

An Accruals Quality Model: A Group Method of Data Handling Approach

Ali Saghafi 

Professor of Accounting Department, Allameh Tabataba'i University, Tehran, Iran

Ghasem Bulu 

Associate Professor of Accounting Department, Allameh Tabataba'i University, Tehran, Iran

HosseinAli Sohrabi
Varzaneh * 

Ph.D. Student in Accounting, Allameh Tabataba'i University, Tehran, Iran.

Abstract

Development of Earnings quality measures, especially Accruals quality measures, has been a critical line of research for more than three decades. Literature indicates that linear-regression-based measures are subject to significant estimation error in non-discretionary accruals estimation. Therefore, recent research has used some machine learning algorithms in order to address the issue. However, being founded on the Black box approach limits future development and applicability of these methods. So, to address the limitations, we have used the Group Method of Data Handling (GMDH) approach, as a White box approach to estimate the accruals. Findings from the use of data from 164 Tehran Securities Exchange listed companies during 1385 to 1397 suggest that GMDH-based models perform superior to regression models and multilayer perceptron neural networks in terms of estimation error measured by mean squared error. Moreover, the Cash flow approach in total accruals calculation leads to less estimation error compared to the balance sheet approach. As a result, the model developed in this study can be used by market participants such as regulators, analyst, and auditors to detect probable financial reporting misstatements.


Keywords: Group Method of Data Handling (GMDH), Earnings quality, Neural networks, Accruals, Accruals quality.


* Corresponding Author: Hosseinsohrabi@gmail.com


How to Cite: Saghafi, A., Bulu, Gh., Sohrabi Varzaneh, H. A. (2022). An Accruals Quality Model: A Group Method of Data Handling Approach, *Empirical Studies in Financial Accounting Quarterly*, 19(75), 1-40.

مدل کیفیت اقلام تعهدی با رویکرد الگوریتم مدیریت گروهی

داده‌ها

علی ثقفی  استاد گروه حسابداری، دانشگاه علامه طباطبائی، تهران، ایران

قاسم بولو  دانشیار گروه حسابداری، دانشگاه علامه طباطبائی، تهران، ایران

حسینعلی سهرابی ورزنه  * دانشجوی دکتری رشته حسابداری، دانشگاه علامه طباطبائی، تهران، ایران

چکیده

مطالعه و توسعه معیارهای سنجش کیفیت سود و به‌طور ویژه کیفیت اقلام تعهدی یکی از موضوعات کلیدی در طی حدود سه دهه اخیر بوده است. ادبیات موضوعی حاکی از این است که معیارهای مبتنی بر رگرسیون خطی، عمدتاً دارای خطای بالا بوده و لذا در سال‌های اخیر، پژوهش‌هایی جهت اعمال الگوریتم‌های یادگیری ماشین صورت گرفته است. با این حال، بنیان‌گذاری این الگوریتم‌ها بر رویکرد جعبه سیاه، توسعه‌پذیری و همچنین درجه کاربردی بودن این مدل‌ها را با محدودیت مواجه می‌نماید. لذا در این پژوهش از الگوریتم مدیریت گروهی داده‌ها که یک نوع مدل‌سازی جعبه سفید تلقی می‌شود جهت پیش‌بینی اقلام تعهدی استفاده شده است. نتایج به‌دست آمده با استفاده از داده‌های ۱۶۴ شرکت پذیرفته شده در بورس اوراق بهادار تهران طی سال‌های ۱۳۸۵ تا ۱۳۹۷ و متغیرهای «تغییر در درآمد»، «دارایی‌های ثابت مشهود»، «جریان‌های نقدی عملیاتی دوره جاری»، «جریان‌های نقدی عملیاتی دوره گذشته» و «جریان‌های نقدی عملیاتی دوره آتی»، حاکی از بهبود سطح خطای این مدل‌ها نسبت به مدل‌های خطی می‌باشد. با توجه به برتری مدل‌های خروجی تحقیق حاضر از نظر خطای پیش‌بینی نسبت به مدل‌های رایج، یافته‌های این پژوهش می‌تواند توسط نهادهای نظارتی، تحلیلگران و حساب‌رسان در شناسایی موارد احتمالی تحریف اطلاعات مالی شرکت‌ها مورد استفاده قرار گیرد.

کلیدواژه‌ها: مدیریت گروهی داده‌ها، کیفیت سود، شبکه‌های عصبی، اقلام تعهدی.

مقدمه

داده‌های حسابداری که به استفاده‌کنندگان گزارش می‌شود باید دارای کیفیت مناسب باشند. بنابر مفاهیم نظری گزارشگری مالی، اطلاعات باکیفیت باید هم مربوط و هم قابل اتکا باشد. با این حال، از منظر حسابداری، این دو مفهوم کم‌وبیش مانع‌الجمع می‌باشند؛ یعنی هر چه مربوط بودن داده افزایش می‌یابد قابلیت اتکای آن داده کاهش می‌یابد و بالعکس. جهت ایجاد تعادل بین قابلیت اتکا و مربوط بودن داده‌های حسابداری، اصول پذیرفته‌شده حسابداری تا حدی به مدیران شرکت‌ها اجازه استفاده از قضاوت و انجام ارزیابی‌های ذهنی را در فرایند تهیه صورت‌های مالی می‌دهد. برای مثال، مدیریت شرکت در تعریف عمر مفید یک دارایی و یا در تعیین مقدار حساب‌های دریافتنی غیرقابل وصول، باید از قضاوت خود استفاده نماید. همچنین استفاده از حسابداری تعهدی نیز منجر به ضرورت اعمال قضاوت در بسیاری از رویدادهای مرتبط با گزارشگری مالی می‌گردد. علاوه بر این، با متمایل شدن گزارشگری مالی به سوی حسابداری ارزش منصفانه، درجه اعمال قضاوت افزایش خواهد یافت. استفاده از قضاوت و ذهنیت در حسابداری هم مهم و هم مطلوب است زیرا مدیریت شرکت در هنگام تعیین ثبت صحیح تراکنش‌های تجاری، قادر به اعمال دانش خود و همچنین در نظر گرفتن شرایط کلی اقتصاد و سایر عوامل داخلی شرکت خواهد بود (Parfet, 2000).

هرگاه قضاوت به درستی در حسابداری به کار گرفته شود، از طریق ارتقاء مربوط بودن و به موقع بودن گزارش‌های مالی منجر به کاهش عدم تقارن اطلاعاتی می‌گردد. متأسفانه، مدیری که به دنبال منافع شخصی است می‌تواند از این اختیار سوءاستفاده نماید که این امر منجر به کاهش قابلیت اتکای گزارشگری مالی و در نتیجه کاهش کیفیت سود می‌گردد. این استفاده فرصت‌طلبانه از قضاوت در گزارشگری مالی، مدیریت سود نامیده می‌شود (Hoglund, 2010). لذا طیف گسترده‌ای از پژوهشگران در حوزه مالی و حسابداری که به دنبال به دنبال تعیین کیفیت سود هستند، معمولاً به این موضوع که آیا استفاده‌کنندگان گزارش‌های مالی می‌توانند مدیریت سود را تشخیص دهند یا اینکه توسط

آن گمراه می‌شوند علاقه‌مند بوده‌اند. هرچند شواهد همه یک نتیجه واحد را نشان نمی‌دهند اما به هر حال شواهدی وجود دارد مبنی بر اینکه استفاده‌کنندگان گزارش‌های مالی تا حدی توسط سوءاستفاده از قضاوت حسابداری گمراه می‌شوند. برای مثال، Teoh et. Al. (1998) نشان می‌دهند که سرمایه‌گذاران توسط مدیریت پیش از انتشار سود در عرضه‌های فصلی سهام گمراه می‌شوند. Sloan (1996) بیان می‌کند که بازار قادر به اعمال کامل اطلاعات ارقام تعهدی در قیمت سهام نمی‌باشد. از آنجا که سود اغلب از طریق ارقام تعهدی مدیریت می‌شود، این مهم حاکی از این است که داده‌های حسابداری گزارش شده قادر به انعکاس سوءاستفاده از قضاوت حسابداری نمی‌باشند.

با توجه به جایگاه و اهمیت موضوع مدیریت سود، تلاش گسترده‌ای برای مدل‌سازی آن صورت گرفته است. بخش عمده‌ای از ادبیات موضوع در این حوزه، به پژوهش‌هایی تعلق دارد که مدیریت سود را از طریق پیش‌بینی ارقام تعهدی غیراختیاری، برآورد می‌نمایند. این بخش از ادبیات، اغلب بر مدل ارائه‌شده توسط Jones (1991) مبتنی بوده که تغییرات ارقام تعهدی را از طریق تغییر در درآمدها و دارایی‌های استهلاک‌پذیر قابل توضیح می‌داند.

با این حال، برخی پژوهش‌ها از جمله Ball & Shivakumar (2006) و Kothari et. Al. (2005) شواهدی در خصوص نامناسب بودن فرض رابطه خطی بین متغیرهای توضیحی و ارقام تعهدی در مدل‌های مزبور ارائه می‌نمایند. Ball & Shivakumar (2006) نشان می‌دهد که وارد نمودن مفهوم محافظه‌کاری در مدل McNichols (2002) به بهبود کارایی مدل مزبور منجر می‌گردد. در ادامه، برخی مطالعات اقدام به استفاده از شبکه‌های عصبی جهت پیش‌بینی ارقام تعهدی یا ارقام تعهدی اختیاری نموده‌اند (Hoglund, 2012؛ کردستانی و همکاران، ۱۳۹۲؛ مشایخی و همکاران، ۱۳۹۱؛ حجازی و همکاران، ۱۳۹۱). با این حال، اصلی‌ترین رویکردهای مورد استفاده در این تحقیقات، شبکه‌های عصبی پرسپترون چندلایه و همچنین شبکه‌های مبتنی بر توابع پایه شعاعی بوده‌اند. مدل‌های مبتنی بر این الگوریتم‌ها در دسته مدل‌های

جعبه سیاه قرار می‌گیرند. باوجوداینکه نتایج حاصل از این مطالعات، حاکی از بهبود عملکرد پیش‌بینی ارقام تعهدی نسبت به مدل‌های خطی می‌باشد، اما این مدل‌ها عموماً به دلیل وجود لایه مخفی، تفسیرپذیری روابط بین متغیرها را دچار محدودیت می‌کنند (Guo et. Al. 2018). لذا در این پژوهش تلاش شده است تا با استفاده از رویکرد الگوریتم مدیریت گروهی داده‌ها^۱ که در زمره رویکردهای جعبه سفید قرار می‌گیرند نسبت به برآورد مدل روابط بین ارقام تعهدی و متغیرهای بنیادی و عملکردی شرکت اقدام شود.

پیشینه پژوهش

از زمان پژوهش Healy (1985) و پس‌از آن ارائه مدل Jones (1991)، بخش قابل‌توجهی از ادبیات مرتبط با کیفیت سود و مدیریت سود به بهبود مدل Jones (1991) مربوط بوده است. ایده اساسی در مدل جونز این است که ارقام تعهدی توسط تغییرات در درآمدها و دارایی‌های استهلاک‌پذیر قابل توضیح است.

باوجوداینکه مدل جونز و سایر مدل‌های مبتنی بر آن در سطح وسیعی به کار برده شده‌اند اما این مدل‌ها مورد انتقادات زیادی نیز قرار دارند. برای مثال، Dechow et. Al. (1995) نشان می‌دهند که توان مدل‌های موجود در مورد شناسایی مقادیری از مدیریت سود که به لحاظ اقتصادی باورپذیرتر و محتمل‌ترند پایین می‌باشد. علاوه بر این، Dechow & Skinner (2000) بیان می‌کنند که روش‌شناسی‌های تحقیق مورد اشاره به‌اندازه کافی در تعیین و تشخیص مدیریت سود و کیفیت ارقام تعهدی اختیاری شرکت‌ها خوب و کارآمد نیستند.

یکی از مفروضات مدل‌های فوق این است که بین متغیرهای مستقل مدل و ارقام تعهدی شرکت رابطه خطی وجود دارد. بااین‌حال، شواهدی وجود دارد که رابطه بین ارقام تعهدی و عملکرد غیرخطی می‌باشد. Ball & Shivakumar (2006) ادعا می‌کنند که مدلی که رابطه بین ارقام تعهدی و جریان‌های نقدی را خطی فرض می‌کند مدل مناسبی

1. Group Method of Data Handling (GMDH)

نیست زیرا اصل محافظه کاری حسابداری را در نظر نگرفته است. آن‌ها برای غلبه بر این مشکل، استفاده از رگرسیون خطی منقطع^۱ را پیشنهاد می‌دهند به این شکل که مثبت یا منفی بودن جریان‌های نقدی حاصل از عملیات منجر به شیب‌های متفاوتی می‌گردد. (Kothari et. Al. (2005) نیز مشکل غیرخطی بودن را تأیید می‌کنند. جدول ۱ تعدادی از مدل‌های مبتنی بر رگرسیون در برآورد ارقام تعهدی را نشان می‌دهد.

مدل‌های مورد اشاره در این جدول به استثنای مدل توسعه یافته توسط Ball & Shivakumar (2006) همگی مبتنی بر این مفروض قرار دارند که رابطه بین متغیرهای توضیحی و ارقام تعهدی، یک رابطه خطی می‌باشد. مطالعه مزبور، مدل را بابت در نظر گرفتن محافظه کاری حسابداری تعدیل می‌نماید.

یک رویکرد بدیل که در تعدادی از پژوهش‌ها به منظور حل مشکل غیرخطی بودن رابطه بین متغیرهای توضیحی و ارقام تعهدی از آن استفاده شده است، مدل‌های مبتنی بر شبکه‌های عصبی می‌باشد. پژوهش‌های متعدد در حوزه حسابداری نشان داده‌اند که شبکه‌های عصبی در موارد متعددی بهتر و یا حداقل به همان خوبی روش‌های آماری سنتی عمل می‌کنند (برای نمونه، Al-Dmour & Al-Dmour, 2018؛ Dameri et. Al. 2020؛ Calderon & Cheh, 2002). در مواردی که مشکلاتی در تحلیل رگرسیون وجود دارد، شبکه‌های عصبی چندین مزیت در مقایسه با روش‌های آماری سنتی دارند. مهم‌ترین مزیت آن‌ها این است که شبکه‌های عصبی قادر به مدل‌سازی روابط غیرخطی بین متغیرهای مستقل و وابسته هستند. مزیت دیگر این است که شبکه‌های عصبی نیاز به مفروضات در خصوص داده‌های زیربنایی خود ندارند زیرا این شبکه‌های عصبی اصولاً مشتق از داده‌ها هستند. در نهایت این که در مقایسه با رگرسیون خطی، شبکه‌های عصبی نسبتاً در برخورد با داده‌های پرت و اخلال در داده‌ها، توان بالایی دارند (DeTienne et. Al. 2003)؛ بنابراین، می‌توان انتظار داشت که شبکه‌های عصبی بتواند برخی از کاستی‌های مدل‌های رگرسیون خطی را در برآورد ارقام تعهدی غیراختیاری رفع نماید.

1. Piecewise linear regression
2. DeTienne et al.

جدول ۱. پرکاربردترین مدل‌های پیش‌بینی ارقام تعهدی

پژوهش	مبانی نظری	مدل
مدل جونز (۱۹۹۱)	ارقام تعهدی سرمایه در گردش، تابعی از رشد درآمد (ΔRev) و استهلاک، تابعی از اموال، ماشین‌آلات و تجهیزات (PPE) است.	$Acc_t = \alpha + \beta_1 \Delta Rev_t + \beta_2 \Delta PPE_t + \varepsilon_t$ <p>که در آن، Acc_t جمع ارقام تعهدی در پایان سال t، ΔRev رشد درآمد شرکت و ΔPPE تغییرات در دارایی‌های ثابت مشهود است.</p>
مدل جونز تعدیل‌شده توسط دجو و همکاران (۱۹۹۵)	این مدل با این فرض که ممکن است فروش نسبه شرکت با توجه به اختیارات مدیریت تحت تأثیر قرار گیرد، نسبت به تعدیل متغیر رشد درآمد در مدل جونز اقدام می‌نماید.	$Acc_t = \alpha + \beta_1 (\Delta Rev_t - \Delta Rec_t) + \beta_2 \Delta PPE_t + \varepsilon_t$ <p>که در آن ΔRev تغییرات در حساب‌های دریافتی را نشان می‌دهد.</p>
مدل دجو و دیچف (۲۰۰۲)	ارقام تعهدی به‌عنوان تابعی از جریان‌های نقدی گذشته، حال و آینده مدل می‌شود. این امر جهت در نظر گرفتن زمان‌بندی متفاوت شناخت و انعکاس جریان نقدی در سود انجام می‌شود.	$\Delta WC = \alpha + \beta_1 CFO_{t-1} + \beta_2 CFO_t + \beta_3 CFO_{t+1} + \varepsilon_t$ <p>که در آن CFO_t جریان نقد عملیاتی برای دوره t، CFO_{t-1} جریان نقد عملیاتی برای دوره $t-1$ و CFO_{t+1} جریان نقد عملیاتی برای دوره $t+1$ است.</p>
مدل مک نیکولز (۲۰۰۲)	در این مدل، متغیرهای دو مدل جونز (۱۹۹۱) و دجو و دیچف (۲۰۰۲) به‌صورت هم‌زمان جهت پیش‌بینی ارقام تعهدی استفاده می‌شود.	$\Delta WC = \alpha + \beta_1 CFO_{t-1} + \beta_2 CFO_t + \beta_3 CFO_{t+1} + \beta_4 \Delta Rev_t + \beta_5 \Delta PPE_t + \varepsilon_t$
مدل فرانسیس و همکاران (۲۰۰۵)	این مدل، کیفیت سود را مطابق رویکرد مک نیکولز در نظر می‌گیرد. با این حال، این محققان بر این باورند که کیفیت سود، صرفاً تابعی از اختیارات مدیریت نبوده و دارای دو جزء ذاتی و اختیاری می‌باشد.	$TCA_t = \alpha + \beta_1 CFO_{t-1} + \beta_2 CFO_t + \beta_3 CFO_{t+1} + \beta_4 \Delta Rev_t + \beta_5 PPE_t + \varepsilon_t$ $\sigma(\varepsilon_t) = \alpha + \lambda_1 Size_t + \lambda_2 \sigma(CFO)_t + \lambda_3 \sigma(Rev)_t + \lambda_4 \log(OperCycle)_t + \lambda_5 NegEarn_t + v_t$

پژوهش	مبانی نظری	مدل
مدل بال و شیواکومار (۲۰۰۶)	مبنای این مدل آن است که به دلیل وجود محافظه کاری حسابداری، مدل خطی به تنهایی نمی تواند نحوه پیش بینی اقلام تعهدی توسط جریان نقد عملیاتی را تبیین نماید و لذا با در نظر گرفتن یک متغیر مجازی، از رویکرد رگرسیون خطی منقطع استفاده می نماید.	$ACC_t = \alpha_0 + \alpha_1 CFO_{t-1} + \alpha_2 CFO_t + \alpha_3 CFO_{t+1} + \alpha_4 \Delta Rev_t + \alpha_5 PPE_t + \alpha_6 \Delta CFO_t + \alpha_7 Dum \Delta CFO_t + \alpha_8 \Delta CFO_t \times Dum \Delta CFO_t + \varepsilon_t$ <p>که در آن، Dum متغیر مصنوعی بوده که در صورت مثبت بودن CFO_t مقدار صفر و در صورت منفی بودن آن، مقدار ۱ به خود می گیرد.</p>

با وجود برتری شبکه های عصبی بر رویکردهای رگرسیون از نظر بهبود کارایی و خطای پیش بینی کمتر، این شبکه ها عمدتاً منجر به تفسیر پذیری کمتر مدل های خروجی می گردد. این امر ناشی از بنیان گذاری بسیاری از این شبکه های عصبی بر رویکرد جعبه سیاه می باشد. در شبکه های عصبی مبتنی بر جعبه سیاه، عمدتاً به دلیل وجود یک یا چند لایه مخفی در الگوریتم یادگیری ماشین، مدل خروجی قابلیت بیان در قالب یک رابطه ریاضی چند متغیره را ندارد.

اصطلاح رویکرد جعبه سیاه^۱، اساساً برای اشاره به کلیه مدل های یادگیری ماشین که به سختی قابل تبیین و قابل فهم توسط خبرگان در حوزه های عملی می باشد استفاده می گردد (Dong, 2019؛ Gonzalez, 2019؛ Rudin, 2019). Gonzalez (2019) این مدل های جعبه سیاه را در چهار دسته (الف) مدل های مبتنی بر ابر صفحه^۲، (ب) مدل های مبتنی بر شبکه های عصبی زیستی که مغز موجودات زنده را شکل می دهد، (ج) مبتنی بر احتمالات و منطق ترکیبی و (د) مدل های مبتنی بر نمونه ها (مدل های یادگیری تنبل^۳) طبقه بندی می کند.

1. White Box
2. Hyperplane
3. Lazy Learning

عموماً، متخصصان در حوزه‌های عملی، دانش خود را به‌جای تمرکز بر کسب اطلاعات در خصوص یادگیری ماشین، بر یادگیری پدیده‌های حوزه تخصصی خود متمرکز می‌کنند. در نتیجه، برای عمده این متخصصان، فهم مدل‌هایی که دارای توابع ریاضی پیچیده باشند دشوار است. برای بسیاری از این متخصصان، هرگونه به‌کارگیری یادگیری ماشین باید به شکلی واضح و دقیق به تصمیم‌گیری‌ها کمک نماید (Rudin, 2019).

اصطلاح مدل جعبه سفید که گاه به آن مدل قابل فهم نیز اطلاق می‌شود برای اشاره به آن دسته از مدل‌هایی استفاده می‌گردد که نتایج آن مدل‌ها توسط متخصصان حوزه عملی مربوطه، به راحتی قابل فهم می‌باشد. این مدل‌ها عموماً یک تعادل مناسب بین دقت و توضیح‌پذیری را ارائه می‌کنند (Gonzalez, 2019). معمولاً واژه‌های تفسیرپذیر و قابل فهم برای اشاره به کلیه مدل‌هایی که توضیحی را برای خبرگان حوزه مربوطه فراهم می‌آورند استفاده می‌گردد. باین حال، طبق (Rudin, 2019)، یک مدل قابل فهم به دسته‌ای از یادگیری ماشین اطلاق می‌گردد که نیازمند یک مدل یا ویژگی اضافی برای فراهم آوردن توضیح به متخصصان حوزه عملی می‌باشد و در مقابل، یک مدل تفسیرپذیر قادر به ارائه توضیح به خبرگان بدون استفاده از یک مدل اضافی می‌باشد.

به هر جهت، در مقایسه رویکرد جعبه سفید با رویکرد جعبه سیاه، بسیاری از محققان اعتقاد دارند که عملکرد مدل‌ها به زمینه تخصصی مورد استفاده بستگی دارد (Gonzalez, 2019). به عبارتی، در برخی از حوزه‌ها، عملکرد مدل‌های جعبه سیاه و در برخی از حوزه‌ها، عملکرد مدل‌های جعبه سفید عملکرد مناسب‌تری دارد. علاوه بر این، برخی از محققان نیز به استفاده از مدل‌های ترکیبی جهت دستیابی به بهترین عملکرد اقدام می‌کنند (Gonzalez, 2019).

عمده اشکال شبکه‌های عصبی استفاده‌شده در ادبیات حوزه‌های حسابداری در ذیل شبکه‌های مبتنی بر رویکرد جعبه سیاه طبقه‌بندی می‌گردند. (برای نمونه، Duan et. Al. 2018؛ Hoglund, 2012؛ Abdoh et. Al. 2020؛ کردستانی و همکاران، ۱۳۹۲؛

مشایخی و همکاران، ۱۳۹۱). برخی از نمونه‌های این گروه از شبکه‌های عصبی، شبکه‌های عصبی پرسپترون چندلایه^۱، توابع پایه شعاعی^۲، ماشین بردار پشتیبان^۳ و الگوریتم نگاشت خودسامانده^۴ می‌باشد. شواهد تجربی حاصل از مشاهدات پیشین حاکی از عملکرد رضایت‌بخش این مدل‌ها و برتری این مدل‌ها در بیشتر موارد نسبت به مدل‌های رگرسیون از منظر خطای پیش‌بینی می‌باشند (به‌عنوان نمونه، کردستانی و همکاران، ۱۳۹۲؛ مشایخی و همکاران، ۱۳۹۱؛ Duan et. Al. 2018؛ Al-Dmour & Al-Dmour, 2018). با این حال، حتی پس از اثبات عملکرد بهتر این مدل‌ها نسبت به روش‌های سنتی، تلاش‌ها جهت بهبود آن نسبت به مدل‌های سنتی با اقبال کمتری مواجه شده است. شاید تفسیرپذیری پایین نتایج این مدل‌ها را بتوان یکی از اصلی‌ترین دلایل این عدم استقبال تلقی نمود. به عبارتی، این مدل‌ها به دلیل داشتن یک یا چندلایه مخفی، روابط را به نحوی ارائه نمی‌کنند که سایر پژوهشگران بتوانند به نقش نسبی عوامل و متغیرها در پیش‌بینی مدل آگاه شده و روابط بین متغیرها را به نحو مناسبی تحلیل نمایند.

به نظر می‌رسد استفاده از الگوریتم‌های یادگیری ماشین مبتنی بر رویکرد جعبه سفید بتواند به محققان در ایجاد تعادل بین کارایی مدل و توضیح‌پذیری آن کمک نماید. الگوریتم مدیریت گروهی داده‌ها به‌عنوان یکی از رویکردهای جعبه سفید در یادگیری ماشین، یک شبکه عصبی پیش‌خور جهت مدل‌سازی سیستم‌های پیچیده است که در دهه ۱۹۶۰ توسط ایواکنکو^۵ معرفی شد. این الگوریتم در واقع با تجزیه یک مسئله پیچیده به تعدادی روابط چندجمله‌ای ساده‌تر به درک روابط بین متغیرها کمک می‌نماید. این مدل در تحقیقات حسابداری تاکنون چندان مورد توجه قرار نگرفته است. در یکی از معدود مطالعات داخلی صورت گرفته با این روش در حوزه حسابداری، انواری رستمی و همکاران (۱۳۹۲) این الگوریتم را جهت پیش‌بینی سود هر سهم در شرکت‌های بورسی به کار

-
1. Multi-Layer Perceptron (MLP)
 2. Radial Basis Function (RBF)
 3. Support Vector Matrix (SVM)
 4. Self-Organizing Map (SOM)
 5. A. G. Ivakhnenko

گرفتند. نتایج پژوهش مزبور نشان‌دهنده برتری شبکه عصبی مدیریت گروهی داده‌ها از نظر دقت پیش‌بینی سود هر سهم نسبت به روش خودرگرسیون می‌باشد. به‌هرحال کاربرد این الگو در ادبیات حوزه‌های مرتبط مانند مالی و اقتصاد مورد توجه گسترده‌تری قرار گرفته است. به‌عنوان نمونه، (Nazemi & Rafiean (2020 از این الگوریتم در پیش‌بینی قیمت مسکن تا سال ۲۰۲۲ استفاده می‌نماید. Hassan et. Al. (2010 از الگوریتم مدیریت گروهی داده‌ها برای ارزیابی اقتصادی آلودگی هوا استفاده می‌نماید. محققان پژوهش مزبور، این الگوریتم را جهت کشف الگوهای مخفی در روابط بین متغیرهای اقتصادی (شامل درآمد ناخالص ملی، تورم، بدهی، شاخص فساد، رشد تولید ناخالص داخلی، جمعیت و خالص سرمایه‌گذاری خارجی) و داده‌های کیفیت هوا (از جمله گوگرد دی‌اکسید، نیتروژن دی‌اکسید) به کار می‌گیرند.

مقایسه رویکردهای رگرسیون خطی، الگوریتم مدیریت گروهی داده‌ها و شبکه عصبی چندلایه

(DeTienne et. Al. (2003 مزایا و معایب شبکه‌های عصبی در مقایسه با رگرسیون خطی را مطرح می‌کنند. اول اینکه روش‌های رگرسیون خطی نمی‌تواند روابط غیرخطی بین متغیرهای مستقل و متغیرهای وابسته را در نظر بگیرد در حالی که شبکه‌های عصبی به‌خوبی و با اثربخشی مناسب روابط غیرخطی را مدل‌سازی می‌کنند. دوم اینکه عملکرد مدل‌های رگرسیون خطی تا حد زیادی به مفروضات متعدد مانند عدم هم خطی چندگانه و توزیع نرمال باقیمانده‌ها وابسته است. سوم هم اینکه در مدل‌های رگرسیون خطی، مدل زیربنایی باید از قبل مشخص شود. به‌عنوان نمونه، مفروض مدل این است که متغیر وابسته، ترکیبی خطی از متغیرهای مستقل است. این در حالی است که در شبکه‌های عصبی چنین الزامی وجود ندارد زیرا این شبکه‌ها کاملاً مشتق از داده‌ها هستند. در مقابل، بخش

قابل ملاحظه‌ای از شبکه‌های عصبی دارای لایه‌های پنهان بوده و دارای تفسیرپذیری به مراتب پایین‌تری نسبت به مدل‌های رگرسیون می‌باشد.

همچنین استفاده از الگوریتم مدیریت گروهی داده‌ها در مقایسه با شبکه‌های عصبی سنتی، دو مزیت عمده به همراه دارد: (۱) این الگوریتم به‌طور خودکار تعداد لایه‌ها و نورون‌های هر لایه را مشخص می‌کند که باعث می‌شود موفقیت مصنوعی و غیرواقعی در فرایند شبیه‌سازی اتفاق نیفتد و (۲) بین پارامترهای منتخب و خروجی روابطی در قالب معادلات چندمتغیره تعریف می‌کند درحالی‌که در شبکه‌های عصبی، برآورد شبکه در قالب جعبه سیاه رخ می‌دهد.

پیشینه تحقیق

شمار زیادی از تحقیقات در حوزه حسابداری و سایر حوزه‌ها از رویکردها و گونه‌های مختلف شبکه عصبی استفاده کرده‌اند. (Calderon & Cheh (2002) به ۲۷ مطالعه در حسابرسی و مدیریت ریسک می‌پردازند که در آن‌ها از شبکه‌های عصبی استفاده شده است. بیش از نیمی از این مطالعات مرتبط با بحران‌های مالی و ورشکستگی می‌باشند. پر استفاده‌ترین شکل شبکه‌های عصبی که در تحقیقات مزبور استفاده شده است، شبکه‌های عصبی پس‌انتشار می‌باشند. آن‌گونه که (Calderon & Cheh (2002) اشاره می‌کنند در اکثر تحقیقات مزبور در نهایت مدل‌های مبتنی بر شبکه عصبی با مدل‌های آماری متنوع مقایسه شده‌اند. با وجود اینکه در بسیاری از این تحقیقات، شواهد حاکی از بهتر عمل کردن این مدل‌ها نسبت به رویکردهای آماری سنتی است اما در برخی موارد شواهد متضاد آن نیز مشاهده می‌شود.

Hoglund (2012) در یکی از اولین پژوهش‌ها جهت ارزیابی توانایی مدل‌های شبکه عصبی در پیش‌بینی ارقام تعهدی و مدیریت سود، به مقایسه نتایج حاصل از این مدل‌ها پرداخت. در پژوهش مزبور، از بین مدل‌های پیشین، رویکردهای Dechow & Dichev (2002) و Ball & Shivakumar (2006) و در مقابل، شبکه‌های عصبی پرسپترون چندلایه (MLP)، شبکه عصبی نگاشت خودسامانده (SOM) و شبکه عصبی

رگرسیون عمومی (GRNN) استفاده شد. نتایج حاصل از تحقیق نشان می‌دهد که شبکه عصبی پرسپترون چندلایه در مقایسه با سایر شبکه‌های عصبی و همچنین در مقایسه با دو رویکرد و رگرسیون خطی (Dechow & Dichev, 2002) و رگرسیون خطی منقطع (Ball & Shivakumar, 2006) عملکرد بهتری از نظر میانگین مجذورات خطا دارد. همچنین از بین مدل‌های رگرسیون، مدل رگرسیون منقطع عملکرد مناسب‌تری نشان می‌دهد.

Duan et. Al. (2018) در مطالعه خود در صنعت نفت و گاز از شبکه‌های عصبی جهت مدل‌سازی نتایج حسابداری شرکت‌ها استفاده نمودند. در پژوهش مزبور با استفاده از چهار ساختار شبکه عصبی و دو گونه شرایط آموزش شبکه، شواهد نشان می‌دهد که اثرگذاری قراردادهای آتی نفت و نرخ بهره بر نتایج حسابداری تقریباً دوبرابر سایر عوامل می‌باشد. همچنین میزان اثرگذاری انتخاب‌های حسابداری مشابه میزان اثرگذاری نرخ تسعیر ارز، ذخایر نفتی و سهم بازار شرکت می‌باشد.

Abdou et. Al. (2018) به مطالعه رابطه بین ویژگی‌های حاکمیت شرکتی و مدیریت سود با استفاده از شبکه‌های عصبی پرداختند. در پژوهش مزبور که با استفاده از داده‌های شرکت‌های فعال در بریتانیا و مصر صورت گرفت، نتایج نشان می‌دهد که وضعیت نظام حاکمیتی کشور بر رابطه بین حاکمیت شرکتی و مدیریت سود اثرگذار است. این نتایج حاکی از این است که اندازه هیئت‌مدیره کوچک‌تر، مدیریت شدن توسط مدیران خارجی و دارا بودن درصد پایین‌تری از مدیران زن، منجر به مدیریت سود کمتری در شرکت‌های بریتانیایی نسبت به شرکت‌های مصری می‌گردد.

Nazemi & Rafiean (2020) پیش‌بینی قیمت مسکن با استفاده از الگوریتم مدیریت گروهی داده‌ها را مورد پژوهش قرار دادند. نتایج پژوهش حاکی از عملکرد مطلوب و قابل قبول این الگوریتم در پیش‌بینی قیمت مسکن در شهر اصفهان می‌باشد.

حجازی و همکاران (۱۳۹۱) به بررسی پیش‌بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم اقدام نمودند. مطالعه مزبور با بررسی داده‌های فصلی مربوط به ۵۵

شرکت از بهار ۱۳۸۵ تا تابستان ۱۳۸۸ و با استفاده از متغیرهای «نسبت اعضای غیرموظف هیئت‌مدیره»، «حساسیت در پرداخت بابت عملکرد»، «تداوم سود»، «فعالیت تأمین مالی»، «آستانه عملکرد»، «نسبت اهرمی»، «ریسک شرکت»، «اقدام تعهدی اختیاری دوره قبل»، «اندازه شرکت»، «عملکرد شرکت» و «فصل» صورت گرفت. نتایج حاصل از پیاده‌سازی روش رگرسیون خطی و مقایسه آن با رویکردهای شبکه عصبی پرسپترون با یک لایه پنهان که از الگوریتم ژنتیک جهت آموزش مدل استفاده شده است، حاکی از عملکرد بهتر رویکرد شبکه عصبی نسبت به رویکرد خطی از منظر ضریب تعیین (۹۵ درصد در مقابل ۶۳ درصد) و از منظر میانگین مجذورات خطا (۰,۰۲۲۷۱ در مقابل ۰,۱۲۳۹) می‌باشد.

مشایخی و همکاران (۱۳۹۱) نیز اقدام به پیاده‌سازی بخشی از آزمون‌های مورد اشاره توسط Hoglund (2012) با استفاده از داده‌های شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران در بازه سال‌های ۱۳۸۱ تا ۱۳۸۹ نمودند. در مطالعه مزبور، از دو نوع شبکه عصبی پرسپترون چندلایه و شبکه مبتنی بر توابع پایه شعاعی استفاده شده و نتایج حاصل از این دو شبکه با نتایج حاصل از رویکردهای رگرسیون خطی و رگرسیون خطی منقطع مقایسه گردیده است. نتایج مطالعه نشان می‌دهد که از منظر ضریب تعیین، شبکه عصبی مبتنی بر توابع پایه شعاعی، شبکه عصبی پرسپترون چندلایه پیشرو، رگرسیون خطی منقطع و رگرسیون خطی به ترتیب بهترین عملکرد را دارا می‌باشند. ضریب تعیین این رویکردها به ترتیب ۷۷ درصد، ۵۳ درصد، ۴۲ درصد و ۴ درصد می‌باشد. با این حال، از نظر میزان خطا با استفاده از معیار میانگین مجذور خطاها، رگرسیون‌های خطی و خطی غیرمنقطع با ۰,۰۲۸ میانگین مجذور خطا عملکرد مناسب‌تری نسبت به پرسپترون چندلایه پیشرو (۰,۰۳۵) دارند. از این نظر نیز شبکه عصبی مبتنی بر توابع پایه شعاعی بهترین عملکرد را دارا می‌باشند به نحوی که میانگین مجذور خطا برای این شبکه عصبی معادل ۰,۰۱۵ می‌باشد.

کردستانی و همکاران (۱۳۹۲) نیز توانایی شبکه‌های عصبی را جهت پیش‌بینی مدیریت سود مورد بررسی قرار دادند. در مطالعه مزبور، از شبکه عصبی پرسپترون سه لایه و

متغیرهای «حد عملکرد»، «نسبت اهرمی»، «ریسک شرکت»، «اقلام تعهدی اختیاری سال قبل»، «پایداری سود»، «اندازه شرکت»، «عملکرد شرکت» و «فعالیت‌های تأمین مالی» استفاده شد. با آزمون و خطا از بین شبکه‌های با تعداد نورون و تعداد سیکل تکرار مختلف، شبکه متشکل از ۲۴ نورون و ۲۰۰۰ تکرار، دارای بالاترین میزان دقت در برآورد مدیریت سود تشخیص داده شد. این شبکه در دسته داده‌های آموزش دارای دقت ۹۴ درصد و در دسته داده‌های آزمون دارای دقت ۶۹ درصد می‌باشد.

کاشانی‌پور و همکاران (۱۳۹۸) در مطالعه خود با استفاده از تجزیه و تحلیل سلسله‌مراتبی به مطالعه ابعاد کیفیت سود و سنجه‌های آن پرداخته و مدلی برای رتبه‌بندی سنجه‌های کیفیت سود ارائه نمودند. در پژوهش مزبور، از بین ابعاد حسابداری، بازار و مدیریتی کیفیت سود، بعد حسابداری دارای بیشترین وزن تشخیص داده شد. با این حال، از بین سنجه‌های مورد مطالعه، سنجه آگاهی‌دهندگی از بعد بازار دارای بیشترین اثر بر کیفیت سود می‌باشد.

عباس‌زاده و همکاران (۱۳۹۸) به مطالعه رابطه بین ارتباطات سیاسی، معاملات با اشخاص وابسته و مدیریت سود در شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران پرداختند. در پژوهش مزبور از مدل Kothari et. Al. (2005) برای سنجش مدیریت سود استفاده گردیده و نتایج حاکی از وجود رابطه معنادار بین ارتباطات سیاسی و معاملات با اشخاص وابسته و عدم مشاهده رابطه معنادار بین معاملات با اشخاص وابسته و مدیریت سود بوده است.

سؤالات پژوهش

با توجه به مبانی نظری و پیشینه تجربی، سؤال این پژوهش به شرح زیر می‌باشد:
«مدل پیش‌بینی اقلام تعهدی بر مبنای الگوریتم مدیریت گروهی داده‌ها چیست؟»

روش تحقیق

پژوهش حاضر از منظر نتایج، از نوع کاربردی، از نظر هدف پژوهش از نوع تحلیلی و از نظر زمانی از نوع گذشته‌نگر و پس‌رویدادی است. هدف این تحقیق ارزیابی توانایی الگوریتم مدیریت گروهی داده‌ها در ارزیابی کیفیت سود یا به بیان دقیق‌تر، توانایی این شبکه‌ها در برآورد بخش غیر اختیاری ارقام تعهدی می‌باشد. با توجه به استفاده تحقیقات پیشین از مدل‌های شبکه عصبی که در قالب رویکرد جعبه سیاه عمل می‌نمایند، در این پژوهش جهت بهره‌مندی از منافع مدل‌های جعبه سفید، از الگوریتم مدیریت گروهی داده‌ها استفاده می‌شود. همچنین عملکرد این مدل با دو مدل رگرسیون خطی مقایسه خواهد شد. مدل‌های مبتنی بر رگرسیون که در این تحقیق مورد استفاده قرار می‌گیرند عبارت‌اند از: رگرسیون خطی ارائه‌شده توسط McNichols (2002) و رگرسیون خطی منقطع ارائه‌شده توسط Ball & Shivakumar (2006)

الگوریتم مدیریت گروهی داده‌ها

الگوریتم مدیریت گروهی داده‌ها به‌عنوان یکی از رویکردهای جعبه سفید در یادگیری ماشین، یک شبکه عصبی جهت مدل‌سازی سیستم‌های پیچیده است. شکل کلی این شبکه می‌تواند در قالب یک رشته از توابع چندمتغیره پیچیده موسوم به چندجمله‌ای کولموگوروف-گابور بیان شود: (Farlow, 1981)

$$\hat{y} = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n a_{ijk} x_i x_j x_k + \dots \quad (1)$$

با این حال، معمولاً رابطه کلی فوق، در هر لایه به صورت چندجمله‌ای دودویی زیر به کار گرفته می‌شود:

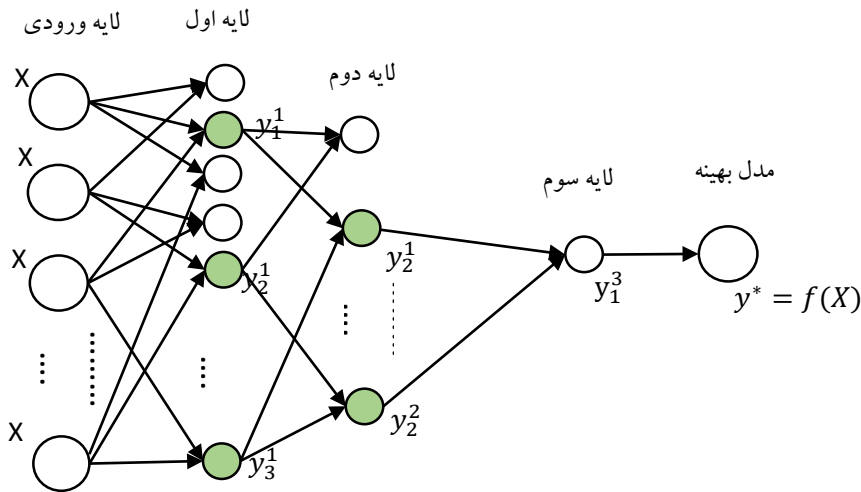
$$\hat{y} = a_1 + a_2 x_1 + a_3 x_2 + a_4 x_1^2 + a_5 x_2^2 + a_6 x_1 x_2 \quad (2)$$

هر جفت نورون در هر لایه به شکل رابطه (۲) محاسبه می‌شود و سپس تفاوت بین مقادیر واقعی خروجی یعنی y و مقدار به‌دست‌آمده \hat{y} به دست می‌آید. این امر به ایجاد نورون‌های جدید در لایه دوم منجر می‌گردد. جهت جلوگیری از ازدیاد نورون‌ها و به جهت نزدیک شدن به مدل بهینه، انتخاب نورون‌های برتر از بین کلیه نورون‌های لایه جدید

صورت می‌گیرد. این غربالگری بر اساس یک عامل مناسب (معمولاً میانگین مجذور خطاها) انجام می‌شود.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad \text{رابطه (۳)}$$

سپس انتخاب جفت‌های نورون‌های برگزیده و اعمال رویکرد مشابه در خصوص آن‌ها در لایه بعد تکرار می‌شود و این امر تا زمانی ادامه می‌یابد که روند کاهشی میزان خطا متوقف گردد یا تعداد نورون‌های برگزیده در یکی از لایه‌ها به یک نورون کاهش یابد.



شکل ۱. الگوی کلی الگوریتم مدیریت گروهی داده‌ها (GMDH)

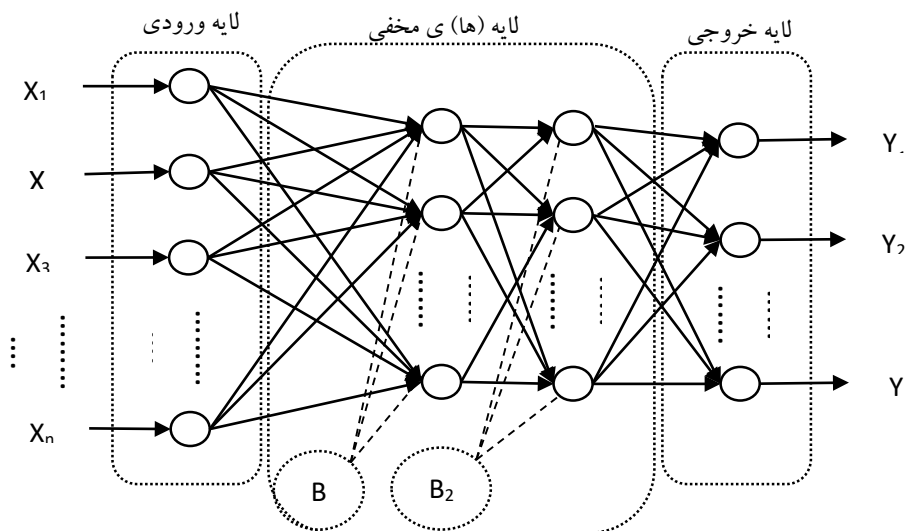
شبکه‌های عصبی

یکی از پر استفاده‌ترین گونه‌های شبکه‌های عصبی در مطالعات حسابداری، شبکه عصبی پس انتشار^۱ می‌باشد (Hoglund, 2010). یک شبکه عصبی پس انتشار حداقل از سه لایه تشکیل می‌شود: یک لایه ورودی، یک یا چند لایه مخفی و یک لایه خروجی (شکل ۱). تعداد گره‌ها در لایه ورودی معمولاً معادل تعداد متغیرهای مستقل و تعداد گره‌ها در لایه خروجی معمولاً معادل تعداد متغیرهای وابسته است. در خصوص تعیین تعداد گره‌ها در

1. Back-propagation neural network (BPNN)

لایه‌های مخفی، هیچ قاعده کلی وجود ندارد و این امر عمدتاً از طریق آزمون و خطا صورت می‌گیرد (DeTienne et. Al. 2003).

توابع انتقال (فعال‌سازی) در گره‌های لایه مخفی و خروجی نیز بخشی از معماری شبکه عصبی پس‌انتشار می‌باشند. اصولاً هر تابع قابل تمایز می‌تواند به عنوان یک تابع انتقال استفاده شود. عموماً از توابع انتقال غیرخطی در گره‌های لایه مخفی استفاده می‌شود زیرا این امر به شبکه عصبی پس‌انتشار، امکان سروکار داشتن با روابط غیرخطی بین متغیرهای ورودی و خروجی را می‌دهد (Zhang et. Al. 1999).



شکل ۲. الگوی کلی مدل شبکه عصبی پرسپترون

پس‌ازاینکه معماری شبکه تعیین شد، شبکه می‌تواند با استفاده از مجموعه داده‌های شامل زوج‌های متغیرهای ورودی و خروجی آموزش داده شود. هدف نهایی فرایند آموزش، به حداقل رساندن خطا بین مقدار خروجی محاسبه‌شده و مقدار خروجی واقعی برای یک بردار خاص از متغیرهای ورودی می‌باشد. خطا عموماً در قالب جمع مجذور خطاها^۱ یا

1. Sum Squared Errors (SSE)

میانگین مربعات خطاها^۱ اندازه‌گیری می‌شود. طی فرایند آموزش، مشاهدات موجود در مجموعه داده‌های آموزش به صورت مکرر به شبکه عصبی وارد می‌شود. پس از شروع آموزش، یک سؤال کلیدی این است که چه زمانی باید این آموزش متوقف شود. یک شرط بدیهی برای توقف آموزش، زمانی است که خطا به صفر رسیده باشد. باین‌حال، در اکثر موارد خطا هرگز به صفر نمی‌رسد. سایر شرایط برای توقف آموزش زمان‌هایی است که خطا به یک مقدار از پیش تعیین شده رسیده و یا زمانی که مجموعه داده‌های آموزش به تعداد دفعات مشخصی به شبکه ارائه شده باشد؛ اما تعیین این مقادیر از پیش تعیین شده نیز می‌تواند دشوار باشد. شاید کاراترین روش برای تعیین زمان توقف آموزش، استفاده از مجموعه داده‌های اعتبارسنجی مجزا باشد. معمولاً مجموعه داده‌های اعتباردهی با استفاده از درصد مشخصی از مشاهده‌ها از مجموعه داده آموزش ایجاد می‌شود. علاوه بر این، مشاهدات مجموعه داده‌های اعتبارسنجی برای تعدیل اوزان یک شبکه عصبی پس‌انتشار استفاده نمی‌شود. با استفاده از مجموعه داده‌های اعتبارسنجی، آموزش زمانی متوقف می‌شود که خطای اعتباردهی ثابت مانده و یا شروع به افزایش می‌کند (Hoglund, 2010).

درنهایت نیز کارایی یک شبکه عصبی پس‌انتشار بر اساس کیفیت تعمیم به داده‌های جدید تعیین می‌شود یعنی داده‌هایی که در طی فرایند آموزش استفاده نشده باشند. برای آزمون توانایی تعمیم شبکه عصبی، یک مجموعه داده آزمون استفاده می‌شود. مشابه آنچه در خصوص مجموعه داده اعتباردهی مطرح بود، مجموعه داده آزمون معمولاً از درصدی مشخص از مشاهدات در مجموعه داده آموزش تشکیل می‌شود. مشاهدات در مجموعه داده آزمون به هیچ‌وجه در طی فرایند آموزش استفاده نمی‌شود. عموماً ۱۰ الی ۳۰ درصد از کل مجموعه داده آموزش برای هر دو مجموعه داده اعتباردهی و آزمون استفاده می‌شود.

اقلام تعهدی کل

گام نخست در استفاده از مدل‌های ارقام تعهدی، تعیین ارقام تعهدی کل است. در ادبیات، دو شیوه برای محاسبه ارقام تعهدی کل مورداستفاده قرار گرفته است: رویکرد ترازنامه‌ای و رویکرد صورت جریان وجوه نقد. در صورت استفاده از رویکرد ترازنامه‌ای که رویکرد پرکاربردتر نیز می‌باشد، ارقام تعهدی کل از تغییرات در سرمایه در گردش به دست می‌آید:

$$TACC_t = \Delta CA - \Delta CL - \Delta Cash + \Delta STDEBT - Dep \quad \text{رابطه (۴)}$$

که در آن، $TACC$ ارقام تعهدی کل، ΔCA تغییرات در دارایی‌های جاری، ΔCL تغییر در بدهی‌های جاری، $\Delta Cash$ تغییر در سطح موجودی نقد، $\Delta STDEBT$ تغییر در حصة جاری بدهی‌های بلندمدت و سایر بدهی‌های کوتاه‌مدت در سرفصل بدهی‌های جاری و DEP هزینه استهلاک می‌باشد. بنابر رویکرد دوم یعنی رویکرد صورت جریان وجوه نقد، جریان‌های نقدی عملیاتی ناشی از عملیات در حال تداوم از سود قبل از ارقام غیرمترقبه کسر می‌شود:

$$TACC_t = Earnings_t - CFO_t \quad \text{رابطه (۵)}$$

که در آن، $Earnings_t$ سود خالص در دوره t و CFO_t جریان نقد عملیاتی در دوره t می‌باشد. باوجوداینکه استفاده از رویکرد ترازنامه‌ای رایج‌تر است، بااین حال موقعیت‌هایی نیز وجود دارد که این رویکرد به معیاری سویه‌دار از ارقام تعهدی کل منجر می‌شود. Collins & Hribar (2002) ادعا می‌کنند که رویکرد صورت جریان وجوه نقد نسبت به رویکرد دیگر ترجیح دارد زیرا رویکرد ترازنامه‌ای ممکن است به خطای اندازه‌گیری در مورد ترکیب‌های تجاری و تسعیر ارز منجر شود. این مهم به این دلیل است که رویدادهای غیرعملیاتی بر حساب‌های دارایی‌ها و بدهی‌های جاری اثر می‌گذارند بدون اینکه بر سود اثر گذاشته باشند. بااین حال، دیدگاه مزبور می‌تواند منجر به نوعی ارزیابی غیرواقع‌بینانه عملکرد مدل بیانجامد. به عبارتی، از آنجایی که متغیر جریان نقد عملیاتی به‌عنوان متغیر کنترلی در مدل‌ها در نظر گرفته می‌شود این امر می‌تواند به ارزیابی بالاتر از

واقع عملکرد مدل‌ها بیانجامد. لذا در این پژوهش، از هر دو رویکرد برای برآزش مدل پیش‌بینی ارقام تعهدی استفاده می‌شود.

DeFond (2010) اعتقاد دارد که مدل‌های بسیاری برگرفته از مدل Jones (1991) جهت پیش‌بینی ارقام تعهدی ارائه گردیده‌اند اما اغلب این مدل‌ها در ادبیات دوام نیافته‌اند. از نظر ایشان، دو مدلی که می‌توان از این موضوع آن‌ها را مستثنی نمود مدل‌های ارائه‌شده توسط Dechow & Dichev (2002) و Francis et. Al. (2005) می‌باشد. از آنجایی که مدل Francis et. Al. (2005) شکل بسط‌یافته‌ای از مدل Dechow & Dichev (2002) می‌باشد در این پژوهش از متغیرهای مورد استفاده در پیش‌بینی ارقام تعهدی توسط Francis et. Al. (2005) استفاده شده است. مدل پیش‌بینی ارقام تعهدی در پژوهش مزبور به شرح رابطه ۶ می‌باشد.

$$TACC_t = \alpha + \beta_1 CFO_{t-1} + \beta_2 CFO_t + \beta_3 CFO_{t+1} + \beta_4 \Delta Rev_t + \beta_5 PPE_t + \varepsilon_t \quad \text{رابطه (۶)}$$

که در این رابطه، CFO_t نشان‌دهنده جریان نقد عملیاتی در سال t ، CFO_{t-1} جریان نقد عملیاتی در سال $t-1$ ، CFO_{t+1} جریان نقد عملیاتی برای سال $t+1$ ، ΔRev تغییرات در درآمد نسبت به درآمد سال گذشته و PPE ناخالص دارایی‌های ثابت مشهود می‌باشد. همچنین کلیه متغیرها از طریق تقسیم بر جمع کل دارایی‌های سال گذشته، مقیاس‌زدایی می‌شوند.

علاوه بر این، در مقایسه مدل‌ها به عملکرد مدل ارائه‌شده توسط Ball & Shivakumar (2006) به شرح رابطه ۷ نیز پرداخته خواهد شد.

$$TACC_t = \alpha + \beta_1 CFO_{t-1} + \beta_2 CFO_t + \beta_3 CFO_{t+1} + \beta_4 \Delta Rev_t + \beta_5 PPE_t + \beta_6 DCFO_t + \beta_7 DCFO_t \times CFO_t + \varepsilon_t \quad \text{رابطه (۷)}$$

که در این رابطه، $DCFO_t$ یک متغیر ساختگی بوده که در صورت منفی بودن CFO_t ، مقدار ۱ و در صورت مثبت بودن آن، مقدار صفر را به خود می‌گیرد.

روش اصلی مورد استفاده در این پژوهش جهت ایجاد مدل کیفیت سود، الگوریتم مدیریت گروهی داده‌ها می‌باشد. هر ترکیب دوتایی از متغیرهای مستقل تحقیق، منجر به ایجاد یک نورون در لایه اول می‌شوند؛ بنابراین، با توجه به استفاده از ۵ متغیر توضیحی و یک عرض از مبدأ، تعداد نورون قابل تشکیل در لایه اول، تعداد ۱۵ نورون خواهد بود. با این حال، جهت اجتناب از پیچیدگی ساختار مدل، بیشینه تعداد نورون‌ها در هر لایه به ۵ نورون محدود خواهد شد. سپس، هر نورونی که میزان خطای برآورد در آن، از α درصد فاصله بین خطای بیشینه و خطای کمینه، کمتر باشد به عنوان نورون برگزیده جهت انتقال به لایه بعد گزینش می‌شود. به عبارتی، در صورتی که خطای نورون مورد بررسی از خطای آستانه‌ای به شرح رابطه ۸ کمتر باشد، به مرحله بعد منتقل می‌گردد (Dong, 2019).

$$\text{خطای آستانه‌ای} = \alpha e_{\min} + (1 - \alpha)e_{\max} \quad \text{رابطه (۸)}$$

در این مدل، عدد α فشار انتخاب، e_{\min} کمترین خطا در بین نورون‌های لایه و e_{\max} بیشترین خطا در بین نورون‌های لایه را نشان می‌دهد. لذا هر چه عدد α بالاتر در نظر گرفته شود به منزله سخت‌گیرانه‌تر بودن گزینش از بین نورون‌های یک لایه جهت ورود به لایه بعدی می‌باشد. همچنین فشار انتخاب (مقدار α) معادل ۶۰ درصد تعیین گردید به این معنا که در هر لایه، نورون‌هایی که در بازه ۴۰ درصد برتر از نظر خطای برآورد قرار می‌گیرند واجد شرایط انتقال به لایه بعد تلقی می‌شوند.

در لایه‌های بعدی نیز این محاسبات تکرار شده و زمانی فرایند متوقف می‌شود که روند کاهشی میزان خطا از یک لایه به لایه بعد متوقف شود.

جامعه و نمونه آماری

جامعه آماری این تحقیق شامل شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران است. قلمرو زمانی تحقیق نیز سال‌های ۱۳۸۵ تا ۱۳۹۷ می‌باشد. برای انتخاب نمونه شرایط زیر در نظر گرفته شده است:

۱. داده‌ها و اطلاعات موردنیاز برای محاسبه متغیرها در مورد شرکت‌ها برای دوره زمانی موردتحقیق در دسترس باشد.

۲. شرکت جزو شرکت‌های سرمایه‌گذاری، واسطه‌گری مالی و بانک‌ها نباشد.

۳. شرکت حداقل برای یک سال قبل و یک سال پس از قلمرو زمانی تحقیق در بورس پذیرفته‌شده باشد.

۴. پایان سال مالی شرکت، پایان اسفندماه باشد و طی دوره پژوهش، تغییر سال مالی نداشته باشد.

با اعمال شرایط فوق، ۱۶۴ شرکت به‌عنوان نمونه برای پیش‌بینی سطح اقلام تعهدی انتخاب گردید و تعداد مشاهدات برابر با ۲۱۳۲ مشاهده در طی ۱۳ سال می‌باشد.

جهت برازش مدل‌های رگرسیون از نرم‌افزار ایویوز نسخه ۱۰ و جهت مدل‌سازی روابط متغیرها بر اساس الگوریتم مدیریت گروهی داده‌ها از نرم‌افزار متلب ۲۰۱۸ استفاده گردیده است.

یافته‌ها

آماره‌های توصیفی داده‌های شرکت‌های عضو نمونه به شرح جدول ۲ می‌باشد. نتایج نشان می‌دهد که تغییرات درآمد شرکت‌ها به‌طور متوسط معادل ۱۶ درصد جمع دارایی‌های سال قبل شرکت‌ها می‌باشد. همچنین، بیشترین افت درآمد گزارش شده در بین مشاهدات معادل ۲,۱۶ برابر جمع دارایی‌های سال قبل و حداکثر رشد درآمد معادل ۲۱,۹۸ برابر جمع دارایی‌ها بوده است. علاوه بر این، دارایی‌های ثابت شرکت‌ها بین صفر تا ۳۳ درصد جمع دارایی‌های سال قبل آن‌ها را شکل داده است.

جدول ۲. آماره‌های توصیفی

متغیر	$\Delta Rev/TA$	PPE/TA	$CFO(t)/TA$	$CFO(t-1)/TA$	$CFO(t+1)/TA$
-------	-----------------	----------	-------------	---------------	---------------

۰,۱۶	۰,۰۹	۰,۱۴	۰,۳۰	۰,۱۶	میانگین
۰,۱۲	۰,۱۰	۰,۱۱	۰,۲۴	۰,۱۰	میانه
۰,۲۳	۱,۴۸	۰,۱۷	۰,۳۳	۰,۵۸	انحراف معیار
-۲,۵۶	-۶۷,۹۵	-۲,۲۳	۰,۰۰	-۲,۱۶	حداقل
۱,۶۹	۰,۸۲	۱,۱۵	۷,۶۰	۲۱,۹۸	حداکثر

در انتخاب مدل رگرسیون، جهت تعیین تناسب رگرسیون تابلویی یا تجمیعی از آزمون F لیمر استفاده شده است. نتایج حاصل از این آزمون در مورد هر چهار رویکرد رگرسیون مورد آزمون، حاکی از احتمال کمتر از ۱ درصد بوده و در نتیجه رگرسیون تابلویی گزینش می‌گردد. در ادامه، نتایج حاصل از آزمون هاسمن حاکی از ارجحیت رگرسیون اثرات ثابت نسبت به رگرسیون اثرات تصادفی در هر چهار مدل رگرسیون می‌باشد. همچنین جهت بررسی وجود هم‌خطی بین متغیرهای توضیحی از آزمون عامل تورم واریانس استفاده شده است. نتایج این آزمون نیز مقدار عامل تورم واریانس برای کلیه متغیرهای توضیحی را کمتر از ۲,۵۰ نشان می‌دهد که حاکی از عدم هم‌خطی بین متغیرهای توضیحی است.

جدول ۳. آزمون‌های انتخاب مدل پژوهش

آزمون	مؤلفه‌های آزمون	رویکرد جریان نقدی		رویکرد ترازنامه‌ای	
		رگرسیون خطی	رگرسیون خطی منقطع	رگرسیون خطی	رگرسیون خطی منقطع
آزمون چاو	آماره آزمون (F) لیمر	۱,۴۵	۱,۳۵	۱,۷۱	۱,۶۱
	احتمال	۰,۰۰	۰,۰۰	۰,۰۰	۰,۰۰
	نتیجه	تابلویی	تابلویی	تابلویی	تابلویی
آزمون هاسمن	آماره آزمون	۴۲,۴	۳۱,۹	۵۸,۶	۴۱,۰۲
	احتمال	۰,۰۰	۰,۰۰	۰,۰۰	۰,۰۰
	نتیجه	اثرات ثابت	اثرات ثابت	اثرات ثابت	اثرات ثابت

ضرایب رگرسیون برای مدل رگرسیون خطی در جدول ۴ ارائه شده است. در این جدول، نتایج برازش هر مدل یک‌بار با استفاده از رویکرد جریان نقدی و یک‌بار استفاده از رویکرد ترازنامه‌ای در محاسبه اقلام تعهدی ارائه گردیده است. علامت ضریب متغیر تغییرات فروش و دارایی‌های ثابت مشهود مطابق انتظار، به ترتیب مثبت و منفی می‌باشد. نتایج حاصل از پیاده‌سازی هر دو رویکرد رگرسیون خطی و رگرسیون خطی منقطع جهت برآورد اقلام تعهدی نشان می‌دهد که ضریب تعیین این مدل‌ها در استفاده از رویکرد ترازنامه‌ای (به ترتیب ۹ درصد و ۱۶ درصد) نسبت به استفاده از رویکرد جریان نقدی (به ترتیب ۳۰ درصد و ۳۶ درصد) کاهش معناداری را نشان می‌دهد. همچنین میزان خطای برآورد مدل‌های مزبور (بر اساس میانگین مربعات خطا) در استفاده از رویکرد ترازنامه‌ای، به میزان معناداری بالاتر از رویکرد جریان نقدی می‌باشد.

جدول ۴. نتایج مدل‌های رگرسیون

رویکرد ترازنامه‌ای		رویکرد جریان نقدی		آماره‌های رگرسیون
رگرسیون خطی منقطع	رگرسیون خطی	رگرسیون خطی منقطع	رگرسیون خطی	
۰,۲۷۸ (۰,۲۵۸)	۰,۳۰۸ (۰,۲۲۹)	-۰,۲۰۱ (۰,۲۵۷)	-۰,۱۷۰ (۰,۳۵۷)	$1/TA$
۰,۰۵۲ (۰,۰۰۰)	۰,۰۵۲ (۰,۰۰۰)	۰,۰۶۶ (۰,۰۰۰)	۰,۰۶۷ (۰,۰۰۰)	$\Delta Rev/TA$
-۰,۱۴۵ (۰,۰۰۰)	-۰,۱۱۵ (۰,۰۰۰)	-۰,۰۸۵ (۰,۰۰۰)	-۰,۰۶۳ (۰,۰۰۰)	$\frac{PPE}{TA}$
-۰,۱۹۹ (۰,۰۰۰)	-۰,۳۸۵ (۰,۰۰۰)	-۰,۵۰۶ (۰,۰۰۰)	-۰,۶۷۴ (۰,۰۰۰)	$\frac{CFO_t}{TA}$
-۰,۰۰۱ (۰,۸۰۲)	۰,۰۰۰ (۰,۹۲۴)	۰,۰۰۱ (۰,۷۵۰)	۰,۰۰۱ (۰,۷۲۹)	$\frac{CFO_{t-1}}{TA}$
۰,۲۳۲ (۰,۰۰۰)	۰,۲۲۲ (۰,۰۰۰)	۰,۲۷۰ (۰,۰۰۰)	۰,۲۶۸ (۰,۰۰۰)	$\frac{CFO_{t+1}}{TA}$
-۰,۰۴۹ (۰,۰۰۴)		-۰,۰۰۹ (۰,۴۳۸)		DCFO

رویکرد ترازنامه‌ای		رویکرد جریان نقدی		آماره‌های رگرسیون
رگرسیون خطی منقطع	رگرسیون خطی	رگرسیون خطی منقطع	رگرسیون خطی	
-۱,۰۴۵ (۰,۰۰۰)		-۰,۷۶۹ (۰,۰۰۰)		$DCFO \times \frac{CFO_t}{TA}$
۰,۱۶۳	۰,۰۹۳	۰,۳۶۶	۰,۳۰۹	ضریب تعیین
۰,۱۶۰	۰,۰۹۰	۰,۳۶۳	۰,۳۰۶	ضریب تعیین تعدیل شده
۰,۲۰۰	۰,۲۰۸	۰,۱۴۴	۰,۱۵۰	خطای معیار رگرسیون
۰,۰۴۰	۰,۰۴۳	۰,۰۲۰۶	۰,۰۲۲۵	میانگین مربعات باقیمانده‌ها
۰,۲۴۳	۰,۲۳۵	۰,۲۴۴	۰,۲۴۰	جذر میانگین مربعات خطا (داده‌های ارزیابی)
۰,۱۴۳	۰,۱۴۲	۰,۱۳۸	۰,۱۳۸	میانگین قدرمطلق خطا (داده‌های ارزیابی)

استفاده از رویکرد الگوریتم مدیریت گروهی داده‌ها، این فرصت را به وجود می‌آورد که ضمن حفظ قابلیت بیان مدل در قالب رابطه ریاضی چندجمله‌ای، روابط غیرخطی بین متغیرها نیز امکان شناسایی را داشته باشد. در این پژوهش، حداکثر تعداد لایه‌های این الگوریتم ۵ لایه و حداکثر تعداد نورون‌های هر لایه نیز معادل ۵ نورون تنظیم گردید. این امر در راستای جلوگیری از پیچیدگی زیاد از حد نتیجه الگوریتم صورت می‌گیرد. همچنین فشار انتخاب معادل ۶۰ درصد تعیین گردید به این معنا که در هر لایه، نورون‌هایی که در بازه ۴۰ درصد برتر از منظر میانگین مربعات خطا قرار می‌گیرند واجد شرایط انتقال به لایه بعد تلقی می‌شوند. انتخاب داده‌های مورد استفاده جهت آموزش الگوریتم مدیریت گروهی داده‌ها به دو شکل صورت می‌گیرد. در حالت اول، داده‌های ۱۱ سال ابتدایی دوره پژوهش، جهت برآورد مدل انتخاب شده و داده‌های دو سال پایانی جهت آزمون عملکرد مدل مورد استفاده قرار گرفتند. نتایج حاصل از این بررسی به شرح جدول ۵ می‌باشد.

جدول ۵. نتایج الگوریتم مدیریت گروهی داده‌ها با تفکیک غیرتصادفی داده‌ها

مجموعه داده	رویکرد جریان نقدی	رویکرد ترازنامه‌ای
-------------	-------------------	--------------------

کل داده‌ها	داده ارزیابی	داده آموزش	کل داده‌ها	داده ارزیابی	داده آموزش	
۰,۰۳۹۸۸	۰,۰۵۵۰۵	۰,۰۳۷۱۲	۰,۰۲۶۵۰	۰,۰۶۰۲۱	۰,۰۲۰۳۷	میانگین مربعات خطا (MSE)
۰,۱۹۹۷	۰,۲۳۴۶۴	۰,۱۹۲۶۷	۰,۱۶۲۸	۰,۲۴۵۳۹	۰,۱۴۲۷۳	جذر میانگین مربعات خطا (RMSE)
-۰,۰۰۵۷۴	-۰,۰۳۴۰۵	-۰,۰۰۰۶۰	-۰,۰۰۵۸۰	-۰,۰۳۸۴۰	۰,۰۰۰۱۲	میانگین خطا
۰,۱۹۹۶۷	۰,۲۳۲۵۱	۰,۱۹۲۷۲	۰,۱۶۲۷۳	۰,۲۴۲۷۸	۰,۱۴۲۷۷	انحراف معیار خطا
۰,۴۲	۰,۱۸	۰,۴۷	۰,۵۴	۰,۳۰	۰,۶۱	ضریب همبستگی بین خروجی مدل و متغیر هدف (R)

همان‌گونه که مشاهده می‌گردد در استفاده از رویکرد جریان نقدی، ضریب همبستگی بین خروجی مدل و متغیر هدف برای داده‌های آموزش و ارزیابی به ترتیب معادل ۶۱ و ۳۰ درصد می‌باشد. همچنین، جذر میانگین مربعات خطا بین داده‌های گروه آموزش معادل ۰,۱۴۳ و در بین داده‌های ارزیابی معادل ۰,۲۴۵ می‌باشد. در این الگوریتم، فرایند ایجاد لایه‌ها در لایه پنجم متوقف شده و مدل ایجاد می‌گردد. همچنین در صورت استفاده از رویکرد ترازنامه‌ای محاسبه اقلام تعهدی، ضریب همبستگی هم در ارتباط با داده‌های آموزش و هم در ارتباط با داده‌های ارزیابی به میزان معناداری کاهش و میانگین مربعات خطا نیز به میزان معناداری افزایش می‌یابد.

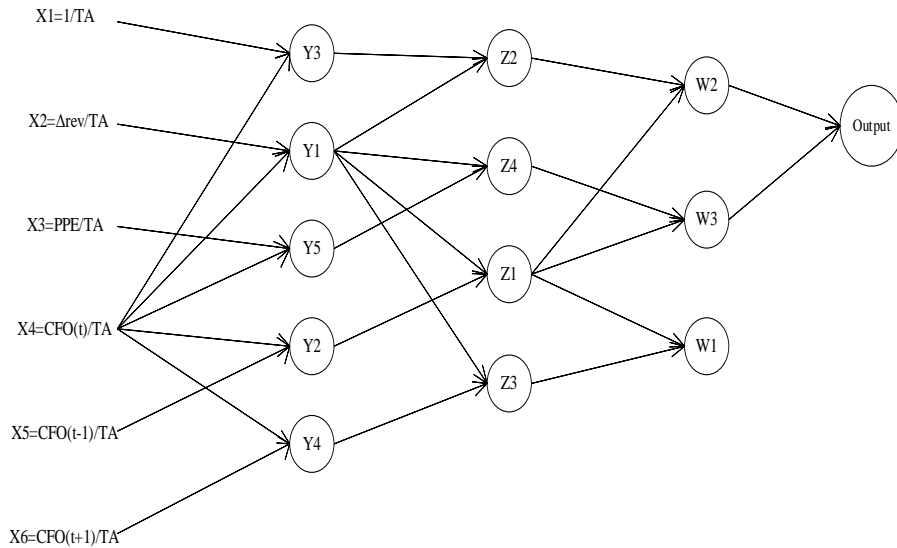
در نوبت دوم، تفکیک مشاهدات به داده‌های آموزش و داده‌های ارزیابی به صورت تصادفی صورت می‌گیرد. در این حالت، جهت حفظ قابلیت مقایسه با آزمون پیشین، تعداد ۱۸۰۴ مشاهده از کل ۲۱۳۲ مشاهده موجود (نسبت ۱۱ به ۱۳) به صورت تصادفی گزینش شده و جهت آموزش مورد استفاده قرار می‌گیرد. باقیمانده مشاهدات جهت ارزیابی مدل خروجی استفاده می‌شود. در این حالت، نتایج حاصل از مدل‌سازی به شرح جدول ۶ می‌باشد.

جدول ۶. نتایج الگوریتم مدیریت گروهی داده‌ها با تفکیک تصادفی داده‌ها

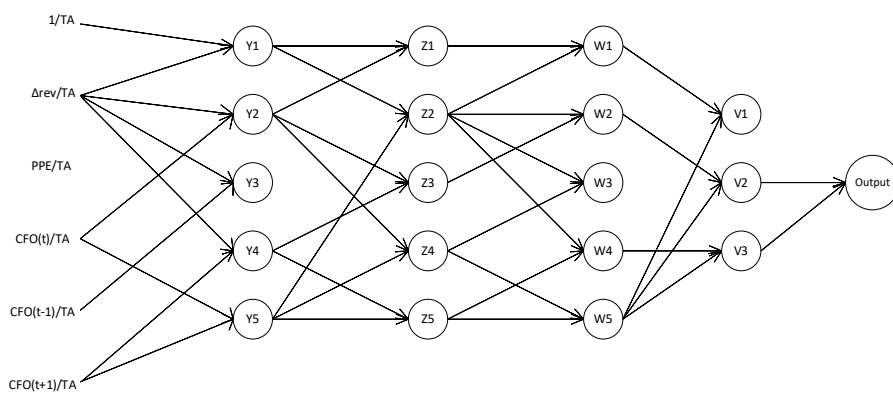
رویکرد ترازنامه‌ای			رویکرد جریان نقدی			مجموعه داده
کل داده‌ها	داده ارزیابی	داده آموزش	کل داده‌ها	داده ارزیابی	داده آموزش	

رویکرد ترازنامه‌ای			رویکرد جریان نقدی			مجموعه داده
کل داده‌ها	داده ارزیابی	داده آموزش	کل داده‌ها	داده ارزیابی	داده آموزش	
۰,۰۳۹۲۱	۰,۰۲۷۳۰	۰,۰۴۱۳۸	۰,۰۲۵۵۳	۰,۰۲۹۷۱	۰,۰۲۴۷۷	میانگین مربعات خطا (MSE)
۰,۱۹۸۰۲	۰,۱۶۵۲۳	۰,۲۰۳۴	۰,۱۵۹۷۹	۰,۱۷۲۳۸	۰,۱۵۷۳۹	جذر میانگین مربعات خطا (RMSE)
۰,۰۰۱۳۸ -	-۰,۰۰۳۸۴	-۰,۰۰۰۹۳	-۰,۰۰۰۶۷	-۰,۰۰۱۵۵	-۰,۰۰۰۵۱	میانگین خطا
۰,۱۹۸۰۶	۰,۱۶۵۴۴	۰,۲۰۳۴۷	۰,۱۵۹۸۲	۰,۱۷۲۶۳	۰,۱۵۷۴۳	انحراف معیار خطا
۰,۴۴	۰,۳۱	۰,۴۵	۰,۵۶	۰,۴۴	۰,۵۸	ضریب همبستگی بین خروجی مدل و متغیر هدف (R)

مقایسه نتایج دو آزمون فوق در رویکرد الگوریتم گروهی داده‌ها حاکی از این است که زمانی که داده‌های مورد استفاده جهت آموزش شبکه به صورت تصادفی می‌باشد قدرت تعمیم نتیجه مدل به داده‌های آزمون بالاتر است نسبت به زمانی که مشاهدات مربوط به سال‌های ابتدایی جهت آموزش شبکه استفاده شود. یکی از دلایل احتمالی چنین تفاوتی، می‌تواند اثرگذاری برخی عوامل متغیر بین سال‌های مختلف بر روابط موجود در مدل باشد که کنترلی برای آن‌ها در نظر گرفته نشده است. مدل‌های به دست آمده از پیاده‌سازی الگوریتم مدیریت گروهی داده‌ها با تفکیک تصادفی داده‌ها به داده‌های آموزش و ارزیابی و با دو رویکرد جریان نقدی و ترازنامه‌ای در محاسبه ارقام تعهدی به شرح شکل‌های ۳ و ۴ می‌باشد.



شکل ۳: خروجی حاصل از الگوریتم مدیریت گروهی داده‌ها (رویکرد جریان نقدی و تفکیک تصادفی)



شکل ۴: خروجی حاصل از الگوریتم مدیریت گروهی داده‌ها (رویکرد ترازنامه‌ای و تفکیک تصادفی)

در شکل‌های فوق، هر یک از متغیرهای Y یک چندجمله‌ای درجه دوم از دو متغیر X بوده که آن نورون را شکل می‌دهد. این موضوع برای لایه‌های بعدی یعنی متغیرهای Z ، W و Output نیز صادق بوده و هر یک از متغیرهای مزبور، یک چندجمله‌ای

کولموگروف-گابور از یک زوج متغیر مرحله قبل می‌باشد. روابط ۹ تا ۱۳، روابط متغیرهای شکل ۲ را نشان می‌دهد.^۱

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \end{bmatrix} = \begin{bmatrix} 0.03 & 0.15 & -0.54 & 0.01 & 0.24 & 0.02 \\ 0.02 & -0.66 & 0.41 & 0.20 & 0.01 & 0.12 \\ -0.08 & 0.78 & 0.76 & -0.41 & 0.26 & -7.26 \\ 0.03 & -0.66 & 0.24 & 0.27 & 0.05 & -0.07 \\ 0.07 & -0.09 & -0.55 & 0.01 & 0.27 & 0.20 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ x_2 & x_4 & x_1 & x_4 & x_3 \\ x_4 & x_5 & x_4 & x_6 & x_4 \\ x_2^2 & x_4^2 & x_1^2 & x_4^2 & x_3^2 \\ x_4^2 & x_5^2 & x_4^2 & x_6^2 & x_4^2 \\ x_2x_4 & x_4x_5 & x_1x_4 & x_4x_6 & x_3x_4 \end{bmatrix} \quad \text{رابطه (۹)}$$

نتیجه رابطه فوق یک ماتریس با پنج سطر و پنج ستون می‌باشد. در این رابطه، هر یک از مقادیر Y به صورت یک چندجمله‌ای از جمع آرایه‌های سطر متناظر خود در ماتریس حاصله به دست می‌آید. به عنوان نمونه، مقدار Y_1 به صورت رابطه ۱۰ نمایش می‌یابد.

$$Y_1 = 0.03 + 0.15x_2 - 0.54x_4 + 0.01x_2^2 + 0.24x_4^2 + 0.02x_2x_4 \quad \text{رابطه (۱۰)}$$

۱. جهت اجتناب از افزایش صفحات مقاله، نمودارهای مربوط به خروجی الگوریتم مدیریت گروهی داده‌ها تنها برای حالت تفکیک تصادفی مشاهدات ارائه گردیده است. همچنین روابط ریاضی تنها برای یکی از رویکردها (رویکرد جریان نقدی و حالت تفکیک تصادفی مشاهدات) ارائه گردیده است.

$$Z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} 0.00 & 0.60 & 0.65 & 0.30 & 0.68 & -1.08 \\ 0.00 & 0.86 & 0.24 & 0.04 & -0.77 & 0.70 \\ 0.00 & 0.62 & 0.57 & 0.29 & -0.60 & 0.21 \\ 0.00 & 0.97 & 0.06 & -0.61 & 2.29 & -1.66 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ Y_1 & Y_1 & Y_1 & Y_1 \\ Y_2 & Y_3 & Y_4 & Y_5 \\ Y_1^2 & Y_1^2 & Y_1^2 & Y_1^2 \\ Y_2^2 & Y_3^2 & Y_4^2 & Y_5^2 \\ Y_1Y_2 & Y_1Y_3 & Y_1Y_4 & Y_1Y_5 \end{bmatrix} \quad \text{رابطه (۱۱)}$$

$$W = \begin{bmatrix} W_1 \\ W_2 \\ W_3 \end{bmatrix} = \begin{bmatrix} -0.00 & 0.62 & 0.42 & 1.91 & 1.73 & -3.67 \\ 0.00 & 1.00 & -0.01 & 0.37 & -0.33 & -0.03 \\ 0.00 & 0.97 & 0.03 & -0.82 & -1.22 & 2.05 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ X_1 & X_1 & X_1 \\ X_3 & X_2 & X_4 \\ X_1^2 & X_1^2 & X_1^2 \\ X_3^2 & X_2^2 & X_4^2 \\ X_1X_3 & X_1X_2 & X_1X_4 \end{bmatrix} \quad \text{رابطه (۱۲)}$$

$$\text{Output} = [0.00 \quad 0.39 \quad 0.59 \quad -21.04 \quad -25.97 \quad 47.09] \begin{bmatrix} 1 \\ W_2 \\ W_3 \\ W_2^2 \\ W_3^2 \\ W_2W_3 \end{bmatrix}$$

$$= 0.00 + 0.39W_2 + 0.59W_3 - 21.04W_2^2 - 25.97W_3^2 + 47.09W_2W_3 \quad \text{رابطه (۱۳)}$$

همان‌گونه که در شکل‌های ۳ و ۴ مشخص می‌باشد در استفاده از رویکرد جریان نقدی، چندجمله‌ای‌هایی که یکی از زوج متغیر ورودی آن، متغیر جریان نقدی دوره جاری می‌باشد در ایجاد مدل بیشترین نقش را دارند در حالی که در استفاده از رویکرد ترازنامه‌ای، چندجمله‌ای‌هایی که یکی از زوج متغیر ورودی، متغیر تغییرات در درآمد شرکت می‌باشد جهت ورود به مراحل بعدی گزینش می‌گردند. همان‌گونه که پیش‌تر بیان گردید (Collins & Hribar, 2002) استفاده از رویکرد جریان نقدی را نسبت به رویکرد ترازنامه‌ای ارجح می‌دانند. با این حال، به نظر می‌رسد در این تحلیل، یک توضیح احتمالی این باشد که استفاده از متغیرهای مشترک به عنوان متغیر توضیحی و متغیر مورد استفاده در محاسبه متغیر وابسته، باعث نمایش بهتر از واقع عملکرد مدل‌های مبتنی بر رویکرد جریان نقدی می‌گردد.

نتایج حاصل از استفاده از رویکرد شبکه عصبی پرسپترون چندلایه نیز به شرح جدول ۷ می‌باشد. جهت آموزش این شبکه، از الگوریتم لاونبرگ-مارکوات و جهت بررسی عملکرد شبکه، از متغیر میانگین مربعات خطا استفاده می‌شود. همچنین تعداد دفعات نمونه‌گیری تصادفی جهت آموزش شبکه، ۱۰۰۰ بار تعیین می‌گردد. در این آزمون، شبکه‌هایی با یک یا دولایه مخفی و تعداد ۶، ۱۲ و ۱۸ نورون در هر لایه آزمون شدند. بهترین نتیجه در استفاده از رویکرد جریان نقدی محاسبه اقلام تعهدی، در حالت یک‌لایه مخفی و ۱۲ نورون و در استفاده از رویکرد ترازنامه‌ای، در حالت دولایه مخفی و ۶ نورون در هر لایه حاصل گردید.

جدول ۷. نتایج شبکه عصبی مصنوعی پرسپترون

رویکرد ترازنامه‌ای				رویکرد جریان نقدی				
کل داده‌ها	داده آزمون	داده اعتبارسنجی	داده آموزش	کل داده‌ها	داده آزمون	داده اعتبارسنجی	داده آموزش	
۰,۰۲۶۵ ۲	۰,۰۲۰۰۸	۰,۰۳۴۷۸	۰,۰۲۶۲۹	۰,۰۲۳۴۷	۰,۰۱۸۲۸	۰,۰۲۵۲۴	۰,۰۲۳۸۹	میانگین مربعات خطا (MSE)
۰,۱۶۲۸ ۵	۰,۱۴۱۶۸	۰,۱۸۶۵	۰,۱۶۲۱۵	۰,۱۵۳۱۸	۰,۱۳۵۲۱	۰,۱۵۸۸۶	۰,۱۵۴۵۷	جذر میانگین مربعات خطا (RMSE)

رویکرد ترازنامه‌ای				رویکرد جریان نقدی				
کل داده‌ها	داده آزمون	داده اعتبارسنجی	داده آموزش	کل داده‌ها	داده آزمون	داده اعتبارسنجی	داده آموزش	
۰,۰۰۲۶ ۰	۰,۰۰۷۰۲	۰,۰۳۱۷۹	-۰,۰۰۱۶۰	۰,۰۰۴۵۸	۰,۰۰۸۷۳	۰,۰۱۰۹۶	۰,۰۰۳۲۷	میانگین خطا
۰,۱۶۲۸ ۷	۰,۱۴۱۸۴	۰,۱۸۴۲	۰,۱۶۲۱۹	۰,۱۵۳۱۵	۰,۱۳۵۲۵	۰,۱۵۸۸۶	۰,۱۵۴۵۸	انحراف معیار خطا
۰,۶۷	۰,۳۰	۰,۵۲	۰,۷۱	۰,۶۱	۰,۷۰	۰,۶۳	۰,۶۰	ضریب همبستگی بین خروجی مدل و متغیر هدف (R)

نتایج به دست آمده از چهار الگوی فوق، حاکی از برتری رگرسیون خطی منقطع نسبت به رگرسیون خطی از نظر میانگین مربعات خطا و همچنین برتری هر دو رویکرد الگوریتم مدیریت گروهی داده‌ها و شبکه عصبی پرسپترون چندلایه نسبت به دو مدل رگرسیون می‌باشد. بررسی نتیجه مدل‌سازی بر اساس الگوریتم مدیریت گروهی داده‌ها، مطابق انتظار نشان‌دهنده آن است که نحوه گزینش داده‌ها جهت برازش مدل، در عملکرد مدل از منظر خطا در داده‌های ارزیابی نقش معناداری دارد. به عبارتی، زمانی که داده‌های دوره ۱۱ ساله ابتدایی جهت برازش مدل و داده‌های دو سال پایانی جهت ارزیابی استفاده می‌شود عملکرد مدل از نظر میانگین مربعات خطا به میزان معناداری ضعیف‌تر از زمانی است که داده‌های مورد استفاده جهت برازش مدل، به صورت تصادفی از میان داده‌های دوره ۱۳ ساله انتخاب می‌شوند.

بحث و نتیجه‌گیری

مدل‌های رگرسیون پیش‌بینی اقلام تعهدی غیراختیاری، به دلیل سادگی تبیین، امکان بررسی مبانی نظری روابط بین متغیرهای مستقل و اقلام تعهدی اختیاری را به محققان می‌دهد. با این حال، تحقیقات متعدد صورت گرفته نشان می‌دهد که این مدل‌ها از منظر خطای آماری، عملکرد مناسبی ندارند. یکی از دلایل این امر، فرض خطی بودن روابط بین

متغیرهای مستقل و ارقام تعهدی می‌باشد. در راستای مقابله با این محدودیت، رویکرد استفاده از شبکه‌های عصبی چندلایه جهت پیش‌بینی ارقام تعهدی توسط چندین پژوهش مورد استفاده قرار گرفت. با وجود عملکرد مناسب‌تر این دسته از مدل‌ها نسبت به مدل‌های رگرسیونی از منظر خطاهای آماری، یکی از محدودیت‌های اساسی این مدل‌ها، مبتنی بودن بر رویکرد جعبه سیاه می‌باشد. به عبارتی، امکان بیان روابط بین متغیرهای مستقل و وابسته در قالب یک رابطه چند متغیره وجود ندارد و لذا امکان توسعه آتی ادبیات تحقیق در حوزه کیفیت ارقام تعهدی بر مبنای نتایج شبکه‌های عصبی چندلایه را محدود می‌سازد. در این پژوهش، با استفاده از الگوریتم مدیریت گروهی داده‌ها تلاش شد تا ضمن ایجاد امکان بیان چندجمله‌ای روابط بین متغیرها، خطاهای آماری را نیز در سطح قابل قبولی حفظ نماید. در این راستا، نتایج حاصل از الگوریتم مدیریت گروهی داده‌ها با مدل‌های رگرسیون خطی و رگرسیون خطی منقطع و همچنین شبکه عصبی چندلایه مقایسه می‌گردد. مشابه یافته‌های Hoglund (2012) و قادری و همکاران (۱۳۹۷)، عملکرد بهتری در شبکه عصبی پرسپترون چندلایه در مقایسه با رگرسیون خطی مشاهده می‌شود. همچنین رگرسیون خطی منقطع عملکرد مناسب‌تری نسبت به رگرسیون خطی نشان می‌دهد. (مطابق یافته‌های Hoglund, (2012)). علاوه بر این، عملکرد الگوریتم مدیریت گروهی داده‌ها در حالت تفکیک غیر تصادفی مشاهدات، نزدیک به رگرسیون خطی و خطی منقطع می‌باشد ولی در مقایسه با شبکه عصبی پرسپترون چندلایه خطای بالاتری دارد. باین حال، با گزینش تصادفی داده‌ها جهت برازش مدل در الگوریتم مدیریت گروهی داده‌ها، عملکرد این مدل به مراتب بهبود یافته و عملکرد بسیار بهتری نسبت به مدل‌های رگرسیون دارد. همچنین عملکرد این الگو در صورت تفکیک تصادفی مشاهدات به داده‌های آموزش و ارزیابی، عملکردی نزدیک به شبکه عصبی پرسپترون نشان می‌دهد. همچنین، در کلیه موارد، استفاده از رویکرد جریان نقدی در محاسبه ارقام تعهدی منجر به خطای برآورد کمتری نسبت به رویکرد ترازنامه‌ای می‌شود. باین حال، این امر را نمی‌توان به‌طور قطع به تناسب بیشتر این مدل‌ها با واقعیت نسبت داد بلکه این احتمال وجود دارد که اتکا به دیدگاه Collins &

Hribar (2002) در خصوص ارجحیت رویکرد جریان نقدی نسبت به رویکرد ترازنامه‌ای در خصوص برخی از مدل‌های مبتنی بر رابطه جونز (و به خصوص مدل‌هایی که متغیر جریان نقدی دوره را به‌عنوان متغیر توضیحی در نظر می‌گیرند) منجر به سویه مدل‌ها به نمایش بهتر از واقع عملکرد از منظر خطا و همبستگی بین متغیرها می‌گردد.

با توجه به نتایج حاصل از این پژوهش، استفاده از الگوریتم مدیریت گروهی داده‌ها به‌عنوان یک روش کارآمد در تحلیل‌های مرتبط با ارزش‌گذاری شرکت‌ها بر مبنای ارقام حسابداری توسط تحلیلگران پیشنهاد می‌گردد. همچنین، این روش می‌تواند توسط حساب‌برسان در برخی از روش‌های تحلیلی به‌عنوان نشانه احتمالی از تحریف ارقام حسابداری مورد توجه قرار گیرد.

با توجه به یافته‌ها و محدودیت‌های تحقیق حاضر، موضوعات زیر برای تحقیقات آتی پیشنهاد می‌گردد:

الف- با توجه به اینکه این پژوهش معطوف به نقش متغیرهای پرکاربرد ادبیات در پیش‌بینی اقلام تعهدی می‌باشد، می‌توان در پژوهش‌های آتی به نقش سایر عوامل در تبیین اقلام تعهدی پرداخت.


ب- در مطالعه نقش سایر عوامل در تغییرات اقلام تعهدی، به مطالعه مجراهای اثرگذاری هر عامل پرداخته و نقش مدیریت سود را از سایر جنبه‌های کیفیت اقلام تعهدی تفکیک نمود.

ج- با توجه به این که در این پژوهش، کارایی مدل در صورت تفکیک تصادفی مشاهدات به داده‌های آموزش و داده‌های ارزیابی بهتر از تفکیک مشاهدات به ترتیب زمان می‌باشد، به نظر می‌رسد نقش برخی عوامل وابسته به زمان از جمله عوامل محیطی در تغییرات اقلام تعهدی محتمل باشد. لذا مطالعه نقش این عوامل در پیش‌بینی اقلام تعهدی پیشنهاد می‌گردد.


از آنجایی که به دلیل عدم وجود تعداد مناسب از شرکت‌ها در صنایع مختلف، امکان انجام تحلیل‌های آماری به تفکیک صنایع مختلف وجود نداشته است، این موضوع از محدودیت‌های این تحقیق به شمار می‌رود. همچنین، تفاوت در ویژگی‌های شرکت‌ها از جمله ساختار مالکیت، میزان سهام شناور و نظایر آن ممکن است بر برآوردهای حاصل از مدل خروجی این تحقیق اثرگذار باشد.

ORCID


Ali Saghafi

 <https://orcid.org/0000-0003-4240-9130>

Ghasem Bulu

 <https://orcid.org/0000-0002-6984-6368>

HosseinAli Sohrabi

 <https://orcid.org/0000-0002-9874-6892>

Varzaneh

منابع

- انواری رستمی، علی‌اصغر؛ آذر، عادل و نوروزی، محمد. (۱۳۹۲). الگوسازی و پیش‌بینی EPS شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران با رویکرد شبکه عصبی GMDH. *بررسی‌های حسابداری و حسابرسی*. ۲۰ (۱). صص ۱۸-۱.
- حجازی، رضوان؛ محمدی، شاپور؛ اصلانی، زهرا و آقاجانی، مجید. (۱۳۹۱). «پیش‌بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم در شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران». *بررسی‌های حسابداری و حسابرسی*. ۱۹ (۲). صص ۴۶-۳۱.
- عباس‌زاده، محمدرضا؛ رجبعلی‌زاده، جواد و قناد، مصطفی. (۱۳۹۸). ارتباطات سیاسی، معاملات با اشخاص وابسته و مدیریت سود در شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران. *فصلنامه مطالعات تجربی حسابداری مالی*. ۱۶ (۶۳). صص ۱۵۵-۱۲۹.
- قادری، اقبال؛ امینی، پیمان؛ نوروش، ایرج و محمدی، عطا. (۱۳۹۷). «تبیین الگوی اندازه‌گیری مدیریت سود با استفاده از روش ترکیبی هوشمند شبکه‌های عصبی و الگوریتم‌های فراابتکاری (ژنتیک و ازدحام ذرات)». *مجله مهندسی مالی و مدیریت اوراق بهادار*. ۳۶. صص ۱۲۷-۹۹.
- کاشانی‌پور، محمد؛ کرمی، غلامرضا و مرادی‌جزء، محسن. (۱۳۹۸). مدلی برای رتبه‌بندی سنج‌های کیفیت سود در ایران. *فصلنامه مطالعات تجربی حسابداری مالی*. ۱۶ (۶۲). صص ۱۵۰-۱۲۷.
- کردستانی، غلامرضا؛ معصومی، جواد و بقایی، وحید. (۱۳۹۲). پیش‌بینی مدیریت سود با استفاده از شبکه عصبی مصنوعی. *مجله پیشرفت‌های حسابداری دانشگاه شیراز*. ۵ (۶۴). صص ۱۹۰-۱۶۹.
- مشایخی، بیتا؛ بیرامی، هانیه؛ بیرامی، هانی و اخلاقی، ساراسادات. (۱۳۹۱). «کشف مدیریت سود با استفاده از شبکه‌های عصبی». *مجله مهندسی مالی و مدیریت اوراق بهادار*. ۱۱. صص ۷۹-۶۳.

References

- Abdou, H. A., Ellelly, N. N., Elamer, A. A., Hussainey, K., & Yazdifar, H. (2020). Corporate governance and earnings management nexus: Evidence from the UK and Egypt using neural networks. *International*

- Journal of Finance & Economics*, ijfe.2120.
<https://doi.org/10.1002/ijfe.2120>
- Al-Dmour, A. H., & Al-Dmour, R. H. (2018). Applying Multiple Linear Regression and Neural Network to Predict Business Performance Using the Reliability of Accounting Information System. *International Journal of Corporate Finance and Accounting*, 5(2), 12–26. <https://doi.org/10.4018/IJCFA.2018070102>
- BALL, R., & SHIVAKUMAR, L. (2006). The Role of Accruals in Asymmetrically Timely Gain and Loss Recognition. *Journal of Accounting Research*, 44(2), 207–242. <https://doi.org/10.1111/j.1475-679X.2006.00198.x>
- Calderon, T. G., & Cheh, J. J. (2002). A roadmap for future neural networks research in auditing and risk assessment. *International Journal of Accounting Information Systems*. [https://doi.org/10.1016/S1467-0895\(02\)00068-4](https://doi.org/10.1016/S1467-0895(02)00068-4)
- Dag, O., Karabulut, E., & Alpar, R. (2019). GMDH2: Binary Classification via GMDH-Type Neural Network Algorithms—R Package and Web-Based Tool. *International Journal of Computational Intelligence Systems*, 12(2), 649. <https://doi.org/10.2991/ijcis.d.190618.001>
- Dameri, R. P., Garelli, R., & Resta, M. (2020). Neural Networks in Accounting: Clustering Firm Performance Using Financial Reporting Data. *Journal of Information Systems*, 34(2), 149–166. <https://doi.org/10.2308/isis-18-002>
- Dechow, P. M., & Dichev, I. D. (2002). Quality Earnings : The The Accruals Accrual Estimation Errors. *The Accounting Review*, 77(2002), 35–59.
- Dechow, P. M., & Skinner, D. J. (2000). Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons*, 14(2), 235-250. <https://doi.org/10.2308/acch.2000.14.2.235>
- Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting Earnings Management Author(s): Detecting Earnings Management. *The Accounting Review*, 70(2), 193-225.
- DeFond, M. L. (2010). Earnings quality research: Advances, challenges and future research. *Journal of Accounting and Economics*, 50(2–3), 402–409. <https://doi.org/10.1016/j.jacceco.2010.10.004>
- DeTienne, K. B., DeTienne, D. H., & Joshi, S. A. (2003). Neural Networks as Statistical Tools for Business Researchers. *Organizational Research Methods*, 6(2), 236-265. <https://doi.org/10.1177/1094428103251907>
- Dong, G. (2019). Exploiting the Power of Group Differences: Using Patterns to Solve Data Analysis Problems. *Synthesis Lectures on Data Mining and Knowledge Discovery*, 11(1), 1–146.

- <https://doi.org/10.2200/S00897ED1V01Y201901DMK016>
- Duan, Y., Yeh, C.-H., & Dowe, D. L. (2018). Accounting Results Modelling with Neural Networks: The Case of an International Oil and Gas Company. *International Conference on Neural Information Processing 2018*, 275–285. https://doi.org/10.1007/978-3-030-04179-3_24
- Farlow, S. J. (1981). The GMDH Algorithm of Ivakhnenko. *The American Statistician*, 35(4), 210–215. <https://doi.org/10.1080/00031305.1981.10479358>
- Francis, J., LaFond, R., Olsson, P., & Schipper, K. (2005). The market pricing of accruals quality. *Journal of Accounting and Economics*, 39(2), 295–327. <https://doi.org/10.1016/j.jacceco.2004.06.003>
- Guo, J., Huang, W., Mao, Q., Wang, X., Wang, X., & Song, T. (2018). Modified GMDH networks for oilfield production prediction. *Geosystem Engineering*, 21(4), 217–225. <https://doi.org/10.1080/12269328.2017.1398110>
- Hassan, M. N., Sulaiman, M. N., Ibrahim, N. A., & Lukman, I. (2010). Group Method of Data Handling (GMDH) for Economic Evaluation of Air Pollution. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1557193>
- Healy, P. M. (1985). The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics*, 7(1–3), 85–107. <https://econpapers.repec.org/RePEc:eee:jaecon:v:7:y:1985:i:1-3:p:85-107>
- Höglund, H. (2010). *Detecting Earnings Management Using Neural Networks*. Hanken School of Economics.
- Höglund, H. (2012). Detecting earnings management with neural networks. *Expert Systems with Applications*, 39(10), 9564–9570. <https://doi.org/10.1016/j.eswa.2012.02.096>
- Jones, J. J. (1991). Earnings Management During Import Relief Investigations. *Journal of Accounting Research*, 29(2), 193. <https://doi.org/10.2307/2491047>
- Kothari, S. P., Leone, A. J., & Wasley, C. E. (2005). Performance matched discretionary accrual measures. *Journal of Accounting and Economics*, 39(1), 163–197. <https://doi.org/10.1016/j.jacceco.2004.11.002>
- Loyola-Gonzalez, O. (2019). Black-Box vs. White-Box: Understanding Their Advantages and Weaknesses From a Practical Point of View. *IEEE Access*, 7, 154096–154113. <https://doi.org/10.1109/ACCESS.2019.2949286>
- McNichols, M. F. (2002). Discussion of the quality of accruals and earnings: The role of accrual estimation errors. *Accounting Review*, 77(1), 61–

69. <https://doi.org/10.2308/accr.2002.77.s-1.61>
- Nazemi, B., & Rafiean, M. (2020). Forecasting house prices in Iran using GMDH. *International Journal of Housing Markets and Analysis*, 14(3), 555-568. <https://doi.org/10.1108/IJHMA-05-2020-0067>
- Parfet, W. U. (2000). Accounting Subjectivity and Earnings Management: A Preparer Perspective. *Accounting Horizons*, 14(4), 481-488. <https://doi.org/10.2308/acch.2000.14.4.481>
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206-215. <https://doi.org/10.1038/s42256-019-0048-x>
- Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review*, 71(3), 289-315.
- Teoh, S. H., Welch, I., & Wong, T. J. (1998). Earnings management and the underperformance of seasoned equity offerings. *Journal of Financial Economics*, 50(1), 63-99. [https://doi.org/10.1016/s0304-405x\(98\)00032-4](https://doi.org/10.1016/s0304-405x(98)00032-4)

References [In Persian]

- Abbaszadeh, Mohammad Reza., Javad Rajabalizadeh., & Mostafa Ghannad. (1398). Political Connections, Related Party Transactions and Earnings Management In Listed Companies in Tehran Stock Exchange. *Empirical Studies in Financial Accounting*, 16(63), 129-155. [In Persian]
- Anvari Rostami, Ali Asghar., Adel Azar., & Mohammad Norozi. (1392). EPS Modeling and Prediction of Listed Companies in Tehran Stock Exchange with GMDH Neural Network Approach. *Accounting and Auditing Reviews*, 20(1), 1-18. [In Persian]
- Ghaderi, Eghbal., Peiman Amini., Iraj Noravesh., & Ata Mohammadi Moqrni. (1397). Explaining the model of earning management measurement using an intelligent hybrid method of neural networks and meta heuristic algorithms (Genetic and particle swarm optimization). *Financial Engineering and Securities Management (Portfolio Management)*, 36. 99-127. [In Persian]
- Hejazi, Rezvan., Shapoor Mohamadi., Zahra Aslani., & Majid Aghajani. (1391). Earnings Management Prediction Using Neural Networks and Decision Tree in TSE. *Accounting and Auditing Reviews*. 19(2). 31-46. [In Persian]
- Kashanipour, Mohammad., Gholamreza Karami., & Mohsen Moradi Joz. (1398). A Model for Ranking of Earnings Quality Measures in Iran. *Empirical Studies of Financial Accounting*. 16(62). 127-150. [In Persian]

- Kordestani, Gholamreza., Javad Masomi., & Vahid Baghaee. (1392). Predicting Earnings Management level By Using Artificial Neural Networks. *Journal of Accounting Advances*, 5(1), 169-190. [In Persian]
- Mashayekhi, Bitaa., Haniyeh Beyrami., Hani Beyrami., & Sara Sadat Akhlaghi. (1391). Detecting Earnings Management Using Neural Networks. *Financial Engineering and Securities Management (Portfolio Management)*, 3(11), 63-79. [In Persian]

استناد به این مقاله: ثقفی، علی، بولو قاسم، سهرابی ورزنده، حسینعلی. (۱۴۰۱). مدل کیفیت اقلام تعهدی با رویکرد الگوریتم مدیریت گروهی داده‌ها، فصلنامه مطالعات تجربی حسابداری مالی، ۱۹(۷۵)، ۱-۴۰.

DOI: 10.22054/qjma.2021.58242.2219



Empirical Studies in Financial Accounting is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.