مقایسه عملکرد مدلهای شبکه عصبی مصنوعی واتورگرسیون برداری در پیش بینی شاخص قیمت و بازده نقدی

مولود فامیلیان[1]

عضو هیات علمی دانشگاه شریعتی

سیما یزدانی

دانشجوی دکتری حسابداری دانشگاه تهران

تاریخ دریافت: ۱۳۹۲/۰۲/۲۲، تاریخ پذیرش: ۱۳۹۲/۰۷/۱۳

چکیده

هدف این مقاله تجزیه و تحلیل های اقتصادی، پیش بینی صحیح و دقیق متغیرهای اقتصادی است. در این زمینه، روشهای مختلفی برای پیش بینی در اقتصاد وجود دارد، که از جمله آنها میتوان به مدلهای رگرسیون ، معادلات همزمان و... اشاره کرد. مدلهای سری زمانی نیز از جمله مدلهای اقتصادی می باشند که در آن پیش بینی مقادیر سری، بیش از هر چیز به عهده خودشان گذاشته می شود اما استفاده از روش های غیر کلاسیک در شناسایی مدل و پیش بینی رفتار سیستم های پیچیده، مدتهاست در محافل علمی و حتی حرفه ای متداول و معمول شده است. در بسیاری از سیستم های پیچیده و خصوصاً غیر خطی که مدل سازی و به دنبال آن پیش بینی و کنترل آنها از طریق روش های کلاسیک و تحلیلی امری بسیار دشوار و حتی بعضاً غیر ممکن می نماید، از روش های غیر کلاسیک که از ویژگی هایی همچون هوشمندی، مبتنی بر معرفت و خبرگی برخوردار هستند، استفاده می شود. شبکه های عصبی، یکی از این روش های بدیع و در حال تحول است که در موضوعات متنوعی از قبیل الگوسازی، شناخت الگو، خوشه بندی و پیش بینی به کار رفته و نتایج مفیدی داشته

---

[1] نویسنده مسئول

است. در این مقاله، از شبکه های عصبی در پیش بینی سری های زمانی داده های اقتصادی استفاده کرده ایم. در این رابطه عوامل مختلف ساختاری، روش های مختلف یادگیری شبکه های عصبی و انتخاب و کاربرد مناسب داده ها در فرایند پیش بینی، مورد ارزیابی و بررسی قرار گرفته است و با مدل‌های ساختاری و سری زمانی مانند اتورگرسیون برداری مقایسه گردیده است. در این پژوهش، از ابزارهای محاسباتی نرم افزار MATLAB و شاخص قیمت و بازده نقدی بورس اوراق بهادار تهران مابین سال های ۱۳۸۵ تا ۱۳۹۰ استفاده شده است. خلاصه نتیجه گیری و پیشنهادات حاکی از آن است که مدل شبکه های عصبی مصنوعی از عملکرد بهتری در پیش بینی مقادیر متغیرها برخوردار است.
واژگان کلیدی: شبکه های عصبی، مدلهای اتورگرسیو، شاخص نقدی و قیمت

مقدمه

مجموعه اقتصاد هر کشور از بخش های مختلفی تشکیل شده که چگونگی روابط بین این بخش ها سمت و سوی اقتصاد آن کشور را مشخص می کند بازارهای سرمایه نیز که در کنار بازارهای پول، تشکیل دهنده بازارهای مالی می باشند. به عنوان بخشی از مجموعه اقتصاد، تابع آن بوده و چنانچه رابطه منطقی با سایر بخش ها نداشته باشند. احتمال وجود معضلات و کاستی هایی در سازکار و عملکرد آنها می رود. با رشد روزافزون بازارهای سرمایه و سهام در دهه های اخیر این بازارها نقش اساسی در اقتصاد کشورهای مختلف ایفا نموده، مهمترین جزء این بازارها را می توان شاخص بازده نقدی و قیمتی سهام قلمداد کرد. که آینه تمام نمای بورس اقتصاد کشور تلقی می شود و به عنوان یکی از متغیر پایه در بسیاری از پژوهش ها اقتصادی مد نظر قرار می گیرد. این نوع شاخص‌ها،همواره به عنوان یک ابزار مهم در ارائه بازخوردهای درونی و بیرونی اقتصاد، هم چنین در سایر سیاست های سرمایه گذاری توسط افراد،موسسه های خصوصی و دولتی و شرکت ها و نهادها و... مورد توجه قرار گرفته است [۴].

استفاده از روشهایی برای برآورد دقیق‌تر پیش‌بینی ها، به سرمایه‌گذاران نیز در تصمیمگیری

کمک می نماید.اگرچه استفاده از مدلهای خطی پیشرفته، پیش بینی های مناسبی در دوره های زمانی میان مدت و کوتاه مدت دارند اما بررسی ها در بازار سرمایه نشان میدهد رفتار سهام از یک الگوی خطی تبعیت نمی کند و الگوهای خطی تنها بخشی از رفتار بازار را نشان می دهند. شبکه های عصبی مصنوعی بعنوان یک سیستم هوشمند می توانند رابطه غیر خطی بودن ورودی ها و خروجی ها را بر اساس مجموعه داده ها، تشخیص و روابط بنیادین بین آنها را شناسایی نمایند. مسأله اصلی این تحقیق، یافتن روشی برای پیش بینی بهتر قیمت سهام است که با استفاده از این روش به دنبال پاسخ به این سوال است که: آیا استفاده از شبکه عصبی مصنوعی، پیش بینی بهتری از بازده نقدی سهام نسبت به مدل های اتورگرسیو ارائه می کند؟

مبانی نظری تحقیق

جهت درک ویژگی های مدلسازی به کمک شبکه های عصبی، لازم است ابتدا مدلسازی کلاسیک را در نظر بگیریم: مدلسازی کلاسیک از نخستین قدم، خطای بزرگی را مرتکب می شود که فقط در سیستم های ساده (خطی یا نزدیک به خطی) قابل صرف نظر است. نخستین قدم در روش کلاسیک برای بررسی داده ها بررسی شاخص های تمایل به مرکز (میانگین،...) و شاخص های پراکندگی (انحراف معیار،...) است. از این مرحله به بعد در روش کلاسیک، کاری با تک تک نمونه ها نداریم و این یک اشکال مهم است. در واقع، روش کلاسیک با عملی شبیه آسیاب کردن داده ها، پیچیدگی روابط آنها را محو می کند و به این دلیل از کشف این پیچیدگی ها باز می ماند. همچنین در روش کلاسیک، یک سیستم خواهید داشت که داده های جدید را بدون در نظر گرفتن اثر همراهی پارامترهایش با هم استفاده می کند و مجدداً این خطا در پیشگویی اثر R توسط سیستم شما تاثیر خواهد داشت. به این ترتیب سیستم کلاسیک در «استخراج» معنی از داده ها، ضعیف و با بازده پایین عمل می کند و در بسیاری از موارد از کشف روابط بین داده ها ناکام می ماند. اگر می توانستیم سیستمی داشته باشیم که با اهمیت دادن به فرد فرد مثال ها تجزیه و تحلیل کند و نیز بدون پیشداوری در مورد شکل تابع هر پارامتر (خطی بودن یا غیر خطی) آن را ذخیره و ارزیابی کند، چنین سیستمی می توانست نتایج بیشتری را از عمق داده ها بیرون بکشد.

عملکرد شبکه های عصبی مصنوعی

در حالت کلی شبکه های عصبی می توانند به عنوان جعبه های سیاهی در نظر گرفته شوند که ورودی را دریافت نموده و خروجی ایجاد می کنند. شبکه های عصبی اطلاعات ورودی را پردازش می کنند. این فرایند مبتنی بر تجزیه موازی اطلاعات پیچیده به اجزای اصلی آن است. در یک شبکه عصبی اطلاعات پیچیده به اجزای اصلی آن تجزیه شده و این اجزاء و رابطه های آنها با یکدیگر در وزن های شبکه که به مثابه حافظه شبکه عمل می کنند، ذخیره می شوند. [۱۳].

برای حل مسائل شبکه های عصبی ۳ مرحله را طی می کنند. آموزش، آزمایش و اجرا. آموزش فرآیندی است که در آن شبکه «می آموزد» تا الگوی موجود در ورودی را که به صورت مجموعه داده های آموزشی است «بشناسد». برای این منظور هر شبکه عصبی از مجموعه ای از قوانین یادگیری که نحوه یادگیری را تعریف می کنند استفاده می کنند. «تعمیم» یا آزمایش، توانایی شبکه را برای ارائه جواب قابل قبول در قبال ورودی هایی که جزو مجموعه آموزشی نبوده اند مورد سنجش قرار می دهد. استفاده از شبکه برای انجام وظیفه ای که به آن منظور طراحی شده است را اجرا می گویند.

روش های مختلف پیش بینی با شبکه عصبی

در سال های اخیر در منابع مختلف بیان شده است که بازده سهام رابطه ای غیر خطی با متغیرهای مالی و اقتصادی دارد و شبکه های عصبی ابزار بسیار مناسبی برای حل مسائل غیر خطی می باشند. در این مراجع، برای تشریح راحت تر و امکان مقایسه ساده تر مقالات مختلف در گروه های جداگانه ای بررسی شده و در هر گروه منابع موجود توضیح داده شده اند. این ۴ گروه عبارت اند از ۱. دادگان، ۲. متغیرهای ورودی، ۳. روش های پیش بینی و ۴. معیارهای ارزیابی.

الف. دادگان: در بعضی از منابع، شاخص بازار را برای کار پیش بینی مورد استفاده قرار داده اند. یکی از دلایل این کار این است که شاخص بازار نماد کاملی از تمام شرکت های موجود در بازار است، کارایی بازار براساس آن تعیین می شود، دید کلی تری را به سرمایه

گذاران می دهد و نمایانگر روند کلی بازار می باشد. جدول ۱ تعدادی از بازارهایی که توسط شبکه های عصبی پیش بینی شده اند را نشان می دهد.

ب. ورودی ها: انتخاب ورودی های درست یکی از عوامل اصلی تاثیرگذار در کارایی مدل پیش بینی می باشد. شبکه های عصبی در صورتی قادر به کشف درست اطلاعات پنهان هستند که ورودی های مناسبی انتخاب گردد.

جدول ۱ - بررسی منابع از نظر دادگان مورد استفاده

مقاله ها	دادگان مورد استفاده
ریفنس،۱۹۹۷	Euronext Paris Stock Exchange(CAC)
هاکریتر،۱۹۹۷	German Stock Exchange(DAX)
براون ،۱۹۹۸وجنکی،۱۹۹۸	Dow Jones Industrial Average(DJIA)
لم،۲۰۰۰	Hang Seng Stock Exchange in Hong Kong(HSI)
کیم،۱۹۹۸	Korea Stock Exchange(KOSPI)
لیونگ،۲۰۰۰	New York Stock Exchange(NYSE)
کایو،۲۰۰۱وکی،۱۹۹۹	Standard and Poor's 500 Stock Exchange(S&P 500)
کیم،۲۰۰۰	Singapore Stock Price
خالوزاده،۱۹۹۹	Tehran Stock Price Index(TEPIX)

در بازار سهام متغیرهای زیادی برای آنالیز بازار مورد استفاده قرار می گیرد. بسیاری از این متغیرها برای پیش بینی نامناسب بوده، اضافی هستند یا کارایی چندانی ندارند. جهت کمینه کردن اثر متغیرهای اضافی و نادرست در یادگیری شبکه های عصبی بایستی متغیرهایی را انتخاب نمود که دارای ارتباط معناداری با خروجی باشند. جدول ۲ متغیرهای ورودی مورد استفاده در منابع مختلف را نشان می دهد.

جدول ۲ - متغیرهای ورودی مورد استفاده در پیش بینی

مقاله	متغیرهای ورودی
براون ،۱۹۹۸	شاخص داوجونز، بازده داوجونز، اندیکاتور افزایش/کاهش قیمت
چاندرا،۱۹۹۹	شاخص
جنکی،۱۹۹۸	سیگنال بازده، سیگنال خرید و فروش
لم،۲۰۰۰	میانگین متحرک و تفاضل سود
کیم،۱۹۹۸	مقدار شاخص
لیونگ،۲۰۰۰	سود شاخص، سطح قیمت مصرف کننده، میزان تولید صنعتی
کی،۱۹۹۹	نسبت P/E، نرخ تورم، نرخ تولید صنعتی، نرخ رشد پول
ریفنس،۱۹۹۷	میزان بازده، سود سهام، نرخ بهره، سود هر سهم، نرخ مبادله ارز
دسای،۱۹۹۸و دسای،۱۹۹۸	T-bill، بازده NYSE، شاخص S&P، سودها
هاکریتر،۱۹۹۷	MACD,RSI,DAX index
سعد،۱۹۹۸	قیمت سهام بر اساس فرمول خاص
وود،۱۹۹۶	میانگین متحرک شاخص

ج. روش های پیش بینی: روش های مختلف هوش مصنوعی برای پیش بینی بورس مورد استفاده قرار می گیرد. جدول ۳ روش های هوش مصنوعی ارائه شده در مقالات مختلف را نشان می دهد. ANN اشاره به روش هایی دارد که از شبکه عصبی استفاده می کنند. ARIMA روشی است که در آن یک مدل به سری های زمانی برازش شده و پیش بینی مشاهدات آینده با استفاده از آن مدل انجام می شود.

جدول ۳ - روش ها مختلف شبکه عصبی که در مراجع مختلف استفاده شده است

	داربلای، ۱۹۹۸	کیم، ۱۹۹۸	او، ۲۰۰۲	لیونگ، ۲۰۰۰	استینر، ۱۹۹۷	سهی، ۱۹۹۸
ANNs		✓	✓	✓	✓	✓
ARIMA	✓			✓		

مدل های مختلف شبکه های عصبی برای پیش بینی سری های زمانی مورد استفاده قرار گرفته است. برای اینکار مدل ها را از زوایا و دیدگاه های مختلفی مورد بررسی قرار می دهیم که شامل خروجی های شبکه، پیش پردازش داده ها، نوع شبکه ها، ساختار شبکه ها و روش آموزش می باشد. در جدول ۴ خلاصه ای از این روش ها آورده شده است. در این جدول MLP شبکه های عصبی پرسپترون چند لایه، TNN شبکه های عصبی مبتنی بر تاخیر زمانی[1] و RNN شبکه عصبی های تکرار پذیر[2] هستند.

جدول ٤ - خلاصه ای از انواع شبکه های عصبی مورد استفاده.

مقاله	نوع شبکه	لایه ها	تابع انتقال
براون، ۱۹۹۸	RNN	49-10-1	N/P
چاندرا، ۱۹۹۸	MLP	10-{3-10}-1	N/P
دسای، ۱۹۹۸	MLP	11-15-15-1	Sig-Sin-Sig
جنکی، ۱۹۹۸	MLP	6-{1-10}-1	Sig-Sig
هاکریتر، ۱۹۹۷	MLP	3-8-1,5-8-1,12-9-1	N/P
کیم، ۱۹۹۸	TNN	{5-8}-{2-15}-1	N/P

1-Time Delay Neural Network
2-Recurrent Neural Network

Sig-Sig	12-{4-10}-1	MLP	کیم، ۲۰۰۰
Sig-Sig	{5-6}-{5-6}-1	MLP	او، ۲۰۰۲
Tan-Tan	6-5-1	MLP	لم، ۲۰۰۰
N/P	5-{6-10}-1	MLP	لیونگ، ۲۰۰۰
Sig-Sig	8-3-1	MLP	وود، ۱۹۹۶

د. معیارهای ارزیابی: برای ارزیابی کارایی روش های پیش بینی سری های زمانی بایستی از یک معیار ارزیابی استفاده کرد. معیارهای ارزیابی متنوعی توسط مولفان مختلف مورد استفاده قرار گرفته اند که در جدول ۵ آمده است.

جدول ۵ – خلاصه ای از معیارهای ارزیابی

نام اختصاری	معیار ارزیابی	مرجع استفاده کننده
MSE	میانگین مربعات خطا	دسای،۱۹۹۸
NMSE	میانگین مربعات خطای نرمالیزه شده	کیم،۱۹۹۸
RMSE	ریشه میانگین مربع خطا	او،۲۰۰۲؛کی،۱۹۹۹؛ ریفنس،۱۹۹۷؛ سیکمن،۲۰۰۱؛ وود،۱۹۹۶
MAE	میانگین قدر مطلق خطا	او،۲۰۰۲ ؛ کی،۱۹۹۹
MAPE	میانگین قدر مطلق درصد خطا	او،۲۰۰۲
Sharpe ratio	ریشه میانگین مربع خطا	او،۲۰۰۲ ؛کی،۱۹۹۹
RETURN	بازگشت سرمایه	کی،۱۹۹۹؛ چاندرا،۱۹۹۹؛ لیونگ،۲۰۰۰
Index Profit	میزان بازگشت سرمایه نسبت به پیش بینی	او،۲۰۰۲

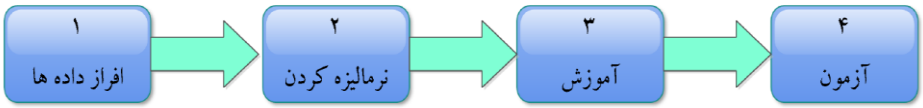
پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

روش های مبتنی بر شبکه های عصبی

فرض می کنیم داده های بازده نقدی قیمت سهام بازار بورس تهران را به صورت زمانی در دست داریم. شبکه عصبی با استفاده از نرم افزار متلب به گونه ای طراحی می شود که

شاخص قیمت سهام را از شروع دوره تا زمان t دریافت کند سپس با استفاده از این داده ها آموزش دیده و قادر است شاخص قیمت سهام در بازه زمانی t تا پایان دوره ای که مایل به پیش بینی آن هستیم را پیش بینی کند. این پیش بینی به عنوان آزمونی برای عملکرد مدل طراحی شده می باشد. شمای کلی کار در شکل (۱) نشان داده شده است.

شکل ۱ – نمودار مراحل پیش بینی بازده نقدی قیمت سهام توسط شبکه های عصبی.



الف. افراز داده ها: برای اینکه بتوانیم از داده ها برای پیش بینی توسط شبکه عصبی استفاده کنیم، داده ها را به دو بخش داده های آزمایشی و داده های آزمون تقسیم می کنیم. داده های آزمایش داده هایی هستند که شبکه توسط آنها آموزش داده می شود و داده های آزمون، داده هایی است که برای آزمون کارایی شبکه به کار برده می شود. در واقع داده های آزمون، داده هایی است که به شبکه داده می شود تا شبکه میزان کارایی خود را برای پیش بینی آینده مشخص سازد. برای این کار بسیاری از محققین، ۸۰ درصد داده ها را برای آموزش و ۲۰ درصد داده ها را برای آزمون پیشنهاد می کنند [۶و۱۲و۹].

ج) آموزش: مرحله آموزش وقتی قابل انجام است که داده های آموزش به همراه پاسخ آنها موجود باشند. این داده ها را داده های ورودی آموزش و پاسخ آنها را داده های خروجی آموزش می نامیم. در مساله پیش بینی بورس که هدف این است که قیمت سهام را با استفاده از قیمت سهام t روز قبل به دست آوریم باید به سیستم یادگیرنده، داده های t روز متوالی را به عنوان داده های ورودی آموزش و داده های روز $t+1$ های متوالی برای هر یک از این روزها را به عنوان داده های خروجی پیش بینی وارد کنیم تا سیستم یادگیری بتواند از مشاهده t روز متوالی و $t+1$ روز متوالی، مدل را یاد بگیرد.

د) آزمون: در مرحله آزمون ما می خواهیم مطمئن شویم که آیا مدل یادگیری می تواند با دریافت داده هایی به غیر از داده های وارد شده به مدل یادگیری در مرحله قبل، پیش بینی

را به درستی انجام دهد؟ برای این کار قسمتی از داده ها را که به سیستم وارد نکرده ایم و اصطلاحاً داده های ورودی آزمون نامیده می شود را به مدل شبکه عصبی وارد می کنیم. سیستم با توجه به آنچه در مرحله یادگیری آموخته است، خروجی را محاسبه می کند. جواب های بدست آمده را ستاده های خروجی آزمون می نامیم.

مبانی نظری روش اتورگرسیون برداری (VAR)

با گسترش روزافزون علم اقتصادسنجی طی دهه های گذشته، رویکرد به سمت الگوهای پویا برای پیوند دادن روابط کوتاه مدت با بلندمدت و تعیین و پیش بینی اثرات و شوک های احتمالی وارد بر سیستم اقتصادی از اهمیت بالایی برخوردار گشته است. مدل های VAR یکی از ابزارهای اقتصادی می باشد که از قدرت خوبی برای پیش بینی متغیرهای اقتصادی برخوردار است. این روش بر تئوری خاص اقتصادی دلالت ندارد و مدلی غیر ساختاری نامیده می شود. این نکته از نظر اقتصاددانان کلاسیک ضعف عمده مدل شمرده می شود، ولی سیمز در مطالعه خود توانست بر ایرادات مطرح شده پاسخ دهد. امروزه از مدل های VAR به طور گسترده در پیش بینی و تعیین متغیرهای اقتصادی استفاده می شود. از ابزار تابع عکس العمل برای پیگیری ، نحوه و دوره تأثیر شوک ها در مدل استفاده می شود و از ابزار تجزیه واریانس به تأثیر متغیرهای دیگر سیستم بر متغیر مورد مطالعه استفاده می شود.

در تحلیل چند متغیره سری های زمانی باید ارتباط درونی سری های زمانی بررسی شود. وقتی تعداد متغیرهای مدل بیش از دو باشد، ممکن است بیش از یک بردار هم انباشتگی بین متغیرها وجود داشته باشد. در این شرایط برای اینکه ارتباط تمامی متغیرها با همدیگر در نظر گرفته شود، از روش های تک معادله ای نمی توان استفاده کرد. یکی از روش های مرسوم برای این امر، استفاده از مدل معادلات هم زمان است. ویژگی که این مدل دارد، این است که متغیرهایی باوقفه درون زا، به عنوان متغیرهای از پیش تعیین شده وارد مدل شده اند و به همین دلیل به چنین مدلی «سیستم معادلات هم زمان پویا» گویند. در این مدل دو مشکل اساسی داریم: یکی تقسیم بندی متغیرهای سری زمانی به درون زا و برون زا و دیگری اعمال قیود بر روی پارامترهای مدل به منظور تشخیص و شناسایی الگو. سیمز با

استفاده از قیود اعمال شده برای شناسایی الگوها معتقد است که این قیود با تئوری اقتصادی پویا سازگار نیستند. وی بر مبنای تحقیقات خود روشی را معرفی کرد که گروه متغیرهای برون زا وجود نداشته و تمامی متغیرها را به صورت درون زا در نظر گرفت. فقط یک دسته معادلات با وقفه های مساوی برای همه متغیرها تخمین زده می شود. سیمز این روش را خودرگرسیون برداری نامید.

پیشینه تحقیق:

پژوهش های انجام شده را می توان در سه گروه طبقه بندی کرد: ۱) پیش بینی قیمت با استفاده از مدل های خطی ۲) پیش بینی قیمت با استفاده از مدل های شبکه عصبی و ۳) مقایسه پیش بینی های انجام شده توسط دو دسته مدل های فوق.

فرلاند و لالاسنت(۲۰۱۰) در مقاله ای تحت عنوان "مدلسازی دینامیک نوسانات واقعی" نشان دادند مدل سری های زمانی با پایه ARIMA، ضعیف تر از شبکه های عصبی پیش خور عمل می کنند و این نتیجه را از مقایسه ناپایداری ها و همبستگی های نتایج روش های متفاوت پیش بینی بدست آورده اند.

هوارنگ و آنگ (۲۰۱۱) در مقاله "کاربرد شبکه عصبی برای مدلسازی سری های زمانی" نشان دادند شبکه های عصبی چند لایه پس خور از مدل ARMA بهتر عمل می کند.

خانسا (۲۰۱۱) در تحقیق خود با عنوان " پیش‌بینی بازده بازار سهام شرکت های امنیت اطلاعات با استفاده ازشبکه های عصبی مصنوعی و مدل اتورگرسیو" با اشاره به اهمیت رو به رشد کسب و کار مبتنی بر اینترنت، عنوان می کند که حملات سایبری مخرب بر روی زیرساخت های فناوری اطلاعات در حال افزایش بوده است و این حملات مخرب با اثرات مضر اقتصادی بر شرکت های مورد حمله همراه خواهد بود.از سوی دیگر آنها بیان کردند که حملات شدیدتر با افزایش قیمت سهام شرکت های امنیت اطلاعات همراه خواهد بود. علاوه بر این، ایشان با استفاده از شبکه های عصبی مصنوعی و مدل VAR به عنوان روش مکمل برای بررسی ارتباط بین بازده بازار سهام شرکت های امنیت اطلاعات و شدت حملات مخرب، به عنوان محصول حملات مخرب و میزان شدت آنها پرداخته است.

حسین هادی پور در تحقیق "تعیین بهترین مدل برای پیش بینی قیمت سهام در گروه صنایع

غذایی و آشامیدنی بورس اوراق بهادار تهران" به پیش بینی قیمت سهام در گروه صنایع غذایی پرداخته است. به این ترتیب که نخست قیمت هفتگی سهام در سال های ۷۷ و ۷۸ و ۹ ماهه اول سال ۷۹ را برای این شرکت ها جمع آوری و سپس به پیش بینی با مدل های هموارسازی نمایی، میانگین متحرک و باکس جنکینز برای ۱۴ هفته انتهایی سال ۷۹ پرداخته و به مقایسه شاخص MSE این روش ها با یکدیگر پرداخته است. در نهایت این پژوهش بر این نکته اصرار می ورزد که مدل مشخصی برای پیش بینی قیمت سهام در گروه صنایع غذایی و آشامیدنی وجود ندارد و هر سری داده های قیمت سهام، دارای روند، ویژگی ها و محدودیت های ویژه ای به خود می باشد و برای پیش بینی قیمت سهام هر شرکت نخست باید با استفاده از روند و ویژگی های آن سری زمانی داده ها، مدل مناسب را با استفاده از روش شناسی مدل های پیش بینی گزینه و سپس با استفاده از آن مدل به پیش بینی قیمت سهام آن شرکت اقدام نمود.

احمد حاتمی در تحقیق "شناسایی عوامل موثر بر قیمت سهام در بورس اوراق بهادار تهران و ارائه مدلی ریاضی برای پیش بینی تغییرات آن" به بررسی عوامل موثر بر شاخص قیمت سهام پرداخته و در نهایت مدلی با استفاده از رگرسیون خطی برای پیش بینی قیمت سهام با متغیرهای نرخ ارز، تورم و وقفه اول شاخص قیمت ارائه داده است.

علی رجب زاده قطرمی در تحقیق "ارزیابی ترکیبی روش های پیش بینی و ارائه مدل بهینه برای پیش بینی قیمت سهام در بورس اوراق بهادار تهران" به بررسی روش های پیش بینی و ترکیب این روش ها با هدف کاهش خطای پیش بینی پرداخته است. روش های پیش بینی، روش های پیش بینی سری زمانی (تک متغیره) هستند که به داده های تاریخی برای برآورد مقادیر آینده اتکا دارند. در این بررسی، هر یک از روش های پیش بینی سری زمانی روش های فردی و ترکیب این روش ها، روش های ترکیبی نام دارند. جهت ترکیب روش های فردی از مدل رگرسیون چند متغیره (چند گانه)، که توانایی انجام آزمون های آماری مختلف برای تایید مدل را دارد، استفاده شده و در آن مقادیر نتایج روش های فردی، متغیرهای مستقل و پیش بینی ترکیبی، متغیر وابسته تعریف شده است. داده های جمع آوری شده، قیمت سهام شرکت پارس الکتریک برای سه سال با دوره های هفتگی

داده های واقعی در بورس تهران بوده و پیش بینی قیمت سهام با روش های مختلف برای ۱۴ دوره انجام شده، که از بین این روش ها، ۶ روش که سازگاری بیشتر با داده ها داشته و دارای خطای کمتری اند، انتخاب شده و در ترکیب استفاده شده اند.

شهاب الدین مقرب الحق در مقاله "رویکرد شبکه های عصبی در بهینه سازی پیش بینی سری های زمانی" با استفاده از جعبه ابزار شبکه عصبی نشان داده است که مدل های شبکه های عصبی، اگر به طور مناسب از لحاظ ساختار درونی و همچنین استفاده از تئوری های آماری و اقتصادی طراحی شوند، در پیش بینی سری های زمانی، بر سایر مدل های رقیب برتری خواهند داشت.

مسعود عزیزخانی در مقاله "بررسی روش های ترکیبی پیش بینی و ارائه مدل بهینه برای پیش بینی قیمت سهام در بازار بورس تهران" از روش های خطی ARIMA، روند درجه ۲، نمو همدار، براون درجه ۲ و ترکیبی از روش های پیش بینی براساس مدل های خطی و پیش بینی براساس مدل های غیر خطی (شبکه های عصبی غیر خطی) استفاده کرده است. از تکنیک های مدلسازی براساس شبکه های عصبی برای مدلسازی رفتار قیمت سهام به صورت های مختلف استفاده شده است. در نهایت ساختاری برای پیش بینی دراز مدت ارائه گردیده است.

محمد قوام زاده در پژوهش "پیش بینی در بازارهای سازمان یافته معاملات" به پیش بینی شاخص بورس اوراق بهادار تهران پرداخته است. برای این کار قیمت دو سهم در بورس اوراق بهادار تهران، قیمت جهانی نفت و قیمت جهانی طلا به کار گرفته شده است.

حسین پناهیان در پایان نامه خود "استفاده از شبکه های عصبی برای پیش بینی روند شاخص قیمت سهام در بورس اوراق بهادار تهران" از یک شبکه پرسپترون با سه لایه پنهان جهت پیش بینی شاخص بورس اوراق بهادار تهران استفاده کرده است. متغیرهای ورودی شبکه عوامل برون زا و درون زای بازار بورس از جمله قیمت طلا، دلار، حجم معاملات و شاخص بورس است.

حمید خالوزاده در پایان نامه دکترای خود "مدل سازی غیر خطی و پیش بینی رفتار قیمت سهام در بازار بورس تهران" را به دو قسمت اصلی قابلیت پیش بینی (یا پیش بینی پذیری)

و پیش بینی تفکیک کرده است. نتیجه این تحقیق نشان می دهد که شبکه های عصبی بهتر از مدل های ARIMA عمل می کند، این تحقیق بر روی سهام شهد ایران انجام شده است.

داده های تحقیق

شاخص قیمت و بازدۀ نقدی یا همان شاخص درآمد کل با نماد TEDPIX از فروردین ۱۳۷۷ در بورس تهران محاسبه و منتشر شده است. تغییرات این شاخص نشانگر بازده کل بورس است و از تغییرات قیمت و بازده نقدی پرداختی، متأثر می‌شود. این شاخص کلیۀ شرکت‌های پذیرفته شده در بورس را در بردارد و شیوۀ وزن‌دهی و محاسبۀ آن همانند شاخص کل قیمت (TEPIX) است. تنها تفاوت میان آن دو در شیوۀ تعدیل آن‌ها است. داده‌های مورد استفاده در این پژوهش شاخص قیمت و بازده سهام بصورت روزانه از ۱۳۸۵/۱/۱ تا ۱۳۹۰/۱۲/۲۹ بوده که در مجموع ۱۳۶۷ مشاهده را در بر می گیرد. این داده ها از دو بخش تشکیل شده اند. نخست داده های مربوط به دوره ۱۳۹۰/۶/۱ تا ۱۳۹۰/۱۲/۲۹ برای اعتبار سنجی و مقایسه مدل شبکه عصبی پیشخور با فرآیند VAR مورد استفاده قرار می گیرد.در این مطالعه از شبکه عصبی پیشخور چند لایه، دارای ۲۰ نرون در لایه مخفی و تابع فعال سازی سیگموئید و لایه خروجی آن استفاده شده است. پس از تعیین تعداد وقفه های بهینه، برای انتخاب تعداد نرون های لایه مخفی شبکه، شبکه های مختلف با تعداد نرون های مخفی متفاوت طراحی شده و آموزش داده شد. از میان این شبکه ها با توجه به معیار MSE، شبکه بهینه انتخاب گردید. یعنی شبکه با کمترین MSE که دارای ۲۰ نرون مخفی بود به کار گرفته شد. لایه ورودی با توجه به داده های ورودی با ۴، ۵ و ۱۰ نرون در نظر گرفته شد. تابع فعال سازی استفاده شده برای شبکه مذکور از نوع سیگموئید است. خروجی های شبکه، دارای یک نرون و تابع فعال سازی خطی می باشد.

از میان الگوریتم های مختلفی که برای آموزش شبکه های عصبی مورد استفاده قرار می گیرد، الگوریتم لونبرگ-مارکوات انتخاب گردید چراکه این الگوریتم در مقایسه با سایر الگوریتم های پس انتشار خطا دارای سرعت بیشتری است. دوره آموزش و آزمایش، شامل

۱۳۶۷ داده می باشد. تعداد داده های آموزش و آزمایش مدل به نسبت ۰/۸۵ به ۰/۱۵ تقسیم گردید و از نرخ یادگیری ۰/۰۱ استفاده شد. برای توقف فرآیند آموزش از روش Early Stopping بهره گرفته شد. در نهایت شبکه عصبی با استفاده از نرم افزار MATLAB طراحی و تخمین زده شده است.

ضمناً بردارهای ورودی به صورت تصادفی به سه مجموعه به شرح زیر تقسیم می شوند: ۶۰٪ داده ها در مجموعه آموزشی، ۲۵٪ داده ها در مجموعه ارزیابی در راستای جلوگیری از بیش برازش شبکه، ۱۵٪ داده ها در مجموعه تست برای بررسی کارایی شبکه نهایی. دراین تحقیق از تابع train جهت آموزش شبکه استفاده شده است. روال آموزش در صورتی که خطای مجموعه ارزیابی در ۶ تکرار متوالی افزایش یابد، متوقف می شود.

برای پیش بینی شاخص بازده نقدی سهام روشهای مختلفی وجود دارد. در یک تقسیم بندی کلی آنها را به سه دسته روشهای کلاسیک، هوش مصنوعی و ترکیبی تقسیم بندی می شوند. بنابراین در تحلیل پیش بینی شاخص بازده نقدی سهام دو سناریو مورد بررسی قرار گرفته است:

سناریو اول: پیش بینی شاخص بازده نقدی سهام با رویکرد روشهای کلاسیک.

سناریو دوم: پیش بینی شاخص بازده نقدی سهام با رویکرد روشهای هوش مصنوعی.

آماده سازی شبکه عصبی: برای پیاده سازی شبکه عصبی باید مراحل تهیه داده های تاریخی، پیش پردازش داده ها، انتخاب معماری، آموزش و آزمایش شبکه را طی کرد.

تهیه داده های گذشته و قدیمی: داده های مورد نیاز پس از جمع آوری اطلاعات بورس اوراق بهادار در فایل های مربوطه قرار گرفته و توسط برنامه شبیه سازی شده است. برای شناخت بیشتر شاخص بازده قیمتی سهام در سالهای مختلف و همچنین دستیابی به آماره‌های تلخیصی به انجام آمار توصیفی بر روی این متغیرها می‌پردازیم.

جدول ۶ – آمار توصیفی متغیر تحقیق

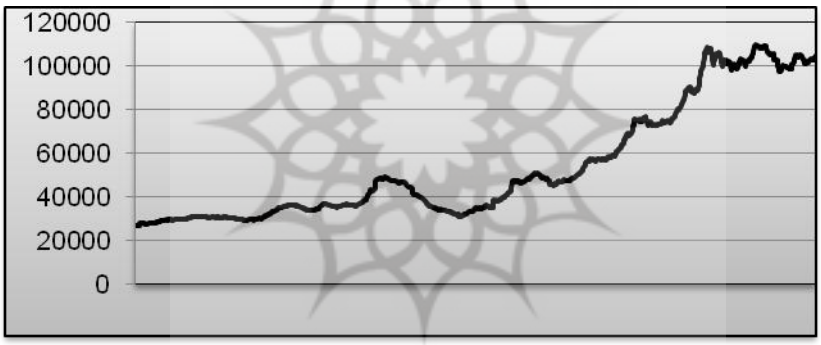
	حداقل	حداکثر	میانگین	انحراف معیار
شاخص بازده قیمتی سهام	۱۰۱۸۰,۹۰	۱۱۰۰۰۰	۵۴۱۴۴	۲۶۶۴۹

پیش پردازش داده ها: در این مقاله از روش نرمال سازی استفاده شده است به طوریکه بعد از جمع آوری داده ها، یک سری عملیات آماده سازی روی آنها انجام گردید. ابتدا کنترل شد که اعداد شاخص کامل باشند و احتمالاً عددی از قلم نیفتاده است. همچنین نقاط اخلال که در فرایند پیش بینی تاثیر نامطلوب دارند، در صورت وجود حذف گردند. در مرحله بعد داده ها مقیاس بندی یا نرمال شدند. برای این منظور از فرمول زیر استفاده گردید.

$$X_n = 2\left(\frac{X - X_{min}}{X_{max} - X_{min}}\right) - 1$$

نمودار شماره (۱) روند شاخص بازده قیمتی سهام را قبل و بعد از نرمال سازی نشان می دهد.

نمودار (۱) روند شاخص بازده نقدی و قیمتی سهام بورس اوراق بهادار تهران از ۱۳۸۵/۱/۱ تا ۱۳۹۰/۱۲/۲۹



الف) قبل از نرمال سازی

ب) بعد از نرمال سازی

توپولوژی و ساختار شبکه عصبی : در این تحقیق، شبکه عصبی چند لایه پیش خور MLFN با الگوریتم یادگیری پس انتشار BP با استفاده از نرم افزار MATLAB 7.12 پیاده سازی شده است:

- ساختار شبکه ۱-۱۰-۱۰-۴. این ساختار شبکه را ANN_1 نام گذاری کرده ایم.
- ساختار شبکه ۱-۱۰-۱۰-۵. این ساختار شبکه را ANN_2 نام گذاری کرده ایم.
- ساختار شبکه ۱-۱۰-۱۰-۱۰. این ساختار شبکه را ANN_3 نام گذاری کرده ایم.

همانطور که ملاحظه می گردد سه ساختار فوق فقط از لحاظ تعداد نرون های لایه ورودی با یکدیگر متفاوت می باشند و تعداد لایه های میانی همراه با تعداد نرون های هر لایه ثابت در نظر گرفته شده است. در حقیقت از سه الگوی ورودی متفاوت جهت پیش بینی شاخص استفاده شده است اما با مقایسه هر سه ساختار تنها به ذکر خروجی مطلوب اکتفا شده است.

در الگوی اول مقدار شاخص در پنجمین روز آینده $(t+5)$ با استفاده از مقدار شاخص در زمانهای t، $t-5$، $t-10$ و $t-15$ پیش بینی می شود. به عبارت دیگر:

$$y = x_{t+5} = f\left(x_t, x_{t-5}, x_{t-10}, x_{t-15}\right)$$

در الگوی دوم مقدار شاخص برای روز بعد با استفاده از سری پنج روز قبل آن پیش بینی می گردد:

$$y = x_{t+1} = f\left(x_t, x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}\right)$$

و در الگوی سوم مقدار شاخص در روز بعد با استفاده از سری ۱۰ روز قبل آن پیش بینی می گردد:

$$y = x_{t+1} = f\left(x_t, x_{t-1}, x_{t-2}, ..., x_{t-9}\right)$$

علت انتخاب شبکه عصبی چند لایه پیش خور $(MLFN)$ این است که در اکثر تحقیقات مشابه از این تیپ شبکه استفاده گردیده است. این تیپ شبکه قادر است که

ارتباطات پیچیده را به خوبی توابع ساده مدل سازی نماید (جانوس کویشیوس، ۲۰۰۳) این امر طی تحقیقات کاربردی بسیاری اثبات گردیده است. جدول (۷) چگونگی طراحی و مدل سازی سری های زمانی شاخص بازده نقدی و قیمتی سهام را در شبکه عصبی نشان می دهد.

جدول ۷ - طراحی و مدل سازی شاخص بازده قیمتی سهام در شبکه عصبی

نوع شبکه عصبی	پیشخور چند لایه ای	الگوریتم آموزش شبکه های عصبی	لونبرگ – مارکوات
تابع فعال سازی	سیگموئید	متد توقف فرآیند آموزش	Early Stopping
تعداد نرون ورودی	۴ – ۵ – ۱۰	دوره زمانی آموزش و آزمایش	۸۵/۱/۱ تا ۹۰/۱۲/۲۹
تعداد نرون خروجی	۱	نسبت تعداد داده های آموزش و آزمایش	۰/۸۵ به ۰/۱۵
معیار تعیین تعداد نرون های مخفی	MSE	نرخ یادگیری	۰/۰۱
تعداد لایه پنهان	۲	دوره زمانی پیش بینی	۱۰ روز آینده
تعداد نرون های پنهان	۲۰		

آموزش شبکه عصبی: قبل از آموزش شبکه عصبی برای هر یک ساختارهای فوق، ابتدا داده های تحقیق متناسب با الگوهای ورودی هر ساختار، به صورت آرایه های ورودی- خروجی تبدیل شد. این عمل از طریق کد نویسی در Toolbox نرم افزار MATLAB انجام گردید.

بعد از تبدیل داده ها به صورت رکوردهای فوق، کل رکوردهای داده به سه قسمت تقسیم شدند؛ مجموعه آموزش یا یادگیری شبکه شامل ۶۰ درصد کل رکوردها، مجموعه ارزیابی شامل ۲۵ درصد کل رکوردها و مجموعه آزمون یا تست شبکه که شامل ۱۵ درصد کل رکوردها می باشند. رکوردهای مجموعه آموزش و مجموعه ارزیابی از بین ۸۵

درصد اول داده ها به صورت تصادفی انتخاب می گردند و مجموعه تست تقریباً شامل ۲۰۰ رکورد (۱۵ درصد) آخری داده ها است.

فرآیند یادگیری شبکه به این صورت است که ابتدا رکوردهای مربوط به یادگیری به شبکه وارد می گردد. وزنه های اولیه اتصالات بین نرونها به صورت تصادفی توسط شبکه تنظیم می گردند. پس از بارگذاری داده های مجموعه یادگیری و تنظیم وزنهای شبکه، اولین رکورد داده به عنوان ورودی به شبکه اعمال شده و خروجی شبکه با استفاده از توابع و الگوریتم یادگیری محاسبه شده و با خروجی مطلوب مقایسه می گردد. در اینجا شبکه با استفاده از خطای حاصله وزنهای شبکه را تغییر می دهد. این عملیات برای کل داده های آزمایش اعمال می گردد. پس از اعمال هر ورودی وزنهای شبکه به هنگام می گردند و هر بار تکرار فرآیند فوق برای کل داده های آموزش یک epoch نام دارد. بعد از هر epoch میانگین مربعات خطا محاسبه شده و با MSE هدف مقایسه می شود. در صورتی که خطا بزرگتر از خطای هدف باشد دوباره یک epoch دیگر شروع می گردد. در اینجا شرط توقف یادگیری، رسیدن به خطای هدف می باشد. برای جلوگیری از طولانی شدن زمان یادگیری عدد ۱۰۰۰ به عنوان شرط حداکثر تکرار epoch جهت توقف مرحله یادگیری تعیین گردید.

از طرف دیگر ممکن بود قبل از رسیدن به شروط فوق، فرآیند یادگیری بر روی داده ها آموزش بیش از حد باشد در نتیجه علیرغم کاهش خطا در مجموعه آموزش، کارایی شبکه در پیش بینی داده های مجموعه آزمون کاهش یابد، لذا از مجموعه داده های ارزیابی جهت کنترل خطای شبکه استفاده گردید. به این ترتیب بعد از هر epoch، کارایی شبکه بر روی مجموعه داده ارزیابی توسط شبکه بررسی می گردد. در صورتی که خطای مجموعه داده های ارزیابی نیز رو به کاهش باشد فرایند یادگیری ادامه می یابد. در غیر این صورت یادگیری شبکه متوقف گردیده و مجموعه داده های آزمون پیش بینی می گردد.

خروجی شبکه: پیاده سازی شبکه بصورتی انجام شد که خروجی شبکه پس از هربار پیش بینی شامل نمودار خطای پیش بینی شامل خطای مجموعه آموزش، ارزیابی و تست شبکه، نمودار ارقام واقعی و پیش بینی مجموعه آموزش شبکه، نمودار ارقام واقعی و پیش بینی

مجموعه تست شبکه، و نمودار ارقام واقعی و پیش بینی مربوط به مجموعه آموزش شبکه باشد.

معیار ارزیابی و آزمون نتایج: معیار ارزیابی مدلهای پیش بینی در این تحقیق ریشه دوم میانگین مربعات خطا (RMSE) می‌باشد که از فرمول زیر محاسبه گردیده است.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_t - \hat{y}_t)}{n}}$$

n تعداد ارقام پیش بینی شده و $\hat{y}_t$: عدد پیش بینی شده می باشد.

با توجه به اینکه وزنهای اولیه شبکه به صورت تصادفی تعیین می گردند، هر بار که پیش بینی از طریق مدل شبکه عصبی صورت می‌گیرد، جواب های متفاوتی بدست می آید. در این تحقیق فرآیند پیش بینی برای هر یک از ساختارهای شبکه ۱۰ بار تکرار گردید. در ادامه کار عملیات پیش بینی کل داده‌ها به صورت انباشته (۱۳۹۰-۱۳۸۵) با استفاده از مدلهای شبکه عصبی در سه ساختار ANN_1، ANN_2 و ANN_3 و برای هر ساختار ۱۰ مرتبه انجام شد در جدول ۸ اطلاعات ارائه شده توسط نرم افزار متلب در خصوص هر پیش بینی به طور خلاصه نشان داده شده است. لازم به ذکر است که معیار خطا در جدول، MSE است و با توجه به مقیاس داده ها، در محدوده $[-1,+1]$ قابل تفسیر است.

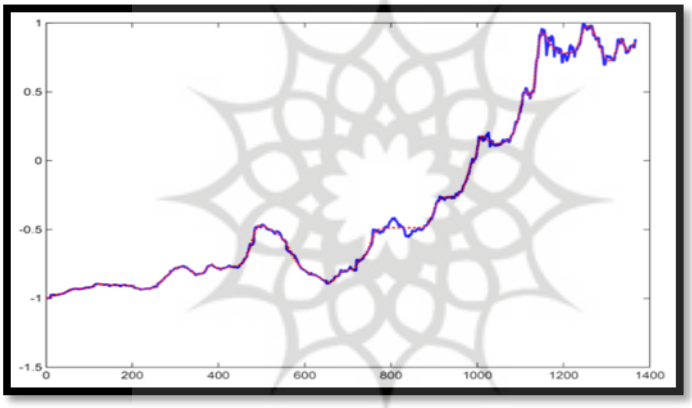
جدول ۸ - خلاصه اطلاعات استخراج شده در مورد پیش بینی های شبکه عصبی

ساختار شبکه	ANN_1			ANN_2			ANN_3		
شماره پیش بینی	Epoch	Train	Test	Epoch	Train	Test	Epoch	Train	Test
اول	125	0.0080	0.00610	129	0.007	0.0055	74	0.0076	0.0054
دوم	96	0.0135	0.012	105	0.0144	0.0112	57	0.0026	0.0022
سوم	41	0.0169	0.012	111	0.0059	0.0046	90	0.0079	0.006
چهارم	200	0.0071	0.0054	49	0.0147	0.0121	81	0.0034	0.0033
پنجم	133	0.0112	0.008	53	0.001	0.0074	186	0.0004	0.0037
ششم	170	0.0149	0.0101	118	0.0076	0.0062	49	0.0031	0.0027
هفتم	41	0.0148	0.0104	57	0.0117	0.0095	170	0.0071	0.0057

هشتم	178	0.0122	0.0101	89	0.007	0.0057	49	0.005	0.0037
نهم	114	0.0109	0.0094	105	0.0077	0.0063	51	0.003	0.0023
دهم	237	0.0083	0.0062	27	0.0126	0.0105	36	0.004	0.0033

جهت بررسی عملکرد شبکه انتخابی آموزش، نمودار مربوط به خطای آموزشی، ارزیابی و آزمایشی رسم می شود. این موارد در نمودار ۲ نمایش داده شده اند. همانطور که در نمودار مشاهده می شود خروجی مجموعه های آموزشی، ارزیابی و تست به خوبی بر روی بردارهای هدف منطبق شده اند و R-Value کل مدلسازی بیشتر از ۹۹٪ می باشد. در این مورد عکس العمل شبکه مناسب به نظر می رسد و می توان شبکه را با استفاده از تابع Sim و در مورد ورودی های جدید شبیه سازی کرد.

نمودار ۲. بررسی عملکرد شبکه



البته به لحاظ فنی با توجه به ساختار شبکه های عصبی، با افزایش تعداد دوره ها، دقت پیش بینی شبکه های عصبی کاهش می یابد. همچنین به لحاظ تئوری این امکان وجود دارد که با در نظر گرفتن انتظارات، با افزایش تعداد روزهای پیش بینی ضریب اطمینان صحت پیش بینی کمتر گردد. زیرا پیش بینی روزهای دورتر بر اساس پیش بینی روزهای قبل می باشد نه قیمت های واقعی.

آزمون و برآورد مدل VAR

برای گزینش بهترین مدل، توضیح برداری روی ۸۵ درصد داده ها اعمال گردید، سپس روی ۱۵ درصد حجم داده ها پیش بینی انجام می گردد و نهایتاً خطای حاصل از پیش بینی با استفاده از فرمول $MSE = \sum e_i^2 / n$ محاسبه خواهد شد.

بررسی مانایی متغیرهای مدل

روند این آزمون به این شکل می‌باشد که ابتدا آزمون مانایی را براساس سطح متغیر و با وقفه‌های مختلف انجام داده و وقفه‌ای که دارای کمترین مقدار معیارهای AIC و SCمی‌باشد انتخاب می‌نماییم، سپس با مقایسه آماره دیکی- فولر و مقادیر بحرانی مک‌کینون، مانایی متغیر را تعیین می‌کنیم. نتایج آزمون مانایی حاکی از این بود که متغیر قیمت و بازده نقدی پس از یک بار تفاضل گیری به متغیرهای ایستا تبدیل شدند. نتایج آزمون های فوق در جدول (۹) آمده است.

جدول شماره ۹ - آزمون ریشه واحد دیکی‌فولر تعمیم یافته برای داده‌های سری زمانی

نام سری	ADF	Mackinnon Critical Values			نتیجه آزمون
		%1	%5	%10	
Y	1.03	-3.43	-2.86	-2.56	عدم پایایی
D(Y)	-24.46	-3.43	-2.86	-2.56	پایایی در تفاضل مرتبه اول

همان‌طور که ملاحظه می شود، با توجه به اینکه مقدار قدرمطلق آماره ADF از مقادیر بحرانی در سطح ۵٪ بزرگ تر است، لذا نتیجه گیری می شود که فرضیه H_0مبنی بر وجود ریشه واحد (عدم پایایی) در سطح بالایی از درجه اطمینان (۹۵٪) رد می‌گردد.

آزمون انگل گرنجر دو مرحله ای

نتایج حاصل از پاپایی ، حاکی از آن است که متغیر تحقیق در سطح، ساکن نیست و نشانگر این است که متغیر تحقیق دارای انباشتگی از مرتبه اول می‌باشند. با توجه به درجه انباشتگی متغیر، از روش‌های معادلات همزمان متعارف و روش VAR بدون تعیین روابط

بلندمدت نمی‌توان استفاده نمود. به همین دلیل ابتدا بایستی تعیین شود که آیا روابط بلندمدت بین متغیرها وجود دارد یا خیر؟ یعنی بایستی مشخص گردد که آیا پسماندهای حاصل از مدل رگرسیون (به عنوان ترکیب خطی از دو سری ناپایای X و Y) پایاست؟
در این صورت فرضیه صفر و مقابل را می توان به صورت زیر بیان کرد:

H_0= جمله پسماند ناپایاست

H_1= جمله پسماند پایاست

بر طبق نتایج بدست آمده از آزمون ریشه واحد در باقیمانده ها که در جدول ۱۰آمده است بدلیل آنکه قدرمطلق آماره‌های به‌دست آمده بزرگتر از قدرمطلق مقادیر بحرانی در سطح ۵٪ می‌باشند، فرضیه صفر مبنی بر ناپایی جملات پسماند در سطح ۹۵٪ رد شده و در نتیجه وجود یک رابطه بلندمدت میان متغیر و وقفه های آن تایید و پذیرفته می‌شود.

جدول شماره ۱۰ - آزمون ریشه واحد بر روی جملات پسماند مدل

Augmented Dickey – Fuller test statistic		-4.761943	0.0052
Test critical Values:	1% level	-4.297073	
	5% level	-3.212696	
	10% level	-2.747676	

تعیین وقفه بهینه در مدل تحقیق

قبل از برآورد الگوی VAR و بررسی پویایی آن لازم است که طول وقفه بهینه تعیین گردد. برای این منظور از دو معیار آکائیک (AIC) و شواتز- بیزین (SBC) استفاده می‌شود. برای از دست ندادن درجه آزادی بیشتر از آماره SBC استفاده می‌شود به همین دلیل آماره SBC از دقت بیشتری برخوردار است. در نتیجه در این تحقیق طول وقفه‌ای به عنوان طول وقفه بهینه انتخاب می‌شود که مقدار آماره SBC مقدار بیشتری را نشان می‌دهد.

جدول شماره ۱۱ - تعیین تعداد وقفه های بهینه مدل

طول وقفه (Lag Intervals)	معیار شوارزبیزن (SBC)
0	1.9404
1	-6.2561
2	-6.4119*
3	-6.4065
4	-6.4026

مطابق جدول (۱۱)، برای مدل، بیشترین مقدار معیار شوارتز- بیزین در وقفه شماره ۲ قرار داشته و لذا وجود دو وقفه برای هر دو مدل تأیید می شود.

آزمون هم انباشتگی

مفهوم اقتصادی هم‌انباشتگی این است که وقتی دو یا چند سری‌زمانی بر اساس مبانی نظری با یکدیگر ارتباط داده می‌شوند تا یک رابطه تعادلی بلندمدت را شکل دهند، هرچند ممکن است خود این سری‌های زمانی دارای روند تصادفی بوده و نامانا باشند، یکدیگر را در طول زمان به‌گونه‌ای دنبال می‌کنند که تفاضل بین آنها مانا گردد. در رابطه با این آزمون ها وقتی فرضیه صفر مربوط به وجود r بردار همجمعی پذیرفته می شود که کمیت آماره آزمون مربوطه از مقدار بحرانی ارائه شده توسط جوهانسون و جوسیلیوس کمتر باشد. بنابراین با توجه به خروجی نرم افزار متلب و بر اساس هر دو آماره اثر و حداکثر مقادیر ویژه وجود یک بردار همگرایی برای مدل ارائه شده در بلند مدت پذیرفته می شود.

تخمین مدل تحقیق: در جدول ۱۲ نتیجه تخمین مدل تحقیق بیان شده است.

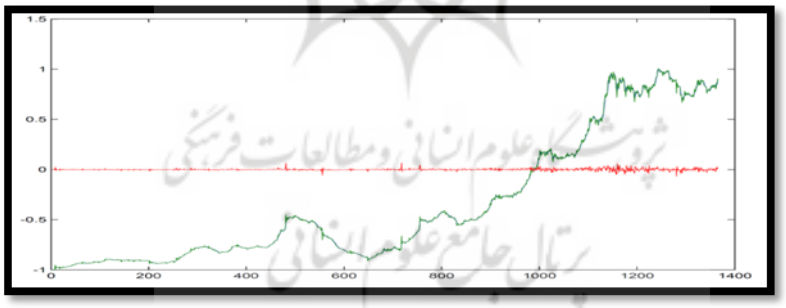
پرتال جامع علوم انسانی

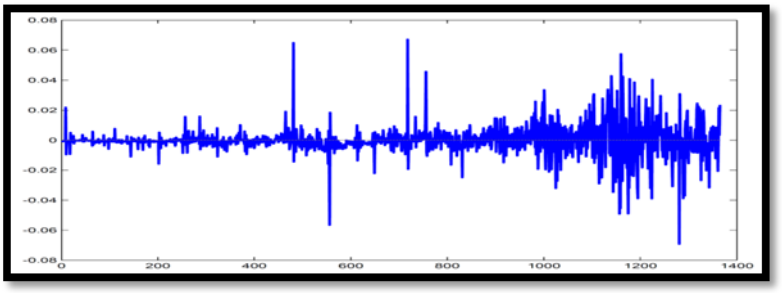
جدول ۱۲- نتایج تخمین مدل تحقیق

متغیرهای مستقل	نتایج تخمین
C	0.001
Yt-1	1.3859
Yt-2	-0.3855
R2	0.9997
ADJ. R2	0.9997
F-Stat	2927122

$$Y = 0.001 + 1.3859Y_{t-1} - 0.3855Y_{t-2}$$

با توجه به ضرایب بدست آمده، می توان مشاهده نمود که شاخص بازده قیمتی سهام در سال گذشته تاثیر مثبت و معنی دار و شاخص بازده نقدی و قیمتی سهام در دو سال گذشته تاثیر منفی و معنی دار بر شاخص بازده نقدی و قیمتی سهام در طول دوره مورد بررسی داشته است.





به منظور مقایسه قدرت پیش بینی فرآیند VAR و شبکه عصبی مصنوعی از معیارهای میانگین مربع خطای استاندارد (MSE)، مجذور میانگین مربع خطا استاندارد (RMSE)، میانگین قدر مطلق خطا (MAE) و میانگین درصد قدر مطلق خطا (MAPE) به صورت گام به گام (روز یکم تا روز دهم) محاسبه شده است. دلیل استفاده از پیش بینی گام به گام بیان این نکته است که برای مثال یک مدل، در طول ۱۰ روز پیش بینی، ممکن است در چند روز اول نسبت به مدل دیگری از قدرت پیش بینی بالاتری برخوردار باشد اما در ارزیابی مجموع ۱۰ روز این گونه نباشد، که توانایی پیش بینی این مدل در نوع دوم نادیده گرفته می شود در اینجا هر گام معرف یک روز می باشد. از نتایج بدست آمده مشخص می شود که مدل شبکه عصبی پیشخور در تمام گام ها و از نظر تمامی معیارهای عملکرد، بر روش VAR برتری قابل توجهی دارد. بر اساس نتایج بدست آمده، مدل شبکه عصبی پیشخور دارای خطای کمتر و در نتیجه کارایی بیشتری در پیش بینی روزهای آینده قیمت دارد. نباید فراموش نمود که شبکه های عصبی معمولاً به تعداد زیادی داده برای آموزش و آزمایش نیاز دارند. لذا تغییر در تعداد داده های آموزشی و آزمایش و خصوصاً کاهش تعداد داده های آموزشی، نتایج متفاوتی را به دنبال خواهد داشت. جدول ۱۳ آماره های توصیفی را در هر دو وضعیت ارایه می کند. همان طور که مشاهده می شود میانگین بازده پیش بینی شده با استفاده از مدل ANN بیشتر از مدل دیگر است.

جدول ۱۳- آمار توصیفی مدلهای تحقیق

	Mean	Std. Deviation	N
ANN	.00366520	.001748781	10
VAR	.00115390	.000131805	10

اما خروجی زیر (جدول ۱۴) کل تغییرات را در دو منبع فاکتور (۲ مدل پیش بینی) و خطا نشان می دهد. همانطور که ملاحظه می شود Sig این نگاره در هر چهار آزمون کوچک تر از سطح معنی داری ۵ درصد می باشد. بنابراین می توان گفت بین میانگین بازده نقدی سهام بر اساس مدل های پیش بینی مختلف تفاوت معنی داری وجود دارد و نوع مدل پیش بینی تاثیر معنی داری بر میانگین بازده پیش بینی دارد.

جدول ١٤- تغییرات دو منبع فاکتور و خطا

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	3.153E-5	1	3.153E-5	23.245	.001
	Greenhouse-Geisser	3.153E-5	1.000	3.153E-5	23.245	.001
	Huynh-Feldt	3.153E-5	1.000	3.153E-5	23.245	.001
	Lower-bound	3.153E-5	1.000	3.153E-5	23.245	.001
Error(factor1)	Sphericity Assumed	1.221E-5	9	1.357E-6		
	Greenhouse-Geisser	1.221E-5	9.000	1.357E-6		
	Huynh-Feldt	1.221E-5	9.000	1.357E-6		
	Lower-bound	1.221E-5	9.000	1.357E-6		

نتایج پژوهش:

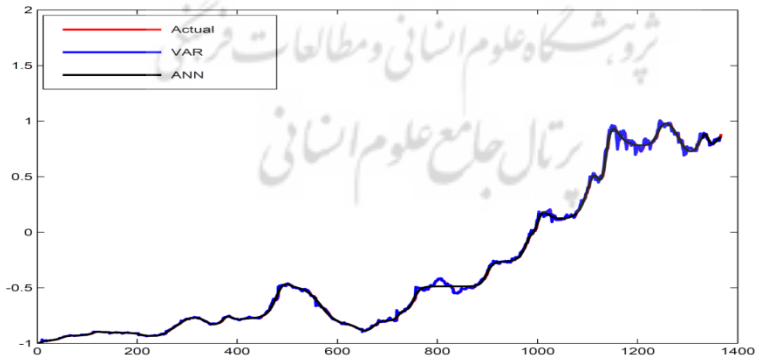
۱. هر دو مدل شبکه عصبی و مدل VAR در پژوهش اخیر موفقیت نسبی در پیش بینی شاخص دارند.

۲. از بین شبکه های عصبی طراحی شده با داده های وقفه های مختلف از شاخص شبکه عصبی سه لایه ۱-۱۰-۱۰-۵ با ورودی پنج وقفه از شاخص و ۲۰۰ تکرار آموزش با $R^2 = 0.999$ بهترین شبکه مدل شده برای پیش بینی شاخص بورس اوراق بهادار است.

۳. شبکه های عصبی طراحی شده برای پیش بینی شاخص به همراه وقفه هایی از خود شاخص عملکرد بهتری از شبکه های عصبی داشته اند که تنها ورودی آنها خود شاخص به تنهایی بوده اند و این امر نشان می دهد که شاخص بیشترین تاثیر را از مقادیر تاریخی خود می پذیرد.

۴. شبکه های عصبی، در این پژوهش تخمینی بهتر از مدل VAR داشته اند، در نمودار (۲) پیش بینی این دو مدل و تطبیق آنها با واقعیت نشان داده شده است و همچنین معیارهای مقایسه در جدول (۱۳) آمده است. به این معنا که هر دو مدل همبستگی X_t ، X_{t-1} و X_{t-2} را شناسایی کرده اند ولی به دلیل رابطه غیر خطی حاکم بر حرکات شاخص شبکه عصبی که یک مدل غیر خطی است پیش بینی صحیح تری نسبت به مدل VAR ارایه داده است.

نمودار ۲. مقایسه مقادیر دو مدل پیش بینی با واقعیت در داده های دوره برآورد



جدول ۱۵- مقایسه معیارهای مختلف برای دو مدل پیش بینی

مدل	MSE	RMSE	R^2
مدل شبکه عصبی	0.000328	0.01811	0.9996

۵. موفقیت نسبی مدل VAR و شبکه های عصبی با استفاده از دو وقفه برای پیش بینی سری زمانی TEPIX، و در این بین برتری مدل شبکه های عصبی نشان دهنده دو مطلب است: اول، وجود روند آشوبناک غیر تصادفی به دلیل همبستگی بین قیمت های متوالی در سری زمانی شاخص و دوم، وجود حرکات غیر خطی در سری زمانی شاخص، که اولی نشان می دهد تغییرات قیمت در بورس تهران تصادفی نیست و تغییرات قیمت های این هفته، هفته قبل و دو هفته قبل تاثیر بسزایی بر تغییرات قیمت هفته آینده می گذارند و دومی با فرضیه مبنی بر حرکات غیر خطی شاخص همخوانی دارد.

۶. نتایج نشان میدهند که جو روانی، حاکم بر تغییرات قیمت در بازار بورس اوراق بهادار تهران است و هنوز در بازار بورس تهران قیمت ها به صورت اصولی تعیین نمی شوند. بلکه بر اساس نظریه چارتیست ها قیمت دیروز، جهت تغییرات قیمت امروز را مشخص می کند.

منابع و ماخذ

۱. احمدی قراچه، احسان، (۱۳۸۵)، "ارائه یک مدل شبکه عصبی برای پیش بینی قیمت ماهانه نفت خام با در نظر گرفتن شوک های نفتی"، پایان نامه کارشناسی ارشد، دانشگاه تربیت مدرس.

۲. اصفهانیان، مجید، (۱۳۸۲)، "ارائه یک مدل شبکه عصبی جهت پیش بینی قیمت نفت خام"، پایان نامه کارشناسی ارشد، دانشگاه تربیت مدرس، دانشکده فنی و مهندسی، بخش مهندسی صنایع.

۳. بت شکن، محمود، (۱۳۷۹)، "پیش بینی قیمت سهام با استفاده از شبکه های عصبی – فازی و مقایسه آن با الگوهای خطی پیش بینی"، پایان نامه کارشناسی ارشد، دانشگاه تهران، دانشکده مدیریت.

۴. برزنده، محمد، (۱۳۷۶)، "بررسی متغیرهای اقتصادی اثرگذار بر شاخص قیمت سهام بورس اوراق بهادار"، **مجله برنامه و بودجه**، شماره ۴۰ و ۴۱.

۵. خالوزاده، حمید، (۱۳۷۷)، "مدلسازی غیر خطی و پیش بینی رفتار قیمت سهام در بازار بورس تهران"، پایان نامه دکترا، دانشکده تربیت مدرس.

۶. رشیدی، پ.، (۱۳۸۷)، "پیش بینی قیمت سهام بازار بورس اوراق بهادار بااستفاده از شبکه های عصبی"، پایان نامه کارشناسی ارشد، دانشگاه علم و صنعت ایران.

۷. عرب مازار یزدی، محمد و قاسمی، مهسا، (۱۳۸۸)، "برآورد قیمت عرضه های عمومی اولیه با استفاده از شبکه های عصبی مصنوعی"، **مجله تحقیقات حسابداری و حسابرسی**، انجمن حسابداری ایران، شماره ۱

۸. فدایی نژاد، محمد اسماعیل و حسن نژاد، محمد، (۱۳۹۱)، "پیش بینی بازده بورس اوراق بهادار با استفاده از مدل های یانگین متحرک و میانگین متحرک

با ورودی های خارجی"، **مجله تحقیقات حسابداری و حسابرسی**، انجمن حسابداری ایران، شماره ۱۶.

۹. نیکوکار، وهاب، (۱۳۸۶)، "طراحی محیط یادگیری برای پیش بینی بازار سهام با استفاده از شبکه عصبی"، پایان نامه کارشناسی ارشد، دانشگاه تربیت مدرس.

10. Chi-Ming Lin, Jih-Jeng Huang, Mitsuo Gen, Gwo-Hshiung Tzeng, (2006), "Recurrent neural network for dynamic portfolio selection", **Applied Mathematics and Computation**, Vol. 175, Issue 2, PP. 1139-1146.

11. Fabio D. Freitas, Alberto F. De Souza, Ailson R. de Almeida, (2009), "Prediction - based portfolio optimization model using neural networks", **Neuro computing**, Vol. 72, Issues 10-12, PP. 2155-2170.

12. G. Peter Zhang, (2004), **Neural Networks in Business Forecasting**, IDEA GROUP PUBLISHING, USA.

13. H. Wang, A. S. Weigend, (2004), "Data mining for financial decision making", **Decision Support Systems**, Vol. 37, No. 4, PP. 457-460.

14. Po-ChangKo, Ping - ChenLin, (2008), "Resource allocation neural network in portfolio selection", **Expert Systems with Applications**, Vol. 35, Issues 1-2, PP. 330-337.

15. Saad, W. E., Prokhorov, D. V. & Wunsch, C. D., (1998), "Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks", **IEEE Transactions on Neural Networks**, Vol. 9, No. 6, PP. 1456-1470.
16. White H., (1993), **Economic prediction using neural networks: the case of IBM daily stock returns**, In Trippi R. R. & Truban E. Neural network in finance and investing. Chicago, Ihhonois, and Cambridge, England: probes Publication Company, PP. 315-329.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

Comparison of neural networks and Auto regression performance in prediction of Dividend and Price Index

Abstract:

This paper present two models for short and long term stock price forecasting using Artificial Neural Networks and Vector Auto regression analyses as complementary methods. Presented models also will be able to forecast the stock price of future day. The models forecast the stock price of future day with using of previous stock price information. In presented models, it is possible to use the calendar parameters, like the day of week, special days and etc to forecast the stock price. A major contribution of this work is the resulting time-delayed artificial neural network model that allows stock return predictions and is particularly useful as an investment decision support system for hedge funds and other investors, whose portfolios are at risk of losing market value.

After designing mentioned models, and specifying construction of multi-layer perception neural network, designed network is trained using available data. Presented method for training the perception neural network is, Error back propagation algorithm, which has variable learning rate and momentum factor, and also is fast. Finally, the models have been applied on Tehran-stock Exchange Price Index (TEPIX) and Tehran- stock Exchange-Dividend and Price Index (TEDPIX) and result are compared together and advantages and disadvantages of these models are described. The results show validity of presented models.

Key words: neural networks, Dividend and Price Index, Auto regress Models