https://amf.ui.ac.ir
Journal of Asset Management and Financing
E-ISSN: 2383-1189
Vol. 11, Issue 2, No. 41, Summer 2023, p 53-74
Received: 05.02.2023 Accepted: 04.10.2023

Research Paper

Comparing the Efficiency of Statistical Models and Machine-Learning Models and Choosing the Optimal Model for Predicting Net Profit and Operating Cash Flows

Sajjad Mirzaei
MSc. of Financial Management, Faculty of Economics and Management, Urmia University, Urmia, Iran
st_s.mirzaei@urmia.ac.ir
Ali Ashtab *
Assistant Professor, Department of Accounting, Faculty of Economics and Management, Urmia University, Urmia, Iran
a.ashtab@urmia.ac.ir
Akbar Zavari Rezaei
Assistant Professor, Department of Accounting, Faculty of Economics and Management, Urmia University, Urmia, Iran
a.zavarirezaei@urmia.ac.ir

Abstract
The present study compared the predictive performance of machine-learning models and statistical models for forecasting profit and operational cash flow by using a combination of accrual and cash variables. The research method encompassed 3 main stages: data set and variable selection, modeling, and estimation. The study focused on companies listed on the Tehran Stock Exchange (TSE), analyzing data from 184 companies over the period of 2012-2021. The findings indicated that accrual variables exhibited greater explanatory power than cash variables in predicting net profit and future operating cash flow. Furthermore, the comparison of machine-learning and statistical models for forecasting net profit and future operating cash flow revealed that the artificial intelligence approach exhibited superior capability. Specifically, symbolic regression among the machine-learning models and the probit model among the statistical models demonstrated higher performance. Additionally, the results indicated that certain statistical models outperformed some machine-learning models while, on average, machine-learning models outperformed statistical models.
Keywords: Classification, Data Mining, Machine Learning, Net Profit Forecasting, Operating Cash Flow Forecasting.

Introduction

In the current intensely competitive business environment, precise prediction of financial outcomes has emerged as a pivotal element in organizational triumph. Projecting crucial financial indicators, such as net profit and operating cash flows, equips businesses with the insight needed to make well-informed choices regarding investment strategies, resource distribution, and comprehensive financial strategizing. The capacity to anticipate future financial performance enables organizations to streamline operations and mitigate risks. Consequently, there is an escalating need for effective forecasting models.

This study had two primary objectives: firstly, assessing the predictive capability of accrual and cash variables for forecasting profit and future cash flows and secondly, comparing the efficacy of statistical models and machine-learning models in predicting net profit and operating cash flows. Statistical models seek to scrutinize historical data patterns and underlying relationships to anticipate future financial outcomes. Conversely, machine-learning models have emerged as a potent alternative, employing advanced computational techniques to glean insights from data and make predictions without explicit programming. This research was guided by four hypotheses:

First hypothesis: The predictive capability of accrual variables for future net profit significantly exceeds that of cash variables.

Second hypothesis: The predictive capacity of accrual variables for future operational cash flow significantly surpasses that of cash variables.

Third hypothesis: Machine-learning models outperform statistical models significantly in predicting net profit.

*Corresponding author
Mirzaei, S., Ashtab, A., Zavari Rezaei, A. (2023). Comparing the Efficiency of Statistical Models and Machine-Learning Models and Choosing the Optimal Model for Predicting Net Profit and Operating Cash Flows. *Journal of Asset Management and Financing*, 11(2), 53-74.

2383-1189 © University of Isfahan

This is an open access article under the BY-NC-ND/4.0/ License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

http://dx.doi.org/10.22108/AMF.2023.136720.1784

Fourth hypothesis: Machine-learning models outperform statistical models significantly in predicting operational cash flows.

Materials & Methods

This study utilized the Bourseview software database, Rahavard Novin, and the Codal website for analyzing and drawing conclusions regarding the hypotheses. Additionally, data-mining software, such as Weka, SPM, RapidMiner, SPSS Modeler, and Eureqa, were employed for modeling, while Stata econometric and statistical software was used for the Vuong test, EViews for descriptive statistics, SPSS for mean comparison test, and Excel for data sorting and categorization. Following the application of these specified tools, 184 companies listed on the Tehran Stock Exchange (TSE) were examined. Initially, the study investigated the ability to explain each category of cash and accrual variables for net profit and future operating cash flow through special regression estimation of panel data and the Vuong test. Subsequently, the superior model was utilized for modeling and the average performance of the machine-learning models was compared with that of statistical models.

Findings

The significance of Vuong statistic in predicting net profit at a 1% significance level suggested a notable difference in the explanatory power of the two models with the model of accrual variables demonstrating higher explanatory power than that of the cash flow statement variables. Conversely, the non-significance of the Vuong statistic at the 5% significance level for predicting operational cash flow indicated no significant difference in the explanatory power of the two models. The performance results of both statistical and machine-learning models indicated that the symbolic regression classifier, utilizing the genetic algorithm to predict net profit, exhibited the best overall performance and provided valuable results in the longitudinal test sample. Following symbolic regression, the linear support vector machine and MARS ranked second and third, respectively, in overall performance. Similarly, the symbolic regression classifier, employing the genetic algorithm to predict operating cash flow, demonstrated the best overall performance in the longitudinal test samples. After symbolic regression, the deep learning classifier and MARS ranked second and third, respectively, in overall performance.

Discussion & Conclusions

In accordance with testing of the first and second hypotheses of the research, which posited that accrual variables have a greater explanatory capacity for net profit and future operating cash flow compared to cash variables, the coefficients of determination of the models were compared after estimating the appropriate panel data approach. The investigation results indicated that accrual variables indeed possessed greater explanatory power for net profit, thus providing no grounds for rejecting the first hypothesis of the study. However, in the case of operating cash flow, while the explanatory value of accrual variables surpassed that of cash variables, there was no statistically significant difference in the explanation between accrual and cash variables. Consequently, the second hypothesis of the research was rejected. In accordance with testing of the third and fourth hypotheses of the current study, which posited that machine-learning models outperform statistical models in predicting net profit and operating cash flow, the AUC criterion was derived through the implementation of both statistical and machine-learning models. By comparing the success rates of the statistical and machine-learning models, it was observed that the machine-learning models significantly outperformed statistical models in predicting net profit and operational cash flow. Therefore, there was no basis for rejecting the third and fourth hypotheses of the study.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مدیریت دارایی و تأمین مالی
سال یازدهم، شماره دوم، شماره پیاپی (۴۱)، تابستان ۱۴۰۲، ص ۵۳-۷۴
تاریخ دریافت: ۱۴۰۱/۱۱/۱۶ تاریخ پذیرش: ۱۴۰۲/۰۷/۱۲

مقاله پژوهشی

مقایسهٔ کارایی مدل‌های آماری و یادگیری ماشین و انتخاب مدل بهینه در پیش‌بینی سود خالص و جریان‌های نقدی عملیاتی

سجاد میرزایی
کارشناس ارشد مدیریت مالی، دانشکدهٔ اقتصاد و مدیریت، دانشگاه ارومیه، ارومیه، ایران
st_s.mirzaei@urmia.ac.ir
علی آشتاب *
استادیار، گروه حسابداری، دانشکدهٔ اقتصاد و مدیریت، دانشگاه ارومیه، ارومیه، ایران
a.ashtab@urmia.ac.ir
اکبر زواری رضائی
استادیار، گروه حسابداری، دانشکدهٔ اقتصاد و مدیریت، دانشگاه ارومیه، ارومیه، ایران
a.zavarirezaei@urmia.ac.ir

چکیده

هدف: در پژوهش حا ضر، مقایسهٔ عملکرد مدل‌های یادگیری ما شین و مدل‌های آماری در پیش‌بینی سود و جریان نقد عملیاتی با ا ستفاده از مجموعه متغیرهای تعهدی و نقدی بررسی شده است.

روش: روش شنا سی پژوهش به سه مرحلهٔ گزینش مجموعه داده و متغیرها، مدل سازی و قیاس تق سیم‌بندی می شود. جامعهٔ آماری پژوهش حاضر، شرکت‌های بورس اوراق بهادار تهران و داده‌های ۱۸۴ شرکت طی بازهٔ زمانی ۱۳۹۱ تا ۱۴۰۰ بررسی شده است.

یافته‌ها: نتایج این پژوهش نشان‌دهندهٔ آن بود که متغیرهای تعهدی توان تبیین بیشتری نسبت به متغیرهای نقدی برای پیش‌بینی سود خالص و جریان نقد عملیاتی آتی دارد. علاوه بر این، مقایسهٔ عملکرد مدل‌های یادگیری ماشین و آماری در پیش‌بینی سود خالص و جریان نقد عملیاتی آتی نشان‌دهندهٔ آن بود که رویکرد هوش مصنوعی توانایی بیشتری دارد و بین مدل‌های یادگیری ماشین، رگرسیون نمادین و مدل‌های آماری، مدل پروبیت از عملکرد بیشتری برخوردار است؛ همچنین نتایج نشان‌دهندهٔ آن بود که اگرچه به‌طور میانگین مدل‌های یادگیری ماشین عملکرد بیشتری نسبت به مدل‌های آماری دارد، مدل‌های آماری نیز عملکرد بیشتری از برخی مدل‌های یادگیری ماشین ارائه می‌دهد.

کلیدواژه‌ها: پیش‌بینی جریان نقد عملیاتی، پیش‌بینی سود خالص، داده‌کاوی، طبقه‌بندی، یادگیری ماشین

مقدمه

در چشم‌انداز تجاری بسیار رقابتی امروزی، پیش‌بینی دقیق نتایج مالی به عاملی حیاتی در موفقیت سازمان تبدیل شده است. پیش‌بینی معیارهای مالی کلیدی مانند سود خالص و جریان‌های نقدی عملیاتی به کسب‌وکارها قدرت می‌دهد تا تصمیمات آگاهانه دربارهٔ استراتژی‌های سرمایه‌گذاری، تخصیص منابع و برنامه‌ریزی مالی کلی اتخاذ کنند. توانایی پیش‌بینی عملکرد مالی آتی به سازمان‌ها اجازه

* نویسندهٔ مسئول

میرزایی، سجاد.، آشتاب، علی.، زواری رضائی، اکبر. (۱۴۰۲). مقایسهٔ کارایی مدل‌های آماری و یادگیری ماشین و انتخاب مدل بهینه در پیش‌بینی سود خالص و جریان‌های نقدی عملیاتی. *مدیریت دارایی و تأمین مالی*، ۱۱(۲)، ۵۳-۷۴.

2383-1189 © University of Isfahan
This is an open access article under the BY-NC-ND/4.0/ License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

http://dx.doi.org/10.22108/AMF.2023.136720.1784

می‌دهد تا عملیات را بهینه کنند، ریسک‌ها را کاهش دهند و فرصت‌ها را به‌موقع استفاده کنند؛ درنتیجه تقاضای فزاینده‌ای برای مدل‌های پیش‌بینی کارا وجود دارد که پیش‌بینی‌های دقیق و قابل‌اعتمادی را برای این شاخص‌های مالی مهم ارائه می‌دهد (Parlina & Budianto, 2021).

اجزای مختلف سود عامل توضیحی مهمی برای پیش‌بینی جریان نقد و سود آتی است. منابع موجود در تجزیه‌وتحلیل صورت‌های مالی اغلب از بر رسی اجزای تعهدی و جریان نقدی سود جاری به‌منظور پیش‌بینی سودهای آتی حمایت می‌کند (Oz et al., 2021). اطلاعات سود شامل اقلام تعهدی و جریان‌های نقدی ا ست که هر دو بر عملکرد شرکت در آینده تأثیر می‌گذارد (Sharawi, 2021). به‌طور سنتی، تجزیه‌وتحلیل مالی برای مدت طولانی به عملکرد حسابداری از طریق متغیرهای تعهدی بستگی دارد؛ با این حال، این اشکال متغیرها متأثر از اشکالات اساسی هستند که ویژگی‌های حسابداری تعهدی است. بیشتر نسبت‌های محاسبه‌شده به‌طور معمول فقط بر صورت وضعیت مالی و صورت سود و زیان تمرکز دارد و این در حالی است که صورت جریان‌های نقدی نیز بینش مفیدی از تجزیه‌وتحلیل متغیرها ارائه می‌دهد. متغیرهای صورت وضعیت مالی فقط چشم‌اندازی از تاریخ در زمان ارائه می‌دهد؛ در حالی که صورت جریان وجوه نقد نشان‌دهندهٔ فعالیت برای دوره‌ای مداوم ا ست. صورت سود و زیان نتایج عملیات را برای یک دورهٔ زمانی گزارش می‌کند؛ اما سایر تغییرات مهم در منابع را که نا شی از فعالیت‌های تأمین مالی و سرمایه‌گذاری ا ست، اف شا نمی‌کند. صورت جریان وجوه نقد با ارائهٔ اطلاعات اضافی نسبت به صورت وضعیت مالی و سود و زیان توانایی سازمان را برای عملکرد کارآمد، تأمین مالی و پرداخت تعهدات تکمیل می‌کند (Barua & Saha, 2015).

هدف اول این پژوهش، ارائهٔ نمایی کلی از متغیرهای جریان نقدی و تعهدی به‌عنوان یک ابزار تحلیلی قدرتمند و مؤثر است. اگرچه محتوای اطلاعاتی متغیرهای تعهدی و صورت جریان‌های نقدی ممکن است عاملی توضیحی برای سودها و جریان‌های نقدی آینده باشد، توانایی آنها در شناسایی سود و جریان نقد آتی بررسی نشده است؛ از این رو، این پژوهش در مرحلهٔ اول، به‌دنبال بررسی توانایی توضیح‌دهندگی متغیرهای تعهدی و نقدی برای سود و جریان نقد آتی است. به‌طور خاص، این پژوهش پیش‌بینی سود خاص و جریان نقد عملیاتی آتی را با مجموعه‌ای از متغیرهای نقدی و تعهدی بررسی می‌کند که توانایی توضیح‌دهندگی بیشتری دارد.

در قلمرو مدل سازی پیش‌بینی، دو رویکرد برجسته جذابیت جالب‌توجهی به د ست آورده‌اند: مدل‌های آماری و یادگیری ما شین. مدل‌های آماری مدت‌هاست که به‌عنوان ابزاری اساسی در تحلیل و پیش‌بینی داده‌ها استفاده می‌شود. این مدل‌ها براساس مفروضات از پیش تعریف‌شده و الگوریتم‌های ریاضی، با استفاده از فن‌هایی مانند تحلیل رگرسیون، تحلیل سری‌های زمانی و مدل‌سازی اقتصادسنجی ساخته شده است (Monahan, 2018). هدف مدل‌های آماری با بررسی الگوهای داده‌های تاریخی، گرفتن روابط و پویایی‌های اساسی به‌منظور ایجاد پیش‌بینی برای نتایج مالی آینده است. (Martins, 2022).

از سوی دیگر، مدل‌های یادگیری ما شین به‌عنوان جایگزینی قدرتمند ظاهر می شود و از فن‌های محا سباتی پیشرفته برای یادگیری از داده‌ها و پیش‌بینی بدون برنامه‌نویسی صریح استفاده می‌کند (Petropoulos et al., 2020). الگوریتم‌های یادگیری ماشین مانند شبکه‌های عصبی، درخت‌های تصمیم‌گیری و ماشین‌های بردار پشتیبان، قابلیت‌های چشمگیری را در مدیریت مجموعهٔ داده‌های پیچیده و با ابعاد زیاد نشان داده است (Ashtab et al., 2017). این مدل‌ها در گرفتن روابط غیرخطی، شناسایی الگوهای پیچیده و تطبیق با دینامیک داده‌های در حال تغییر برتری دارند. حوزهٔ یادگیری ما شینی در سال‌های اخیر شاهد پیشرفت‌های جالب‌توجهی بوده که به افزایش کاربرد آن در حوزه‌های مختلف ازجمله امور مالی منجر شده است (Dastile et al., 2020).

سؤال کلیدی که مطرح می شود، این ا ست: کدام رویکرد، مدل‌های آماری یا یادگیری ما شینی راه‌حل بهینه را برای پیش‌بینی سود خالص و جریان‌های نقد عملیاتی ارائه می‌دهد؟ در حالی که هر دو روش مزایای خود را دارد و مزایا و محدودیت‌های متمایزی را نیز ارائه می‌کند که باید به‌دقت بررسی شود. مدل‌های آماری پایه‌ای محکم را مبتنی بر نظریهٔ آماری و امکان تفسیر و استنتاج فراهم می‌کند. آنها اغلب به فرضیات صریح دربارهٔ داده‌ها نیاز دارند و برای تخمین پارامتر به فن‌های آماری تثبیت‌شده تکیه می‌کنند

(Mullainathan & Spiess, 2017). در مقابل، مدل‌های یادگیری ما شین با اجازه‌دادن به داده‌ها برای هدایت فرآیند تصمیم‌گیری مدل، پتانسیل افزایش دقت و انعطاف‌پذیری را ارائه می‌کند؛ با این حال، این انعطاف‌پذیری گاهی اوقات به قیمت تفسیرپذیری تمام می‌شود؛ زیرا مدل‌های یادگیری ماشـــینی به دلیل معماری پیچیده و عملکرد داخلی پیچیده‌ای که دارند، اغلب به‌عنوان مدل‌های جعبه ســـیاه در نظر گرفته می‌شوند (Anand et al., 2019).

هدف این پژوهش، برر سی چالش مهم مقایسهٔ کارایی مدل‌های آماری و یادگیری ما شین در پیش‌بینی سود خالص و جریان‌های نقدی عملیاتی اســت. با انجام تحلیلی جامع از این دو رویکرد، هدف، ارائهٔ بینشـــی دربارهٔ نقاط قوت و ضـــعف مربوط به آنهاســـت و تصمیم‌گیرندگان را قادر می‌کند تا هنگام انتخاب مدل بهینه برای پیش‌بینی مالی، انتخاب‌های آگاهانه‌ای دا شته با شند. برای ر سیدن به این هدف، معیار ارزیابی AUC در نظر گرفته شده است. به‌طور خلاصه، این مقاله به گفتمان جاری پیرامون انتخاب مدل‌های پیش‌بینی منا سب برای پیش‌بینی مالی کمک می‌کند. با برر سی کارایی مدل‌های آماری و یادگیری ما شین در پیش‌بینی سود خالص و جریان‌های نقدی عملیاتی، هدف این ا ست که شکاف بین نظریه و عمل پر شود و متخ صـ صان و پژوهـ شگران با دانش لازم برای تصمیم‌گیری آگاهانه دربارهٔ انتخاب مدل تجهیز شوند.

در پژوهش‌های گذشته اگرچه پیش‌بینی از طریق مدل‌های یادگیری ماشین و آماری انجام شده است، تاکنون این مدل‌ها در رابطه با پیش‌بینی سود خالص و جریان نقد عملیاتی با یکدیگر مقایسه نشده است؛ همچنین برر سی نتایج پژوهش‌های پیشین نشان‌دهندهٔ آن اســت که تعداد معدودی از مدل‌ها در هر پژوهش بررســی شــده اســت که در این پژوهش برای جامعیت بیشــتر موضــوع از لیســت جامع‌تری از مدل‌های آماری و یادگیری ماشین استفاده شده است. علاوه بر این، در پژوهش حاضر توان متغیرهای تعهدی و نقدی در پیش‌بینی سود خالص و جریان نقد عملیاتی آتی برر سی شده و برای دقیق‌بودن نتایج برخلاف پژوهش‌های پی شین که فقط از نمونهٔ آزمایشی مقطعی استفاده شده است، در این پژوهش از نمونهٔ آزمایشی طولی نیز در بررسی مدل‌ها استفاده شده است؛ بنابراین پژوهش حاضــر با اســتفاده از ۳۵ طبقه‌بندی‌کنندهٔ غیر پارامتری و یادگیری ماشــین، از طریق شــش طبقه‌بندی‌کنندهٔ پارامتری و مدل‌های آماری، پیش‌بینی سود و جریان نقدی عملیاتی آتی را بررسی کرده است.

مبانی نظری

واژهٔ سود یکی از مفاهیم پیچیدهٔ جهان تجارت است که شاید نتوان تعریف دقیقی از سود ارائه کرد که موردقبول همگان قرار گیرد. در میان تعاریف مختلفی که تاکنون از سود ارائه شده ا ست، تعریف ذیل یکی از جامع‌ترین آنها ست: « سود از تغییر در حقوق صاحبان سهام یا تغییر در خالص دارایی‌های یک واحد اقتصادی طی دوره‌ای مالی نا شی می شود. به‌عبارت دیگر، سود برآیند کلیهٔ تغییرات در حقوق صاحبان سهام طی یک دورهٔ مالی، به‌جز تغییرات نا شی از سرمایه‌گذاری تو سط صاحبان سهام و توزیع منابع بین آنان ا ست» (Curtis et al., 2021).

نالاردی[1] (2020) بر این باور ا ست که واژهٔ سود مفهومی انتزاعی ا ست. در صورتی که وجه نقد مفهومی عینی ا ست. وی معتقد است که تداوم فعالیت بنگاه اقتصادی از اولین موارد موردتوجه بوده است و وجه نقد از مهم‌ترین عوامل ادامهٔ حیات هر بنگاه اقتصادی به‌حســـاب می‌آید. تنها شـــرکت‌هایی به فعالیت خود ادامه می‌دهند که علاوه بر ســـودآور بودن، امکان تأمین نیازهای نقدی خود را نیز داشته باشند. توجه به وضعیت نقدینگی بنگاه برای پرداخت‌های روزمره از ضروریات است. این در حالی است که سود به‌خوبی این اطلاعات را بیان نمی‌کند؛ اما صـــورت جریان‌های نقدی امکان فراهم‌کردن چنین اطلاعاتی را دارد. یک عقیده در تصـــدیق ســـود حسابداری این است که سود در طول زمان به‌طور پیوسته توانایی این را داشته است که سرافراز بماند.

[1]. Nallareddy

یک اصل اساسی در حسابداری این است که صورت‌های مالی تهیه‌شده براساس روش‌های اقلام تعهدی، اطلاعات بیشتری نسبت به حسابداری نقدی دارد که نتیجهٔ آن این است که اقلام تعهدی سود را بیشتر از جریان‌های نقدی تبیین می‌کند؛ همان‌طور که در ادبیات دانشگاهی بیان شده است، اصل این است که اقلام تعهدی این مشکل را بهبود می‌بخشد که جریان‌های نقدی تحقق‌یافته دارای مشکلات زمان‌بندی و تطبیق هستند که باعث می شود، آنها معیاری پیچیده و پرحرف و حدیث برای عملکرد شرکت با شند. درواقع، مدل‌های اقلام تعهدی تحلیلی به این نتیجه می‌رسند که سود، پیش‌بینی‌کنندهٔ بهتر برای جریان‌های نقدی عملیاتی آتی نسبت به جریان‌های نقدی عملیاتی فعلی ا ست (Ball & Nikolaev, 2020). ادبیات گزارش هیئت ا ستانداردهای حسابداری مالی (1978) بیان می‌کند که: علاقهٔ کاربران به جریان‌های نقدی آتی شرکت و توانایی آن برای ایجاد جریان‌های نقدی مطلوب اغلب به علاقه‌مندی به اطلاعات مربوط به سود شرکت منجر می‌شود تا اطلاعات مستقیم دربارهٔ جریان‌های نقدی شرکت. بخش عمده‌ای از اطلاعات مربوط به جریان‌های نقدی با سود تأمین می شود؛ به‌عنوان‌مثال، با کسرکردن اقلام تعهدی از سود مقدار جریان‌های نقدی به د ست می‌آید. به‌طور کلی جزء تعهدی سود به نسبت جزء نقدی آن عدم قطعیت بیشتری دارد؛ به این دلیل که اقلام تعهدی با قضاوت‌ها، برآوردها و تخصیص‌ها (از جریان‌های نقد تولیدشده در دیگر دوره‌ها) ایجاد می‌شوند، در صورتی که جزء نقدی سود عینی‌تر است (Nguyen & Nguyen, 2020). با توجه به پژوهش‌های قبلی در رابطه با ویژگی‌های کیفی سود، کیفیت اقلام تعهدی نسبت به سایر ویژگی‌های کیفی سود برای تعیین ریسک اطلاعاتی مربوط به جریان‌های نقدی مناسب‌تر است (Khodamipour et al., 2013). اقلام تعهدی، بخش مهم سیستم گزارشگری مالی است که شامل هر چیزی است که بین سود و جریان نقدی فاصله ایجاد می‌کند. آنها منعکس‌کنندهٔ طیف و سیعی از تصمیمات شرکتی، ازجمله فروش، تولید، سرمایه‌گذاری، حسابداری و انتخاب‌های مدیریت نقدی شرکت هستند. به‌خوبی ثابت شده است که از دو شرکت با سود یکسان جاری، شرکتی که اقلام تعهدی بیشتری گزارش می‌کند، سود کمتری در آینده خواهد داشت. این پیوند بین اقلام تعهدی و سودآوری آتی که اغلب با این جمله خلاصه می‌شود که «اقلام تعهدی نسبت به جریان‌های نقدی دوام کمتری دارد»، برای ارزیابی شرکت، تجزیه‌وتحلیل صورت‌های مالی و طیف و سیعی از مسائل در حسابداری مهم است. آیا شرکت‌ها از اقلام تعهدی برای مدیریت سود استفاده می‌کنند؟ آیا اقلام تعهدی مثبت یا منفی بزرگ منعکس‌کنندهٔ شرایط اقتصادی شرکت است یا اطلاعاتی را دربارهٔ کیفیت سود نشان می‌دهد؟ دو توضیح اصلی برای تداوم پایین اقلام تعهدی، اولین بار از سوی اسلون بیان شد (Sloan, 1996). اول، این‌که ذهنیت و تحریف در گزارشگری مالی به خطای اندازه‌گیری گذرا در اقلام تعهدی و سود منجر می‌شود؛ دوم، اقلام تعهدی ارتباط نزدیکی با سرمایه‌گذاری دارند و به دلیل کاهش بازده به مقیاس، هزینه‌های تعدیل مرتبط با سرمایه‌گذاری یا محافظه‌کاری در حسابداری، سودآوری آتی کمتر را پیش‌بینی می‌کند (Christensen et al., 2022). اقلام تعهدی بیشتر با تغییرات گذرا در حاشیهٔ سود مرتبط است و سودهای بعدی کمتر را پیش‌بینی می‌کند؛ زیرا امروزه افزایش قیمت نهاده‌ها هزینه‌های تولید و موجودی شرکت را افزایش می‌دهد؛ اما زمانی که موجودی به فروش می‌رسد، سودهای آتی را کاهش می‌دهد. علاوه بر این، افزایش تقاضا به افزایش موقت سود و سرمایه در گردش و به‌دنبال آن بازگشت میانگین در متغیرها منجر می شود؛ زیرا رقابت قیمت‌ها و سودآوری را به سطوح تعادلی بلندمدت خود بازمی‌گرداند؛ درنتیجه اقلام تعهدی به‌طور مثبت با سود جاری مرتبط است؛ اما با کنترل این رابطه، با سودهای بعدی ارتباط منفی دارد. به‌طور خلاصه، تداوم پایین اقلام تعهدی به دلیل واکنش تولید، فروش و سود به تغییرات در بازارهای محصول شرکت است (Lewellen & Resutek, 2019).

بال و نیکولایف[1](2020) این سؤال را که آیا طبق چارچوب مفهومی هیئت استانداردهای حسابداری مالی (FASB, 1978) سودهای مبتنی بر اقلام تعهدی اطلاعات بهتری دربارهٔ جریان‌های نقدی عملیاتی آتی نسبت به خود جریان‌های نقدی عملیاتی ارائه می‌دهد، بین شرکت‌های پذیرفته‌شده در بورس اوراق بهادار هلند از سال 1999 تا 2019 بررسی کردند و به این نتیجه رسیدند که سود بهتر از جریان‌های نقدی عملیاتی عمل می‌کند .بر این اساس فرضیه‌های اول و دوم پژوهش به‌صورت زیر تدوین می‌شود:

[1]. Ball & Nikolaev

فرضیهٔ اول: توانایی توضیح‌دهندگی متغیرهای تعهدی برای سود خالص آتی به‌طور معناداری بیشتر از متغیرهای نقدی است.

فرضهٔ دوم: توانایی توضیح‌دهندگی متغیرهای تعهدی برای جریان نقد عملیاتی آتی به‌طور معناداری بیشتر از متغیرهای نقدی است.

مدل‌های آماری از دیرباز به‌عنوان سنگ‌بنای تحلیل و پیش‌بینی داده‌ها عمل کرده است. این مدل‌ها مبتنی بر نظریهٔ آماری هستند و از فن‌های تثبیت‌شده مانند تحلیل رگرسیون و سری‌های زمانی برای ثبت روابط بین متغیرها و ایجاد پیش‌بینی استفاده می‌کنند (Chatfield, 2020). مدل‌های آماری چندین مزیت ازجمله تفسیرپذیری، استنتاج، توانایی ترکیب دانش و فرضیات قبلی دربارهٔ داده‌ها ارائه می‌دهد (Barboza et al., 2017). تفسیرپذیری مدل‌های آماری به متخصصان اجازه می‌دهد تا تأثیر متغیرها را درک کرده و اهمیت ضرایب آنها را ارزیابی کنند، تصمیم‌گیری را تسهیل می‌کند و بینش‌های ارزشمندی دربارهٔ پویایی‌های اساسی داده‌های مالی ارائه می‌دهد (Montgomery et al., 2019). علاوه بر این، مدل‌های آماری اغلب به نقاط دادهٔ کمتری برای تولید پیش‌بینی‌های قابل‌اعتماد نیاز دارد و آنها را برای سناریوهایی با در دسترس بودن داده‌های محدود مناسب می‌کند (Ahn et al., 2020).

با این حال، مدل‌های آماری محدودیت‌های خاصی دارند که باید در نظر گرفته شوند. این مدل‌ها اغلب خطی‌بودن بین متغیرها را فرض می‌کند و ممکن است برای گرفتن روابط غیرخطی پیچیده تلاش کند (Bishop & Nasrabadi, 2006). آنها بر مفروضاتی مانند استقلال، نرمال‌بودن و همسانی واریانس تکیه می‌کنند که ممکن است در مجموعهٔ داده‌های مالی دنیای واقعی وجود نداشته باشد (Mills, 2019)؛ درنتیجه مدل‌های آماری ممکن است در صورت مواجهه با توزیع داده‌های غیرخطی یا غیر نرمال، پیش‌بینی‌های کمتر از حد بهینه را ایجاد کند. به‌علاوه، مدل‌های آماری ممکن است با مجموعه داده‌های با ابعاد بالا که حاوی متغیرهای متعددی است، مشکل داشته باشد؛ زیرا اغلب به انتخاب دقیق متغیر و ساده‌سازی مدل برای جلوگیری از برازش بیش‌ازحد نیاز دارد (Gelman et al., 2020).

از سوی دیگر، مدل‌های یادگیری ماشینی، رویکردی انعطاف‌پذیر و مبتنی بر داده‌ها را برای پیش‌بینی مالی ارائه می‌دهد. این مدل‌ها ازجمله شبکه‌های عصبی، درخت‌های تصمیم‌گیری و ماشین‌های بردار پشتیبان از فن‌های محاسباتی پیشرفته برای یادگیری خودکار الگوها و روابط از داده‌ها بدون تکیه بر مفروضات صریح استفاده می‌کند (Raschka & Mirjalili, 2020). مدل‌های یادگیری ماشین قابلیت‌های پیش‌بینی استثنایی را با گرفتن روابط غیرخطی پیچیده و انطباق با دینامیک داده‌ها در حال تغییر نشان داده است (Bengio et al., 2021) آنها در سناریوهایی که روابط بین متغیرها پیچیده یا به‌خوبی درک نشده است، برتری دارند.

یکی از مزیت‌های اولیهٔ مدل‌های یادگیری ماشین، توانایی آنها در مدیریت مجموعه داده‌های با ابعاد بالا با تعداد زیادی متغیر است (Hastie et al., 2009). این مدل‌ها ویژگی‌های مرتبط را به‌طور خودکار شناسایی می‌کنند و بازنمایی‌های داخلی خود را برای ثبت الگوهای زیربنایی در داده‌ها تطبیق می‌دهند (Goodfellow et al., 2016). مدل‌های یادگیری ماشینی انواع مختلفی از داده‌ها ازجمله داده‌های عددی، مقوله‌ای و متنی را مدیریت می‌کند و آنها را برای طیف گسترده‌ای از وظایف پیش‌بینی مالی قابل‌اجرا می‌سازد (Chen & Guestrin, 2016). علاوه بر این، مدل‌های یادگیری ماشینی از فن‌های مجموعه مانند بسته‌بندی و تقویت برای بهبود دقت پیش‌بینی با ترکیب خروجی‌های مدل‌های متعدد استفاده می‌کنند (Breiman, 1996).

با این حال، مدل‌های یادگیری ماشینی بدون چالش نیستند. یکی از نگرانی‌های کلیدی، ماهیت جعبه سیاه آنهاست که مانع از تفسیرپذیری و درک مکانیسم‌های اساسی مدل می‌شود (Murdoch et al., 2019). برخلاف مدل‌های آماری، مدل‌های یادگیری ماشین دقت پیش‌بینی را بر تفسیرپذیری اولویت می‌دهند و توضیح چگونگی و چرایی پیش‌بینی‌های خاص را دشوار می‌کنند (Caruana et al., 2015). این عدم تفسیرپذیری نگرانی‌های قانونی و اخلاقی ایجاد می‌کند، به‌ویژه در حوزه‌هایی مانند امور مالی که شفافیت و پاسخگویی بسیار مهم است (Rudin, 2019). به‌علاوه، مدل‌های یادگیری ماشینی ازنظر محاسباتی فشرده است، به منابع محاسباتی قابل‌توجهی نیاز دارد و آنها را برای برنامه‌های پیش‌بینی بلادرنگ یا محیط‌هایی باقابلیت‌های محاسباتی محدود کمتر مناسب می سازد (Huang et al., 2017).

رجب‌زاده و همکاران (2022)، طی پژوهشی با عنوان پیش‌بینی جریان وجه نقد عملیاتی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از روش یادگیری ماشین جریان نقد عملیاتی را با رویکرد هوش مصنوعی PLSVM و CART در بازهٔ زمانی ۱۳۹۰ تا ۱۳۹۷ بررسی کردند و نشان دادند که رویکرد هوش مصنوعی قانون‌گرا و غیرخطی پارامتریک توانایی زیادی در پیش‌بینی نقدینگی شرکت‌های بورس اوراق بهادار تهران دارد. حبیب‌زاده و ایزدپور (2020)، تأثیر ۳۴ متغیر را بر صحت پیش‌بینی سودآوری شرکت‌های پذیرفته‌شده توسط بورس اوراق بهادار تهران بررسی کردند. نتایج مقایسه نشان‌دهندهٔ آن بود که استفاده از درخت تصمیم C5 با دقت ۹۳/۵۴ درصد بهترین پیش‌بینی را دارد و سپس مدل شبکهٔ عصبی با دقت ۸۱/۴۵ درصد دقیق‌تر از ماشین بردار پشتیبان با دقت ۶۹/۳۵ در صد شده است و خطای کمتر دارد. مهربان‌پور و حبیب‌زاده (2017)، سودآوری شرکت‌ها را خوشه‌بندی و پیش‌بینی کردند. شرکت‌ها به سه خوشه تبدیل شدند و با استفاده از درخت تصمیم C5، تجزیه‌وتحلیل خوشه‌ای و متغیرهای مؤثر بر سودآوری شناسایی و نشان داده شد که ۸ متغیر شامل: سود ناخالص به کل دارایی‌ها، فروش به‌کل دارایی‌ها، سود به حقوق صاحبان سهام، سود عملیاتی به فروش خالص، سود و زیان انباشته به حقوق صاحبان سهام، سود خالص به فروش خالص، کل بدهی‌ها به کل دارایی‌ها و دارایی‌های جاری به کل دارایی‌ها بر سودآوری شرکت‌ها تأثیر می‌گذارد؛ درنتیجه با استفاده از این متغیرها، پیش‌بینی سه خوشه انجام شد و دقت آنها به‌ترتیب برابر با ۸۶/۳۴، ۸۸/۱۵ و ۸۱/۶۸ در صد است. آناند و همکاران[1](2019) به‌دقت طبقه‌بندی ۵۷ تا ۶۴ در صد برای معیارهای سودآوری در مقایسه با ۵۰ درصد گام تصادفی دست یافتند. آنها دریافتند که دقت طبقه‌بندی در افق‌های یک تا پنج‌ساله یکسان است و نشان دادند که روش‌های یادگیری ماشین عملکرد پیش‌بینی بهتری نسبت به روش‌های سنتی مبتنی بر رگرسیون ارائه می‌دهد. شینیو و همکاران[2](2020) به‌طور جامع امکان‌پذیری و مناسب‌بودن اتخاذ مدل‌های یادگیری ماشین را بر پیش‌بینی مبانی شرکت (سود) ارزیابی کردند و نشان دادند که در مقایسه با مدل‌های آماری سنتی قبلی مانند رگرسیون لجستیک روش آنها در دقت و سرعت پیش‌بینی به پیشرفت رضایت‌بخشی دست یافته است. با توجه به مبانی نظری و پیشینهٔ پژوهش، فرضیهٔ سوم و چهارم پژوهش به شرح زیر بیان می‌شود:

فرضیهٔ سوم: عملکرد مدل‌های یادگیری ماشین نسبت به مدل‌های آماری در پیش‌بینی سود خالص به‌طور معناداری بیشتر است.

فرضیهٔ چهارم: عملکرد مدل‌های یادگیری ماشین نسبت به مدل‌های آماری در پیش‌بینی جریان نقد عملیاتی به‌طور معناداری بیشتر است.

پژوهشگاه علوم انسانی و مطالعات فرهنگی

روش پژوهش

در این پژوهش برای تجزیه‌وتحلیل و نتیجه‌گیری درخصوص فرضیه‌ها از بانک اطلاعاتی نرم‌افزار بورس‌ویو برای داده‌های مربوط به متغیرهای صورت جریان وجوه نقد، ره‌آورد نوین برای داده‌های صورت وضعیت مالی و صورت سود و زیان و برای متغیر استهلاک و نواقص موجود از سایت کدال (سایت اطلاع‌رسانی ناشران اوراق بهادار) استفاده شده است. از نرم‌افزارهای داده‌کاوی Weka، SPM، Rapidminer، SPSS Modeler، Eureqa برای مدل‌سازی و نرم‌افزارهای اقتصادسنجی و آماری Stata برای آزمون ونگ، EViwes برای آمار توصیفی، SPSS برای آزمون مقایسهٔ میانگین‌ها و برای مرتب‌سازی و دسته‌بندی داده‌ها از Excel استفاده شده است.

جامعهٔ آماری، کلیهٔ شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران در بازهٔ زمانی ۱۳۹۱ تا ۱۴۰۰ را شامل می‌شود که شرایط زیر را دارند: شرکت در گروه شرکت‌های سرمایه‌گذاری یا واسطه‌گری مالی نباشد؛ علت این امر، ماهیت متفاوت عملیات شرکت‌های مذکور است؛ اطلاعات مالی آن برای دورهٔ مدنظر در اختیار باشد؛ زیرا شرکت‌هایی که داده‌های مالی گمشده یا ناقص دارد، باعث سوگیری یا نادرستی در یافته‌ها می شود؛ در طول دورهٔ مدنظر پژوهش تغییر سال مالی نداشته باشد؛ زیرا به یکپارچگی و سازگاری پژوهش کمک می‌کند؛ در طول دورهٔ مدنظر پژوهش در بورس حضور داشته باشد؛ زیرا شرکت‌هایی که در بورس نباشد، ممکن است

[1]. Anand et al
[2]. Xinyue et al

شفافیت و الزامات گزارشگری مالی مانند شرکت‌های سهامی عام نداشته و دورهٔ مالی شرکت‌ها منتهی به ۱۲/۲۹ هر سال باشد تا اثر چرخه‌های تجاری خنثی شود. پس از اعمال محدودیت‌های مذکور ۱۸۴ شرکت بورس اوراق بهادار تهران بررسی شد.

از داده‌های ۵۹ متغیر تعهدی و نقدی استفاده شده است که تعریف عملیاتی متغیرها در جدول (۱) بیان شده است. همهٔ متغیرها به‌جز نسبت‌ها بر دارایی ابتدای سال تقسیم شده است. در ابتدا، توانایی تبیین هر دسته از متغیرهای نقدی و تعهدی برای سود خالص و جریان نقد عملیاتی آتی از طریق تخمین رگرسیون خاص داده‌های تابلویی و آزمون ونگ بررسی شده و سپس مدل برتر برای مدل‌سازی مورداستفاده قرار گرفته است. مدل‌های پیش‌بینی مورداستفاده در این پژوهش اگرچه فهرست جامعی نیست، ۴۱ مدل انتخاب شده برای پژوهش حاضر نشان‌دهندهٔ پرکاربردترین و مورد استنادترین مدل‌های طبقه‌بندی در ادبیات است. در یک سمت طیف، مدل‌های خطی به نسبت ساده‌ای وجود دارد که در ادبیات شایع است؛ اما توانایی محدودی برای مدل‌سازی روابط غیرخطی و ناهمسانی در مجموعه داده‌ها دارد. در وسط طیف مدل‌هایی قرار گرفته که برای کنترل غیرخطی بودن و ناهمسانی مجهزتر است. انعطاف‌پذیری بیشتر این مدل‌ها اغلب به قیمت تفسیرپذیری کمتر ختم می‌شود. در سمت دیگر طیف، مدل‌های غیرخطی و کلی وجود دارد که به شیوه‌ای طراحی شده است که همهٔ روابط و تعاملات غیرخطی را در مجموعهٔ داده نمایش دهند (برای مروری بر کاربردهای مدرن و مطالعهٔ بیشتر به هستی و همکاران (2009) مراجعه کنید). برای مدل سازی و اعتبار سنجی ۷۰ در صد از داده‌ها به‌عنوان دادهٔ آموزش و ۳۰ درصد از داده‌ها به‌عنوان داده آزمایش در نظر گرفته شده است.

جدول (۱) تعاریف عملیاتی متغیرهای پژوهش

Table (1) Operational definitions of research variables

متغیر	تعریف عملیاتی	منبع
نسبت جاری x1	دارایی جاری / بدهی جاری	باروا و سها (2015)
نسبت سریع x2	(موجودی کالا – دارایی جاری) / بدهی جاری	باروا و سها (2015)
گردش دارایی‌ها x3	فروش / کل دارایی	باروا و سها (2015)
نسبت بدهی به کل دارایی x4	کل بدهی‌ها / کل دارایی‌ها	باروا و سها (2015)
نسبت پوشش هزینهٔ ثابت x5	سود قبل از بهره و مالیات / هزینهٔ بهره + بدهی کوتاه‌مدت	باروا و سها (2015)
بازدهٔ فروش x6	سود خالص / فروش	باروا و سها (2015)
بازدهٔ حقوق صاحبان سهام x7	سود خالص / حقوق صاحبان سهام	باروا و سها (2015)
بازدهٔ سرمایه به کار گرفته‌شده x8	سود / حقوق صاحبان سهام + بدهی‌های بلندمدت	باروا و سها (2015)
سرمایه، سرمایه‌گذاری شده به ازای هر ریال تأمین مالی x9	سرمایهٔ سرمایه‌گذاری‌شده / کل منابع تأمین مالی	باروا و سها (2015)
حاشیهٔ سود ناخالص x10	تغییرات حاشیهٔ سود ناخالص – تغییرات فروش	سنگ و هنکاک (2012)

متغیر	تعریف عملیاتی	منبع
هزینه‌های عمومی اداری و فروش x11	تغییرات هزینهٔ عمومی اداری و فروش – تغییرات فروش	سنگ و هنکاک (2012)
مالیات x12	هزینهٔ مالیات که از صورت سود و زیان قابل‌استخراج است.	آرتور و همکاران (2010)
هزینهٔ بهره x13	هزینهٔ مالی که از صورت سود و زیان قابل‌استخراج است.	آرتور و همکاران (2010)
سایر درآمدها x14	سایر درآمدها که از صورت سود و زیان قابل‌استخراج است.	آرتور و همکاران (2010)
بهای تمام‌شده x15	بهای تمام‌شده که از صورت سود و زیان قابل‌استخراج است.	مولنگا و باتیا (2017)
فروش x16	فروش که از صورت سود و زیان قابل‌استخراج است.	مولنگا، باتیا (2017)
تغییرات بدهی x17	تغییرات کل بدهی‌ها (کوتاه‌مدت + بلندمدت)	حسینی مهر و نوری فرد (2014)
انتشار سهام x18	تغییرات در سود انباشته – تغییرات در حقوق صاحبان سهام	حسینی مهر و نوری فرد (2014)
خالص سرمایه در گردش x19	بدهی جاری – دارایی جاری	حسینی مهر و نوری فرد (2014)
سود خالص قبل از اقلام فوق‌العاده به‌اضافهٔ استهلاک x20	استهلاک + سود خالص قبل از اقلام غیرمترقبه	الضمری و اسماعیل (2014)
سود خالص به‌علاوه استهلاک x21	سود خالص + استهلاک	مولنگا، باتیا (2017)
اقلام تعهدی بلندمدت x22	تعهدات بلندمدت غیر بدهی – دارایی جاری – کل دارایی‌ها	لولن و رسوتک (2019)
سرمایه در گردش عملیاتی x23	بدهی جاری عملیاتی – دارایی جاری عملیاتی	مولنگا، باتیا (2017)
سود عملیاتی x24	سود قبل از بهره و مالیات	اوز و همکاران (2021)
استهلاک دارایی‌های مشهود x25	استهلاک دارایی‌های ثابت	اوز و همکاران (2021)
اقلام تعهدی کل x26	جریان نقد عملیاتی – سود خالص	اوز وهمکاران (2021)
تغییرات حساب‌های دریافتنی x27	تغییرات حساب‌های دریافتنی	اوز وهمکاران (2021)
تغییرات حساب‌های پرداختنی x28	تغییرات حساب‌های پرداختنی	اوز و همکاران (2021)
تغییرات موجودی کالا x29	تغییرات موجودی کالا	اوز و همکاران (2021)
سود تقسیمی پرداختنی x30	سود تقسیمی پرداختنی که از صورت وضعیت مالی استخراج می‌شود	کوورسکی (2013)
گردش دارایی‌های ثابت x31	$\frac{\text{فروش}}{\text{خالص دارایی ثابت}}$	باروا و سها (2015)
بازپرداخت وام x32	$\frac{\text{بازپرداخت وام}}{\text{سود خالص}}$	باروا و سها (2015)
نسبت پوشش بدهی x33	$\frac{\text{کل بدهی}}{\text{سود خالص}}$	باروا و سها (2015)
نسبت پوشش بهره x34	$\frac{\text{سود قبل از بهره و مالیات}}{\text{هزینهٔ مالی}}$	باروا و سها (2015)
سود خالص x35	سود خالص که از صورت سود و زیان قابل‌استخراج است.	اوز و همکاران (2021)
نسبت وجوه نقد به بدهی کوتاه‌مدت x36	$\frac{\text{وجه نقد عملیاتی}}{\text{بدهی کوتاه‌مدت}}$	باروا و سها (2015)
نسبت پوشش نیازهای حیاتی شرکت x37	$\frac{\text{وجه نقد عملیاتی + هزینهٔ مالی پرداختی}}{\text{هزینهٔ مالی پرداختی + بدهی جاری + سود تقسیمی پرداختی}}$	باروا و سها (2015)

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

متغیر	تعریف عملیاتی	منبع
نسبت وجوه نقد به دارایی ثابت x38	وجه نقد عملیاتی / دارایی ثابت	باروا و سها (2015)
نسبت سرمایه‌گذاری مجدد x39	جریان نقد عملیاتی / سرمایهٔ به کار گرفته‌شده	باروا و سها (2015)
نسبت وجه نقد به بدهی x40	جریان نقد عملیاتی / بدهی کوتاه‌مدت	باروا و سها (2015)
نسبت پوشش هزینهٔ نقدی ثابت x41	جریان نقد عملیاتی / هزینهٔ مالی + پرداخت بدهی+ سود سهام ممتاز	باروا و سها (2015)
نسبت وجه نقد به فروش x42	جریان نقد عملیاتی / فروش	باروا و سها (2015)
نسبت وجه نقد به حقوق صاحبان سهام x43	جریان نقد عملیاتی / حقوق صاحبان سهام	باروا و سها (2015)
نسبت وجه نقد به سرمایهٔ به کار گرفته‌شده x44	جریان نقد عملیاتی / بدهی بلند مدت + سهام صاحبان حقوق	باروا و سها (2015)
نسبت پرداخت سود تقسیمی x45	سود تقسیمی پرداختی / جریان نقد عملیاتی	باروا و سها (2015)
سرمایهٔ سرمایه‌گذاری‌شده به‌ازای هر ریال تأمین مالی x46	نقد پرداختی برای سرمایه‌گذاری / کل منابع تأمین مالی	باروا و سها (2015)
سود تقسیمی دریافتی x47	سود تقسیمی دریافت‌شده که از صورت جریان وجوه نقد قابل‌استخراج است.	آرتور و همکاران (2010)
بهره پرداختی نقدی x48	هزینهٔ بهرهٔ پرداختی که از طریق صورت جریان‌های نقدی قابل‌استخراج است.	آرتور و همکاران (2010)
بهرهٔ دریافتی نقدی x49	بهرهٔ دریافتی که از صورت جریان‌های نقدی قابل‌استخراج است.	آرتور و همکاران (2010)
مالیات پرداختی نقدی x50	مالیات بر درآمد پرداختی که از صورت جریان نقدی قابل‌استخراج است.	آرتور و همکاران (2010)
فروش نقدی x51	تغییرات در حساب‌های دریافتنی – فروش	چنگ و هالی (2008)
بهای تمام‌شدهٔ نقدی x52	تغییرات حساب پرداختنی - تغییرات موجودی کالا - بهای تمام‌شدهٔ کالای فروش‌رفته	چنگ و هالی (2008)
جریان نقد عملیاتی x53	جریان نقدی حاصل از عملیات که از صورت جریان‌های نقدی قابل‌استخراج است.	شوبیت (2021)
جریان نقد سرمایه‌گذاری x54	جریان نقدی حاصل از فعالیت‌های سرمایه‌گذاری که از صورت جریان‌های نقدی قابل دریافت است.	شوبیت (2021)
جریان نقد تأمین مالی x55	جریان نقدی حاصل از تأمین مالی که از صورت جریان‌های نقدی قابل‌استخراج است.	شوبیت (2021)
جریان نقد آزاد x56	خالص مخارج سرمایه‌ای – وجه نقد عملیاتی	شوبیت (2021)
نسبت پوشش بدهی نقدی x57	کل بدهی / جریان نقد عملیاتی	باروا و سها (2015)

متغیر	تعریف عملیاتی	منبع
بازپرداخت وام نقدی x58	بازپرداخت وام / جریان نقد عملیاتی	باروا و سها (2015)
نسبت پوشش بهره نقدی x59	جریان نقد عملیاتی / بهره پرداختی	باروا و سها (2015)

رگرسیون نمادین یک مدل رگرسیون عملکردی است که الگوریتم ژنتیک را بهبود می‌بخشد. رگرسیون نمادین شبیه به یک دانشمند رباتیک به‌طور خودکار روابطی را براساس ویژگی‌های داخلی داده‌ها برقرار می‌کند. مدل، عملکردی با بیشترین برازش را برای نشان‌دادن این رابطه پیدا و پارامترها و ساختار هر مدل رگرسیونی را تعیین می‌کند. هستهٔ رگرسیون نمادین براساس نظریهٔ تکامل داروین است که عوامل مهم را انتخاب می‌کند تا به‌تدریج مدلی را شکل دهد و به‌طور خودکار عوامل غیر مهم را از مدل جدا کند. این نظریه براساس اصل برنامه‌ریزی ژنتیکی است. این رویکرد به پژوهشگران کمک می‌کند تا سهم هر عامل را در یک مدل تعیین کنند. به عبارت دیگر، وجود عوامل، بیانگر فراوانی و وقوع نشان‌دهندهٔ اهمیت است (Liu & Zhang, 2022).

برای آزمون فرضیهٔ پژوهش، مقایسهٔ مدل‌های آماری و یادگیری ماشین از معیار AUC[1] استفاده شده است. معیار AUC کل ناحیهٔ دوبعدی زیر منحنی ROC را از (۰،۰) تا (۱،۱) اندازه‌گیری می‌کند. AUC معیاری کلی از عملکرد در تمام آستانه‌های طبقه‌بندی ممکن ارائه می‌دهد. AUC به‌عنوان معیاری تک عددی جایگزینی برای ارزیابی و مقایسهٔ الگوریتم‌های یادگیری ماشین پیشنهاد شده و معیار بهتری نسبت به دقت است. علاوه بر این، ازنظر آماری نیز سازگار است (Ling et al., 2003).

بنابراین مطابق با جونز و همکاران عملکرد پیش‌بینی‌کنندهٔ همهٔ طبقه‌بندی‌کننده‌ها با استفاده از ناحیهٔ زیر منحنی (AUC) مقایسه و نتایج برای پیش‌بینی سود خالص و جریان نقد عملیاتی به‌صورت کلی، مقطعی و طولی اندازه‌گیری شده است (Jones et al., 2017).

یافته‌ها

ابتدا، برای متغیرهای سود خالص و جریان نقد عملیاتی افزایش و کاهش نسبت به سال قبل به‌عنوان یک و صفر کدگذاری شده است؛ زیرا از الگوریتم‌های طبقه‌بندی یادگیری ماشین استفاده شده است؛ بنابراین متغیرهای وابسته از نوع باینری و متغیرهای مستقل کمی است؛ به همین دلیل دو جدول جداگانه برای آمار توصیفی در جدول‌های (۲) و (۳) ارائه شده است.

جدول (۲) تحلیل توصیفی متغیرهای پژوهش

Table (2) Descriptive analysis of variables

متغیر	میانگین	میانه	انحراف معیار	بیشینه	کمینه	چولگی	کشیدگی	آمارهٔ جارک برا	احتمال آماره
x1	۱/۵۰۵	۱/۳۱۱	۰/۷۴۰	۳/۵۳۹	۰/۶۰۱	۱/۳۱۱	۴/۲۱۷	۶۷۰/۲۵۹	۰/۰۰۰
x2	۰/۹۶۷	۰/۸۴۶	۰/۵۵۷	۲/۴۸۱	۰/۲۸۵	۱/۲۲۸	۴/۰۶۷	۴۹۲/۷۱۶	۰/۰۰۰
x3	۱/۱۵۱	۰/۹۴۰	۰/۷۲۲	۳/۲۱۳	۰/۳۰۱	۱/۳۵۲	۴/۲۷۷	۷۳۵/۷۱۵	۰/۰۰۰
x4	۰/۵۶۱	۰/۵۶۳	۰/۱۹۰	۰/۸۸۶	۰/۲۱۱	−۰/۱۳۰	۲/۱۱۷	۴۵/۹۵۲	۰/۰۰۰
x5	۰/۶۷۶	۰/۳۴۱	۰/۸۵۱	۳/۱۴۸	−۰/۰۹۹	۱/۷۲۹	۵/۱۲۷	۱۳۲۲/۹۱۷	۰/۰۰۰
x6	۰/۱۸۲	۰/۱۴۲	۰/۱۸۳	۰/۶۰۶	−۰/۰۸۴	۰/۷۴۴	۲/۷۸۴	۱۳۴/۵۶۸	۰/۰۰۰
x7	۰/۳۰۸	۰/۳۰۷	۰/۲۴۹	۰/۷۳۸	−۰/۱۶۷	−۰/۰۲۶	۲/۱۲۳	۴۷/۸۹۹	۰/۰۰۰
x8	۰/۳۳۱	۰/۳۱۰	۰/۲۵۳	۰/۸۲۵	−۰/۱۱۸	۰/۲۱۹	۲/۲۸۳	۴۶/۳۱۷	۰/۰۰۰

[1]. Area under the ROC Curve

متغیر	میانگین	میانه	انحراف معیار	بیشینه	کمینه	چولگی	کشیدگی	آمارهٔ جارک برا	احتمال آماره
x9	۰/۰۵۰	۰/۰۱۷	۰/۰۹۳	۰/۳۲۱	−۰/۰۵۶	۱/۶۴۹	۵/۱۳۶	۱۰۶۶/۵۳۵	۰/۰۰۰
x10	۰/۲۴۰	۰/۱۴۷	۰/۳۳۴	۱/۰۳۹	−۰/۲۲۳	۰/۹۰۸	۳/۰۷۰	۶۷۰/۱۵۵	۰/۰۰۰
x11	−۰/۲۲۶	−۰/۱۳۷	۰/۳۲۵	۰/۲۳۰	−۱/۰۱۰	−۰/۹۱۰	۳/۱۱۰	۷۶۶/۳۲۱	۰/۰۰۰
x12	۰/۰۲۷	۰/۰۱۷	۰/۰۳۰	۰/۱۰۲	۰/۰۰۰	۱/۰۹۴	۳/۱۶۸	۵۵۴/۵۶۱	۰/۰۰۰
x13	۰/۰۴۱	۰/۰۳۲	۰/۰۳۷	۰/۱۲۵	۰/۰۰۰	۰/۸۴۷	۲/۷۲۹	۸۰/۷۹۸	۰/۰۰۰
x14	۰/۰۲۹	۰/۰۱۲	۰/۰۴۲	۰/۱۵۴	−۰/۰۱۵	۱/۶۵۷	۵/۰۷۳	۸۰۳/۸۶۰	۰/۰۰۰
x15	۰/۹۳۹	۰/۷۰۳	۰/۶۸۱	۲/۷۷۳	۰/۲۳۵	۱/۳۵۱	۳/۹۸۸	۱۳۳۱/۸۶۱	۰/۰۰۰
x16	۱/۲۵۸	۱/۰۳۷	۰/۷۵۹	۳/۲۰۶	۰/۳۷۹	۱/۱۳۸	۳/۵۲۴	۱۰۶/۷۷۴	۰/۰۰۰
x17	۰/۱۱۰	۰/۰۷۰	۰/۱۶۹	۰/۵۱۱	−۰/۱۳۶	۰/۸۴۷	۳/۰۳۰	۳۹/۲۳۷	۰/۰۰۰
x18	۰/۰۴۸	۰/۰۰۰	۰/۰۹۳	۰/۳۳۵	۰/۰۰۰	۲/۰۸۷	۶/۲۴۰	۴۰/۸۶۳	۰/۰۰۰
x19	۰/۲۱۲	۰/۱۹۵	۰/۲۹۶	۰/۸۱۳	−۰/۳۱۶	۰/۲۰۹	۲/۴۴۹	۲۶۴/۵۰۳	۰/۰۰۰
x20	۰/۲۶۶	۰/۲۱۶	۰/۲۱۷	۰/۷۸۱	−۰/۰۳۳	۰/۸۲۴	۲/۹۲۲	۸۳۶/۲۵۴	۰/۰۰۰
x21	۰/۲۴۱	۰/۱۸۳	۰/۲۲۵	۰/۸۰۵	−۰/۰۵۰	۱/۰۰۸	۳/۲۸۶	۱۴/۹۲۶	۰/۰۰۰
x22	۰/۴۳۵	۰/۳۷۴	۰/۳۰۸	۱/۲۲۹	۰/۰۶۲	۱/۰۱۸	۳/۴۴۰	۱۴۱/۹۰۸	۰/۰۰۰
x23	۰/۳۶۳	۰/۳۵۴	۰/۲۹۹	۰/۹۳۳	−۰/۱۶۷	۰/۱۲۵	۲/۲۷۷	۳۶۳/۱۷۲	۰/۰۰۰
x24	۰/۲۳۳	۰/۱۷۹	۰/۲۱۰	۰/۷۳۰	−۰/۰۵۹	۰/۸۲۳	۲/۸۹۸	۱۸۳/۱۹۸	۰/۰۰۰
x25	۰/۰۳۰	۰/۰۲۳	۰/۰۲۴	۰/۰۹۲	۰/۰۰۵	۱/۲۵۸	۳/۸۰۵	۱۱۰۹/۹۳۸	۰/۰۰۰
x26	۰/۰۷۶	۰/۰۴۶	۰/۱۷۰	۰/۴۶۸	−۰/۱۹۸	۰/۶۳۱	۲/۸۹۳	۷۱۰/۳۵۵	۰/۰۰۰
x27	−۰/۰۱۱	−۰/۰۰۴	۰/۱۵۹	۰/۳۱۴	−۰/۳۵۲	−۰/۱۲۹	۳/۰۳۰	۲۱۷/۴۶۸	۰/۰۰۰
x28	−۰/۰۰۸	−۰/۰۰۳	۰/۱۳۱	۰/۲۵۰	−۰/۳۰۶	−۰/۳۰۲	۳/۲۷۱	۲۷۷۶/۹۹۹	۰/۰۰۰
x29	−۰/۰۰۱	۰/۰۰۱	۰/۱۰۸	۰/۲۱۸	−۰/۲۳۴	−۰/۱۶۴	۳/۰۷۹	۳۴۲/۲۷۱	۰/۰۰۰
x30	۰/۳۶۷	۰/۱۰۰	۰/۵۱۲	۱/۵۵۶	۰/۰۰۰	۱/۴۱۵	۳/۵۹۰	۲۶۳/۷۸۱	۰/۰۰۰
x31	۴/۸۹۹	۴/۰۲۲	۳/۰۷۵	۹/۵۲۷	۱/۵۱۸	۰/۴۰۸	۱/۶۰۹	۴۳۹/۱۸۵	۰/۰۰۰
x32	۰/۰۳۶	۰/۰۱۴	۰/۰۹۹	۰/۱۹۵	−۰/۰۸۷	۰/۴۵۲	۱/۹۹۸	۶۹۷/۹۰۶	۰/۰۰۰
x33	۴/۵۳۷	۳/۴۸۷	۳/۱۵۲	۹/۰۰۶	۱/۲۵۹	۰/۳۹۹	۱/۴۹۳	۴۶۴/۹۷۷	۰/۰۰۰
x34	۵/۹۷۷	۴/۵۸۷	۳/۹۷۴	۱۱/۱۵۳	۱/۹۱۹	۰/۳۰۷	۱/۳۲۷	۱۱۶/۲۱۷	۰/۰۰۰
x35	۰/۲۰۷	۰/۱۴۷	۰/۲۱۷	۰/۷۴۵	−۰/۰۸۰	۱/۰۰۲	۳/۲۴۶	۱۲۷۳/۲۳۳	۰/۰۰۰
x36	۰/۳۱۰	۰/۱۹۴	۰/۳۸۲	۱/۳۳۴	−۰/۱۶۱	۱/۳۴۰	۴/۱۵۲	۲۴/۲۲۳	۰/۰۰۰
x37	۰/۲۴۵	۰/۱۲۴	۰/۳۸۷	۱/۲۹۸	−۰/۲۲۸	۱/۳۸۶	۴/۲۹۹	۲۰۳/۲۲۶	۰/۰۰۰
x38	۰/۶۰۱	۰/۳۸۸	۰/۸۱۰	۲/۸۲۲	−۰/۶۲۵	۱/۲۳۱	۴/۳۰۱	۳۳۱/۹۹۸	۰/۰۰۰
x39	۰/۱۱۷	۰/۰۹۷	۰/۱۱۸	۰/۳۷۳	−۰/۰۶۹	۰/۵۱۵	۲/۵۷۱	۲۹۳/۲۴۹	۰/۰۰۰
x40	۰/۲۶۸	۰/۱۶۷	۰/۳۳۲	۱/۱۸۳	−۰/۱۲۹	۱/۴۲۵	۴/۴۳۵	۳۳/۷۱۴	۰/۰۰۰
x41	۰/۴۵۹	۰/۲۲۵	۰/۶۲۸	۲/۳۰۲	−۰/۱۸۰	۱/۷۰۳	۵/۱۸۲	۲۰۶/۸۵۷	۰/۰۰۰
x42	۰/۱۴۶	۰/۱۱۲	۰/۱۵۹	۰/۵۰۴	−۰/۱۰۹	۰/۶۲۰	۲/۷۵۰	۵۶۱/۱۴۴	۰/۰۰۰
x43	۰/۲۷۷	۰/۲۳۹	۰/۲۹۸	۰/۹۲۶	−۰/۲۶۲	۰/۳۳۹	۲/۶۵۵	۱۱۴/۷۶۹	۰/۰۰۰
x44	۰/۲۳۳	۰/۲۰۴	۰/۲۴۲	۰/۷۳۸	−۰/۱۹۳	۰/۳۱۶	۲/۵۱۰	۲۱۴/۹۷۴	۰/۰۰۰
x45	۰/۳۴۲	۰/۲۰۴	۰/۴۷۶	۱/۵۲۳	−۰/۴۳۳	۰/۹۲۱	۳/۳۹۴	۹۴۹/۴۲۱	۰/۰۰۰
x46	۰/۰۴۰	۰/۰۲۵	۰/۰۴۲	۰/۱۵۸	۰/۰۰۲	۱/۵۶۲	۴/۶۲۴	۱۰/۰۵۵	۰/۰۰۷
x47	۰/۰۵۰	۰/۰۱۸	۰/۰۶۶	۰/۲۳۰	۰/۰۰۰	۱/۵۳۲	۴/۲۲۹	۸۱/۶۶۱	۰/۰۰۰
x48	۰/۰۳۷	۰/۰۲۷	۰/۰۳۵	۰/۱۱۹	۰/۰۰۰	۰/۹۳۸	۲/۹۰۷	۵/۳۶۴	۰/۰۶۸

متغیر	میانگین	میانه	انحراف معیار	بیشینه	کمینه	چولگی	کشیدگی	آمارهٔ جارک برا	احتمال آماره
x49	۰/۰۰۹	۰/۰۰۳	۰/۰۱۴	۰/۰۵۳	۰/۰۰۰	۲/۰۲۵	۶/۱۶۲	۲۹/۲۴۵	۰/۰۰۰
x50	۰/۰۲۱	۰/۰۱۶	۰/۰۲۱	۰/۰۷۴	۰/۰۰۰	۱/۰۴۴	۳/۲۳۱	۶/۳۵۷	۰/۰۴۲
x51	۰/۹۹۶	۰/۸۴۵	۰/۷۳۱	۲/۷۹۶	۰/۰۰۰	۰/۸۴۱	۳/۱۸۹	۴۶۰/۵۸۵	۰/۰۰۰
x52	۰/۶۸۱	۰/۵۰۸	۰/۶۰۰	۲/۲۷۲	۰/۰۰۰	۱/۱۹۸	۳/۸۲۳	۱۶۰/۵۲۵	۰/۰۰۰
x53	۰/۱۵۷	۰/۱۲۲	۰/۱۷۵	۰/۵۸۱	-۰/۱۱۱	۰/۸۲۱	۳/۱۸۷	۱۱۷/۰۰۴	۰/۰۰۰
x54	-۰/۰۵۶	-۰/۰۲۷	۰/۱۰۰	۰/۰۹۰	-۰/۳۳۶	-۱/۳۷۲	۴/۶۶۰	۱۵۳/۶۵۹	۰/۰۰۰
x55	-۰/۰۸۳	-۰/۰۷۱	۰/۱۳۰	۰/۱۶۸	-۰/۳۶۸	-۰/۳۰۳	۲/۹۲۱	۱۸۵/۴۹۳	۰/۰۰۰
x56	۰/۰۹۳	۰/۰۷۰	۰/۱۵۷	۰/۴۵۴	-۰/۱۷۴	۰/۵۸۲	۲/۹۹۴	۱۹۱/۲۴۵	۰/۰۰۰
x57	۴/۶۱۹	۳/۷۱۱	۳/۴۲۸	۹/۷۵۰	۰/۸۱۰	۰/۳۹۷	۱/۶۱۱	۱۷۷/۶۸۸	۰/۰۰۰
x58	۱/۳۳۵	۰/۶۹۲	۱/۴۴۶	۳/۷۱۰	۰/۰۰۰	۰/۶۹۸	۱/۸۶۰	۱۹۴/۴۰۰	۰/۰۰۰
x59	۵/۰۹۴	۳/۳۷۶	۴/۱۸۱	۱۱/۰۷۸	۰/۹۳۶	۰/۴۵۷	۱/۴۹۹	۳۲۷/۷۳۷	۰/۰۰۰

یکی از کاربردهای مهم جدول، آمار توصیفی بیان نوع توزیع داده‌هاست. با توجه به این‌که احتمال آمارهٔ آزمون جارک‌برا کمتر از ۵ درصد است، به‌جز متغیر بهرهٔ پرداختی نقدی، سایر متغیرهای این پژوهش دارای توزیع نرمال نیست. گفتنی است که برای خنثی‌کردن اثرات داده‌های پرت از تکنیک وینسورایزین (در سطح ۵ و ۹۵ درصد) استفاده شده است. انحراف معیار سود خالص و سود عملیاتی به‌ترتیب با مقدار ۰/۲۱۷ و ۰/۲۱۰ بیشتر از جریان نقدی ناشی از عملیات با مقدار ۰/۱۷۵ است که نشان‌دهندهٔ آن است که سود نوسان‌های بیشتری نسبت به جریان نقدی حاصل از عملیات دارد. میانهٔ جریان نقدی عملیاتی دارای مقدار ۰/۱۲۲ است که نشانهٔ خوبی در نظر گرفته می‌شود و به این معناست که این شرکت‌های بورس تهران با فعالیت‌های اساسی عملیاتی خود پول نقد تولید می‌کنند. علامت منفی برای جریان نقدی حاصل از فعالیت‌های سرمایه‌گذاری با مقدار ۰/۰۲۷- به این معنی است که شرکت‌های موردمطالعهٔ بورس تهران برای توسعهٔ کار خود در خرید دارایی‌های ثابت سرمایه‌گذاری می‌کنند. درنهایت، جریان نقدی منفی ناشی از فعالیت‌های مالی با مقدار ۰/۰۷۱- به این معناست که شرکت‌ها سرمایهٔ خود را کاهش می‌دهند یا گزینهٔ محتمل‌تر، نشان‌دهندهٔ پرداخت تعهدات این شرکت‌هاست.

جدول (۳) تحلیل توصیفی متغیرهای وابستهٔ پژوهش

Table (2) Descriptive analysis of dependent variables

نام متغیر	مقدار	تعداد	درصد
کاهش سود خالص	۰	۶۹۰	۴۵
افزایش سود خالص	۱	۸۴۳	۵۵
کاهش جریان نقد عملیاتی	۰	۷۴۸	۴۹
افزایش جریان نقد عملیاتی	۱	۷۸۵	۵۱

به‌منظور آزمون فرضیهٔ اول و دوم، متغیرهای وابسته با هرکدام از دستهٔ متغیرهای تعهدی و نقدی از طریق رگرسیون خاص داده‌های تابلویی برازش شده و سپس ضریب تعیین استخراج شده حاصل از برازش مدل‌ها از طریق آزمون ونگ مقایسه شده است. متغیرهای تعهدی پژوهش حاضر از X35 - X1 و متغیرهای نقدی از X59 - X36 است. معناداری آمارهٔ ونگ (۹/۲۰۸۲) برای پیش‌بینی سود خالص در سطح معناداری یک درصد نشان‌دهندهٔ این موضوع است که بین توان تبیین دو مدل، اختلاف معناداری وجود دارد و مدل با متغیرهای تعهدی نسبت به مدل با متغیرهای صورت جریان‌های نقدی از توان بیشتری برای تبیین برخوردار است. عدم معناداری آمارهٔ ونگ در سطح معناداری پنج درصد برای پیش‌بینی جریان نقد عملیاتی نشان‌دهندهٔ این موضوع است که اختلاف معناداری بین توان تبیین دو مدل وجود ندارد؛ اما هنوز هم ضریب متغیرهای تعهدی بیشتر از متغیرهای نقدی است؛ درنتیجه برای

پیش‌بینی سود و جریان نقد عملیاتی از طریق مدل‌های یادگیری ماشین و آماری از مجموعهٔ متغیرهای تعهدی استفاده شده است. نتایج حاصل از مقایسهٔ ضریب تعیین مدل‌ها در جدول (۴) ارائه شده است.

جدول (۴) نتایج ضریب تعیین حاصل از تخمین مدل‌ها و مقایسهٔ مدل‌ها با آزمون ونگ

Table (4) Results of R-squared obtained from estimation of models and comparison of models with Vuong test

	ضریب تعیین	ضریب تعیین تعدیل‌شده	آمارهٔ آزمون ونگ	احتمال آماره
تخمین سود خالص با مجموعهٔ متغیرهای تعهدی	۰/۶۰۸۸	۰/۵۹۸۵	۹/۲۰۸۲	۰/۰۰۰۰
تخمین سود خالص با مجموعهٔ متغیرهای نقدی	۰/۴۳۱۶	۰/۴۲۲۴		
تخمین جریان نقد عملیاتی با مجموعهٔ متغیرهای تعهدی	۰/۳۲۱۱	۰/۳۰۳۲	۱/۲۴۸۴	۰/۲۱۱۹
تخمین جریان نقد عملیاتی با مجموعهٔ متغیرهای نقدی	۰/۳۰۲۲	۰/۲۹۰۹		

نمونهٔ آزمایشی فقط برای ارزیابی عملکرد پیش‌بینی خارج از نمونه استفاده می‌شود؛ در حالی که مجموعهٔ آموزشی فقط برای تخمین مدل استفاده می شود؛ برای نمونه، در آزمون مقطعی، هیچ توجهی به ترتیب زمانی داده‌ها نمی شود. ممکن است داده‌های سال ۱۳۹۸ در نمونهٔ تخمینی و داده‌های سال ۱۳۹۱ در نمونهٔ آزمایشی وجود داشته باشد. جونز و همکاران[1] (2017) خاطرنشان می‌کنند، نمونه‌های آزمایشی که از همان دورهٔ زمانی نمونهٔ آموزشی برآورد گرفته شده است، هیچ اعتبار بین زمانی ارائه نمی‌دهد؛ از این رو، قادر نیست، آزمون مناسبی از توانایی پیش‌بینی یک مدل ارائه دهد. به‌منظور مقایسهٔ عملکرد طبقه‌بندی‌کننده‌ها در یک محیط پیش‌بینی واقعی‌تر، از نمونهٔ آزمایشی طولی نیز استفاده شده است؛ بنابراین مطابق با جونز و همکاران (2017) عملکرد پیش‌بینی‌کنندهٔ همهٔ طبقه‌بندی‌کننده‌ها با استفاده از ناحیهٔ زیر منحنی (AUC) مقایسه می‌شود. نتایج برای پیش‌بینی سود خالص و جریان نقد عملیاتی به‌صورت کلی که میانگین تخمین حالت مقطعی و طولی است، در جدول (۵) گزارش شده است.

جدول (۵) نتایج عملکرد مدل‌های آماری و یادگیری ماشین

Table (5) Performance results of statistical models and machine learning

مدل	پیش‌بینی سود، عملکرد کلی		پیش‌بینی جریان نقد، عملکرد کلی	
	AUC	رتبه	AUC	رتبه
بیز ساده [2]	۰/۶۴۶	۲۸	۰/۶۶۷	۱۶
پروبیت [3]	۰/۸۰۳	۵	۰/۷۹۰	۶
رگرسیون لجستیک [4]	۰/۷۹۹	۶	۰/۷۹۰	۴
حاشیه بزرگ سریع [5]	۰/۸۰۶	۴	۰/۷۹۰	۵
یادگیری عمیق [6]	۰/۷۸۸	۱۱	۰/۸۰۰	۲
درخت تصمیم [7]	۰/۶۰۹	۳۵	۰/۵۹۹	۲۴
جنگل تصادفی [8]	۰/۷۹۰	۹	۰/۷۱۷	۱۴
تقویت گرادیان [9]	۰/۷۹۸	۷	۰/۷۸۵	۸

1. Jones et al
2. Naive Bayes
3. probit
4. Logistic Regression
5. Fast Large Margin
6. Deep Learning
7. Decision Tree
8. Random Forest
9. Gradient Boosted Trees

مدل	پیش‌بینی سود، عملکرد کلی		پیش‌بینی جریان نقد، عملکرد کلی	
	AUC	رتبه	AUC	رتبه
ماشین‌بردار پشتیبان [1]	۰/۷۷۶	۱۳	۰/۷۸۹	۷
ماشین بردار پشتیبان کتابخانه‌ای [2]	۰/۶۴۵	۲۹	۰/۵۲۷	۳۸
ماشین‌بردار پشتیبان خطی [3]	۰/۸۱۸	۲	۰/۷۸۴	۹
ماشین‌بردار پشتیبان تکاملی [4]	۰/۶۲۰	۳۳	۰/۵۳۳	۳۶
ماشین‌بردار پشتیبان ازدحام ذرات [5]	۰/۶۰۶	۳۶	۰/۵۴۴	۳۵
آنالیز تشخیص خطی [6]	۰/۷۶۲	۱۵	۰/۷۰۹	۱۵
آنالیز تشخیص درجه دوم [7]	۰/۳۵۲	۴۱	۰/۵۰۴	۴۰
آنالیز تشخیص منظم [8]	۰/۳۵۴	۴۰	۰/۵۱۴	۳۹
شبکهٔ عصبی [9]	۰/۷۶۹	۱۴	۰/۷۲۶	۱۳
پروسپترون چندلایه [10]	۰/۷۹۰	۱۰	۰/۷۶۳	۱۰
پرسپترون [11]	۰/۷۷۹	۱۲	۰/۷۳۹	۱۱
درخت چاید [12]	۰/۷۰۱	۱۹	۰/۶۴۴	۱۸
درخت کوئست [13]	۰/۶۸۶	۲۱	۰/۶۳۶	۱۹
لیست تصمیم [14]	۰/۶۴۱	۳۰	۰/۶۱۳	۲۱
شبکه بیزین [15]	۰/۶۸۱	۲۴	۰/۵۶۴	۳۱
درخت سی ۵[16]	۰/۶۱۰	۳۴	۰/۵۸۱	۲۷
کا نزدیک‌ترین همسایه [17]	۰/۶۳۷	۳۱	۰/۵۶۰	۳۲
درخت کارت [18]	۰/۶۰۴	۳۷	۰/۵۷۹	۲۹
رگرسیون نمادین [19]	۰/۸۳۱	۱	۰/۸۱۱	۱
آی بی کا [20]	۰/۵۹۹	۳۸	۰/۵۲۸	۳۷
کا استار [21]	۰/۶۳۲	۳۲	۰/۵۴۶	۳۴
یادگیری با وزن محلی [22]	۰/۷۱۲	۱۸	۰/۶۱۳	۲۲
آدابوست [23]	۰/۷۱۴	۱۷	۰/۶۰۳	۲۳
جدول تصمیم [24]	۰/۶۸۴	۲۲	۰/۵۷۳	۳۰

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

1. Support Vector Machine
2. Library support vector machine
3. Linear SVM
4. Evolutionary SVM
5. PSO SVM
6. Linear discriminant analysis
7. Quadratic discriminant analysis
8. regularized discriminant analysis
9. neural network
10. A multilayer perceptron
11. perceptron
12. chaid
13. quest
14. decision list
15. bayesian network
16. C.
17. KNN Algorithm
18. CART tree
19. Symbolic Regression
20. IBK
21. Kstar
22. Locally weighted learning
23. AdaBoost
24. DecisionTable

مدل	پیش‌بینی سود، عملکرد کلی		پیش‌بینی جریان نقد، عملکرد کلی	
	AUC	رتبه	AUC	رتبه
هرس افزایشی مکرر[1]	۰/۶۶۶	۲۵	۰/۵۸۰	۲۸
پارت [2]	۰/۷۰۰	۲۰	۰/۶۱۶	۲۰
استامپ تصمیم [3]	۰/۶۶۵	۲۶	۰/۵۵۴	۳۳
درختان مدل لجستیک [4]	۰/۷۵۶	۱۶	۰/۶۵۳	۱۷
درخت هوفدینگ [5]	۰/۶۸۲	۲۳	۰/۵۸۵	۲۵
سی ۴,۵ [6]	۰/۶۴۸	۲۷	۰/۵۸۳	۲۶
ایکس جی بوست [7]	۰/۷۹۵	۸	۰/۷۳۳	۱۲
مارس [8]	۰/۸۱۳	۳	۰/۷۹۱	۳
کاکس [9]	۰/۴۲۹	۳۹	۰/۴۹۹	۴۱

نتایج حا صل از عملکرد مدل‌های آماری و یادگیری ما شین ن شان‌دهندهٔ آن بود که طبقه‌بندی‌کنندهٔ رگر سیون نمادین با ا ستفاده از الگوریتم ژنتیک برای پیش‌بینی سـود خالص، در عملکرد کلی با AUC= ۰/۸۳۱ و در نمونهٔ آزمایشـی طولی بهترین عملکرد را داشـته اسـت. این میزان از AUC نشـان‌دهندهٔ عملکرد طبقه‌بندی خوب و مفید اسـت. در رتبه‌بندی عملکرد کلی پس از رگرسـیون نمادین، به‌ترتیب ماشـین بردار پشـتیبان خطی با AUC = ۰/۸۱۸ در رتبهٔ دوم قرار می‌گیرد که عملکرد خوب و مفیدی ارائه می‌دهد؛ رگر سیون کمانک تطبیقی چند متغیره (مارس) با AUC= ۰/۸۱۳ در رتبهٔ سوم قرار می‌گیرد؛ طبقه‌بندی‌کنندهٔ حا شیهٔ بزرگ سریع نیز با AUC=۰/۸۰۶ در رتبهٔ چهارم قرار می‌گیرد؛ مدل آماری پروبیت و رگر سیون لجستیک نیز به‌ترتیب با AUC= ۰/۸۰۳ و AUC= ۰/۷۹۹ در رتبه‌های پنجم و شـشـم قرار می‌گیرند که عملکرد بسـیار نزدیک به یکدیگر دارند. سـایر طبقه‌بندی‌کننده‌ها در رتبه‌های بعدی قرار می‌گیرند که AUC کمتر از ۰/۸ دارند و نشـان از عملکرد ضـعیف و یا تصـادفی اسـت؛ در حالی که برخی فن‌های یادگیری ماشـین عملکرد برتری ارائه می‌دهد، روش‌های سـادهٔ آماری نیز عملکرد قابل‌قبول و بهتری را از برخی فن‌های معروف یادگیری ماشـین ارائه می‌دهد. طبقه‌بندی‌کنندهٔ رگرسـیون نمادین با اسـتفاده از الگوریتم ژنتیک برای پیش‌بینی جریان نقد عملیاتی، در عملکرد کلی با AUC= ۰/۸۱۱ و در نمونهٔ آزمایشی طولی، بهترین عملکرد را داشته است. این میزان از AUC نشان‌دهندهٔ عملکرد طبقه‌بندی خوب و مفید ا ست. در رتبه‌بندی عملکرد کلی پس از رگر سیون نمادین به‌ترتیب طبقه‌بندی‌کنندهٔ یادگیری عمیق با AUC= ۰/۸۰۰ در رتبهٔ دوم قرار می‌گیرد که عملکرد خوب و مفید ارائه می‌دهد. رگرسیون مارس با AUC=۰/۷۹۱ در رتبهٔ سوم قرار می‌گیرد که عملکرد به نسبت ضـعیفی را ارائه می‌دهد؛ اما از یک طبقه‌بندی تصـادفی بهتر عمل می‌کند. سـایر طبقه‌بندی‌کننده‌ها در رتبه‌های بعدی قرار می‌گیرند که نشان از عملکرد ضعیف یا تصادفی است.

درنهایت، مقایسه‌ای کلی میان مدل‌های یادگیری ما شین و آماری انجام شده ا ست که در جدول (۶) نتایج آن گزارش شده ا ست. نتایج جدول (۶) نشـان‌دهندهٔ آن اسـت که با توجه به این‌که احتمال آمارهٔ آزمون یک‌سـویه کمتر از پنج درصـد اسـت، دقت عملکرد مدل‌های یادگیری ماشین از مدل‌های آماری به‌طور معناداری بیشتری است.

1. Repeated Incremental Pruning
2. PART
3. Decision Stump
4. logistic model trees
5. HoeffdingTree
6. C.4.5
7. Xgboost
8. MARS
9. COX

جدول (۶) نتایج مقایسهٔ میانگین مدل‌های آماری و یادگیری ماشین

Table (6) Results of comparing the average of statistical models and machine learning

	آمارهٔZ	احتمال آمارهٔ آزمون یک‌سویه	احتمال آمارهٔ آزمون دوسویه
مقایسهٔ عملکرد مدل برای پیش‌بینی جریان نقد عملیاتی	۲/۴۴۵	۰/۰۰۷	۰/۰۱۴
مقایسهٔ عملکرد مدل‌ها برای پیش‌بینی سود خالص	۴/۳۴۵	۰/۰۰۰	۰/۰۰۰

نتایج جدول (۶) نشان‌دهندهٔ آن است که با توجه به این‌که احتمال آمارهٔ آزمون یک سویه کمتر از پنج درصد است، دقت عملکرد مدل‌های یادگیری ماشین از مدل‌های آماری به‌طور معناداری بیشتری است.

نتایج و پیشنهادها

توانایی بنگاه اقتصادی در تولید سود و جریان نقد آتی جزء مهمی از فرآیند و برنامه‌ریزی سرمایه‌گذاران، اعتباردهندگان و سایر ذی‌نفعان است. از گذشتهٔ دور تاکنون موضوع پیش‌بینی سود و جریان نقد آتی از موضوعات داغ موردبررسی پژوهشگران و فعالان حوزهٔ مالی بوده است. این امر به این دلیل است که از شاخص‌های عملکردی سود و جریان نقد عملیاتی می‌توان در حوزه‌های مختلف ازجمله مدل‌های ارزش‌گذاری بنگاه اقتصادی، بررسی توانایی پرداخت تعهدات، بررسی میزان ریسک بنگاه، تجزیه‌وتحلیل عملکرد شرکت و تصمیمات مدیریت برای بهبود کارایی شرکت استفاده کرد. در این صورت اگر بتوان با دقت قابل‌قبولی سود و جریان نقد عملیاتی آتی را پیش‌بینی کرد، بخش قابل‌توجهی از نیاز به محتوای اطلاعاتی در سرمایه‌گذاران، اعتباردهندگان و سایر ذی‌نفعان پاسخ داده می‌شود.

در راستای آزمون فرضیهٔ اول و دوم پژوهش مبنی بر این‌که توانایی توضیح‌دهندگی متغیرهای تعهدی برای سود خالص و جریان نقد عملیاتی آتی بیشتر از متغیرهای نقدی است، پس از تخمین مدل با رویکرد مناسب داده‌های تابلویی، ضریب تعیین مدل‌ها با یکدیگر مقایسه شده است. نتایج بررسی‌ها نشان‌دهندهٔ آن بود که برای سود خالص متغیرهای تعهدی توانایی توضیح‌دهندگی بیشتری دارد؛ بنابراین دلیلی بر رد فرضیهٔ اول پژوهش وجود ندارد؛ اما دربارهٔ جریان نقد عملیاتی اگرچه مقدار تبیین متغیرهای تعهدی بیشتر از متغیرهای نقدی است، ازنظر آماری تفاوت معناداری بین تبیین متغیرهای تعهدی و نقدی وجود ندارد؛ بنابراین فرضیهٔ دوم پژوهش رد شد. این نتایج با یافته‌های بال و نیکولایف و بیانیهٔ شمارهٔ یک مفاهیم حسابداری مبنی بر بهتربودن سیستم تعهدی از سیستم نقدی به‌عنوان معیاری برای سنجش عملکرد شرکت در یک راستاست (Ball & Nikolaev, 2020). در تشریح این موضوع می‌توان گفت که حسابداری تعهدی سعی می‌کند، آثار مالی رویدادها را در زمان تحقق رویداد ثبت کند، نه‌تنها در زمانی که وجه نقد دریافت شود؛ بنابراین گفته می‌شود، حسابداری تعهدی مشکلات زمان‌بندی و تطابق موجود در جریان‌های نقدی را ندارد؛ اما در هر حال اقلام تعهدی به احتمال حاوی خطای اندازه‌گیری است.

در راستای آزمون فرضیهٔ سوم و چهارم پژوهش حاضر مبنی بر این‌که مدل‌های یادگیری ماشین نسبت به مدل‌های آماری عملکرد برتری را در پیش‌بینی سود خالص و جریان نقد عملیاتی ارائه می‌دهد، معیار AUC با اجرای مدل‌های آماری و یادگیری ماشین استخراج شده و سپس درصد موفقیت مدل‌های آماری و یادگیری ماشین توسط آزمون مقایسهٔ نسبت‌ها مقایسه شد. بررسی نتایج نشان‌دهندهٔ آن بود که مدل‌های یادگیری ماشین به‌طور معناداری از مدل‌های آماری در پیش‌بینی سود خالص و جریان نقد عملیاتی عملکرد برتری دارد؛ بنابراین دلیلی بر رد فرضیهٔ سوم و چهارم پژوهش وجود ندارد. نتایج پژوهش با یافته‌های اوز و همکاران (2020)، چن و همکاران[1] (2022)، آشتاب و همکاران (2017)، حجازی (2012) مطابقت دارد. بهترین عملکرد بین مدل‌های یادگیری ماشین مربوط به رگرسیون نمادین با استفاده از الگوریتم ژنتیک و بهترین عملکرد بین مدل‌های آماری مربوط به مدل پروبیت است.

پیش‌بینی دقیق سود خالص و جریان‌های نقد عملیاتی برای ارزیابی ریسک، مدیریت ریسک مالی و برنامه‌ریزی مالی استراتژیک بسیار مهم است. درک نقاط قوت و محدودیت‌های این مدل‌ها به مدیران در تصمیم‌گیری آگاهانه و کاهش خطرات کمک می‌کند. بنا

[1]. Chen et al

بر نتایج پژوهش حاضر پیشنهاد می‌شود، مدیران نهادهای مالی به مدل‌های یادگیری ماشین و به‌طور خاص مدل رگرسیون نمادین توجه داشته باشند.

در ادامه، برای انجام پژوهش‌های هم‌راستا با موضوع این پژوهش و توسعهٔ آن پیشنهادهایی به پژوهشگران آتی ارائه می شود. در پژوهش حاضر، از متغیرهای مالی و حسابداری استفاده شده است. به پژوهشگران توصیه می‌شود، پیش‌بینی سود را با نقش متغیرهای کلان اقتصادی در عملکرد شرکت بررسی کنند. پژوهش‌های آینده پتانسیل ترکیب مدل‌های آماری و یادگیری ماشین را برای استفاده از نقاط قوت و کاهش محدودیت‌های آن‌ها بررسی می‌کند. رویکردهای ترکیبی روشی جدید برای بهبود نتایج پیش‌بینی است.

عملکرد مدل‌های آماری و یادگیری ماشین در پیش‌بینی سود خالص و جریان‌های نقد عملیاتی ممکن است در صنایع، بخش‌ها یا شرایط اقتصادی مختلف متفاوت باشد. اذعان به این نکته مهم است که یافته‌های این مطالعه ممکن است به‌طورکلی قابل‌اجرا نباشد و ممکن است نیاز به سفارشی سازی یا اعتبار سنجی در زمینه‌ها یا صنایع خاص داشته باشد. مدل‌های یادگیری ماشین به علت فرآیند آموزش و آزمایش به تعداد داده‌های در دسترس متکی هستند؛ بنابراین محدودیت پژوهش حاضر در کمبود داده‌هاست.

منابع

آشتاب، علی، حقیقت، حمید، و کردستانی، غلامرضا (۱۳۹۶). مقایسهٔ دقت مدل‌های پیش‌بینی بحران مالی و تأثیر آن بر ابزارهای مدیریت سود. *بررسی‌های حسابداری و حسابرسی*، ۲۴(۲)، ۱۴۷-۱۷۲. https://doi.org/10.22059/acctgrev.2017.231176.1007585

حبیب‌زاده، ملیحه، و ایزدپور، مصطفی (۱۳۹۹). پیش‌بینی سودآوری با رویکرد شبکهٔ عصبی و مقایسهٔ آن با ماشین‌بردار پشتیبان (SVM) و درخت تصمیم C5. *دانش مالی تحلیل اوراق بهادار*، ۱۳(۴۶)، ۳۹-۵۶.

حجازی، رضوان، محمدی، شاپور، اصلانی، زهرا، و آقاجانی، مجید (۱۳۹۱). پیش‌بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. *بررسی‌های حسابداری و حسابرسی*, ۱۹(۲), ۳۱-۴۶. http://doi/org/10.22059/acctgrev.2012.29198

خدامی‌پور، احمد، پوراحمد، رضا، و ترک‌زاده ماهانی، علی (۱۳۹۲). بررسی تأثیر وقفه‌های زمانی متغیرهای مالی در پیش‌بینی سود عملیاتی و جریان‌های نقدی عملیاتی. *حسابداری مدیریت*، ۶(۳)، ۴۵-۵۶.

رجب‌زاده، حامد، گرگانلی دوجی، جمادوردی، نادریان، آرش، و اشرفی، مجید (۱۴۰۱). پیش‌بینی جریان وجه نقد عملیاتی شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران با استفاده از روش یادگیری ماشین. *حسابداری مدیریت*، ۱۵(۵۲)، ۵۹-۷۸. https://doi.org/10.30495/jma.2022.20205

مهربان‌پور، محمدرضا، و حبیب‌زاده، ملیحه (۱۳۹۷). خوشه‌بندی و پیش‌بینی سودآوری شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران با رویکرد درخت تصمیم C5. *مطالعات تجربی حسابداری مالی*. ۱۵(۵۹)، ۱۳۵-۱۵۷.

References

Ahn, S. C., Horenstein, A. R., & Ma, Y. (2020). Efficient principal component regression estimation with heteroscedastic errors and general singular covariance matrix. *Journal of Econometrics*, 214(1), 30-52. https://doi.org/10.1016/j.jeconom.2019.12.007

Al-Dhamari, R. A., & Ismail, K. N. I. K. (2014). An investigation into the effect of surplus free cash flow, corporate governance and firm size on earnings predictability. *International Journal of Accounting and Information Management*, 22(2), 118-133. https://doi.org/10.1108/IJAIM-05-2013-0037

Anand, V., Brunner, R., Ikegwu, K., & Sougiannis, T. (2019). Predicting profitability using machine learning. *Available at SSRN 3466478*. https://dx.doi.org/10.2139/ssrn.3466478

Arthur, N., Cheng, M., & Czernkowski, R. (2010). Cash flow disaggregation and the prediction of future earnings. *Accounting & Finance*, 50(1), 1-30. https://doi.org/10.1111/j.1467-629X.2009.00316.x

Ashtab, A., Haghighat, H., & kordestani, G. (2017). Comparison of financial distress prediction models accuracy and its effect on earnings management tools. *Accounting and Auditing Review*, 24(2), 147-172. https://doi.org/10.22059/acctgrev.2017.231176.1007585 [In Persian].

Ball, R., & Nikolaev, V. V. (2020). FASB was right: Earnings beat cash flows when predicting future cash flows. *Chicago Booth Research Paper*, 20-23. https://dx.doi.org/10.2139/ssrn.3689802

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, 83, 405-417. https://doi.org/10.1016/j.eswa.2017.04.006

Barua, S., & Saha, A. K. (2015). Traditional ratios vs. cash flow-based ratios: Which one is better performance indicator. *Advances in Economics and Business*, 3(6), 232-251.

Bengio, Y., Courville, A., & Vincent, P. (2021). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1798-1828. https://doi.org/10.1109/TPAMI.2013.50

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* 4, (4), 738). springer.

Bradshaw, M. T., Drake, M. S., Myers, J. N., & Myers, L. A. (2012). A re-examination of analysts' superiority over time-series forecasts of annual earnings. *Review of Accounting Studies*, 17(4), 944-968. https://doi.org/10.1007/s11142-012-9185-8

Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123-140. https://doi.org/10.1007/BF00058655

Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M., & Elhadad, N. (2015). Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (1721-1730).* https://doi.org/10.1145/2783258.2788613

Chatfield, C. (2020). *The Analysis of Time Series: An Introduction*. CRC Press.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (785-794).* https://doi.org/10.1145/2939672.2939785

Chen, X., Cho, Y. H., Dou, Y., & Lev, B. (2022). Predicting future earnings changes using machine learning and detailed financial data. *Journal of Accounting Research*, 60(2), 467-515. https://doi.org/10.1111/1475-679X.12429

Cheng, C. S., & Hollie, D. (2008). Do core and non-core cash flows from operations persist differentially in predicting future cash flows?. *Review of Quantitative Finance and Accounting*, 31(1), 29-53. https://doi.org/10.1007/s11156-007-0062-7.

Christensen, T. E., Huffman, A., Lewis-Western, M. F., & Scott, R. (2022). Accruals earnings management proxies: Prudent business decisions or earnings manipulation? *Journal of Business Finance & Accounting*, 49(3-4), 536-587. https://doi.org/10.1111/jbfa.12585

Curtis, A., Li, V., & Patrick, P. H. (2021). The use of adjusted earnings in performance evaluation. *Review of Accounting Studies*, 26(4), 1290-1322. https://doi.org/10.1007/s11142-021-09580-1

Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, 91, 106263. https://doi.org/10.1016/j.asoc.2020.106263

FASB (1978). Statement of financial accounting concepts no. 1, objectives of financial reporting by business enterprises. *Norwalk, CT: Financial Accounting Standards Board*.

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

Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and Other Atories*. Cambridge University Press.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Habibzade, M., & Ezadpour, M. (2020). Using neural network approach to predict company's profitability and comparison with decision tree c5 and Support Vector Machine (SVM). *Financial Knowledge of Securities Analysis*, 13(46), 39-56. [In Persian].

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer.

Hejazi, R., Mohamadi, S., aslani, Z., & Aghajani, M. (2012). Earnings management prediction using neural networks and decision tree in TSE. *Accounting and Auditing Review*, 19(2), 31-46. [In Persian]. https://www.doi.org/10.22059/acctgrev.2012.29198

Hosseinimehr, S. G., & Nourifard, Y. (2014). The persistence of accruals and investment in operating and investment cycle. *Accounting and Finance Research*, 3(2), 92-111.

Huang, G. B., Zhou, H., Ding, X., & Zhang, R. (2017). Extreme learning machines: A survey. *International Journal of Machine Learning and Cybernetics*, 8(2), 513-531. https://doi.org/10.1007/s13042-016-0559-2

Jones, S., Johnstone, D., & Wilson, R. (2017). Predicting corporate bankruptcy: An evaluation of alternative statistical frameworks. *Journal of Business Finance & Accounting*, 44(1-2), 3-34. https://doi.org/10.1111/jbfa.12218

Khodamipour, A., Pourahmad, R. & Turkzadeh M. A. (2013). The influence of financial variables' time lags in predicting operating profit and operating cash flows. *Management Accounting*, 6(3), 45-56. [In Persian].

Kowerski, M. (2013). Dividends and earnings quality in Poland. *E-Finance*: *Financial Internet Quarterly*, 9(3), 42-51.

Lewellen, J., & Resutek, R. J. (2019). Why do accruals predict earnings?. *Journal of Accounting and Economics*, 67(2-3), 336-356. https://doi.org/10.1016/j.jacceco.2018.12.003

Li, F. (2010). The information content of forward-looking statements in corporate filingsa na¨ıve bayesian machine learning approach, *Journal of Accounting Research* 48(5): 1049– 1102. https://doi.org/10.1111/j.1475-679X.2010.00382.x

Ling, C. X., Huang, J., & Zhang, H. (2003). AUC: A statistically consistent and more discriminating measure than accuracy. In *Ijca*, 3, 519-524.

Liu, H., & Zhang, Z. (2022). Probing the carbon emissions in 30 regions of China based on symbolic regression and Tapio decoupling. *Environmental Science and Pollution Research*, 29(2), 2650-2663. https://doi.org/10.1007/s11356-021-15648-x

Martins, A. I. (2022). *Earnings prediction using machine learning methods and analyst comparison*. [Un Published Doctoral dissertation]. Universidade Católica Portuguesa.

MehrabanPour, M. R., & Habibzadeh, M. (2017). Clustering and forecasting the profitability of companies listed in the Tehran Stock Exchange with the C5 decision tree approach. *Empirical Studies of Financial Accounting*. 15(59), 135-157. [In Persian].

Mills, T. C. (2019). *The Econometric Modelling of Financial Time Series*. Cambridge University Press.

Monahan, S. J. (2018). Financial statement analysis and earnings forecasting. *Foundations and Trends in Accounting*, 12(2), 105-215.

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2019). *Introduction to Time Series Analysis and Forecasting*. John Wiley & Sons.

Mulenga, M., & Bhatia, M. (2017). The review of literature on the role of earnings, cash flows and accruals in predicting of future cash flows. *Accounting and Finance Research*, 6(2), 59-70.

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87-106.

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi Asl, R., & Yu, B. (2019). Interpretable machine learning: Definitions, methods, and applications. *Proceedings of the National Academy of Sciences*, 116(44), 22071-22080. https://doi.org/10.1073/pnas.1900654116

Nallareddy, S., Sethuraman, M., & Venkatachalam, M. (2020). Changes in accrual properties and operating environment: Implications for cash flow predictability. *Journal of Accounting and Economics*, 69(2-3), 101313. https://doi.org/10.1016/j.jacceco.2020.101313

Nguyen, H., & Nguyen, T. (2020). The prediction of future operating cash flows using accrual-based and cash-based accounting information: Empirical evidence from Vietnam. *Management Science Letters*, 10(3), 683-694. http://dx.doi.org/10.5267/j.msl.2019.9.010

Oz, I. O., Yelkenci, T., & Meral, G. (2021). The role of earnings components and machine learning on the revelation of deteriorating firm performance. *International Review of Financial Analysis*, 77, 101797. https://doi.org/10.1016/j.irfa.2021.101797

Parlina, N. D., & Budianto, E. (2021). Implementation of cash flow as a measuring tool in predicting future net income:(Case study at Kedai Nyobian 8 daily period September–October 2020). *Journal of Management, Accounting, General Finance and International Economic Issues*, 1(1), 16-24. https://doi.org/10.55047/marginal.v1i1.6

Petropoulos, A., Siakoulis, V., Stavroulakis, E., & Vlachogiannakis, N. E. (2020). Predicting bank insolvencies using machine learning techniques. *International Journal of Forecasting*, 36(3), 1092-1113. https://doi.org/10.1016/j.ijforecast.2019.11.005

Rajabzadeh, H., Gorganli davaji, J., Naderian, A., & Ashrafi, M. (2022). Forecast the operating cash flow of accepted companies In Tehran Stock Exchange using machine learning method. *Management Accounting*, 15(52), 59-78. https://www/doi.org/10.30495/jma.2022.20205 [In Persian].

Raschka, S., & Mirjalili, V. (2020). *Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-Learn, and TensorFlow 2*. Packt Publishing.

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

Seng, D., & Hancock, J. R. (2012). Fundamental analysis and the prediction of earnings. *International Journal of Business and Management*, 7(3), 32.

Sharawi, H. (2021). Earnings versus cash flows in predicting future cash flows: Evidence from Egypt and KSA. *Alexandria Journal of Accounting Research*, 1(1), 18-34.

Shubita, M. F. (2021). The ability of cash flows to predict the earning: Evidence from Jordan. *Investment Management & Financial Innovations*, 18(4), 36.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings?, *The Accounting Review*, 289–315.

Xinyue, C., Zhaoyu, X., & Yue, Z. (2020). Using Machine Learning to Forecast Future Earnings. *Atlantic Economic Journal*, 48(4), 543-545. https://doi.org/10.1007/s11293-020-09691-1

Zhang, X. F. (2007). Accruals, investment, and the accrual anomaly. *The Accounting Review*, 82(5), 1333-1363. https://doi.org/10.2308/accr.2007.82.5.1333

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی