



## Explaining the Earnings Management Prediction Model Using the Hybrid of Machine Learning Methods

**Hassan Hassani** : Ph.D. Student in Accounting, Faculty of Economics and Administrative Sciences, University of Mazandaran, Babolsar, Iran.  
[h.hassani13@yahoo.com](mailto:h.hassani13@yahoo.com)

**Esfandiar Malekian Kallehbasti** \* : Professor, Department of Accounting, Faculty of Economics and Administrative Sciences, University of Mazandaran, Babolsar, Iran.  
[e.malekian@umz.ac.ir](mailto:e.malekian@umz.ac.ir)

**Yahya Kamyabi** : Professor, Department of Accounting, Faculty of Economics and Administrative Sciences, University of Mazandaran, Babolsar, Iran.  
[kamyabi@umz.ac.ir](mailto:kamyabi@umz.ac.ir)

### Abstract

Knowledge of earnings management is essential for users of accounting information due to performance evaluation, profitability forecasting, and determining the true value of the company. The purpose of this research is to provide a model to diagnose accrual-based earnings management and real earnings management through performance evaluation of machine learning methods including decision tree, support vector machine, k-nearest neighbor, deep learning, and combining them with feature selection methods based on relief and principal component analysis. To achieve this goal, 180 companies admitted to the Tehran Stock Exchange were selected as a statistical sample from 2010 to 2021. Also, to test the hypotheses, the criteria of average accuracy and type I and type II errors were used. The results show that the performance of accrual-based earnings management forecasting methods based on the relief-based feature selection model is better than the feature selection model based on principal component analysis. This result was confirmed in all prediction methods. However, the results did not show the superiority of the relief-based feature selection model over the principal component analysis-based feature selection model in predicting real earnings management. Also, the findings showed that accrual earnings management can be more accurately predicted than real earnings management. The research results can be of interest to investors, creditors, financial analysts, and auditors. Incorporating machine learning methods can help identify potential earnings management activities.

**Keywords:** Accruals Earnings Management, Real Earnings Management, Earnings Management Forecast, Machine Learning, Feature Selection.

### Introduction

Earnings management can be described as the discretion utilized by managers to provide generally accepted accounting principles (GAAP)-based financial reports that can affect the relevance and reliability of the presented accounting information. EM can be performed either (1) through deviations from normal business practices to purposefully manipulate earnings; this is called real earnings management (Roychowdhury, 2006), and it affects cash flow from operating activities; or (b) by manipulating reported earnings through accruals, that is accrual-based earnings management, to achieve a suitable earnings figure. As a corporation's earnings are used

\* Corresponding author

2322-3405 © The author(s) Publisher: University of Isfahan



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



[10.22108/FAR.2024.142152.2058](https://doi.org/10.22108/FAR.2024.142152.2058)

by different financial statement users (such as shareholders, creditors, and financial analysts) to gauge its performance, detection of earnings management can be interesting and crucial for them. In this context, this study attempts to present prediction tools that aid in detecting earnings management activities. For this purpose, six machine learning methods have been discussed to predict earnings management.

### **Methods & Material**

A sample of 180 companies listed on the Tehran Stock Exchange during the period 2010-2021 was selected for testing hypotheses. The performance of each machine learning method at predicting accrual-based earnings management and real earnings management was evaluated based on three criteria: type I error, type II error, and average accurac.

### **Finding**

The results show that the performance of accrual-based earnings management forecasting methods based on the relief-based feature selection model is better than the feature selection model based on principal component analysis. This result was confirmed in all prediction methods. However, the results did not show the superiority of the relief-based feature selection model over the principal component analysis-based feature selection model in predicting real earnings management. Also, the findings showed that accrual earnings management can be more accurately predicted than real earnings management.

### **Conclusion & Results**

Earnings management would affect accounting data, in particular, the earnings reported in accounting other than the actual earnings of an enterprise. Therefore, the prediction of earnings management is still an issue of great importance. The purpose of this research is to use machine learning methods such as decision tree, support vector machine, k-nearest neighbor, and deep learning to predict earnings management. Also, this research relies on feature selection to identify the most optimal features for use in the prediction model. Even though the determinants of earnings management have been studied for a long time, the ability of these factors to predict earnings management has received less attention.

The results show that the combination of feature selection based on relief with each of the forecasting methods provides a more accurate performance for predicting accruals earnings management than the feature selection based on principal components analysis. However, the results did not show the superiority of the relief-based feature selection model over the principal component analysis-based feature selection model in predicting real earnings management. Also, the findings showed that accrual earnings management can be more accurately predicted than real earnings management. In addition, the results indicated that the most important features for prediction are related to the auditor's features in the first place and then to the features of the company's ownership structure. In other words, investors should pay a lot of attention to the features of the auditor and the ownership structure of companies in predicting earnings management. Based on the obtained results, the hybrid method based on deep learning and relief feature selection has the highest prediction accuracy (89/62) among other hybrid methods for forecasting accruals earnings management, and the hybrid method based on deep learning and principal component analysis feature selection has the highest prediction accuracy (82/65) among other hybrid methods for forecasting real earnings management.

The findings of this study can expedite earnings management detection for financial statement users by improving earnings management prediction accuracy. These results may be applied to reduce investment risks and losses and increase investment benefits for investors and creditors if they are better able to predict misleading financial reports due to earnings management. In summary, the results of this study suggest tools to decision makers that help in predicting earnings management with relatively high accuracy.

## تبیین الگوی پیش‌بینی مدیریت سود با استفاده از ترکیب روش‌های یادگیری ماشین

حسن حسینی <sup>id</sup>: دانشجوی دکتری حسابداری، دانشکده علوم اقتصادی و اداری، دانشگاه مازندران، بابلسر، ایران.

[h.hassani13@yahoo.com](mailto:h.hassani13@yahoo.com)

اسفندیار ملکیان کله بستی <sup>id</sup>\*: استاد، گروه حسابداری، دانشکده علوم اقتصادی و اداری، دانشگاه مازندران، بابلسر، ایران.

[e.malekian@umz.ac.ir](mailto:e.malekian@umz.ac.ir)

یحیی کامیابی <sup>id</sup>: استاد، گروه حسابداری، دانشکده علوم اقتصادی و اداری، دانشگاه مازندران، بابلسر، ایران.

[y.kamyabi@umz.ac.ir](mailto:y.kamyabi@umz.ac.ir)

### چکیده

شناخت مدیریت سود برای استفاده‌کنندگان از اطلاعات حسابداری به دلیل ارزیابی عملکرد، پیش‌بینی سودآوری و تعیین ارزش واقعی شرکت بسیار حائز اهمیت است. هدف از این پژوهش ارائه مدلی جهت تشخیص مدیریت سود اقلام تعهدی و مدیریت سود واقعی از طریق ارزیابی عملکرد با استفاده از روش‌های یادگیری ماشین از جمله درخت تصمیم، ماشین بردار پشتیبان، k-نزدیک‌ترین همسایه، یادگیری عمیق و ترکیب آنها با روش انتخاب ویژگی ریلیف و تحلیل مؤلفه‌های اصلی است. برای دستیابی به این هدف، تعداد ۱۸۰ شرکت پذیرفته‌شده در بورس تهران به‌عنوان نمونه آماری برای سال‌های ۱۳۸۹ تا ۱۴۰۰ انتخاب گردید. همچنین برای آزمون فرضیه‌ها از معیارهای میانگین صحت پیش‌بینی، خطاهای نوع اول و دوم استفاده گردید یافته‌های پژوهش بیانگر آن است عملکرد روش‌های پیش‌بینی مدیریت سود اقلام تعهدی بر اساس الگوی انتخاب ویژگی مبتنی بر ریلیف نسبت به الگوی انتخاب ویژگی مبتنی بر تحلیل مؤلفه‌های اصلی از توانای بهتری برخوردار است. این نتیجه در کلیه روش‌های پیش‌بینی مورد تایید قرار گرفت. اما نتایج برتری الگوی انتخاب ویژگی مبتنی بر ریلیف نسبت به الگوی انتخاب ویژگی مبتنی بر تحلیل مؤلفه‌های اصلی را در پیش‌بینی مدیریت سود واقعی نشان نداد. همچنین، یافته‌ها نشان دادند مدیریت سود اقلام تعهدی را می‌توان بادقت بالاتری نسبت به مدیریت سود واقعی پیش‌بینی کرد. نتایج پژوهش می‌تواند موردتوجه سرمایه‌گذاران، اعتباردهندگان، تحلیلگران مالی و حسابرسان قرار گیرد. استفاده از ترکیب روش‌های یادگیری ماشین، می‌تواند به شناسایی فعالیت‌های بالقوه مدیریت سود کمک کند.

**واژگان کلیدی:** مدیریت سود اقلام تعهدی، مدیریت سود واقعی، پیش‌بینی مدیریت سود، یادگیری ماشین، انتخاب ویژگی.



## مقدمه

شرکت‌های بزرگ سعی در جذب سرمایه از سرمایه‌گذاران و اعتباردهندگان با افشای منظم عملکرد مالی خود از طریق گزارش سود و با برآورده کردن پیش‌بینی‌های سود تحلیلگران مالی دارند (Degeorge et al., 1999). برای برآوردن صحیح انتظارات ذی‌نفعان مختلف (مانند سرمایه‌گذاران، اعتباردهندگان و تحلیلگران مالی) و جلوگیری از هر گونه عواقب نامطلوب ناشی از دست دادن این انتظارات مانند افت قیمت سهام و نقض قراردادهای بدهی (Braswell & Daniels, 2017)، مدیریت یک سازمان ممکن است تلاش کند عمداً بر سود گزارش‌شده در صورت‌های مالی تأثیر بگذارد (Rodriguez-Ariza et al., 2016). این روش به طور کلی به عنوان مدیریت سود شناخته می‌شود. با این حال، مدیریت سود ممکن است به از دست دادن شهرت (Rodriguez-Ariza et al., 2016) یا حتی دعاوی حقوقی (Gunny, 2010) منجر شود؛ زیرا ممکن است به عنوان شاخصی از فساد مدیریت یا تلاش برای گمراه کردن سرمایه‌گذاران در نظر گرفته شود. انرون و ورلدکام دو مورد از فاجعه‌بارترین موارد مدیریت سود فرصت‌طلبانه را نشان می‌دهند که به کلاهبرداری و بزرگ‌ترین ورشکستگی در تاریخ ایالات متحده منجر شدند (Jiraporn et al., 2008; Tort, 2013; Bajra & Cadez, 2018). اگرچه مدیریت سود معمولاً به دلیل تأثیر بد آن بر کیفیت سود به عنوان یک مفهوم منفی در نظر گرفته می‌شود، برخی از پژوهشگران به جنبه‌های مثبت مدیریت سود نیز اشاره کرده‌اند (Bajra & Cadez, 2018; Jiraporn et al., 2008; Tort, 2013). مدیریت سود را می‌توان: (۱) از طریق انحراف از رویه‌های تجاری عادی برای دست‌کاری هدفمند سود انجام داد، که مدیریت سود واقعی نامیده می‌شود (Roychowdhury, 2006) و جریان نقدی حاصل از فعالیت‌های عملیاتی را تحت تأثیر قرار می‌دهد؛ یا (۲) با دست‌کاری سود گزارش‌شده از طریق اقلام تعهدی برای دستیابی به رقم سود مناسب انجام داد، که مدیریت سود مبتنی بر اقلام تعهدی نامیده می‌شود. به طور کلی، مدیریت سود وضعیت مالی واقعی شرکت و حقایقی را که ذی‌نفعان حق دانستن آن را دارند، پنهان می‌کند.

از آنجا که سود یک شرکت توسط استفاده‌کنندگان مختلف از صورت‌های مالی (مانند سهام‌داران، اعتباردهندگان و تحلیلگران مالی) برای سنجش عملکرد آن استفاده می‌شود، تشخیص مدیریت سود می‌تواند برای آن‌ها جالب و حیاتی باشد. در واقع، ایجاد ابزارهایی برای پیش‌بینی و جلوگیری از وقوع چنین دست‌کاری‌هایی سودمند است (Jiraporn et al., 2008; Tort, 2013; Bajra & Cadez, 2018). در این زمینه، این پژوهش تلاش می‌کند تا ابزارهای پیش‌بینی را ارائه کند که به شناسایی فعالیت‌های مدیریت سود کمک می‌کنند. برای این منظور، با استفاده از شش روش یادگیری ماشین، پیش‌بینی مدیریت سود بررسی شده است.

توانایی پیش‌بینی مدیریت سود برای ناظران شرکت‌ها مانند حسابرسان مهم است؛ زیرا به آن‌ها اجازه می‌دهد تا سطوح ریسک و روش‌های آزمون حسابرسی خود را برای فعالیت‌های احتمالی مدیریت سود تعدیل کنند (Hammami & Hendijani, 2022). مقاله‌ای که توسط حسابداران حرفه‌ای خبره کانادا و انجمن حسابداران رسمی آمریکا منتشر شد به توانایی الگوریتم‌های یادگیری ماشین در کشف تقلب و در عین حال، افزایش کارایی و بهره‌وری آن‌ها در کشف تقلب اشاره می‌کند (CPA Canada and AICPA, 2019). علاوه بر این، پژوهش میلنر و برگ (۲۰۱۷) به مزایای داده‌کاوی و یادگیری ماشین اشاره می‌کند که می‌تواند برای مثال توسط حسابرسان مالیاتی برای شناسایی خطاها، برآورد ریسک حسابرسی و حتی پیشنهاد راهبردهای مالیاتی کارآمد استفاده شود. همچنین، آن‌ها بیان می‌کنند قانون‌گذاران و مقامات مالیاتی می‌توانند با استفاده از

یادگیری ماشین برای یافتن اختلاف در بازده گزارش شده منتفع شوند (Milner & Berg, 2017). داده‌کاوی، به عنوان یک جنبه کلیدی تجزیه و تحلیل داده، رویه‌ای است که داده‌های خام، بزرگ و نامفهوم را به دانش منظم، قابل درک و پیش‌بینی تبدیل می‌کند (Gupta & Modise, 2012). استفاده از یک رویکرد داده‌کاوی با استفاده از روش‌های یادگیری ماشین برای شناسایی مدیریت سود نسبت به روش‌های آماری مرسوم (مانند تجزیه و تحلیل تمایزی چندگانه، لاجیت و پروبیت، رگرسیون و روش‌های آماری تک‌متغیره) مزیت دارد؛ زیرا داده‌کاوی برخلاف رویکردهای مرسوم، به مفروضات محدودکننده‌ای مانند وجود خطی بودن و نرمال بودن متغیرها و چولگی داده‌ها که معمولاً در داده‌های مالی دنیای واقعی نقض می‌شوند، محدود نمی‌شود (Höglund, 2012).

به طور کلی، پژوهش حاضر از چندین جنبه می‌تواند حائز اهمیت باشد: اول، در بیشتر پژوهش‌های انجام‌شده خارجی و داخلی در حوزه پیش‌بینی مدیریت سود (از جمله Hammami & Hendijani Zadeh, 2022؛ Chen & Shen, 2020؛ قادری و همکاران، ۱۳۹۷؛ صالحی و فرخی پله‌رود، ۱۳۹۷؛ سعید مقدم و همکاران، ۱۳۹۶؛ Ezazi et al., 2013؛ Najari et al., 2014؛ گرد و همکاران، ۱۳۹۴) از متغیرهای نسبت مالی و برخی از ویژگی هیئت‌مدیره و حسابرسان استفاده شده است و سایر ویژگی‌ها از جمله ساختار مالکیت، ویژگی‌های مدیرعامل و سایر متغیرهای تأثیرگذار مورد توجه قرار نگرفته‌اند؛ از این رو، پژوهش حاضر به دنبال بررسی جامع عوامل تأثیرگذار بر پیش‌بینی مدیریت سود است. بر این اساس، متغیرهایی که در پژوهش حاضر استفاده شده‌اند، هم از نظر نوع و هم از نظر نحوه انتخاب متغیرهای پیش‌بین بهینه، با متغیرهای پژوهش‌های قبلی متفاوت هستند. دوم، با توجه به اهمیت موضوع مدیریت سود، پژوهشگران زیادی مشتاق بودند تا عوامل تعیین‌کننده مدیریت سود را به منظور شناسایی عواملی که به طرزی جالب توجه بر این شیوه‌ها تأثیر می‌گذارند، بررسی کنند. با این حال، تسای و چیو (2009) استدلال می‌کنند مطالعات قبلی در این زمینه بر همبستگی بین ویژگی خاص شرکت، حاکمیت شرکتی، مالکیت، سایر عوامل و مدیریت سود تأکید بیشتری دارند (Tsai & Chiou, 2009) و توانایی این عوامل به طور مستقیم برای پیش‌بینی مدیریت سود بررسی نشده است؛ بنابراین، آن‌ها تأکید کردند ساخت مدل پیش‌بینی مدیریت سود با استفاده از عوامل مختلف برای اطمینان از دریافت گزارش عملکرد مالی دقیق و قابل اعتماد سهام‌داران شرکت مهم است؛ از این رو، پژوهش حاضر تلاش می‌کند تا ابزارهای پیش‌بینی را ارائه کند که به شناسایی فعالیت‌های مدیریت سود کمک می‌کنند. برای رسیدن به این هدف، از رویکرد داده‌کاوی با استفاده از روش‌های یادگیری ماشین استفاده شده است. سوم، پژوهش حاضر اولین پژوهش در ایران (با بهترین دانش ما) در حوزه کاربرد روش‌های مختلف یادگیری ماشین است که هر دو نوع مدیریت سود را در یک پژوهش در نظر می‌گیرد؛ در نتیجه، دقت پیش‌بینی را در بین دو نوع مدیریت سود مقایسه می‌کند. علاوه بر این، برخی از روش‌های یادگیری ماشین از جمله یادگیری عمیق و  $k$ -نزدیک‌ترین همسایه برای اولین بار در پژوهش‌های پیش‌بینی مدیریت سود استفاده شده‌اند. همچنین، با استفاده از دو نوع روش انتخاب ویژگی، از یک طرف، امکان مقایسه روش‌های انتخاب ویژگی با هم فراهم می‌شود و از طرف دیگر، می‌توان متغیرهایی (ویژگی‌هایی) را شناسایی کرد که اطلاعات بیشتری دارند و قدرت پیش‌بینی بیشتری برای پیش‌بینی مدیریت سود دارند.

## مبانی نظری و پیشینه پژوهش

### مدیریت سود

مدیران به طور مداوم برای دستیابی به انتظارات سود سهام‌داران و تحلیلگران مالی تحت فشار هستند؛ زیرا عدم تحقق این اهداف پیامدهایی نامطلوب دارد (Braswell & Daniels, 2017). در این زمینه، مدیریت سود روشی رایج است که توسط مدیران برای برآورده کردن انتظارات عملیاتی و عملکرد تجاری ذی‌نفعان استفاده می‌شود (Braswell & Daniels, 2017). هیلی و والن (1919) به مدیریت سود به عنوان فرایندی اشاره می‌کنند که توسط مدیران برای تأثیرگذاری بر محتوای صورت‌های مالی به دلیل ملاحظات و انگیزه‌های مالی استفاده می‌شود (Healy & Wahlen, 1999). باجرا و کادز (2018) بیان کردند مدیریت سود به اقدامات مدیریتی برای افزایش (کاهش) درآمدها یا سودها برای دسته‌های مختلف سهام از طریق تاکتیک‌های حسابداری تهاجمی اشاره دارد (Bajra & Cadez, 2018). شن و چیه (۲۰۰۷) اظهار داشتند مدیریت سود تغییر اطلاعات اقتصادی و عملکرد گزارش‌شده شرکت‌ها برای گمراه کردن ذی‌نفعان به منظور کاهش مداخله خارجی یا محافظت از مزایای کنترل مالکیت داخلی است (Shen & Chih, 2007). به طور خلاصه، مدیریت سود را می‌توان به عنوان اختیاری برای مدیران به منظور ارائه گزارش‌های مالی مبتنی بر اصول عمومی پذیرفته‌شده حسابداری که می‌تواند بر مربوط بودن و قابلیت اطمینان اطلاعات حسابداری ارائه‌شده تأثیر بگذارد، توصیف کرد. اگر ذی‌نفعان استفاده فرصت‌طلبانه از اختیارات یادشده را کشف کنند، این امر ممکن است به دعاوی حقوقی (Gunny, 2010) یا از دست دادن شهرت (Rodriguez-Ariza et al., 2016) منجر شود.

در ادبیات اخیر برای مثال (Roychowdhury, 2006; McVay, 2006; Cohen & Zarowin, 2010)، مدیریت سود به سه نوع مدیریت سود اقلام تعهدی، مدیریت سود واقعی و تغییر طبقه‌بندی، دسته‌بندی شده است. رویچودوری (2006) تأکید می‌کند مدیریت ممکن است از دو روش برای مدیریت سود گزارش‌شده استفاده کند (Roychowdhury, 2006). اول، سود شرکت را می‌توان با استفاده از روش حسابداری اقلام تعهدی، که به عنوان مدیریت سود تعهدی شناخته می‌شود، مدیریت کرد. این دست‌کاری اعداد حسابداری هیچ تأثیر مستقیمی بر جریان نقدی شرکت ندارد (نوبخت و برادران حسن‌زاده، ۱۳۹۶). دوم، مدیریت ممکن است اعداد حسابداری را با استفاده از فعالیت‌های تجاری واقعی، که به عنوان مدیریت سود واقعی شناخته می‌شود، مدیریت کند. مدیریت سود واقعی بر جریان‌های نقدی شرکت‌ها تأثیر می‌گذارد؛ زیرا مدیران اعداد سود را با استفاده از اقدامات اقتصادی واقعی مدیریت می‌کنند (Roychowdhury, 2006). نتایج پژوهش برخی از پژوهشگران نشان می‌دهد مدیران بیشتر تمایل به مدیریت سود واقعی نسبت به مدیریت سود اقلام تعهدی دارند؛ زیرا آن‌ها مدیریت سود واقعی را به عنوان یک گزینه کم‌خطر و غیرقابل حدس توسط حسابرسان می‌بینند (Roychowdhury, 2006; Zang, 2012).

پژوهش‌های پیشین عوامل تعیین‌کننده مختلف تأثیرگذار بر مدیریت سود را بررسی کرده‌اند. برای مثال، سطوح پایین‌تر مدیریت سود با سطوح پایین‌تر عدم تقارن اطلاعات (Schipper, 1989)، کیفیت حسابرسی بالاتر (Cohen et al., 2008)؛ شاه مرادی و طباطبایی (نسب، ۱۴۰۰)، عملکرد مسئولیت اجتماعی شرکتی بالاتر (Pyo & Lee, 2013)، قدرت سیستم قانونی (Francis et al., 2016) و رتبه‌بندی ریسک اعتباری بهتر (Ge & Kim, 2014) مرتبط است. فان و همکاران (2019) این موضوع را بررسی کردند که چگونه حضور مدیران زن در هیئت‌مدیره بر مدیریت سود بانک‌ها تأثیر می‌گذارد. با استفاده از احتمال انتصاب

مدیران زن توسط هیئت‌مدیره (بر اساس شاخص Blau تنوع جنسیتی)، آن‌ها یک رابطه U شکل معکوس بین حضور مدیران زن در هیئت‌مدیره و مدیریت سود بانک‌ها پیدا کردند (Fan et al., 2019). نتایج پژوهش دارابی و اژدری (۱۳۹۷) بیانگر رابطه مثبت و معنادار بین حضور زنان در بین اعضای هیئت‌مدیره و مدیریت سود واقعی است. لاسود و همکاران (2017) دریافتند بانک‌هایی که مالکیت متمرکزتری دارند از مفاد ضرر وام اختیاری برای مدیریت سود خود استفاده می‌کنند (Lassoued et al., 2017). پژوهشگران همچنین دریافتند مالکان دولتی و نهادی مدیریت سود را تشویق می‌کنند (دارابی و اژدری، ۱۳۹۷)؛ در حالی که مالکان خانوادگی این عمل را کاهش می‌دهند (Lassoued et al., 2017). دئو و همکاران (2017) در نمونه‌ای از شرکت‌های چینی به بررسی نقش نظارتی مدیران بیرونی در بازدارندگی مدیریت سود اقلام تعهدی پرداختند. یافته‌های آن‌ها نشان می‌دهد که مدیریت سود اقلام تعهدی به طور معنادار و منفی با حضور و نسبت مدیران بیرونی در هیئت‌مدیره شرکت‌ها مرتبط است. آن‌ها همچنین دریافتند که مدیریت سود اقلام تعهدی در شرکت‌های دولتی نسبت به سایر شرکت‌ها کمتر مشخص است (Du et al., 2017). عارف‌منش و عموزادی ریزی (۱۳۹۹) در پژوهش خود دریافتند اندازه و درصد اعضای غیرموظف هیئت‌مدیره به ترتیب تأثیر منفی و مثبت معناداری بر مدیریت سود تعهدی دارند و درصد مالکیت نهادی و مالکیت دولتی تأثیر مثبت و معنادار بر مدیریت سود واقعی دارد؛ اما بر مدیریت سود تعهدی تأثیر معناداری ندارد. همچنین، در شرکت‌های دولتی، اندازه و استقلال هیئت‌مدیره و تمرکز مالکیت رابطه‌ای مثبت و معنادار با مدیریت سود تعهدی دارند و درصد سرمایه‌گذاران نهادی رابطه‌ای منفی و معنادار با مدیریت سود تعهدی دارد.

لوپز و وگا (2019) دریافتند حسابرسی‌های انجام‌شده توسط شرکت‌هایی با مدت زمان تخصص در صنعت طولانی‌تر با سطوح پایین‌تر مدیریت سود اقلام تعهدی و سطوح بالاتر مدیریت سود اقلام واقعی مرتبط است (Lopez & Vega, 2019). قلی‌پور خانقاه و همکاران (۱۴۰۰) در پژوهشی نتیجه گرفتند اندازه و استقلال حسابرس رابطه‌ای مثبت و معنادار با مدیریت سود دارد؛ اما بین دوره تصدی و تخصص حسابرس با مدیریت سود رابطه معناداری مشاهده نشد. نتایج پژوهش نمازی و همکاران (۱۳۹۰) و داغانی و همکاران (۱۳۹۸) بیانگر رابطه مثبت بین اندازه و دوره تصدی حسابرس با مدیریت سود است. علی و ژانگ (2015) تغییرات در انگیزه‌های مدیران عامل برای مدیریت سود گزارش‌شده شرکت‌هایشان را در طول دوره تصدی خود بررسی کردند. آن‌ها دریافتند بیش‌نمایی سود در سال‌های اولیه بیشتر از سال‌های آخر خدمت مدیران عامل است و این رابطه برای شرکت‌هایی با نظارت بیرونی و داخلی بیشتر ضعیف‌تر است (Ali & Zhang, 2015). احمدپور و همکاران (۱۳۹۴) نشان دادند مدیران عامل با دوره تصدی بلندمدت در مقایسه با مدیران عامل با دوره تصدی کوتاه‌مدت گزارشگری جسورانه‌تری دارند. به علاوه، مدیران عامل با دوره تصدی بلندمدت در نیمه اول دوره تصدی خود در مقایسه با نیمه دوم، گزارشگری محافظه‌کارانه‌تری دارند. بوعزیز و همکاران (2020) نشان دادند بین دوگانگی و ملیت مدیرعامل با مدیریت سود اقلام تعهدی رابطه‌ای مثبت و معنادار وجود دارد. با این حال، هیچ رابطه معناداری بین اعضای هیئت‌مدیره و تغییر مدیرعامل با مدیریت سود اقلام تعهدی یافت نشد (Bouaziz et al., 2020). نتایج پژوهش استا (۱۳۹۰) بیانگر این است که رابطه منفی معناداری بین مالکیت نهادی و مالکیت مدیریتی با مدیریت سود وجود دارد؛ اما بین مالکیت شرکتی و مدیریت سود رابطه مثبت و معناداری وجود دارد.

با وجود اهمیت تشخیص مدیریت سود، مطالعات کمی چگونگی تشخیص آن را بررسی کرده‌اند. حتی در میان این معدود، بیشتر آن‌ها از رویکردهای آماری مرسوم مانند تجزیه و تحلیل تمایزی چندگانه، تحلیل‌های لجیت و پروبیت، رگرسیون و روش‌های آماری تک‌متغیره استفاده کرده‌اند (Haga et al., 2015; Höglund, 2012; Tsai, 2009). از آنجا که رویکردهای آماری مرسوم مقید به مفروضاتی ساده و محدودکننده هستند که در دنیای واقعی داده‌های مالی نقض می‌شوند (مانند نرمال بودن متغیرهای مستقل و خطی بودن و چولگی نقاط داده)، شک و تردیدهایی درباره اعتبار و اثربخشی تشخیص مدیریت سود از طریق این رویکردهای آماری وجود دارد. راه‌حل برای این معضل را می‌توان در تکنیک‌های داده‌کاوی یافت که به مفروضات یادشده در بالا محدود نمی‌شوند و در نتیجه، آن‌ها را برای رسیدگی به مشکلات تشخیص مدیریت سود مناسب‌تر می‌کند. علاوه بر این، مطالعات قبلی انعطاف‌پذیری و عملکرد مناسب رویکردهای داده‌کاوی را در مواجهه با مجموعه داده‌های نامتعادل با مقادیر زیاد و ناقص نشان می‌دهد؛ نتیجه‌ای که برای رویکردهای آماری مرسوم قابل اجرا نیست (Chye Koh & Kee Low, 2004; Liou, 2008)؛ بنابراین، پژوهش حاضر بر پیش‌بینی مدیریت سود اقلام تعهدی و مدیریت سود واقعی با استفاده از تکنیک‌های داده‌کاوی متمرکز است.

### داده‌کاوی

داده‌کاوی فرایندی است که داده‌های خام را که ممکن است بسیار بزرگ، نامفهوم یا حتی ناقص باشند، به دانش تبدیل می‌کند؛ از این رو، خروجی داده‌کاوی اطلاعات منظم، قابل درک و بالقوه مفیدی است (Gupta & Modise, 2012). استفاده از داده‌کاوی برای تشخیص مدیریت سود می‌تواند بسیار سودمند باشد؛ زیرا امکان ساخت مدل‌های غیرخطی را می‌دهد که با مفروضات محدودکننده‌ای که در روش‌های آماری مرسوم و تحلیل‌های رگرسیونی مانند نرمال بودن و استقلال متغیرهای مستقل وجود دارند، محدود نمی‌شوند. رویکرد داده‌کاوی با موفقیت برای بسیاری از مسائل مالی در پژوهش‌های قبلی به کار گرفته شده است (e.g., Amani & Fadlalla, 2017; Chen et al., 2017; Fallahpour et al., 2017; Oreski & Oreski, 2014; Yu et al., 2014). در پژوهش حاضر، برای اجرای رویکرد داده‌کاوی از روش‌های یادگیری ماشین از جمله ماشین بردار پشتیبان، درخت تصمیم، k-نزدیک‌ترین همسایه و یادگیری عمیق برای پیش‌بینی مدیریت سود استفاده شده است. عملکرد عالی روش‌های یادشده به عنوان یک الگوریتم طبقه‌بندی در پژوهش‌هایی مانند مدل‌سازی ارزیابی ریسک اعتباری (Danenas & Garsva, 2012؛ میرغفوری و امین آشوری، ۱۳۹۴؛ محمدیان حاجی کرد و همکاران، ۱۳۹۵)، پیش‌بینی سری زمانی مالی (Tay & Cao, 2001)، پیش‌بینی حرکت بازار سهام و بازده سهام (Huang et al., 2005؛ ستایش و کاظم‌نژاد، ۱۳۹۸)، پیش‌بینی ورشکستگی شرکت (Veganzones & Séverin, 2018؛ حسینی و رشیدی، ۱۳۹۲)، پیش‌بینی درماندگی مالی (Ninh et al., 2018؛ Matin et al., 2019؛ نمازی و ابراهیمی، ۱۴۰۰) و پیش‌بینی مدیریت سود (Hammami & Hendijani Zadeh, 2022؛ قادری و همکاران، ۱۳۹۷؛ صالحی و فرخی پیلهرود، ۱۳۹۷؛ کاردان و همکاران، ۱۳۹۶) نشان داده شده است. در ادامه، روش‌های اشاره‌شده به طور مختصر شرح داده می‌شوند.

### ماشین‌های بردار پشتیبان

ماشین‌های بردار پشتیبان روشی است که برای دسته‌بندی داده‌های خطی و غیرخطی استفاده می‌شود. به طور خلاصه، بردارهای پشتیبان به زبان ساده، مجموعه‌ای از نقاط در فضای  $n$  بعدی داده‌ها هستند که مرز دسته‌ها را مشخص می‌کنند و مرزبندی و دسته‌بندی داده‌ها بر اساس آن‌ها انجام می‌شود و با جابه‌جایی یکی از آن‌ها، خروجی دسته‌بندی ممکن است تغییر کند. الگوریتم ماشین بردار پشتیبان این ابرصفحه را با کمک بردارهای پشتیبان و حاشیه‌ها پیدا می‌کند (اسماعیلی، ۱۳۹۳). به دلیل اینکه الگوریتم‌های ماشین‌های پشتیبان قابلیت مدل‌سازی کران‌های تصمیم‌گیری غیرخطی پیچیده را دارند، حتی سریع‌ترین آن‌ها نیز ممکن است دارای سرعت کمی در زمان آموزش باشند؛ اما صحت آن‌ها بسیار بالاست. افزون بر این، نسبت به دیگر روش‌ها کمتر دچار مشکل بیش‌پردازش داده‌ها می‌شوند (اسماعیلی، ۱۳۹۳). مطالعات قبلی برتری هسته‌های مختلف از جمله چندجمله‌ای<sup>۱</sup>، تابع پایه شعاع<sup>۲</sup> و تابع هسته سیگموئید<sup>۳</sup> را نشان داده‌اند (Dash & Dash, 2016; Nair et al., 2010) و در میان آن‌ها، هسته تابع پایه شعاع محبوب‌ترین تابع است (Veganzones & Séverin, 2018).

### درخت تصمیم

درخت تصمیم یک مدل طبقه‌بندی است که با استفاده از رویکرد یادگیری استقرایی ایجاد شده است. تفاوت آن با مدل‌های سنتی، از جمله تحلیل ممیزی چندگانه و لاجیت، در این واقعیت است که از محدودیت فرضیه‌های آماری عاری است و قواعد قضاوت درختی شکل می‌تواند بر اساس نتایج پس از بررسی متغیرهای گسسته و پیوسته شکل بگیرد (Viaene et al., 2005; Jan, 2018). علاوه بر این، همان‌طور که توسط چن و همکاران (2015) بیان شده است، نقاط قوت اصلی درخت تصمیم، توانایی آن در پردازش داده‌های ناقص و کشف روابط بالقوه بین متغیرهای ورودی و خروجی عظیم و پیچیده است (Chen et al., 2015). در حال حاضر، یکی از رایج‌ترین الگوریتم‌های درخت تصمیم‌گیری روش C5.0 است (Viaene et al., 2005). C5.0 یک ساختار درختی مانند نمودار جریان است که توسط یک الگوریتم بازگشتی تقسیم و غلبه ساخته شده است که پارتیشنی از داده‌ها را تولید می‌کند. برای روش تقسیم گره ویژگی عددی پیوسته، C5.0 ابتدا اشیاء را جمع‌آوری می‌کند و آن‌ها را بر اساس ویژگی مرتب می‌کند و سپس مقدار مشخصه را در نقطه میانی دو شیء مجاور پیدا می‌کند که به آن نقطه برش می‌گویند. آن‌هایی که می‌توانند مقدار بهینه را پس از محاسبه تابع ارزیابی به دست آورند، می‌توانند از نقطه میانی صفت پیروی کنند تا تقسیم باینری کنند. مقادیر مشخصه معیوب و نامطمئن معمولاً با متداول‌ترین مقادیر مشخصه جایگزین می‌شوند یا با روش احتمال برآورد خوش‌بینانه حل می‌شوند (Chen & Howard, 2016).

### k- نزدیک‌ترین همسایه

الگوریتم k- نزدیک‌ترین همسایه (KNN) روشی برای طبقه‌بندی اشیاء بر اساس نزدیک‌ترین نمونه‌های آموزشی در فضای ویژگی است. k- نزدیک‌ترین همسایه نوعی یادگیری مبتنی بر نمونه یا یادگیری تنبل است که در آن تابع فقط به صورت محلی

<sup>1</sup> Polynomial

<sup>2</sup> radius basic function

<sup>3</sup> sigmoid kernel function

تقریبی می‌شود و تمام محاسبات تا طبقه‌بندی به تعویق می‌افتند (Cover, 1986; Devroye, 1981). هنگامی که دانش قبلی یا اصلاً دانشی درباره توزیع داده‌ها وجود نداشته باشد،  $k$ -نزدیک‌ترین همسایه اساسی‌ترین و ساده‌ترین تکنیک طبقه‌بندی است (Devroye, 1981; Devroye & Wagner, 1982). این قانون به سادگی کل مجموعه آموزشی را در حین یادگیری حفظ می‌کند و به هر پرسش  $y$  یک کلاس اختصاص می‌دهد که با اکثریت برچسب  $k$  نزدیک‌ترین همسایه‌های آن در مجموعه آموزشی نمایش داده می‌شود. قانون نزدیک‌ترین همسایه (NN) ساده‌ترین شکل KNN است که در آن  $K = 1$  است.

$k$ -نزدیک‌ترین همسایه به عنوان یک تکنیک داده‌کاوی دارای کاربردهایی گسترده در طبقه‌بندی و همچنین رگرسیون است. برخی از کاربردهای  $k$ -نزدیک‌ترین همسایه در زمینه‌های استخراج متن، کشاورزی، مالی و پزشکی است (Imandoust & Bolandraftar, 2013). پیش‌بینی بازار سهام یکی از اصلی‌ترین وظایف مالی  $k$ -نزدیک‌ترین همسایه است. پیش‌بینی بازار سهام شامل کشف روندهای بازار، برنامه‌ریزی راهبردهای سرمایه‌گذاری، شناسایی بهترین زمان برای خرید سهام و تعیین سهام برای خرید است (Imandoust & Bolandraftar, 2013).

### یادگیری عمیق

در سال ۲۰۰۶، هیتون و همکاران یادگیری عمیق را ارائه کردند که مبتنی بر مفهوم شبکه عصبی مصنوعی بود (Hinton et al., 2006). پس از آن، یادگیری عمیق به موضوعی مهم تبدیل شد که به یک رنسانس در پژوهش‌های شبکه‌های عصبی منجر شد؛ از این رو، اصطلاح شبکه‌های عصبی نسل جدید نام گرفت. دلیل این امر آن است که شبکه‌های عصبی عمیق وقتی به درستی آموزش داده می‌شوند، در طیفی وسیع از مشکلات طبقه‌بندی و رگرسیون بسیار خوب عمل می‌کنند (Karhunen et al., 2015). شبکه‌های عصبی عمیق زیرشاخه‌ای از یادگیری ماشین را تشکیل می‌دهند. یادگیری عمیق از توپولوژی‌های شبکه عصبی به عنوان پایه خود استفاده می‌کند؛ در نتیجه، آن‌ها به عنوان شبکه‌های عصبی شناخته می‌شوند (Schmidhuber, 2015). الگوریتم‌های یادگیری عمیق از افزایش تولید داده، قدرت پردازش بهتر در حال حاضر و رشد هوش مصنوعی به عنوان یک سرویس سود می‌برند. حتی با جمع‌آوری داده‌های بسیار متنوع، بدون ساختار و متصل، یادگیری عمیق ماشین‌ها را قادر می‌سازد تا مسائل پیچیده را حل کنند. الگوریتم‌های یادگیری عمیق هرچه بیشتر یاد بگیرند، عملکردی بهتر دارند (Ahmad et al., 2019; Janiesch et al., 2021). معروف‌ترین الگوریتم‌های یادگیری عمیق شامل شبکه عصبی بازگشتی (RNN)<sup>۴</sup>، شبکه عصبی پیچشی (CNN)<sup>۵</sup>، حافظه کوتاه‌مدت ماندگار (LSTM)<sup>۶</sup>، شبکه باور عمیق (DBN)<sup>۷</sup>، شبکه عصبی عمیق (DNN)<sup>۸</sup> و یادگیری تقویتی عمیق (DRL)<sup>۹</sup> هستند. در پژوهش حاضر از روش شبکه عصبی عمیق استفاده شده است.

در ادامه، برخی از پیشینه‌های داخلی و خارجی بررسی شده‌اند:

<sup>۴</sup> Recurrent Neural Networks

<sup>۵</sup> Convolutional Neural Networks

<sup>۶</sup> Long Short Term Memory

<sup>۷</sup> Deep Belief Networks

<sup>۸</sup> Deep Neural Networks

<sup>۹</sup> Deep Reinforcement Learning

همامی و هنديجانی‌زاده (2022) پژوهشی تحت عنوان پیش‌بینی مدیریت سود از طریق طبقه‌بندی‌کننده‌های گروه یادگیری ماشین را انجام دادند. آن‌ها از سه تکنیک انتخاب ویژگی تکی (کسب اطلاعات، تجزیه و تحلیل مؤلفه اصلی و ریلیف) و سه تکنیک انتخاب ویژگی گروهی (بهینه‌سازی ازدحام ذرات، الگوریتم ژنتیک و کلونی زنبورهای مصنوعی) استفاده کردند. این تکنیک‌های انتخاب ویژگی با ماشین بردار پشتیبان ترکیب شده‌اند تا شش طبقه‌بندی‌کننده گروهی تولید کنند. نتایج نشان داد روش ترکیبی کلونی زنبورهای مصنوعی و ماشین بردار پشتیبان بهترین عملکرد را برای پیش‌بینی مدیریت سود اقلام تعهدی و مدیریت سود واقعی دارد (Hammami & Hendijani Zadeh, 2022). جاسپرسن و همکاران (2021) در پژوهشی پیش‌بینی مدیریت سود از افشاهای کیفی را بررسی کردند. آن‌ها از مدل‌های یادگیری ماشینی نظارت‌شده برای پیش‌بینی مدیریت سود توسط بیمه‌گران اموال و تلفات از پرونده‌های مورد بحث و تحلیل مدیریت استفاده کردند. آن‌ها از یک الگوریتم جدید استفاده کردند که داده‌های متنی را مشروط به وضعیت مالی گزارش‌شده شرکت تفسیر می‌کرد. نتایج نشان داد افشای کیفی می‌تواند مدیریت سود را پیش‌بینی کند که بیانگر این است که مدیران قادر به حذف همه پیام‌های فرعی از آن‌ها نیستند (Jaspersen et al., 2021).

چن و شن (2020) پژوهشی با عنوان تأثیر مدل پیش‌بینی مدیریت سود برای توسعه بازار سرمایه در کشور تایوان را انجام دادند. نتایج نشان داد مدل تشخیص مدیریت سود ایجادشده با استفاده از شبکه الاستیک و درخت تصمیم C5.0 بهترین عملکرد طبقه‌بندی را ارائه می‌دهد؛ به طوری که دقت متوسط پیش‌بینی آن به ۹۷/۳۲ درصد می‌رسد (Chen & Shen, 2020). در پژوهشی، دبوک و زارنور (2017) تشخیص دست‌کاری سود صورت‌های مالی با استفاده از یک لایه یادگیری ماشین را بررسی کردند. نتایج نشان داد مدل‌های ریاضی از مدل حساب‌برسان بهتر عمل می‌کنند. آن‌ها نرخ طبقه‌بندی استفاده‌شده با استفاده از مدل بنیش را ۸۶/۸۴ درصد و روش‌های حساب‌برسان را ۶۰/۵۳ درصد نشان دادند. یافته‌ها حاکی از آن است که روش‌های حساب‌برسان برای تشخیص دست‌کاری سود در صورت‌های مالی دشوار است (Dbouk & Zaarour, 2017).

چن و همکاران (2015) در پژوهشی با عنوان تشخیص مدیریت سود صنعت بیوتکنولوژی با استفاده از شبکه‌های بایس، تجزیه و تحلیل مؤلفه‌های اصلی، شبه عصبی مصنوعی پس از انتشار و درخت تصمیم، نشان دادند ترکیب روش غربالگری شبکه بایس با درخت تصمیم‌گیری بهترین عملکرد را با میزان دقت ۹۸/۵۱ درصد در تشخیص مدیریت سود نشان می‌دهد (Chen et al., 2015). در پژوهشی، نجاری و همکاران (2014) اقدام به پیش‌بینی مدیریت سود با استفاده از الگوریتم ماشین بردار پشتیبان کردند که نتایج پژوهش بیانگر آن است که الگوریتم یادشده توانایی مناسبی برای پیش‌بینی مدیریت سود دارد (Najari et al., 2014).

آزادی و همکاران (۱۴۰۰) پژوهشی با عنوان به‌کارگیری پدیده تونلینگ برای افزایش توانایی پیش‌بینی مدیریت سود در مدل بنیش بر مبنای تکنیک شبکه‌های عصبی مصنوعی و الگوریتم بهینه‌سازی حرکت تجمعی ذرات را انجام دادند. نتایج حاصل از تحلیل مدل نشان داد کلیه نسبت‌های مالی بر پیش‌بینی مدیریت سود بنیش تأثیر معنادار داشته‌اند و بیشترین تأثیر در پیش‌بینی مدیریت سود بنیش را شاخص پدیده تونلینگ INE و کمترین تأثیر را شاخص اهرم مالی داشته است. نتایج حاصل از برآورد شبکه‌های عصبی طراحی‌شده نشان داد استفاده از الگوریتم بهینه‌سازی تجمعی ذرات برای پیش‌بینی مدیریت سود برای شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران از عملکردی قابل قبول برخوردار است.

پورعلی و کوچکی تاجانی (۱۳۹۹) دقت پیش‌بینی دست‌کاری سود شرکت‌ها با استفاده از الگوریتم رقابت استعماری و الگوریتم ژنتیک را مقایسه کردند. نتایج این پژوهش نشان داد الگوریتم رقابت استعماری با دقت ۹۳ درصد و خطای ۷ درصد و الگوریتم ژنتیک با دقت ۷۶ درصد و خطای ۲۴ درصد توان پیش‌بینی ضرایب متغیرهای مدل دست‌کاری سود را داشته‌اند. همچنین، نتایج نشان‌دهنده این بود که توان پیش‌بینی دقت ضرایب مدل دست‌کاری سود توسط الگوریتم‌های رقابت استعماری و ژنتیک بیشتر از دقت پیش‌بینی مدل اولیه بنیش (Beneish, 1999) و مدل تعدیل‌شده بنیش توسط کردستانی و تاتلی (۱۳۹۵) است.

قادری و همکاران (۱۳۹۷) در پژوهشی دقت شبکه‌های عصبی مصنوعی و الگوریتم بهینه‌سازی کلونی مورچگان در پیش‌بینی مدیریت سود را بررسی کردند. برای این منظور، از ۲۸ متغیر تأثیرگذار بر مدیریت سود در قالب چهار گروه (مالی، مدیریتی، شرکتی و حسابرسی) در طی سال‌های ۱۳۹۰ الی ۱۳۹۵ در ۱۲۴ شرکت پذیرفته‌شده بورس اوراق بهادار تهران استفاده شد. نتایج کلی حاصل از این پژوهش نشان داد روش شبکه عصبی مصنوعی و الگوریتم کلونی مورچگان در پیش‌بینی مدیریت سود نسبت به روش خطی دقیق‌تر و دارای سطح خطای کمتری است.

صالحی و فرخی پیلرود (۱۳۹۷) پیش‌بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم در چهار صنعت کشاورزی، دارویی، نساجی و فراورده‌های نفتی را بررسی کردند. نتایج حاصل از پژوهش آنان نشان داد مدیریت سود با متغیرهای اقلام تعهدی اختیاری دوره قبل، اقلام تعهدی غیراختیاری دوره قبل یا آستانه عملکرد و ریسک در چهار روش مدل‌های خطی، شبکه عصبی، درخت‌های C5.0 و Cart دارای بیشترین ارتباط است و دریافتند روش شبکه عصبی و درخت تصمیم‌گیری در پیش‌بینی مدیریت سود نسبت به روش خطی دقیق‌تر و دارای سطح خطای کمتری است. کاردان و همکاران (۱۳۹۶)، با بررسی دقت الگوریتم خطی - تکاملی ICDE و BBO و الگوریتم‌های غیرخطی SVR و CART در پیش‌بینی مدیریت سود، نشان دادند به طور کلی الگوریتم‌های غیرخطی از دقت بیشتری نسبت به الگوریتم‌های خطی برخوردار هستند و الگوریتم رگرسیون پشتیبان، مدیریت سود را بهتر از سایر الگوریتم‌ها پیش‌بینی می‌کند.

آرسته و نصیرزاده (۱۳۹۵) دقت الگوریتم‌های تخمین‌گر بردار پشتیبان، تخمین‌گر حداقل درجه و شبکه عصبی فازی در کشف مدیریت سود را مقایسه کردند. جامعه آماری این پژوهش شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران طی سال‌های ۱۳۸۵ الی ۱۳۹۱ بود. نتایج بررسی فرضیه‌ها حاکی از آن است که الگوریتم تخمین‌گر بردار پشتیبان و تخمین‌گر حداقل درجه و شبکه عصبی فازی به‌ترتیب بیشترین دقت را در پیش‌بینی مدیریت سود را دارند. گرد و همکاران (۱۳۹۴)، با استفاده از الگوریتم‌های مورچگان و غذایابی باکتری، اقدام به پیش‌بینی مدیریت سود کردند. برای اندازه‌گیری مدیریت سود در این پژوهش از اقلام تعهدی اختیاری استفاده شده است. نتایج نشان از دقت زیاد پیش‌بینی مدیریت سود توسط این الگوریتم‌ها داشت.

### فرضیه‌های پژوهش

با توجه به مبانی نظری و پیشینه تجربی، فرضیه‌ها به صورت زیر تدوین شده‌اند:

**فرضیه اول:** عملکرد پیش‌بینی مدیریت سود ارقام تعهدی شرکت‌ها بر اساس الگوی روش انتخاب ویژگی مبتنی بر ریلیف نسبت به روش انتخاب ویژگی مبتنی بر تحلیل مؤلفه‌های اصلی بهتر است.

**فرضیه دوم:** عملکرد پیش‌بینی مدیریت سود واقعی شرکت‌ها بر اساس الگوی روش انتخاب ویژگی مبتنی بر ریلیف نسبت به روش انتخاب ویژگی مبتنی بر تحلیل مؤلفه‌های اصلی بهتر است.

**فرضیه سوم:** صحت پیش‌بینی مدیریت سود ارقام تعهدی نسبت به مدیریت سود واقعی بالاتر است.

### روش پژوهش

این پژوهش برحسب هدف، کاربردی و بر مبنای روش، استنباطی و از نوع همبستگی است. پژوهش حاضر از نظر ماهیت جزء پژوهش‌های اثباتی و از نظر نوع داده‌ها کمی محسوب می‌شود. همچنین، از نظر جمع‌آوری اطلاعات و بررسی از نوع پس‌رویدادی (با استفاده از اطلاعات گذشته) است و بر اساس اطلاعات مندرج در صورت‌های مالی و گزارش هیئت‌مدیره انجام شد. برای تجزیه و تحلیل نهایی، به منظور واکاوی داده‌ها از نرم‌افزارهای RapidMiner نسخه ۱۰-۹ و SPSS نسخه ۲۶ استفاده شد. همچنین، برای آزمون فرضیه‌های پژوهش از معیارهای میانگین صحت پیش‌بینی و خطاهای نوع اول و دوم استفاده شد.

### جامعه آماری و نمونه پژوهش

جامعه آماری پژوهش حاضر را شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران تشکیل می‌دهند. برای انتخاب نمونه آماری پژوهش از روش غربالگری استفاده خواهد شد. نمونه نهایی پژوهش شامل شرکت‌هایی است که حائز شرایط زیر باشند:

- ۱- جزو صنایع واسطه‌گری، سرمایه‌گذاری، لیزینگ و بیمه نباشند.
- ۲- داده‌های مدنظر آن‌ها در دسترس باشد.
- ۳- طی دوره پژوهش، تغییر فعالیت یا سال مالی نداده و حذف نشده باشند.
- ۴- ارزش دفتری حقوق صاحبان سهام آن‌ها در دوره تحت بررسی مثبت باشد.
- ۵- به دلیل ویژگی متفاوت صنایع به لحاظ شرایط اقتصادی و حجم فعالیت، تکنیک‌های مختلف مدیریت سود توسط مدیران با توجه به شرایط هر یک از صنایع، به منظور تعمیم بهتر نتایج مدیریت سود برای هر سال و در سطح هر صنعت به طور جداگانه محاسبه شد. برای این منظور، به گفته ریچودری (2006) و کوتاری و همکاران (2005) تعداد شرکت‌های هر صنعت باید دست‌کم دارای ۱۵ عضو باشد (Kothari et al., 2005; Roychowdhury, 2006). همچنین، بازه زمانی پژوهش پیش‌رو یک دوره ۱۲ ساله از سال ۱۳۸۹ تا ۱۴۰۰ را تشکیل می‌دهد.

با توجه به شرایط بالا، در نهایت ۱۸۰ شرکت انتخاب شدند که داده‌ها و اطلاعات مالی آن‌ها در تجزیه و تحلیل‌های انجام‌شده استفاده شد.

### متغیر وابسته

در این پژوهش، از مدیریت سود واقعی و مدیریت سود ارقام تعهدی به عنوان متغیرهای وابسته استفاده و به صورت زیر محاسبه می‌شوند:

### مدیریت سود واقعی

به منظور سنجش مدیریت سود واقعی، مطابق پژوهش رویچودھاری (2006)، کوهن و همکاران (2008) و کوهن و زاروین (2010) از جریان‌های نقدی عملیاتی غیرعادی، هزینه‌های تولید غیرعادی و هزینه‌های اختیاری غیرعادی استفاده شده است که به صورت زیر اندازه‌گیری و محاسبه می‌شوند (Cohen et al., 2008; Cohen & Zarowin, 2010; Roychowdhury, 2006):

#### اندازه‌گیری جریان‌های نقد عملیاتی غیرعادی

از مدل (۱) برای برآورد جریان‌های نقدی عملیاتی غیرعادی استفاده شده است؛ به نحوی که باقی‌مانده مدل به عنوان معیار جریان‌های نقدی عملیاتی غیرعادی (AB\_CASH) در نظر گرفته شده است.

$$\frac{CFO_{i,t}}{TA_{i,t-1}} = \beta_0 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{Sales_{i,t}}{TA_{i,t-1}} + \beta_2 \frac{\Delta Sales_{i,t}}{TA_{i,t-1}} + \varepsilon_{i,t} \quad \text{مدل (1)}$$

$CFO_{i,t}$ : جریان‌های نقد عملیاتی شرکت  $i$  در پایان سال  $t$ ;  $TA_{i,t-1}$ : جمع کل دارایی‌های شرکت  $i$  در پایان سال  $t-1$ ؛  $Sales_{i,t}$ : فروش خالص شرکت  $i$  در طی سال  $t$ ;  $\Delta Sales_{i,t}$ : تغییرات در فروش خالص شرکت  $i$  در پایان سال  $t$  که برابر است با فروش در سال  $t$  منهای فروش در سال  $t-1$ ;  $\varepsilon_{i,t}$ : باقیمانده مدل.

#### اندازه‌گیری هزینه‌های تولید غیرعادی

هزینه‌های تولید غیرعادی از طریق مدل (۲) برآورد شده‌اند؛ به نحوی که باقیمانده مدل به عنوان معیار هزینه تولید غیرعادی (AB\_COST) در نظر گرفته شده است.

$$\frac{PROD_{i,t}}{TA_{i,t-1}} = \beta_0 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{Sales_{i,t}}{TA_{i,t-1}} + \beta_2 \frac{\Delta Sales_{i,t}}{TA_{i,t-1}} + \beta_3 \frac{\Delta Sales_{i,t-1}}{TA_{i,t-1}} + \delta_{i,t} \quad \text{مدل (2)}$$

$PROD_{i,t}$ : هزینه تولید شرکت  $i$  در پایان سال  $t$  که برابر است با بهای تمام‌شده کالای فروش‌رفته + تغییرات در موجودی کالا؛  $\delta_{i,t}$ : باقیمانده مدل.

#### اندازه‌گیری هزینه‌های اختیاری غیرعادی

هزینه‌های اختیاری غیرعادی از طریق مدل (۳) برآورد شده‌اند؛ به نحوی که باقیمانده مدل به عنوان معیار هزینه اختیاری غیرعادی (AB\_EXP) در نظر گرفته شده است.

$$\frac{DISEXP_{i,t}}{TA_{i,t-1}} = \beta_0 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{Sales_{i,t}}{TA_{i,t-1}} + \lambda_{i,t} \quad \text{مدل (3)}$$

$DISEXP_{i,t}$ : هزینه اختیاری شرکت  $i$  در پایان سال  $t$  که برابر است با هزینه‌های فروش، اداری و عمومی؛  $\lambda_{i,t}$ : باقیمانده مدل. حاصل جمع باقیمانده مدل‌های (۱) تا (۳) نشان‌دهنده مدیریت سود واقعی است.

## مدیریت سود اقلام تعهدی

در پژوهش حاضر، اقلام تعهدی اختیاری (DA) به عنوان نماینده متغیر مدیریت سود اقلام تعهدی در نظر گرفته شده است که با استفاده از مدل (۴) از طریق باقیمانده مدل جونز تعدیل شده توسط کوتاری و همکاران (2005) محاسبه می‌شود. این مدل عملکرد شرکت یعنی بازده دارایی‌ها را شامل می‌شود و به دلیل تفاوت مدیریت سود با عملکرد، قابلیت اطمینان را افزایش می‌دهد (Kothari et al., 2005). علاوه بر این، این مدل می‌تواند مشکل ناهمسانی و مسائل نادرست تعیین شده را که در سایر مدل‌های اقلام تعهدی کل وجود دارد، کاهش دهد (Sun et al., 2010). اندازه‌گیری و محاسبه آن به صورت زیر است:

$$TACC_{i,t} = \beta_0 \frac{1}{TA_{i,t-1}} + \beta_1 \frac{PPE_{i,t}}{TA_{i,t-1}} + \beta_2 \frac{\Delta Sales_{i,t} - \Delta RECI_{i,t}}{TA_{i,t-1}} + \beta_3 ROA_{i,t} + \gamma_{i,t} \quad \text{مدل (4)}$$

$TACC_{i,t}$ : مجموع اقلام تعهدی است و تفاوت بین جریان نقدی ناشی از فعالیت‌های عملیاتی و سود خالص پس از مالیات است؛  $\Delta RECI_{i,t}$ : تغییرات در حساب‌های دریافتی شرکت  $i$  در پایان سال  $t$  که برابر است با حساب‌های دریافتی در سال  $t$  منهای حساب‌های دریافتی در سال  $t-1$ ؛  $PPE_{i,t}$ : اموال، ماشین‌آلات و تجهیزات شرکت  $i$  در پایان سال  $t$ ؛  $ROA_{i,t}$ : بازده دارایی‌های شرکت  $i$  در پایان سال  $t$  است که برابر است با نسبت سود خالص به مجموع دارایی‌های شرکت  $i$  در پایان سال  $t$ ؛  $\gamma_{i,t}$ : باقیمانده مدل که نشان‌دهنده اقلام تعهدی غیرعادی (اختیاری) است.

## متغیرهای پیش‌بین (مستقل)

در مرحله اول، با بررسی ادبیات و پیشینه، حدود ۱۶۰ متغیر پیش‌بین اولیه شناسایی شدند. از میان متغیرهای شناسایی شده، ۶۴ متغیری که بیشتر در ادبیات مربوط به تبیین یا پیش‌بینی مدیریت سود استفاده شده بودند و داده‌های آن‌ها در دسترس بود، انتخاب شدند. در این راستا، سعی شد ویژگی‌هایی از جمله نسبت‌های مالی، شرکت، حسابر، حاکمیت شرکتی، ساختار مالکیت، مدیریتی و کمیته حسابرسی در نظر گرفته شوند. نگاره (۱) نشان‌دهنده متغیرهای پیش‌بین استفاده شده است.

## نگاره ۱. متغیرهای پیش‌بین (مستقل)

Table 1. Predictor variables (independent)

| ردیف | ویژگی‌ها      | نام متغیر                    | تعریف عملیاتی متغیر                                                                                |
|------|---------------|------------------------------|----------------------------------------------------------------------------------------------------|
| ۱    |               | Growth rate total assets     | نرخ رشد مجموع دارایی‌ها برابر با تغییر در کل دارایی‌ها تقسیم بر مجموع دارایی‌ها در پایان سال $t-1$ |
| ۲    |               | ROA                          | بازده دارایی‌ها برابر با سود خالص تقسیم بر کل دارایی‌ها                                            |
| ۳    |               | ROA growth rate              | نرخ رشد بازده دارایی‌ها برابر با تغییر در بازده دارایی‌ها تقسیم بر بازده دارایی‌ها در سال $t-1$    |
| ۴    | نسبت‌های مالی | ROE                          | بازده سرمایه برابر با سود خالص تقسیم بر ارزش دفتری حقوق صاحبان سرمایه                              |
| ۵    |               | ROE growth rate              | نرخ رشد بازده سرمایه برابر با تغییر در بازده سرمایه تقسیم بر بازده سهام در سال $t-1$               |
| ۶    |               | Total assets turnover        | گردش مجموع دارایی‌ها برابر با فروش خالص تقسیم بر میانگین مجموع دارایی‌ها                           |
| ۷    |               | Inventory turnover           | گردش موجودی‌ها برابر با بهای تمام‌شده کالا تقسیم بر میانگین موجودی کالا                            |
| ۸    |               | Accounts receivable turnover | گردش حساب‌های دریافتی برابر با فروش خالص تقسیم بر میانگین حساب‌های دریافتی                         |

|                                                                                                                                                                                                                                                                                         |                                     |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|----|
| فروش خالص تقسیم بر ارزش دفتری حقوق صاحبان سرمایه                                                                                                                                                                                                                                        | Sales-to-equity ratio               | ۹  |
| نسبت مجموع بدهی‌ها به مجموع دارایی‌ها                                                                                                                                                                                                                                                   | Debt ratio                          | ۱۰ |
| نسبت دارایی‌های آنی به بدهی‌های جاری                                                                                                                                                                                                                                                    | Quick ratio                         | ۱۱ |
| نسبت دارایی‌های جاری به بدهی‌های جاری                                                                                                                                                                                                                                                   | Current ratio                       | ۱۲ |
| نسبت سرمایه در گردش به کل دارایی‌ها                                                                                                                                                                                                                                                     | WC/TA                               | ۱۳ |
| نسبت هزینه‌های عملیاتی به فروش خالص                                                                                                                                                                                                                                                     | Operating expenses ratio            | ۱۴ |
| نسبت سود عملیاتی به فروش خالص                                                                                                                                                                                                                                                           | Operating profit margin             | ۱۵ |
| نسبت سود خالص به فروش خالص                                                                                                                                                                                                                                                              | Profit margin                       | ۱۶ |
| نسبت سود قبل از مالیات به سرمایه                                                                                                                                                                                                                                                        | Pre-tax income-to-capital ratio     | ۱۷ |
| نسبت سود عملیاتی به سرمایه                                                                                                                                                                                                                                                              | Operating income-to-capital ratio   | ۱۸ |
| نسبت سود سال جاری به سال قبل                                                                                                                                                                                                                                                            | PROFITV                             | ۱۹ |
| تغییر در فروش خالص تقسیم بر فروش خالص در پایان سال t-1                                                                                                                                                                                                                                  | Sales revenue growth rate           | ۲۰ |
| تغییر در حاشیه سود عملیاتی تقسیم بر حاشیه سود عملیاتی در پایان سال t-1                                                                                                                                                                                                                  | Operating profit margin growth rate | ۲۱ |
| تغییر در سود خالص تقسیم بر سود خالص در پایان سال t-1                                                                                                                                                                                                                                    | Net income growth rate              | ۲۲ |
| تغییر در ارزش دفتری حقوق صاحبان سهام تقسیم بر ارزش دفتری حقوق صاحبان سهام در پایان سال t-1                                                                                                                                                                                              | Equity growth rate                  | ۲۳ |
| نسبت جریان نقد عملیاتی به مجموع دارایی‌ها                                                                                                                                                                                                                                               | Operating cash flow ratio           | ۲۴ |
| نسبت کیفیت برابر با نسبت سود خالص به جریان نقد عملیاتی                                                                                                                                                                                                                                  | SMOOTH                              | ۲۵ |
| نسبت سود قبل از بهره و مالیات به هزینه بهره                                                                                                                                                                                                                                             | Times interest earned               | ۲۶ |
| نسبت مجموع بدهی‌ها به ارزش دفتری حقوق صاحبان سرمایه                                                                                                                                                                                                                                     | Debt-to-equity ratio                | ۲۷ |
| سود هر سهم                                                                                                                                                                                                                                                                              | Earnings per share                  | ۲۸ |
| نسبت ارزش بازار به ارزش دفتری                                                                                                                                                                                                                                                           | Growth                              | ۲۹ |
| اگر سود گزارش شده شرکت منفی باشد، برابر با یک و در غیر این صورت، برابر با صفر                                                                                                                                                                                                           | Loss                                | ۳۰ |
| حد عملکرد برابر با اقلام تعهدی غیراختیاری شرکت i در سال t-1 منهای اقلام تعهدی غیراختیاری شرکت i در سال t-2                                                                                                                                                                              | THOD                                | ۳۱ |
| اقلام تعهدی اختیاری سال قبل شرکت                                                                                                                                                                                                                                                        | DA <sub>n-1</sub>                   | ۳۲ |
| فعالیت‌های تأمین مالی: اگر افزایش سرمایه شرکت بیشتر از ۱۰ درصد باشد، برابر با یک و در غیر این صورت، برابر با صفر                                                                                                                                                                        | SHARVAR                             | ۳۳ |
| اندازه شرکت برابر با لگاریتم طبیعی مجموع دارایی‌ها                                                                                                                                                                                                                                      | Size firm                           | ۳۴ |
| سرمایه‌بر بودن فعالیت برابر با نسبت خالص دارایی‌های ثابت مشهود به کل دارایی‌ها                                                                                                                                                                                                          | Capital-intensive activities        | ۳۵ |
| تعداد سال‌هایی که شرکت در سازمان بورس اوراق بهادار پذیرفته شده است.                                                                                                                                                                                                                     | AGE                                 | ۳۶ |
| رقابت در بازار محصول برابر با نسبت فروش شرکت به کل فروش صنعت                                                                                                                                                                                                                            | competition market                  | ۳۷ |
| رتبه نقدشوندگی شرکت طبق تابلوی بورس به صورت سالانه                                                                                                                                                                                                                                      | Liquidity rating                    | ۳۸ |
| درماندگی مالی با استفاده از فرمول Z-score آلتمن محاسبه می‌شود، به پیروی از دمیرکان و پلات (Demirkan & Platt, 2009) شرکت‌هایی که امتیاز Z-score آنها کمتر از ۱/۸۱ باشد به عنوان شرکت‌هایی که با درماندگی مالی مواجه هستند طبقه‌بندی می‌شوند، برابر با یک و در غیر این صورت، برابر با صفر | Financial Distress                  | ۳۹ |

ویژگی‌های  
شرکت

|                                                                                                                                               |                                |                      |
|-----------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|----------------------|
| جریان نقد آزاد برابر با جریان وجه نقد عملیاتی منهای بهره پرداختی و سود سهام پرداختی به سهامداران ممتاز و عادی تقسیم بر جمع دارایی‌ها در سال t | FCF                            | ۴۰                   |
| اگر شرکت توسط مؤسسه حسابرسی و مفید راهبر حسابرسی شده باشد، برابر با یک و در غیر این صورت، برابر با صفر                                        | Auditor Size                   | ۴۱                   |
| اگر گزارش حسابرس مقبول باشد، برابر با یک و در غیر این صورت، برابر با صفر                                                                      | Auditor OPINION                | ۴۲                   |
| تخصص حسابرس برابر با سهم بازار مؤسسه حسابرسی در یک صنعت خاص                                                                                   | Auditor expertise              | ویژگی‌های حسابرسی ۴۳ |
| تعداد سال‌هایی که یک مؤسسه حسابرسی، حسابرس شرکت است                                                                                           | Auditor tenure                 | ۴۴                   |
| اگر حسابرس نسبت به دوره قبل تغییر کرده باشد، برابر با یک و در غیر این صورت، برابر با صفر                                                      | Change of auditor              | ۴۵                   |
| اندازه هیئت‌مدیره برابر با تعداد اعضای هیئت‌مدیره                                                                                             | Board Size                     | ۴۶                   |
| استقلال هیئت‌مدیره برابر با نسبت تعداد اعضای غیرموظف به مجموع اعضای هیئت‌مدیره                                                                | Board Ind                      | ۴۷                   |
| تنوع جنسیتی هیئت‌مدیره: در صورتی که یکی از اعضای هیئت‌مدیره زن باشد، برابر با یک و در غیر این صورت، برابر با صفر                              | Board gender diversity         | حاکمیت شرکتی ۴۸      |
| نسبت تعداد مدیران زن در هیئت‌مدیره به مجموع اعضای هیئت‌مدیره                                                                                  | Percentage of women directors  | ۴۹                   |
| اگر مدیر مالی شرکت زن باشد، برابر با یک و در غیر این صورت، برابر با صفر                                                                       | CFO                            | ۵۰                   |
| نسبت تعداد مدیران خانوادگی هیئت‌مدیره به مجموع اعضای هیئت‌مدیره                                                                               | FAMILY PERCENT                 | ۵۱                   |
| میزان سهام متعلق به سهامداران نهادی                                                                                                           | Institutional Ownership        | ۵۲                   |
| مالکیت کنترلی برابر با درصد سهام متعلق به بزرگ‌ترین سهام‌دار عمده                                                                             | Controlling Ownership          | ۵۳                   |
| سهامداران عمده برابر با درصد سهام متعلق به سهامداران بالای ۵ درصد                                                                             | Concentration Ownership        | ۵۴                   |
| تعداد مالکان بالای ۵ درصد                                                                                                                     | Concentration Ownership Number | ۵۵                   |
| درصد مالکیت سهام شرکت توسط دولت یا بنگاه‌های دولتی                                                                                            | GOVERNMENT Ownership           | ساختار مالکیت ۵۶     |
| میزان مالکیت سهام‌دار حقیقی یا حقوقی که اعضای هیئت‌مدیره نماینده اوست                                                                         | CEO Ownership                  | ۵۷                   |
| اگر میزان سهام در دست سهامداران خانوادگی بیشتر از ۲۰ درصد باشد، شرکت خانوادگی است و برابر با یک و در غیر این صورت، برابر با صفر               | Family firms                   | ۵۸                   |
| اگر مجموع مالکیت دولت در یک شرکت بالای ۵۰ درصد باشد، برابر با یک و در غیر این صورت، برابر با صفر                                              | GOVer FIMR                     | ۵۹                   |
| دوره تصدی مدیرعامل                                                                                                                            | CEO Tenure                     | ۶۰                   |
| اگر مدیرعامل نسبت به دوره قبل تغییر کرده باشد، برابر با یک و در غیر این صورت، برابر با صفر                                                    | CEO changes                    | ویژگی‌های مدیریتی ۶۱ |
| اگر اعضای هیئت‌مدیره تحصیلات مالی و حسابداری یا تجربه کاری مالی و حسابداری داشته باشند، برابر با یک و در غیر این صورت، برابر با صفر           | CEO financial expertise        | ۶۲                   |
| اندازه کمیته حسابرسی برابر با لگاریتم تعداد مدیران کمیته حسابرسی شرکت                                                                         | AC_size                        | کمیته ۶۳             |
| استقلال کمیته حسابرسی برابر با تعداد مدیران غیرموظف تقسیم بر مجموع مدیران در کمیته حسابرسی                                                    | AC_independence                | حسابرسی ۶۴           |

## انتخاب ویژگی

انتخاب ویژگی تکنیکی است که برای کاهش ابعاد بردار ویژگی در مجموعه داده‌های با ابعاد بالا به کار می‌رود (Oreski & Oreski, 2014)؛ یعنی ویژگی‌های نويزدار و اضافی (متغیرها) را برای یادگیری مدل در مجموعه داده‌های بزرگ که ویژگی‌های زیادی از متغیرها را در خود جای می‌دهند، حذف می‌کند. انتخاب ویژگی در مسائل با ابعاد بالا مانند مدیریت سود، تلاش می‌کند تا زیرمجموعه‌ای از ویژگی‌ها را تعیین کند که برای دستیابی به دقت پیش‌بینی بهینه برای تشخیص مدیریت سود کمک می‌کنند. به طور خلاصه، هدف این تکنیک افزایش دقت پیش‌بینی و قدرت توضیحی مدل طبقه‌بندی و در عین حال، کاهش هزینه‌های محاسباتی و پیچیدگی دستیابی به یک راه‌حل بهینه است (Oreski & Oreski, 2014). مطالعات قبلی اهمیت و برتری استفاده از انتخاب ویژگی همراه با الگوریتم‌های طبقه‌بندی را برای ایجاد طبقه‌بندی‌کننده‌های یادگیری گروهی برای غلبه بر مسائل مالی مختلف مانند پیش‌بینی ورشکستگی شرکت (Tian et al., 2015)، امتیازدهی اعتباری (Koutanaei et al., 2015) و مدیریت سود (Chen et al., 2015) نشان می‌دهند.

تکنیک‌های انتخاب ویژگی به دو رویکرد تقسیم می‌شوند: روش‌های فیلترینگ/تکی (Filtering) و روش‌های بسته‌بندی/گروهی/ترکیبی (Wrapper) (Li et al., 2014; Yu et al., 2014). تکنیک‌های فیلتر از یک تابع فیلتر برای رتبه‌بندی ویژگی‌ها و انتخاب زیرمجموعه بهینه از ویژگی‌ها بر اساس روش رتبه‌بندی با کاوش در فضای امکان‌پذیر در میان ویژگی‌ها استفاده می‌کنند؛ از این رو، تکنیک‌های انتخاب ویژگی فیلتر مشخصات اولیه ویژگی‌ها را در نظر می‌گیرند و به طبقه‌بندی‌کننده اصلی مدل پیش‌بینی (طبقه‌بندی) وابسته نیستند (Li et al., 2014; Yu et al., 2014). این عدم وابستگی به الگوریتم طبقه‌بندی اصلی روش یادگیری ماشین تکنیک‌های فیلتر کردن را سریع می‌کند (Oreski & Oreski, 2014). در این پژوهش، از دو تکنیک انتخاب ویژگی مبتنی بر روش تکی از جمله ریلیف و تحلیل مؤلفه‌های اصلی برای انتخاب متغیرهای پیش‌بین بهینه برای پیش‌بینی مدیریت سود استفاده شده است. استفاده از رویکرد انتخاب ویژگی این امکان را فراهم می‌کند تا متغیرهای دارای محتوای اطلاعاتی بیشتر را برای تشخیص مدیریت سود شناسایی کنیم.

## ریلیف

انتخاب ویژگی ریلیف به عنوان یک روش مؤثر و ساده در ارزیابی کیفیت ویژگی‌ها شناخته شده است (Kira & Rendell, 1992; Wei et al., 2017) که در آن برای هر ویژگی، یک امتیاز وزنی بر اساس میزان تفاوت ارزش ویژگی با جفت‌های نمونه نزدیک‌ترین همسایه محاسبه می‌شود. این امتیاز بر اساس تفاوت بین ارزش ویژگی و جفت‌های نمونه همسایه مختلف تا زمانی که امتیاز ویژگی نهایی به دست آید، به تغییر ادامه می‌دهد (Kira & Rendell, 1992; Wei et al., 2017).

## تحلیل مؤلفه‌های اصلی

انتخاب ویژگی تحلیل مؤلفه‌های اصلی با تبدیل متعامد ویژگی‌های همبسته به زیرمجموعه‌ای از ویژگی‌های غیرهمبسته، ویژگی‌های اضافی را که بسیار همبسته و مرتبط با سایر ویژگی‌ها هستند، حذف می‌کند (Geiger & Kubin, 2012; Roweis, 2012).

(1998). مطالعات قبلی قدرت این تکنیک را در پژوهش‌های مختلف، مانند مشکل هماهنگ‌سازی گزارش‌های مالی (Robu & Istrate, 2015)، پیش‌بینی ورشکستگی شرکت‌ها (Tsai, 2009) و کشف تقلب مالی (Ngai et al., 2011) نشان می‌دهند.

### روش اجرای پژوهش

پس از مشخص شدن متغیرهای پیش‌بین بهینه به وسیله هر یک از روش‌های انتخاب ویژگی، از روش‌های یادگیری ماشین برای پیش‌بینی مدیریت سود استفاده شده است. به منظور ارزیابی عملکرد هر یک از روش‌های یادگیری ماشین برای پیش‌بینی مدیریت سود از معیارهای ارزیابی میانگین صحت پیش‌بینی، خطاهای نوع اول و خطاهای نوع دوم استفاده می‌شود. به عبارت دیگر، میانگین صحت پیش‌بینی مدیریت سود و خطای نوع اول و دوم با استفاده از هر یک از روش‌های پیش‌بینی مقایسه می‌شود. برای ارزیابی عملکرد هر یک از روش‌های استفاده‌شده برای تشخیص مدیریت سود، از ماتریس دقت پیش‌بینی نشان‌داده‌شده در نگاره (۲) استفاده می‌شود. در این ماتریس، خطای نوع اول به وضعیتی اشاره دارد که مدیریت سود واقعاً «بسیار رو به پایین» یا «بسیار رو به بالا» است؛ اما طبقه‌بندی‌کننده وضعیت خروجی نادرست مدیریت سود جزئی را پیش‌بینی می‌کند. در مقابل، یک خطای نوع ۲ نشان می‌دهد وضعیت مدیریت سود نسبتاً جزئی به اشتباه به عنوان بسیار رو به بالا یا پایین طبقه‌بندی می‌شود. رابطه‌های (۱) تا (۳) به ترتیب فرمول‌های محاسبه خطای نوع اول، خطای نوع دوم و میانگین صحت پیش‌بینی را نشان می‌دهند. با توجه به طبقه‌بندی درجه مدیریت سود ارقام تعهدی، خطای نوع اول بسیار مهم‌تر از خطای نوع دوم است؛ زیرا فعالیت‌های مدیریت سود بسیار رو به بالا/پایین ممکن است برای شرکت‌ها و سهام‌داران پرریسک‌تر از مدیریت سود متوسط باشند.

$$TypeI = \frac{FMUD}{FMUD + TEUD} \quad (1)$$

$$TypeII = \frac{FEUD}{FEUD + TMUD} \quad (2)$$

$$Average\ accuracy = \frac{TEUD + TMUD}{FEUD + TMUD + TEUD + FMUD} \quad (3)$$

نگاره ۲. ماتریس صحت پیش‌بینی  
Table 2. Prediction accuracy matrix

| وضعیت واقعی                                        |                                                  | وضعیت مدیریت سود                            |
|----------------------------------------------------|--------------------------------------------------|---------------------------------------------|
| وضعیت نسبتاً جزئی                                  | وضعیت بسیار رو به بالا یا بسیار رو به پایین      | وضعیت بسیار رو به بالا یا بسیار رو به پایین |
| وضعیت نادرست بسیار رو به بالا یا بسیار رو به پایین | وضعیت درست بسیار رو به بالا یا بسیار رو به پایین |                                             |
| <sup>۱۱</sup> (FEUD)                               | <sup>۱۱</sup> (TEUD)                             | وضعیت نسبتاً جزئی                           |
| وضعیت درست نسبتاً جزئی <sup>۱۳</sup> (TMUD)        | وضعیت نادرست نسبتاً جزئی <sup>۱۲</sup> (FMUD)    |                                             |

<sup>10</sup> True extremely upward or downward

<sup>11</sup> False extremely upward or downward

<sup>12</sup> False moderately upward or downward

<sup>13</sup> True moderately upward or downward

در پژوهش حاضر، به منظور بررسی تعمیم‌پذیری پیش‌بینی‌های انجام‌شده از روایی متقابل ۱۰ بخشی برای به حداقل رساندن خطای احتمالی پیش‌بینی طبقه‌بندی‌کننده‌ها استفاده شد. در این روش، مجموعه داده مدیریت سود به ۱۰ زیرمجموعه مساوی تقسیم می‌شود. همچنین، مجموعه داده به دو نمونه آموزش و آزمایش تقسیم می‌شود. هر تکرار با ۹ زیرمجموعه که نقش مجموعه داده آموزشی را ایفا می‌کنند آغاز می‌شود و آخرین زیرمجموعه به عنوان مجموعه داده آزمایشی در نظر گرفته می‌شود. این روش ۱۰ بار تکرار می‌شود تا هر یک از ۱۰ مجموعه داده اولیه یک بار نقش مجموعه داده آزمایشی را ایفا کند. صحت پیش‌بینی کلی به عنوان میانگین در تمام ۱۰ گروه محاسبه می‌شود. استفاده از روایی متقابل از بروز مشکل بیش‌برازش<sup>۱۴</sup> و مشکلات مربوط به نتایج برون‌نمونه‌ای<sup>۱۵</sup> جلوگیری می‌کند (Arlot & Celisse, 2010). روش اجرای پژوهش در شکل (۱) ارائه شده است.

### یافته‌های پژوهش

#### طبقه‌بندی درجه مدیریت سود اقلام تعهدی و مدیریت سود واقعی

در پژوهش حاضر، برای طبقه‌بندی درجه مدیریت سود اقلام تعهدی و مدیریت سود واقعی از گره گسسته‌سازی استفاده شده است. در این گره، از تکنیک میانگین و انحراف معیار برای گسسته‌سازی استفاده می‌شود که به این صورت عمل می‌کند: از مجموعه نمونه‌ها که در اینجا مدیریت سود است، میانگین و انحراف معیار آن‌ها را محاسبه می‌کند و سپس در مرحله اول یک درجه از مقدار انحراف معیار را به مقدار میانگین اضافه می‌کند. مقدار میانگین به علاوه یک درجه انحراف معیار را به عنوان سقف (مدیریت سود بسیار به سمت بالا) در نظر می‌گیرد و در مرحله دوم، یک درجه از مقدار انحراف معیار را از مقدار میانگین کم می‌کند، مقدار میانگین منهای یک درجه انحراف معیار را به عنوان کف (مدیریت سود بسیار به سمت پایین) در نظر می‌گیرد. با این حال، مقداری از مدیریت سود که بین سقف و کف باشد، رفتار مدیریت سود جزئی (اندک) تلقی می‌شود. آمار توصیفی مدیریت سود در نگاره (۳) و فواصل طبقه‌بندی مدیریت سود اقلام تعهدی و مدیریت سود واقعی به ترتیب در نگاره‌های (۴) و (۵) ارائه شده‌اند.

نگاره ۳. آمار توصیفی مدیریت سود

Table 3. Descriptive statistics of earnings management

| نوع مدیریت سود         | میانگین | انحراف معیار | میانگین + انحراف معیار | میانگین - انحراف معیار |
|------------------------|---------|--------------|------------------------|------------------------|
| مدیریت سود اقلام تعهدی | -۰/۰۰۵۳ | ۰/۱۳۸۶       | ۰/۱۳۳۳                 | -۰/۱۴۳۹                |
| مدیریت سود واقعی       | ۰/۰۰۲۵  | ۰/۱۵۵۷       | ۰/۱۵۸۲                 | -۰/۱۵۳۲                |

<sup>14</sup> Overfitting

<sup>15</sup> Out-of- Sample

نگاره ۴. فواصل طبقه‌بندی مدیریت سود اقلام تعهدی

Table 4. Accrual earnings management categorization intervals

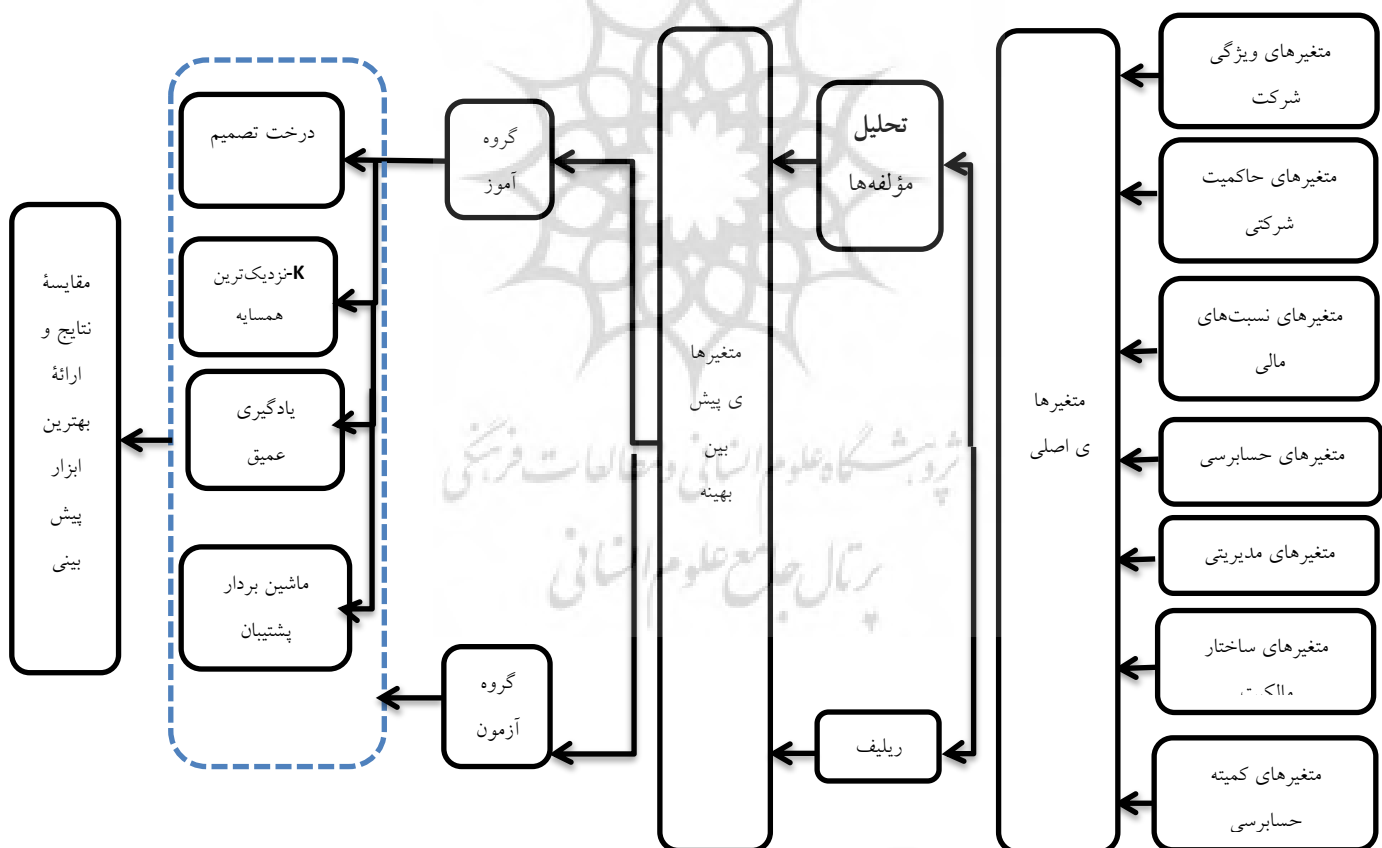
| برچسب | فاصله طبقه‌بندی                                   | نام طبقه‌بندی            | تعداد مشاهدات |
|-------|---------------------------------------------------|--------------------------|---------------|
| -۱    | مدیریت سود بسیار رو به پایین (مدیریت سود کاهنده)  | $AEM < -0.1439$          | ۱۷۳           |
| ۰     | مدیریت سود جزئی                                   | $-0.1439 < AEM < 0.1333$ | ۱۱۱۸          |
| ۱     | مدیریت سود بسیار رو به بالا (مدیریت سود افزایشده) | $AEM > 0.1333$           | ۱۷۳           |

نگاره ۵. فواصل طبقه‌بندی مدیریت سود واقعی

Table 5. Real earnings management categorization intervals

| برچسب | فاصله طبقه‌بندی                                   | نام طبقه‌بندی            | تعداد مشاهدات |
|-------|---------------------------------------------------|--------------------------|---------------|
| -۱    | مدیریت سود بسیار رو به پایین (مدیریت سود کاهنده)  | $REM < -0.1532$          | ۱۶۹           |
| ۰     | مدیریت سود جزئی                                   | $-0.1532 < REM < 0.1582$ | ۱۱۱۷          |
| ۱     | مدیریت سود بسیار رو به بالا (مدیریت سود افزایشده) | $REM > 0.1582$           | ۱۷۸           |

منبع: یافته‌های پژوهش



شکل ۱. روش اجرای پژوهش

Figure 1. Research implementation method

## انتخاب متغیرهای پیش‌بین بهینه

در این پژوهش، در مرحله دوم، از روش‌های انتخاب ویژگی تحلیل مؤلفه‌های اصلی و ریلیف برای کاهش ابعاد داده‌ها و به دست آوردن متغیرهای مستقل بااهمیت استفاده شد. متغیرهای مستقل بهینه انتخاب شده با هر یک از روش‌ها به همراه وزن آن‌ها که حاکی از سطح اهمیت هر ویژگی است، در نگاره‌های (۶) و (۷) ارائه شده‌اند.

## تحلیل مؤلفه‌های اصلی (PCA)

در این قسمت، پس از وارد کردن داده‌های به دست آمده از مراحل قبل، تابع PCA با آستانه واریانس ۸۵ درصد را انتخاب می‌کنیم. نتایج حاصل از تابع PCA در نگاره (۶) ارائه شده است.

## نگاره ۶. واریانس و واریانس تجمعی

Table 6. Variance and cumulative variance

| مؤلفه | انحراف معیار | نسبت واریانس | واریانس تجمعی | مؤلفه | انحراف معیار | نسبت واریانس | واریانس تجمعی |
|-------|--------------|--------------|---------------|-------|--------------|--------------|---------------|
| PC1   | ۲/۸۷۵        | ۰/۱۳۱        | ۰/۱۳۱         | PC16  | ۱/۰۷۹        | ۰/۰۱۸        | ۰/۶۵۹         |
| PC2   | ۲/۳۶۹        | ۰/۰۸۹        | ۰/۲۲۰         | PC17  | ۱/۰۴۴        | ۰/۰۱۷        | ۰/۶۷۶         |
| PC3   | ۲/۰۰۲        | ۰/۰۶۴        | ۰/۲۸۴         | PC18  | ۱/۰۲۳        | ۰/۰۱۷        | ۰/۶۹۳         |
| PC4   | ۱/۷۹۸        | ۰/۰۵۱        | ۰/۳۳۵         | PC19  | ۱/۰۰۹        | ۰/۰۱۶        | ۰/۷۰۹         |
| PC5   | ۱/۵۹۶        | ۰/۰۴۰        | ۰/۳۷۶         | PC20  | ۱/۰۰۳        | ۰/۰۱۶        | ۰/۷۲۵         |
| PC6   | ۱/۵۶۶        | ۰/۰۳۹        | ۰/۴۱۵         | PC21  | ۰/۹۹۷        | ۰/۰۱۶        | ۰/۷۴۱         |
| PC7   | ۱/۴۸۲        | ۰/۰۳۵        | ۰/۴۴۹         | PC22  | ۰/۹۷۴        | ۰/۰۱۵        | ۰/۷۵۶         |
| PC8   | ۱/۳۶۰        | ۰/۰۲۹        | ۰/۴۷۹         | PC23  | ۰/۹۷۱        | ۰/۰۱۵        | ۰/۷۷۱         |
| PC9   | ۱/۳۳۳        | ۰/۰۲۸        | ۰/۵۰۷         | PC24  | ۰/۹۶۰        | ۰/۰۱۵        | ۰/۷۸۶         |
| PC10  | ۱/۲۷۴        | ۰/۰۲۶        | ۰/۵۳۳         | PC25  | ۰/۹۳۹        | ۰/۰۱۴        | ۰/۸۰۰         |
| PC11  | ۱/۲۴۸        | ۰/۰۲۵        | ۰/۵۵۸         | PC26  | ۰/۹۳۴        | ۰/۰۱۴        | ۰/۸۱۳         |
| PC12  | ۱/۲۲۷        | ۰/۰۲۴        | ۰/۵۸۱         | PC27  | ۰/۹۱۶        | ۰/۰۱۳        | ۰/۸۲۷         |
| PC13  | ۱/۱۵۰        | ۰/۰۲۱        | ۰/۶۰۲         | PC28  | ۰/۹۰۰        | ۰/۰۱۳        | ۰/۸۴۰         |
| PC14  | ۱/۱۰۷        | ۰/۰۱۹        | ۰/۶۲۲         | PC29  | ۰/۸۸۶        | ۰/۰۱۲        | ۰/۸۵۲         |
| PC15  | ۱/۰۸۹        | ۰/۰۱۹        | ۰/۶۴۱         |       |              |              |               |

منبع: یافته‌های پژوهش

با توجه به واریانس تجمعی به دست آمده در نگاره (۶)، روش تحلیل مؤلفه‌های اصلی به استخراج ۲۹ عامل منجر شد که این عامل‌ها حدود ۸۵ درصد از تغییرپذیری (واریانس) متغیرهای اولیه را نشان می‌دهند. به عبارت دیگر، می‌توان ۶۴ متغیر اولیه را در ۲۹ عامل کلی خلاصه و حدود ۸۵ درصد از اطلاعات اولیه را نیز حفظ کرد.

## روش ریلیف

روش انتخاب متغیر ریلیف برای هر یک از ویژگی‌ها وزنی را تعیین و کلیه متغیرهای پیش‌بین را بر اساس وزن اختصاص‌یافته رتبه‌بندی می‌کند. در این پژوهش، مطابق پژوهش‌های قبلی حداقل وزن ۰/۱۵ به عنوان معیار انتخاب یک ویژگی در زیرمجموعه بهینه ویژگی‌ها به کار گرفته شد (Fallahpour et al., 2017; Koutanaei et al., 2015; Oreski & Oreski, 2014; Hammami & Hendijani Zadeh, 2022). بر این اساس و به منظور قابلیت مقایسه بهتر با روش تحلیل مؤلفه‌های اصلی، از ۲۹ متغیر اول (بهتر) رتبه‌بندی‌شده در این روش استفاده می‌شود که این متغیرها نیز به ترتیب وزن اختصاص‌یافته و رتبه در نگاره (۷) ارائه شده‌اند.

نگاره ۷. وزن و رتبه ویژگی‌ها بر اساس روش ریلیف

Table 7. The weight and rank of the features based on the relief method

| وزن اختصاص‌یافته |                        | ویژگی                          | ردیف |
|------------------|------------------------|--------------------------------|------|
| مدیریت سود واقعی | مدیریت سود اقلام تعهدی |                                |      |
| ۰/۶۵۵            | ۰/۵۹۰                  | Auditor expertise              | ۱    |
| ۰/۴۱۷            | ۰/۵۴۸                  | Auditor Size                   | ۲    |
| ۰/۳۶۵            | ۰/۵۲۴                  | Financial Distress             | ۳    |
| ۰/۴۹۲            | ۰/۵۰۴                  | OPINION                        | ۴    |
| —                | ۰/۵۰۱                  | Operating cash flow ratio      | ۵    |
| ۰/۴۰۲            | ۰/۴۹۰                  | CEO financial expertise        | ۶    |
| ۰/۵۳۸            | ۰/۴۵۷                  | GOVer FIMR                     | ۷    |
| —                | ۰/۴۲۷                  | FCF                            | ۸    |
| ۰/۴۶۱            | ۰/۳۹۵                  | GOVERNMENT Ownership           | ۹    |
| ۰/۲۹۰            | ۰/۳۸۶                  | Controlling ownership          | ۱۰   |
| ۰/۳۶۳            | ۰/۳۲۸                  | Institutional ownership        | ۱۱   |
| ۰/۲۰۴            | ۰/۳۲۲                  | Capital-intensive activities   | ۱۲   |
| ۰/۱۶۴            | ۰/۲۹۸                  | SHARVAR                        | ۱۳   |
| ۰/۲۷۴            | ۰/۲۸۹                  | CEO changes                    | ۱۴   |
| ۰/۲۴۷            | ۰/۲۷۲                  | Change of auditor              | ۱۵   |
| ۰/۲۲۰            | ۰/۲۵۵                  | AGE                            | ۱۶   |
| ۰/۳۳۴            | ۰/۲۵۳                  | Family firms                   | ۱۷   |
| ۰/۲۵۷            | ۰/۲۴۷                  | Board gender diversity         | ۱۸   |
| ۰/۲۲۴            | ۰/۲۴۶                  | CEO Ownership                  | ۱۹   |
| ۰/۲۰۲            | ۰/۲۴۳                  | Debt ratio                     | ۲۰   |
| ۰/۲۲۶            | ۰/۲۰۲                  | Concentration Ownership Number | ۲۱   |
| ۰/۱۶۰            | ۰/۱۹۵                  | WC/TA                          | ۲۲   |
| ۰/۱۶۱            | ۰/۱۸۵                  | ROA                            | ۲۳   |
| ۰/۱۶۳            | ۰/۱۸۳                  | Total assets turnover          | ۲۴   |
| ۰/۱۵۹            | ۰/۱۸۲                  | Liquidity rating               | ۲۵   |
| ۰/۲۱۱            | ۰/۱۸۲                  | Loss                           | ۲۶   |

|       |       |                               |    |
|-------|-------|-------------------------------|----|
| ۰/۱۸۲ | ۰/۱۸۱ | Concentration Ownership       | ۲۷ |
| —     | ۰/۱۷۳ | ROE                           | ۲۸ |
| —     | ۰/۱۶۳ | Total assets growth rate      | ۲۹ |
| ۰/۱۶۵ | —     | Size firm                     | ۳۰ |
| ۰/۱۵۶ | —     | AC_independence               | ۳۱ |
| ۰/۱۵۳ | —     | Board Ind                     | ۳۲ |
| ۰/۱۵۰ | —     | Percentage of women directors | ۳۳ |

منبع: یافته‌های پژوهش

همان‌طور که در نگاره (۷) مشاهده می‌شود، بیشترین رتبه متغیرهای بهینه انتخاب‌شده مربوط به ویژگی‌های حسابرس و بعد از آن، مربوط به ویژگی‌های ساختار مالکیت شرکت‌ها است. در وهله بعدی، متغیرهای مربوط به ویژگی نسبت‌های مالی حائز اهمیت هستند. به بیان دیگر، سرمایه‌گذاران باید توجه زیادی به ویژگی‌های حسابرس و ساختار مالکیت شرکت‌ها در پیش‌بینی مدیریت سود داشته باشند؛ این در حالی است که در پژوهش‌های قبلی انجام‌شده در حوزه پیش‌بینی مدیریت سود عمدتاً به نسبت‌های مالی توجه شده است.

### نتایج آزمون فرضیه‌ها

#### فرضیه اول

بر اساس فرضیه اول پژوهش، پیش‌بینی می‌شود عملکرد پیش‌بینی مدیریت سود اقلام تعهدی بر اساس الگوی روش انتخاب ویژگی مبتنی بر ریلیف نسبت به روش انتخاب ویژگی مبتنی بر تحلیل مؤلفه‌های اصلی بهتر باشد. در ادامه، به منظور آزمون فرضیه اول، نتایج هر یک از روش‌های ارائه‌شده در نگاره (۸) را به طور جداگانه با هم مقایسه می‌کنیم. به منظور مقایسه هر یک از روش‌ها از میانگین صحت پیش‌بینی، خطای نوع اول و دوم استفاده می‌شود.

#### نگاره ۸. عملکرد روش‌های مختلف پیش‌بینی مدیریت سود اقلام تعهدی

Table 8. The performance of different forecasting methods of accrual earnings management

| SVM   |        | Knn   |        | DL    |        | DT    |        | روش معیار        |          |
|-------|--------|-------|--------|-------|--------|-------|--------|------------------|----------|
| PCA   | relief | PCA   | relief | PCA   | relief | PCA   | relief | روش انتخاب ویژگی |          |
| ۷۷/۳۲ | ۸۳/۴۷  | ۸۱/۸۳ | ۸۲/۹۹  | ۸۱/۸۳ | ۸۹/۶۲  | ۷۷/۵۳ | ۸۲/۹۹  | صحت پیش‌بینی     | عملکرد   |
| ۹۲/۳۷ | ۹۴/۰۷  | ۸۳/۵۱ | ۸۵/۹۸  | ۸۳/۵۱ | ۸۹/۹۴  | ۸۱/۳۸ | ۸۸/۱۲  | دقت پیش‌بینی     | روش      |
| ۳۶/۰۳ | ۵۳/۳۷  | ۵۱/۱۹ | ۵۳/۹۸  | ۵۱/۱۹ | ۷۴/۶۲  | ۳۶/۶۱ | ۵۳/۰۰  | پوشش             | پیش‌بینی |
| ۰     | ۰      | ۱۳/۶۳ | ۹/۹۱   | ۱۳/۶۳ | ۹/۹۱   | ۰     | ۵/۵۰   | خطای نوع اول     |          |
| ۲۲/۸۹ | ۱۷/۷۹  | ۱۸/۵۳ | ۱۷/۶۴  | ۱۸/۴۱ | ۱۰/۴۷  | ۲۲/۷۳ | ۱۷/۹۳  | خطای نوع دوم     |          |

منبع: یافته‌های پژوهش

بر اساس نگاره (۸)، با استفاده از روش درخت تصمیم، در حالت استفاده از متغیرهای انتخاب‌شده در روش مبتنی بر ریلیف و تحلیل مؤلفه‌های اصلی، صحت پیش‌بینی به ترتیب برابر با ۸۲/۹۹ و ۷۷/۵۳ است. همچنین، خطای نوع اول (خطای نوع دوم)

نیز به‌ترتیب برابر با ۵/۵۰ (۱۷/۹۳) و ۰ (۲۲/۷۳) است. با استفاده از روش یادگیری عمیق، در حالت استفاده از متغیرهای انتخاب‌شده در روش مبتنی بر ریلیف و تحلیل مؤلفه‌های اصلی، صحت پیش‌بینی به‌ترتیب برابر با ۸۹/۶۲ و ۸۱/۸۳ است. همچنین، خطای نوع اول (خطای نوع دوم) نیز به‌ترتیب برابر با ۹/۹۱ (۱۰/۴۷) و ۱۳/۶۳ (۱۸/۴۱) است. نتایج در نگاره (۸) نشان‌دهنده آن است که با استفاده از روش K- نزدیک‌ترین همسایه، در حالت استفاده از متغیرهای انتخاب‌شده در روش مبتنی بر ریلیف و تحلیل مؤلفه‌های اصلی، صحت پیش‌بینی به‌ترتیب برابر با ۸۲/۹۹ و ۸۱/۸۳ است. همچنین، خطای نوع اول (خطای نوع دوم) نیز به‌ترتیب برابر با ۹/۹۱ (۱۷/۶۴) و ۱۳/۶۳ (۱۸/۵۳) است. با استفاده از روش ماشین بردارهای پشتیبان، در حالت استفاده از متغیرهای انتخاب‌شده در روش مبتنی بر ریلیف و تحلیل مؤلفه‌های اصلی، صحت پیش‌بینی به‌ترتیب برابر با ۸۳/۴۷ و ۷۷/۳۲ است. همچنین، خطای نوع اول (خطای نوع دوم) نیز به‌ترتیب برابر با ۰ (۱۷/۷۹) و ۰ (۲۲/۸۹) است. با توجه به عملکرد بهتر کلیه روش‌های پیش‌بینی مدیریت سود اقلام تعهدی با استفاده از متغیرهای انتخاب‌شده در روش مبتنی بر ریلیف نسبت به عامل‌های استخراج‌شده در روش‌های تحلیل مؤلفه‌های اصلی، فرضیه اول پژوهش، مبنی بر بهتر بودن عملکرد پیش‌بینی مدیریت سود اقلام تعهدی شرکت‌ها بر اساس الگوی روش انتخاب ویژگی مبتنی بر ریلیف نسبت به روش انتخاب ویژگی مبتنی بر تحلیل مؤلفه‌های اصلی، تأیید می‌شود. نکته دیگری که می‌توان به آن اشاره کرد این است که خطای نوع اول در همه روش‌های پیش‌بینی از خطای نوع دوم کمتر است. به عبارت دیگر، روش‌های پیش‌بینی، مدیریت سود اقلام تعهدی بسیار رو به بالا یا پایین را کمتر به‌اشتباه در وضعیت مدیریت سود نسبتاً جزئی طبقه‌بندی می‌کنند.

### فرضیه دوم

بر اساس فرضیه دوم پژوهش، پیش‌بینی می‌شود عملکرد پیش‌بینی مدیریت سود واقعی بر اساس الگوی روش انتخاب ویژگی مبتنی بر ریلیف نسبت به روش انتخاب ویژگی مبتنی بر تحلیل مؤلفه‌های اصلی بهتر باشد.

نگاره ۹. عملکرد روش‌های مختلف پیش‌بینی مدیریت سود واقعی

| Table 9. The performance of different forecasting methods of real earnings management |        |       |        |       |        |       |        |
|---------------------------------------------------------------------------------------|--------|-------|--------|-------|--------|-------|--------|
| SVM                                                                                   |        | Knn   |        | DL    |        | DT    |        |
| PCA                                                                                   | relief | PCA   | relief | PCA   | relief | PCA   | relief |
| ۷۶/۸۴                                                                                 | ۷۸/۳۵  | ۸۲/۶۵ | ۸۲/۰۴  | ۸۲/۶۵ | ۸۰/۲۶  | ۷۶/۷۸ | ۷۸/۳۵  |
| ۸۸/۵۵                                                                                 | ۹۲/۶۳  | ۷۸/۶۹ | ۸۴/۲۲  | ۷۸/۶۹ | ۶۷/۴۸  | ۸۲/۷۳ | ۸۷/۰۸  |
| ۳۵/۰۰                                                                                 | ۳۹/۰۹  | ۵۳/۵۹ | ۵۰/۶۹  | ۵۳/۵۹ | ۵۶/۵۰  | ۳۴/۶۶ | ۳۹/۴۲  |
| ۱۰/۰۰                                                                                 | ۰      | ۱۲/۱۹ | ۸/۰۰   | ۱۲/۱۹ | ۳۵/۹۲  | ۰     | ۵/۸۸   |
| ۲۳/۲۴                                                                                 | ۲۲/۱۰  | ۱۷/۸۲ | ۱۸/۶۹  | ۱۰/۱۸ | ۱۷/۰۹  | ۲۳/۳۳ | ۲۲/۰۲  |

منبع: یافته‌های پژوهش

بر اساس نگاره (۹)، با استفاده از روش درخت تصمیم، در حالت استفاده از متغیرهای انتخاب‌شده در روش مبتنی بر ریلیف و تحلیل مؤلفه‌های اصلی، صحت پیش‌بینی به‌ترتیب برابر با ۷۸/۳۵ و ۷۶/۷۸ است. همچنین، خطای نوع اول (خطای نوع دوم)

نیز به ترتیب برابر با ۵/۸۸ (۰) و ۲۲/۰۲ (۲۳/۳۳) است. با استفاده از روش یادگیری عمیق، در حالت استفاده از متغیرهای انتخاب شده در روش مبتنی بر ریلیف و تحلیل مؤلفه‌های اصلی، صحت پیش‌بینی به ترتیب برابر با ۸۰/۲۶ و ۸۲/۶۵ است. همچنین، خطای نوع اول (خطای نوع دوم) نیز به ترتیب برابر با ۳۵/۹۲ (۱۲/۱۹) و ۱۷/۰۹ (۱۰/۱۸) است.

نتایج در نگاره ۹) نشان‌دهنده آن است که با استفاده از روش K- نزدیک‌ترین همسایه، در حالت استفاده از متغیرهای انتخاب شده در روش مبتنی بر ریلیف و تحلیل مؤلفه‌های اصلی، صحت پیش‌بینی به ترتیب برابر با ۸۲/۰۴ و ۸۲/۶۵ است. همچنین، خطای نوع اول (خطای نوع دوم) نیز به ترتیب برابر با ۸/۰۰ (۱۲/۱۹) و ۱۸/۶۹ (۱۷/۸۲) است. با استفاده از روش ماشین بردارهای پشتیبان، در حالت استفاده از متغیرهای انتخاب شده در روش مبتنی بر ریلیف و تحلیل مؤلفه‌های اصلی، صحت پیش‌بینی به ترتیب برابر با ۷۸/۳۵ و ۷۶/۸۴ است. همچنین، خطای نوع اول (خطای نوع دوم) نیز به ترتیب برابر با ۰ (۱۰/۰۰) و ۲۲/۱۰ (۲۳/۲۴) است. نتایج حاکی از آن است که در پیش‌بینی مدیریت سود واقعی، روش یادگیری عمیق و K- نزدیک‌ترین همسایه با ترکیب با روش تحلیل مؤلفه‌های اصلی بالاترین صحت پیش‌بینی را دارند؛ اما روش درخت تصمیم و ماشین بردارهای پشتیبان با ترکیب با روش ریلیف بالاترین صحت پیش‌بینی مدیریت سود واقعی را دارند؛ بنابراین، فرضیه دوم پژوهش تأیید نمی‌شود. نکته دیگری که می‌توان به آن اشاره کرد این است که خطای نوع اول به جز در روش یادگیری عمیق از خطای نوع دوم کمتر است. به عبارت دیگر، روش‌های پیش‌بینی، مدیریت سود واقعی بسیار رو به بالا یا پایین را کمتر به اشتباه در وضعیت مدیریت سود نسبتاً جزئی طبقه‌بندی می‌کنند.

### فرضیه سوم

بر اساس فرضیه سوم پژوهش، پیش‌بینی می‌شود صحت پیش‌بینی مدیریت سود اقلام تعهدی نسبت به مدیریت سود واقعی بالاتر است. نگاره ۱۰) نشان‌دهنده مقایسه عملکرد روش‌های مختلف پیش‌بینی مدیریت سود با استفاده از ترکیب روش‌های یادگیری ماشین در ۸ حالت است.

نگاره ۱۰. مقایسه عملکرد روش‌های مختلف پیش‌بینی مدیریت سود

**Table 10. Comparing the performance of different forecasting methods of earnings management**

| روش ترکیبی یادگیری ماشین | مدیریت سود اقلام تعهدی | مدیریت سود واقعی |
|--------------------------|------------------------|------------------|
| Relife-DL                | ۸۹/۶۲                  | ۸۰/۲۶            |
| Relife-SVM               | ۸۳/۴۷                  | ۷۸/۳۵            |
| Relife-Knn               | ۸۲/۹۹                  | ۸۲/۰۴            |
| Relife-DT                | ۸۲/۹۹                  | ۷۸/۳۵            |
| PCA-DL                   | ۸۱/۸۳                  | ۸۲/۶۵            |
| PCA-Knn                  | ۸۱/۸۳                  | ۸۲/۶۵            |
| PCA-DT                   | ۷۷/۵۳                  | ۷۶/۷۸            |
| PCA-SVM                  | ۷۷/۳۲                  | ۷۶/۸۴            |

منبع: یافته‌های پژوهش

همان‌طور که در نگاره (۱۰) مشاهده می‌شود، صحت پیش‌بینی مدیریت سود اقلام تعهدی به‌جز در ۲ مورد بالاتر از صحت پیش‌بینی مدیریت سود واقعی است و در ۲ مورد که صحت پیش‌بینی مدیریت سود واقعی بالاتر از صحت پیش‌بینی مدیریت سود اقلام تعهدی است، اختلاف صحت پیش‌بینی زیاد نیست؛ بنابراین، می‌توان ادعا کرد صحت پیش‌بینی مدیریت سود اقلام تعهدی بالاتر از صحت پیش‌بینی مدیریت سود واقعی است؛ در نتیجه، فرضیه سوم پژوهش تأیید می‌شود.

مطابق نتایج مندرج در نگاره (۱۰)، مشاهده می‌شود روش ترکیبی Relife-DL بیشترین میانگین صحت پیش‌بینی (۸۹/۶۲) با خطای نوع اول (۹/۹۱) و خطای نوع دوم (۱۰/۴۷) را در میان سایر روش‌های ترکیبی برای پیش‌بینی مدیریت سود اقلام تعهدی دارد و به عنوان بهترین روش برای پیش‌بینی مدیریت سود اقلام تعهدی پیشنهاد می‌شود. همچنین، روش ترکیبی PCA-DL بیشترین میانگین صحت پیش‌بینی (۸۲/۶۵) با خطای نوع اول (۱۲/۱۹) و خطای نوع دوم (۱۰/۱۸) را در میان سایر روش‌های ترکیبی برای پیش‌بینی مدیریت سود واقعی دارد و به عنوان بهترین روش برای پیش‌بینی مدیریت سود واقعی پیشنهاد می‌شود.

### بحث و نتیجه‌گیری

مدیریت سود بر داده‌های حسابداری، به ویژه سود گزارش‌شده در حسابداری به‌غیر از سود واقعی یک شرکت، تأثیر می‌گذارد؛ بنابراین، پیش‌بینی مدیریت سود همچنان موضوعی بسیار مهم است. هدف پژوهش حاضر استفاده از روش‌های یادگیری ماشین از جمله درخت تصمیم، ماشین بردار پشتیبان،  $k$ -نزدیک‌ترین همسایه و یادگیری عمیق برای پیش‌بینی مدیریت سود است. همچنین، این پژوهش به انتخاب ویژگی برای شناسایی بهینه‌ترین ویژگی‌ها برای استفاده در مدل پیش‌بینی متکی است. با تکنیک انتخاب ویژگی بررسی شد که کدام ویژگی‌ها در پیش‌بینی مدیریت سود کارآمدتر هستند و کدام تکنیک انتخاب ویژگی، انتخاب ویژگی بهینه‌تری را ارائه می‌دهد.

بر اساس فرضیه اول، مشاهده شد عملکرد روش‌های پیش‌بینی مدیریت سود اقلام تعهدی بر اساس الگوی انتخاب ویژگی مبتنی بر ریلیف نسبت به الگوی انتخاب ویژگی مبتنی بر تحلیل مؤلفه‌های اصلی بهتر است. این نتیجه در کلیه روش‌های پیش‌بینی (یادگیری عمیق،  $K$ -نزدیک‌ترین همسایه، ماشین بردارهای پشتیبان، یادگیری تجمیعی آداپوست ماشین بردار پشتیبان) تأیید شد. این نتایج با یافته‌های همایی و هنديجانی‌زاده (2022) مطابقت ندارد و دلیل این امر ممکن است به جامعیت متغیرهای استفاده‌شده در پژوهش حاضر مربوط باشد؛ زیرا در پژوهش همایی و هنديجانی‌زاده (2022) فقط از متغیرهای نسبت مالی استفاده شده است (Hammami & Hendijani Zadeh, 2022).

بر اساس نتایج حاصل از آزمون فرضیه دوم، مشاهده شد در پیش‌بینی مدیریت سود واقعی، روش یادگیری عمیق و  $K$ -نزدیک‌ترین همسایه با ترکیب با روش تحلیل مؤلفه‌های اصلی بالاترین صحت پیش‌بینی را دارند؛ اما روش درخت تصمیم و ماشین بردارهای پشتیبان با ترکیب با روش ریلیف بالاترین صحت پیش‌بینی مدیریت سود واقعی را دارند. این نتایج با یافته‌های همایی و هنديجانی‌زاده (2022) مطابقت دارد (Hammami & Hendijani Zadeh, 2022).

همچنین، بر اساس فرضیه سوم، مشاهده شد مدیریت سود اقلام تعهدی را می‌توان با بیشتری بالاتری نسبت به مدیریت سود واقعی پیش‌بینی کرد. این نتایج، یافته‌های برخی از پژوهشگران که نشان دادند مدیران بیشتر تمایل به مدیریت سود واقعی نسبت

به مدیریت سود اقلام تعهدی دارند، زیرا آن‌ها مدیریت سود واقعی را به عنوان یک گزینه کم‌خطر و غیرقابل حدس توسط حسابرسان می‌بینند، تأیید می‌کند (Roychowdhury, 2006; Zang, 2012).

همچنین، نتایج انتخاب ویژگی بر اساس الگوی ریلیف حاکی از آن بود که مهم‌ترین ویژگی‌ها برای پیش‌بینی در وهله اول مربوط به ویژگی‌های حسابرس و بعد از آن، مربوط به ویژگی‌های ساختار مالکیت شرکت‌ها هستند. در وهله بعدی، متغیرهای مربوط به ویژگی نسبت‌های مالی حائز اهمیت هستند. به بیان دیگر، سرمایه‌گذاران باید توجه زیادی به ویژگی‌های حسابرس و ساختار مالکیت شرکت‌ها در پیش‌بینی مدیریت سود داشته باشند. همچنین، به سازمان بورس و اوراق بهادار پیشنهاد می‌شود به ساختار مالکیت شرکت‌ها و ویژگی‌های حسابرس از جمله تخصص در صنعت و اندازه آن توجه ویژه کنند؛ این در حالی است که در پژوهش‌های قبلی انجام‌شده در حوزه پیش‌بینی مدیریت سود (برای مثال، Hammami & Hendijani Zadeh, 2022; Chen & Shen, 2020; Chen et al., 2015; Najari et al., 2014؛ قادری و همکاران، ۱۳۹۷؛ کاردان و همکاران، ۱۳۹۶؛ آرسته و نصیرزاده، ۱۳۹۵)، عمدتاً به نسبت‌های مالی توجه شده است. بر اساس نتایج به‌دست‌آمده، روش ترکیبی یادگیری عمیق و انتخاب ویژگی ریلیف (Relife-DL) بیشترین میانگین صحت پیش‌بینی (۸۹/۶۲) را در میان سایر روش‌های ترکیبی برای پیش‌بینی مدیریت سود اقلام تعهدی دارد و روش ترکیبی یادگیری عمیق مبتنی و انتخاب ویژگی تحلیل مؤلفه‌های اصلی (PCA-DL) بیشترین میانگین صحت پیش‌بینی (۸۲/۶۵) را در میان سایر روش‌های ترکیبی برای پیش‌بینی مدیریت سود واقعی دارد.

یافته‌های این پژوهش می‌تواند تشخیص مدیریت سود را برای سرمایه‌گذاران شرکت‌های فعال در بورس اوراق بهادار و کاربران صورت‌های مالی با بهبود دقت پیش‌بینی مدیریت سود تسریع کند؛ از این رو، به سرمایه‌گذاران توصیه می‌شود متغیرهای مهم مالی و غیرمالی ارائه‌شده با استفاده از روش‌های انتخاب ویژگی به عنوان شاخص‌های تشخیص برای دست‌کاری سود را برای تصمیم‌گیری در زمینه سرمایه‌گذاری خود مدنظر قرار دهند. همچنین، به سرمایه‌گذاران پیشنهاد می‌شود هنگام تصمیم‌گیری بر اساس صورت‌های مالی، مدیریت سود را با توجه به انگیزه‌های مدیریت مورد توجه قرار دهند تا به کاهش ریسک‌ها و زیان‌های سرمایه‌گذاری و افزایش مزایای سرمایه‌گذاری منجر شود.

نتایج یافته‌ها می‌تواند برای کارکنان حسابرسی (حسابرسان داخلی و خارجی) جالب باشد؛ زیرا به دلیل سرعت روش‌های داده‌کاوی در تشخیص مدیریت سود، حسابرسان با هزینه‌ای کمتر نسبت به روش‌های سنتی به عوامل تأثیرگذار بر تشخیص مدیریت سود دست می‌یابند؛ در نتیجه در زمان صرفه‌جویی می‌شود و در عین حال، این روش اطلاعات بیشتری را فراهم می‌کند که این امر می‌تواند به حسابرسان در وظایف تحلیلی، تعیین ریسک حسابرسی (که بر طراحی برنامه حسابرسی تأثیر می‌گذارد) و رویه‌های حسابرسی کمک کند؛ بنابراین، به کارکنان حسابرسی و مؤسسه‌های حسابرسی توصیه می‌شود موضوع مدیریت سود و انگیزه‌های مدیران برای دست‌کاری سود را مورد توجه قرار دهند و در تدوین برنامه‌های حسابرسی و اجرای روش‌های آن با استفاده از معیارهای مالی ارائه‌شده در پژوهش به گونه‌ای مناسب احتمال مدیریت سود را مدنظر قرار دهند و با استفاده از روش‌های یادگیری ماشین از افشای ناکافی یا حتی همراه‌کننده شرکت‌ها با استفاده گسترده از مدیریت سود در گزارشگری مالی جلوگیری کنند.

به طور کلی، پژوهش حاضر می‌تواند مورد توجه سرمایه‌گذاران، اعتباردهندگان و قانون‌گذاران قرار گیرد. استفاده از روش‌های داده‌کاوی می‌تواند به سرمایه‌گذاران در شناسایی مدیریت سود بالقوه کمک کند تا (الف) هیئت‌مدیره را در انجام بهتر و مؤثر وظایف نظارتی خود تحت فشار قرار دهند و (ب) از سرمایه‌گذاری در سهام شرکت‌هایی که به طور بالقوه ممکن است با افت قیمت هنگام آشکار شدن مدیریت سود توسط این شرکت‌ها به طور عمومی پیش بروند، اجتناب کنند. از طرفی، اعتباردهندگان می‌توانند از این روش‌ها به منظور ارزیابی پروفایل ریسک اعتباری وام‌گیرندگان و احتمال نکول استفاده کنند. در نهایت، این پژوهش می‌تواند مرجعی برای افراد درگیر در پژوهش‌های دانشگاهی مرتبط با مدیریت سود، مدیران شرکت‌ها، انجمن حسابداران خبره، حسابداران رسمی و حساب‌برسان، سرمایه‌گذاران و تحلیلگران اوراق بهادار باشد.

طی انجام هر پژوهش، ابعادی گسترده‌تر و تازه‌تر از موضوع مشخص می‌شوند که می‌توانند نقطه آغازی برای پژوهش‌های بعدی باشند؛ بنابراین، با توجه به نتایج حاصل از پژوهش، پیشنهادهای زیر برای پژوهش‌های آتی ارائه می‌شوند:

۱. استفاده از سایر مدل‌های پیش‌بینی مدیریت سود و مقایسه نتایج به دست آمده با نتایج پژوهش حاضر؛
۲. پیش‌بینی مدیریت سود در سطح صنایع با توجه به ویژگی‌های خاص هر صنعت؛
۳. پیش‌بینی مدیریت سود در شرکت‌های مالی، سرمایه‌گذاری و بانک‌ها؛
۴. پیش‌بینی مدیریت سود با استفاده از سایر متغیرها از جمله متغیرهای کلان اقتصادی، ویژگی‌های شخصی مدیران عامل، ویژگی‌های بیشتر کمیته حسابرسی، ساختار مالکیت و حاکمیت شرکتی.
۵. پیش‌بینی مدیریت سود با استفاده از روش‌های انتخاب ویژگی گروهی از جمله الگوریتم ژنتیک، الگوریتم بهینه‌سازی ازدحام ذرات و مقایسه با نتایج پژوهش حاضر.

## منابع

- احمدپور، احمد، اصابت طبری، عصمت، و طالب تبار آهنگر، میثم (۱۳۹۴). رابطه بین کیفیت سود و دوره تصدی مدیرعامل. *پژوهش‌های تجربی حسابداری*، ۵(۴)، ۱-۱۴. [https://jera.alzahra.ac.ir/article\\_633.html#ar\\_info\\_pnl\\_cite](https://jera.alzahra.ac.ir/article_633.html#ar_info_pnl_cite)
- اسماعیلی، مهدی (برگرداننده) (۱۳۹۳). *داده‌کاوی: مفاهیم و تکنیک‌ها* (چاپ اول). تهران: انتشارات نیاز دانش.
- أستا، سهراب (۱۳۹۰). بررسی رابطه بین ساختار مالکیت و مدیریت سود. *پژوهش‌های حسابداری مالی*، ۳(۲)، ۹۳-۱۰۶. [https://far.ui.ac.ir/article\\_16915.html](https://far.ui.ac.ir/article_16915.html)
- آرسته، قاسم، و نصیرزاده، فرزانه (۱۳۹۵). مقایسه دقت الگوریتم‌های تخمین‌گر بردار پشتیبان، تخمین‌گر حداقل درجه و شبکه عصبی فازی در کشف مدیریت سود. *پژوهش‌های تجربی حسابداری*، ۶(۴)، ۹۵-۱۱۶. [https://jera.alzahra.ac.ir/article\\_2632.html#ar\\_info\\_pnl\\_cite](https://jera.alzahra.ac.ir/article_2632.html#ar_info_pnl_cite)
- آزادی، فرهاد، قنبری، مهرداد، جمشیدی نوید، بابک، و مسعودی، جواد (۱۴۰۰). به‌کارگیری پدیده تونلینگ جهت افزایش توانایی پیش‌بینی مدیریت سود در مدل بنیش بر مبنای تکنیک شبکه‌های عصبی مصنوعی و الگوریتم بهینه‌سازی حرکت جمعیتی ذرات. *چشم‌انداز مدیریت مالی*، ۱۱(۳۳)، ۱۳۹-۱۷۱. <https://doi.org/10.52547/jfmp.11.33.139>

- پورعلی، محمدرضا، و کوچکی تاجانی، محدثه (۱۳۹۹). *مقایسه دقت پیش‌بینی دست‌کاری سود شرکت‌ها با استفاده از الگوریتم رقابت استعماری و الگوریتم ژنتیک*. اولین کنفرانس بین‌المللی چالش‌ها و راهکارهای نوین در مهندسی صنایع و مدیریت و حسابداری، ساری. <https://civilica.com/doc/1045469>
- حسینی، سید محسن، و رشیدی، زینب (۱۳۹۲). پیش‌بینی ورشکستگی شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران با استفاده از درخت تصمیم و رگرسیون لجستیک. *نشریه پژوهش‌های حسابداری مالی*، ۵(۳)، ۱۰۵-۱۲۸. [https://far.ui.ac.ir/article\\_16987.html](https://far.ui.ac.ir/article_16987.html)
- دارابی، رویا، و اژدری، فاطمه (۱۳۹۷). بررسی رابطه نظام راهبری و ویژگی‌های حسابرسی با مدیریت سود واقعی شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران. *حسابداری و منافع اجتماعی*، ۸(۲)، ۱-۲۱. [https://jaacsi.alzahra.ac.ir/article\\_3605.html#ar\\_info\\_pnl\\_cite](https://jaacsi.alzahra.ac.ir/article_3605.html#ar_info_pnl_cite)
- داغانی، رضا، حاجیان، نجمه، و طلوعی، کبری (۱۳۹۸). تأثیر ساختار مالکیتی و نظارتی هیئت‌مدیره و ویژگی‌های حسابرسان بر مدیریت سود. *پژوهش‌های تجربی حسابداری*، ۹(۴)، ۲۹۹-۳۲۶. <https://doi.org/10.22051/jera.2018.20138.2019>
- ستایش، محمد حسین، و کاظم‌نژاد، مصطفی (۱۳۹۸). بررسی سودمندی روش‌های کاهش متغیرها در پیش‌بینی بازده سهام شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران. *مطالعات تجربی حسابداری مالی*، ۱۶(۶۳)، ۸۳-۱۰۷. <https://doi.org/10.22054/qjma.2019.10647>
- سعید مقدم، مرضیه، جاوید، داریوش، و همت فر، محمود (۱۳۹۶). تشخیص مدیریت سود صنعت خودرو و ساخت قطعات با استفاده از ترکیب شبکه بیزین و درخت تصمیم ۵,۰ C. *فصلنامه حسابداری مالی*، ۹(۳۶)، ۱۰۲-۱۲۶. <http://qfaj.mobarakeh.iau.ir/article-1-1080-fa.html>
- شاه مرادی، نسیم، و طباطبایی نسب، زهره (۱۴۰۰). بررسی تأثیر کیفیت حسابرسی بر رابطه نااطمینانی اقتصادی و مدیریت سود ناشی از اقلام تعهدی در شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران. *نشریه پژوهش‌های حسابداری مالی*، ۱۳(۱)، ۶۷-۸۶. <https://doi.org/10.22108/far.2021.125552.1683>
- صالحی، مهدی، و فرخی پیل‌رود، لاله (۱۳۹۷). پیش‌بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم. *پژوهش‌های حسابداری مالی و حسابرسی*، ۱۰(۳۷)، ۱-۲۴. <https://sanad.iau.ir/Journal/faar/Article/1073544>
- عارف‌منش، زهره، و عموزادی ریزی، کبری (۱۳۹۹). بررسی نقش تعدیل‌کننده نوع مالکیت دولتی بر رابطه بین سازوکارهای راهبری شرکت و مدیریت سود (واقعی و تعهدی) در شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران. *پژوهش‌های تجربی حسابداری*، ۱۰(۱)، ۱۱۵-۱۳۷. <https://doi.org/10.22051/jera.2019.19123.1942>
- قادر، اقبال، امینی، پیمان، محمدی ملقرنی، عطا، و نوروش، ایرج (۱۳۹۷). بررسی دقت شبکه‌های عصبی مصنوعی و الگوریتم بهینه‌سازی کلونی مورچگان در پیش‌بینی مدیریت سود. *فصلنامه حسابداری مالی*، ۱۰(۳۹)، ۸۲-۱۱۰. <http://qfaj.mobarakeh.iau.ir/article-1-1347-fa.html>
- قلی‌پور خانقاه، مهدی، جعفرپور، هومن، و صارمی‌نیا، مهسا (۱۴۰۰). ارتباط کیفیت حسابرسی و مدیریت سود در شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران. *دانش حسابداری و حسابرسی مدیریت*، ۱۰(۴۰)، ۳۷۹-۳۹۰. [https://www.jmaak.ir/article\\_18285.html?lang=fa](https://www.jmaak.ir/article_18285.html?lang=fa)

- کاردان، بیتا، قره‌خانی، بیتا، صالحی، مهدی، و منصوری، مرتضی (۱۳۹۶). بررسی دقت الگوریتم‌های خطی تکاملی BBO - و icde و الگوریتم‌های غیرخطی CVR و CART در پیش‌بینی سود. *پژوهش‌های حسابداری مالی*، ۹(۳۱)، ۷۷-۹۵.  
<https://doi.org/10.22108/far.2017.21776>
- کردستانی، غلامرضا و تاتلی، رشید. (۱۳۹۵). پیش‌بینی دستکاری سود: توسعه یک مدل. *بررسی‌های حسابداری و حسابرسی*، ۲۳ (۱)، ۷۳-۹۶.  
<https://doi.org/10.22059/acctgrev.2016.57021>
- گرد، عزیز، وقفی، سید حسام، حبیب‌زاده بایگی، سید جواد، و خواجه‌زاده، سارا (۱۳۹۴). مقایسه دقت پیش‌بینی مدیریت سود با استفاده از الگوریتم‌های مورچگان و غذایابی باکتری. *پژوهش‌های تجربی حسابداری*، ۵(۱)، ۱۸۱-۲۰۳.  
<https://doi.org/10.22051/jera.2015.2073>
- محمدیان حاجی کرد، امین، اصغرزاده زعفرانی، ملیحه، و امام دوست، مصطفی (۱۳۹۵). بررسی ریسک اعتباری مشتریان حقوقی با استفاده از مدل ماشین بردار پشتیبان و مدل هیبریدی الگوریتم ژنتیک - مطالعه موردی بانک تجارت. *مهندسی مالی و مدیریت اوراق بهادار*، ۲۷، ۱۷-۳۲.  
<https://sanad.iau.ir/Journal/fej/Article/1079259>
- میرغفوری، سیدحبیب اله، و امین آشوری، زهره (۱۳۹۴). ارزیابی ریسک اعتباری مشتریان بانک‌ها. *کاوش‌های مدیریت بازرگانی*، ۷(۱۳)، ۱۴۷-۱۶۶.  
[https://bar.yazd.ac.ir/article\\_655.html](https://bar.yazd.ac.ir/article_655.html)
- نمازی، محمد، و ابراهیمی، شهلا (۱۴۰۰). پیش‌بینی درماندگی مالی شرکت‌های پذیرفته‌شده در فرابورس و بورس اوراق بهادار تهران با استفاده از ماشین بردار پشتیبان. *راهبرد مدیریت مالی*، ۹(۱)، ۱۱۵-۱۳۲.  
<https://doi.org/10.22051/jfm.2020.25973.2077>
- نمازی، محمد، بایزیدی، انور، و جبارزاده کنگرلویی، سعید (۱۳۹۰). بررسی رابطه بین کیفیت حسابرسی و مدیریت سود شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران. *تحقیقات حسابداری و حسابرسی*، ۳(۹)، ۴-۲۱.  
<https://doi.org/10.22034/iaar.2011.104758>
- نوبخت، مریم، و برادران حسن‌زاده، رسول (۱۳۹۶). تأثیر جریان‌های نقد آزاد بر مدیریت سود واقعی و تصنعی. *بررسی‌های حسابداری و حسابرسی*، ۲۴(۳)، ۴۲۱-۴۴۰.  
<https://doi.org/10.22059/acctgrev.2017.224708>

## References

- Ahmad, J., Farman, H., & Jan, Z. (2019). Deep learning methods and applications. In *Deep Learning: Convergence to Big Data Analytics; Springer Briefs in Computer Science* (pp. 31-42). Springer, Singapore.  
[https://doi.org/10.1007/978-981-13-3459-7\\_3](https://doi.org/10.1007/978-981-13-3459-7_3)
- Ahmadpor, A., Esabat Tabari, E., & Talebtabar, M. (2015). Earnings Quality and CEO Tenure. *Empirical Research in Accounting*, 5(4), 1-14. [https://jera.alzahra.ac.ir/article\\_633.html#ar\\_info\\_pnl\\_cite](https://jera.alzahra.ac.ir/article_633.html#ar_info_pnl_cite) [In Persian]
- Ali, A., & Zhang, W. (2015). CEO tenure and earnings management. *Journal of Accounting and Economics*, 59(1), 60-79. <https://doi.org/10.1016/j.jacceco.2014.11.004>
- Amani, F. A., & Fadlalla, A. M. (2017). Data mining applications in accounting: A review of the literature and organizing framework. *International Journal of Accounting Information Systems*, 24, 32-58.  
<https://doi.org/10.1016/j.accinf.2016.12.004>
- Arasteh, Gh., & Nassirzadeh, F. (2017). Capability of Models Support Vector Regression, Least Angle Regression and Adaptive Neural Fuzzy Inference System for Earnings Management. *Empirical Research in Accounting*, 6(4), 95-116. [https://jera.alzahra.ac.ir/article\\_2632.html#ar\\_info\\_pnl\\_cite](https://jera.alzahra.ac.ir/article_2632.html#ar_info_pnl_cite) [In Persian]
- Arefmanesh, Z., & Amozadi Rizi, K. (2020). Moderating Role of Ownership Type (state) in the Relationship between Corporate Governance Mechanisms and Earning Management (real, accrual). *Empirical Research in Accounting*, 10(1), 115-137. <https://doi.org/10.22051/jera.2019.19123.1942> [In Persian]
- Arlot, S., & Celisse, A. (2010). A Survey of Cross-Validation Procedures for Model Selection. *Statistics Surveys*, 4, 40-79. <https://doi.org/10.1214/09-SS054>

- Azadi, F., Ghanbari, M., Jamshidi Navid, B., & Masodi, J. (2021). Presenting the developed model of Benish by using tunneling phenomena based on artificial neural network technique and particle swarm optimization algorithm to identifying profit manipulating companies. *Financial Management Perspective*, 11(33), 139-171. <https://doi.org/10.52547/jfmp.11.33.139> [In Persian]
- Bajra, U., & Cadez, S. (2018). The impact of corporate governance quality on earnings management: Evidence from European companies cross-listed in the US. *Australian Accounting Review*, 28(2), 152-166. <https://doi.org/10.1111/auar.12176>
- Beneish, M. D. (1999). The Detection of Earnings Manipulation. *Financial Analysts Journal*, 55 (5), 24-36. <https://doi.org/10.2469/faj.v55.n5.2296>
- Bouaziz, D., Salhi, B., & Jarboui, A. (2020). CEO characteristics and earnings management: Empirical evidence from France. *Journal of Financial Reporting and Accounting*, 18, 77-110. <https://doi.org/10.1108/JFRA-01-2019-0008>
- Braswell, M., & Daniels, R. B. (2017). Alternative earnings management techniques: What audit committees and internal auditors should know. *The Journal of Corporate Accounting and Finance*, 28(2), 45-54. <https://doi.org/10.1002/jcaf.22239>
- Chen, F. H., Chi, D. J., & Wang, Y. C., (2015). Detecting biotechnology industry's earnings management using Bayesian network principal component analysis back propagation neural network and decision tree. *Economic Modelling*, 46, 1-10. <https://doi.org/10.1016/j.econmod.2014.12.035>
- Chen, F. H., & Howard, H. (2016). An alternative model for the analysis of detecting electronic industries earnings management using stepwise regression, random forest, and decision tree. *Soft Computing*, 20, 1945-1960. <https://doi.org/10.1007/s00500-015-1616-6>
- Chen, S., & Shen, Z.-D. (2020). An effective enterprise earnings management detection model for capital market development. *Journal of Economics, Management and Trade*, 26(4), 77-91. <https://doi.org/10.9734/jemt/2020/v26i430250>
- Chen, Y.-J., Wu, C.-H., Chen, Y.-M., Li, H.-Y., & Chen, H.-K. (2017). Enhancement of fraud detection for narratives in annual reports. *International Journal of Accounting Information Systems*, 26, 32-45. <https://doi.org/10.1016/j.accinf.2017.06.004>
- Chye Koh, H., & Kee Low, C. (2004). Going concern prediction using data mining techniques. *Managerial Auditing Journal*, 19(3), 462-476. <https://doi.org/10.1108/02686900410524436>
- CPA Canada and AICPA. (2019). A CPAS Introduction To AI: From Algorithms to Deep Learning, What You Need To Know. Available at: [https://www.cpacanada.ca/en/business-andaccounting-resources/other-general-business-topics/information-management-and-technology/publications/a-cpaintedroduction-to-ai?utm\\_medium=stakeholder\\_outreach%20&utm\\_source=awc%20&utm\\_campaign=ai](https://www.cpacanada.ca/en/business-andaccounting-resources/other-general-business-topics/information-management-and-technology/publications/a-cpaintedroduction-to-ai?utm_medium=stakeholder_outreach%20&utm_source=awc%20&utm_campaign=ai)
- Cohen, D. A., Dey, A., & Lys, T. Z. (2008). Real and accrual based earnings management in the pre and post Sarbanes Oxley periods. *The Accounting Review*, 83(3), 757-787. <https://doi.org/10.2308/accr.2008.83.3.757>
- Cohen, D. A., & Zarowin, P. (2010). Accruals-based and real earnings management activities around seasoned equity offerings. *Journal of Accounting and Economics*, 50(1), 2-19. <https://doi.org/10.1016/j.jacceco.2010.01.002>
- Cover, T. M. (1986). Rates of convergence for nearest neighbor procedures. In *Proceedings of the Hawaii International Conference on System Sciences*. Univ. Hawaii Press, Honolulu, 413-415. <https://isl.stanford.edu/cover/papers/paper009>
- Daghani, R., Hajian, N., & Toloyee, K. (2020). Effects of the Board's Ownership and Oversight Structure and Audit Characteristics On Earning Management. *Empirical Research in Accounting*, 9(4), 299-326. <https://doi.org/10.22051/jera.2018.20138.2019> [In Persian]
- Danenas, P., & Garsva, G. (2012). Credit risk evaluation modeling using evolutionary linear SVM classifiers and sliding window approach. *Procedia Computer Science*, 9, 1324-1333. <https://doi.org/10.1016/j.procs.2012.04.145>
- Darabi, R., & Ajdari, F. (2018). Examination of the Relationship between Governance and Auditing Properties and Real Earnings Management of the Firms Listed in the Tehran Stock Exchange. *Journal of Accounting and Social Interests*, 8(2), 1-21. [https://jaacsi.alzahra.ac.ir/article\\_3605.html#ar\\_info\\_pnl\\_cite](https://jaacsi.alzahra.ac.ir/article_3605.html#ar_info_pnl_cite) [In Persian]

- Dash, R., & Dash, P. K. (2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. *The Journal of Finance and Data Science*, 2(1), 42–57. <https://doi.org/10.1016/j.jfds.2016.03.002>
- Dbouk, B., & Zaarour, I. (2017). Towards a machine learning approach for earnings manipulation detection. *Asian Journal of Business and Accounting*, 10(2), 215–251. <https://ajba.um.edu.my/index.php/AJBA/article/view/9772>
- Degeorge, F., Patel J., & Zeckhauser, R. (1999). Earning Management to Exceed Thresholds. *Journal of Business*, 72(1), 1–33. <https://doi.org/10.1086/209601>
- Demirkan, S. & Platt, H. (2009). Financial Status, Corporate Governance Quality, and the Likelihood of Managers Using Discretionary Accruals. *Accounting Research Journal*, 22 (2), 93–117. <https://doi.org/10.1108/10309610910987475>
- Devroye, L. (1981). On the equality of Cover and Hart in nearest neighbor discrimination, *IEEE Trans. Pattern Anal. Mach. Intell.*, 3, 75–78. <https://doi.org/10.1109/tpami.1981.4767052>
- Devroye, L., & Wagner, T. J. (1982). Nearest 8 neighbor methods in discrimination, In *Classification, Pattern Recognition and Reduction of Dimensionality*. In *Handbook of Statistics* (2; pp. 193–197). [https://doi.org/10.1016/S0169-7161\(82\)02011-2](https://doi.org/10.1016/S0169-7161(82)02011-2)
- Du, X., Jian, W., & Lai, S. (2017). Do foreign directors mitigate earnings management? Evidence from China. *The International Journal of Accounting*, 52(2), 142–177. <https://doi.org/10.1016/j.intacc.2017.04.002>
- Esmaili, M. (Trans.) (2014). *Data Mining: Concepts and Techniques* (First Edition). Tehran: Niaz Danesh Publications. [In Persian]
- Ezazi, M. E., Ghotbi, F. S., & Ghotbi, S. F. (2013). Predicting earning management using RBF, ICA, and SVM in firms listed in Tehran security exchange. *Asian Journal of Management Research*, 4(1), 208–220. <https://ajmjournal.com/Issues.aspx?VID=4&IID=1>
- Fallahpour, S., Lakvan, E. N., & Zadeh, M. H. (2017). Using an ensemble classifier based on sequential floating forward selection for financial distress prediction problem. *Journal of Retailing and Consumer Services*, 34, 159–167. <https://doi.org/10.1016/j.jretconser.2016.10.002>
- Fan, Y., Jiang, Y., Zhang, X., & Zhou, Y. (2019). Women on boards and bank earnings management: From zero to hero. *Journal of Banking & Finance*, 107. <https://doi.org/10.1016/j.jbankfin.2019.105607>
- Francis, B., Hasan, I., & Li, L. (2016). A cross-country study of legal-system strength and real earnings management. *Journal of Accounting and Public Policy*, 35(5), 477–512. <https://doi.org/10.1016/j.jaccpubpol.2016.06.004>
- Imandoust, S. B., & Bolandraftar, M. (2013). Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background. *Journal of Engineering Research and Applications*, 3(5), 605–610. <https://www.ijera.com/pages/v3-no5.html>
- Ge, W., & Kim, J.-B. (2014). Real earnings management and the cost of new corporate bonds. *Journal of Business Research*, 67(4), 641–647. <https://doi.org/10.1016/j.jbusres.2013.01.021>
- Geiger, B. C., & Kubin, G. (2012). Relative information loss in the PCA. In *2012 IEEE Information Theory Workshop* (pp. 562–566). <https://doi.org/10.48550/arXiv.1204.0429>
- Ghaderi, E., Amini, P., Mohammadi Mlqny, A., & Norvash, I. (2019). The Accuracy of Artificial Neural Network and Ant Colony Optimization algorithm in predicting profit management. *Quarterly Financial Accounting*, 10(39), 82–110. <http://qfaj.mobarakeh.iau.ir/article-1-1347-fa.html> [In Persian]
- Gholipour, M., Jafarpour, H., & Saremi Nia, M. (2022). The Relationship between Audit Quality and Earnings Management in Companies Accepted Tehran Stock Exchange. *Journal of Management Accounting and Auditing Knowledge*, 10(40), 379–390. [https://www.jmaak.ir/article\\_18285.html?lang=fa](https://www.jmaak.ir/article_18285.html?lang=fa) [In Persian]
- Gord, A., Vaghfi, H., Habibzade, J., & Khajehzadeh, S. (2015). Comparing the Accuracy of Earnings Management Forecast Using Ant Colony Optimization Algorithm and Bacteria Foraging Algorithm. *Empirical Research in Accounting*, 5(1), 181–203. <https://doi.org/10.22051/jera.2015.2073> [In Persian]
- Gunny, K. A. (2010). The relation between earnings management using real activities manipulation and future performance: Evidence from meeting earnings benchmarks. *Contemporary Accounting Research*, 27(3), 855–888. <https://doi.org/10.1111/j.1911-3846.2010.01029.x>
- Gupta, R., & Modise, M. P. (2012). South African stock return predictability in the context data mining: The role of financial variables and international stock returns. *Economic Modelling*, 29(3), 908–916. <https://doi.org/10.1016/j.econmod.2011.12.013>

- Haga, J., Siekkinen, J., & Sundvik, D. (2015). A neural network approach to measure real activities manipulation. *Expert Systems with Applications*, 42(5), 2313–2322. <https://doi.org/10.1016/j.eswa.2014.10.047>
- Hammami, A.; & Hendijani Zadeh, M. (2022). Predicting earnings management through machine learning ensemble classifiers. *Journal of Forecasting*, 41(8), 1639–1660. <https://doi.org/10.1002/for.2885>
- Healy, P. M., & Wahlen, J. M. (1999). A Review of the Earning Management Literature and Its Implications for Standard Setting. *Accounting Horizons*, 13(4), 365–383. <https://doi.org/10.2308/acch.1999.13.4.365>
- Hinton, G. E., Osindero, S. & Teh, Y.-W. (2006). A fast-learning algorithm for deep belief nets. *Neural Comput*, 18, 1527–1554. <https://doi.org/10.1162/neco.2006.18.7.1527>
- 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>
- Hosseini, S., & Rashidi, Z. (2013). Bankruptcy Prediction of Companies listed Corporations in Tehran Stock Exchange by Using Decision Tree and Logistic Regression. *Financial Accounting Research*, 5(3), 105–128. [https://far.ui.ac.ir/article\\_16987.html](https://far.ui.ac.ir/article_16987.html) [In Persian]
- Huang, W., Nakamori, Y., & Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, 32(10), 2513–2522. <https://doi.org/10.1016/j.cor.2004.03.016>
- Jan, C. L. (2018). An effective financial statements fraud detection model for the sustainable development of financial markets: Evidence from Taiwan. *Sustainability*, 10(2), 1–14. <https://doi.org/10.3390/su10020513>
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electron. Mark*, 31, 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- Jaspersen, J., Richter, A., & Zoller, S. (2021). *Predicting Earnings Management from Qualitative Disclosures*. Available at SSRN: <https://ssrn.com/abstract=3732203> or <http://dx.doi.org/10.2139/ssrn.3732203>
- Jiraporn, P., Miller, G. A., Yoon, S. S., & Kim, Y. S. (2008). Is earnings management opportunistic or beneficial? An agency theory perspective. *International Review of Financial Analysis*, 17(3), 622–634. <https://doi.org/10.1016/j.irfa.2006.10.005>
- Kardan, B., Salehi, M., Gharekhani, B., & Mansouri, M. (2017). The evaluation accuracy of BBO and ICDE as Linear- evolutionary Algorithms and SVR and CART as Non-linear Algorithms to earnings management prediction. *Financial Accounting Research*, 9(1), 77–96. <https://doi.org/10.22108/far.2017.21776> [In Persian]
- Karhunen, J., Raiko, T., & Cho, K. (2015). Unsupervised deep learning: A short review. In *Advances in Independent Component Analysis and Learning Machines* (pp. 125–142). Academic Press: Cambridge, MA, USA, <https://doi.org/10.1016/B978-0-12-802806-3.00007-5>
- Kira, K., & Rendell, L. A. (1992). The feature selection problem: Traditional methods and a new algorithm. In *Aaai* (Vol. 2; pp. 129–134).
- Kordestani, G., & Tatli, R. (2016). The Prediction of Earnings Manipulation: Development of a Model. *Accounting and Auditing Review*, 23(1), 73–96. <https://doi.org/10.22059/acctgrev.2016.57021> [In Persian]
- 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>
- Koutanaei, F. N., Sajedi, H., & Khanbabaei, M. (2015). A hybrid data mining model of feature selection algorithms and ensemble learning classifiers for credit scoring. *Journal of Retailing and Consumer Services*, 27, 11–23. <https://doi.org/10.1016/j.jretconser.2015.07.003>
- Lassoued, N., Attia, M. B. R., & Sassi, H. (2017). Earnings management and ownership structure in emerging market: Evidence from banking industry. *Managerial Finance*, 43(10), 1117–1136. <https://doi.org/10.1108/MF-11-2015-0312>
- Li, H., Li, C.-J., Wu, X.-J., & Sun, J. (2014). Statistics-based wrapper for feature selection: An implementation on financial distress identification with support vector machine. *Applied Soft Computing*, 19, 57–67. <https://doi.org/10.1016/j.asoc.2014.01.018>
- Liou, F.-M. (2008). Fraudulent financial reporting detection and business failure prediction models: A comparison. *Managerial Auditing Journal*, 23(7), 650–662. <https://doi.org/10.1108/02686900810890625>
- Lopez, D. M., & Vega, J. J. (2019). Evaluating the effect of industry specialist duration on earnings management. *Advances in Accounting*, 45, 100412. <https://doi.org/10.1016/j.adiac.2019.02.002>

- Matin, R., Hansen, C., Hansen, C., & Mølgaard, P. (2019). Predicting distresses using deep learning of text segments in annual reports, *Expert Systems with Applications*, 132, 199–208. <https://doi.org/10.1016/j.eswa.2019.04.071>
- McVay, S. E. (2006). Earnings management using classification shifting: An examination of core earnings and special items. *The Accounting Review*, 81(3), 501–531. <https://doi.org/10.2308/accr.2006.81.3.501>
- Milner, C., & Berg, B. (2017). *Tax Analytics Artificial Intelligence and Machine Learning–Level 5*. Available at: <https://www.pwc.no/no/publikasjoner/Digitalisering/artificial-intelligence-andmachine-learning-final1.pdf>. Accessed 19 May 2020
- Mirghafouri, S. H., & Amin, Z. (2015). Presenting a Model for Measuring Credit Risk of Bank Customers using Data Mining Approach. *Journal of Business Administration Researches*, 7(13), 247-266. [https://bar.yazd.ac.ir/article\\_655.html](https://bar.yazd.ac.ir/article_655.html) [In Persian]
- Mohammadian, H. K. A., Asgharzadeh Z. M., & Emam, D. M. (2016). Credit risk assessment of corporate customers using support vector machine and genetic algorithm hybrid model - A case study of Tejarat Bank. *Financial Engineering & Portfolio Management*, (7), 17-32. <https://sanad.iau.ir/Journal/fej/Article/1079259> [In Persian]
- Nair, B. B., Mohandas, V., & Sakthivel, N. (2010). A genetic algorithm optimized decision tree-SVM based stock market trend prediction system. *International Journal on Computer Science and Engineering*, 2(9), 2981–2988. <http://www.enggjournals.com/ijcse/issue.html?issue=20100209>
- Najari, M., Bishak, A. H., Rezaie Pazhand, P., & Habibzadeh Baygi, S. J. (2014). Forecasting of Earning Management by Support Vector Machine: Case Study in Tehran Exchange Stock. *Middle-East Journal of Scientific Research*, 19(7), 1007-1017. [https://idosi.org/mejsr/mejsr19\(7\)14.htm](https://idosi.org/mejsr/mejsr19(7)14.htm)
- Namazi, M., & Ebrahimi, S. (2021). Financial Distress Prediction of the Listed Companies on Tehran Stock Exchange (TSE) and Iran Fara Burse (IFB) Using Support Vector Machine. *Financial Management Strategy*, 9(1), 115-132. <https://doi.org/10.22051/jfm.2020.25973.2077> [In Persian]
- Namazi, M. , Bayazidi, A. & Jabarzadeh Kangarloie, S. (2011). Investigating the Relationship between Audit Quality and Earnings Management of Companies Listed in Tehran Stock Exchange. *Accounting and Auditing Research*, 3(9), 4-21. <https://doi.org/10.22034/iaar.2011.104758> [In Persian]
- Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3), 559–569. <https://doi.org/10.1016/j.dss.2010.08.006>
- Ninh, P. V. B., Do Thanh, T., & Hong, D. V. (2018). Financial distress and bankruptcy prediction: An appropriate model for listed firms in Vietnam. *Economic Systems*, 42(4), 616-624. <https://doi.org/10.1016/j.ecosys.2018.05.002>
- Nobakht, M., & Baradaran Hassanzadeh, R. (2016). The effect of free cash flows on real and artificial profit management. *Accounting and Auditing Reviews*, 24(3), 421-440. <https://doi.org/10.22059/acctgrev.2017.224708> [In Persian]
- Oreski, S., & Oreski, G. (2014). Genetic algorithm-based heuristic for feature selection in credit risk assessment. *Expert Systems with Applications*, 41(4), 2052–2064. <https://doi.org/10.1016/j.eswa.2013.09.004>
- Osta, S. (2011). Investigating the relationship between Ownership Structure and Earning Management. *Financial Accounting Research*, 3(2), 93-106. [https://far.ui.ac.ir/article\\_16915.html](https://far.ui.ac.ir/article_16915.html) [In Persian]
- Poor Ali, M., & Kochaki Tajani, M. (2020). *Comparing the accuracy of predicting the manipulation of corporate earnings using the colonial competition algorithm and the genetic algorithm*. 1st International Conference on Challenges and New Solutions in Industrial Engineering and Management and Accounting, Sari. <https://civilica.com/doc/1045469> [In Persian]
- Pyo, G., & Lee, H.-Y. (2013). The association between corporate social responsibility activities and earnings quality: Evidence from donations and voluntary issuance of CSR reports. *Journal of Applied Business Research*, 29(3), 945–962. <https://doi.org/10.19030/jabr.v29i3.7793>
- Robu, I. B., & Istrate, C. (2015). The analysis of the principal components of the financial reporting in the case of Romanian listed companies. *Procedia Economics and Finance*, 20, 553–561. [https://doi.org/10.1016/S22125671\(15\)00108-2](https://doi.org/10.1016/S22125671(15)00108-2)

- Rodriguez-Ariza, L., Martínez-Ferrero, J., & Bermejo-Sanchez, M. (2016). Consequences of earnings management for corporate reputation: Evidence from family firms. *Accounting Research Journal*, 29(4), 457–474. <https://doi.org/10.1108/ARJ-02-2015-0017>
- Roweis, S. T. (1998). EM algorithms for PCA and SPCA. In *Advances in neural information processing systems* (pp. 626–632).
- Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of Accounting and Economics*, 42(3), 335–370. <https://doi.org/10.1016/j.jacceco.2006.01.002>
- Saeidmoghadam, M., Javid, D., & Hematfar, M. (2018). Detecting automotive and parts manufacturing industry earnings management by combining Bayesian networks and C5.0 decision tree. *Quarterly Financial Accounting*, 9(36), 102-126. <http://qfaj.mobarakkeh.iau.ir/article-1-1080-fa.html> [In Persian]
- Salehi, M., & Farrokhi Pilehroud, L. (2018). Predicting profit management using neural network and decision tree. *Quarterly Journal of Financial Accounting and Auditing Research*, 10(37), 1-24. <https://sanad.iau.ir/Journal/faar/Article/1073544> [In Persian]
- Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. *Neural Networks*, 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- Schipper, K., (1989). Commentary on earnings management. *Accounting Horizons*, 3(4), 91-102. <https://publications.aaahq.org/accounting-horizons/issue/3/4>
- Setayesh, M. H., & Kazemnezhad, M. (2019). The Usefulness of Variables (Dimension) Reduction Methods in Stock Returns of the Companies Listed on Tehran Stock Exchange. *Empirical Studies in Financial Accounting*, 16(63), 83-107. <https://doi.org/10.22054/qjma.2019.10647> [In Persian]
- Shahmoradi, N., & Tabatabaieinasab, Z. (2021). The effect of Audit Quality on the Relationship between Economic Uncertainty and Accrual Based Earnings Management in Listed Companies in Tehran Stock Exchange. *Financial Accounting Research*, 13(1), 67-86. <https://doi.org/10.22108/far.2021.125552.1683> [In Persian]
- Shen, C. H., & Chih, H. L. (2007). Earnings management and corporate governance in Asia's emerging markets. *Corporate Governance: An International Review*, 15(5), 999–1021. <https://doi.org/10.1111/j.1467-8683.2007.00624.x>
- Sun, N., Salama, A., Hussainey, K., & Habbash, M. (2010). Corporate environmental disclosure, corporate governance and earnings management. *Managerial Auditing Journal*, 25(7), 679-700. <https://doi.org/10.1108/02686901011061351>
- Tay, F. E., & Cao, L. (2001). Application of support vector machines in financial time series forecasting. *Omega*, 29(4), 309–317. [https://doi.org/10.1016/S0305-0483\(01\)00026-3](https://doi.org/10.1016/S0305-0483(01)00026-3)
- Tian, S., Yu, Y., & Guo, H. (2015). Variable selection and corporate bankruptcy forecasts. *Journal of Banking & Finance*, 52, 89–100. <https://doi.org/10.1016/j.jbankfin.2014.12.003>
- Tsai, C. F. (2009). Feature Selection in Bankruptcy Prediction. *Knowledge-Based Systems*, 22(2), 120–127. <https://doi.org/10.1016/j.knosys.2008.08.002>
- Tsai, C. F., & Chiou, Y. J. (2009). Earnings Management Prediction: A Pilot Study of Combining Neural Networks and Decision Trees. *Expert Systems with Applications*, 36(3), 7183–7191. <https://doi.org/10.1016/j.eswa.2008.09.025>
- Tort, L. P. (2013). Earnings management under IFRS and PGC. *Revista de Contabilidad Y Direccion*, 16(1), 161–185. <https://accid.org/es/revista-de-contabilidad-y-direccion-num-16/>
- Veganzones, D., & Séverin, E. (2018). An investigation of bankruptcy prediction in imbalanced datasets. *Decision Support Systems*, 112, 111–124. <https://doi.org/10.1016/j.dss.2018.06.011>
- Viaene, S., Dedene, G., & Derrig, R. A. (2005). Auto claim fraud detection using Bayesian learning neural networks. *Expert Systems with Applications*, 29(3), 653-666. <https://doi.org/10.1016/j.eswa.2005.04.030>
- Wei, Z., Wang, Y., He, S., & Bao, J. (2017). A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection. *Knowledge-Based Systems*, 116, 1–12. <https://doi.org/10.1016/j.knosys.2016.10.022>
- Yu, Q., Miche, Y., Séverin, E., & Lendasse, A. (2014). Bankruptcy prediction using extreme learning machine and financial expertise. *Neurocomputing*, 128, 296–302. <https://doi.org/10.1016/j.neucom.2013.01.063>
- Zang, A. Y. (2012). Evidence on the trade-off between real activities manipulation and accrual-based earnings management. *The Accounting Review*, 87(2), 675–703. <https://doi.org/10.2308/accr-10196>