پیش‌بینی سود هر سهم با استفاده از شبکه‌های عصبی در شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران

رضوان حجازی *
شاپور محمدی **
پروانه فایقی ***

چکیده

هدف اصلی تحقیق حاضر بررسی دقت پیش‌بینی سود هر سهم با استفاده از شبکه‌های عصبی و مقایسه آن با مدل‌های خطی می‌باشد. همچنین اثر متغیرهای بنیادی حسابداری بر پیش‌بینی سود هر سهم نیز مورد بررسی قرار گرفت. برای این منظور مدل‌های خطی و غیرخطی به صورت تک‌متغیره و چندمتغیره مورد استفاده قرار گرفتند. لذا از هفت متغیر اثرگذار بر سود هر سهم به عنوان متغیرهای مستقل و سود هر سهم به عنوان متغیر وابسته استفاده شده است. از طرفی برای آموزش شبکه‌های عصبی از دو الگوریتم پس‌انتشار خطا و الگوریتم ژنتیک استفاده گردید و دقت پیش‌بینی این دو الگوریتم مقایسه گردید. در این تحقیق تعداد ۹۱ شرکت از سال ۱۳۸۳ تا سال ۱۳۸۹ به صورت فصلی مورد بررسی قرار گرفت. از روش رگرسیون پنلی جهت مدل خطی و از شبکه عصبی پیشخور تعمیم یافته جهت بررسی از طریق شبکه عصبی استفاده شد. نتایج حاصل از این تحقیق نشان داد شبکه‌های عصبی که در آن از متغیرهای بنیادی حسابداری استفاده گردید، دقت بالاتری در پیش‌بینی سود هر سهم نسبت به دیگر روش‌ها داشت. به طور کل می‌توان گفت افزودن متغیرهای بنیادی حسابداری دقت پیش‌بینی شبکه‌های عصبی را افزایش می‌دهد. در مورد مقایسه دقت پیش‌بینی بین دو الگوریتم آموزشی ژنتیک و پس‌انتشار خطا با توجه به نتایج متفاوتی که از گره‌های مختلف حاصل شد، امکان قضاوت قطعی وجود ندارد.

کلمات کلیدی: سود هر سهم، مدل‌های خطی، شبکه‌های عصبی، الگوریتم پس‌انتشار خطا، الگوریتم ژنتیک.

* دانشیار دانشکده علوم اجتماعی و اقتصاد دانشگاه الزهرا (س)، ایمیل: hejazi۳۳@yahoo.com (نویسنده مسئول)
** دانشیار دانشکده مدیریت دانشگاه تهران
*** کارشناس ارشد حسابداری دانشگاه الزهرا (س)

مقدمه

توجه به کاربرد تکنیک‌های هوش مصنوعی و ابزارهای مدل‌سازی در حوزه کسب و کار به طور فزاینده‌ای در حال افزایش است. در این راستا سیستم‌های خبره جایگاه ویژه‌ای یافته‌اند. در چند دهه گذشته دو عنوان شبکه‌های عصبی و الگوریتم‌های ژنتیک از موضوعاتی بوده‌اند که توجه بسیاری از دانشگاهیان را به خود جلب کرده‌اند. این دو به عنوان ابزاری نیرومند در حل مسائلی که دیگر توسط متدلوژی‌ها و روش‌های سنتی گذشته قابل حل نبودند، شناخته شده و مورد استفاده قرار گرفته‌اند. این روزها استفاده از آنها به زندگی اجتماعی ما نیز تسری یافته تا جایی که کاربرد آنها در تصمیم‌گیری‌ها نقش حیاتی یافته است. با توجه به اهمیت سود هر سهم و این‌که پیش‌بینی سود هر سهم برای سرمایه‌گذاران خارج از شرکت و هم مدیران وظیفه مهمی محسوب می‌شود، بنابراین انتخاب روش پیش‌بینی، به خودی خود، یک تصمیم مهم برای سرمایه‌گذاران و مدیران به حساب می‌آید. تنوع منابع پیش‌بینی سود سبب توجه به دقت در پیش‌بینی می‌شود. منبعی که به درصد اشتباه کمتری بینجامد یعنی پیش‌بینی بر اساس آن به واقعیت نزدیک‌تر باشد، مطمئن‌تر خواهد بود.

اهمیت پیش‌بینی سود هر سهم

اهمیت EPS واضح و آشکار است. قابلیت دوام یک واحد تجاری بستگی به درآمدی است که آن واحد تجاری می‌تواند به دست آورد. اگر یک واحد تجاری توانایی ایجاد درآمد نداشته باشد، سرانجام به سوی ورشکستگی خواهد رفت. بنابراین تنها راه برای بقای بلندمدت واحد تجاری، کسب درآمد است و سود هر سهم اجازه می‌دهد تا قدرت و توانایی کسب درآمد واحدهای تجاری مختلف با هم مقایسه شوند. پیش‌بینی، یک عنصر کلیدی در تصمیم‌گیری‌های مدیریتی است، بنابراین پیش‌بینی‌های سود هر سهم در سرمایه‌گذاری‌ها از اهمیت ویژه‌ای برخوردار است. اهمیت این پیش‌بینی بستگی به میزان انحرافی دارد که با سود واقعی دارد. هر چه میزان این انحراف کمتر باشد، پیش‌بینی از دقت بیشتری برخوردار است.

اهمیت و ضرورت تحقیق

با توجه با اهمیت پیش‌بینی سود هر سهم در این تحقیق با استفاده از مدل‌های خطی و غیرخطی پیش‌بینی برای سود هر سهم را با هم مقایسه نموده و بهترین مدل که دقت بالاتری دارد را انتخاب می‌نماییم.

پیشینه تحقیق

انتخاب متد پیش‌بینی به نوبه خود یک تصمیم مهم برای سرمایه‌گذاران و مدیران محسوب می‌شود. تحقیقات گذشته در مورد دقت متدهای پیش‌بینی در ۲ طبقه قرار داشتند:

۱) مقایسه متدهای مختلف آماری یا متدهای پیش‌بینی مکانیکال و

۲) مقایسه متدهای آماری با پیش‌بینی‌های قضاوتی تحلیلگران.

در اینجا به چند نمونه از تحقیقاتی که با شبکه‌های عصبی انجام گرفته به صورت مختصر اشاره خواهد شد.

ژانگ[۱] و همکارانش (۲۰۰۴)، به بررسی دقت رویکردهای غیرخطی برای پیش‌بینی EPS فصلی پرداختند. برای این منظور، مقایسه بین متدهای خطی و مدل‌های شبکه عصبی را مورد بررسی قرار دادند. نمونه آنها شامل ۲۸۳ شرکت در ۴۱ صنعت مختلف طی دوره زمانی ۲۰۰۲-۱۹۹۲ بود. آنها در پژوهش خود از مدل‌های چندمتغیره که شامل متغیرهای بنیادی حسابداری می‌شد و هم مدل‌های تک‌متغیره استفاده نمودند. نتایج تحقیق آنها نشان داد که مدل شبکه عصبی در پیش‌بینی سود هر سهم عملکرد بهتری نسبت به مدل‌های خطی داشتند. کالن و همکارانش[۲] (۱۹۹۶) بر روی ۲۹۶ شرکت تجاری نیویورک تحقیقی مبنی بر مقایسه توانایی پیش‌بینی مدل‌های ARIMA و مدل شبکه‌های عصبی انجام دادند. آنها دریافتند که مدل‌های خطی پیش‌بینیEPS بادقت‌تری نسبت به مدل‌های شبکه عصبی فراهم می‌کند. مطالعه آنها نشان داد که شبکه‌های عصبی تک‌متغیره لزوما نسبت به مدل‌های خطی برتری ندارند، حتی زمانی که داده‌ها مالی، فصلی و غیرخطی باشند. کینگ کاو و مارک پری[۳] (۲۰۰۹)، با استفاده از داده‌های تحقیق ژانگ و همکاران نیز دقت پیش‌بینی سود هر سهم را با استفاده از مدل شبکه‌هایی عصبی مورد بررسی قرار دادند. نتایج تحقیق آنها نشان داد که مدل شبکه عصبی که وزن آنها با الگوریتم ژنتیک تخمین زده می‌شود، دقت بیشتری نسبت به مدل‌های دیگر دارد.

کینگ کاو و کیوی گان (۲۰۱۰)، تحقیقی بر روی ۷۲۳ شرکت پذیرفته‌شده در بورس اوراق بهادار چین در ۲۲ صنعت مختلف و برای یک دوره ۱۰ ساله انجام دادند. در این پژوهش از مدل شبکه‌های عصبی برای پیش‌بینی سود هر سهم استفاده شد. در این پژوهش آنها به مقایسه بین الگوریتم‌های ژنتیک و الگوریتم پس‌انتشار خطا پرداختند. نتایج حاصل از تحقیق نشان داد، مدل شبکه‌های عصبی که وزن‌های آن با الگوریتم ژنتیک تخمین زده شده موفق‌تر از شبکه عصبی است که وزن‌های آن با الگوریتم پس‌انتشار خطا برآورد شده است. همچنین نتایج فوق نشان داد که افزودن متغیرهای بنیادی حسابداری در مدل شبکه‌های عصبی قدرت و دقت پیش‌بینی را افزایش می‌دهد.

شبکه‌های عصبی

یک شبکه عصبی مصنوعی ایده‌ای است برای پردازش اطلاعات که از سیستم عصبی زیستی الهام گرفته شده و مانند مغز به پردازش اطلاعات می‌پردازد. شبکه‌های عصبی را می‌توان با اغماض زیاد، مدل‌های الکترونیکی از ساختار عصبی مغز انسان نامید. مکانیسم فراگیری و آموزش مغز اساساً بر تجربه استوار است. مدل‌های الکترونیکی شبکه‌های عصبی طبیعی نیز بر اساس همین الگو بنا شده‌اند و روش برخورد چنین مدل‌هایی با مسائل، با روش‌های محاسباتی که به طور معمول توسط سیستم‌های کامپیوتری در

پیش گرفته شده‌اند، تفاوت دارد. این شبکه‌ها از تعداد زیادی عناصر پردازشی (PE) تشکیل شده‌اند که به صورت هماهنگ با یکدیگر عمل کرده و با پردازش داده‌ها یا قوانین نهفته در آنها به ساختار شبکه منتقل می‌نمایند. (موقر، بهزاد. ۱۳۸۸. ص۱۷-۱۶)

الگوریتم آموزش

فرایند آموزش برای پیش‌بینی رفتار مورد انتظار شبکه به ورودی شبکه و خروجی مورد انتظار شبکه نیاز دارد. در طول فرایند آموزش وزن‌ها و بایاس‌ها تنظیم می‌شوند تا تابع کارایی شبکه که برای شبکه‌های پیشخور به صورت پیش فرض MSE است، حداقل شود. الگوریتم‌های آموزشی مختلفی برای شبکه وجود دارند. تمامی این توابع از شیب تابع کارایی برای تنظیم وزن‌ها و بایاس‌ها استفاده می‌کنند. در ساده‌ترین پیاده‌سازی یادگیری، وزن‌ها و بایاس‌ها در جهتی که تابع کارایی کاهش می‌یابد به روز می‌شوند. در این پژوهش برای یادگیری شبکه از دو نوع الگوریتم متداول یعنی الگوریتم پس‌انتشار خطا و الگوریتم ژنتیک استفاده می‌شود.

الگوریتم پس‌انتشار خطا [۴]

این الگوریتم که در سال ۱۹۸۶ توسط روملهارت و مک کلیلاند پیشنهاد گردید، در شبکه‌های عصبی پیشخور مورد استفاده قرار می‌گیرد. واژه پس‌انتشار بدین معنا است که خطاها به سمت عقب در شبکه تغذیه می‌شوند تا وزن‌ها را اصلاح کنند و پس از آن، مجدداً ورودی مسیر پیش سوی خود تا خروجی را تکرار کند. روش پس‌انتشار خطا از روش‌های با سرپرست است. به این مفهوم که نمونه‌های ورودی برچسب خورده‌اند و خروجی مورد انتظار هر یک از آنها از پیش دانسته است. لذا خروجی شبکه با این خروجی‌های ایده‌آل مقایسه شده و خطای شبکه محاسبه می‌گردد. در این الگوریتم ابتدا فرض بر این است که وزن‌های شبکه به طور تصادفی انتخاب شده‌اند. در هر گام خروجی شبکه محاسبه شده و بر حسب میزان اختلاف آن با خروجی مطلوب، وزن‌ها تصحیح می‌گردند تا در نهایت این خطا، مینیمم شود. الگوریتم پس‌انتشار خطا به الگوریتم تکراری معروف است، لذا اوزان شبکه به گونه‌ای تعدیل می‌شوند که خطای بین مقادیر هدف و مقادیر تولید شده توسط شبکه حداقل گردد. بنابراین وقتی اوزان با هر تکرار تغییر می‌کند گفته می‌شود شبکه یاد می‌گیرد. (Principe & Euliano، ۲۰۰۲)

الگوریتم ژنتیک

الگوریتم‌های ژنتیک از اصول انتخاب طبیعی داروین برای یافتن فرمول بهینه جهت پیش‌بینی یا تطبیق الگو استفاده می‌کنند. الگوریتم‌های ژنتیک اغلب گزینه خوبی برای تکنیک‌های پیش‌بینی بر مبنای رگرسیون هستند. مختصرا گفته می‌شود که الگوریتم ژنتیک یک تکنیک برنامه‌نویسی است که از

تکامل ژنتیکی به عنوان یک الگوی حل مسئله استفاده می‌کند. مسئله‌ای که باید حل شود ورودی است و راه‌حل‌ها طبق یک الگو کدگذاری می‌شـوند که تابع fitness نام دارد و هر راه حل کاندید را ارزیابی می‌کند که اکثر آنها به صورت تصادفی انتخاب می‌شوند. الگوریتم ژنتیک (GA) یک تکنیک جستجو در علم رایانه برای یافتن راه حل بهینه و مسائل جستجو است. الگوریتم‌های ژنتیک یکی از انواع الگوریتم‌های تکاملی‌اند که از علم زیست‌شناسی مثل وراثت، جهش، انتخاب طبیعی و ترکیب الهام گرفته شده است، (Beasly ,etal. 1993 ، pp.58-69).

پرسش‌های پژوهش

با توجه به ادبیات و نتایج تحقیقات گذشته پرسش‌های زیر مطرح می‌شود:

۱. آیا مدل شبکه عصبی در پیش‌بینی EPS دقت بالاتری نسبت به مدل‌های خطی دارد؟

۲. آیا الگوریتم ژنتیک نسبت به الگوریتم پس‌انتشار خطا دقت بالاتری در پیش‌بینی EPS دارد؟

فرضیه‌های پژوهش

در این تحقیق به منظور پاسخگویی به سوالات تحقیق فرضیه‌های زیر تدوین می‌گردد:

فرضیه ۱: مدل شبکه‌های عصبی در پیش‌بینی EPS دقت بالاتری نسبت به مدل‌های خطی دارند.

فرضیه ۲: الگوریتم ژنتیک نسبت به الگوریتم پس‌انتشار خطا دقت بالاتری در پیش‌بینی EPS دارد.

برای آزمون این فرضیه‌ها نیازمند مقایسه بین مدل‌های خطی و غیرخطی یک متغیره و مدل‌های خطی و غیرخطی چندمتغیره هستیم، لذا فرضیه‌های تحقیق به فرضیه‌های زیر تقسیم می‌شوند.

فرضیه ۱-۱: مدل‌های غیرخطی یک متغیره در پیش‌بینی EPS دقت بالاتری نسبت به مدل‌های خطی یک متغیره دارند.

فرضیه ۱-۲: مدل‌های غیرخطی چندمتغیره در پیش‌بینی EPS دقت بالاتری نسبت به مدل‌های خطی چندمتغیره دارند.

فرضیه۲-۱: پیش‌بینی با شبکه‌های عصبی تک‌متغیره تخمین زده شده با الگوریتم ژنتیک دقت بالاتری نسبت به پیش‌بینی شبکه‌های عصبی تک‌متغیره تخمین زده شده با الگوریتم پس‌انتشار خطا، دارند.

فرضیه۲-۲: پیش‌بینی با شبکه‌های عصبی چندمتغیره تخمین زده شده با الگوریتم ژنتیک دقت بالاتری نسبت به پیش‌بینی شبکه‌های عصبی چندمتغیره تخمین زده شده با الگوریتم پس‌انتشار خطا، دارند.

متغیرهای پژوهش

سود هر سهم به عنوان متغیر وابسته و موجودی مواد و کالا، حساب‌های دریافتنی، مخارج سرمایه‌ای، هزینه‌های توزیع و فروش، سـود ناخالص، نرخ موثر مالیاتـی و بهره‌وری نیروی کار به عنوان متغیرهای مستقل در نظر گرفته شدند.

جامعه آماری و حجم نمونه

جامعه آماری این تحقیق تمام شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران طی سال‌های ۱۳۸۹-۱۳۸۰ است، که دارای ویژگی‌های زیر می‌باشند:

از سال ۱۳۸۰ صورت‌های مالی میان‌دوره‌ای منتشر نموده باشد. سال مالی آنها منتهی به ۲۹ اسفندماه باشد. دسترسی به اطلاعات وجود داشته باشد. مقدار سود هر سهم آنها صفر نباشد.

با توجه به توضیحات و محدودیت‌های فوق تعداد ۹۱ شرکت در جامعه آماری این پژوهش قرار گرفتند. همچنین با توجه به در دسترس نبودن اطلاعات میان دوره‌ای در بازه زمانی ۱۳۸۲-۱۳۸۰، بازه زمانی مورد استفاده در این پژوهش از سال ۱۳۸۹-۱۳۸۳ می‌باشد.

روش تجزیه و تحلیل اطلاعات

در این تحقیق داده‌ها از دو روش مورد تجزیه و تحلیل قرار گرفته، سپس میانگین قدر مطلق درصد خطاها (MAPE) در هر دو روش محاسبه و مقایسه می‌گردد. روشی که دارای MAPE کمتری باشد روش مناسب‌تری خواهد بود.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right|$$

در روش اول از طریق روش رگرسیونی (مدل خطی) و با استفاده از نرم افزار Eviews و در روش دوم از طریق شبکه عصبی و با استفاده از نرم افزار Neuro Solutions تجزیه و تحلیل صورت می‌پذیرد.

روش تجزیه و تحلیل داده‌ها- رگرسیون

آزمون فرضیه تحقیق

در ذیل مدل‌های مربوط به روش‌های خطی یک‌متغیره و چندمتغیره که در این تحقیق مورد استفاده قرار گرفته است ارائه می‌گردد. مدل خطی تک‌متغیره برگرفته از مدل فاستر (۱۹۷۷) می‌باشد. این مدل در تحقیقات انجام شده توسط کالن و همکاران (۱۹۹۶)، ژانگ و همکاران (۲۰۰۴) و کینگ کاو و مارک پری (۲۰۰۹) نیز به کار گرفته شد. این مدل عبارت است از:

$$(Y_t - Y_{t-4}) = \alpha + \beta(Y_{t-1} - Y_{t-5}) + \varepsilon_t \quad (1)$$

مدل تک‌متغیره دیگر استفاده شده در این تحقیق عبارت است از مدل OLS ساده زیر:

$$E(Y_t) = b_0 + b_1 Y_{t-1} + b_2 Y_{t-4} + \delta \quad (2)$$

از طرف دیگر مدل‌های چندمتغیره که آبرانیل و بوشی (۱۹۹۷)، لورک و ویلینگر (۱۹۹۶)، کالن و

همکاران (۱۹۹۶)، ژانگ و همکاران (۲۰۰۴) و کینگ کاو و مارک پری (۲۰۰۹) در تحقیقات خود استفاده نمودند و نیر در این تحقیق استفاده شده است عبارتند از:

(۳)

$$MLM1: E(Y_t) = \alpha + b_1 Y_{t-1} + b_2 Y_{t-4} + b_3 INV_{t-1} + b_4 AR_{t-1} + b_5 CAPX_{t-1} + b_6 GM_{t-1} + b_7 SA_{t-1} + b_8 ETR_{t-1} + b_9 LFP_{t-1}$$

(۴)

$$MLM2: E(Y_t) = \alpha + b_1 Y_{t-1} + b_2 Y_{t-4} + b_3 INV_{t-4} + b_4 AR_{t-4} + b_5 CAPX_{t-4} + b_6 GM_{t-4} + b_7 SA_{t-4} + b_8 ETR_{t-4} + b_9 LFP_{t-4}$$

CAPX = مخارج سرمایه‌ای

δ = عامل / متغیر ثابت

GM = سود ناخالص

Y_t = سود هر سهم فصلی

SA= هزینه‌های توزیع و فروش

INV= موجودی مواد و کالا

ETR = نرخ موثر مالیاتی

AR= حساب‌های دریافتنی

LFP= لگاریتم بهره‌وری نیروی کار

بررسی معادله (۱)

$$ULM1: (Y_t - Y_{t-4}) = \alpha + \beta(Y_{t-1} - Y_{t-5}) + \varepsilon_t$$

با توجه به این‌که در معادله فوق الگوی تلفیقی (Pool) پذیرفته‌شده نتایج به شرح جدول زیر می‌باشد.

جدول ۱: نتایج حاصل از تخمین مدل ULM۱

R-squared	۰.۲۰۲۶۹۴	تعداد مشاهدات	۲۴۵۷
Adjusted R-squared	۰.۲۰۲۳۶۹	F-statistic	۶۲۴.۱۱۸۳
Durbin-Watson stat	۲.۰۳۶	Prob (F-statistic)	۰.۰۰۰۰۰

نام متغیر	B	Std. Error	مقدار t	Prob (t)
عرض از مبدا	-۳۰.۳۵۰۴۸	۹.۹۸۶۶۱۲	-۳.۰۳۹۱۱۷	۰.۰۰۲۴
$Y_{t-1} - Y_{t-5}$	۰.۴۵۰۴۹۰	۰.۰۱۸۰۳۲	۲۴.۹۸۲۳۶	۰.۰۰۰۰

با توجه به مقدار Prob (F) که کمتر از ۰.۰۵ است، معناداری کل رگرسیون در سطح ۹۵ درصد تایید شد. همچنین در بررسی مقدار Prob (t) به منظور تعیین همبستگی متغیرهای مستقل با وابسته، مشاهده شد که متغیر دارای مقدار کوچک‌تر از ۰.۰۵ بود، لذا ضریب این متغیر معنادار بوده. با استفاده از نتایج فوق داده‌ها برای سال ۸۹ پیش‌بینی گردید. خطای پیش‌بینی با استفاده از معیارهای MAPE و MSE به شرح جدول زیر می‌باشد.

جدول ۲: مقدار خطای پیش‌بینی مدل ULM۱

مدل	فصل اول	فصل دوم	فصل سوم	فصل چهارم	سالانه
ULM۱	MAPE MSE	MAPE MSE	MAPE MSE	MAPE MSE	MAPE MSE
	۰.۵۷۲ ۰.۵۱۵	۰.۵۹۲ ۰.۶۵۴	۰.۵۲۹ ۰.۶۹۷	۰.۵۰۹ ۰.۷۷۸	۰.۵۷۰ ۰.۸۲۳
تعدادمشاهدات	۹۱	۹۱	۹۱	۹۱	۳۶۴

بررسی معادله (۲)

$$ULM2: E(Y_t) = b_0 + b_1 Y_{t-1} + b_2 Y_{t-4} + \delta$$

پس از برآورد معادله فوق مشاهده گردید آماره D-W ۱.۷۱۱۲۳۵ می‌باشد لذا با توجه به پایین بودن آماره D-W و وجود مشکل خود همبستگی، برای رفع آن از وقفه پنجم متغیر وابسته استفاده نموده‌ایم مشاهده نمودیم که مقداره آماره D-W به ۲ رسید و مشکل خود همبستگی رفع گردید. فرم تعدیل شده معادله (۲) به صورت ذیل می‌باشد.

$$RULM2: E(Y_t) = b_0 + b_1 Y_{t-1} + b_2 Y_{t-4} + b_3 Y_{t-5} + \delta$$

با توجه به این‌که در معادله فوق الگوی اثرات ثابت پذیرفته‌شده نتایج به شرح جدول زیر می‌باشد.

جدول ۳: نتایج حاصل از تخمین مدل RULM۲

R-squared	۰,۷۰۱۹۶۱	تعداد مشاهدات	۲۵۴۷
Adjusted R-squared	۰,۶۹۰۲۳۱	F-statistic	۵۹,۸۴۴۰۷
Durbin-Watson stat	۲,۰۴۴۴۳۴	Prob (F-statistic)	۰,۰۰۰۰۰

نام متغیر	B	Std. Error	مقدار t	Prob (t)
عرض از مبدا	۱۵۱,۰۳۱۹	۱۴,۷۰۷۸۳	۱۰,۲۶۸۸۱	۰,۰۰۰۰
Y_{t-1}	۰,۴۰۰۱۲۹	۰,۰۱۸۷۹۹	۲۱,۲۸۴۵۱	۰,۰۰۰۰
Y_{t-4}	۰,۰,۵۶۷۴۸۲	۰,۱۶۸۸۱	۳۳,۶۱۶۰۴	۰,۰۰۰۰
Y_{t-5}	-۰,۲۶۹۲۶۸	۰,۰۱۹۹۹۴	-۱۳,۴۶۷۳۲	۰,۰۰۰۰

با توجه به مقدار Prob (F) که کمتر از ۰.۰۵ است، معناداری کل رگرسیون در سطح ۹۵ درصد تایید شـد. همچنین در بررسـی مقدار Prob (t) به منظور تعیین همبسـتگی متغیرهای مستقل با وابسـته، مشاهده شد که متغیر و دارای مقدار کوچک‌تر از ۰.۰۵ بود، لذا ضریب این متغیرها معنادار بوده. با اسـتفاده از نتایج فوق و معادله مربوطه، داده‌ها برای سـال ۱۳۸۹ پیش‌بینی گردید. خطای پیش‌بینی با استفاده از معیارهای MAPE و MSE به شرح جدول ۴ می‌باشد.

جدول ۴: مقدار خطای پیش بینی از مدل RULM۲

مدل	فصل اول	فصل دوم	فصل سوم	فصل چهارم	سالانه
RULM۲	MAPE MSE	MAPE MSE	MAPE MSE	MAPE MSE	MAPE MSE
	۰.۵۹۴ ۰.۵۰۲	۰.۵۰۸ ۰.۵۱۸	۰.۵۰۰ ۰.۵۰۴	۰.۵۱۱ ۰.۴۵۰	۰.۵۸۱ ۰.۷۴۵
تعداد مشاهدات	۹۱	۹۱	۹۱	۹۱	۳۶۴

بررسی معادله (۳)

$$MLM1: E(Y_t) = \alpha + b_1 Y_{t-1} + b_2 Y_{t-4} + b_3 INV_{t-1} + b_4 AR_{t-1} + b_5 CAPX_{t-1} + b_6 GM_{t-1} + b_7 SA_{t-1} + b_8 ETR_{t-1} + b_9 LFP_{t-1}$$

پس از برآورد معادله فوق مشـاهده گردید آماره D-W ۱.۷۱۸۲۱۶ می‌باشـد لذا با توجه به پایین بودن آماره D-W و وجود مشکل خود همبستگی، برای رفع آن از وقفه پنجم متغیر وابسته استفاده نموده‌ایم و مشاهده نمودیم که مقداره آماره D-W به ۲ رسید و مشکل خود همبستگی رفع گردید. فرم تعدیل شده معادله (۳) به صورت ذیل می‌باشد.

جدول ۵: نتایج حاصل از تخمین مدل RMLM۱

R-squared	۰,۷۰۴۱۰۶	تعداد مشاهدات	۲۴۵۷
Adjusted R-squared	۰,۶۹۱۵۴۷	F-statistic	۵۶,۰۶۳۰۳
Durbin-Watson stat	۲,۰۴۷۷۳۴	Prob (F-statistic)	۰,۰۰۰۰

Prob (t)	مقدار t	Std. Error	B	نام متغیر
۰,۰۲۵۵	۲,۲۳۵۱۷۰	۸۷,۱۴۰۸۱	۱۹۴,۷۷۴۶	عرض از مبدا
۰,۰۰۰۰	۲۰,۱۸۱۱۹	۰,۰۱۹۸۸۹	۰,۴۰۱۳۹۲	Y_{t-1}
۰,۰۰۰۰	۳۲,۵۱۴۳۷	۰,۰۱۷۱۳۵	۰,۵۵۷۱۴۶	Y_{t-4}
۰,۰۰۰۰	-۱۳,۰۵۴۶۶	۰,۰۲۰۱۸۹	-۰,۲۶۳۵۶۷	Y_{t-5}

INV_{t-1}	-۰,۰۰۰۰۵۹۰	۰,۰۰۰۰۸۶۸	-۰,۶۷۹۹۳۵	۰,۴۹۶۶
AR_{t-1}	۰,۰۰۰۰۸۰۵	۰,۰۰۰۰۴۸۶	۱,۶۵۵۱۵۲	۰,۰۹۸۰
CAP_{t-1}	-۰,۰۰۰۰۶۴۲	۰,۰۰۰۰۵۷۸	-۱,۱۱۰۳۰۰	۰,۲۶۷۰
GM_{t-1}	۰,۰۰۰۰۴۸۳	۰,۰۰۰۰۸۱۱	۰,۵۹۵۵۶۴	۰,۵۵۱۵
SA_{t-1}	-۰,۰۰۰۱۰۲۶	۰,۰۰۰۴۶۸	-۲,۱۹۲۸۸۴	۰,۰۲۸۴
ETR_{t-1}	۲۲۰,۶۶۰۱	۹۹,۹۸۹۴۸	۲,۲۰۶۸۳۳	۰,۰۲۷۴
LFP_{t-1}	-۱۵,۹۵۳۸۷	۳۵,۷۶۶۲۵	-۰,۴۴۶۰۵۹	۰,۶۵۵۶

با توجه به مقدار Prob (F) که کمتر از ۰.۰۵ است، معناداری کل رگرسیون در سطح ۹۵ درصد تایید شد. همچنین در بررسی مقدار Prob (t)، مشاهده شد که متغیر Y_{t-4} ، Y_{t-1} و Y_{t-5} و SA_{t-1} و ETR_{t-1} معنادار بوده و بر سود هر سهم اثرگذار هستند. و متغیرهای بی‌معنا باید از مدل خارج گردند. جدول نهایی پس از حذف کلیه متغیرهای بی‌معنا به صورت زیر می‌باشد.

جدول ۶: نتایج حاصل از مدل RMLM۱ پس از خروج متغیر LFP_{t-1} و GM_{t-1} و INV_{t-1} و CAP_{t-1} و AR_{t-1}

R-squared	۰,۷۰۳۵۵۰	تعداد مشاهدات	۲۴۵۷
Adjusted R-squared	۰,۶۹۱۶۲۲	F-statistic	۵۸,۹۸۱۶۳
Durbin-Watson stat	۲,۰۴۷۴۹۹	Prob (F-statistic)	۰,۰۰۰۰

نام متغیر	B	Std. Error	مقدار t	Prob (t)
عرض از مبدا	۱۴۴,۹۶۶۶	۲۰,۳۰۹۱۲	۷,۱۳۸۰۰۴	۰,۰۰۰۰
Y_{t-1}	۰,۴۰۳۰۵۳	۰,۰۱۸۹۰۴	۲۱,۳۲۱۴۶	۰,۰۰۰۰
Y_{t-4}	۰,۵۵۹۳۵۵	۰,۰۱۶۹۹۹	۳۲,۹۰۵۹۹	۰,۰۰۰۰
Y_{t-5}	-۰,۲۶۴۴۰	۰,۰۲۰۰۰۹	-۱۳,۱۹۴۵۸	۰,۰۰۰۰
SA_{t-1}	-۰,۰۰۱۰۳۲	۰,۰۰۰۳۷۹	-۲,۷۲۲۴۶۹	۰,۰۰۶۵
ETR_{t-1}	۲۲۵,۸۴۴۸	۹۹,۷۴۷۵۶	۲,۲۶۴۱۶۳	۰,۰۲۳۷

با حذف متغیرهای بی‌معنا مربوط به معادله فوق، سود هر سهم به ۵ متغیر وقفه اول و چهارم و پنجم متغیر وابسته و هزینه‌های توزیع و فروش و نرخ موثر مالیاتی با وقفه اول ارتباط معنادار پیدا کرد. با استفاده از نتایج فوق و معادله مربوطه، داده‌ها برای سال ۱۳۸۹ پیش‌بینی گردید.

جدول ۷: مقدار خطای پیش بینی مدل RMLM۱					
مدل	فصل اول	فصل دوم	فصل سوم	فصل چهارم	سالانه
RMLM۱	MAPE MSE	MAPE MSE	MAPE MSE	MAPE MSE	MAPE MSE
	۰.۶۲۴ ۰.۵۲۲	۰.۵۴۰ ۰.۵۲۶	۰.۵۶۳ ۰.۵۷۰	۰.۵۵۶ ۰.۵۱۱	۰.۵۹۲ ۰.۷۵۰
تعداد مشاهدات	۹۱	۹۱	۹۱	۹۱	۳۶۴

بررسی معادله (۴)

$$MLM2: E(Y_t) = \alpha + b_1 Y_{t-1} + b_2 Y_{t-4} + b_3 INV_{t-4} + b_4 AR_{t-4} + b_5 CAPX_{t-4} + b_6 GM_{t-4} + b_7 SA_{t-4} + b_8 ETR_{t-4} + b_9 LFP_{t-4}$$

پس از برآورد معادله فوق مشـاهده گردید آماره D-W ۱.۷۳۵۹۱۴ می‌باشد لذا برای رفع مشکل خود همبسـتگی از وقفه پنجم متغیر وابسـته استفاده نموده‌ایم و مشـاهده نمودیم که مقداره آماره D-W به ۲ رسـید و مشـکل خود همبستگی رفع گردید. فرم تعدیل شده معادله (۴) به صورت ذیل می‌باشد.

$$RMLM2: E(Y_t) = \alpha + b_1 Y_{t-1} + b_2 Y_{t-4} + b_3 Y_{t-5} + b_3 INV_{t-4} + b_4 AR_{t-4} + b_5 CAPX_{t-4} + b_6 GM_{t-4} + b_7 SA_{t-4} + b_8 ETR_{t-4} + b_9 LFP_{t-4}$$

جدول ۸: نتایج حاصل از مدل RMLM۲			
R-squared	۰,۷۰۷۵۶۸	تعداد مشاهدات	۲۴۵۷
Adjusted R-squared	۰,۶۹۵۱۵۶	F-statistic	۵۷,۰۰۵۷۰
Durbin-Watson stat	۲,۰۲۷۸۹۳	Prob (F-statistic)	۰,۰۰۰۰

نام متغیر	B	Std. Error	مقدار t	Prob (t)
عرض از مبدا	-۱۹۴,۶۵۵۲	۸۹,۸۰۷۹۵	-۲,۱۶۷۴۶۰	۰,۰۳۰۳
Y_{t-1}	۰,۴۰۹۶۶۷	۰,۰۱۸۷۵۲	۲۱,۸۴۶۰۷	۰,۰۰۰۰
Y_{t-4}	۰,۵۱۱۵۸۰	۰,۰۱۸۹۳۴	۲۷,۰۱۹۳۶	۰,۰۰۰۰
Y_{t-5}	-۰,۲۴۸۲۲۰	۰,۰۲۰۲۱۶	-۱۲,۲۷۸۶۸	۰,۰۰۰۰
INV_{t-4}	-۰,۰۰۰۱۶۷	۰,۰۰۰۰۸۹۴	-۱,۸۶۹۸۱۱	۰,۰۶۱۶
AR_{t-4}	-۰,۰۰۰۰۱۵۹	۰,۰۰۰۰۵۵۲	-۰,۲۸۸۹۶۷	۰,۷۷۲۶
CAP_{t-4}	-۰,۰۰۰۱۴۴	۰,۰۰۰۰۵۶۴	-۲,۵۵۵۴۵۸	۰,۰۱۰۷

GM_{t-4}	۰,۰۰۰۱۷۲	۰,۰۰۰۰۸۹۴	۱,۹۲۱۵۶۳	۰,۰۵۴۸
SA_{t-4}	۰,۰۰۱۱۲۱	۰,۰۰۰۵۰۳	۲,۲۲۹۴۷۹	۰,۰۲۵۹
ETR_{t-4}	-۱۰,۸۵۳۱۱	۹۹,۰۷۵۶۶	-۰,۱۰۹۵۴۴	۰,۹۱۲۸
LFP_{t-4}	۱۴۸,۰۲۳۴	۳۶,۵۳۲۲۰	۴,۰۵۱۸۶۲	۰,۰۰۰۱

با توجه به مقدار Prob (F) که کمتر از ۰.۰۵ است، معناداری کل رگرسیون در سطح ۹۵ درصد تایید شد. همچنین در بررسی مقدار Prob (t)، مشاهده شد که متغیر Y_{t-4} و Y_{t-1} و Y_{t-5} و SA_{t-4} و CAP_{t-4} و GM_{t-4} و LFP_{t-4} معنادار بوده و بر سود هر سهم اثرگذار هستند و مابقی متغیرها اثری بر سود هر سهم ندارند و باید از مدل خارج گردند. جدول نهایی پس از حذف کلیه متغیرهای بی‌معنا به صورت زیر می‌باشد.

جدول ۹: نتایج حاصل از مدل RMLM۲ پس از خروج متغیر ETR_{t-4} و AR_{t-4}

R-squared	۰,۷۰۷۵۵۶	تعداد مشاهدات	۲۴۵۷
Adjusted R-squared	۰,۶۹۵۴۰۲	F-statistic	۵۸,۲۱۵۰۶
Durbin-Watson stat	۲,۰۲۷۷۹۵	Prob (F-statistic)	۰,۰۰۰۰

نام متغیر	B	Std. Error	مقدار t	Prob (t)
عرض از مبدا	-۱۹۴,۴۶۰۱	۸۸,۶۷۵۷۷	-۲,۱۹۲۹۳۴	۰,۰۲۸۴
V_{t-1}	۰,۴۰۹۴۹۲	۰,۰۱۸۷۲۴	۲۱,۸۷۰۰۸	۰,۰۰۰۰
V_{t-4}	۰,۵۱۱۸۲۳	۰,۱۸۸۷۸	۲۷,۱۱۱۴۵	۰,۰۰۰۰
V_{t-5}	-۰,۲۴۸۴۲۰	۰,۰۲۰۱۹۷	-۱۲,۲۹۹۷۴	۰,۰۰۰۰
INV_{t-4}	-۰,۰۰۰۱۷۴	۰,۰۰۰۰۸۶۷	-۲,۰۰۱۹۸۸	۰,۰۴۵۴
CAP_{t-4}	-۰,۰۰۰۱۴۶	۰,۰۰۰۰۵۵۹	-۲,۶۱۰۶۹۲	۰,۰۰۹۱
GM_{t-4}	۰,۰۰۰۱۶۹	۰,۰۰۰۰۸۸۸	۱,۸۹۹۰۰۷	۰,۰۵۱۷
SA_{t-4}	۰,۰۰۱۱۰۷	۰,۰۰۰۴۹۹	۲,۲۱۶۵۸۴	۰,۰۲۶۷
LFP_{t-4}	۱۴۷,۳۳۸۵	۳۶,۴۲۸۵۳	۴,۰۴۴۵۹۱	۰,۰۰۰۱

با حذف متغیرهای بی‌معنا مربوط به معادله فوق، سود هر سهم به ۸ متغیر وقفه اول و چهارم و پنجم متغیر وابسته و موجودی مواد و کالا، دارایی‌های سرمایه‌ای، سود ناخالص، هزینه‌های توزیع و فروش و بهروری نیروی کار با وقفه چهارم ارتباط معنادار پیدا کرد. با استفاده از نتایج فوق و معادله مربوطه، داده‌ها برای سال ۱۳۸۹ پیش‌بینی گردید. جدول ۱۰ مقدار خطای پیش‌بینی را نشان می‌دهد.

جدول ۱۰: مقدار خطای پیش بینی مدل RMLM۲

مدل	فصل اول	فصل دوم	فصل سوم	فصل چهارم	سالانه
	MAPE MSE	MAPE MSE	MAPE MSE	MAPE MSE	MAPE MSE
RMLM۲	۰.۶۷۴ ۰.۵۸۵	۰.۶۰۳ ۰.۷۶۱	۰.۵۷۴ ۰.۶۷۲	۰.۵۶۰ ۰.۵۱۶	۰.۶۰۲ ۰.۷۸۰
تعداد مشاهدات	۹۱	۹۱	۹۱	۹۱	۳۶۴

نتایج ۴ مدل خطی ارائه شده در جدول زیر خلاصه گردیده است.

جدول ۱۱: مقایسه دقت پیش بینی EPS برای ۴ مدل خطی

مدل	فصل اول	فصل دوم	فصل سوم	فصل چهارم	سالانه
	MAPE MSE	MAPE MSE	MAPE MSE	MAPE MSE	MAPE MSE
ULM۱	۰,۵۱۵ ۰,۵۷۲	۰,۶۵۴ ۰,۵۰۸	۰,۶۹۷ ۰,۵۰۰	۰,۷۷۸ ۰,۵۰۹	۰,۸۲۳ ۰,۵۷۰
RULM۲	۰,۵۰۲ ۰,۵۹۴	۰,۵۱۸ ۰,۵۹۲	۰,۵۰۴ ۰,۵۲۹	۰,۴۵۰ ۰,۵۱۱	۰,۷۴۵ ۰,۵۸۱
RMLM۱	۰,۵۲۲ ۰,۶۲۴	۰,۵۲۶ ۰,۵۴۰	۰,۵۷۰ ۰,۵۶۳	۰,۵۱۱ ۰,۵۵۶	۰,۷۵۰ ۰,۵۹۲
RMLM۱	۰,۵۸۵ ۰,۶۷۴	۰,۷۶۱ ۰,۶۰۳	۰,۶۷۲ ۰,۵۷۴	۰,۵۱۶ ۰,۵۶۰	۰,۷۸۰ ۰,۶۰۲

با توجه به نتایج فوق و مقدار آماره MAPE مشاهده شد مدل خطی ULM۱ پایین‌ترین مقدار خطا را دارد (۰.۵۷۰، سالانه). بررسی مقدار خطای مربوط به هر فصل در این مدل نتیجه فوق را نیز تایید می‌کند. همچنین نتایج جدول فوق نشان می‌دهد که دخالت دادن متغیرهای بنیادی حسابداری اثر منفی بر دقت پیش‌بینی مدل‌های رگرسیون خطی دارند. این نتیجه مطابق با نتیجه تحقیق ژانگ و همکاران (۲۰۰۴) و کینگ کاو و مارک پری (۲۰۰۹)، می‌باشد، آنها استنباط نمودند که متغیرهای بنیادی حسابداری رابطه غیرخطی با EPS دارند.

روش تجزیه و تحلیل دادها – شبکه‌های عصبی

در طراحی ساختار و معماری شبکه عصبی، تعداد عناصر بردار ورودی با انتخاب طراح نیست و از صورت مسئله مورد بررسی مشخص شده ولی تعیین تعداد لایه‌های پنهان، تعداد نرون‌ها، نوع تابع فعالسازی و تعداد تکرارها با آزمون و خطا از طریق طراح انتخاب می‌شود. در این تحقیق از شبکه عصبی پیشخور تعمیم یافته[۵] استفاده می‌شود. این شبکه با یک لایه ورودی، یک لایه خروجی و یک لایه پنهان که هر کدام شامل گره‌های ورودی، گره‌های خروجی و گره‌های پنهان می‌باشد، جهت طراحی شبکه عصبی مورد استفاده واقع شده است. گره‌های ورودی در حقیقت همان متغیرهای مستقل تحقیق می‌باشد. تعداد گره‌های پنهان مشخص نبوده، لذا گره‌های پنهان برای مدل‌های چند متغیر از تعداد ۵ گره تا ۲۰ گره و برای مدل‌های تک‌متغیره از ۳۳ گره تا ۴۳ گره مورد آزمون قرار گرفتند. تعداد دفعاتی که مرحله آموزش تکرار می‌شود[۶]، نیز به صورت پیشفرض سیستم ۱۰۰۰ بوده، لذا شبکه‌های عصبی با learning epochs ۱۰۰۰ بررسی شده است. روش الگوریتم ژنتیک و روش الگوریتم پس‌انتشار خطا جهت آموزش شبکه انتخاب شدند و نتایج این دو الگوریتم برای آزمون فرضیه تحقیق نیز مورد مقایسه قرار گرفتند. تابع تانژانت[۷] برای انتقال اطلاعات در آکسون‌ها نیز در سیستم انتخاب شده است.

مجموعه داده‌های مورد بررسی شامل ۹۱ شرکت با ۲۸ دوره از فصل اول سال ۱۳۸۳ تا فصل چهارم سال ۱۳۸۹ تشکیل شده است. این مجموعه داده‌ها به دو گروه آموزش و پیش‌بینی تقسیم شده‌اند. از دوره اول سال ۱۳۸۹ تا دوره چهارم سال ۱۳۸۹ به عنوان داده‌های گروه پیش‌بینی انتخاب گردیدند و برای مقایسه بین مدل‌های مختلف به کار برده می‌شوند. با توجه به مطالب بیان شده، تعداد ۱۶ شبکه با استفاده از داده‌های آموزش برای مدل‌های چندمتغیره و تعداد ۱۱ شبکه برای مدل‌های تک‌متغیره طراحی گردید. و پس از آن بهترین شبکه برای پیش‌بینی سود هر سهم با توجه به معیار خطای MAPE انتخاب گردید. نتایج در جداول ۱۲ و ۱۳ نشان داده شده است.

با توجه به جداول فوق، در مدل چندمتغیره، شبکه با ۹ گره در لایه پنهان و آموزش از طریق الگوریتم ژنتیک دارای کمترین خطا می‌باشد (۰.۴۸۷). و در مدل تک‌متغیره، شبکه با تعداد ۳۶ گره در لایه پنهان و آموزش از طریق الگوریتم پس‌انتشار خطا دارای کمترین خطا می‌باشد (۰.۵۷۶). از جداول فوق می‌توان نتیجه گرفت که اضافه نمودن متغیرهای بنیادی حسابداری، قدرت پیش‌بینی شبکه را افزایش می‌دهد. به عبارت دیگر دخالت دادن متغیرهای بنیادی حسابداری اثر مثبت بر دقت پیش‌بینی شبکه‌های عصبی دارند. این نتیجه نیز منطبق با نتایج تحقیق ژانگ و همکاران (۲۰۰۴) و کینگ کاو و مارک پری (۲۰۰۹)، می‌باشد. از طرف دیگر با توجه به نتایج مختلفی که از دو الگوریتم ژنتیک و پس‌انتشار خطا حاصل شد نمی‌توان مقایسه دقیقی بین این دو الگوریتم انجام داد. لذا نمی‌توان در مورد تایید یا رد فرضیه تحقیق قضاوت نمود.

جدول ۱۲: MSE و MAPE هر شبکه در مدل‌های چند متغیر با تعداد گره‌های مختلف

تعدادگره	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۲۰
MAPE BP	۰.۶۴۳	۰.۵۴۰	۰.۵۷۷	۰.۶۰۱	۰.۵۳۵	۰.۵۷۳	۰.۵۰۲	۰.۶۷۵	۰.۵۰۶	۰.۵۷۶	۰.۵۸۵	۰.۵۳۳	۰.۵۰۶	۰.۵۹۷	۰.۵۶۱	۰.۵۵۶
MAPE GA	۰.۴۹۲	۰.۵۵۰	۰.۵۶۷	۰.۵۹۷	۰.۴۸۷	۰.۵۶۴	۰.۵۶۰	۰.۵۵۹	۰.۵۰۲	۰.۵۵۱	۰.۵۸۱	۰.۵۱۸	۰.۴۹۷	۰.۶۰۰	۰.۶۱۷	۰.۵۳۳
MSE BP	۰.۹۶۹	۰.۶۴۷	۰.۸۱۲	۰.۹۳۲	۰.۶۶۳	۰.۷۳۰	۰.۹۵۸	۰.۸۷۸	۰.۹۹۵	۰.۹۵۱	۰.۹۰۳	۰.۶۴۹	۰.۷۹۰	۰.۸۸۵	۰.۸۷۳	۰.۷۳۳
MSE GA	۰.۶۲۳	۰.۶۵۵	۰.۸۴۰	۰.۷۹۲	۰.۵۷۳	۰.۷۲۴	۰.۶۶۰	۰.۸۵۹	۰.۹۸۱	۰.۸۲۹	۰.۹۲۳	۰.۵۴۲	۰.۶۵۴	۰.۸۱۶	۰.۷۷۶	۰.۸۳۲

جدول ۱۳: MSE و MAPE هرشبکه در مدل‌های تک‌متغیره با تعداد گره‌های مختلف

تعدادگره	۳۳	۳۴	۳۵	۳۶	۳۷	۳۸	۳۹	۴۰	۴۱	۴۲	۴۳	۴۴
MAPE BP	۰,۶۲۵	۰,۶۰۲	۰,۶۰۴	۰,۵۷۶	۰,۶۳۳	۰,۶۱۵	۰,۶۰۶	۰,۶۲۳	۰,۶۱۲	۰,۶۲۷	۰,۶۳۲	۰,۶۳۸
MAPE GA	۰,۵۸۷	۰,۵۹۷	۰,۶۱۵	۰,۵۹۱	۰,۵۸۲	۰,۵۹۵	۰,۶۱۷	۰,۶۰۶	۰,۶۰۳	۰,۶۳۳	۰۶۲۱	۰,۶۲۶
MSE BP	۰,۹۲۳	۰,۸۴۱	۰,۹۳۷	۰,۷۷۱	۰,۹۸۹	۰,۹۰۱	۰,۹۶۹	۰,۹۳۰	۰,۸۹۴	۰,۸۲۱	۰,۸۷۱	۰,۹۰۳
MES GA	۰,۸۳۸	۰,۸۰۵	۰,۹۲۹	۰,۷۸۳	۰,۷۶۷	۰,۸۴۶	۰,۹۵۹	۰,۹۴۹	۰,۸۸۱	۰,۸۳۲	۰,۹۰۴	۰,۸۶۱

تعیین پارامترها و عملگرهای الگوریتم ژنتیک

در تعییـــن پارامترهـــا و عملگرها برای الگوریتـــم ژنتیک با توجه به تصادفی بـــودن مقادیر ضرایب وزن اولیه شـــبکه‌ها و مقادیر انتخاب شـــده برای ژن‌ها در اولین نســـل، تصادفی بودن انتخاب محور در عملگرهـــای One Point و Two Point، تصادفـــی بـــودن تعویض ژن‌هـــای والدین در عملگر Uniform و همچنیـــن تصادفی بودن عمل جهش ژن‌ها در هر نســـل، نمی‌تـــوان تنها با یک بار به کارگیری الگوریتم ژنتیک با اســـتفاده از هر یک از این عملگرها، در مورد نتایج قضاوت کرد و همانند آموزش شبکه‌های عصبی که برای رسیدن به بهترین پاسخ باید عمل آموزش را چند مرتبه تکرار کرد در اینجـــا نیز برای رســـیدن به بهترین پارامترها، باید پـــس از انتخاب عملگر مورد نظر چندین مرتبه عمل جستجوی بهترین پارامترها با الگوریتم ژنتیک را تکرار کرد تا به نتایج مناسب دست یافت.

در بین عملگرهای بررســـی شـــده برای عمل ترکیب، نتایج به دســـت آمـــده از عملگر uniform نســـبت به عملگرهای One Point و Two Point و Heuristic کمترین خطا را داشته و از بین عملگرهای انتخاب عملگر Roulette نسبت به عملگرهای Best و Random و Top Percent و Tournament عملکرد بهتری داشت و بهترین نتیجه هنگام استفاده از این عملگر به همراه عملگر Uniform بوده است.

مقایسه بین مدل‌های خطی و شبکه‌های عصبی (غیرخطی)

برای مقایســـه بین مدل‌های خطی و مدل‌های غیرخطی، مقدار آماره خطای آنها را مورد بررســـی قرار می‌دهیم. با توجه به این‌که از بین مدل‌های خطی مدل ULM۱ کمترین مقدار خطا را داشـــته است، مقدار خطای این مدل را با مدل‌های تک‌متغیره و چندمتغیره شبکه‌های عصبی مقایسه نموده و مدلی که کمترین مقدار خطا را داشته باشد به عنوان مدل برگزیده انتخاب می‌نماییم. نتایج مقایسه آنها در جدول زیر ارائه می‌گردد.

جدول ۱۴: مقایسه بین مدل‌های انتخاب شده

مدل	فصل اول		فصل دوم		فصل سوم		فصل چهارم		سالانه	
	MAPE	MSE	MAPE	MSE	MAPE	MSE	MAPE	MSE	MAPE	MSE
ULM۱	۰.۵۷۲	۰.۵۱۵	۰.۵۰۸	۰.۶۵۴	۰.۵۰۰	۰.۶۹۷	۰.۵۰۹	۰.۷۷۸	۰.۵۷۰	۰.۸۲۳
UBP	۰.۵۹۸	۰.۶۳۵	۰.۵۹۰	۰.۸۲۱	۰.۵۵۱	۰.۵۲۲	۰.۵۵۳	۰.۴۴۳	۰.۵۷۶	۰.۷۸۳
UGA	۰.۶۲۳	۰.۷۳۹	۰.۵۸۰	۰.۷۶۱	۰.۵۶۳	۰.۴۴۳	۰.۵۷۱	۰.۴۸۱	۰.۵۹۱	۰.۷۷۱
MBP	۰.۵۹۴	۰.۶۵۲	۰.۵۲۱	۰.۷۳۸	۰.۵۳۱	۰.۵۹۷	۰.۵۱۳	۰.۵۵۱	۰.۵۳۵	۰.۶۶۳
MGA	۰.۴۹۰	۰.۵۴۸	۰.۴۹۸	۰.۶۰۵	۰.۴۳۳	۰.۴۰۴	۰.۴۶۰	۰.۵۶۲	۰.۴۸۷	۰.۵۷۳

با توجه به نتایج به دست آمده از جدول فوق انتخاب مدل‌ها به ترتیب زیر می‌باشد:

۱. مدل چندمتغیره شبکه عصبی با آموزش از طریق الگوریتم ژنتیک کمترین مقدار خطا (۰.۴۸۷، سالانه) را داشته است، بررسی مقدار خطای مربوط به هر فصل در این مدل نتیجه فوق را نیز تایید می‌کند.

۲. مدل خطی ULM۱ با مقدار خطای ۰.۵۷۰ و در نهایت

۳. مدل تک‌متغیره شبکه عصبی با آموزش از طریق الگوریتم پس‌انتشار خطا با مقدار خطای ۰.۵۷۶

نتیجه‌گیری

برای بررسی فرضیه‌های فوق از دو روش رگرسیون خطی و شبکه‌های عصبی و به منظور مقایسه دو روش از آماره MAPE استفاده گردید. آماره فوق نشان داد شبکه‌های عصبی که در آن از متغیرهای بنیادی حسابداری استفاده گردید دقت بالاتری در پیش‌بینی سود هر سهم نسبت به روش‌های تک‌متغیره شبکه عصبی و تک‌متغیره و چندمتغیره روش‌های خطی دارد. به طور کل می‌توان گفت افزودن متغیرهای بنیادی حسابداری دقت پیش‌بینی شبکه‌های عصبی را افزایش می‌دهد از طرف دیگر اضافه نمودن آنها به روش‌های خطی اثر منفی بر دقت پیش‌بینی توسط آنها داشت. این موضوع نشان می‌دهد که متغیرهای بنیادی حسابداری رابطه غیرخطی با EPS دارند. با توجه به توضیحات فوق فرضیه ۱-۲ که بیان می‌دارد مدل‌های غیرخطی چندمتغیره در پیش‌بینی EPS دقت بالاتری نسبت به مدل‌های خطی چندمتغیره دارند، تایید می‌شود. با توجه به نتایج و مقادیر به دست آمده از آماره MAPE برای مدل تک‌متغیره شبکه عصبی و تک‌متغیره روش‌های خطی که به ترتیب برابر است با ۰.۵۷۶ و ۰.۵۷۰ نشان می‌دهد هنگامی که متغیرهای بنیادی حسابداری در مدل‌ها دخالت داده نمی‌شوند، دقت پیش‌بینی مدل‌های تک‌متغیره خطی از مدل تک‌متغیره شبکه‌های عصبی بالاتر است. لذا این نتایج در جهت رد فرضیه ۱-۱ که بیان‌کننده این است که مدل‌های غیرخطی تک‌متغیره در پیش‌بینی EPS دقت بالاتری نسبت به مدل‌های خطی تک‌متغیره دارند، می‌باشد.

در مورد مقایسه بین دو الگوریتم آموزشی ژنتیک و پس‌انتشار خطا با توجه به نتایج متفاوتی که از گره‌های مختلف حاصل شد نمی‌توان مقایسه دقیقی بین این دو الگوریتم انجام داد. لذا نمی‌توان در مورد تایید یا رد فرضیه ۲ تحقیق به صورت کلی قضاوت نمود. اما با توجه به گره‌های انتخاب شده برتر (گره ۹ در شبکه چندمتغیره و گره ۳۶ در شبکه تک‌متغیره) می‌توان گفت فرضیه فرعی ۲-۱ رد و فرضیه فرعی ۲-۲ تایید می‌شود.

منابع:

۱. موقر، بهزاد. (۱۳۸۸). ارزیابی کاربرد شبکه‌های عصبی مصنوعی در پیش‌بینی بازار مبادلات ارز خارجی. پایان‌نامه کارشناسی ارشد، دانشگاه آزاد اسلامی واحد تهران جنوب.

2. Beasley, D, & Bull, R.D, & R. Martin. (1993). An overview of genetic algorithms: part 1, fundamentals, University Computing, vol.15, pp.58-69

3. Callen, J.L, et al. (1996). Neural network forecasting of quarterly accounting earnings, International Journal of Forecasting 12 (4), pp: 475–482.

4. Cao,Q & Mark, M.E. (2009). Neural network earnings per share forecasting models: A comparison of backward propagation and the genetic algorithm. Decision Support Systems 47, pp:32–41.

5. Cao,Q & Gan,Q . (2010). Forecasting EPS of Chinese listed companies using neural network with GA. International journal of Society Systems Science ,vol.2, no.3, pp.207-225.

6. Principe J. C., Euliano N. R., Lefebvre W. C. (2002), Neural and Adaptive Systems: Fundamentals through Simulations, John Wiley & Sons.

7. Zhang, W., & Cao, Q., & Schniederjans, M. (2004). Neural network earnings per share forecasting models: a comparative analysis of alternative methods , Decision Sciences 35 (2) ,pp:205–237.

پاورقی:

1 . Zhang
2 . Callen
3 . Qing Cao & Mark Parry
4 . Backward Propagation
5 . Generalized feedforward
6 . Leaning Epochs
7 . Tanh Axon
8 .Associate Professor of Alzahra University
9 .Associate Professor of Tehran University
10 .M.A Accounting of Alzahra University