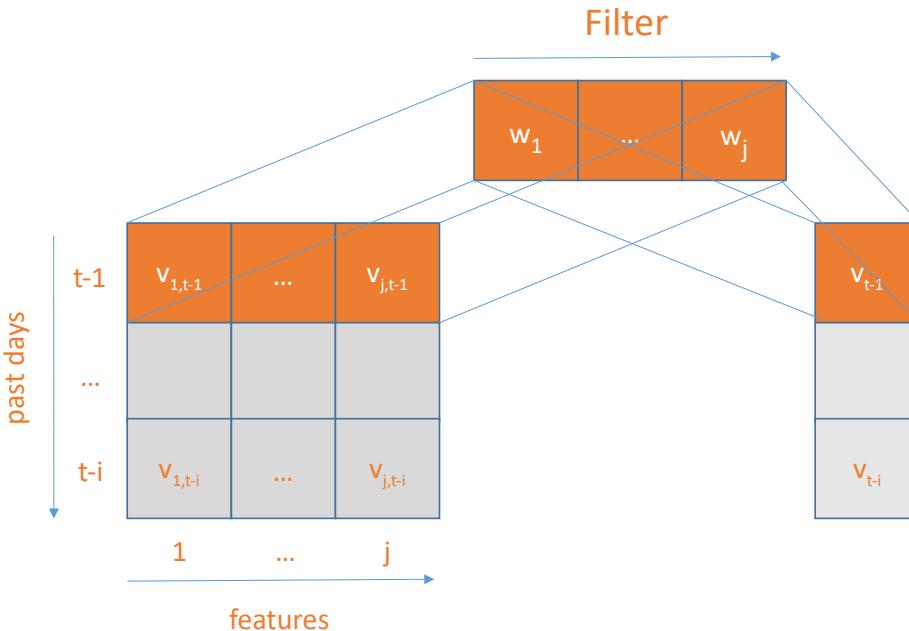


شکل ۲-۴: نمای گرافیکی از ماتریس ورودی در 2D-CNNpred

می‌تواند ترکیب‌های مختلفی از ویژگی‌های ابتدایی مالی بسازد. همچنین، شبکه این امکان را دارد تا با صفر کردن وزن متناظر با ویژگی‌های نامناسب در فیلترها، آن‌ها را کنار بگذارد. بنابراین، این لایه به عنوان یک الگوریتم استخراج / انتخاب ویژگی عمل می‌کند. شکل ۳-۴<sup>۳</sup> نحوه عملکرد این لایه را نمایش می‌دهد.

**استخراج ویژگی‌های بلند مدت:** در حالی‌که اولین لایه در 2D-CNNpred ۲۰ مسئول استخراج ویژگی از داده‌های روزانه است، لایه‌های بعدی وظیفه ترکیب ویژگی‌های استخراج شده در بازه زمانی بیش از یک روز را بر عهده دارند. همانند لایه اول، لایه‌های بعدی نیز از فیلترها برای تبدیل ویژگی‌های سطح پایین به ویژگی‌های سطح بالاتر استفاده می‌کنند. 2D-CNNpred از فیلترهایی با ابعاد  $1 \times 3$  در دومین لایه استفاده می‌کند تا ۳ روز متوالی را پوشش دهد. ایده استفاده کردن از فیلتری که ۳ روز را پوشش بدهد الهام گرفته از این حقیقت است که اکثر الگوهای معناداری نظری سه کلاع سیاه<sup>۲</sup> و سه خط شکست<sup>۳</sup> که با استفاده از نمودار شمعی تولید می‌شوند، رفتار بازارها را در ۳ روز متوالی در نظر می‌گیرند [۹۱، ۹۲، ۹۳]. استفاده از این الگوها در تحلیل تکنیکال کاملاً رایج است و به همین دلیل الگوبرداری از آن‌ها می‌تواند اطلاعات مفیدی برای پیش‌بینی در اختیار شبکه قرار دهد. سومین لایه، یک لایه ادغام با ابعاد  $1 \times 2$  است که در ادبیات شبکه‌های پیچشی کاملاً رایج است. به منظور استخراج ویژگی‌های پیچیده‌تر و طولانی‌مدت‌تر، 2D-CNNpred از یک لایه پیچشی دیگر با فیلترهایی به ابعاد

<sup>2</sup>Three Black Crows<sup>3</sup>Three Line Strike



شکل ۳-۴: اعمال فیلتری با ابعاد  $\times 1$  تعداد ویژگی‌های روزانه، بر ماتریس ورودی

$1 \times 3$  و یک لایه ادغام  $1 \times 2$  نیز بهره می‌برد.

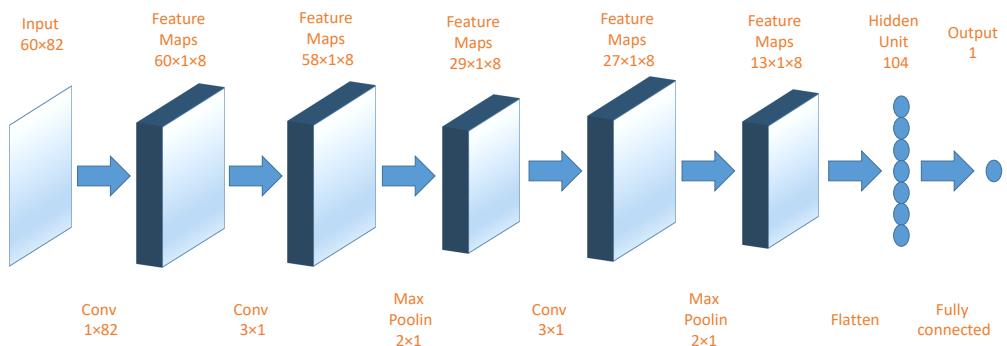
پیش‌بینی نهایی: ویژگی‌های تولید شده در آخرین لایه، تبدیل به برداری ۱-بعدی می‌شوند تا لایه اتصال کامل بتواند با استفاده از آن‌ها خروجی نهایی شبکه را پیش‌بینی کند.تابع فعال‌سازی آخرین لایه، سیگموید<sup>۴</sup> (رابطه ۱-۴) است که در نتیجه استفاده از آن، خروجی شبکه تبدیل به عددی در بازه  $[0, 1]$  خواهد شد. تعبیر این عدد می‌تواند احتمال افزایش قیمت سهام برای روز بعد باشد که اطلاعات مهمی را برای سرمایه‌گذار فراهم می‌کند. هرآنچه این عدد بالاتر باشد، سرمایه‌گذار اطمینان بیشتری نسبت به افزایش قیمت در روز آینده می‌یابد و درنتیجه می‌تواند پول بیشتری را سرمایه‌گذاری کند. با توجه به اینکه هدف این چارچوب پیش‌بینی تغییرات جهتی است که مسئله‌ای از نوع مسائل دسته‌بندی می‌باشد، لازم است تا خروجی با استفاده از مقدار آستانه  $5\%$  به دو برچسب  $1, 0$  گسسته‌سازی شود.

$$f(x) = \frac{1}{1 + \exp(-x)} \quad (1-4)$$

<sup>4</sup>Sigmoid

یک ساختار نمونه از **2D-CNNpred**: همانطوری که قبلاً اشاره شد، ورودی به شبکه در آزمایش ما شامل داده‌های ۶۰ روز قبل از روز پیش‌بینی است که هرکدام از این روزها با ۸۲ ویژگی نمایش داده می‌شوند. بنابراین، ماتریس ورودی دارای ابعاد  $82 \times 60$  خواهد بود. اولین لایه پیچشی از ۸ فیلتر  $8 \times 8 \times 1$  استفاده می‌کند. پس از آن، دو لایه پیچشی هرکدام از ۱۰۴ ویژگی است که از طریق یک لایه اتصال کامل تبدیل به پاسخ نهایی می‌شوند.

شکل ۴-۴ نشان‌دهنده ساختاری است که در بالا ذکر شد.

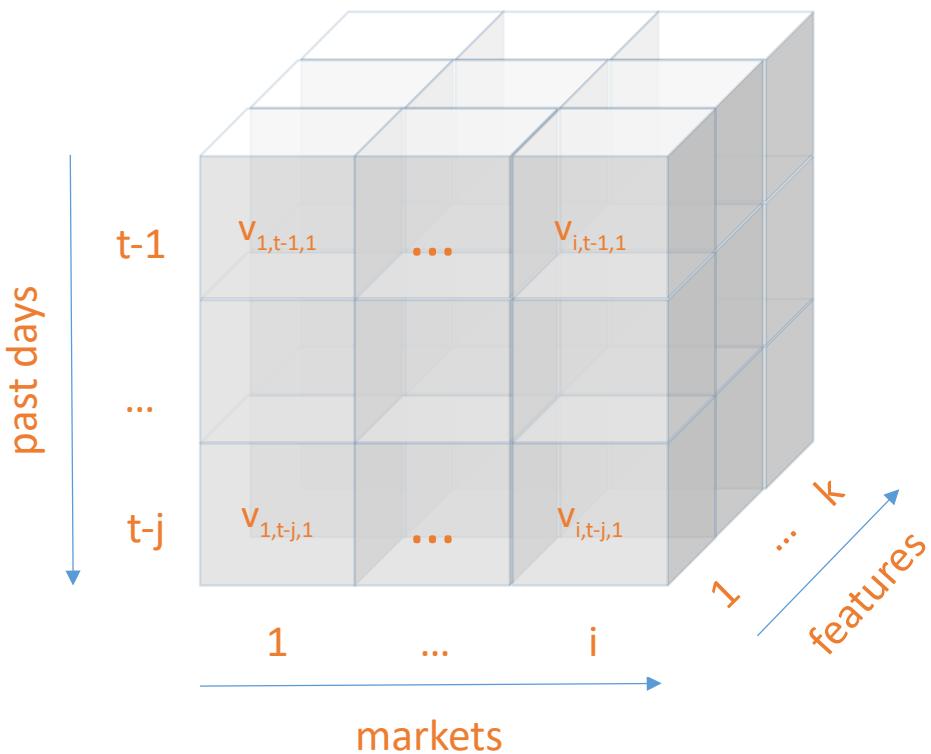


شکل ۴: نمایی گرافیکی از 2D-CNNpred

## 3D-CNNpred ۲-۲-۴

**نمایش داده:** 3D-CNNpred برخلاف 2D-CNNpred از یک تنسور ۳-بعدی برای نمایش داده‌ها استفاده می‌کند. دلیل این است که هر نمونه داده‌ای که استفاده می‌شود، شامل اطلاعاتی از سایر بازارها نیز است. بنابراین، ویژگی‌های روزانه مربوط به بازار، بازارهای دارای تأثیر بر بازار هدف و تغییرات ویژگی‌ها در روزهای قبل، سه بعد تنسور ورودی را تشکیل می‌دهند. بافرض اینکه مجموعه داده ما شامل  $t$  بازار مختلف باشد که هرکدام اشان با  $k$  ویژگی در هر روز نمایش داده می‌شوند و هدف، پیش‌بینی روز  $t$  بر حسب داده‌های زروری گذشته باشد، نحوه نمایش داده‌ها مطابق با شکل ۵-۴ خواهد بود.

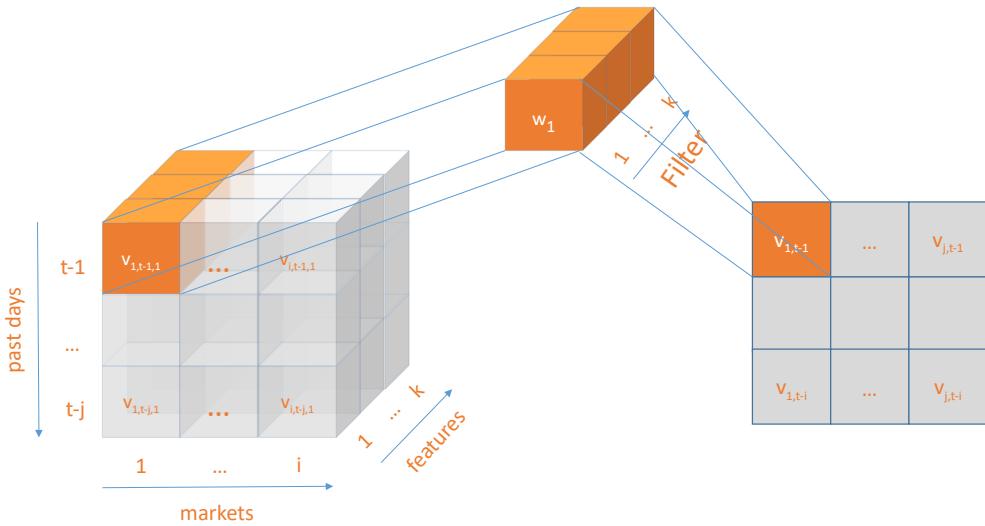
**استخراج ویژگی‌های روزانه:** ویژگی‌های اولیه هر بازار در عمق تنسور ورودی قرارگرفته است و به همین دلیل اولین لایه پیچشی شامل فیلترهایی با ابعاد  $1 \times 1$  است تا تمامی ویژگی‌هایی که نمایش‌دهنده هر روز از یک بازار هستند را پوشش دهد. شکل ۶-۴ نحوه عملکرد این فیلتر را به تصویر می‌کشد. این لایه مسئول ترکیب



شکل ۴-۵: نمای گرافیکی از تنسور ورودی در 3D-CNNpred

زیرمجموعه‌ای از ویژگی‌های ورودی به ویژگی‌هایی سطح بالاتر است. خروجی این لایه تنسوری خواهد بود که از نظر طول و عرض با تنسور ورودی به این لایه مطابقت می‌کند، اما عمق آن به اندازه فیلترهایی است که در این لایه مورد استفاده قرار می‌گیرند. همانند 2D-CNNpred، این لایه به عنوان یک الگوریتم استخراج / انتخاب ویژگی اولیه عمل می‌کند.

**استخراج ویژگی‌های بلندمدت:** علاوه بر ویژگی‌های روزانه، داده‌های ورودی به 3D-CNNpred شامل اطلاعاتی از بازارهای دیگر نیز می‌شود. همانند 2D-CNNpred، چهار لایه بعدی به استخراج ویژگی‌های بلندمدت و سطح بالاتر اختصاص دارند، با این تفاوت که در 3D-CNNpred 3 بجای یک بازار، با اطلاعات چندین بازار مواجه هستیم. بنابراین، ابعاد فیلترهای لایه دوم باید طوری طراحی شود که سایر بازارها را نیز شامل شود. از طرفی دیگر و به همان دلایلی که در مورد 2D-CNNpred ۲ کفته شد، پوشش ۳ روز متوالی از ویژگی‌ها مناسب به نظر می‌رسد. بنابراین، ابعاد فیلترهای لایه پیچشی دوم، تعداد بازارها  $\times$  ۳ خواهد بود. ۳ لایه بعدی نیز همانند 2D-CNNpred یک لایه ادغام بیشینه  $1 \times 2$ ، یک لایه پیچشی با فیلترهای  $1 \times 3$  و یک لایه ادغام بیشینه



شکل ۶-۴: اعمال فیلتری  $1 \times 1$  بر تنسور وردی

$1 \times 2$  خواهد بود.

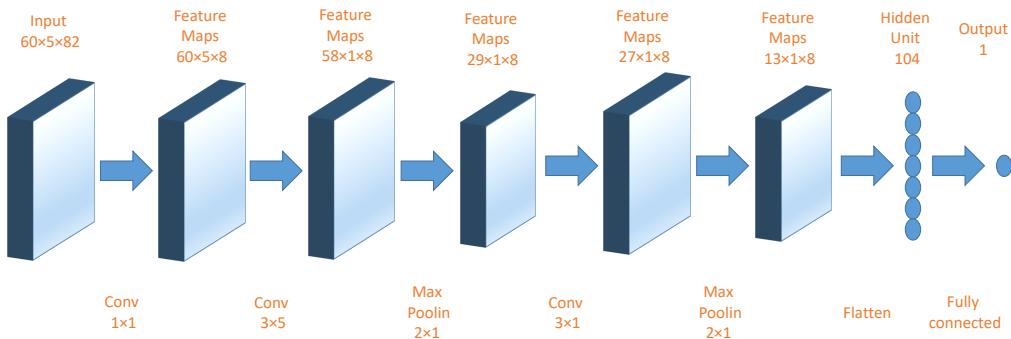
پیش‌بینی نهایی: همانند 2D-CNNpred، خروجی مراحل قبلی به یک بردار ۱-بعدی تبدیل می‌شود تا پس از عبور از لایه اتصال کامل، خروجی نهایی تولید شود.

یک ساختار نمونه از 3D-CNNpred: ورودی در آزمایش ما یک تنسور  $60 \times 60 \times 5$  با عمق ۸۲ است. اولین لایه پیچشی شامل ۸ فیلتر  $1 \times 1$  است. لایه دوم نیز از ۸ فیلتر  $5 \times 3$  استفاده می‌کند. پس از لایه ادغام بیشینه  $1 \times 2$ ، یک لایه پیچشی با ۸ فیلتر  $1 \times 3$  قرار دارد. پس از اعمال یک لایه ادغام بیشینه  $1 \times 2$  دیگر،  $10^4$  ویژگی نهایی به لایه اتصال کامل داده می‌شوند تا خروجی نهایی تولید شود. شکل ۷-۴ نشان‌دهنده ساختاری است که در بالا ذکر شد.

### ۳-۲-۴ ویژگی‌های استفاده شده برای هر شاخص

همانطور که اشاره شد، هدف ما ارائه چارچوبی برای پیش‌بینی تغییرات جهتی بازارهای مالی است. ما از CNpred برای پیش‌بینی ۵ شاخص اصلی آمریکا یعنی S&P500, NASDAQ, Dow Jones, NYSE, RUSSELL استفاده می‌کنیم. از ۸۲ ویژگی برای نمایش هر روز از این شاخص‌ها استفاده می‌کنیم که برخی از آن‌ها متعلق به یک شاخص هستند و برخی دیگر ویژگی‌های کمکی هستند که برای همه شاخص‌ها یکسان هستند. این

#### فصل ۴. چارچوب‌های پیشنهادی



شکل ۷-۴: نمایی گرافیکی از 3D-CNNpred

مجموعه غنی از ویژگی‌ها به ۸ دسته ویژگی‌های ابتدایی، تکنیکال اندیکاتورها، شاخص‌های مهم جهان، نرخ تبدیل دلار آمریکا به سایر ارزها، کالاهای بزرگ آمریکایی، قراردادهای آتی و سایر ویژگی‌های مفید دیگر قابل تقسیم است. اهمیت برخی از این ویژگی‌ها در این است که بازارها را بصورت مستقیم یا غیر مستقیم تحت تأثیر قرار می‌دهند و اهمیت برخی دیگر در آن است که می‌توانند نشانه‌هایی بر تغییرات کوتاه مدت بازارهای مالی باشند. در این پخش دسته‌های ذکر شده مورد بحث قرار می‌گیرند و جزئیات بیشتر در مورد آن‌ها در ضمیمه آن قابل دسترسی است.

**ویژگی‌های ابتدایی:** قیمت پایانی و روزی در هفته که قرار است پیش‌بینی انجام شود، ویژگی‌های ابتدایی استفاده شده برای شاخص‌ها هستند.

**تکنیکال اندیکاتورها:** تحلیل‌گران از تکنیکال اندیکاتورهایی که از داده‌های گذشته بازارها استخراج می‌کنند برای آنالیز رفتار آینده بازارها استفاده می‌کنند. این نوع ویژگی‌ها در پیش‌بینی بازارهای مالی بسیار رایج هستند.

**شاخص‌های مهم جهان:** بدلیل پدیده جهانی شدن اقتصاد، بازارهای مالی در سرتاسر جهان بر یکدیگر تأثیر می‌گذارند. اختلاف زمانی میان کشورهای مختلف این امکان را فراهم می‌سازد تا با مشاهده تغییرات بازارهای مالی کشورهای مختلف، اطلاعاتی درباره روند حرکتی بازار مالی کشور هدف بدست آورد [۹۴، ۹۵]. برای نمونه می‌توان به اثرگذاری بازار بورس کشورهایی مانند چین، ژاپن و کره بر بازار بورس آمریکا اشاره کرد.

**نرخ تبدیل دلار آمریکا به سایر ارزها:** شرکت‌های چندملیتی فراوانی وجود دارند که به سایر کشورها محصولات خود را صادر می‌کنند و یا نیاز خود را از آن‌ها تأمین می‌کنند. بنابراین، نرخ تبدیل دلار آمریکا به سایر ارزها نظیر دلار کانادا و یا یورو، بر میزان سود و زیان این شرکت‌ها اثرگذار است. زمانی که این تغییرات سود و زیان شرکت‌ها اعلام شود، میزان تقاضا برای سهام این شرکت‌ها و به طبع آن قیمت سهام آن‌ها دست‌خوش

تغییر می‌شود. سهام شرکت‌های داخلی نیز با تغییر وضعیت شرکت‌های چندملیتی دچار تغییر می‌شود. بنابراین، تغییرات نرخ ارز بر قیمت سهام شرکت‌ها و شاخص‌ها اثر می‌گذارد [۹۶، ۹۷].

**کالاهای:** عامل دیگری که می‌تواند بر پیش‌بینی وضعیت آینده بازارها تاثیرگذار باشد، قیمت کالاهایی نظیر طلا، نفت، نقره، گندم و غیره است. پژوهشگران نشان داده‌اند که ارتباطی میان قیمت کالاهای و بازار بورس وجود دارد [۹۸]. علاوه براین، کالاهای در کنار سهام شرکت‌ها، بخش مهمی از سبدهای سرمایه‌گذاری را تشکیل می‌دهند و درنتیجه تغییرات قیمت آن‌ها می‌تواند باعث تغییر در بازار سهام شود.

**شرکت‌های بزرگ آمریکایی:** در محاسبه شاخص‌های بازار سهام از سهام شرکت‌هایی که در آن شاخص‌ها حضور دارند استفاده می‌شود. هر شرکت به میزان سهمی که از بازار دارد در محاسبه شاخص کل اثر می‌گذارد. به عبارتی دیگر، شرکت‌های بزرگ تاثیر بسیار بیشتری نسبت به شرکت‌های کوچک بر شاخص‌ها می‌گذارند. اکسون موبیل<sup>۵</sup> و اپل<sup>۶</sup> نمونه‌هایی از این دست شرکت‌ها هستند.

**قراردادهای آتی:** قراردادهای آتی به قراردادهایی گفته می‌شود که در آن یک طرف از قرارداد قرار است تا خدمتی، کالایی، و یا سهامی را در آینده تحویل دهد. این قراردادها نشان دهنده میزان ارزش مورد انتظار کالاهای و یا سهام‌ها برای آینده هستند. سرمایه‌گذاران تمایل دارند تا سهام شرکت‌هایی را بخرند که قیمت مورد انتظار آن‌ها در آینده نسبت به قیمت فعلی آن‌ها بیشتر باشد. برای نمونه، قراردادهای آتی NASDAQ و سایر شاخص‌ها می‌توانند بر وضعیت فعلی شاخص S&P500 اثرگذار باشد.

**سایر ویژگی‌ها مفید:** بنابر پژوهش‌های مختلف، نرخ اوراق قرضه<sup>۷</sup> و برخی موارد مرتبط با آن نیز می‌توانند اطلاعات مفیدی برای پیش‌بینی بازارهای مالی در اختیار مان بگذارند [۹۹، ۱۰۰، ۱۰۱].

## ۳-۴ چارچوب HT-CNNpred

آموزش شبکه‌های عصبی نقشی اساسی در میزان کارایی آن‌ها ایفا می‌کند. اگر این آموزش به درستی انجام نگیرد، شبکه به بهینه محلی همگرا می‌شود و عملکرد آن افت خواهد کرد. وزن‌های اولیه شبکه‌های عصبی و بخصوص شبکه‌های عمیق تاثیر بسیاری در همگرایی شبکه و کیفیت پیش‌بینی آن‌ها دارند، بطوری که عدم انتخاب صحیح آن‌ها حتی ممکن است به واگرایی شبکه منجر شود. برای حل این مشکل، ما در چارچوب HT-CNNpred

<sup>5</sup>Exon Mobil Corporation

<sup>6</sup>Apple Inc.

<sup>7</sup>Treasury bill rate

#### فصل ۴. چارچوب‌های پیشنهادی

و در راستای بهبود چارچوب 2D-CNNpred یک روش سلسله‌مراتبی برای آموزش شبکه‌های عصبی پیشنهاد داده‌ایم. بطور دقیق‌تر، بعد از طراحی ساختار کلی شبکه پیچشی، آموزش آن از طریق یک رویه لایه به لایه انجام می‌شود. در هر مرحله از آموزش شبکه، وزن‌های اولیه از شبکه‌ای که در مرحله قبل آموزش دیده شده است، گرفته می‌شود. شبکه آموزش دیده از این طریق پیش‌بینی کننده عمومی نام دارد.

ایده مهم دیگری که در این چارچوب پوشش داده شده است، پیش‌بینی بازارهای جدیدی است که در آموزش شبکه پیچشی استفاده نشده‌اند از طریق تنظیم مناسب<sup>۸</sup> پیش‌بینی کننده عمومی برای بازار جدید است. این تنظیم می‌تواند در تمام شبکه و یا فقط لایه آخر شبکه صورت پذیرد. پیش‌بینی بازارهای در چارچوب HT-CNNpred به سه حالت زیر امکان‌پذیر است:

- پیش‌بینی کننده عمومی: پیش‌بینی بازارها با استفاده از شبکه پیچشی‌ای که به روش سلسله‌مراتبی آموزش داده شده است.
- پیش‌بینی کننده نیمه تنظیم‌شده: پیش‌بینی بازارهای جدید با استفاده از تنظیم لایه آخر پیش‌بینی کننده عمومی برای بازار جدید.
- پیش‌بینی کننده تمام تنظیم‌شده: پیش‌بینی بازارهای جدید با استفاده از تنظیم تمامی لایه‌های پیش‌بینی کننده عمومی برای بازار جدید.

HT-CNNpred شامل مراحل شبکه پیچشی ابتدایی، نمایش داده‌ها، آموزش سلسله‌مراتبی و یادگیری انتقالی است. درابتدا، ساختار شبکه پیچشی باستثنی متناسب با مسئله‌ای که قرار است این شبکه حل کند انتخاب شود. این مرحله نقش بسزایی در موفقیت چارچوب ایفا می‌کند. پس از این مرحله، لازم است تا داده‌های بازارهای مختلف به شکلی خاص که هم ویژگی‌های روزانه مرتبط با یک بازار و هم تغییرات آنها در روزهای قبل را پوشش می‌دهد، آماده‌سازی شوند. مرحله بعدی، آموزش شبکه پیچشی ابتدایی با استفاده از روش سلسله‌مراتبی است. در یادگیری انتقالی نیز پیش‌بینی بازارهای جدید با استفاده از تنظیم شبکه آموزش دیده انجام می‌شود. الگوریتم ۱-۴ نحوه کار چارچوب HT-CNNpred را نمایش می‌دهد.

<sup>8</sup>Fine tune

---

**Algorithm 4-1** HT-CNN-pred

---

**Data:** stock, new-stock, history

**Result:** results, half-results, full-results

initial CNN = define structure of CNN

representation = []

//Combine data of all the markets

**for**  $i = 1$  to # of stocks **do**

//Representation function returns an array of samples for each market

//each sample is a 2D matrix (# features  $\times$  history)

//history is the number of days used for prediction of each sample

temp-rep = Representation(stock[i], history)

Add temp-rep to representation

**end**

universal predictor = hierarchy training (initial CNN, representation)

//Prediction of test section of stocks used in training

results = predict(universal predictor, representation)

//Transfer learning

**for**  $i = 1$  to # of new stocks **do**

rep-stock = Representation(new-stock[i], history)

half-model = half-fine-tuned(universal predictor, rep-stock)

Add predict(half-model, res-stock) to half-results

full-model = fully-fine-tuned(universal predictor, rep-stock)

Add predict(full-model, res-stock) to full-results

**end**

---

return results, half-results, full-results

### ۱-۳-۴ شبکه‌پیچشی ابتدایی

شبکه‌های پیچشی برای توانمندی‌شان در استخراج ویژگی معروف هستند. با این حال، تنها استفاده از شبکه‌های پیچشی برای گرفتن نتایج خوب کافی نیست، چراکه این شبکه‌ها پارامترهای فراوانی نظیر تعداد لایه‌های مخفی، ابعاد فیلترهای پیچشی، تعداد فیلترهای پیچشی، نرخ حذف تصادفی و غیره دارند و تنها در صورت انتخاب درست و هوشمندانه آن‌ها می‌توان انتظار نتایج خوب را داشت. بنابراین، بسیار مهم است که ساختار و پارامترهای شبکه پیچشی به‌ نحوی انتخاب شوند که تناسب کامل با مسئله‌ای که قرار است این شبکه‌ها حل کنند، داشته باشند. شبکه‌پیچشی طراحی شده در این مرحله که شبکه پیچشی ابتدایی نام دارد، توسط یک روش سلسله‌مراتبی آموزش می‌بیند.

### ۲-۳-۴ نمایش داده‌ها

HT-CNNpred از داده‌های بازارهای مختلف برای آموزش و پس از آن پیش‌بینی بازارها استفاده می‌کند. هر بازار با تعدادی ویژگی روزانه نمایش داده می‌شود و برای پیش‌بینی هر روز، نیاز به داشتن داده‌های چندین روز قبل وجود دارد. بنابراین، همانند آنچه در مورد نمایش داده‌ها در 2D-CNNpred گفته شد، نمایش داده‌ها در HT-CNNpred نیز با استفاده از یک ماتریس ۲-بعدی انجام می‌پذیرد.

### ۳-۳-۴ آموزش سلسله‌مراتبی

روش‌های سنتی آموزش شبکه‌های پیچشی از الگوریتم‌های یادگیری بی‌نظرارت<sup>۹</sup> برای انتخاب وزن‌های اولیه شبکه استفاده می‌کنند. با درنظر گرفتن این موضوع که شبکه عصبی پیچشی یک الگوریتم از نوع بانظرارت است، استفاده از روشی که وزن‌های اولیه شبکه پیچشی را با توجه به برچسب داده‌ها انتخاب کند، هوشمندانه به‌نظر می‌رسد. بنابراین، ما شبکه عصبی پیچشی ابتدایی را به چندین زیرشبکه تقسیم می‌کنیم که درنهایت شبکه پیچشی ابتدایی را می‌سازند. سپس، هرکدام از این زیرشبکه‌ها را با استفاده از زیرشبکه آموزش دیده قبلی آموزش می‌دهیم. به عبارتی دیگر، وزن‌های نهایی زیرشبکه آموزش دیده قبلی به عنوان وزن‌های اولیه زیرشبکه بعدی مورد استفاده قرار می‌گیرند تا از همگرایی شبکه به بهینه محلی جلوگیری شود.

<sup>9</sup>Unsupervised learning

فرض کنیم که شبکه ما شامل لایه ورودی،  $n$  لایه مخفی قابل آموزش (لایه‌های پیچشی و اتصال کامل ونہ لایه ادغام) و لایه خروجی است. فاز آموزش شبکه در  $1 - n$  مرحله انجام می‌پذیرد. در مرحله  $i$ ، زیر شبکه شامل لایه ورودی، تمامی لایه‌ها تا  $1 + i$  امین لایه قابل آموزش از شبکه پیچشی ابتدایی و لایه خروجی خواهد بود. زیرشبکه آموزش دیده در مرحله  $1 - i$  وزن‌های اولیه لایه‌های مشابه در مرحله  $i$  را فراهم می‌کند. این فرآیند تا زمانی که تمامی  $1 - n$  مرحله طی نشده است، ادامه می‌یابد و زیرشبکه مرحله آخر که از نظر ساختار با شبکه پیچشی ابتدایی مطابقت می‌کند، مدل نهایی است. الگوریتم ۲-۴ فرآیند ذکر شده را توضیح می‌دهد. خروجی آموزش سلسله‌مراتبی پیش‌بینی کننده عمومی نام دارد چراکه قادر به پیش‌بینی تمامی بازارهای حاضر در داده آموزشی است.

---

#### Algorithm 4-2 Hierarchical Training

---

**Data:** initial CNN, data

**Result:** new-model

split data into train, validation, test

//save current model in each step to use it's weights for initial weights of next model

prev-model = new-model = []

**for**  $i = 1$  to # of layers **do**

new-model = initial CNN[input: $i$ ] + Output layer

**if**  $i > 1$  **then**

| initial weights new-model[input: $i - 1$ ] = weights prev-model[input: $i$ ]

**end**

initial weights new-model[ $i - 1$ :output] = Randomly

new-model = Train (new-model, train, validation)

prev-mode = new-model

**end**

return new-model

---

### ۴-۳-۴ یادگیری انتقالی

به منظور اندازه‌گیری میزان پایداری پیش‌بینی کننده عمومی، از آن برای پیش‌بینی بازارهای جدید استفاده خواهیم کرد. خوشبختانه چون پیش‌بینی کننده عمومی با استفاده از داده‌های تعداد زیادی بازار آموزش داده می‌شود، قادر خواهد بود تا ویژگی‌ها عمومی بازارهای مالی را استخراج کند. از آن‌جا که بازارهای جدید ممکن است رفتارهای مخصوص به خود را داشته باشند، تنظیم پیش‌بینی کننده عمومی برای این بازارها عاقلانه به نظر می‌رسد. این تنظیم برای بازارهای جدید می‌تواند به دو صورت تنظیم کامل و تنظیم نیمه صورت پذیرد.

• **تنظیم نیمه:** در این رویکرد تمامی لایه‌های پیش‌بینی کننده عمومی بجر آخرین لایه، مسئول استخراج ویژگی‌های عمومی بازارهای مالی هستند و تنها آخرین لایه برای بازار جدید تنظیم می‌شود.

• **تنظیم کامل:** گزینه دیگر برای هماهنگ کردن پیش‌بینی کننده عمومی با بازار جدید این است که تمامی لایه‌های پیش‌بینی کننده عمومی برای بازار جدید و با استفاده از داده‌های آن بازار تنظیم شوند.

### ۴-۴ چارچوب C-CNNpred

این چارچوب تلاشی برای بهبود چارچوب 3D-CNNpred 3D-CNNpred است. سایر بازارها برای پیش‌بینی بازار هدف را فراهم می‌کند، اما نحوه انتخاب این بازارها را پوشش نمی‌دهد و این کار بنابر تجربه تحلیل گران انجام می‌شود. C-CNNpred با استفاده از خوشبندی سایر بازارها سعی در ارائه راه حلی الگوریتمیک برای این موضوع دارد. این چارچوب نیز از چند مرحله خوشبندی بازارها، نمایش داده‌ها، استخراج ویژگی‌های روزانه، استخراج ویژگی‌های بلند مدت و پیش‌بینی نهایی تشکیل شده است.

### ۱-۴-۴ خوشبندی بازارها

از مواردی که برای ساخت داده‌های ورودی اهمیت فراوانی دارد، انتخاب بازارهای مشابهی است که برای پیش‌بینی بازار اصلی مورد استفاده قرار خواهند گرفت. سختی این امر از آن‌جا ناشی می‌شود که گزینه‌های موجود برای استفاده به عنوان بازارهای مشابه بسیار زیاد هستند و علاوه بر این، انتخاب بخشی از آن‌ها ممکن است باعث شود تا شبکه

#### **فصل ۴. چارچوب‌های پیشنهادی**

از در اختیار داشتن بخشی از اطلاعات که در بازارهای انتخاب‌نشده وجود دارد، محروم شود. به همین دلیل، اقدام به خوشبندی بازارهای کاندید با استفاده از الگوریتم  $k$  میانگین می‌کنیم تا بازارهای مشابهی دارند در خوشبندی مشابه قرار گیرند. سپس میانگین بازارهای موجود در یک خوش را به عنوان نماینده آن خوش برای قرارگیری در تنسور سه‌بعدی که در مرحله بعد توضیح داده می‌شود، انتخاب می‌کنیم.

#### **۲-۴-۴ نمایش داده‌ها**

C-CNNpred از داده‌های بازارهای مختلف برای آموزش و پس از آن پیش‌بینی بازارها استفاده می‌کند. هر بازار با تعدادی ویژگی روزانه نمایش داده می‌شود و برای پیش‌بینی هر روز، نیاز به داشتن داده‌های چندین روز قبل وجود دارد. بنابراین، همانند آنچه در مورد نمایش داده‌ها در 3D-CNNpred گفته شد، نمایش داده‌ها در C-CNNpred نیز با استفاده از یک تنسور ۳-بعدی انجام می‌پذیرد.

#### **۳-۴-۴ استخراج ویژگی‌های روزانه**

در این مرحله به همان دلایلی که در مورد 3D-CNNpred گفته شد، ویژگی‌های روزانه در عمق تنسور ۳-بعدی قرار دارند. بنابراین از یک لایه پیچشی با فیلترهای  $1 \times 1$  برای استخراج ویژگی‌های روزانه استفاده می‌شود.

#### **۴-۴-۴ استخراج ویژگی‌های بلند مدت**

لایه‌های این مرحله نیز همانند چارچوب 3D-CNNpred هستند با این تفاوت که به منظور منطبق کردن شبکه با بازارهایی که داده‌های کمتری از آن‌ها موجود است، از عمق شبکه کاسته و یک لایه ادغام بیشینه و یک لایه پیچشی از انتهای قسمت استخراج ویژگی‌های بلند مدت 3D-CNNpred، در این چارچوب حذف شده است. بدین ترتیب، استخراج ویژگی‌های بلند مدت در C-CNNpred توسط یک لایه پیچشی با فیلترهای تعداد بازارها  $\times 3$  و یک لایه ادغام بیشینه  $1 \times 2$  انجام می‌شود.

## ۵-۴-۴ پیش‌بینی نهایی

همانند 3D-CNNpred، خروجی مراحل قبلی به یک بردار ۱-بعدی تبدیل می‌شود تا پس از عبور از لایه اتصال کامل، خروجی نهایی تولید شود.

## ۵-۴ جمع‌بندی

این فصل به توضیح ۳ چارچوب پیشنهادی در این پایان‌نامه برای پیش‌بینی تغییرات جهتی بازارهای مالی اختصاص یافت.

CNNpred در زمرة اولین چارچوب‌های ارائه شده برای پیش‌بینی بازارهای مالی است که از شبکه‌های عصبی پیچشی استفاده می‌کند. این چارچوب به ۲ نوع 2D-CNNpred و 3D-CNNpred تقسیم می‌شود. 2D-CNNpred با ساخت یک مدل عمومی به پیش‌بینی تمامی بازارهایی که در آموزش شبکه استفاده شده‌اند می‌پردازد. از طرفی دیگر، 3D-CNNpred برای هر بازار مدلی جدید می‌سازد که به اطلاعات تمامی بازارها دسترسی دارد. از طیف وسیعی از ویژگی‌ها و منابع نظری تکنیکال اندیکاتورها، قراردادهای آتی، شاخص بورس سایر کشورها و تغییرات نرخ دلار به سایر ارزها برای آموزش این چارچوب استفاده شده است. همچنین، طراحی شبکه پیچشی و فیلترهای آن با درنظر گرفتن خصوصیات مالی این ویژگی‌ها انجام شده است.

دومین چارچوب پیشنهادی HT-CNNpred نام دارد. استفاده از آموزش شبکه که منجر به پیدا کردن وزن‌های ابتدایی مناسب‌تر برای آموزش شبکه‌های عمیق می‌شود، از مزایای این چارچوب است. به کمک این چارچوب امکان پیش‌بینی بازارهایی که در آموزش شبکه قرار نداشتند، با استفاده از یادگیری انتقالی فراهم شده است.

سومین چارچوب پیشنهادی C-CNNpred نام دارد که بهبود یافته 3D-CNNpred است. این چارچوب با استفاده از یک الگوریتم خوشبندی، بازارهای مشابه را دیگر خوشه قرار می‌دهد و پس از آن استفاده از نماینده هر خوشه را به عنوان بازارهای کمکی به چارچوب 3D-CNNpred می‌دهد.

## فصل ۵

### نتایج، بحث و ارزیابی

#### ۱-۵ مقدمه

در این فصل نتایج حاصل از چارچوب‌های ارائه شده در فصل قبل را خواهیم دید. این نتایج را با نتایج سایر الگوریتم‌های موجود در ادبیات پیش‌بینی مقایسه می‌کنیم.

ساختار این فصل بدین صورت است: در ابتدا اشاره‌ای به داده‌ها خواهیم داشت. شرح و بسط کامل داده‌ها در پیوست آآمده است. داده‌های استفاده شده برای چارچوب‌های مختلف، متفاوت است و به همین دلیل، داده‌های استفاده شده در هر چارچوب، بصورت جداگانه توضیح داده خواهند شد. نحوه پیش‌پردازش داده‌های نیز در مرحله بعد ذکر خواهد شد. پس از مباحث مربوط به داده‌ها نوبت به توضیح معیار ارزیابی، پارامترهای استفاده شده در شبکه و الگوریتم‌های پایه می‌رسد. در انتها نیز نتایج و تحلیل آن‌ها بحث می‌شود. آخرین مبحث این فصل نیز جمع‌بندی است.

#### ۲-۵ جمع‌آوری و پیش‌پردازش داده‌ها

در این پایان‌نامه ۳ چارچوب معرفی شده است که برای هر کدام از آن‌ها مجموعه داده متفاوتی جمع‌آوری شده است. بنابرین، توضیحات مربوط به هر بخش بصورت جداگانه آورده شده است. برجسب‌زنی داده‌ها طبق رابطه ۱-۵

انجام می‌شود که در این ربطه  $Close_t$  قیمت پایانی را در روز  $t$  نشان می‌دهد. نرمال‌سازی داده‌ها نیز فرآیندی است که برای داده‌های تمامی چارچوب‌ها اتفاق می‌افتد.

$$target = \begin{cases} 1 & Close_{t+1} > Close_t \\ 0 & \text{else} \end{cases} \quad (1-5)$$

### CNNpred ۱-۰-۲-۵

مجموعه داده استفاده شده برای آزمودن کارایی CNNpred، ۵ شاخص اصلی آمریکا یعنی-S&P 500, NASDAQ, Dow Jones, NYSE, RUSSELL ۱-۵ هستند. توضیحات بیشتر در مورد این شاخص‌ها در جدول ۳-۲-۴ آمده است. هر روز از هرکدام از این شاخص‌ها با ۸۲ ویژگی نمایش داده می‌شود که در ۸ دسته جای می‌گیرند و توضیحات آن‌ها در ۳-۲-۴ آمده است. جزئیات بیشتر نیز در آ ذکر شده است. تنها تکنیکال اندیکاتورها و ویژگی‌های ابتدایی برای هر شاخص منحصر به فرد است و سایر ویژگی‌ها میان شاخص‌ها مشترک است. داده‌های جمع‌آوری شده مربوط به بازه زمانی ۲۰۱۰ تا نوامبر ۲۰۱۷ است. از ۶۰٪ ابتدایی داده‌ها برای آموزش، ۲۰٪ بعدی برای اعتبارسنجی<sup>۱</sup> و از ۲۰٪ پایانی برای آزمون استفاده می‌شود.

### HT-CNNpred ۲-۰-۲-۵

برای آزمودن کارایی HT-CNNpred از دو مجموعه داده استفاده شده است. اولین مجموعه که برای آموزش پیش‌بینی کننده عمومی استفاده می‌شود، شامل اطلاعات قیمتی ۴۵۸ شرکت حاضر در شاخص S&P 500 است. از جمله این شرکت‌ها می‌توان به اکسون موبیل، اپل، مایکروسافت، جی‌پی مورگان چیس، جنرال الکتریک، جانسون و جانسون، ولز فارگو و آمازون اشاره کرد. علاوه بر این، از اطلاعات ۱۴ شاخص معروف جهان برای اندازه‌گیری میزان موفقیت HT-CNNpred در پیش‌بینی بازارهای جدید استفاده خواهیم کرد. اطلاعات بیشتر در این زمینه را می‌توان در جدول ۱-۵ یافت. هر روز از داده همانند CNNpred دارای ۸۲ ویژگی است که در ضمیمه آ قابل مشاهده است.

<sup>1</sup>Validation

توضیحات	نماد
S&P 500 شاخص قیمت سهام بزرگترین ۵۰۰ شرکت آمریکایی از نظر سرمایه S&P 500	S&P 500
شراکت ۳۰ شرکت بزرگ آمریکایی	Average Industrial Jones Dow
NASDAQ شاخص قیمت سهام شرکت‌های حاضر در بازار بورس NASDAQ	Composite NASDAQ
شاخص قیمت سهام شرکت‌های حاضر در بازار بورس نیویورک ۲۰۰۰ شرکت کوچک آمریکایی	Composite NYSE
شاخص قیمت سهام شرکت بزرگ آلمانی ۳۰ شرکت بزرگ آلمانی	RUSSEL 2000
شاخص قیمت سهام بورس کره جنوبی	DAX
شاخص قیمت سهام بازار بورس ترکیبو	KOSPI
شاخص قیمت سهام بورس فرانسه	Nikkei
شاخص قیمت سهام ۱۰۰ شرکت موجود در بورس لندن	CAC
شاخص قیمت سهام بازار هنگ کنگ	FTSE
شاخص قیمت سهام بازار شانگهای	HSI
شاخص قیمت سهام ۳۰ شرکت حاضر در بازار بورس بهبی	Composite SSE
شاخص قیمت سهام ۵۰ شرکت حاضر در بازار بورس ملی هند	BSE 30
شرکت اکسون موبیل	NIFTI 50
شرکت جی بی مورگان چیس	XOM
شرکت ابل	JPM
شرکت مایکروسافت	AAPL
شرکت جنرال الکتریک	MSFT
شرکت جانسون و جانسون	GE
شرکت ولز فارگو	JNJ
شرکت آمازون	WFC
	AMZN

جدول ۱-۵: توضیحات نمادها

داده‌های جمع‌آوری شده مربوط به بازه زمانی ژانویه ۲۰۱۰ تا نوامبر ۲۰۱۷ است. در اولین مجموعه داده، از داده‌های ژانویه ۲۰۱۰ تا جولای ۲۰۱۵ برای آموزش و اعتبارسنجی استفاده شده است. داده‌های جولای ۲۰۱۵ تا آپریل ۲۰۱۶ نیز برای آزمون استفاده شده است. بازه زمانی داده‌های آموزش و اعتبارسنجی مجموعه داده دوم نیز همانند مجموعه داده اول است، اما برای آزمایش پایداری این چارچوب در گذر زمان، بازه زمانی داده‌های آزمون از آپریل ۲۰۱۶ تا نوامبر ۲۰۱۷ تعیین شده است.

### C-CNNpred ۴-۰-۴

داده‌های استفاده شده برای آزمون کارایی C-CNNpred، شامل ۱۰ تکنیکال اندیکاتور (جدول ۲-۵) و قیمت پایانی برای ۲۷ شاخص مرتبط با بازار بورس تهران است. هدف از این پژوهش پیش‌بینی تغییرات جهت شاخص کل برای روز آینده است. اسامی شاخص‌های انتخابی درضمیمه آقابل مشاهده است. داده‌های استخراج شده مربوط به بازه زمانی اردیبهشت ۱۳۹۲ تا دی سال ۱۳۹۷ هستند. از ۶۰٪ ابتدایی داده‌ها برای آموزش، ۲۰٪ بعدی برای اعتبارسنجی و از ۲۰٪ پایانی برای آزمون استفاده می‌شود.

## فصل ۵. نتایج، بحث و ارزیابی

نحوه محاسبه	شماره	نحوه محاسبه	شماره
Relative Strength Index	۶	Simple Moving Average	۱
Williams %R	۷	Weighted Moving Average	۲
Chaikin A/D Oscillator	۸	Moving Average Convergence Divergence	۳
Stochastic %K	۹	Commodity Channel Index	۴
Stochastic %D	۱۰	Momentum	۵

جدول ۲-۵: ویژگی‌های استفاده شده برای نمایش شاخص‌های بازار تهران

### ۱-۲-۵ نرمال‌سازی داده‌ها

ویژگی‌های ورودی به شبکه پیچشی در این پژوهش دارای بازه‌های مختلفی هستند. به منظور بهبود کارایی الگوریتم‌های یادگیری ماشین، لازم است تا با انجام فرآیند نرمال‌سازی، ویژگی‌ها به بازه‌های تقریباً یکسان تبدیل شوند که در آن میانگین مقادیر هر ویژگی  $\bar{x}$  و انحراف معیار آن  $\sigma$  شود. برای این منظور استفاده از رابطه ۲-۵ مناسب است. در این رابطه  $x_{old}$  مقدار ابتدایی،  $x_{new}$  مقدار نرمال‌شده،  $\bar{x}$  میانگین مقادیر ابتدایی و  $\sigma$  انحراف معیار مقادیر ابتدایی است.

$$x_{new} = \frac{x_{old} - \bar{x}}{\sigma} \quad (2-5)$$

### ۳-۵ معیار ارزیابی

به منظور ارزیابی کارایی الگوریتم پیشنهادی و سایر الگوریتم‌ها نیاز به معیاری مناسب برای ارزیابی وجود دارد. صحت<sup>۲</sup> از جمله معیارهای پرکاربرد برای مسائل دسته‌بندی است. با این حال و به دلیل نامتقارن بودن داده‌ها، این معیار برای مسائل حل شده، نامناسب است و نیاز به معیاری است که تمامی دسته‌های موجود را به طور مناسب در ارزیابی لحاظ کند. به همین خاطر، از معیار Macro-Averaged F-Measure که در واقع همان میانگین F-measure برای هر دو کلاس صعود و نزول است، استفاده می‌کنیم [۱۰۲، ۱۷، ۳۰]. این معیار از ماتریس سردرگمی<sup>۳</sup> استخراج می‌شود و شامل دقت<sup>۴</sup> و یادآوری<sup>۵</sup> نیز است. نحوه محاسبه F-Measure در روابط ۳-۵ آمده است.

<sup>2</sup>Accuracy

<sup>3</sup>Confusion matrix

<sup>4</sup>Precision

<sup>5</sup>Recall

$$\begin{aligned}
 Confusion &= \begin{bmatrix} TP & FP \\ TN & FN \end{bmatrix} \\
 Precision &= \frac{TP}{TP + FP} \\
 Recall &= \frac{TP}{TP + FN} \\
 F - Measure &= \frac{2 \times Precision \times Recall}{Precision + Recall}
 \end{aligned} \tag{۳-۵}$$

دو دسته صعود و نزول را به ترتیب Positive و Negative لاحظ می‌کنیم و پیش‌بینی‌های درست و نادرست الگوریتم را با True و False نشان می‌دهیم. به این صورت است که الگوریتم ما چهار خروجی متفاوت (True) شمارش این حالات در هر پیش‌بینی الگوریتم و محاسبه کسرهای رابطه ۳-۵ معیارهای ارزیابی محاسبه می‌شوند. شمارش این حالات در هر پیش‌بینی الگوریتم و محاسبه کسرهای رابطه ۳-۵ معیارهای ارزیابی محاسبه می‌شوند. دو دسته صعود و نزول را بار دیگر به ترتیب Positive و Negative می‌نامیم و معیارهای ذکر شده را محاسبه می‌کنیم. حال از ۲ F - Measure محاسبه شده میانگین می‌گیریم تا Macro-Averaged F-Measure به دست آید.

## ۴-۵ پارامترهای شبکه

چارچوب‌های مطرح شده در این پایان‌نامه تماماً از شبکه‌های عمیق پیچشی برای پیش‌بینی استفاده می‌کنند. تعداد فیلترهایی که در لایه‌های پیچشی استفاده شده است، ۸ فیلتر برای هر لایه است. همچنین، توابع فعال‌سازی در تمامی لایه‌ها به غیر از لایه آخر که خروجی را مشخص می‌کند، تابع یکسوساز<sup>۶</sup> است. در لایه آخر نیز از تابع فعال‌سازی سیگموید<sup>۷</sup> استفاده شده است. همچنین، الگوریتم بهینه‌سازی Adam [۵۳] نیز وظیفه بروزرسانی وزن‌ها را بر عهده دارد. لازم به ذکر است که در میان کتابخانه‌های مختلفی که برای پیاده‌سازی شبکه‌های عصبی و همینطور یادگیری عمیق وجود دارند، در این پایان‌نامه از کتابخانه کراس<sup>۸</sup> [۱۰۳] استفاده شده است.

<sup>6</sup>Relu

<sup>7</sup>Sigmoid

<sup>8</sup>Keras

## ۵-۵ الگوریتم‌های پایه

به منظور بررسی کارایی چارچوب‌های پیشنهادی، ما نتایج حاصل از اجرای آن‌ها را با نتایج حاصل از اجرای الگوریتم‌های زیر مقایسه می‌کنیم:

**PCA+ANN** : اولین الگوریتم پایه، ترکیب شبکه پرسپترون چندلایه و تحلیل مولفه‌های اساسی است. در ابتدا الگوریتم تحلیل مولفه‌های اساسی بر داده‌های ورودی که شامل طیف زیادی از ویژگی‌ها هستند اعمال می‌شود تا ویژگی‌های بهتری استخراج شوند. سپس، یک شبکه کم‌عمق عمل پیش‌بینی را انجام می‌دهد [۹۹].

**Tech+ANN** : دومین الگوریتم پایه از ۱۰ تکنیکال اندیکاتور به عنوان ورودی به یک شبکه پرسپترون چندلایه استفاده می‌کند و این شبکه عمل پیش‌بینی را انجام می‌دهد [۴].

**CNN-cor** : سومین الگوریتم پایه یک شبکه پیچشی است [۱۷]. ابتدا با استفاده از خوشه‌بندی، ویژگی‌ها به دسته‌های مختلف تقسیم شده و ویژگی‌هایی که در یک خوشه قرار دارند، در ماتریس ورودی کنار یکدیگر قرار می‌گیرند. سپس، یک شبکه پیچشی با ساختاری مشخص بر ماتریس ورودی اعمال می‌شود.

**ANN** : یکی از پرکاربردترین و رایج‌ترین الگوریتم‌های موجود در این حوزه است. از یک شبکه پیشرو با یک لایه مخفی شامل ۵ نورون به عنوان یک الگوریتم پایه استفاده شده است [۹۹، ۱۰۴، ۷].

**SVM** : ماشین بردار پشتیبان نیز به کرات در حوزه پیش‌بینی بازارهای مالی دیده شده است و الگوریتم پایه بعدی ای است که برای مقایسه استفاده می‌شود [۱۰۵].

**LR**: رگرسیون لوجستیک نیز در زمرة الگوریتم‌های دسته‌بندی قرار می‌گیرد و از این حیث برای مقایسه با چارچوب پیشنهادی مناسب به نظر می‌رسد. همچنین در حوزه مالی نیز مشاهده شده است [۱۰۶].

**CNN** : الگوریتم پایه بعدی، یک شبکه پیچشی کاملاً مطابق با چارچوب CNNpred است، با این تفاوت که ورودی فقط شامل داده‌های بازار اصلی می‌شود و داده‌های بازارهای مشابه برای پیش‌بینی مورد استفاده قرار نمی‌گیرند.

Market \ Model	Technical	CNN-cor	PCA+ANN	2D-CNNpred	3D-CNNpred
S&P 500	0.4469	0.3928	0.4237	0.4914	0.4837
DJI	0.415	0.39	0.4283	0.4975	0.4979
NASDAQ	0.4199	0.3796	0.4136	0.4944	0.4931
NYSE	0.4071	0.3906	0.426	0.4885	0.4751
RUSSELL	0.4525	0.3924	0.4279	0.5002	0.4846

جدول ۳-۵: میانگین F-measure الگوریتم‌های مختلف

Market \ Model	Technical	CNN-cor	PCA+ANN	2D-CNNpred	3D-CNNpred
S&P 500	0.5627	0.5723	0.5165	0.5408	0.5532
DJI	0.5518	0.5253	0.5392	0.5562	0.5612
NASDAQ	0.5487	0.5498	0.5312	0.5521	0.5576
NYSE	0.5251	0.5376	0.5306	0.5472	0.5592
RUSSELL	0.5665	0.5602	0.5438	0.5463	0.5787

جدول ۴-۵: بهترین F-measure الگوریتم‌های مختلف

## ۶-۵ نتایج

این قسمت به ارائه نتایج اجرای چارچوب‌ها در کنار الگوریتم‌های پایه و همین‌طور تحلیل نتایج اختصاص دارد.  
نتایج هر کدام از ۳ چارچوب بصورت جداگانه آورده و تحلیل می‌شود.

### ۱-۶-۵ CNNpred

بدلیل اینکه یکی از الگوریتم‌های پایه از الگوریتم تحلیل مؤلفه‌های اساسی استفاده می‌کند، عملکرد این الگوریتم با تعداد مختلفی از مؤلفه‌های اساسی سنجیده می‌شود. به‌منظور یکسان کردن شرایط برای سایر الگوریتم‌ها و همین‌طور انجام آزمون‌های آماری، سایر الگوریتم‌ها نیز چندین بار با شرایط یکسان اجرا می‌شوند و در نهایت میانگین این اجراهای گزارش می‌شود.

جدول ۳-۵ میانگین نتایج اجرای الگوریتم‌های پایه و همین‌طور CNNpred را بر روی ۵ شاخص اصلی آمریکا از نظر F-measure نشان می‌دهد. همچنین، نتایج حاصل از بهترین اجرای الگوریتم‌ها نیز در جدول ۴-۵ ذکر شده است.

چارچوب‌های 3D-CNNpred و 2D-CNNpred برتری معنادار آماری نسبت به سایر الگوریتم‌های پایه

دارند. تفاوت میان چارچوب ما و Tech+ANN که تنها از تکنیکال اندیکاتورها استفاده می‌کند کاملاً واضح است. یک دلیل قانع کننده برای این عملکرد بدتر می‌تواند ناکافی بودن اطلاعاتی باشد که از ۱۰ تکنیکال اندیکاتور بدست PCA+ANN می‌آید. حتی اضافه کردن ویژگی‌های بیشتر و استفاده از الگوریتم تحلیل مولفه‌های اساسی که در CNNpred اتفاق افتاده است نیز منجر به بهبود قابل توجه نتایج نشده است. دلیل عدم موفقیت این دو الگوریتم این است که هر دو آن‌ها از یک شبکه عصبی کم عمق استفاده می‌کنند و شبکه‌های عصبی کم عمق توانایی محدودی برای استخراج ویژگی نسبت به شبکه عصبی پیچشی دارد. این مشاهده نشان می‌دهد که افزایش ویژگی‌ها و متنوع کردن آن‌ها به تنهایی و بدون داشتن الگوریتم مناسب استخراج ویژگی، منجر به بهبود پیش‌بینی کافی نمی‌شود. چارچوب CNNpred دو مزیت اساسی نسبت به این دو الگوریتم پایه دارد که عبارتند از: اول، استفاده از مجموعه غنی‌ای از ویژگی‌های مالی. دوم، استفاده از الگوریتمی مبتنی بر یادگیری عمیق که می‌تواند ویژگی‌های پیچیده‌تری را از داده‌های ابتدایی استخراج کند.

الگوریتم پایه بعدی CNN-cor است که بدترین نتیجه را میان سایر الگوریتم‌های پایه دارد. توانایی شبکه‌های عصبی کاملاً به انتخاب هوشمندانه پارامترهای آن بستگی دارد. این پارامترها به نحوی باید انتخاب شوند که شبکه طراحی شده متناسب با مسئله‌ای که قرار است حل کند بشود. با درنظر گرفتن این موضوع که CNN-cor و 2D-CNNpred هر دو از ویژگی‌های اولیه یکسان استفاده کرده‌اند و نحوه آموزششان نیز تقریباً یکسان بوده است، می‌توان به این نتیجه رسید که نتایج ضعیف CNN-cor مربوط به نحوه طراحی آن است. بطورکلی، ایده استفاده از فیلترهای  $3 \times 3$  و  $5 \times 5$  در هر مسئله‌ای که قرار باشد از شبکه‌های پیچشی استفاده شود، قدری ناپرداخته است. محبوب بودن این فیلترها در پردازش تصویر و بینایی ماشین‌الزماء این نتیجه را نمی‌دهد که این سبک از طراحی فیلترها به نتایج خوبی در پیش‌بینی بازارهای ملی نیز منتج شود. همانطوری که نتایج نشان می‌دهد، طراحی ساختار شبکه‌های پیچشی چالشی مهم در استفاده از این نوع شبکه‌ها برای پیش‌بینی در بازارهای مالی است.

## HT-CNNpred ۲-۶-۵

پیش از هر چیز لازم به ذکر است که شبکه عصبی پیچشی ابتدایی که برای آزمودن کارایی HT-CNNpred درنظر گرفته شده است، همان چارچوب 2D-CNNpred است. در این قسمت نتایج ۳ آزمایش گزارش شده است.

- مقایسه نتایج حاصل از اجرای HT-CNNpred و سایر الگوریتم‌های پایه.

Target stock \Model	Tech+ANN	PCA+ANN	CNN-cor	2D-CNNpred	U pred
Whole data	0.4425	0.4501	0.3569	0.4942	<b>0.5111</b>
XOM	0.4266	0.5032	0.3567	0.4969	<b>0.5109</b>
JPM	0.4408	0.4626	0.3689	0.4936	<b>0.5176</b>
AAPL	0.4615	0.4517	0.3670	0.4889	<b>0.5058</b>
MSFT	0.4507	0.4667	0.3416	0.4961	<b>0.5213</b>
GE	0.4077	0.4588	0.3462	0.4973	<b>0.5161</b>
JNJ	0.4290	0.4625	0.3641	0.4840	<b>0.5071</b>
WFC	0.4151	0.4395	0.3637	0.4949	<b>0.5102</b>
AMZN	0.4993	0.4376	0.3531	0.4962	<b>0.5114</b>

جدول ۵: میانگین F-measure الگوریتم‌های مختلف در پیش‌بینی ۴۵۸ شرکت

- بررسی تاثیر عمیق کردن شبکه بر کیفیت پیش‌بینی

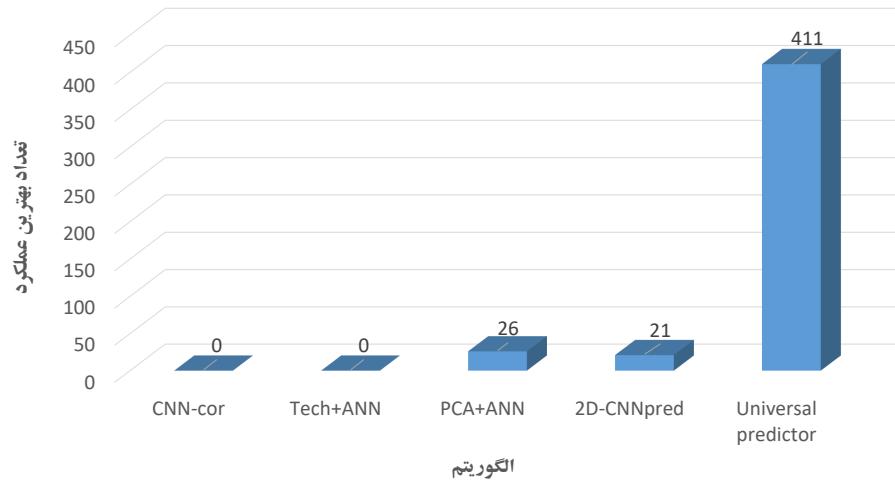
- بررسی میزان پایداری HT-CNNpred در گذر زمان از نظر پیش‌بینی بازارهای جدید.

بدلیل آنکه از داده‌های ۴۵۸ برای آزمون کارایی HT-CNNpred استفاده شده است، حتی یک بهبود کوچک در کیفیت پیش‌بینی نیز از نظر آماری معنادار است. در اینجا نیز به دلایل ذکر شده در قسمت قبل، الگوریتم‌ها چندین بار اجرا می‌شوند و میانگین نتایج آن‌ها گزارش می‌شود. پیش‌بینی بازارها در این چارچوب به ۳ طریق پیش‌بینی کننده عمومی pred U، پیش‌بینی کننده نیمه تنظیم شده pred H و پیش‌بینی کننده تمام تنظیم شده pred G انجام می‌شود.

اولین مشاهده، پیش‌بینی تغییرات جهتی سهام ۴۵۸ شرکت با استفاده از الگوریتم‌های پایه و همینطور -HT-CNNpred است. میانگین F-measure الگوریتم‌ها در جدول ۵-۵ گزارش شده است. همچنین، نتایج مربوط به ۸ شرکت بزرگ نیز ذکر شده است [۹۹]. علاوه بر این، تعداد دفعاتی که هر الگوریتم توانسته بهترین عملکرد را در پیش‌بینی تغییرات جهتی سهام شرکت‌ها کسب کند نیز در شکل ۱-۵ به نمایش گذاشته شده است.

آزمایش دوم به بررسی تاثیر عمیق کردن شبکه در کیفیت پیش‌بینی می‌پردازد. هنگام آموزش شبکه پیچشی ابتدایی به روش سلسله‌مراتبی، تعدادی زیرشبکه ایجاد می‌شود. زیر شبکه تشکیل شده در هر مرحله، از یک لایه عمق بیشتر نسبت به زیرشبکه مرحله قبل برخوردار است. هرکدام از این زیرشبکه‌ها نیز بطور جداگانه قابلیت پیش‌بینی دارند. با درنظر گرفتن 2D-CNNpred به عنوان شبکه پیچشی ابتدایی، می‌توان به این نتیجه رسید که در حین آموزش شبکه، ۳ زیرشبکه با ۲، ۳ و ۴ لایه قابل آموزش تشکیل می‌شود. جدول ۶-۵ عملکرد این ۳ زیرشبکه را مقایسه می‌کند.

## فصل ۵. نتایج، بحث و ارزیابی



شکل ۱-۵: مقایسه الگوریتم‌های مختلف در داشتن بهترین عملکرد از نظر F-measure

آزمایش سوم با هدف بررسی کارایی HT-CNNpred در پیش‌بینی بازارهایی است که در آموزش آن استفاده نشده‌اند. این قسمت به دو زیرآزمایش بررسی شاخص‌های آمریکایی و شاخص‌های غیر آمریکایی تقسیم شد. هدف دیگر از انجام این آزمایش، بررسی پایداری HT-CNNpred در حالتی که اختلاف زمانی میان داده‌های آموزش و آزمون وجود دارد، است. در این حالت، داده‌های آموزشی و اعتبارسنجی مربوط به ژانویه ۲۰۱۵ تا جولای ۲۰۱۵ است و داده آزمون از آپریل ۲۰۱۶ تا نوامبر ۲۰۱۷ است. جدول ۷-۵ نتایج حاصل از پیش‌بینی شاخص‌های آمریکایی با استفاده از یادگیری انتقالی را نشان می‌دهد. همچنین نتایج دو الگوریتم CNN-cor و

Target stock \ Model	2D-CNNpred	U pred (2 layers)	U pred (3 layers)	U pred (4 layers)
Whole data	0.4942	0.5022	0.5050	<b>0.5111</b>
XOM	0.4969	0.5079	0.5059	<b>0.5109</b>
JPM	0.4936	0.5075	0.5017	<b>0.5176</b>
AAPL	0.4889	0.5004	<b>0.5076</b>	0.5058
MSFT	0.4961	0.5039	0.5157	<b>0.5213</b>
GE	0.4973	0.5095	0.4955	<b>0.5161</b>
JNJ	0.4840	0.4921	0.5021	<b>0.5071</b>
WFC	0.4949	<b>0.5104</b>	0.5092	0.5102
AMZN	0.4962	0.5068	0.5069	<b>0.5114</b>

جدول ۵-۶: عملکرد زیرشبکه‌های تشکیل شده در پیش‌بینی وضعیت شرکت‌ها

Target stock \Model	Tech+ANN	PCA+ANN	CNN-cor	2D-CNNpred	H pred	G pred
S&P 500	0.4469	0.4237	0.3928	0.4907	0.4941	<b>0.4979</b>
NASDAQ	0.4199	0.4136	0.3769	0.4964	<b>0.5021</b>	0.5001
Dow 30	0.4150	0.4283	0.39	0.4964	0.4924	<b>0.4976</b>
NYSE	0.4071	0.4260	0.3906	0.4861	0.4865	<b>0.4910</b>
RUSSELL 2000	0.4525	0.4279	0.3924	0.4903	0.4905	<b>0.4961</b>
Average	0.4283	0.4239	0.3885	0.4920	0.4931	<b>0.4965</b>

جدول ۷-۵: الگوریتم‌های مختلف در پیش‌بینی شاخص‌های آمریکایی

Target stock \Model	Tech+ANN	PCA+ANN	CNN-cor	2D-CNNpred	H pred	G pred
DAX	0.4344	0.4455	0.3219	<b>0.4978</b>	0.4956	0.4917
KOSPI	0.4699	<b>0.4985</b>	0.3180	0.4843	0.4936	0.4912
Nikkei	0.4183	0.4114	0.3269	0.4876	<b>0.4888</b>	0.4832
CAC	0.4766	0.4565	0.3421	0.4924	<b>0.4933</b>	0.4901
FTSE	0.4371	0.4566	0.3328	0.4895	0.4925	<b>0.4955</b>
HSI	0.4518	0.4184	0.3156	0.4839	0.4971	<b>0.4990</b>
SSE	0.4520	0.4877	0.3122	0.4884	<b>0.4886</b>	0.4868
BSE 30	0.4196	<b>0.5067</b>	0.3240	0.4839	0.5020	0.5002
NIFTI 50	0.4366	0.4803	0.3274	0.4848	<b>0.4956</b>	0.4918
Average	0.4440	0.4624	0.3246	0.4881	<b>0.4942</b>	0.4922

جدول ۸-۵: الگوریتم‌های مختلف در پیش‌بینی شاخص‌های مهم جهان

2D-CNNpred نیز با استفاده از یادگیری انتقالی گزارش شده است.

از آنجاکه برای آموزش HT-CNNpred از داده‌های شرکت‌های آمریکایی استفاده شده است و ۵ شاخص آزمایش شده متاثر از این شرکت‌ها هستند، از داده‌های شاخص‌های غیرآمریکایی نیز برای بررسی کارایی HT-CNNpred استفاده می‌کنیم. بنابراین، فرآیند طی شده در مورد ۵ شاخص آمریکایی، در مورد ۹ شاخص معروف غیرآمریکایی نیز طی می‌شود که نتایج حاصل از آن را می‌توان در جدول ۸-۵ مشاهده کرد.

همانطور که از نتایج گزارش شده می‌توان دید، چارچوب پیشنهادی ما نسبت به تمامی الگوریتم‌های پایه در تقریباً تمامی آزمایش‌های انجام‌شده برتری دارد. طراحی هوشمندانه شبکه پیچشی ابتدایی یک دلیل قانع‌کننده برای برتری این چارچوب که در واقع همان 2D-CNNpred است، بر سایر الگوریتم‌های پایه می‌تواند باشد. علی‌رغم این برتری، وقتی عملکرد شبکه پیچشی ابتدایی را با پیش‌بینی کننده عمومی مقایسه می‌کنیم، متوجه عملکرد ضعیف‌تر آن نسبت به پیش‌بینی کننده عمومی می‌شویم. تنها تفاوت این ۲ چارچوب، نحوه آموزش آن‌ها است. در آموزش پیش‌بینی کننده عمومی از آموزش سلسله‌مراتبی استفاده شده است که تلاشی برای یافتن وزن‌های اولیه بهتر برای شبکه پیچشی از طریق داده‌های برچسب‌دار است. بنابراین، پیش‌بینی کننده عمومی به نقطه بهینه بهتری همگرا می‌شود.

آزمایش دوم نشان می‌دهد که افزایش عمق پیش‌بینی کننده عمومی باعث بهبود کارایی آن شده است. هرچه تعداد لایه‌های یک شبکه افزایش یابد، توانایی آن برای استخراج ویژگی‌های پیچیده‌تر نیز افزایش می‌یابد. بهمین

## فصل ۵. نتایج، بحث و ارزیابی

دلیل، عمیق‌ترین پیش‌بینی کننده عمومی عملکرد بهتری را نسبت به سایرین نشان می‌دهد. مشاهده جالب دیگر این است که عملکرد شبکه پیچشی ابتدایی حتی از کم عمیق‌ترین پیش‌بینی کننده عمومی نیز ضعیفتر است. به عبارتی دیگر، ترکیب آموزش سلسله‌مراتبی و یک شبکه عمیق باعث بهبود پیش‌بینی می‌شود و عمیق کردن شبکه به تنها یکی کافی نیست.

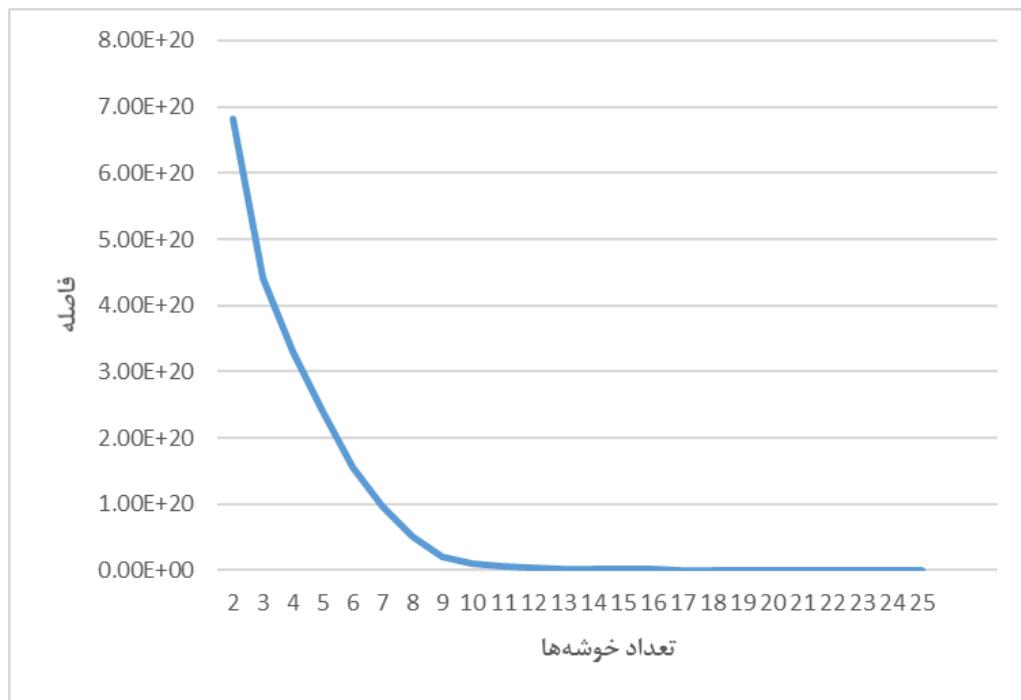
اگرچه از داده‌های ۵ شاخص آمریکایی در آموزش پیش‌بینی کننده عمومی استفاده نشده است، این مدل به خوبی توانسته است ۵ شاخص مطرح شده را پیش‌بینی کند. نتایج گزارش شده این معنی را می‌دهد که بازارهای مختلف، ویژگی‌های رفتاری مشابهی دارند و شبکه عمیقی که توسط داده‌های زیادی از این بازارها آموزش دیده شده باشد، قادر است تا این ویژگی‌های مشترک را استخراج کند و از آن‌ها برای پیش‌بینی بازارهای جدید استفاده کند. نتایج مربوط به پیش‌بینی شاخص‌های غیر آمریکایی اثبات می‌کند که این شباهت‌ها و ویژگی‌های مشترک میان بازارها فراتر از مرز کشورها می‌روند. علاوه بر 2D-CNNpred، HT-CNNpred حتی در برتری خود را نسبت به سایر الگوریتم‌های پایه از طریق یادگیری انتقالی حفظ کند و این نشانه‌ای دیگر بر مشابهت بازارها به یکدیگر است.

ما بطور تعمدی فاصله زمانی میان داده‌های آزمون و سایر داده‌ها قرار دادیم. داده‌های آزمون حدوداً ۹ ماه بعد از داده‌های اعتبارسنجی شروع می‌شود. کیفیت خوب پیش‌بینی نشان می‌دهد که HT-CNNpred حتی در برابر گذشت زمان نیز پایدار است.

## **C-CNNpred ۳-۶-۵**

اولین مرحله مربوط به پیدا کردن تعداد مناسب خوش‌ها و همین‌طور قرار دادن بازارهای مرتبط در یک خوش است. به این منظور، الگوریتم k میانگین را بر روی دادگان ۲۶ شاخص موجود اجرا می‌کنیم. نمودار ۲-۵ میزان مجموع فاصله درون خوش‌های را نسبت به تعداد خوش‌ها نشان می‌دهد. همان‌طور که قابل مشاهده است، میزان کاهش فاصله پس از ۹ خوش تغییر محسوسی نمی‌کند و می‌توان تعداد ۹ خوش را مناسب دانست.

پس از قرار دادن شاخص‌ها در ۹ خوش و محاسبه نمایندگان هر خوش، نوبت به اجرای الگوریتم‌ها می‌رسد. از آنجاکه نتایج اجرای الگوریتم‌ها به ازای هر بار اجرا می‌توانند متفاوت باشند و همین‌طور فراهم کردن شرایط بررسی‌های آماری میان الگوریتم‌ها، هرکدام از آن‌ها چندین مرتبه اجرا شده‌اند و میانگین آن‌ها در جدول ۹-۵ گزارش شده است. همچنین نمودار ۳-۵ نیز بهترین نتیجه را برای هرکدام از الگوریتم‌ها نشان می‌دهد.



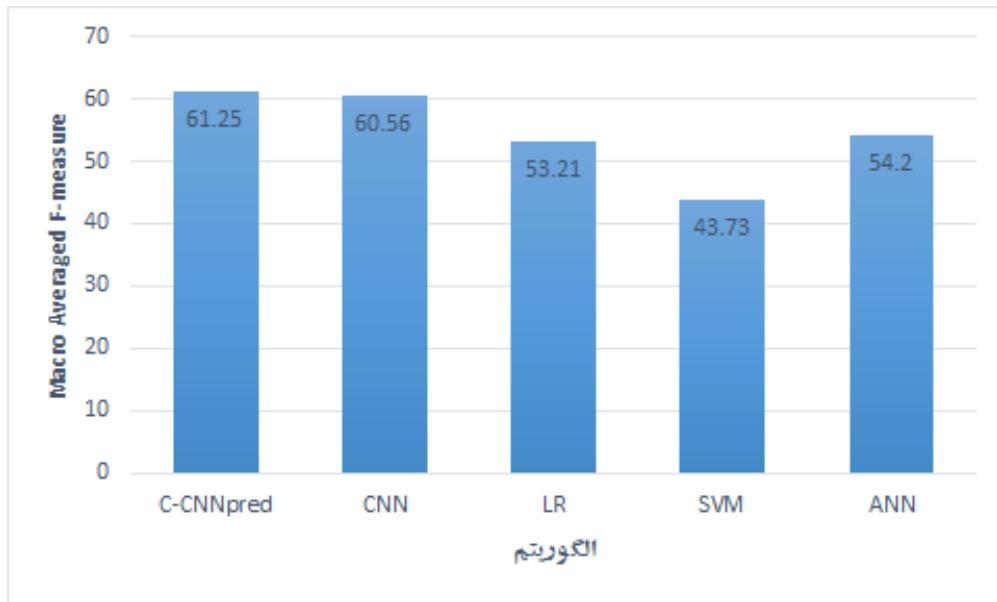
شکل ۵-۵: مجموع فاصله درون خوشها نسبت به تعداد خوشها

Algorithm	Average	p-value against C-CNNpred
C-CNNpred	0.5539	1
ANN	0.5256	less than 0.0001
SVM	0.4373	less than 0.0001
LR	0.5321	less than 0.0001
CNN	0.5077	less than 0.0001

جدول ۹-۵: میانگین F-measure الگوریتم‌های مختلف

همان‌طور که از نتایج مشخص است، بهترین نتیجه مربوط به C-CNNpred است که با اختلاف آماری معناداری از سایر الگوریتم‌ها، چه شبکه‌های عصبی و چه سایر مدل‌های یادگیری ماشینی، بهتر عمل کرده است. در دسته الگوریتم‌های مبتنی بر شبکه‌های عصبی، شبکه عصبی پرسپترون چندلایه که در واقع یک شبکه کم‌عمق است نتیجه بهتری نسبت به شبکه عصبی پیچشی گرفته است که دلیل این امر می‌تواند عدم وجود دادگان کافی برای آموزش شبکه عمیق باشد، در حالی که پس از استفاده از بازارهای مشابه در C-CNNpred که منجر به افزایش داده‌ها شده است، کارایی شبکه عصبی پیچشی بهبود پیداکرده است.

نکته دیگری که از جدول نتایج می‌توان دریافت این است که تفاوت میزان دقیقت شبکه عصبی پیچشی و C-



شکل ۳-۵: بهترین نتایج الگوریتم‌های مختلف

CNNpred کاملاً محسوس است. تفاوت این دو روش در این است که در الگوریتم دوم، بازارهای مشابهی که امکان تأثیرگذاری بر شاخص کل بازار تهران را دارند، در قالب یک تصور<sup>۳</sup>-بعدی به الگوریتم داده شده‌اند. نتایج به دست آمده حکایت از این دارد که حضور بازارهای تأثیرگذار بر بازار هدف تأثیر چشم‌گیری در بهبود کارایی شبکه‌های پیچشی گذاشته‌اند.

## ۷-۵ جمع‌بندی

در این فصل به بررسی نتایج حاصل از چارچوب‌های ارائه شده در فصل ۴ پرداختیم و نتایج را از ابعاد مختلفی تحلیل کردیم.

چارچوب CNNpred با طراحی منحصر به فرد ۲ شبکه پیچشی بنام‌های 2D-CNNpred و 3D-CNNpred توانست ۵ شاخص اصلی بازار آمریکا را با استفاده از طیف وسیعی از ویژگی‌ها پیش‌بینی کند. این چارچوب توانست با کسب F-measure نزدیک ۵۰٪، برتری معنادار آماری خود را نسب به سایر الگوریتم‌های پایه نشان دهد.

چارچوب بعدی که مورد بررسی قرار گرفت HT-CNNpred است. این چارچوب توانست با استفاده از

آموزش سلسله مراتبی، عملکرد 2D-CNNpred در پیش‌بینی تغییرات جهتی ۴۵۸ شرکت آمریکایی را حدود ۱٪ بهبود دهد. همچنین، استفاده از یادگیری انتقالی به این چارچوب قدرت پیش‌بینی ۱۴ شاخص معروف جهان را بخشدید.

چارچوب C-CNNpred نیز با استفاده از خوشبندی بازارها، از هر خوش نماینده آن خوش را انتخاب و از آن به عنوان بازار کمکی در چارچوب 3D-CNNpred استفاده می‌کند. این چارچوب با عملکرد حدود ۵۵٪ بهترین نتیجه را در میان الگوریتم‌های پایه در پیش‌بینی شاخص کل بازار بورس تهران به دست آورد.

## فصل ۶

# جمع‌بندی، نتیجه‌گیری و پیشنهادهایی برای تحقیقات آینده

در این پایان‌نامه چارچوب‌هایی برای پیش‌بینی تغییرات جهتی بازارهای مالی برای روز آینده ارائه شد. نتایج به دست آمده از اجرای چارچوب CNNpred مشخص کرد که فقط استفاده از تکنیکال اندیکاتورها برای پیش‌بینی کافی نیست و نیاز به ویژگی‌هایی با تنوع بیشتر برای بهبود کارایی الگوریتم‌های پیش‌بینی کننده حس می‌شود. همچنین دریافتیم که استفاده از مجموعه غنی‌ای از ویژگی‌های ابتدایی به‌نهایی کافی نیست و استفاده از الگوریتمی مناسب برای استخراج ویژگی‌های مناسب نیز ضروری به‌نظر می‌رسد. یادگیری عمیق و بطور خاص، شبکه‌های عصبی پیچشی گزینه مناسبی برای استخراج ویژگی‌های مناسب از ویژگی‌های ابتدایی هستند. این شبکه‌ها دارای متغیرهای فراوانی هستند و طراحی درست آن‌ها در افزایش کارایی آن‌ها تاثیر چشم‌گیری دارد. موارد ذکر شده تماماً در چارچوب CNNpred پوشش داد شده‌است.

چارچوب HT-CNNpred نیز با ارائه روشی سلسله‌مراتبی برای آموزش شبکه‌های عصبی، سعی بر این دارد تا از مزیت داده‌های برچسب‌دار برای یافتن وزن‌های اولیه مناسب برای شبکه‌های عصبی استفاده کند. نتایج کسب شده در پیش‌بینی تغییرات جهتی ۴۵۸ شرکت حاضر در بازار آمریکا سودمند بودن این روش آموزش را تایید کرده‌است. استفاده از یادگیری انتقالی برای پیش‌بینی بازارهای جدید، این چارچوب را به یک پیش‌بینی کننده قوی تبدیل کرده‌است که توانایی پیش‌بینی بازارهای جدید را بدون آموزش مجدد کل شبکه از ابتدا فراهم می‌کند و از این حیث می‌توان به آن یک پیش‌بینی کننده عمومی تمام بازارهای مالی نیز گفت. نتایج کسب شده در پیش‌بینی

۱۴ شاخص مطرح در سرتاسر جهان، شاهدی بر این ادعاست.

تأثیرات متقابل شاخص‌ها و بازارهای مختلف بر یکدیگر باعث می‌شود تا بتوان از آن‌ها برای پیش‌بینی یکدیگر استفاده کرد. چارچوب C-CNNpred با طراحی روالی الگوریتمیک سعی در انتخاب بازارهایی مناسب برای پیش‌بینی یک بازار هدف دارد. استفاده از خوشبندی طیف وسیعی از بازارهایی که می‌توانند بر بازار هدف اثرگذار باشند، باعث می‌شود تا بازارهای با رفتار مشابه در یک خوشبند قرار بگیرند و از هر خوشبند آن برای حضور در پیش‌بینی نهایی انتخاب شود. از این چارچوب برای پیش‌بینی شاخص کل بازار بورس تهران استفاده شد. اضافه کردن دادگان مربوط به بازارهای مشابه منجر به افزایش حدود ۵ درصدی در دقت پیش‌بینی شد. علاوه بر این، مقایسه با سایر الگوریتم‌های پایه برتری و کارایی چارچوب ارائه شده را به خوبی نشان داد.

## ۱-۶ پیشنهادهایی برای پژوهش‌های آینده

این پژوهش امکانات گسترش فراوانی را پیش‌روی خود دارد که به مواردی اشاره خواهیم داشت:

۱. در این پایان‌نامه سعی شد تا از طیف وسیعی از ویژگی‌های ابتدایی برای پیش‌بینی تغییرات جهتی شرکت‌ها و شاخص‌های حاضر در بازار آمریکا استفاده شود. با این حال، منابع دیگری وجود دارند که می‌توان از آن‌ها برای بهبود پیش‌بینی بهره برد. استفاده از خبرها و همچنین توییت‌های شبکه اجتماعی توییتر و تحلیل اثر آن‌ها بر بازارهای مالی می‌تواند در پژوهش‌های آتی به چارچوب‌های پیشنهادی اضافه شود.

۲. الگوریتم‌های پیش‌بینی بازارهای مالی نقشی بسیار مهم در موفقیت سیستم‌های خرید و فروش الگوریتمیک سهام ایفا می‌کنند. این پایان‌نامه به توسعه چارچوب‌هایی کارا برای پیش‌بینی بازارهای مالی پرداخته است. قدم بعدی می‌تواند استفاده از خروجی چارچوب‌های پیشنهادی در یک سیستم خرید و فروش باشد.

۳. مورد بعدی می‌تواند الهام گرفتن از چارچوب‌های طراحی شده برای طراحی چارچوبی مبتنی بر الگوریتم‌های یادگیری تقویتی عمیق باشد. بدین‌ترتیب، خرید و فروش و حتی میزان خرید و فروش سهام با استفاده از یک الگوریتم یادگیری تقویتی عمیق خواهد بود و نقش مهندسان مالی به مرتب کمرنگ‌تر از قبل خواهد شد.

# مراجع

- [1] [http://cs231n.github.io/convolutional-networks/.](http://cs231n.github.io/convolutional-networks/)
- [2] [https://cambridgespark.com/content/tutorials/convolutional-neural-networks-with-keras/index.html/.](https://cambridgespark.com/content/tutorials/convolutional-neural-networks-with-keras/index.html/)
- [3] [https://en.wikipedia.org/wiki/Convolutional\\_neural\\_network.](https://en.wikipedia.org/wiki/Convolutional_neural_network)
- [4] [http://colah.github.io/posts/2015-08-Understanding-LSTMs/.](http://colah.github.io/posts/2015-08-Understanding-LSTMs/)
- [5] Setak, Mostafa. Introduction to genetic algorithm. University Lecture.
- [6] [https://en.wikipedia.org/wiki/Autoencoder.](https://en.wikipedia.org/wiki/Autoencoder)
- [7] Kara, Yakup, Boyacioglu, Melek Acar, and Baykan, Ömer Kaan. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the istanbul stock exchange. *Expert systems with Applications*, 38(5):5311–5319, 2011.
- [8] Kyoung-jae Kim, Ingoo Han. Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems with Applications*, 2000.
- [9] Mingyue Qiu, Yu Song, Fumio Akagi. Application of artificial neural network for the prediction of stock market returns: The case of the japanese stock market. *Elsevier*, 2016.
- [10] Mingyue Qiu, Yu Song, Li cheng. Application of the artifical neural network in predicting the direction of stock market index. *10th International Conference on Complex, Intelligent, and Software Intensive Systems , IEEE*, 2016.

- [11] Xiao Zhong, David Enke. Forecasting daily stock market return using dimensionality reduction. *Expert Systems with Applications*, Elsevier, 67:126–139, 2017.
- [12] Amin Hedayati Moghaddam, Moein Hedayati Moghaddam, Morteza Esfandyari. Stock market index prediction using artificial neural network. *Journal of Economics, Finance and Administrative Science*, Elsevier, 21:89–93, 2016.
- [13] Andrés Arévalo, Jaime Niño, German Hernández Javier Sandoval. High-frequency trading strategy based on deep neural networks. *Intelligent Computing Methodologies*, Springer, pp. 424–436, 2016.
- [14] Bang Xiang, authorMohd Rozaini, Abdul Rahim Ahmad Shahidan Abdullah. Forecasting daily stock market return using dimensionality reduction. *Modeling, Design and Simulation of Systems. AsiaSim 2017. Communications in Computer and Information Science*, Springer, 751:356–364, 2017.
- [15] Gao, Qiyuan. Stock market forecasting using recurrent neural network. *Master Thesis , University of Missouri-Columbia*, 2016.
- [16] Nelson, David MQ, Pereira, Adriano CM, and de Oliveira, Renato A. Stock market's price movement prediction with lstm neural networks. in *Neural Networks (IJCNN), 2017 International Joint Conference on*, pp. 1419–1426. IEEE, 2017.
- [17] Gunduz, Hakan, Yaslan, Yusuf, and Cataltepe, Zehra. Intraday prediction of borsa istanbul using convolutional neural networks and feature correlations. *Knowledge-Based Systems*, 137:138–148, 2017.
- [18] Sezer, Omer Berat and Ozbayoglu, Ahmet Murat. Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach. *Applied Soft Computing*, 2018.
- [19] Eunsuk Chong a, Chulwoo Han b, Frank C. Park. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, Elsevier, 83:187–205, 2017.
- [20] Isaac Madan, Shaurya Saluja, Aoja Zhao. Automated bitcoin trading via machine learnin algorithms. tech. rep., Stanford University, Department of Computer Science, 2014.

- [21] Potoski, Megan. Predicting gold prices. tech. rep., Stanford University, Department of Computer Science, 2013.
- [22] Yang Zhaoab, Jianping Liab, Lean Yu. A deep learning ensemble approach for crude oil price forecasting. *Energy Economics* , Elsevier, 2017.
- [23] Ming Hao Eng, Yang Li, Qing-Guo Wang Tong Heng Lee. Forecast forex with ann using fundamental data. *International Conference on Information Management, Innovation Management and Industrial Engineering, IEEE*, 2008.
- [24] Yao, Jingtao and Tan, Chew Lim. A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing*, 34(1):79–98, 2000.
- [25] Kia, Arash Negahdari, Haratizadeh, Saman, and Shouraki, Saeed Bagheri. A hybrid supervised semi-supervised graph-based model to predict one-day ahead movement of global stock markets and commodity prices. *Expert Systems with Applications*, 105:159–173, 2018.
- [26] Li, Wei and Liao, Jian. A comparative study on trend forecasting approach for stock price time series. in *Anti-counterfeiting, Security, and Identification (ASID), 2017 11th IEEE International Conference on*, pp. 74–78. IEEE, 2017.
- [27] Patel, Jigar, Shah, Sahil, Thakkar, Priyank, and Kotecha, K. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications*, 42(1):259–268, 2015.
- [28] Imandoust, Sadegh Bafandeh and Bolandraftar, Mohammad. Forecasting the direction of stock market index movement using three data mining techniques: the case of tehran stock exchange. *International Journal of Engineering Research and Applications*, 4(6):106–117, 2014.
- [29] Huang, Wei, Nakamori, Yoshiteru, and Wang, Shou-Yang. Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, 32(10):2513–2522, 2005.
- [30] Hoseinzade, Ehsan and Haratizadeh, Saman. Cnnpred: Cnn-based stock market prediction using a diverse set of variables. *Expert Systems with Applications*, 129:273–285, 2019.

- [31] Moghaddam, Amin Hedayati, Moghaddam, Moein Hedayati, and Esfandyari, Morteza. Stock market index prediction using artificial neural network. *Journal of Economics, Finance and Administrative Science*, 21(41):89–93, 2016.
- [32] Chong, Eunsuk, Han, Chulwoo, and Park, Frank C. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, 83:187–205, 2017.
- [33] Zhong, Xiao and Enke, David. A comprehensive cluster and classification mining procedure for daily stock market return forecasting. *Neurocomputing*, 267:152–168, 2017.
- [34] Bjoern Krollner, Bruce Vanstone, Gavine Finnie. Financial time series forecasting with machine learning techniques: A survey. *European Symposium on Artificial Neural Networks - Computational Intelligence and Machine Learning*, 2010.
- [35] Kim, Kyoung-jae and Han, Ingoo. Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert systems with Applications*, 19(2):125–132, 2000.
- [36] Qiu, Mingyue, Song, Yu, and Akagi, Fumio. Application of artificial neural network for the prediction of stock market returns: The case of the japanese stock market. *Chaos, Solitons & Fractals*, 85:1–7, 2016.
- [37] Qiu, Mingyue and Song, Yu. Predicting the direction of stock market index movement using an optimized artificial neural network model. *PloS one*, 11(5):e0155133, 2016.
- [38] Mahfoud, Sam and Mani, Ganesh. Financial forecasting using genetic algorithms. *Applied artificial intelligence*, 10(6):543–566, 1996.
- [39] Brown, Michael Scott, Pelosi, Michael J, and Dirska, Henry. Dynamic-radius species-conserving genetic algorithm for the financial forecasting of dow jones index stocks. in *International Workshop on Machine Learning and Data Mining in Pattern Recognition*, pp. 27–41. Springer, 2013.
- [40] LeCun, Yann, Bengio, Yoshua, and Hinton, Geoffrey. Deep learning. *nature*, 521(7553):436, 2015.

- [41] Collobert, Ronan and Weston, Jason. A unified architecture for natural language processing: Deep neural networks with multitask learning. in *Proceedings of the 25th international conference on Machine learning*, pp. 160–167. ACM, 2008.
- [42] He, Kaiming, Zhang, Xiangyu, Ren, Shaoqing, and Sun, Jian. Deep residual learning for image recognition. in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [43] Cai, Xianggao, Hu, Su, and Lin, Xiaola. Feature extraction using restricted boltzmann machine for stock price prediction. in *Computer Science and Automation Engineering (CSAE), 2012 IEEE International Conference on*, vol. 3, pp. 80–83. IEEE, 2012.
- [44] Fischer, Thomas and Krauss, Christopher. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2):654–669, 2018.
- [45] Bao, Wei, Yue, Jun, and Rao, Yulei. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PloS one*, 12(7):e0180944, 2017.
- [46] Li Deng, Dong Yu. *Deep Learning: Methods and Applications*, vol. 7. Foundations and Trends in Signal Processing, 2013.
- [47] Bengio, Yoshua. *Learning Deep Architectures for AI*, vol. 2. Foundations and Trends in Machine Learning, 2009.
- [48] A Graves, N Jaitly, A Mohamed. Hybrid speech recognition with deep bidirectional lstm. *Automatic Speech Recognition and Understanding (ASRU)*, 2013.
- [49] Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65:386–408, 1958.
- [50] Y LeCun, Y Bengio. Convolutional networks for images, speech, and time series. in *The handbook of brain theory and neural networks*, vol. 4, pp. 255–258. MIT Press, 1995.
- [51] Bottou, Léon. Large-scale machine learning with stochastic gradient descent. in *Proceedings of COMPSTAT’2010*, pp. 177–186. Springer, 2010.

- [52] Robbins, Herbert and Monro, Sutton. A stochastic approximation method. *The annals of mathematical statistics*, pp. 400–407, 1951.
- [53] Kingma, Diederik P and Ba, Jimmy. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [54] Duchi, John, Hazan, Elad, and Singer, Yoram. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(Jul):2121–2159, 2011.
- [55] Zeiler, Matthew D. Adadelta: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*, 2012.
- [56] Xie, Xin-Kai and Wang, Hong. Recurrent neural network for forecasting stock market trend. in *Computer Science, Technology and Application: Proceedings of the 2016 International Conference on Computer Science, Technology and Application (CSTA2016)*, p. 397. World Scientific, 2016.
- [57] Sepp Hochreiter, Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9:1735–1780, 1997.
- [58] Hinton, Geoffrey E, Srivastava, Nitish, Krizhevsky, Alex, Sutskever, Ilya, and Salakhutdinov, Ruslan R. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*, 2012.
- [59] Srivastava, Nitish, Hinton, Geoffrey, Krizhevsky, Alex, Sutskever, Ilya, and Salakhutdinov, Ruslan. Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958, 2014.
- [60] Pan, Sinno Jialin and Yang, Qiang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2010.
- [61] Shin, Hoo-Chang, Roth, Holger R, Gao, Mingchen, Lu, Le, Xu, Ziyue, Nogues, Isabella, Yao, Jianhua, Mollura, Daniel, and Summers, Ronald M. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE transactions on medical imaging*, 35(5):1285–1298, 2016.
- [62] Veisi, Hadi. Genetic algorithm. University Lecture.

- [63] Martin Längkvist, Lars Karlsson, Amy Loutf. A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recognition Letters*, Elsevier, 42:11–24, 2014.
- [64] Esmaili, Mahdi. *Data Mining Concept and Techniques*.
- [65] Hinton, Geoffery. A practical guide to training restricted boltzman machine. tech. rep., University of Toronto, Department of Computer Science, 2010.
- [66] Jain, Anil K. Data clustering: 50 years beyond k-means. *Pattern recognition letters*, 31(8):651–666, 2010.
- [67] Jain, Anil K, Murty, M Narasimha, and Flynn, Patrick J. Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3):264–323, 1999.
- [68] Johnson, Stephen C. Hierarchical clustering schemes. *Psychometrika*, 32(3):241–254, 1967.
- [69] Hartigan, John A and Wong, Manchek A. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.
- [70] Krollner, Bjoern, Vanstone, Bruce, and Finnie, Gavin. Financial time series forecasting with machine learning techniques: A survey. 2010.
- [71] Dai, Wensheng, Wu, Jui-Yu, and Lu, Chi-Jie. Combining nonlinear independent component analysis and neural network for the prediction of asian stock market indexes. *Expert systems with applications*, 39(4):4444–4452, 2012.
- [72] Ticknor, Jonathan L. A bayesian regularized artificial neural network for stock market forecasting. *Expert Systems with Applications*, 40(14):5501–5506, 2013.
- [73] de Oliveira, Fagner A, Nobre, Cristiane N, and Zárate, Luis E. Applying artificial neural networks to prediction of stock price and improvement of the directional prediction index—case study of petr4, petrobras, brazil. *Expert Systems with Applications*, 40(18):7596–7606, 2013.
- [74] Hecht-Nielsen, Robert. Theory of the backpropagation neural network. in *Neural networks for perception*, pp. 65–93. Elsevier, 1992.

- [75] Hagan, Martin T and Menhaj, Mohammad B. Training feedforward networks with the marquardt algorithm. *IEEE transactions on Neural Networks*, 5(6):989–993, 1994.
- [76] Mingyue Qiu, Yu Song. Predicting the direction of stock market index movement using an optimized artificial neural network model. *PLOS ONE*, 2016.
- [77] Zahedi, Javad and Rounaghi, Mohammad Mahdi. Application of artificial neural network models and principal component analysis method in predicting stock prices on tehran stock exchange. *Physica A: Statistical Mechanics and its Applications*, 438:178–187, 2015.
- [78] Chatzis, Sotirios P, Siakoulis, Vassilis, Petropoulos, Anastasios, Stavroulakis, Evangelos, and Vlachogiannakis, Nikos. Forecasting stock market crisis events using deep and statistical machine learning techniques. *Expert Systems with Applications*, 112:353–371, 2018.
- [79] Rather, Akhter Mohiuddin, Agarwal, Arun, and Sastry, VN. Recurrent neural network and a hybrid model for prediction of stock returns. *Expert Systems with Applications*, 42(6):3234–3241, 2015.
- [80] Shen, Guizhu, Tan, Qingping, Zhang, Haoyu, Zeng, Ping, and Xu, Jianjun. Deep learning with gated recurrent unit networks for financial sequence predictions. *Procedia computer science*, 131:895–903, 2018.
- [81] Balaji, A Jayanth, Ram, DS Harish, and Nair, Binoy B. Applicability of deep learning models for stock price forecasting an empirical study on bankex data. *Procedia computer science*, 143:947–953, 2018.
- [82] Sang, Chenjie and Di Pierro, Massimo. Improving trading technical analysis with tensorflow long short-term memory (lstm) neural network. *The Journal of Finance and Data Science*, 2018.
- [83] Chung, Hyejung and Shin, Kyung-shik. Genetic algorithm-optimized long short-term memory network for stock market prediction. *Sustainability*, 10(10):3765, 2018.

- [84] Baek, Yujin and Kim, Ha Young. Modaugnet: A new forecasting framework for stock market index value with an overfitting prevention lstm module and a prediction lstm module. *Expert Systems with Applications*, 113:457–480, 2018.
- [85] Sagheer, Alaa and Kotb, Mostafa. Time series forecasting of petroleum production using deep lstm recurrent networks. *Neurocomputing*, 323:203–213, 2019.
- [86] Kim, Ha Young and Won, Chang Hyun. Forecasting the volatility of stock price index: A hybrid model integrating lstm with multiple garch-type models. *Expert Systems with Applications*, 103:25–37, 2018.
- [87] Lahmiri, Salim and Bekiros, Stelios. Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons & Fractals*, 118:35–40, 2019.
- [88] Tsantekidis, Avraam, Passalis, Nikolaos, Tefas, Anastasios, Kanniainen, Juho, Gabouj, Moncef, and Iosifidis, Alexandros. Forecasting stock prices from the limit order book using convolutional neural networks. in *Business Informatics (CBI), 2017 IEEE 19th Conference on*, vol. 1, pp. 7–12. IEEE, 2017.
- [89] Niño, Jaime, Hernandez, German, Arévalo, Andrés, Leon, Diego, and Sandoval, Javier. Cnn with limit order book data for stock price prediction. in *Proceedings of the Future Technologies Conference*, pp. 444–457. Springer, 2018.
- [90] Di Persio, Luca and Honchar, Oleksandr. Artificial neural networks architectures for stock price prediction: Comparisons and applications. *International Journal of Circuits, Systems and Signal Processing*, 10:403–413, 2016.
- [91] Nison, Steve. *Beyond candlesticks: New Japanese charting techniques revealed*, vol. 56. John Wiley & Sons, 1994.
- [92] Bulkowski, Thomas N. *Encyclopedia of candlestick charts*, vol. 332. John Wiley & Sons, 2012.
- [93] Achelis, Steven B. *Technical Analysis from A to Z*. McGraw Hill New York, 2001.
- [94] Brzeszczyński, Janusz and Ibrahim, Boulis Maher. A stock market trading system based on foreign and domestic information. *Expert Systems with Applications*, 118:381–399, 2019.

- [95] Ibrahim, Boulis Maher and Brzeszczynski, Janusz. How beneficial is international stock market information in domestic stock market trading? *The European Journal of Finance*, 20(3):201–231, 2014.
- [96] Bahmani-Oskooee, Mohsen and Sohrabian, Ahmad. Stock prices and the effective exchange rate of the dollar. *Applied economics*, 24(4):459–464, 1992.
- [97] Aggarwal, Raj. Exchange rates and stock prices: A study of the us capital markets under floating exchange rates. 2003.
- [98] Creti, Anna, Joëts, Marc, and Mignon, Valérie. On the links between stock and commodity markets' volatility. *Energy Economics*, 37:16–28, 2013.
- [99] Zhong, Xiao and Enke, David. Forecasting daily stock market return using dimensionality reduction. *Expert Systems with Applications*, 67:126–139, 2017.
- [100] Niaki, Seyed Taghi Akhavan and Hoseinzade, Saeid. Forecasting s&p 500 index using artificial neural networks and design of experiments. *Journal of Industrial Engineering International*, 9(1):1, 2013.
- [101] Enke, David and Thawornwong, Suraphan. The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with applications*, 29(4):927–940, 2005.
- [102] Özgür, Arzucan, Özgür, Levent, and Güngör, Tunga. Text categorization with class-based and corpus-based keyword selection. in *International Symposium on Computer and Information Sciences*, pp. 606–615. Springer, 2005.
- [103] Chollet, François et al. Keras. <https://keras.io>, 2015.
- [104] Zarei, Ghasem, Rana, Mohamadiyan, Hazeri, Hatef, and Bashokouh Ajirlou, Mohammad. The comparison of fuzzy neural network methods with wavelet fuzzy neural network in predicting stock prices of banks accepted in tehran stock exchange. *Journal of Financial Management Strategy*, 6(3):109–138, 2018.
- [105] Bajalan, Saeed, Fallahpour, Saeed, and Dana, Nahid. The stock trend prediction using volume weighted support vector machine with a hybrid feature selection method to predict the stock price trend in tehran stock exchange. 4(3):121–148, 2016.

## فصل ۶. جمع‌بندی، نتیجه‌گیری و پیشنهادهایی برای تحقیقات آینده

[106] Ou, Phichhang and Wang, Hengshan. Prediction of stock market index movement by ten data mining techniques. *Modern Applied Science*, 3(12):28, 2009.

# آپیوست پیوست

## شناخت داده‌ها

این پیوست به توضیح بیشتر ویژگی‌های و داده‌های استفاده شده در این پایان‌نامه اختصاص دارد.

جدول آ-۱ ویژگی‌های استفاده شده برای پیش‌بینی بازار آمریکا و همچنین سایر شاخص‌های مالی سراسر جهان را نشان می‌دهد. لیست شاخص‌های استفاده شده در پیش‌بینی شاخص کل بازار تهران نیز در جدول آ-۲ قابل مشاهد است.

منبع / محاسبه	نوع	توضیحات	ویژگی	#
Pandas	ابتداي	week of day which	Day	۱
Finance Yahoo	ابتداي	price Close	Close	۲
TA-Lib	تکنيکال انديكاتور	volume of change Relative	Vol	۳
TA-Lib	تکنيکال انديكاتور	before days $\gamma$ of Return	MOM- $\gamma$	۴
TA-Lib	تکنيکال انديكاتور	before days $\gamma$ of Return	MOM- $\gamma$	۵
TA-Lib	تکنيکال انديكاتور	before days $\gamma$ of Return	MOM- $\gamma$	۶
TA-Lib	تکنيکال انديكاتور	Change of Rate days $\delta$	ROC- $\delta$	۷
TA-Lib	تکنيکال انديكاتور	Change of Rate days $\delta$	ROC- $\delta$	۸
TA-Lib	تکنيکال انديكاتور	Change of Rate days $\delta$	ROC- $\delta$	۹
TA-Lib	تکنيکال انديكاتور	Change of Rate days $\delta$	ROC- $\delta$	۱۰
TA-Lib	تکنيکال انديكاتور	Average Moving Exponential days $\gamma$	EMA- $\gamma$	۱۱
TA-Lib	تکنيکال انديكاتور	Average Moving Exponential days $\gamma$	EMA- $\gamma$	۱۲
TA-Lib	تکنيکال انديكاتور	Average Moving Exponential days $\delta$	EMA- $\delta$	۱۳
TA-Lib	تکنيکال انديكاتور	Average Moving Exponential days $\gamma$	EMA- $\gamma$	۱۴
FRED	سایر	Rate Market Secondary Bill: Treasury Week- $\gamma$	DTB $\gamma$ WK	۱۵
FRED	سایر	Rate Market Secondary Bill: Treasury Month- $\gamma$	DTB $\gamma$	۱۶
FRED	سایر	Rate Market Secondary Bill: Treasury Month- $\gamma$	DTB $\gamma$	۱۷
FRED	سایر	Rate Maturity Constant Treasury Year- $\delta$	DGS $\delta$	۱۸
FRED	سایر	Rate Maturity Constant Treasury Year- $\delta$	DGS $\delta$	۱۹
FRED	سایر	Yield Bond Corporate Aaa Seasoned s' Moody	DAAA	۲۰
FRED	سایر	Yield Bond Corporate Baa Seasoned s' Moody	DBAA	۲۱
FRED	سایر	DGS $\gamma$ -DTB $\gamma$ WK	TE $\gamma$	۲۲
FRED	سایر	DGS $\gamma$ -DTB $\gamma$	TE $\gamma$	۲۳
FRED	سایر	DGS $\gamma$ -DTB $\gamma$	TE $\gamma$	۲۴

**پیوست آ. شناخت داده‌ها**

منبع / محاسبه	نوع	توضیحات	ویرگی	#
FRED	سایر	DTB۳-DTB۴WK	TE۵	۲۵
FRED	سایر	DTB۶-DTB۴WK	TE۶	۲۶
FRED	سایر	DBAA-BAAA	DE۱	۲۷
FRED	سایر	DBAA-DGS۱۰	DE۲	۲۸
FRED	سایر	DBAA-DTB۶	DE۴	۲۹
FRED	سایر	DBAA-DTB۴	DE۵	۳۰
FRED	سایر	DBAA-DTB۴WK	DE۶	۳۱
FRED	سایر	at securities Treasury U.S. on yield market the in Change basis investment on quoted maturity , constant month-۳	CTB۷M	۳۲
FRED	سایر	at securities Treasury U.S. on yield market the in Change basis investment on quoted maturity , constant month-۶	CTB۸M	۳۳
FRED	سایر	at securities Treasury U.S. on yield market the in Change basis investment on quoted maturity , constant year-۱	CTB۱Y	۳۴
FRED	کالا	Oklahoma price(WTI), oil of change Relative	Oil	۳۵
Investing.com	کالا	price(Brent) oil of change Relative	Oil	۳۶
Investing.com	کالا	price(WTI) oil of change Relative	Oil	۳۷
FRED	کالا	market) (London price gold of change Relative	Gold	۳۸
Investing.com	کالا	futures price gold of change Relative	Gold-F	۳۹
Investing.com	کالا	dollar U.S. spot gold of change Relative	XAU-USD	۴۰
Investing.com	کالا	dollar U.S. spot silver of change Relative	XAG-USD	۴۱
Investing.com	کالا	price gas of change Relative	Gas	۴۲
Investing.com	کالا	price silver of change Relative	Silver	۴۳
Investing.com	کالا	future copper of change Relative	Copper	۴۴
Finance Yahoo	شاخص‌های جهان	index Composite NASDAQ of Return	IXIC	۴۵
Finance Yahoo	شاخص‌های جهان	index ۵۰ S&P of Return	GSPC	۴۶
Finance Yahoo	شاخص‌های جهان	Average Industrial Jones Dow of Return	DJI	۴۷
Finance Yahoo	شاخص‌های جهان	index exchange stock NY of Return	NYSE	۴۸
Finance Yahoo	شاخص‌های جهان	index ۲۰۰ RUSSELL of Return	RUSSELL	۴۹
Finance Yahoo	شاخص‌های جهان	index Seng Hang of Return	HSI	۵۰
Finance Yahoo	شاخص‌های جهان	index Composite Exchange Stock Hai Shang of Return	SSE	۵۱
Finance Yahoo	شاخص‌های جهان	¥ CAC of Return	FCHI	۵۲
Finance Yahoo	شاخص‌های جهان	۱۰۰ FTSE of Return	FTSE	۵۳
Finance Yahoo	شاخص‌های جهان	DAX of Return	GDAXI	۵۴
Finance Yahoo	نرخ تبدیل	rate exchange yen Japanese to dollar US in change Relative	USD-Y	۵۵
Finance Yahoo	نرخ تبدیل	rate exchange pound British to dollar US in change Relative	USD-GBP	۵۶
Finance Yahoo	نرخ تبدیل ارز	rate exchange dollar Canadian to dollar US in change Relative	USD-CAD	۵۷
Finance Yahoo	نرخ تبدیل ارز	rate exchange yuan Chinese to dollar US in change Relative	USD-CNY	۵۸
Investing.com	نرخ تبدیل ارز	rate exchange dollar Australian to dollar US in change Relative	USD-AUD	۵۹
Investing.com	نرخ تبدیل ارز	rate exchange dollar Zealand New to dollar US in change Relative	USD-NZD	۶۰
Investing.com	نرخ تبدیل ارز	rate exchange franc Swiss to dollar US in change Relative	USD-CHF	۶۱
Investing.com	نرخ تبدیل ارز	rate exchange Euro to dollar US in change Relative	USD-EUR	۶۲
Investing.com	نرخ تبدیل ارز	index dollar US in change Relative	USDX	۶۳
Finance Yahoo	شرکت‌های آمریکایی	Corporation Mobil Exon of Return	XOM	۶۴
Finance Yahoo	شرکت‌های آمریکایی	Co. & Chase JPMorgan of Return	JPM	۶۵
Finance Yahoo	شرکت‌های آمریکایی	Inc. Apple of Return	AAPL	۶۶
Finance Yahoo	شرکت‌های آمریکایی	Corporation Microsoft of Return	MSFT	۶۷
Finance Yahoo	شرکت‌های آمریکایی	Company Electric General of Return	GE	۶۸
Finance Yahoo	شرکت‌های آمریکایی	Johnson & Johnson of Return	JNJ	۶۹
Finance Yahoo	شرکت‌های آمریکایی	Company & Fargo Wells of Return	WFC	۷۰
Finance Yahoo	شرکت‌های آمریکایی	Inc. Amazon.com of Return	AMZN	۷۱
Investing.com	بازار آتی	Futures ¥ CAC of Return	FCHI-F	۷۲
Investing.com	بازار آتی	Futures ۱۰۰ FTSE of Return	FTSE-F	۷۳

## پیوست آ. شناخت داده‌ها

منبع / محاسبه	نوع	توضیحات	ویژگی	#
Investing.com	بازار آتی	Futures DAX of Return	GDAXI-F	۷۴
Investing.com	بازار آتی	Futures index Seng Hang of Return	HSI-F	۷۵
Investing.com	بازار آتی	Futures index Nikkei of Return	Nikkei-F	۷۶
Investing.com	بازار آتی	Futures exchange stock Korean of Return	KOSPI-F	۷۷
Investing.com	بازار آتی	Futures index Composite NASDAQ of Return	IXIC-F	۷۸
Investing.com	بازار آتی	Futures Average Industrial Jones Dow of Return	DJI-F	۷۹
Investing.com	بازار آتی	Futures index ۵۰۰ S&P of Return	S&P-F	۸۰
Investing.com	بازار آتی	Futures RUSSELL of Return	RUSSELL-F	۸۱
Investing.com	نرخ تبدیل ارز	futures index dollar US in change Relative	USDX-F	۸۲

جدول آ-۱: توضیحات ویژگی‌های استفاده شده

شاخص	شماره	شاخص	شماره	شاخص	شماره
محصولات کاغذی	۱۹	استخراج نفت جز کشف	۱۰	شاخص کل	۱
مالی	۲۰	شاخص اول فرابورس	۱۱	شاخص ۳۰ شرکت بزرگ	۲
مالی	۲۱	بازار دوم فرابورس	۱۲	شاخص ۵۰ شرکت فعالتر	۳
مواد دارویی	۲۲	فرآوردهای نفتی	۱۳	شاخص قیمت ۵۰ شرکت	۴
رادیوئی	۲۳	قند و شکر	۱۴	انبوسازی	۵
شاخص صنعت	۲۴	دستگاه‌های برقی	۱۵	شاخص آزاد مشاور	۶
سرمایه‌گذاری‌ها	۲۵	سایر معدن	۱۶	شاخص بازار اول	۷
سیمان	۲۶	محصولات چوبی	۱۷	شاخص بازار دوم	۸
زراعت	۲۷	محصولات فلزی	۱۸	بیمه و بازنشستگی	۹

جدول آ-۲: لیست شاخص‌های حاضر در بازار ایران

## پیوست ب

### دستاوردهای پژوهشی پایان نامه

- Hoseinzade, E., & Haratizadeh, S. (2019). CNNpred: CNN-based stock market prediction using a diverse set of variables. *Expert Systems with Applications*, 129, 273-285. **ISI: Impact Factor: 3.768**
- Hoseinzade, E., Haratizadeh, S., & Khoeini, A. (2019). HT-CNNpred: An Universal CNN-Based Predictor of Stock Markets Using Hierarchy Training and Transfer Learning. *Expert Systems with Applications*, **ISI, Ready to submit**
- حسینزاده احسان. هراتیزاده، سامان (۱۳۹۸). C-CNNpred: چارچوبی مبتنی بر شبکه‌های پیچشی و خوشبندی سهام برای پیش‌بینی بازار بورس. *فصلنامه علمی-پژوهشی راهبرد مدیریت مالی*. آماده ارسال

# واژه‌نامه فارسی به انگلیسی

ماشین بولتزمن محدود . . .	Restricted Boltzman machine	استخراج ویژگی . . . . .
پرسپترون چندلایه . . . . .	Multi layer perceptron	اعتبارسنجی . . . . .
یادگیری انتقالی . . . . .	Transfer learning	الگوریتم ژنتیک . . . . .
یادگیری عمیق . . . . .	Deep learning	انتشار به جلو . . . . .
کاهش بعد . . . . .	Dimension reduction	انتشار به عقب . . . . .
کاهش گرادیان تصادفی . . .	Stochastic gradient descent	Fundamental . . . . .
حافظه کوتاه-مدت ماندگار . . .	Long short-term memory	Overfitting . . . . .
		تابع فعال‌سازی . . . . .
		تحلیل مؤلفه‌های اساسی . . . . .
		analysis
		حذف تصادفی . . . . .
		خودمرکزدارها . . . . .
		خوشه‌بندی . . . . .
		سلسله‌مراتبی . . . . .
		شبکه‌های بازگشتی . . . . .
		شبکه‌های عصبی . . . . .
		شبکه‌های پیچشی . . . . .
		networks
		صحت . . . . .
		لایه اتصال کامل . . . . .
		لایه ادغام . . . . .
		Convolutional layer . . . . .
		ماتریس سردرگمی . . . . .
		ماشین بردار پشتیبان . . .

# واژه‌نامه انگلیسی به فارسی

## F

Feature extraction .....	استخراج ویژگی .....
Forward propagation .....	انتشار به جلو .....
Fully connected layer .....	لایه اتصال کامل .....
Fundamental .....	بنیادی .....

## A

Accuracy .....	صحت .....
Activation function .....	تابع فعال‌سازی .....
Autoencoders .....	خودمرمزگذارها .....
Backward propagation .....	انتشار به عقب .....

## B

Genetic algorithm .....	الگوریتم ژنتیک .....
	Backward propagation .....

## C

Hierarchical .....	سلسله‌مراتبی .....
	Clustering .....
	Confusion matrix .....
	Convolutional layer .....
L	Convolutional neural networks .....

حافظه کوتاه-مدت ماندگار .. .  
memory

## D

M	یادگیری عمیق .....
	Dimension reduction .....
Multi layer perceptron .. .	پرسپترون چندلایه .. .

**V**

اعتبارسنجی . . . . . Validation . . . . . شبکه‌های عصبی . . . . . Neural networks . . . . .

**O**

بیشبرازش . . . . . Overfitting . . . . .

**P**

لایه ادغام . . . . . Pooling layer . . . . .  
تحلیل مؤلفه‌های اساسی . . . . . Principal component analysis . . . . .

**R**

شبکه‌های بازگشتی . . . . . Recurrent neural networks . . . . .  
ماشین بولتزمن محدود . . . . . Restricted Boltzman machine . . . . .

**S**

کاهش گرادیان تصادفی . . . . . Stochastic gradient descent . . . . .  
ماشین بردار پشتیبان . . . . . Support vector machine . . . . .

**T**

یادگیری انتقالی . . . . . Transfer learning . . . . .

# نمایه

- تابع فعال‌سازی، ۶  
کاهش بعد، ۲۱  
کاهش گرادیان تصادفی، ۱۲  
یادگیری انتقالی، ۵  
یادگیری عمیق، ۲  
استخراج ویژگی، ۲۲  
اعتبارسنجی، ۷۶  
الگوریتم ژنتیک، ۲  
انتشار به جلو، ۱۰  
انتشار به عقب، ۱۰  
بنیادی، ۲  
بیش‌برازش، ۹  
تحلیل مؤلفه‌های اساسی، ۵  
حافظه کوتاه-مدت ماندگار، ۱۴  
حذف تصادفی، ۵  
خودرمنذارها، ۳  
خوشبندی، ۵  
سلسله‌مراتبی، ۲۸  
شبکه‌های بازگشتی، ۳  
شبکه‌های عصبی، ۲  
شبکه‌های پیچشی، ۳  
صحت، ۷۸  
لایه اتصال کامل، ۷  
لایه ادغام، ۷  
لایه پیچشی، ۷  
ماتریس سردرگمی، ۷۸  
ماشین بردار پشتیبان، ۲  
ماشین بولتزمن محدود، ۳  
پرسپترون چندلایه، ۳

**Abstract:**

Prediction of financial markets has always been an important topic for scientist of different fields including statistics, finance, mathematics and computer science. The aim of this thesis to represent deep learning based frameworks for prediction of directional movements of financial markets. Among various deep learning algorithms, convolutional neural networks (CNN) are the newer ones. In this thesis, three CNN-based frameworks have been developed. The first framework, CNNpred, takes advantage of rich initial feature set which includes technical indicators, the exchange rate of currencies, future contracts, famous indices around the world and so on to forecast different financial markets. Being uniquely designed to handle utilized financial data and paying attention to the correlation between different markets are two other advantages of CNNpred. CNNpred has been applied for prediction of the 5 major U.S. indices. Results show its statistically significant superiority over the other baseline algorithm. The second framework, HT-CNNpred, improves the CNNs performance by wisely selection of the initial weights using labeled data. Introducing hierarchy training procedure has resulted in successfully prediction of 458 stocks in the U.S. market and outperforming the other baseline algorithms. Incorporating transfer learning gives the HT-CNNpred the ability to forecast new markets and successfully prediction of 14 famous indices around the world is a clear sign of that. The last framework, C-CNNpred, utilizes a clustering algorithm to put similar markets in one cluster and then selects the representative of that cluster to be used as extra information in prediction of the target market. Having this approach in prediction of Tehran Stock Exchange Index, has caused CNNpred to outperform the other baseline algorithm.

**Keywords:** Financial Markets Prediction, Deep Learning, Convolutional Neural Networks, Transfer Learning , Clustering



**University of Tehran  
Faculty of New Sciences and Technologies**

# **Using Deep Learning in Financial Markets Prediction**

A Thesis submitted to the Graduate Studies Office  
In partial fulfillment of the requirements for  
The degree of Master of Science  
in Decision Science and Knowledge Engineering

**By:**  
**Ehsan Hoseinzade**

**Supervisor:**  
**Dr. Saman Haratizadeh**

Summer 2019