

## کاربرد روش شبکه عصبی مصنوعی و مدل‌های واریانس ناهمسانی شرطی

### در محاسبه ارزش در معرض خطر

رضا نریمانی<sup>۱</sup>

نادر حکیمی پور<sup>۲</sup>

اسعد اله‌رضایی<sup>۳</sup>

دریافت: ۱۳۹۲/۰۴/۰۶ پذیرش: ۱۳۹۲/۰۶/۰۸

#### چکیده

ریسک بازار از عدم اطمینان در خصوص بازدهی آتی دارایی‌ها در بازار نشأت می‌گیرد. امروزه معیارهای مختلفی برای بررسی انواع ریسک مرتبط با بازار، سبدهای مختلف دارایی، صنایع و ... به کار می‌روند. اما هر چند این معیارهای مختلف، اطلاعات ارزشمندی را برای فعالان بازار به همراه می‌آورند، لیکن هر یک به تنهایی نمی‌توانند اطلاعات جامع و کاملی را در خصوص ریسک بازار و یا سبد سهام به دست دهند. به همین منظور، «ارزش در معرض خطر» که به اختصار VaR نامیده می‌شود، تلاشی است برای اینکه عدد معینی در خصوص ریسک بازار و یا سبد سهام به تحلیل‌گران و فعالان بازار ارائه کند.

در این تحقیق از کلیه مدل‌های مبتنی بر واریانس ناهمسانی شرطی شامل GARCH، EGARCH، CGARCH، TARCH و GARCH-M و روش شبکه عصبی مصنوعی برای پیش‌بینی ارزش در معرض خطر پرتفویی متشکل از ۵۰ شرکت با نقدشوندگی بالا استفاده شده است. سپس نتایج به دست آمده با استفاده از آزمون‌های پوشش غیرشرطی کوپیک مورد بررسی قرار گرفت. در نهایت مدل شبکه عصبی در مقایسه با سایر روش‌های مبتنی بر واریانس ناهمسانی شرطی، عملکرد بهتری را بر اساس آزمون کوپیک داشته است.

JEL: G14, C59, G11, G17

۱. کارشناس ارشد مهندسی مالی، دانشگاه علوم اقتصادی، تهران، ایران. (نویسنده مسئول)

۲. عضو هیأت علمی پژوهشکده آمار ایران، تهران، ایران.

۳. کارشناسی ارشد اقتصاد، مرکز آمار ایران، تهران، ایران. Email: alahrezaee@yahoo.com

## ۱- مقدمه

به طور کلی، موسسات و بنگاه‌های مالی و اقتصادی با چهار نوع ریسک اعتباری، عملیاتی، نقدینگی و بازار مواجه‌اند. ریسک اعتباری، ریسک یا زیان بالقوه‌ای است که از ناتوانی طرف‌های تجاری بنگاه در ایفای تعهداتشان ناشی می‌شود. ریسک عملیاتی به ریسکی اطلاق می‌شود که از طریق امکان بروز خطا یا تقلب در تسویه قراردادها و مبادله اسناد می‌تواند متوجه بنگاه شود. ریسک نقدینگی زمانی مطرح می‌شود که بنگاه از منابع مالی کافی جهت رفع نیازهای فوری خود برخوردار نباشد. ریسک بازار نیز از عدم اطمینان در خصوص بازدهی آتی دارائی‌ها در بازار نشأت می‌گیرد. البته این تقسیم‌بندی به مفهوم جدایی این ریسک‌ها از یکدیگر نیست، بلکه کلیه این ریسک‌ها با یکدیگر در ارتباط بوده و مدیریت ریسک در بنگاه‌ها و موسسات بایستی به نحوی صورت گیرد که همه آنها را پوشش دهد. در این تحقیق به طور خاص، به دنبال بررسی و تعیین ریسک بازار مرتبط با پرتفویی متشکل از ۵۰ شرکت با نقدشوندگی بالا در بازار اوراق بهادار تهران هستیم.

امروزه معیارهای مختلفی برای بررسی انواع ریسک مرتبط با سبدهای مختلف دارائی، صنایع و ... به کار می‌روند. اما هر چند این معیارهای مختلف، اطلاعات ارزشمندی را برای فعالان بازار به همراه می‌آورند، لیکن هر یک به تنهایی نمی‌توانند اطلاعات جامع و کاملی را در خصوص ریسک صنعت و یا سبد سهام به دست دهند. به همین منظور، «ارزش در معرض خطر»<sup>۱</sup> که به اختصار VaR نامیده می‌شود، تلاشی است برای اینکه عدد معینی در خصوص ریسک صنعت و یا سبد سهام به تحلیل‌گران و مدیران ارشد مؤسسات ارائه کند. سال‌هاست که محاسبه ارزش در معرض خطر در بخش‌های مختلف اقتصادی اعم از بازار اوراق بهادار، بانک‌ها، شرکت‌های سرمایه‌گذاری و ... در سطح جهانی با استفاده از مدل‌های مبتنی بر واریانس ناهمسانی شرطی، شبیه‌سازی تاریخی و شبیه‌سازی مونت کارلو صورت می‌گیرد. از سوی دیگر، با پیشرفت روزافزون استفاده از روش‌های مبتنی بر هوش مصنوعی، استفاده از روش‌هایی از قبیل شبکه عصبی مصنوعی رواج یافته است که نتایج نسبتاً بهتری را نیز ارائه می‌کنند.

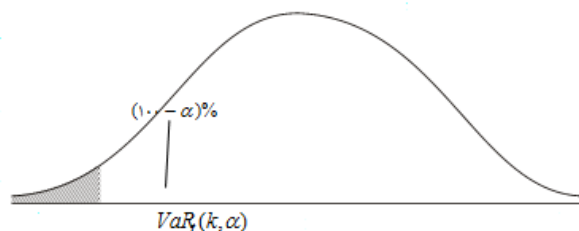
در ایران نیز تحقیقات گسترده و متنوعی در این خصوص با استفاده از مدل های مبتنی بر واریانس ناهمسانی شرطی، شبیه سازی تاریخی و شبیه سازی مونت کارلو صورت گرفته است. اما علی رغم استفاده گسترده از روش شبکه عصبی مصنوعی در پیش بینی ارزش در معرض خطر در خارج از کشور، تاکنون هیچ تحقیقی در این زمینه در کشور صورت نگرفته است. هدف از این تحقیق بررسی عملکرد مدل های اقتصادسنجی مبتنی بر واریانس ناهمسانی شرطی و روش شبکه عصبی مصنوعی در محاسبه مقادیر ارزش در معرض خطر پرتفوی ۵۰ شرکت برتر بورس اوراق بهادار است. بر این اساس فرضیه تحقیق نیز بر این اساس قرار دارد که روش شبکه عصبی مصنوعی به دلیل ماهیت کاملاً غیرخطی خود نتایج بهتری نسبت به مدل های نیمه غیرخطی مبتنی بر واریانس ناهمسانی شرطی در پیش بینی ارزش در معرض خطر دارد.

## ۲- ارزش در معرض خطر

در حال حاضر یکی از متداول ترین معیارهای سنجش ریسک بازار، محاسبه ارزش در معرض خطر یا VaR متناسب با سبد سهام یا پرتفوی مورد نظر است. به طور خلاصه، VaR یا مبلغ تحت ریسک، بیانگر حداکثر زیان مورد انتظار در خصوص سبد سهام یا مجموعه سرمایه گذاری ها در افق زمانی معین (مانند یک روز، یک هفته و یا یک ماه) در سطح اطمینان معین است. طبق این تعریف، معیار VaR دربرگیرنده دو پارامتر افق زمانی (k) و سطح اطمینان ( $\alpha$ ) است.

به طور ساده تر، با فرض اینکه افق زمانی k روزه و سطح اطمینان  $\alpha$  درصد باشد، مبلغ تحت ریسک یا VaR، مقدار زیان متناظر با  $1-\alpha$  درصد منحنی توزیع احتمال تغییرات ارزش سبد دارائی ها در طی k روز آتی است. برای مثال، هنگامی که  $k=10$  و  $\alpha=99$  درصد باشد، ارزش در معرض خطر، معادل ۱٪ سطح زیر منحنی توزیع احتمال تغییرات ارزش سبد سهام در طی ۱۰ روز آتی است. نمودار زیر، VaR را برای موقعیتی که تغییرات در ارزش سبد دارائی ها تقریباً به صورت نرمال توزیع شده باشند را نشان می دهد.

شکل ۱. محاسبه VaR با استفاده از توزیع احتمالات تغییرات در ارزش سبد دارایی‌ها در سطح اطمینان  $\alpha$  درصد



بنابراین ارزش در معرض خطر با استفاده از رابطه زیر تعریف می‌شود:

$$\int_{-\infty}^{VaR_t(k, \alpha)} f_{t+k}(x) dx = \frac{\alpha}{100}$$

همچنین از آنجا که بازده یک سبد دارایی به صورت  $r_t = \log(p_t / p_{t-1})$  محاسبه می‌شود،

ارزش در معرض خطر را به صورت زیر نیز می‌توان تعریف کرد:

$$p(r_t \leq VaR_t(\alpha, k)) = 1 - \alpha$$

که در آن  $p_t$  شاخص قیمت سبد سهام در زمان  $t$  است.

#### روش‌های محاسبه ارزش در معرض خطر

در سال‌های اخیر، روش‌های مختلفی برای محاسبه ارزش در معرض خطر ارائه شده است. این روش‌ها به دو نوع پارامتریک و ناپارامتریک تقسیم می‌شوند. روش پارامتریک که شامل مدل‌های مبتنی بر واریانس ناهمسانی شرطی و روش شبکه عصبی مصنوعی است، با استفاده از داده‌های تاریخی به برآورد پارامترهای تابع چگالی می‌پردازد. اما در روش ناپارامتریک، که به طور عمده شامل روش‌های شبیه‌سازی تاریخی و شبیه‌سازی مونت کارلو است، به جای برآورد پارامترهای تابع چگالی، مستقیماً از خود داده‌های تاریخی برای محاسبه ارزش در معرض خطر استفاده می‌شود.

در این تحقیق از مدل‌های مبتنی بر واریانس ناهمسانی شرطی و شبکه عصبی مصنوعی استفاده خواهد شد. به همین جهت، در ادامه پس از بیان پیشینه تحقیقات صورت گرفته در این زمینه در داخل و خارج از کشور، به بیان مدل‌سازی نوسان‌پذیری و یا واریانس

ناهمسانی شرطی در داده های سری زمانی مالی و روش های مختلف محاسبه آن و همچنین نحوه عمل روش شبکه عصبی مصنوعی خواهیم پرداخت.

### ۳- مروری بر مطالعات تجربی

مدل های مبتنی بر واریانس ناهمسانی شرطی یا ARCH، در ابتدا توسط انگل در سال ۱۹۸۲ معرفی شد. بولرسلف (۱۹۸۶) مدل انگل را تعمیم داد و گروهی از مدل ها را که به مدل های تعمیم یافته خودرگرسیو واریانس ناهمسان (GARCH) شهرت یافتند، ارائه کرد. از این پس، این مدل ها با تأکید بر ویژگی های مختلف داده های مالی گسترش یافتند، که از آن جمله می توان به مدل های EGARCH، IGARCH و ... اشاره کرد که در بخش روش شناسی تحقیق به طور کامل مورد بررسی قرار خواهند گرفت. اما در ادامه به بررسی تعدادی از تحقیقاتی که اخیراً در این زمینه صورت گرفته است می پردازیم.

چان و همکاران (۲۰۰۷) در مقاله ای به محاسبه ارزش در معرض خطر در بازار انرژی با استفاده از مدل های ARCH و GARCH می پردازند. نکته اصلی تحقیق آنها، تأکید بر ساخت فاصله اطمینان برای برآورد VAR شرطی است که می تواند تاحدی سبب افزایش قابلیت اعتماد این روش در برآورد ریسک گردد.

ماریموتو و دیگران (۲۰۰۹) در مقاله ای به بررسی ریسک مرتبط با بازار انرژی می پردازند. آنها اظهار می کنند که افزایش های اخیر در قیمت های انرژی، به خصوص قیمت نفت، تبدیل به یک نگرانی عمده برای مصرف کنندگان، شرکت ها و دولت ها گردیده است. بر این اساس، از آنجا که بسیاری از تحلیلگران بر این باورند که نوسانات قیمت نفت پیامدهای قابل توجهی در فعالیتهای اقتصادی داشته و این در حالی است که بازارهای نفت به طور طبیعی نسبت به تغییرات قابل توجه قیمت نفت، آسیب پذیرند، نویسندگان این مقاله معتقدند پیش بینی و در نهایت محافظت در برابر ریسک بازار این محصول به یک ضرورت تبدیل شده است. از این رو آنها به محاسبه ارزش در معرض خطر با استفاده از روش های GARCH، شبیه سازی تاریخی و شبیه سازی تاریخی فیلتر شده پرداخته و نتایج آن با یکدیگر مقایسه شده است. نتایج به دست آمده اهمیت فرآیند فیلتر کردن داده ها برای موفقیت روش های استاندارد را مورد تأیید قرار می دهد.

با وجود گسترش چشم‌گیر مدل‌های اقتصادسنجی و استفاده روزافزون از آنها در بازارهای مالی دنیا، در ایران این مدل‌ها در بررسی ریسک بازار مورد توجه قرار نگرفته‌اند. با این وجود، در ادامه به بیان برخی از مطالعات صورت گرفته در این زمینه در ایران می‌پردازیم. شاهرادی و زنگنه (۱۳۸۶) در مقاله‌ای با استفاده از چهار نوع مدل GARCH، ارزش در معرض خطر را برای پنج شاخص عمده بورس اوراق بهادار تهران (شاخص کل، شاخص وزنی، شاخص صنعت، شاخص قیمت و بازده نقدی و شاخص پنجاه شرکت فعال‌تر) که واریانس ناهمسانی شرطی در آنها مشاهده می‌شود، برآورد کردند. نتایج حاکی از آن است که این گروه از مدل‌ها، رفتار میانگین و واریانس داده‌ها را نسبت به سایر مدل‌های کلاسیک محاسبه VaR به نحو مطلوب‌تری توضیح می‌دهند. همچنین نتایج نشان می‌دهد که شاخص‌های قیمت و بازده نقدی، صنعت و ۵۰ شرکت فعال‌تر نسبت به شاخص‌های دیگر، ارزش در معرض خطر کمتری دارند.

اسلامی بیدگلی، راعی و کمال‌زاده (۱۳۹۲) در تحقیقی عملکرد روش پارامتریک در پیش‌بینی مقادیر ارزش در معرض خطر در قیمت سبد نفتی اوپک مورد بررسی قرار دادند. محققین پس از محاسبه مقادیر ارزش در معرض خطر یک و ده روزه با استفاده از برخی مدل‌های خانواده حافظه بلندمدت گارچ مانند FIGARCH، HYGARCH و FIEGARCH بر روی سه توزیع آماری نرمال،  $t$  استیودنت و  $t$  استیودنت چوله، نتایج به دست آمده را با استفاده از آزمون‌های کوپیک و صدک پویا، در سطوح اطمینان پایین و بالا مورد مقایسه و تحلیل قرار دادند. نتایج به دست آمده نشان می‌دهند که دنباله پهن و خاصیت حافظه بلندمدت در نوسانات بازده قیمت در سبد نفتی اوپک وجود دارد، پیش‌بینی مقادیر ارزش در معرض خطر یک روزه و ده روزه با استفاده از توزیع چوله از دقت و عملکرد بالاتری برخوردار است. در نهایت، مدل FIEGARCH در پیش‌بینی ارزش در معرض خطر در هر دو بازه ۱ و ۱۰ روزه بهتر از سایر مدل‌ها عمل می‌کند.

سجاد و هدایتی (۱۳۹۲) در مقاله خود مقدار ارزش در معرض خطر را با استفاده از هفت روش مختلف از جمله نظریه ارزش فرین و برای سه سطح اطمینان، برای بازده لگاریتمی شاخص کل بورس تهران، نرخ برابری دلار و یورو به صورت روزانه محاسبه می‌کنند. همچنین به منظور پیش‌بینی نوسانات بازده از مدل GARCH استفاده شده است. برای

بررسی کفایت دقت مدل های به کار گرفته شده، آزمون های نسبت شکست های کوپیک، کریستوفرسن و تابع زیان لویز را به کار گرفته شده اند. نتایج حاصل نشان می دهد که محاسبه ارزش در معرض خطر با استفاده از روش های سنتی لزوماً به نتایج مناسبی نمی انجامد و در برخی از موارد استفاده از نظریه ارزش فرین و در نظر گرفتن نوسانات شرطی برای داده ها موجب نتایج بهتری می شود این نتایج بیشتر در سطح اطمینان های بالاتر قابل مشهودتر است.

## ۴- روش شناسی تحقیق

### ۴-۱- مدل های مبتنی بر واریانس ناهمسانی شرطی

در اقتصادسنجی متعارف، برای بررسی و تبیین عوامل اثرگذار بر یک متغیر مانند  $Y_t$  و همچنین مباحث مربوط به پیش بینی آن، از الگویی به شکل زیر استفاده می شود:

$$Y_t = a_1 + a_2 X_{2t} + a_3 X_{3t} + \dots + a_k X_{kt} + u_t \quad (1)$$

در نهایت آنچه که برآورد می شود، معادله میانگین  $Y_t$  است که به شرط غیر تصادفی بودن مقادیر متغیرهای توضیحی و  $E(u_t) = 0$ ، به صورت زیر تعریف می شود:

$$E(Y_t | X_t) = a_1 + a_2 X_{2t} + a_3 X_{3t} + \dots + a_k X_{kt} \quad (2)$$

معادله فوق، به معادله میانگین<sup>۱</sup> معروف است. به طور مشابه، در اقتصادسنجی سری های زمانی، از الگوی ARMA (و یا ARIMA) به شکل زیر برای پیش بینی یک سری زمانی مانند  $Y_t$  استفاده می شود که به تعبیری همان معادله میانگین در سری های زمانی است:

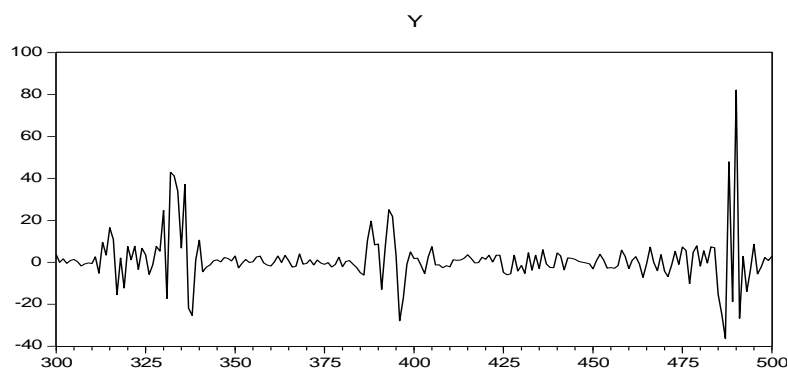
$$Y_t = a_0 + \sum_{l=1}^m a_l Y_{t-l} + u_t - \sum_{h=1}^n b_h u_{t-h} \quad (3)$$

مدل فوق به ARMA(m,n) معروف است که چنانچه  $Y_t$  نامانا باشد، با تفاضل گیری به مدل ARIMA(m,n) می رسیم. به طور معمول مدل فوق، با برقراری فروض استاندارد کلاسیک از قبیل نرمال بودن جزء خطا، عدم وجود خودهمبستگی و واریانس ناهمسانی در جزء خطا و ... قابل برآورد خواهد بود.

اما در طی دو دهه اخیر، مطالعات گسترده‌ای در خصوص مدل‌سازی و پیش‌بینی سری‌های زمانی مانند بازار سهام، نرخ ارز، تورم و ... صورت گرفته است. در این مطالعات، نوع خاصی از نوسان‌پذیری موسوم به نوسان‌پذیری خوشه‌ای<sup>۱</sup> مورد توجه قرار گرفته است. در این حالت جزء خطا معادله رگرسیون به صورت شرطی ثابت نبوده و نسبت به اطلاعات تغییر می‌کند. روش‌های مختلفی برای مدل‌سازی چنین پدیده‌ای مطرح شده است که در ادامه به آن می‌پردازیم.

#### ۴-۱-۱- مدل واریانس ناهمسانی شرطی خودرگرسیون<sup>۲</sup>

در سال ۱۹۸۲، اقتصادسنجی به نام رابرت انگل<sup>۳</sup> در مطالعه روند داده‌های تورم، دریافت که در تورم‌های بالا، نوسانات بیشتر و در تورم‌های پائین، نوسانات کمتر است. به عبارت دیگر، تغییرات بزرگ در متغیر مورد نظر، منجر به تغییرات بزرگ و تغییرات کوچک، منجر به تغییرات کوچک می‌شود.



شکل ۲. نوسان‌پذیری خوشه‌ای در داده‌های سری زمانی

- 
1. Volatility Clustering
  2. Autoregressive Conditional Heteroskedasticity (ARCH)
  3. Robert Engel



در واقع انگل با تمایز گذاشتن بین واریانس شرطی و غیرشرطی<sup>۱</sup>، فصل جدیدی را در کاراتر ساختن تخمین ضرایب و بهبود نتایج پیش‌بینی کرد. حال برای مدل‌سازی نوسان‌پذیری خوشه‌ای و یا واریانس شرطی اجزاء خطا از رابطه زیر استفاده می‌شود:

$$\sigma_t^2 = \text{var}(u_t | u_{t-1}, u_{t-2}, \dots) = E[(u_t - E(u_t))^2 | u_{t-1}, u_{t-2}, \dots] \quad (4)$$

که با فرض  $E(u_t) = 0$  داریم:

$$\sigma_t^2 = E(u_t^2 | u_{t-1}, u_{t-2}, \dots) \quad (5)$$

بر اساس رابطه فوق، واریانس شرطی  $u_t$  برابر است با امید ریاضی شرطی مجذور  $u_t$ . لذا محاسبه واریانس جزء خطا در زمان  $t$ ، به شرط معلوم بودن مقادیر خطا در دوره‌های گذشته امکان‌پذیر خواهد بود.

رابطه زیر، مدل ARCH را که در ساده‌ترین حالت، تنها به مجذور جمله خطا در دوره قبل بستگی دارد نشان می‌دهد که تحت عنوان ARCH(1) شناخته می‌شود<sup>۲</sup>:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 \quad (6)$$

معادله فوق به معادله واریانس<sup>۳</sup> معروف است که مطابق معمول می‌توان آن را تا وقفه دلخواه گسترش داد. لذا در نهایت برای ارائه برآوردهای بهتر از ضرایب و همچنین پیش‌بینی‌های دقیق‌تر، بایستی دو معادله میانگین و واریانس به طور همزمان برآورد گردند. لذا چنانچه رابطه ۴ را تا وقفه  $q$  ام گسترش دهیم، مدل ARCH(q) به صورت زیر خواهد بود:

۱. بحث‌هایی که در اقتصاد سنجی کلاسیک در خصوص واریانس ناهمسانی مطرح می‌شود، مربوط به واریانس غیرشرطی است و نه واریانس شرطی.

۲. در برخی از کتب و مقالات به جای استفاده از معادله واریانس در رابطه ۶، از دو رابطه زیر استفاده می‌کنند:

$$u_t = v_t \sigma_t \quad ; \quad v_t \sim N(0,1)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2$$

که به راحتی می‌توان نشان داد که دو رابطه فوق با معادله واریانس در رابطه ۶ برابر است:

$$E(u_t) = E(v_t \sigma_t) = E(v_t) E(\sigma_t) = 0$$

$$\text{var}(u_t) = \text{var}(v_t \sigma_t) = \sigma_t^2 \text{var}(v_t) = \sigma_t^2$$

### 3. Variance Equation

$$Y_t = a_0 + \sum_{l=1}^m a_l Y_{t-l} + u_t - \sum_{h=1}^n b_h u_{t-h} ; u_t \sim N(0, \sigma_t^2) \quad (7)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 ; \alpha_0 > 0, \alpha_i \geq 0$$

#### ۴-۱-۲- مدل واریانس ناهمسانی شرطی خودرگرسیونی تعمیم یافته<sup>۱</sup>

مدل ARCH دارای دو نقطه ضعف اساسی است؛ اول اینکه تعیین مقدار  $q$  (وقفه‌های معادله واریانس) بسیار دشوار بوده و دوم اینکه برای جلوگیری از منفی شدن مقادیر برآورد شده واریانس، نیاز است که ساختار خاص و از پیش تعیین شده‌ای بر مدل اعمال گردد. برای رفع این مشکلات، اقتصادسنجی به نام بولرسلف<sup>۲</sup> در سال ۱۹۸۶ مدل ARCH تعمیم یافته یا GARCH را ارائه کرد. این مدل دارای معادله میانگین به شکل مرسوم ولی معادله واریانس متفاوت از مدل ARCH است که شکل ساده آن به صورت زیر است:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (8)$$

رابطه فوق با GARCH(1,1) نشان داده می‌شود. به راحتی می‌توان نشان داد که معادله فوق معادل است با ARCH( $\infty$ )؛ برای این منظور با نوشتن رابطه ۸ با یک وقفه جایگذاری آن در رابطه فوق، داریم:

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \beta(\alpha_0 + \alpha_1 u_{t-2}^2 + \beta \sigma_{t-2}^2) \\ &= \alpha_0(1 + \beta) + \alpha_1 u_{t-1}^2 + \beta \alpha_1 u_{t-2}^2 + \beta^2 \sigma_{t-2}^2 \end{aligned} \quad (9)$$

به طور مشابه، با جایگذاری‌های مکرر خواهیم داشت:

$$\begin{aligned} \sigma_t^2 &= \alpha_0(1 + \beta + \beta^2 + \dots) + \alpha_1(u_{t-1}^2 + \beta u_{t-2}^2 + \beta^2 u_{t-3}^2 + \dots) \\ &= \alpha'_0 + \alpha'_1 u_{t-1}^2 + \alpha'_2 u_{t-2}^2 + \alpha'_3 u_{t-3}^2 + \dots ; \alpha'_0 = \alpha_0 \sum_{i=0}^{\infty} \beta^i, \alpha'_i = \alpha_i \beta^i \end{aligned} \quad (10)$$

بنابراین GARCH(1,1) معادل ARCH( $\infty$ ) خواهد بود. اما در حالت کلی، مدل GARCH( $q, p$ ) به صورت زیر تعریف می‌شود:

- 
1. Generalized Autoregressive Conditional Heteroskedasticity (GARCH)
  2. Bollerslev

$$Y_t = a_0 + \sum_{l=1}^m a_l Y_{t-l} + u_t - \sum_{h=1}^n b_h u_{t-h} ; u_t \sim N(0, \sigma_t^2) \quad (11)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 ; \alpha_0 > 0, \alpha_i, \beta_j \geq 0$$

اما با این وجود، نتایج بیشتر مطالعات تجربی، بر کفایت GARCH(1,1) تأکید دارند.

#### ۴-۱-۳- مدل واریانس ناهمسانی شرطی با هم‌انباشتگی<sup>۱</sup>

فرض ضمنی مدل GARCH، کوچک‌تر از یک بودن مجموع ضرایب  $\alpha_1$  و  $\beta_j$  است. به عبارت دیگر فرض بر این است که به علت نوسانات خوشه‌ای، واریانس شرطی  $u_t$  در طی زمان در حال تغییر بوده ولی واریانس غیرشرطی آن بایستی ثابت باشد. لذا برای مثال، چنانچه برای محاسبه واریانس غیرشرطی یک مدل GARCH(1,1)، با گرفتن میانگین معادله میانگین آن، خواهیم داشت:

$$E(\sigma_t^2) = \alpha_0 + \alpha_1 E(u_{t-1}^2) + \beta E(\sigma_{t-1}^2)$$

که با فرض واریانس همسانی غیرشرطی، داریم:

$$E(\sigma_t^2) = E(u_{t-1}^2) = E(\sigma_{t-1}^2) = \sigma^2$$

، لذا واریانس  $u_t$  به صورت زیر قابل محاسبه خواهد بود:

$$\text{var}(u_t) = \frac{\alpha_0}{1 - (\alpha_1 + \beta)} \quad (12)$$

به راحتی می‌توان دریافت که عبارت فوق در صورتی قابل تعریف است که  $\alpha_1 + \beta < 1$  باشد؛ اما اگر  $\alpha_1 + \beta = 1$  باشد، در اصطلاح گفته می‌شود که ریشه واحد وجود دارد و آن را با IGARCH نشان می‌دهند.

نتایج بسیاری از مطالعات تجربی نیز حاکی از وجود چنین پدیده‌ای (ریشه واحد) در مدل‌های GARCH می‌باشند. لذا برای کاراتر شدن مدل، قید زیر بر ضرایب مدل GARCH(q,p) تحمیل می‌شود:

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j = 1 \quad (13)$$

#### 1. Integrated GARCH (IGARCH)

برای مثال، معادله واریانس مدل GARCH(1,1) که در بسیاری از مطالعات نیز مورد استفاده قرار گرفته است، به صورت زیر خواهد بود:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + (1 - \alpha_1) \sigma_{t-1}^2 \quad (14)$$

#### ۴-۱-۴- مدل واریانس ناهمسانی شرطی آستانه‌ای<sup>۱</sup>

در مدل‌هایی که تاکنون بررسی شد، اثر شوک‌های مثبت و منفی بر معادله واریانس، یکسان و متقارن در نظر گرفته شده است. اما به طور معمول انتظار داریم اثر شوک‌های مثبت و منفی که به بازدهی سهام وارد می‌شود و یا اثر کاهش و افزایش قیمت نفت بر اقتصاد متفاوت از یکدیگر عمل کنند. برای این منظور، یکی از مدل‌هایی که برای در نظر گرفتن چنین عدم تقارنی در اخبار و اطلاعات وارده بر بازارهای مالی و سطح کلان اقتصاد در نظر گرفته شده است، مدل ARCH آستانه‌ای یا TARARCH است. در واقع مدل TARARCH به دنبال تبیین اثرات وقایعی است که در گذشته روی داده است ولی اثر آنها در زمان حال ظاهر می‌شود. این مدل نیز دارای معادله میانگین به شکل قبل، ولی معادله واریانس به شکل رابطه زیر است:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{i=1}^r \gamma_i u_{t-i}^2 I_{t-i} \quad (15)$$

$$\begin{cases} I_{t-i} = 1 & ; \quad u_{t-i} < 0 \\ I_{t-i} = 0 & ; \quad u_{t-i} \geq 0 \end{cases}$$

در این مدل  $u_{t-i} \geq 0$  بیانگر اخبار خوب و  $u_{t-i} < 0$  بیانگر اخبار بد می‌باشند. همچنین در این مدل، مقدار  $r$  بیانگر تعداد وقفه‌هایی است که طی آن، اخبار گذشته (خوب یا بد) بر واریانس اثر می‌گذارند. در این مدل تأثیر اخبار خوب به میزان  $\alpha_i$  و تأثیر اخبار بد به میزان  $\alpha_i + \gamma_i$  است. از سوی دیگر، به راحتی می‌توان دریافت که چنانچه ضرایب  $\gamma_i$  به ازای  $i=1, 2, \dots, r$  معنی‌دار نباشند، مدل متقارن خواهد بود. همچنین چنانچه  $\gamma_i > 0$  باشد، اثر شوک‌های منفی بیشتر از اثر شوک‌های مثبت در دوره  $i$  ام بوده است و بالعکس.

## 2. Threshold ARCH (TARARCH)

۴-۱-۵- مدل واریانس ناهمسانی شرطی نمایی<sup>۱</sup>

مدل EGARCH نیز به شکل دیگری، به دنبال تبیین اثرات نامتقارن شوک های مثبت و منفی بر واریانس شرطی است. این مدل در سال ۱۹۹۱ توسط نلسون<sup>۲</sup> مطرح شد که فرم عمومی آن به صورت زیر است:

$$\log(\sigma_t^2) = \alpha_0 + \sum_{j=1}^q \beta_j \log(\sigma_{t-j}^2) + \sum_{j=1}^p \alpha_j \left| \frac{u_{t-i}}{\sigma_{t-i}} \right| + \sum_{j=1}^r \gamma_j \left| \frac{u_{t-i}}{\sigma_{t-i}} \right| \quad (16)$$

مزیت اساسی این مدل نسبت به مدل TARARCH، لگاریتمی بودن متغیر وابسته یعنی  $\sigma_t^2$  است که در نتیجه ضرایب متغیرهای سمت راست می توانند مثبت یا منفی برآورد گردند<sup>۳</sup>.

۴-۱-۶- مدل واریانس ناهمسانی شرطی توانی

تیلور<sup>۴</sup> (۱۹۸۶) و شوارت<sup>۵</sup> (۱۹۸۹) مدل GARCH را بر اساس انحراف معیار معرفی کردند. در این مدل، به جای واریانس، انحراف معیار مدل سازی می گردد. این مدل در سال ۱۹۹۳ با عنوان ARCH توانی عمومیت یافت. در مدل PARARCH، پارامتر توانی از انحراف معیار قابل برآورد است و پارامترهای انتخابی  $\gamma$  نیز برای به دست آوردن عدم تقارن مرتبه بالاتر از  $r$  به مدل اضافه شده اند:

$$\sigma_t^\delta = w + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta \quad (17)$$

که در آن  $\delta > 0$  بوده و نیز به ازاء  $i=1, 2, \dots, r$ ،  $|\gamma_i| \leq 1$  است و برای تمامی  $i > r$  و  $r \leq p$  داریم:  $\gamma_i = 0$ . در یک مدل نامتقارن  $\gamma_i$  برای تمامی آنها صفر خواهد بود. لازم به ذکر است که اگر برای تمامی آنها،  $\delta = 2$  و  $\gamma_i = 0$  باشد، مدل PARARCH مشخصه GARCH استاندارد را دارد و نیز همانند مدل های قبلی، چنانچه  $\gamma_i \neq 0$  باشد، اثرات عدم تقارن برقرار خواهند بود.

1. Exponential GARCH (EGARCH)
2. Nelson

۳. در این حالت نیازی به اعمال قید غیرمنفی بودن ضرایب مدل نیست.

1. Taylor
2. Schwert

۴-۱-۷- مدل واریانس شرطی ترکیبی<sup>۱</sup>

واریانس شرطی در مدل GARCH(1,1) عبارت است از:

$$\sigma_t^2 = \bar{w} + \alpha(\varepsilon_{t-1}^2 - \bar{w}) + \beta(\sigma_{t-1}^2 - \bar{w}) \quad (18)$$

که بازگشت میانگین به  $\bar{w}$  را نشان می‌دهد که برای تمامی دوره‌ها ثابت است. در مقابل، مدل CGARCH بازگشت میانگین را به سطح متغیر جایز می‌شمارد که به صورت زیر مدل‌سازی می‌گردد:

$$\sigma_t^2 - m_t = \alpha w_t - 1 + \beta \sigma_{t-1}^2 - m_{t-1} \quad (19)$$

که در آن

$$m_t = w + \rho(m_{t-1} - w) + \phi(\varepsilon_{t-1}^2 - \sigma_{t-1}^2) \quad (20)$$

در اینجا  $\sigma_t^2$  نوسان است و  $q_t$  جایگزین  $w$  می‌شود و نوسان بلندمدت متغیر در طی زمان را نشان می‌دهد. اولین معادله، جزء ناپایدار و موقت  $\sigma_t^2 - q_t$  را توصیف می‌کند که با درجه (توان)  $\alpha + \beta$  هم‌گرا به صفر است. معادله دوم، جزء بلندمدت  $m_t$  را توصیف می‌کند که با توان  $p$  به سمت  $w$  هم‌گرا است.  $p$  نوعاً بین ۰/۹۹ و ۱ بوده و لذا  $m_t$  بسیار آهسته به  $w$  نزدیک می‌شود. می‌توان معادلات ناپایدار و دائمی را ترکیب کرده و چنین نوشت:

$$\sigma_t^2 = (1 - \alpha - \beta)(1 - \rho)w + (\alpha + \phi)\varepsilon_{t-1}^2 - (\alpha\rho + (\alpha + \beta)\phi)\varepsilon_{t-2}^2 + (\beta - \phi)\sigma_{t-1}^2 - (\beta\rho - (\alpha + \beta)\phi)\sigma_{t-2}^2 \quad (21)$$

این معادله حاکی از این است که مدل CGARCH یک مدل غیرخطی محدود GARCH(2,2) است.

۴-۱-۸- مدل اثر ناهمسانی واریانس شرطی در معادله میانگین<sup>۲</sup>

در برخی از موارد، ممکن است نیاز باشد واریانس شرطی را نیز به عنوان یکی از متغیرهای توضیحی وارد معادله میانگین کرد. برای مثال، وجود همبستگی مثبت میان ریسک و بازده، به عنوان یکی از تئوری‌های اساسی در مباحث مالی مطرح است.<sup>۳</sup> رابطه زیر، بیانگر یک

3. Component GARCH (CGARCH)

1. GARCH-M

۳. به عبارت دیگر، صرف ریسک بیشتر، بازدهی بیشتری را به همراه خواهد داشت.

مدل GARCH-M است که در آن واریانس شرطی با یک وقفه، وارد معادله میانگین شده است:

$$Y_t = a_0 + \sum_{l=1}^m a_l Y_{t-l} + u_t - \sum_{h=1}^n b_h u_{t-h} + \delta \sigma_{t-1}^2 ; u_t \sim N(0, \sigma_t^2) \quad (22)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 ; \alpha_0 > 0, \alpha_i, \beta_j \geq 0$$

در مدل فوق، چنانچه ضریب  $\delta$  معنی دار باشد، بیانگر وجود رابطه میان  $Y_t$  و واریانس شرطی (برای مثال بازده و ریسک) است.

### تخمین مدل های مبتنی بر واریانس شرطی

از آنجا که هیچ یک از مدل های مبتنی بر واریانس شرطی، خطی نیستند، نمی توان آنها را به روش حداقل مربعات برآورد نمود<sup>۱</sup>. به همین جهت، برای تخمین این مدل ها از روش حداکثر درستنمایی استفاده می شود. برای نشان دادن نحوه عملکرد این روش، دوباره معادله میانگین را به شکل کلی آن در رابطه ۱ در نظر بگیرید:

$$Y_t = Xa + u_t ; u_t \sim N(0, \sigma_t^2)$$

حال از آنجا که  $u_t$  دارای توزیع نرمال با میانگین صفر و واریانس  $\sigma_t^2$  است<sup>۲</sup>، دارای تابع چگالی احتمال به صورت زیر خواهد بود:

$$f(u_t) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left(-\frac{u_t^2}{2\sigma_t^2}\right) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left(-\frac{(Y_t - Xa)^2}{2\sigma_t^2}\right) \quad (23)$$

حال می توان تابع درستنمایی را برای  $n$  مشاهده به شکل زیر تشکیل داد<sup>۳</sup>:

۱. زیرا روش OLS به دنبال حداقل سازی مجموع مربعات باقیمانده های است که از تخمین معادله میانگین شرطی به دست آمده است و هیچ گونه وابستگی به واریانس شرطی ندارد.

۲. که می تواند توسط هر یک از فرآیندهای ARCH(q)، GARCH(q,p) و ... توضیح داده شود.

۳. توجه داشته باشید که در اینجا به منظور ارزیابی کارایی مدل ها و بررسی قدرت پیش بینی آنها، از  $N$  مشاهده موجود، تعداد  $n$  مشاهده را برای برآورد ضرایب مدل و همچنین انجام پیش بینی برای  $N-n=d$  دوره باقیمانده استفاده نموده و سپس نتایج پیش بینی برای  $d$  دوره را با مقادیر واقعی متناظر با آن، مقایسه خواهیم کرد.

$$L = \prod_{t=1}^n f(u_t) = \frac{1}{(\sqrt{2\pi})^n (\sigma_1 \times \sigma_2 \times \dots \times \sigma_n)} \exp\left(-\sum_{t=1}^n \left(\frac{(Y_t - Xa)^2}{2\sigma_t^2}\right)\right) \quad (24)$$

حال با گرفتن لگاریتم از طرفین رابطه فوق داریم:

$$\ln(L) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \sum_{t=1}^n \ln \sigma_t^2 - \frac{1}{2} \sum_{t=1}^n \left(\frac{(Y_t - Xa)^2}{\sigma_t^2}\right) \quad (25)$$

حال بایستی ضرایب معادله میانگین و واریانس به گونه‌ای برآورد گردند که مقدار تابع فوق حداکثر گردد. اما در بسیاری از مطالعات تجربی صورت گرفته در این زمینه، پهن‌تر بودن توزیع احتمال داده‌های مالی نسبت به تابع توزیع نرمال مورد تأکید قرار گرفته است. به همین جهت، در داده‌های مالی، از توزیع t-student که دارای دنباله‌های کشیده‌تری نسبت به توزیع نرمال است، استفاده می‌کنند. لذا چنانچه جملات خطای معادله میانگین از تابع توزیع t-student پیروی کنند، تابع حداکثر درست‌نمایی آن به شکل زیر خواهد بود:

$$\ln(L = \eta) = \Gamma\left(\frac{\eta + 1}{\sqrt{(\eta - 1)^2 + 1}}\right) - \frac{1}{2} \left[ \ln \sigma_t^2 + \eta + \ln \left(\frac{1 + u_t^2}{(\eta - 1)^2 + 1}\right) \right] \quad (26)$$

که در آن  $\Gamma(0)$  تابع چگالی گاما و  $\eta$  درجه آزادی توزیع t-student هستند.

#### ۴-۲- شبکه‌های عصبی مصنوعی<sup>۱</sup>

شبکه عصبی مصنوعی در سال ۱۹۶۲ توسط روزن بلات<sup>۲</sup> مطرح و در سال ۱۹۸۶ توسط رومل‌هارت<sup>۳</sup> و مک‌کلند<sup>۴</sup> با ابداع و ارائه مدل پرسپترون چندلایه، به طور گسترده‌ای در علوم مختلف مورد استفاده قرار گرفت. شبکه عصبی مصنوعی، از ساختاری نرونی و هوشمند، با الگوبرداری مناسب از نرون‌های موجود در مغز انسان سعی می‌کند تا از طریق توابع تعریف شده ریاضی، رفتار درون‌سلولی نرون‌های مغز را شبیه‌سازی کند و از طریق وزن‌های محاسباتی موجود در خطوط ارتباطی نرون‌های مصنوعی، عملکرد سیناپسی را در نرون‌های طبیعی به مدل درآورد. ماهیت و ذات تجربی و منعطف این روش باعث می‌شود

3. Artificial Neural Network (ANN)

4. Rosenblatt

5. Rumelhart

1. Maccland



تا در مسائلی مانند طبقه بندی، پیش بینی و ... که عموماً از رفتاری غیرخطی برخوردار هستند، به خوبی قابل استفاده باشد.

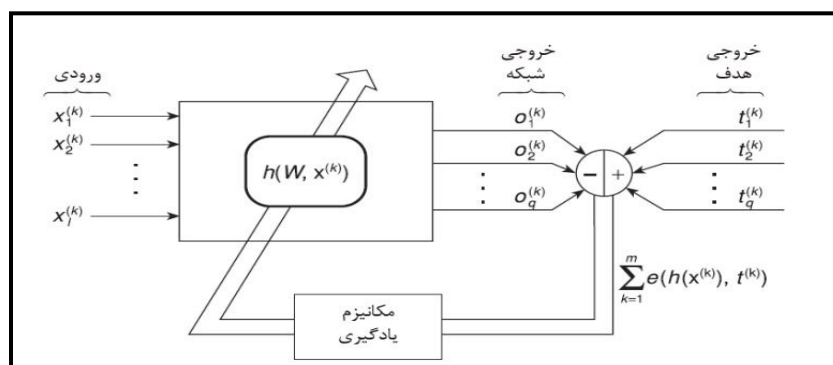
نحوه یادگیری در شبکه های عصبی به منظور پیش بینی سری های زمانی به صورت یادگیری هدایت شونده<sup>۱</sup> است. در این نوع یادگیری، شبکه عصبی مصنوعی با دریافت سیگنال ورودی سعی در تنظیم و تصحیح پارامترهای شبکه با توجه به داده های هدف دارد که این امر در فازی به نام آموزش انجام می گیرد. در فاز آموزش داده هایی با ورودی و خروجی معلوم به شبکه داده می شود و با استفاده از روش های بهینه سازی، سیستم سعی در تخمین بردار پاسخ مطلوب با استفاده از خروجی الگوریتم یادگیری دارد.

این امر با حل یک مسئله بهینه سازی با دریافت  $m$  الگو از پاسخ های مطلوب  $t^{(1)}, t^{(2)}, \dots, t^{(m)}$ ، ورودی های متناظر آن  $t^{(1)}, t^{(2)}, \dots, t^{(m)}$  و همچنین پارامترهای بهینه سازی معلوم پیشین  $W = [w_1, w_2, \dots, w_p]$  که وزن های اتصالات نامیده می شوند، انجام می گیرد. اندیس  $m$  نشان دهنده تعداد نمونه های آموزشی و اندیس  $p$  نشان دهنده تعداد وزن های اتصالات بین نرون ها<sup>۲</sup> است. هر بردار ورودی  $x$  دارای  $l$  جزء و هر بردار خروجی دارای  $q$  جزء است. الگوریتم بهینه سازی به دنبال حداقل کردن مجموع خطاهای بین بردار خروجی مطلوب  $k$  ام  $(t^{(k)})$  و  $k$  امین بردار خروجی  $(o^{(k)})$  که تابعی از وزن ها و بردار خروجی است، یعنی  $f(W, x^{(k)})$ . شکل زیر نمایی از ورودی و خروجی شبکه عصبی و فرایند یادگیری آن را نشان می دهد.

---

## 2. Supervised Learning

۲. مجموعه ای از واحدهای محاسباتی پخش شده و موازی که معمولاً در لایه هایی که توسط ارتباطات یک طرفه وزن داده شده اند، قرار می گیرند.



شکل ۳. نمایش یک ANN با مکانیزم یادگیری هدایت‌شونده

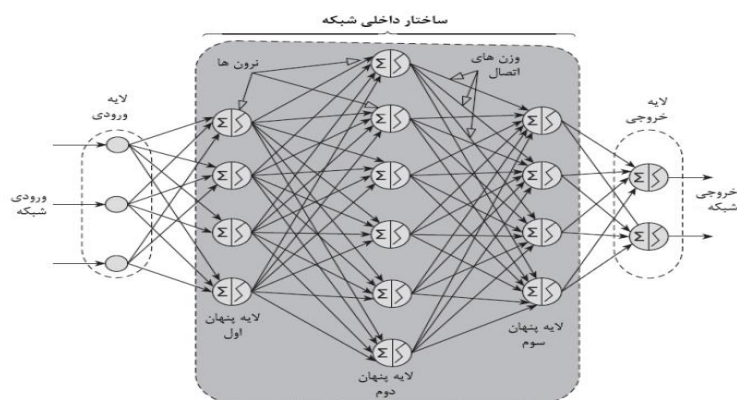
حال مسئله بهینه‌سازی به صورت زیر نشان داده می‌شود:

$$\min_w \sum_{k=1}^m e(o^{(k)}, t^{(k)}) = \min_w \sum_{k=1}^m e(f(w, x^{(k)}), t^{(k)}) \quad (27)$$

هر چه داده‌های نمونه بیشتر باشد، آنگاه الگوریتم وزن‌ها را با توجه به شرایط جدید تصحیح می‌کند. چندین تکنیک برای حل این مسئله بهینه‌سازی استفاده شده است که از جمله آن می‌توان به الگوریتم ژنتیک و الگوریتم‌های مبتنی بر بردار گرادبان اشاره کرد.

#### ساختار شبکه عصبی

ساختار نشان‌دهنده ترتیب و نحوه قرار گرفتن نرون‌ها از لایه ورودی به لایه خروجی است. نحوه استفاده از ساختار را شرایط مسئله تعیین می‌کند. یکی از پرکاربردترین ساختارهای مورد استفاده، ساختار شبکه‌های پیشخور<sup>۱</sup> است.



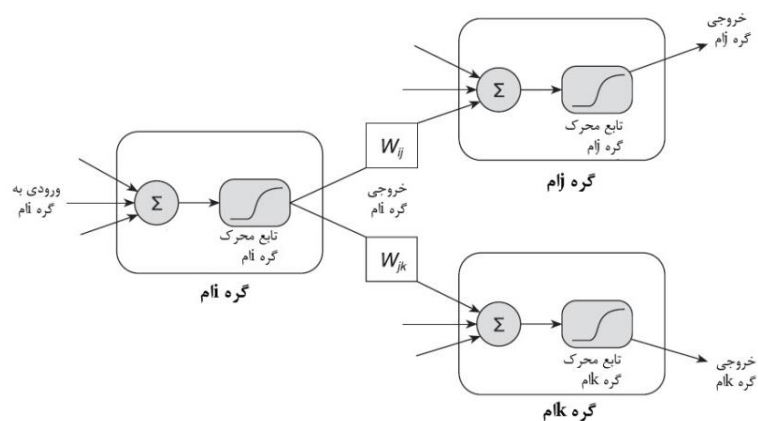
شکل ۴. نمایش شبکه عصبی مصنوعی پیش‌خور با سه ورودی و دو خروجی

در شکل فوق که نمایی از یک شبکه پیش‌خور است، نرون‌ها در یک ساختار سلسله‌مراتبی در لایه‌هایی که از لایه ورودی شروع و به لایه خروجی ختم می‌شود، قرار می‌گیرند. لایه‌های داخلی که بین دو لایه ورودی و خروجی قرار می‌گیرند، لایه‌های پنهان نامیده می‌شوند. نرون‌ها در هر لایه (مبدأ) به لایه بعدی (مقصد) از طریق مسیر یک طرفه رو به جلو مرتبط می‌شوند. بدین معنی که خروجی هر لایه، ورودی لایه بعدی خواهد بود. این ساختار به واسطه این که به الگوریتم یادگیری پسانتشار خطا<sup>۱</sup> مجهز است، بسیار مورد استفاده قرار می‌گیرد.

پس از تعیین ساختار، بایستی به تعیین تابع محرک شبکه عصبی پردازیم. هر نرون در شبکه عصبی، به مانند پردازنده ساده‌ای است که مجموع وزن‌های ورودی از دیگر نرون‌ها دریافت و با یک تابع غیرخطی که تابع محرک نامیده می‌شود آن را تبدیل به ورودی نرون بعدی می‌کند.

---

## 1. Backpropagation Learning Algorithm (BPL)



شکل ۵. ورودی و خروجی شبکه









برای مثال، خروجی  $O_k$  از نرون  $k$  ام که دارای  $l$  ورودی است به صورت زیر است:

(۲۸)

به نرون  $k$  ام،  $w_{1k}, w_{2k}, \dots, w_{lk}$  وزنهای اتصالی بین آنها و  $\theta_k$  نیز حد آستانه<sup>۱</sup> است. در این حالت، پردازشها در هر لایه به صورت همزمان انجام می‌شود، از این رو این ساختار، شبکه عصبی را با توانایی بالای محاسبات همزمان مجهز می‌کند. بر اساس نوع مسئله و مکان قرارگیری لایه، تابع محرک می‌تواند به مانند یکی از حالت‌های جدول زیر باشد:

## 2. Threshold

جدول ۱. انواع تابع محرک

ردیف	نام	تعریف تابع	علامت قراردادی
۱	آستانه دو مقداره	$y(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$	
۲	آستانه دو مقداره متقارن	$y(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}$	
۳	خطی	$y = x$	
۴	آستانه خطی متقارن	$y(x) = \begin{cases} 1 & x > 1 \\ x & -1 < x < 1 \\ -1 & x < -1 \end{cases}$	
۵	آستانه های خطی	$y(x) = \begin{cases} 1 & x > 1 \\ x & 0 < x < 1 \\ 0 & x < 0 \end{cases}$	
۶	زیگموئیدی	$y(x) = \frac{1}{1 + e^{-n}}$	
۷	تانزانته هیپربولیک	$y(x) = \frac{e^n - e^{-n}}{e^n + e^{-n}}$	
۸	خطی مثبت	$y(x) = \begin{cases} x & x > 0 \\ 0 & x < 0 \end{cases}$	

در شبکه های پرسپترون چندلایه، با امکان تغییر تعداد لایه های پنهان و همچنین تعداد نرون ها، بایستی در خصوص تعداد لایه های پنهان و تعداد نرون های هر لایه تصمیم گیری شود. اگرچه لایه های پنهان تعامل مستقیمی با محیط خارج ندارند اما تأثیر فزاینده ای بر روی خروجی نهایی دارند. سایبکوا ۱ نشان داده است که یک شبکه پرسپترون چند لایه با یک لایه پنهان که متشکل از تعداد مناسبی نرون باشد، با تابع محرک زیگموئید می تواند هر نوع تابع پیوسته را تخمین بزند و تاکنون هیچ توجیهی تئوریکی در مورد انتخاب تعداد لایه پنهان بالاتر از دو وجود ندارد.

مسئله بعدی در تعیین ساختار شبکه عصبی، انتخاب تعداد نرون در هر لایه است. اگر تعداد نرون خیلی کم در نظر گرفته شود، عدم تطابق<sup>۱</sup> رخ می‌دهد. در واقع زمانی عدم تطابق به وجود می‌آید که تعداد نرون کافی برای تخمین سیگنال پیچیده استفاده نشده باشد. از سوی دیگر، استفاده از تعداد نرون‌های بسیار زیاد در لایه‌های پنهان نیز می‌تواند منجر به چندین مشکل شود: اول این که تعداد زیاد نرون موجب بیش‌برازش<sup>۲</sup> می‌شود که در این صورت با ورود داده جدید، خطای مدل بسیار افزایش خواهد یافت و دوم اینکه با افزایش تعداد نرون‌ها، زمان آموزش آنها افزایش می‌یابد. بنابراین بایستی تعداد نرون بهینه‌میزانی بین خیلی کم و خیلی زیاد انتخاب شود.

انتخاب تعداد نرون‌های بهینه می‌تواند از طریق سعی و خطا نیز حاصل شود. البته این بدان معنا نیست که به صورت تصادفی این کار انجام شود چون در این صورت زمان زیادی صرف خواهد شد. چند روش مختلف بر اساس سعی و خطا، برای به دست آوردن بهترین ساختار ارائه شده است که پرکاربردترین آن، روش انتخاب پیش‌رو<sup>۳</sup> است. در این روش، با انتخاب تعداد نرون پایین (معمولاً از دو عدد) شبکه آموزش داده می‌شود و سپس خطای شبکه مورد ارزیابی قرار می‌گیرد، اگر خطا قابل قبول بود که تعداد نرون مشخص شده است و گرنه به تعداد نرون‌ها، یکی اضافه کرده و سپس تا آنجایی ادامه می‌دهیم که به خطای از پیش تعیین شده‌ای برسیم.

#### مکانیزم یادگیری پس‌انتشار خطا

پس از تعیین تعداد لایه‌ها و نرون‌های شبکه، بایستی الگوریتم یادگیری آن تشریح شود. در مدل پرسپترون چند لایه، به طور معمول از تکنیک کاهش گرادینان برای حل مسئله بهینه‌سازی استفاده می‌شود که شامل حداقل‌سازی خطای تجمعی شبکه<sup>۴</sup> است که آن را با  $E_c$  نشان می‌دهیم. در واقع  $E_c$  بیانگر مجموع  $n$  مربع خطای  $E(k)$  با رابطه

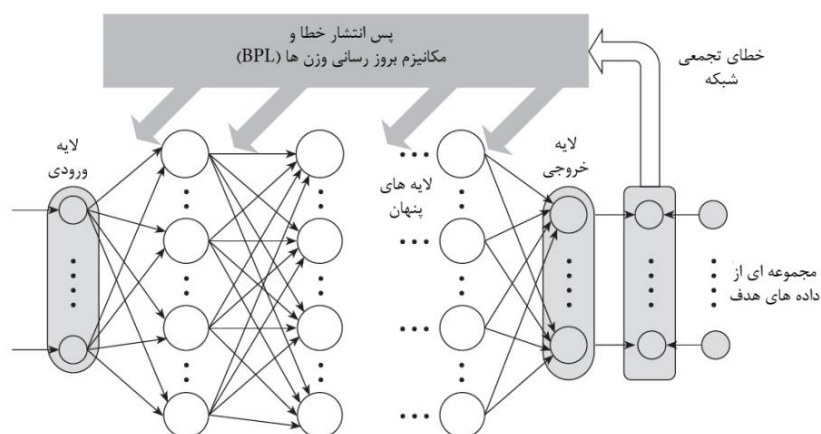
$$E(k) = \frac{1}{2} \sum_{i=1}^q [t_i(k) - o_i(k)]^2$$

است که  $E(k)$  مجذور برداری تفاوت بین  $k$  امین بردار

- 
1. Underfitting
  2. Overfitting
  3. Forward Selection Method
  4. Network Cumulative Error

خروجی شبکه  $O(k)$  و بردار خروجی هدف  $t(k)$  و  $n$  نیز تعداد اعضای نمونه در فاز آموزش است.

به روز رسانی وزن ها در این الگوریتم به گونه ای طراحی شده است که در آن وزن ها در جهت عکس حرکت بردار گرادیان خطای تجمعی، تعدیل شوند. بر این اساس، با ورود داده آموزشی، سیگنال بازخوردی (خطای بین خروجی شبکه و هدف) ساخته شده و سپس با انتشار این خطا به عقب (به واسطه الگوریتم پس انتشار خطا)، وزن ها به روز رسانی می شوند. شکل زیر نحوه عملکرد این مکانیزم را نشان می دهد.



شکل ۶. مکانیزم یادگیری پس انتشار خطا

در ادامه با فرض اینکه تابع محرک زیگموئیدی است، به بررسی این الگوریتم می پردازیم. از آنجایی که می خواهیم خطای تجمعی را کمینه کنیم:

$$E_c = \sum_{k=1}^n E(k) = \frac{1}{2} \sum_{k=1}^n \sum_{i=1}^q [t_i(k) - o_i(k)]^2 \quad (29)$$

در این قسمت، بسته به این که بخواهیم بهینه سازی را به صورت آنلاین (آموزش نمونه به نمونه صورت می پذیرد) و یا آفلاین (تمامی نمونه های آموزش به سیستم ارائه می شود و سپس با توجه به خطای تجمعی به روز رسانی وزن ها صورت می گیرد) انجام دهیم، مسئله بهینه سازی به یکی از صورت های زیر در می آید:

$$\text{offline: } \min_w E_c = \min_w \frac{1}{2} \sum_{k=1}^n \sum_{i=1}^q [t_i(k) - o_i(k)]^2 \quad (30)$$

$$\text{online: } \min_w E(k) = \min_w \frac{1}{2} \sum_{i=1}^q [t_i(k) - o_i(k)]^2 \quad (31)$$

در هر دو رابطه فوق،  $w$  نشان‌دهنده بردار وزن‌های اتصالات بین نرون‌ها است. در این مقاله به علت بهتر بودن نتایج استفاده از بهینه‌سازی آنلاین وزن‌ها (بر اساس سایر مطالعات صورت گرفته)، از همین روش استفاده خواهیم کرد. در این حالت، تغییرات وزن‌ها بر اساس رابطه زیر انجام می‌گیرد:

$$\Delta w^{(l)} = -\eta \frac{\partial E(k)}{\partial w^{(l)}} \quad (32)$$

به طوری که  $\frac{\partial E(k)}{\partial w^{(l)}}$  بردار گرادیان  $E(k)$  نسبت به بردار  $w^{(l)}$  است که نشان‌دهنده تمام وزن‌هایی است که بین لایه 1 و لایه ماقبل آن، یعنی  $l-1$ ، وجود دارد. همچنین  $\eta$  نشان‌دهنده نرخ یادگیری است که معمولاً مقداری کوچک مثبت (معمولاً بین 0 و 1) در نظر گرفته می‌شود که در همگرایی و پایداری الگوریتم یادگیری نقش مهمی دارد. همچنین  $\Delta w^{(l)}$  نشان‌دهنده اختلاف بین بردار  $w^{(l)}$  و  $w^{(l)(k)}$  است که بیانگر وزن اتصالی به لایه  $l$  قبل و بعد از ورود نمونه آموزشی  $k$ ام، است. این الگوریتم در شکل زیر نشان داده شده است:



شکل ۷. نمایش ارتباطات در پرسپترون چند لایه



$\Delta w_{ij}^{(l)}$  میزان تغییر وزن ارتباط بین نرون  $j$  از لایه  $l-1$  و نرون  $i$  واقع در لایه  $l$  را بیان می کند. همچنین  $O_j^{(l-1)}$  نمایش خروجی زامین نرون واقع در لایه  $l-1$  است. سیگنال  $tot_i^l$  نشان دهنده مجموع وزنی تمام سیگنال‌هایی است که به نرون  $i$  واقع در لایه پنهان  $l$ ، از لایه قبلی آن یعنی  $l-1$ ، می رسد. حال برای بازکردن رابطه ۳۲، با استفاده قانون مشتق زنجیره ای داریم:

$$\Delta w_{ij}^{(l)} = \Delta w_{ij}^{(l)} = -\eta \left[ \frac{\partial E(k)}{\partial o_i^{(l)}} \right] \left[ \frac{\partial o_i^{(l)}}{\partial tot_i^{(l)}} \right] \left[ \frac{\partial tot_i^{(l)}}{\partial w_{ij}^{(l)}} \right] \quad (33)$$

حال چنانچه لایه مورد بررسی، لایه خروجی باشد، به جای استفاده از  $l$  از  $L$  استفاده خواهیم کرد که در نتیجه در این حالت رابطه فوق را به صورت زیر می توان بازنویسی کرد:

$$\Delta w_{ij}^{(L)} = -\eta (t_i - o_i^{(L)}) f'(tot_i^{(L)}) o_j^{(L-1)} \quad (34)$$

که در آن

$$f'(tot_i^{(l)}) = \frac{\partial f(tot_i^{(l)})}{\partial tot_i^{(l)}} \quad (35)$$

حال اگر کمیت  $\delta_i^{(L)} = (t_i - o_i^{(L)}) f'(tot_i^{(L)})$  را به عنوان سیگنال خطای لامین نرون لایه خروجی معرفی کنیم، می توان رابطه به روزرسانی به صورت زیر نوشت:

$$\Delta w_{ij}^{(L)} = \eta \delta_i^{(L)} o_j^{(L-1)} \quad (36)$$

دوباره با فرض تابع محرک زیگموئیدی، خطای سیگنال را می توان به صورت زیر نوشت:

$$\delta_i^{(L)} = (t_i - o_i^{(L)}) o_i^{(L)} (1 - o_i^{(L)}) \quad (37)$$

حال با استفاده از رابطه زیر می توان خطا را به عقب انتشار داد که در آن  $l$  نشان دهنده یک لایه پنهان ( $l < L$ ) است و عبارت  $\Delta w_{ij}^{(l)}$  نیز از طریق رابطه زیر به دست می آید:

$$\Delta w_{ij}^{(l)} = \eta \delta_i^{(l)} o_j^{(l-1)} \quad (38)$$

سیگنال خطا به عنوان تابع خروجی لایه های قبلی، به صورت زیر محاسبه می شود:

$$\delta_i^{(l)} = f'(tot_i^{(l)}) \sum_{p=1}^{n_l} \delta_i^{(l+1)} w_{pi}^{(l+1)} \quad (39)$$

حال دوباره با فرض اینکه  $f$ ، به صورت زیگموئیدی باشد،  $\delta_i^{(l)}$  را می توان به صورت زیر نوشت:

$$\delta_i^{(l)} = o_i^{(l)} (1 - o_i^{(l)}) \sum_{p=1}^{n_l} \delta_i^{(l+1)} w_{p_i}^{(l+1)} \quad (40)$$

عبارت  $n_l$  در رابطه‌های ۳۷ و ۳۸ نشان دهنده مجموع کل نرون‌ها در لایه  $l+1$ ، اندیس  $i$  نشان دهنده نرون‌های لایه  $l$  و اندیس  $p$  نشان دهنده نرون‌های لایه  $l+1$  است. در واقع، این حقیقت که  $\delta_i^{(l)}$  در رابطه ۳۶، مقدارش را از دانش موجود در  $\delta_p^{(l+1)}$  لایه بعدی می‌گیرد، باعث نام‌گذاری این روش به پس‌انتشار خطا شده است. مراحل اجرای الگوریتم:

گام اول: با وزن‌ها و اریب اولیه که تصادفی و کوچک می‌باشند، شروع می‌کنیم.  
گام دوم: یک الگو ورودی - خروجی از داده‌های آموزشی انتخاب می‌کنیم:  $(x(k), t(k))$ .  
گام سوم:  $k$  امین سیگنال را در طول شبکه رو به جلو عبور داده و خروجی هر نرون  $i$  در هر لایه  $l$  را با استفاده از رابطه زیر به دست می‌آوریم:

$$o_i^l(k) = f \left( \sum_{p=0}^{n_{l-1}} w_{ip}^l o_p^{l-1} \right) \quad (41)$$

گام چهارم: خطای کل را با استفاده از  $E = E(k) + E$  و سیگنال خطای  $\delta_i^{(L)}$  با استفاده از رابطه زیر محاسبه می‌کنیم:

$$\delta_i^{(L)} = (t_i - o_i^L) f'(tot_i^{(L)}) \quad (42)$$

گام پنجم: وزن‌ها را بر اساس  $\Delta w_{ij}^{(L)} = \eta \delta_i^{(L)} o_j^{(L-1)}$  برای  $i=1, 2, \dots, L$  با استفاده از رابطه زیر به دست می‌آوریم:

$$\delta_i^{(l)} = f'(tot_i^{(l)}) \sum_{p=1}^{n_l} \delta_i^{(l+1)} w_{p_i}^{(l+1)} \quad \text{for } l < L \quad (43)$$

گام ششم: مراحل بالا را از گام دوم به بعد، برای سایر داده‌های آموزشی اجرا کرده و هرگاه تمام نمونه‌ها مورد استفاده قرار گرفت، یک Epoch به اتمام می‌رسد.  
گام هفتم: اگر خطای تجمعی  $E$  در لایه خروجی از یک میزان از پیش تعیین شده‌ای کمتر باشد، به این نتیجه می‌رسیم که آموزش به اتمام رسیده است وگرنه دوباره این مراحل را تا رسیدن به خطای مورد نظر، تکرار می‌کنیم.

استفاده از روش کاهش گرادیان، پایه بسیاری از الگوریتم های بهینه سازی است و از آن در یادگیری شبکه های عصبی و حداقل سازی مقدار خطای شبکه برای تنظیم پارامترهای شبکه استفاده می شود. دسته وسیعی از تحقیقات برای افزایش سرعت این الگوریتم در آموزش شبکه های عصبی چند لایه پیشخور انجام شده است که از نتایج آن می توان به انواع روش های گرادیان مرتبه اول (که در قسمت قبل به آن اشاره شد) و دوم اشاره کرد. روش های مبتنی بر گرادیان مرتبه دوم، برای سرعت دادن به روش های گرادیان مرتبه اول و برطرف کردن معایب آن ها مطرح شده اند که در آن از مشتقات جزئی مرتبه دوم تابع خطا یا خروجی شبکه استفاده می شود. رابطه زیر روش نیوتن کلاسیک را که پایه بسیاری از روش های نیوتنی و گرادیان مرتبه دوم است، نشان می دهد:

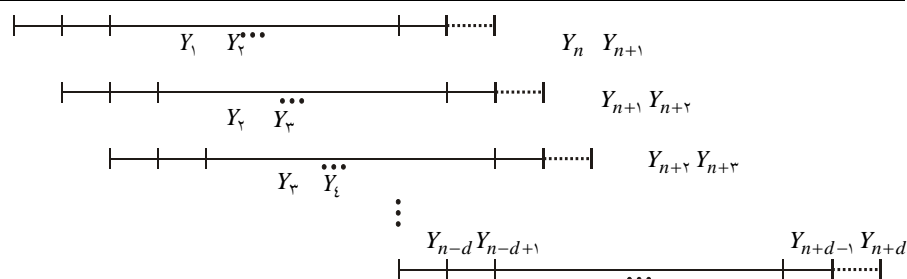
$$w_{k+1} = w_k - \varepsilon H^{-1}g \quad (44)$$

در رابطه بالا،  $H$  ماتریس هشین<sup>۱</sup>،  $g$  بردار گرادیان تابع خطای شبکه و  $\varepsilon$  نرخ آموزش است. در این روش تغییر وزن ها، با محاسبه مستقیم ماتریس هشین انجام می شود. نکته کلیدی و تفاوت روش های مبتنی بر گرادیان مرتبه دوم در محاسبه معکوس ماتریس هشین است.

#### ۴-۳- نحوه انجام پیش بینی

نحوه انجام پیش بینی در هر دو روش شبکه عصبی مصنوعی و مدل های مبتنی بر واریانس ناهمسانی شرطی به این صورت است که در ابتدا یک حجم نمونه انتخابی ثابت ( $n$ ) به صورت پنجره ای رو به جلو<sup>۲</sup> به صورت دوره به دوره، در طول دوره پیش بینی حرکت داده می شود، سپس مدل با استفاده از داده های هر پنجره مدل با استفاده از روش شبکه عصبی مصنوعی و مدل های مبتنی بر واریانس ناهمسانی شرطی برآورد و سپس یک مقدار متغیر وابسته برای یک دوره بعد پیش بینی می شود. این روند تا جایی که  $d$  پیش بینی انجام گیرد، ادامه خواهد یافت. شکل زیر نمایی از چنین روش تخمینی است:

۱. ماتریس هشین، گرادیان مرتبه دوم تابع خطا نسبت به وزن های شبکه می باشد.



شکل ۸. استفاده از پنجره رو به جلو در تخمین مدل

با توجه به شکل فوق، برای مثال پیش‌بینی  $n+1$  امین مشاهده متغیر  $Y_t$ ، با استفاده از  $n$  مشاهده قبلی آن یعنی  $Y_1, Y_2, \dots, Y_n$ ، امکان‌پذیر است. در نتیجه برای پیش‌بینی  $d$  دوره به صورت مشاهده به مشاهده، می‌توان از الگوی زیر استفاده نمود:

جدول ۲. الگوی پیشنهادی برای پیش‌بینی مشاهده به مشاهده داده‌ها

مشاهده مورد نظر برای پیش‌بینی	داده‌های مورد استفاده برای پیش‌بینی
$Y_{n+1}$	$Y_1, Y_2, \dots, Y_n$
$Y_{n+2}$	$Y_2, Y_3, \dots, Y_{n+1}$
$\vdots$	$\vdots$
$Y_{n+d}$	$Y_{n-d}, Y_{n-d+1}, \dots, Y_{n+d-1}$

#### ۴-۴- نحوه ارزیابی مدل‌ها

برای ارزیابی مدل‌های مختلف مبتنی بر واریانس ناهمسانی شرطی و شبکه عصبی مصنوعی به منظور پیش‌بینی VaR، می‌توان از تابع زیان صفر و یک استفاده کرد. در این تابع که توسط لویز<sup>۱</sup> ارائه گردید، در دوره پیش‌بینی، هر مقدار زیان واقعی که بیشتر از ارزش در معرض خطر باشد، به عنوان یک استثناء<sup>۲</sup> تلقی گردیده و به آن عدد یک اختصاص می‌یابد و در غیر این صورت تابع مقدار صفر به خود می‌گیرد. با این تعریف، تابع زیان مورد نظر را می‌توان به صورت زیر نوشت:

1. Lopez
2. Exception

$$I_{t+1} = \begin{cases} 1 & ; r_t < VaR_t(k, \alpha) \\ 0 & ; r_t \geq VaR_t(k, \alpha) \end{cases}$$

بنابراین هر مدلی که دارای تعداد استثنائات کمتری باشد، مدل بهتری ارزیابی می‌گردد. همچنین بر این اساس، کوپیک<sup>۱</sup> آزمون پوشش غیرشرطی<sup>۲</sup> را مطرح ساخت که طی آن، در صورتی مدل‌های تبیین‌کننده ارزش در معرض خطر، مناسب تشخیص داده می‌شوند که با در نظر گرفتن سطح اطمینان  $\alpha$ ، نسبت استثنائات<sup>۳</sup> در دوره پیش‌بینی برابر با  $1-\alpha$  محاسبه گردد. به عبارت دیگر، اگر سطح معنی‌داری برای محاسبه ارزش در معرض خطر برابر  $\alpha$  درصد منظور گردد، نسبت  $\frac{x}{d}$  (که در آن  $x$  تعداد استثنائات و  $d$  طول دوره پیش‌بینی است) بایستی نزدیک به  $1-\alpha$  باشد. به راحتی می‌توان دریافت که اگر نسبت مذکور بزرگ‌تر از  $1-\alpha$  درصد باشد، معیار ارزش در معرض خطر، ریسک را به درستی نشان نمی‌دهد و به تعبیری آن را دست کم می‌گیرد. همچنین در حالت عکس، یعنی  $\frac{x}{d} < 1-\alpha$  باشد، ریسک بیش از حد واقعی برآورد شده است.

برای محاسبه نسبت استثنائات که آن را با  $\hat{\alpha}$  نشان می‌دهیم، می‌توان از رابطه زیر استفاده نمود:

$$\hat{\alpha} = \frac{x}{d} = \frac{1}{d} \sum_{i=1}^d I_i$$

در نهایت، فرضیه  $H_0 = \hat{\alpha} = \alpha$  که بیانگر برابر بودن سطح احتمال در نظر گرفته شده برای محاسبه ارزش در معرض خطر و احتمال وقوع استثنائات است، با استفاده از آماره نسبت درستنمایی زیر مورد آزمون قرار می‌گیرد:

$$LR_{UC} = 2[\log \hat{\alpha}^x (1-\hat{\alpha})^{d-x} - \log \alpha^x (1-\alpha)^{d-x}]$$

آماره فوق دارای توزیع مجانبی  $\chi^2$  را درجه آزادی یک است.

3. Kupiec

4. Unconditional Coverage

۳. تعداد استثنائات تقسیم بر کل دوره پیش‌بینی.

## ۵- داده‌های پژوهش و توصیف آماری مقدماتی آنها

داده‌های مورد استفاده در این پژوهش، داده‌های شاخص ۵۰ شرکت برتر بورس ۱ است که به طور هفتگی از تاریخ ۱۳۸۲/۰۱/۰۵ تا ۱۳۹۰/۰۴/۲ در نظر گرفته شده است که مشتمل بر ۴۲۹ مشاهده است. سپس با استفاده از این شاخص، بازده‌های هفتگی با استفاده از رابطه  $r_t = \log(P_t / P_{t-1})$  محاسبه گردیده‌اند که در آن  $r_t$  و  $P_t$  به ترتیب، بازده هفتگی و شاخص قیمت ۵۰ شرکت برتر بورس در هفته  $t$  ام هستند. در جدول زیر خلاصه‌ای از ویژگی‌های آماری بازده لگاریتمی شاخص مذکور ( $r_t$ ) آمده است:

جدول ۳. ویژگی‌های آماری شاخص ۵۰ شرکت برتر

آماره چارک-بر	حداکثر	حداقل	کشیدگی	چولگی	انحراف معیار	میانگین	شاخص ۵۰ شرکت برتر
۴۹۶/۷۱	۰/۱۲۰۷	-۰/۱۰۱۰	۸/۶۶۱۴	۰/۰۲۲۶	۰/۰۲۳۵	۰/۰۰۰۷	

مطابق جدول فوق، در طی دوره مورد بررسی، میانگین بازده‌ها تقریباً برابر صفر است. همچنین میزان انحراف معیار در مقایسه با میانگین، نسبتاً زیاد بوده است که نشان از نوسان بالای شاخص در دوره مورد نظر دارد. همچنین میزان کشیدگی و چولگی به طور قابل توجهی بیشتر از توزیع نرمال است که به همین دلیل آماره چارک‌بر مبنی بر نرمال بودن توزیع بازده‌ها در ناحیه رد قرار گرفته است.

## ۶- نتایج تجربی حاصل از برآورد مدل‌ها

پیش از برآورد مدل‌های مبتنی بر واریانس ناهمسانی شرطی، بایستی متغیر یا ثابت بودن واریانس جمله خطا با استفاده از آزمون اثرات ARCH، بررسی و در صورت رد فرضیه صفر

۱. شاخص ۵۰ شرکت برتر، میانگین قیمت سهام ۵۰ شرکت برتر به لحاظ نقدشوندگی است که در محاسبه آن، تعداد سهام منتشره شرکت تأثیر داده نمی‌شود و تنها از جمع قیمت سهام این شرکت‌ها تقسیم بر تعداد آنها (تعدیل شده) به دست می‌آید. شرکت‌های حاضر در جمع ۵۰ شرکت برتر در وقفه‌های سه ماهه توسط سازمان بورس اوراق بهادار معرفی می‌شوند.

واریانس همسانی در این آزمون، به مرحله دوم برویم. در این آزمون، (با فرض واریانس همسانی) بهترین مدل ARIMA بر روی داده‌ها رگرسی می‌شود و سپس با استفاده از باقیمانده‌های آن  $(\hat{u}_{t-i}; i=0,1,\dots,q)$ ، مدل رگرسیون کمکی زیر تشکیل می‌شود:

$$\hat{u}_t^2 = \alpha_0 + \alpha_1 \hat{u}_{t-1}^2 + \dots + \alpha_q \hat{u}_{t-q}^2 + v_t$$

حال با محاسبه ملاک آزمون اثرات آرچ به شکل  $nR^2$ ، فرضیه  $H_0: \alpha_i; i=1,2,\dots,q$  در مقابل فرضیه  $H_1: \alpha_i \neq 0$  آزمون می‌شود که در صورت غیر صفر بودن حداقل یکی از  $\alpha_i$  ها، مدل دچار واریانس ناهمسانی شرطی خواهد بود. نتایج آزمون به صورت زیر است که بیانگر لزوم استفاده از مدل‌های مبتنی بر واریانس ناهمسانی شرطی است:

جدول ۴. نتایج آزمون آرچ

آماره	مقدار	احتمال
F	۱۶/۲۰۰۳	۰/۰۰۰۰
$nR^2$	۲۷/۰۱۶۹	۰/۰۰۰۰

منبع: یافته‌های پژوهشگر

در مرحله دوم، در ابتدا هر یک از مدل‌های مبتنی بر واریانس ناهمسانی شرطی با وقفه‌های مختلف AR و MA در معادله میانگین و وقفه‌های مختلف  $\sigma_{t-i}^2$  و  $u_{t-i}^2$  در معادله واریانس، به طور همزمان برآورد و سپس بهترین مدل با استفاده از معیار اطلاعاتی آکایکه انتخاب و پیش‌بینی بازده و انحراف معیار آن برای یک دوره بعد انجام می‌شود. توجه داشته باشید که مراحل ذکر شده با استفاده از مفهوم پنجره رو به جلو، به صورت مشاهده به مشاهده برای ۱۰۰ دوره انجام می‌گیرد.

در روش شبکه عصبی مصنوعی نیز تعداد لایه‌ها و نرون‌ها با استفاده از معیار حداقل مربعات خطا<sup>۱</sup> انتخاب و سپس شبکه بهینه برای پیش‌بینی مشاهده به مشاهده مقادیر ارزش در معرض خطر برای ۱۰۰ دوره مورد استفاده قرار خواهد گرفت. در نهایت مقدار ارزش در معرض خطر برای هر یک از دوره‌ها (دوره پیش‌بینی) و مدل‌ها با استفاده از رابطه  $VaR_t^N = \hat{\mu}_t + z_\alpha \hat{\sigma}_t$  محاسبه خواهد شد.

1. Mean Square Error (MSE)

همانطور که در قسمت‌های قبل نیز اشاره شد، اولین مرحله برای ارزیابی روش‌های مختلف محاسبه ارزش در معرض خطر، استفاده از یک معیار آماری برای تعیین اینکه روش‌های ذکر شده، ریسک شاخص یا سبد سهام مورد بررسی را دست کم نگرفته و یا بیش از حد برآورد نکنند، است. برای این منظور از آزمون پوشش غیرخطی کوپیک که در آن فرضیه صفر مبنی بر برابری نرخ شکست محاسبه شده و سطح معنی‌داری مورد نظر (در اینجا ۰.۵٪) است، استفاده می‌شود:

$$H_0: \hat{\alpha} = \alpha$$

$$H_1: \hat{\alpha} \neq \alpha$$

نتایج انجام آزمون برای هر یک از روش‌های مبتنی بر واریانس ناهمسانی شرطی و روش شبکه عصبی مصنوعی در جدول زیر آمده است:

جدول ۵. مقایسه روش‌های مختلف محاسبه VaR با استفاده از آزمون کوپیک

نتیجه آزمون	مقدار بحرانی $\chi^2_{(1)}$	آماره LR	نرخ شکست	مدل
قبول	۶/۶۳۵	۰/۰۹۷۹	۰/۰۴	$GARCH(1,1)$
قبول	۶/۶۳۵	۰/۰۹۷۹	۰/۰۴	$EARCH(1,1)$
قبول	۶/۶۳۵	۰/۰۹۷۹	۰/۰۴	$TARCH(1,1)$
قبول	۶/۶۳۵	۰/۰۹۷۹	۰/۰۴	$PARCH(1,1)$
قبول	۶/۶۳۵	۰/۰۸۶۲	۰/۰۶	$CGARCH(1,1)$
قبول	۶/۶۳۵	۰/۳۲۷۰	۰/۰۷	$ARCHM(1,1)$
قبول	۶/۶۳۵	۰	۰/۰۵	neural network

منبع: یافته‌های پژوهشگر

مطابق جدول فوق، کلیه روش‌های محاسبه ارزش در معرض خطر از قدرت توضیحی مناسبی در برآورد VaR برخوردار هستند و آزمون LR مبنی بر تصریح درست ریسک، برای همه آنها مورد قبول واقع شده است اما روش شبکه عصبی مصنوعی با آماره LR محاسباتی صفر، بهترین عملکرد را در برآورد ارزش در معرض خطر در طی دوره پیش‌بینی به روش‌های مبتنی بر واریانس ناهمسانی شرطی دارد.



### کاربرد روش شبکه عصبی مصنوعی و مدل های واریانس ناهمسانی شرطی در محاسبه ارزش در معرض خطر ۱۳۳

حال برای ارزیابی و درجه بندی دقیق تر روش های مذکور، می توان با محاسبه مقادیر ارزش در معرض خطر واقعی و مقادیر پیش بینی شده آن (در دوره پیش بینی) توسط هر یک از روش ها، به این مهم دست یافت که نتایج آن در جدول زیر آمده است:

جدول ۶. مقایسه روش های مختلف محاسبه VaR با استفاده از مقایسه مقادیر واقعی و

پیش بینی آن

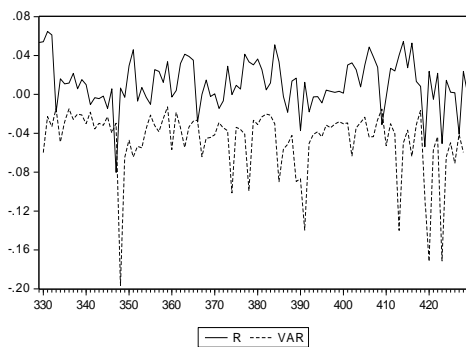
مدل	$VaR_{5\%}$	مدل	$VaR_{5\%}$
مقدار واقعی	-/۰.۳۳۹	$PARCH(1,1)$	-/۰.۴۸۶
$GARCH(1,1)$	-/۰.۴۵۳	$CGARCH(1,1)$	-/۰.۳۹۷
$EGARCH(1,1)$	-/۰.۴۷۲	$ARCHM(1,1)$	-/۰.۳۶۳
$TARCH(1,1)$	-/۰.۴۵۳	شبکه عصبی	-/۰.۳۸۲

منبع : یافته های پژوهشگر

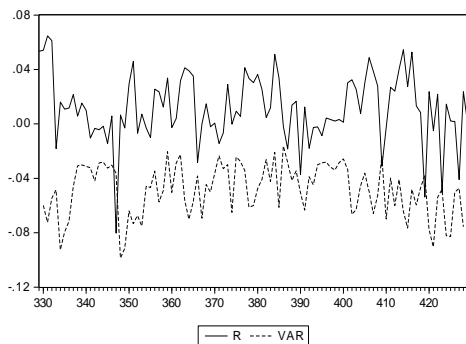
همانطور که در جدول فوق مشاهده می شود، VaR محاسبه شده به روش  $ARCHM(1,1)$  بهترین عملکرد را (حتی شبکه عصبی مصنوعی) داشته است و مقدار آن نزدیک تر به مقدار واقعی VaR است. اما با این وجود، نتیجه روش شبکه عصبی نسبت سایر روش های مبتنی بر واریانس ناهمسانی شرطی بهتر بوده و همچنین مطابق جدول ۵، تعداد خطای آن در محاسبه VaR به حالت نرمال نزدیک تر است.

نمودارهای زیر نیز شاهی بر بهتر بودن نتایج VaR محاسباتی روش شبکه عصبی به میزان واقعی آن نسبت به سایر روش ها است.

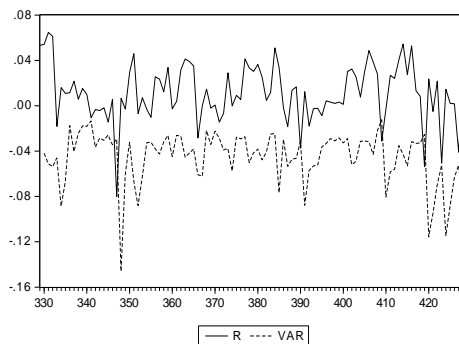
EGARCH(1,1) Forecasting



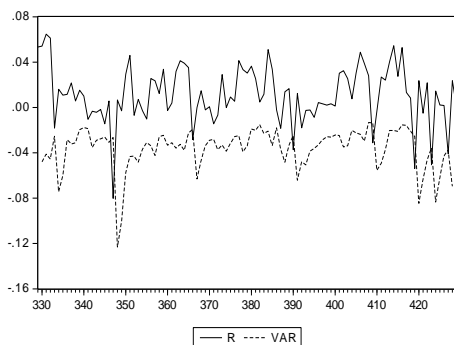
PARCH(1,1) Forecasting



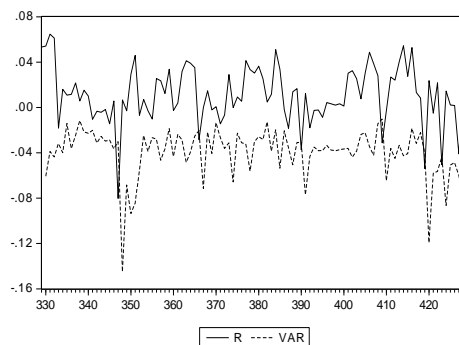
TARCH(1,1) Forecasting



ARCHM(1,1) Forecasting



CGARCH(1,1) Forecasting



## ۷- نتیجه گیری

امروزه محاسبه ارزش در معرض خطر، یکی از معیارهای متداول برای بررسی و اندازه گیری ریسک های مرتبط با سبدهای مختلف دارائی، صنایع و ... است. ارزش در معرض خطر، تلاشی است برای اینکه عدد معینی در خصوص ریسک صنعت و یا سبد سهام به تحلیل گران و مدیران ارشد مؤسسات ارائه کند.

در این تحقیق به پیش بینی ارزش در معرض خطر مرتبط با پرتفوی متشکل از ۵۰ شرکت با نقدشوندگی بالا در بازار اوراق بهادار تهران به صورت پنجره ای رو به جلو، با استفاده از دو رویکرد شبکه عصبی مصنوعی و مدل های مختلف مبتنی بر واریانس ناهمسانی شرطی شامل GARCH، EGARCH، CGARCH، TARCH و GARCH-M پرداختیم.

نتایج نشان می دهد که روش شبکه عصبی مصنوعی در مقایسه با سایر روش های مبتنی بر واریانس ناهمسانی شرطی، به نحو مناسب تری توانسته میزان ریسک را پیش بینی کند، به گونه ای که بر اساس آزمون پوشش غیرشرطی کوپیک، در دوره پیش بینی به میزان ۱۰۰ مشاهده، این روش در ۵ مورد خطا داشته که همین امر بیانگر بهتر بودن نتایج این روش نسبت به سایر روش ها بر اساس این آزمون است. از سوی دیگر در مقایسه VAR واقعی در دوره پیش بینی با مقادیر محاسباتی نیز روش شبکه عصبی عملکرد نسبتاً بهتری داشته است.

## منابع

- محمدی، شاپور و راعی، رضا و فیض‌آباد، آرش، ۱۳۸۷، محاسبه ارزش در معرض خطر پارامتریک با استفاده از مدل‌های واریانس ناهمسانی شرطی در بورس اوراق بهادار تهران، تحقیقات مالی، دوره ۱۰، شماره ۲۵، بهار و تابستان ۱۳۸۷، صفحه ۱۰۹ تا ۱۲۴.
- سجاد، رسول و هدایتی، شهره، ۱۳۹۳، برآورد ارزش در معرض خطر با استفاده از نظریه ارزش فرین در بورس اوراق بهادار تهران، فصلنامه علمی پژوهشی دانش سرمایه‌گذاری، سال سوم، شماره نهم.
- اسلامی بیدگلی، غلامرضا و راعی، رضا و کمالزاده، سحر، ۱۳۹۲، محاسبه ارزش در معرض خطر قیمت سبد نفتی اوپک با استفاده از مدل‌های حافظه بلندمدت گارچ، فصل‌نامه مطالعات اقتصاد انرژی، سال دهم، شماره ۳۹، صفحات ۱-۱۹.
- Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, Vol. 31, No. 3, PP. 307-327.
- Christiansen, C., 1999, "Value at Risk Using the Factor-ARCH Model", *The Journal of Finance*, Vol. 1, No. 2.
- Engle, Robert F., 1982, "Autoregressive Conditional Heteroscedasticity with Estimation of the Variance of United Kingdom Inflation", *Econometrica*, Vol. 50, No. 4, PP. 987-1007.
- Giot, P., and Laurent, S., 2003a, "Value at Risk for Long and Short Trading Positions", *Journal of Applied Econometrics*, Vol. 18, PP. 641-664.
- Glosten, L., Jaganathan, R., and Runkle, D., 1993, "Relationship between the Expected Value and Volatility of the Normal Excess Returns on Stocks", *Journal of Finance*, Vol. 48, PP. 1779-1802.
- Jorion, P., 2000, *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill, New York.
- Kupiec, P., 1995, "Techniques for Verifying the Accuracy of Risk Measurement Models", *Journal of Derivatives*, Vol. 3, PP. 73-84.
- Lopez, J., 1999, "Methods for Evaluating Value-at-Risk estimates, Federal Reserve Bank of San Francisco", *Economic Review*, Vol. 2, PP. 3-17.
- Nelson, D., 1991, "Conditional Heteroscedasticity in Asset Returns: A new Approach", *Econometrica*, 59, 347-370.

- Marimoutou V, Raggad, B, Trabelsi, A, 2009, Extreme Value Theory and Value at Risk: Application to oil market, Energy Economics, 519-530.
- Chan, N. H., S. Deng, L. Peng, and Z. Xia, 2007, Interval estimation of Value at Risk based on GARCH models with heavy-tailed innovations. Journal of Econometrics, 137, 556-576.