فصلنامه حسابداری مالی/سال نهم/ شماره ۳۶ /زمستان ۹۶/صفحات ۱۰۰- ۷۶

پیش‌بینی فرار مالیاتی با استفاده از الگوریتم‌های داده‌کاوی درخت تصمیم

محمد نمازی*
محمد صادق زاده مهارلوئی**

تاریخ دریافت: ۱۳۹۶/۰۹/۲۱ تاریخ پذیرش: ۱۳۹۷/۰۲/۳۱

چکیده

پژوهش حاضر به بررسی قابلیت پیش‌بینی فرار مالیاتی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از الگوهای درخت تصمیم پرداخته است. جامعه آماری پژوهش حاضر کلیه شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران در بازه زمانی ۱۳۸۴ تا ۱۳۹۴ است و نمونه پژوهش برابر با ۱۰۸۱ سال-شرکت است. برای تحلیل داده‌ها از روش‌های آماری تحلیل واریانس یک‌طرفه و الگوریتم‌های داده‌کاوی درخت تصمیم استفاده شد. در این راستا، داده‌های پژوهش با استفاده از نرم‌افزارهای SPSS و Weka مورد تجزیه‌وتحلیل آماری قرار گرفتند. نتایج حاصل از بررسی‌ها نشان می‌دهد که به ترتیب، روش‌های جنگل تصادفی، کاهش خطای هرس، J48، LMT، ریشه تصمیم و درخت تصادفی از دقت و کارایی بیشتری در پیش‌بینی فرار مالیاتی برخوردار هستند. هم‌چنین، نتایج تحلیل واریانس یک‌طرفه نشان داد که تفاوت در کارایی پیش‌بینی‌های روش‌های مختلف درخت تصمیم از لحاظ آماری نیز معنادار است.

واژه‌های کلیدی: پیش‌بینی فرار مالیاتی، درخت تصمیم، الگوریتم درخت تصادفی، الگوریتم جنگل تصادفی.

*استاد حسابداری، دانشگاه شیراز، شیراز، ایران (نویسنده مسئول)
Email: mnamazi@rose.shirazu.ac.ir
**دانشجوی دکترای حسابداری، دانشگاه شیراز، شیراز، ایران
Email: sadeghzadeh.mohammad@gmail.com

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

۱- مقدمه

یکی از منابع لازم و اساسی برای یک دولت به منظور رسیدن به اهداف خاص آن، درآمدهای مالیاتی است. با وجود این، برخی از شرکت‌ها از پرداخت مالیات به میزان واقعی، چه از راه‌های قانونی و چه از راه‌های غیرقانونی، خودداری می‌کنند که این امر تأثیر اساسی بر روی درآمدهای مالیاتی دولت خواهد داشت (وو، او، لین، چانگ و ین[1]، ۲۰۱۲). در این خصوص دولت و مقامات مالیاتی، مجبور به تحمل هزینه‌های کشف و جلوگیری از فرار مالیاتی[2] می‌شوند. به این دلیل همیشه پیدا کردن راه‌ها و شیوه‌هایی که بتوانند فرار مالیاتی را شناسایی و کشف کنند، مسئله‌ای مهم و چالش‌برانگیز بوده است. اگر دولت نتواند فرار مالیاتی را کشف کند، سرمایه‌گذاری‌های عمومی به دلیل کاهش مالیات‌ها و در پی آن کاهش درآمدهای دولت به‌طور منفی تحت تأثیر قرار خواهد گرفت (وو و همکاران، ۲۰۱۲). آمار و تصاویر ایجاد شده در یک نظام مالیاتی کارآمد و عادلانه می‌تواند خود یک کنترل‌کننده نرم‌افزاری و علمی جهت ارتقای سطح کارآمدی اقتصادی و شفافیت اطلاعات منتشر شده در حوزه‌های مختلف اقتصادی هر کشور باشد (عبدلی، شوروزی و میرزائی، ۱۳۹۲). افزون بر این، در بسیاری از کشورهای در حال توسعه از جمله ایران، به دلیل نبود اطلاعات کافی و عدم کارایی نظام اجرایی وصول مالیات، بسیاری از مشاغل پردرآمد از پرداخت مالیات متناسب با درآمدشان می‌گریزند و بخش قابل‌توجهی از درآمد دولت به این ترتیب از بین می‌رود (شیرخانی، ۱۳۸۵). این رفتار مؤدیان مالیاتی در پرداخت نکردن مالیات را می‌توان به دو طبقه اجتناب مالیاتی[3] و فرار مالیاتی تقسیم کرد.

اجتناب مالیاتی به معنای استفاده از نقاط ضعف قانونی و استفاده از آن‌ها به‌منظور پرداخت مالیات کمتر است؛ درصورتی‌که در فرار مالیاتی، قوانین نقض شده و شرکت به دنبال کاهش درآمد مشمول مالیات است (استانکویشس و لئوناس[4]، ۲۰۱۵). فرار مالیاتی امروزه به عنوان یک مشکل جهانی که هزینه‌های اقتصادی زیادی را بر همه جوامع تحمیل می‌کند در نظر گرفته می‌شود که در این میان می‌توان به کاهش رشد اقتصادی به دلیل کاهش توانایی دولت برای فراهم کردن کالاهای عمومی، مؤسسات پشتیبان بازار، زیرساخت‌ها، توسعه سرمایه انسانی و تحقیق و توسعه اشاره کرد (جانسون، کافمن، مک‌میلان و وودرف[5]، ۲۰۰۰). افزون بر این، فرار مالیاتی، منابع را به سمت فعالیت‌های غیر تولیدی از جمله ایجاد شعبه‌های مالی برای پوشش

1 Wu, Ou, Lin, Chang and Yen
2 Tax Evasion
3 Tax Avoidance
4 Stankevicius and Leonas
5 Johnson, Kaufmann, McMillan and Woodruff

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

فرار مالیاتی هدایت می‌کند. همچنین، می‌توان فرار مالیاتی را باعث ایجاد عدم کارایی در تولیدات شرکت دانست؛ چرا که شرکت‌های دارای فرار مالیاتی تمایل به کوچک باقی ماندن دارند تا راحت‌تر بتوانند اقدام به فرار مالیاتی کنند. در نهایت، فرار مالیاتی باعث عدم برابری میان شرکتهای دارای فرار مالیاتی و شرکت‌های فاقد فرار مالیاتی می‌شود؛ این موضوع، انگیزه فرار مالیاتی را افزایش می‌دهد. در این میان، به دلیل هزینه‌های مرتبط با حسابرسی مالیاتی هر شرکت به صورت انفرادی، پیدا کردن الگوهایی که به گونه دقیق مالیات‌های ابرازی اشتباه را مشخص کنند از اهمیت بسیاری برخوردار است.

از سوی دیگر، تاکنون پژوهش های اندکی در زمینه فرار مالیاتی با توجه به اطلاعات حسابداری انجام شده است و اغلب پژوهش‌های انجام شده در این زمینه (به عنوان نمونه سید زارع، ۱۳۸۰؛ فلاحتی، نظیفی و عباسپور، ۱۳۹۱؛ هادیان و تحویلی، ۱۳۹۲) با توجه به رویکردهای اقتصادی به پیش‌بینی فرار مالیاتی پرداخته‌اند. با وجود این، حتی در پژوهش‌هایی که از منظر ارقام اقتصادی به مسئله فرار مالیاتی پرداخته‌اند، جای خالی استفاده از رویکردهای نوینی چون داده‌کاوی[1] مشهود است. این مسئله بدان علت مهم است که در پژوهش‌های بسیاری سودمندی، کارایی و عملکرد بهتر تکنیک‌های داده‌کاوی در مقابل روش‌های خطی به وضوح مورد تائید قرار گرفته است (به‌عنوان نمونه، مور، اشنایدر و دنگ[2]، ۱۹۹۷؛ مین و لی[3]، ۲۰۰۴؛ موکامالا، سانگ، ریبیرو و وی‌یرا[4]، ۲۰۰۶؛ آلفارو، گارسیا، گامز و الیزوندو[5]، ۲۰۰۸؛ لی و تاو[6]، ۲۰۱۰). این پژوهش به دنبال پاسخگویی به این پرسش است که آیا می‌توان با استفاده از روش‌های داده‌کاوی به پیش‌بینی فرار مالیاتی در بورس اوراق بهادار تهران پرداخت و آیا میان الگوریتم های مختلف این روش ها تفاوتی در کارایی وجود دارد یا خیر؟ بنابراین، با توجه به اهمیت پیش‌بینی فرار مالیاتی در ایران، این پژوهش به مطالعه پیش‌بینی فرار مالیاتی با استفاده از اطلاعات حسابداری از طریق الگوهای داده‌کاوی درخت تصمیم می‌پردازد.

ساختار مقاله حاضر بدین صورت است که در بخش اول به مبانی نظری فرار مالیاتی و داده‌کاوی پرداخته می‌شود. در بخش دوم، پیشینه پژوهش مورد بررسی قرار می‌گیرد. در ادامه، به توصیف روش پژوهش و متغیرهای پژوهش پرداخته می‌شود. سپس، روش‌های تجزیه‌وتحلیل

1 Data Mining
2 Moore, Schneider and Deng
3 Min and Lee
4 Mukkamala, Sung, Ribeiro, and Vieira
5 Alfaro, García, Gámez and Elizondo
6 Lee and To

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

داده‌ها و یافته‌های پژوهش مطرح می‌شود. سرانجام، نتیجه‌گیری، پیشنهادها و محدودیت‌های پژوهش بیان می‌گردد.

۲- مبانی نظری پژوهش

۲-۱- فرار مالیاتی

نخستین تجزیه‌وتحلیل فرار از مالیات توسط آلینگهام و ساندمو[1] (۱۹۷۲) و کلم[2] (۱۹۷۳) انجام شد (به نقل از نویدی قاضیانی، طیب‌نیا و اسکندری دورباطی، ۱۳۸۹). قسمت اعظم الگوهای ساخته‌شده در این زمینه مربوط به فرار از مالیات بر درآمد است. از نظر کیم[3] (۲۰۰۸) مطالعات و بررسی‌های انجام‌شده بر روی فرار مالیاتی و تمکین را می‌توان به چهار گروه وسیع زیر تقسیم کرد: بررسی نظریه‌های تمکین، تخمین تجربی نظریه‌های تمکین، تحلیل روندهای مختلف حسابرسی مالیاتی و تخمین تجربی فرار مالیاتی (نویدی قاضیانی و همکاران، ۱۳۸۹).

اجتناب و فرار مالیاتی، هر دو باعث کاهش درآمد مالیاتی می‌شوند، اما از نظر ماهیت با یکدیگر متفاوت‌اند. اجتناب مالیاتی به معنای استفاده از نقاط ضعف قانونی به منظور پرداخت مالیات کمتر است. درصورتی‌که در فرار مالیاتی، قوانین نقض شده و شرکت به دنبال کاهش درآمد مشمول مالیات است؛ این امر می‌تواند از راه‌های زیر انجام پذیرد (ساندمو[4]، ۲۰۰۵):

۱) عدم بیان درآمد مشمول مالیات در اظهارنامه.
۲) گزارش هزینه‌های قابل‌قبول مالیاتی به میزان بیش از واقع.
۳) خودداری از تسلیم اظهارنامه.
۴) کاستن از رقم بدهی مالیاتی با توسل به اظهارات خلاف واقع.

به گونه کلی، فرار مالیاتی به عدم پرداخت غیرقانونی تمام یا قسمتی از مالیات اطلاق می‌شود (موسوی جهرمی، طهماسبی بلداجی و خاکی، ۱۳۸۸). این اقدام می‌تواند در قالب تنظیم حساب‌های مالی، ارائه اطلاعات مالی نادرست و یا ارائه نکردن اطلاعات صورت گیرد. افزون بر این، فرار مالیاتی به گستره اقتصاد غیررسمی مانند معاملات خانگی و فعالیت‌های غیرقانونی نظیر فعالیت‌های جنایی و بزهکاری مربوط می‌شود که از این فعالیت‌ها به دلیل فقدان اطلاعات لازم، مالیاتی اخذ نمی‌گردد (موسوی جهرمی و همکاران، ۱۳۸۸).

1 Allingham and Sandemo
2 Klem
3 Kim
4 Sandmo

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

فرار مالیاتی را می‌توان یکی از مظاهر فساد اقتصادی دانست. از این رو، تمامی کشورها تلاش خود را برای کاهش این پدیده به کار می‌گیرند و در این راستا همواره به اصلاح نظام مالیاتی خود می‌پردازند. فرار از مالیات چه به صورت مستقیم و چه به صورت غیرمستقیم بر فعالیت‌های اقتصادی از جمله سرمایه‌گذاری مؤثر بوده و می‌تواند در بلندمدت مسیر رشد و توسعه اقتصادی را دچار اختلال کند (شهریاری‌فر، ۱۳۹۰).

فرار مالیاتی به معنای شکستن حریم قانون است. وقتی مؤدی از گزارش درآمدهای مشمول مالیات خویش عدول کند یا به شیوه‌های عامدانه‌ای نظیر گزارش کمتر از واقع درآمد و فروش، گزارش بیش از واقع کسورات و یا تنظیم اظهارنامه غیرواقعی از تعهدات مالیاتی خود طفره رود، درحقیقت مرتکب عمل غیرقانونی شده است که مستوجب تنبیه مدنی یا جزایی است و از این رو، کسی که از پرداخت مالیات فرار می‌کند نگران کشف اعمال خلاف قانون خویش است (عبداله‌میلانی و اکبرپور روشن، ۱۳۹۱).

در واقع عوامل مؤثر رفتار فرار مالیاتی در همه مالیات‌ها وجود دارد و محدود به مالیات خاصی نیست و با وجود این عوامل، فرار مالیاتی نیز صورت می‌گیرد. برای حداقل کردن فرصت‌های فرار مالیاتی نیاز به یک نظام مالیاتی قوی است که در آن برای کسانی که از پرداخت مالیات طفره می‌روند جرائمی در نظر گرفته شده باشد (ضیایی بیگدلی و طهماسبی بلداجی، ۱۳۸۴).

۲-۲- داده‌کاوی

بسیاری از افراد، داده‌کاوی را مترادف واژه‌های رایج کشف دانش از داده‌ها می‌دانند، درحالی‌که برخی دیگر، داده‌کاوی را یکی از گام‌های اساسی در فرآیند کشف دانش تلقی می‌کنند. فرآیند کشف دانش را می‌توان پیمودن گام‌های زیر دانست (اسماعیلی، ۱۳۹۳):

۱. **پالایش داده‌ها:** این مرحله برای حذف داده‌های با پارازیت و ناسازگار استفاده می‌شود.

۲. **یکپارچه‌سازی داده‌ها:** در این مرحله ممکن است چندین منبع داده‌ای ترکیب شوند.

۳. **انتخاب داده‌ها:** در این مرحله داده‌های مرتبط با تحلیل از پایگاه‌های داده بازیابی می‌شوند.

۴. **تبدیل داده‌ها:** داده‌ها با کمک عملیاتی مانند تلخیص و تجمیع به شکل‌های مناسبی تبدیل می‌شوند.

۵. **داده‌کاوی:** یک فرآیند ضروری که با اعمال روش‌های هوشمند به استخراج الگوها در میان داده‌ها می‌پردازد.

۶. **ارزیابی الگو:** جهت معرفی الگوهای جالب در میان حجم انبوهی از الگوها که بر اساس مجموعه‌ای از سنجه‌ها انجام می‌شود.

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

۷. ارائه دانش: گامی که در آن با کمک تکنیک‌های دیداری متفاوت، دانش کسب‌شده به کاربر ارائه می‌شود.

گام‌های اول تا چهارم شکل‌های متفاوتی از پیش‌پردازش داده‌ها محسوب می‌شوند. درواقع در این گام‌ها، داده‌ها جهت داده‌کاوی آماده می‌شوند. گام‌های داده‌کاوی نیز ممکن است با کمک کاربر یا پایگاه دانش انجام شود. الگوهای جالب به کاربر ارائه می‌شوند و شاید به‌عنوان دانش جدید در پایگاه دانش نگهداری شوند. در این دیدگاه داده‌کاوی به‌عنوان یک گام در فرآیند کشف دانش شناخته می‌شود، گامی اساسی که در آن الگوهای نهان، آشکار و ارزیابی می‌شوند (اسماعیلی، ۱۳۹۳).

سه ویژگی مهم که داده‌کاوی را از دیگر فن‌های آماری متمـایز می‌کند عبارت‌اند از:

۱- داده‌کاوی قادر به طبقه‌بندی میزان زیادی از داده‌ها برای تجزیه‌وتحلیل ثانویه است.

۲- داده‌کاوی قادر به خلاصه کردن اطلاعات از داده‌های سیستم‌های عملیاتی است و بر پایه آزمایش‌های قبلی نیست.

۳- داده‌کاوی می‌تواند الگوها و روابط بین داده‌ها را کشف کند. هم‌چنین، بیشترین فن‌های مورداستفاده در داده‌کاوی عبارت‌اند از درخت تصمیم، شبکه‌های عصبی و یادگیری استقرایی، شبکه‌های بیزین و قوانین رابطه‌ای (هان و فو[1]، ۱۹۹۹).

۳- پیشینه پژوهش

۳-۱- پیشینه داخلی

اخیراً تقوی‌فرد، رئیسی وانانی و پناهی (۱۳۹۶) به بررسی تحلیل آینده‌نگر تشخیص فرار مالیاتی مودیان مالیات بر ارزش افزوده با استفاده از الگوریتم‌های طبقه‌بندی و خوشه‌بندی پرداخته‌اند. دوره زمانی پژوهش آنان سال‌های ۱۳۸۸ تا ۱۳۹۳ و الگوریتم‌های مورد بررسی شامل سه روش Naive Bayes، درخت تصمیم و KNN بود. همچنین الگوریتم‌های خوشه‌بندی شامل دو روش k-means و k-medoids بود. نتایج پژوهش نشان داد که از میان الگوریتم‌های طبقه‌بندی، روش درخت تصمیم با خطای ۰/۰۵ و خطای آموزش (اعتبارسنجی) ۰/۱۷ نتیجه بهتری کسب کرد.

لاری دشت بیاض، قائم‌مقامی و کهرمی (۱۳۹۵) به بررسی عوامل مؤثر بر فرار مالیاتی در استان خراسان جنوبی با تأکید بر مؤلفه‌های فرهنگی پرداختند. نتایج این پژوهش نشان داد که بین عوامل فرهنگی، اقتصادی و مؤلفه‌های سیاسی- قانونی با فرار مالیاتی رابطه مثبت و معناداری وجود دارد.

1 Han and Fu

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

سامعی راد و شاه‌بهرامی (۱۳۹۵) با استفاده از الگوهای پردازش موازی به بهبود کارایی الگوریتم‌های تشخیص تقلب مالیاتی پرداختند. نتایج پژوهش آنان نشان داد که با استفاده از الگوهای پردازش موازی می‌توان کارایی برنامه‌های کشف تقلب‌های مالیاتی را به‌طور قابل‌ملاحظه‌ای بهبود بخشید.

رحیمی‌کیا، محمدی و غضنفری (۱۳۹۴) به بررسی و تشخیص فرار مالیاتی با استفاده از سیستم هوشمند ترکیبی در صنایع مواد غذایی و نساجی پرداختند. متغیرهای ورودی شامل ۲۱ متغیر مالی بود که با توجه به پژوهش‌های پیشین و نظرات خبرگان انتخاب شدند. دوره زمانی پژوهش آنان سال‌های ۱۳۸۹ تا ۱۳۹۱ و روش داده‌کاوی آن‌ها شبکه‌های عصبی مصنوعی بود که در آن از الگوریتم جستجوی هارمونی به‌منظور بهینه‌سازی هم‌زمان پارامترهای شبکه عصبی مصنوعی استفاده شد. افزون بر این، نتایج حاصل از مقایسه شبکه عصبی و رگرسیون لجستیک نشان داد که استفاده از شبکه عصبی دارای دقت بالاتری بوده و این تفاوت از لحاظ آماری معنادار است.

دستگیر و غریبی (۱۳۹۴) به بررسی و استفاده از روش‌های داده‌کاوی برای ارتقای عملکرد تشخیص فرار مالیاتی در بورس اوراق بهادار تهران پرداختند. نمونه مورد بررسی این پژوهش شامل ۱۲۵ شرکت در دوره زمانی ۱۳۸۳ تا ۱۳۹۰ و روش مورد استفاده آن‌ها قواعد وابستگی و به‌کارگیری الگوریتم پیشینار بود. نتایج پژوهش آنان نشان داد که روش داده‌کاوی مبتنی بر قواعد وابستگی با ایجاد دو مدل با درصد صحت ۹۱٪ بر روی داده‌های آموزش، با صحت ۸۸٪ بر روی داده‌های اعتبارسنجی و با درصد صحت ۸۶٪ بر روی داده‌های آزمون توانسته است موفق به تشخیص فرار مالیاتی شود.

باقرپور ولاشانی، باقری، خادم و حسینی پور (۱۳۹۱) به بررسی عوامل مالی و غیرمالی مؤثر بر فرار مالیاتی با استفاده از فن‌های داده‌کاوی در صنعت خودرو و ساخت قطعات پرداختند. در این پژوهش تأثیر ۲۳ متغیر مالی و غیرمالی بر گریز مالیاتی در شرکت‌های فعال در صنعت خودرو و ساخت قطعات در بازه زمانی ۱۳۸۴ تا ۱۳۸۸ مورد بررسی قرار گرفت. نتایج این مطالعه نشان داد که متغیرهای نسبت دارایی‌ها به درآمد خالص، قدر مطلق نسبت هزینه بهره به درآمد خالص و استقلال هیئت‌مدیره رابطه مستقیم و متغیرهای نسبت سود و زیان خالص به دارایی‌ها و سود یا زیان رابطه عکس با احتمال فرار مالیاتی دارند.

باقری و عموری (۱۳۸۸) به شناسایی و اولویت‌بندی نسبت‌های مالی اثرگذار بر درآمد مشمول مالیات پرداختند. آن‌ها با استفاده از پرسش‌نامه و تجزیه‌وتحلیل داده‌های آن بر اساس تکنیک

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

تحلیل سلسله مراتبی به این نتیجه دست یافتند که سود خالص، سود ناخالص، سود هر سهم، بازده سهام و بازده سرمایه در گردش دارای اهمیت بیشتری بر درآمد مشمول مالیات می‌باشند.

۳-۲- پیشینه خارجی

دیدیمو، گیامینونی، لیوتا، مونتچیانی و پاگلیوکا[1] (۲۰۱۸) به توصیف سیستم پشتیبانی تصمیم برای کشف فرار مالیاتی بر مبنای زبان دیداری و تکنیک‌های پیشرفته مصورسازی شبکه پرداختند. این سیستم به استفاده‌کنندگان از آن، این امکان را می‌دهد که گراف‌های زیرمجموعه نمونه‌های مشکوک را با توجه به تطابق با الگوهای موجود رسم کرده و باعث ادغام نتایج و ایجاد شاخص‌هایی برای طبقه‌بندی کردن مالیات‌دهندگان بر مبنای ریسک مالی شوند. نتایج نشان داد که سیستم پیشنهادی مؤثر است.

رحیمی‌کیا، محمدی، رحمانی و غضنفری (۲۰۱۷) به کشف فرار مالیاتی با استفاده از سیستم‌های هوشمند ترکیبی در صنایع مواد غذایی و نساجی بورس اوراق بهادار تهران پرداختند. دوره زمانی پژوهش آنان سال‌های ۱۳۸۹ تا ۱۳۹۱ بود. روش داده‌کاوی آن‌ها شبکه‌های عصبی مصنوعی و ماشین‌های بردار پشتیبان بود که در آن از الگوریتم جستجوی هارمونی به‌منظور بهینه‌سازی هم‌زمان پارامترهای شبکه عصبی مصنوعی استفاده شد. افزون بر این، نتایج حاصل از مقایسه شبکه عصبی و رگرسیون لجستیک نشان داد که استفاده از شبکه عصبی دارای دقت بالاتری بوده و این تفاوت از لحاظ آماری معنادار است.

وو و همکاران (۲۰۱۲) نیز با استفاده از داده‌کاوی درصدد افزایش عملکرد کشف فرار مالیاتی برآمدند. آن‌ها با استفاده از داده‌کاوی یک چارچوب نظارتی ارائه کردند تا مغایرت‌های بین گزارش‌های ارزش‌افزوده را که نیازمند حسابرسی بیشتری بود، مشخص نمایند. نتایج این پژوهش نشان داد که تکنیک داده‌کاوی مورد استفاده (قوانین انجمنی) به‌درستی توانست میزان تشخیص فرار مالیاتی را افزایش دهد و درنتیجه با به‌کارگیری مؤثر آن می‌توان زیان‌های ناشی از فرار مالیاتی را کاهش داد.

۳-۳- نقد پژوهش‌های پیشین

بررسی پژوهش‌های انجام‌شده در داخل و خارج از کشور، نشان می‌دهد که در اکثر این مطالعات فرار مالیاتی با استفاده از چند متغیر محدود اقتصادی انجام شده و بیشتر پژوهش‌های صورت گرفته در زمینه فرار مالیاتی در سطح کلان بوده و از متغیرهای اقتصادی استفاده کرده‌اند و کمتر به بررسی اطلاعات حسابداری شامل متغیرهای مالی، حاکمیت شرکتی و استفاده از گزارش

1 Didimo, Giamminonni, Liotta, Montecchiani and Pagliuca

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

حسابرس در این باره پرداخته‌اند. سرانجام، در معدود پژوهش‌های داخلی انجام‌شده در زمینه فرار مالیاتی با استفاده از نسبت‌های مالی، متغیرهای انتخاب‌شده همان متغیرهایی است که پژوهشگران خارجی برای گزارشگری مالی متقلبانه استفاده کرده‌اند. در حالی که با اجرای طرح جامع مالیاتی و استفاده از نمونه‌گیری در بررسی اظهارنامه‌های شرکت‌ها، اهمیت انتخاب متغیرهای مربوط و پیش‌بین دوچندان می‌شود. به گونه کلی، این مطالعه درصدد رفع کاستی‌های مورد اشاره و گسترش مدارک تجربی در زمینه فرار مالیاتی صورت پذیرفته است.

۴- روش پژوهش

این پژوهش کاربردی است. طرح پژوهش آن از نوع شبه تجربی و با استفاده از رویکرد پس‌رویدادی (از طریق اطلاعات گذشته) است. از روش پس‌رویدادی زمانی استفاده می‌شود که پژوهشگر پس از وقوع رویدادها به بررسی موضوع می‌پردازد. افزون بر این، امکان دستکاری متغیرهای مستقل وجود ندارد (نمازی، ۱۳۸۹). جامعه آماری این پژوهش، شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران است که دارای شرایط زیر است:

- در بازه ۱۳۸۴-۱۳۹۴ در بورس فعالیت داشته و اطلاعات مربوط به صورت‌های مالی و یادداشت‌های توضیحی آن‌ها موجود باشد.
- فعالیت‌های از نوع سرمایه‌گذاری نداشته باشند و در زمینه استخراج نفت و نیز کشاورزی (به دلیل ماده ۱۴۱ قانون مالیات‌های مستقیم) فعالیت نکنند.
- پایان سال مالی آن‌ها منتهی به پایان اسفند باشد و طی سال‌های ۱۳۸۴ تا ۱۳۹۴، تغییر سال مالی نداده باشد.

در این پژوهش برای جمع‌آوری داده‌ها و اطلاعات از روش‌های کتابخانه‌ای استفاده می‌شود. مبانی نظری پژوهش از کتب، مجلات و سایت‌های تخصصی فارسی و لاتین و داده‌های مالی مورد نیاز با مراجعه به سایت سازمان بورس اوراق بهادار تهران، صورت‌های مالی شرکت‌ها و هم‌چنین با استفاده از نرم‌افزار ره‌آورد نوین گردآوری شد.

۵- متغیرهای پژوهش

متغیرهای استفاده شده در این پژوهش برای پیش‌بینی فرار مالیاتی را می‌توان به سه دسته متغیرهای مربوط به اعداد صورت‌های مالی، متغیرهای نسبت‌های مالی و متغیر مربوط به بند مالیاتی در گزارش حسابرسی تقسیم کرد. در این خصوص از بین کل متغیرهای شناسایی شده، متغیرهایی که بیشتر در ادبیات مربوط به پیش‌بینی فرار مالیاتی استفاده شده و داده‌های مورد نیاز برای سنجش آن‌ها از طریق پایگاه‌های اطلاعاتی سازمان بورس و اوراق بهادار و همچنین نرم‌افزار ره‌آورد نوین در دسترس بود، انتخاب شدند. این متغیرها به شرح زیر هستند:

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

۵-۱- اعداد صورت‌های مالی

اعداد صورت‌های مالی شامل سود عملیاتی، سود خالص، فروش خالص، رشد فروش خالص، مجموع دارایی‌ها، رشد دارایی‌ها، حقوق صاحبان سهام، رشد حقوق صاحبان سهام، رشد سود عملیاتی، رشد سود خالص و مالیات ابرازی می‌باشد. این متغیرها در پژوهش‌های گالمور و لابرو[1] (۲۰۱۵)، لیزوسکی[2] (۲۰۱۰)، مهرانی و سیدی (۱۳۹۳)، دستگیر و غریبی (۱۳۹۴) به‌منظور پیش‌بینی فرار مالیاتی و اجتناب مالیاتی استفاده شده است.

۵-۲- نسبت‌های مالی

نسبت‌های مالی استفاده شده در این پژوهش به پنج دسته نسبت‌های نقدینگی، نسبت‌های سودآوری، نسبت‌های فعالیت، نسبت‌های اهرمی و نسبت‌های ارزش بازار تقسیم می‌شوند.

نسبت‌های نقدینگی شامل نسبت جاری، نسبت آنی و نسبت خالص سرمایه در گردش به دارایی‌ها می‌باشد. از این نسبت‌ها در پژوهش‌های مو[3] (۲۰۰۳)، باقرپور ولاشانی و همکاران (۱۳۹۱)، دستگیر و غریبی (۱۳۹۴)، به‌منظور پیش‌بینی فرار مالیاتی و اجتناب مالیاتی استفاده شده است.

نسبت‌های سودآوری مورد استفاده در این پژوهش شامل بازده دارایی‌ها، نسبت حاشیه سود ناخالص، نسبت حاشیه سود عملیاتی، نسبت حاشیه سود خالص، سود عملیاتی به مجموع دارایی‌ها، رشد نسبت سود عملیاتی به مجموع دارایی‌ها است. از این نسبت‌ها در پژوهش‌های گالمور و لابرو (۲۰۱۵)، مهرانی و سیدی (۱۳۹۳)، دستگیر و غریبی (۱۳۹۴)، رحیمی‌کیا و همکاران (۱۳۹۴) به‌منظور پیش‌بینی فرار مالیاتی و اجتناب مالیاتی استفاده شده است.

نسبت‌های فعالیت مورد استفاده در این پژوهش شامل گردش موجودی کالا، نسبت گردش دارایی‌های ثابت، گردش مجموع دارایی‌ها است. از این نسبت‌ها در پژوهش‌های مو (۲۰۰۳) و رحیمی‌کیا و همکاران (۱۳۹۴) به‌منظور پیش‌بینی فرار مالیاتی و اجتناب مالیاتی استفاده شده است.

نسبت‌های اهرمی مورد استفاده در این پژوهش عبارت‌اند از: نسبت بدهی، نسبت بدهی به حقوق صاحبان سهام و نسبت توانایی پرداخت بهره (نسبت پوشش هزینه‌های مالی). از این نسبت‌ها در پژوهش‌های گالمور و لابرو (۲۰۱۵)، باقرپور ولاشانی و همکاران (۱۳۹۱)، پورحیدری

1 Gallemore and Labro
2 Lisowsky
3 Mo

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

و سروستانی (۱۳۹۱)، مهرانی و سیدی (۱۳۹۳)، دستگیر و غریبی (۱۳۹۴) و رحیمی‌کیا و همکاران (۱۳۹۴)، تیلور و ریچاردسون[1] (۲۰۱۳)، ایزگیارتا[2] (۲۰۱۵) به‌منظور پیش‌بینی فرار مالیاتی و اجتناب مالیاتی استفاده شده است.

نسبت‌های ارزش بازار شامل ارزش بازار، نسبت ارزش بازار به ارزش دفتری سهام و سود هر سهم می‌باشد. از این نسبت‌ها در پژوهش‌های گراهام و همکاران (۲۰۱۴)، گالمور و لابرو (۲۰۱۵)، باقرپور ولاشانی و همکاران (۱۳۹۱)، پورحیدری و سروستانی (۱۳۹۱)، مهرانی و سیدی (۱۳۹۳)، دستگیر و غریبی (۱۳۹۴)، رضایی و جعفری نیارکی (۱۳۹۴) به‌منظور پیش‌بینی فرار مالیاتی و اجتناب مالیاتی نیز استفاده شده است.

۵-۳- متغیر گزارش حسابرسی

با توجه به این که گزارش حسابرسی یکی از منابع اطلاعات حسابداری درباره مالیات شرکت است که توسط اشخاص برون‌سازمانی تائید می‌شود و می‌توان از آن جهت کمک در پیش‌بینی استفاده کرد، گزارش حسابرسی نیز به عنوان یک متغیر مورد مطالعه قرار می‌گیرد. گزارش حسابرسی به عنوان یک متغیر دو وجهی در نظر گرفته می‌شود که در صورت وجود بند مالیاتی در گزارش حسابرسی عدد یک و نبود بند مالیاتی در گزارش حسابرسی عدد صفر منظور می‌گردد. از این متغیر در پژوهش باقرپور ولاشانی و همکاران (۱۳۹۱) نیز استفاده شده است.

۵-۴- متغیر وابسته

متغیر وابسته در این پژوهش، فرار مالیاتی است که با توجه به درصد اختلاف سود مشمول مالیات ابرازی و سود مشمول مالیات قطعی شرکت شناسایی می‌گردد. مطابق با ماده ۱۹۴ قانون مالیات‌های مستقیم مصوب ۱۳۹۴/۴/۳۱ درصورتی‌که سود مشمول مالیات قطعی شده با سود مشمول مالیات ابرازی واحد تجاری بیش از ۱۵٪ اختلاف داشته باشد، واحد تجاری دارای فرار مالیاتی است. از این متغیر در پژوهش دستگیر و غریبی (۱۳۹۴) نیز استفاده شده است.

۶- روش تجزیه‌وتحلیل داده‌ها

در این پژوهش، تکنیک مورد استفاده الگوریتم درخت تصمیم می‌باشد که یکی از مشهورترین و قدیمی‌ترین روش‌های ساخت الگوی دسته‌بندی است. این روش، یکی از روش‌های ناپارامتریک برای رده‌بندی کردن متغیرهاست و دارای الگوریتم‌های متنوعی است (افلاطونی، ۱۳۹۴). در الگوریتم‌های دسته‌بندی مبتنی بر درخت تصمیم، دانش خروجی به صورت یک درخت از حالات

1 Taylor and Richardson
2 Isgiyarta

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

مختلف مقادیر ویژگی‌ها ارائه می‌شود (صنیعی آباده و محمودی، ۱۳۹۴). نمایش دانش به شکل درخت باعث شده است که دسته‌بندی‌های مبتنی بر درخت تصمیم کاملاً قابل تفسیر باشند (صنیعی آباده و محمودی، ۱۳۹۴). این روش یکی از روش‌های ناپارامتری در مسائل طبقه‌بندی و پیش‌بینی است که با توجه به نوع متغیر وابسته به دو دسته طبقه‌بندی درختی[1] (برای متغیر اسمی) و رگرسیون درختی[2] (برای متغیر پیوسته) تقسیم می‌شود (لی[3]، ۲۰۰۶). در حالت کلی، درخت تصمیم رسم شده برای یک مجموعه داده آموزشی، واحد و یکتا نیست و بر اساس یک مجموعه داده می‌توان درخت‌های تصمیم مختلفی به دست آورد (صنیعی آباده و محمودی، ۱۳۹۴). درخت تصمیم، سه مورد را نشان می‌دهد (جعفری و آذر، ۱۳۹۲):

- هر گره[4] داخلی که یک ویژگی را آزمون می‌کند.
- هر شاخه[5] که بیانگر ارزش ویژگی (متغیر) مورد نظر است.
- هر گره برگ[6] که یک طبقه را نشان می‌دهد.

به منظور اعتبارسنجی نتایج الگوریتم‌های داده‌کاوی از دو روش Holdout و روایی متقابل استفاده می شود. در اغلب پژوهش‌های حسابداری و مالی (به‌ویژه در ایران) از روش Holdout استفاده شده است. در این روش اغلب، دو سوم داده‌ها به عنوان مجموعه آموزشی و یک سوم آن به عنوان مجموعه آزمایشی طبقه‌بندی می‌شود. بنابراین، با فرض این که دقت روش پیش‌بینی با افزایش تعداد نمونه‌های مشاهده شده، افزایش یابد، روش Holdout یک تخمین‌گر بدبینانه[7] است، زیرا تنها بخشی از داده‌ها برای آموزش به روش پیش‌بینی ارائه شده است. هر چه تعداد نمونه بیشتری برای مجموعه آزمایشی خارج شود، سوگیری برآورد بیشتر می‌شود. از طرفی، نمونه‌های آزمایشی کوچک‌تر (با تعداد کمتر) به معنای بیشتر بودن فاصله اطمینان دقت است. بنابراین، روش مزبور روش مناسبی نیست (کوهاوی[8]، ۱۹۹۵).

در مقابل، روش روایی متقابل، به دلیل سادگی، شفافیت و جامعیت، یک راهبرد متداول و مناسب است و نتایج بسیاری از پژوهش‌های انجام شده حاکی از عملکرد بهتر این روش است

1 Classification Tree
2 Regression Tree
3 Lee
4 Node
5 Beach
6 Leaf
7 Pessimistic
8 Kohavi

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

(آرلوت و سلیسه[1]، ۲۰۱۰: ۴۰). در این روش در نهایت از تمام نمونه‌ها هم به‌عنوان داده‌های آموزشی و هم به‌عنوان داده‌های آزمایشی استفاده می‌شود. افزون براین، استفاده از این روش، از بروز مشکل بیش برازش[2] و مشکلات مربوط به نتایج برون‌نمونه‌ای[3] جلوگیری می‌کند (آرلوت و سلیسه، ۲۰۱۰: ۴۲ و ۵۲). از این‌رو، در پژوهش حاضر به‌منظور بررسی تعمیم‌پذیری پیش‌بینی‌های انجام شده از روایی متقابل استفاده شده است. از طرف دیگر، بررسی روایی ۱۰ بخشی[4] برای برآورد نرخ خطای واقعی کاملاً قابل‌اتکا و کافی است (هو[5]، ۲۰۱۰: ۲۵۳۵). یافته‌های اغلب پژوهش‌ها (به عنوان نمونه، بریمن و اسپکتور[6]، ۱۹۹۲ و کوهاوی، ۱۹۹۵) نیز حاکی از آن است که در مسائل دنیای واقعی، روایی متقابل ۱۰ بخشی، بهترین روش انتخاب الگو است. ۱۰ بار تکرار بررسی روایی ۱۰ بخشی، در مورد هر الگو منجر به ایجاد ۱۰۰ پیش‌بینی می‌شود که می‌توان آن را برای بررسی تفاوت بین عملکرد الگوهای مختلف و آزمون داده‌های پژوهش مورد استفاده قرار داد. پس از مشخص شدن معیارهای مختلف عملکرد الگوریتم‌های درخت تصمیم به منظور مقایسه عملکرد روش‌های مختلف انتخاب متغیرهای پیش‌بین، از تحلیل واریانس استفاده شد.

۷- یافته‌های پژوهش

۷-۱- آماره های توصیفی

آماره های توصیفی متغیرهای پژوهش در جدول شماره(۱) ارائه شده است. نتایج نشان می‌دهد که به‌طور متوسط نیمی از شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران دارای فرار مالیاتی هستند. همچنین، به‌طور متوسط کمتر از ۴۰٪ حسابرسان به بیان مشکل مالیاتی شرکت‌ها در گزارش حسابرسی پرداخته‌اند. آماره میانگین مربوط به رشد شرکت‌ها (رشد فروش، رشد کل دارایی‌ها، رشد حقوق صاحبان سهام، رشد نسبت سود عملیاتی به مجموع دارایی‌ها، رشد سود عملیاتی و رشد سود خالص) حاکی از آن است که شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران دارای رشد مثبت بوده و در میان متغیرهای ذکر شده، رشد متغیر سود عملیاتی بیشتر از سایر متغیرها است. همچنین میانگین نسبت بدهی نشان‌دهنده آن است که بیش از نیمی از دارایی‌های شرکت‌ها از محل بدهی تأمین شده است.

1 Arlot and Celisse
2 Overfitting
3 Out-of- Sample
4 10-Fold Cross Validation
5 Hu
6 Breiman and Spector

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

جدول (۱). آماره های توصیفی متغیرهای پژوهش

متغیر	کمینه	بیشینه	میانگین	انحراف معیار
ارزش بازار	۱۰/۰۲	۱٤	۱۱/٦٥	۰/٦۳۸
نسبت ارزش بازار به ارزش دفتری	۰/۲۱	۷۰/۸۷	٥/۸۱	۷/٤۷
نسبت بدهی	۰/۱	۱/٥۷	۰/٦۳	۰/۱۸
نسبت بدهی به حقوق صاحبان سهام	-۳۹۷/٦۱۷	۳۰۳/۸۲٤۲	۲/۰٥	۲۰/٥۲۸
بازده دارایی‌ها	-۰/۳۱	۰/٦٤	۰/۱۲	۰/۱۳۲
سود هر سهم	-٤۰۱۹/۹۳٥	۹۳۲۹/۲٥	۹۱۷/۹۱	۱۳۱۱/۷۹
نسبت جاری	۰/۲۲۲	۱۰/۸٥	۱/۳٦	۰/۷۸۸
نسبت آنی	۰/۰۷۱	۸/۳٥۸	۰/۸۳۷	۰/٥٦۰
نسبت گردش موجودی کالا	۰/۲٥۱	۲۲/۲٤	۳/۲٦	۲/٦۸۷
نسبت گردش دارایی ثابت	۰	٦۷/۳٥	٦/۱۳	٥/۷۸
نسبت گردش کل دارایی‌ها	۰	٤/۸۹	۰/۹۰٥	۰/٤۸۹
نسبت حاشیه سود ناخالص	-۰/۸٦	۰/۸۳	۰/۲٦	۰/۱٥۷
نسبت حاشیه سود عملیاتی	-۱/۱٤	۲/۰۷	۰/۱۹٥	۰/۱۸۹
نسبت حاشیه سود خالص	-۱/۱۳	۲/۰۳	۰/۱٥۳	۰/۲۱
نسبت توانایی پرداخت بهره	-۱۰/۳۱	۱٤۳۱٥۰/٦۲	۲۱۰/۹۹	٤٤٦٥/۳۱
نسبت سود عملیاتی به مجموع دارایی‌ها	-۰/۳۲	۰/٦۷۱	۰/۱٦	۰/۱۳
سود عملیاتی	-۲٥٤۰٦۷۳	۱٤۳۳٤۲٦٤	۳۹٥٦۰٦/٤۱	۱۳۳۳۷۰۲/٦۲
سود خالص	-٤٦۹۲۲۲۲	۱٥۷٦۰٥۱۲	۲٦٦٦۹٥/۷۳	۱۱٥۲٥٤۲/۲
نسبت خالص سرمایه در گردش به مجموع دارایی‌ها	-۰/٥۷	۰/۸۲	۰/۱۱	۰/۲۱٤
فروش	۰	۱۰۷٤۲۰۹٦۱	۲۱۸٥٤۸۹/٤٥	۷۹٦۸۰٤٤/۲۹
رشد فروش	-۰/۷۷	٤/٦٥	۰/۱۹٦	۰/۳٤۹۰۱
کل دارایی‌ها	۲۲۷۲٥	۱۰۳٤٤٥٥۲٦	۲٦۱۳٤۸۲/۹٥	۸۹٤۳۲۳٥
رشد کل دارایی‌ها	-۰/٤۹۰	۲/۱٦۷	۰/۱۷	۰/۲٤٥
حقوق صاحبان سهام	-۳٥۳۲٥۱٥	۳۱۱۳۲٥۱۸	۷٥۹۰۲۲/٤۱	۲۳۰٦٦۲٤/٥٦
رشد حقوق صاحبان سهام	-۱۱/۸۳۱	٥٦/۸۳	۰/٤۸٤٦	۲/۹٦۰۷
رشد نسبت سود عملیاتی به مجموع دارایی‌ها	-۲٥/٤۸	۲۹/٥٥۱	۰/۲۳۷	۲/٥۷۳
رشد سود عملیاتی	-۳٦/۷۸	۳۷/۸۹۰	۰/٤۳۳	۲/۹۲
رشد سود خالص	-۱۹۱/۷۳۳	۱۲٥/۲۷	۰/۲۲٥	۸/۷٥۳
مالیات ابرازی	۰	۱۰٥۸۹٦۳	۲۸۷٦۷/٤۰	۷۷٤۲۷/۲٤
بند مالیاتی	۰	۱	۰/۳۹	۰/٤۸۹
فرار مالیاتی	۰	۱	۰/٥٤	۰/٤۹۸

۷-۲- آمار استنباطی

نتایج حاصل از الگوریتم‌های مختلف به کار برده شده در این پژوهش به منظور پیش‌بینی فرار مالیاتی در جدول شماره ۲ نشان داده شده است. میزان صحت الگوها از طریق آماره «درصد طبقه‌بندی صحیح» اندازه‌گیری می‌شود. با توجه به نتایج مندرج در جدول شماره ۲، بالاترین میزان صحت الگوهای مورد استفاده در زمینه پیش‌بینی فرار مالیاتی را روش جنگل تصادفی و

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

کمترین آن را درخت تصادفی دارا است. آماره Kappa یک معیار تصحیح شده بر مبنای تصادف (شانس) برای تطابق بین طبقه‌بندی و کلاس‌های صحیح می‌باشد که مقدار بالاتر از صفر این آماره نشان‌دهنده این موضوع است که طبقه‌بندی کننده مورد نظر بهتر از حالت تصادفی کار می‌کند. هرچه این میزان بیشتر باشد نشان‌دهنده عملکرد بهتر الگوریتم استفاده شده است. با توجه به آماره‌های مندرج در جدول شماره ۲، الگوریتم جنگل تصادفی دارای بهترین عملکرد در پیش‌بینی فرار مالیاتی و الگوریتم درخت تصادفی دارای ضعیف‌ترین عملکرد است.

جدول (۲). نتایج حاصل از الگوریتم‌های پژوهش

	کاهش خطای هرس[1]	درخت تصادفی	جنگل تصادفی	LMT	.J48	ریشه تصمیم[2]
درصد طبقه‌بندی صحیح	٦٧/٥٣	٦٠/٥٩	٧٠/٢١	٦٦/٩٧	٦٨/٣٦	٦٥/٥٨
درصد طبقه‌بندی نادرست	٣٢/٤٦	٣٩/٤٠	٢٩/٧٨	٣٣/٠٢	٣١/٦٣	٣٤/٤١
آماره Kappa	٠/٣٤٥	٠/٢٠٦٥	٠/٤٠٠٦	٠/٣٣٦	٠/٣٦٣٨	٠/٣٠٨٦
میانگین قدر مطلق خطا	٠/٣٩٦١	٠/٣٨٩٤	٠/٣٩٢٥	٠/٣٨٥٨	٠/٣٧٠١	٠/٤٤٧٦
ریشه میانگین مجذور خطا	٠/٤٧٢٣	٠/٦١٥٢	٠/٤٤١٧	٠/٤٧٨٦	٠/٤٩٨٥	٠/٤٧٤٤
خطای مطلق نسبی	٧٩/٨٠	٧٨/٤٤	٧٩/٠٨	٧٧/٧٣	٧٤/٥٧	٩٠/١٧
ریشه مجذور خطای نسبی	٩٤/٨١	١٢٣/٥٠	٨٨/٦٦	٩٦/٠٧	١٠٠/٠٧	٩٥/٢٢

جدول شماره ۳ نشان‌دهنده ماتریس درهم‌ریختگی حاصل از بررسی داده‌های پژوهش می‌باشد. این ماتریس، چگونگی عملکرد الگوریتم دسته‌بندی را با توجه به مجموعه داده ورودی به تفکیک انواع دسته‌های مسئله طبقه‌بندی نشان می‌دهد. علامت مثبت بدین معناست که شرکت دارای فرار مالیاتی بوده و علامت منفی نشان‌دهنده این مورد است که شرکت بدون فرار مالیاتی می‌باشد؛ بنابراین عددی که در هر دو سمت آن دارای علامت مثبت می‌باشد، بدین معناست که شرکت دارای فرار مالیاتی بوده و روش پیش‌بینی نیز به‌طور صحیح اقدام به پیش‌بینی آن نموده است که اصطلاحاً به این مورد مثبت‌های واقعی[3] گفته می‌شود. همین مورد نیز در رابطه با علامت منفی صادق است. بدین معنا که شرکت دارای فرار مالیاتی نبوده و الگوریتم داده‌کاوی نیز به‌طور صحیح اقدام به پیش‌بینی آن نموده است. در مواردی که اعداد دارای علامت-های متفاوت در سطرها و ستون‌های ماتریس باشند بدین معناست که الگوریتم نتوانسته به‌طور صحیح اقدام به پیش‌بینی فرار مالیاتی کند؛ به بیان دیگر، اعداد مندرج در جدول شماره ۳ که محل تلاقی سطرها و ستون‌های با علامت متفاوت می‌باشند خطای نوع اول و نوع دوم در پیش‌بینی فرار مالیاتی را نشان می‌دهند.

1 Reduced Error Purning
2 Decision Stump
3 True Positive Numbers

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

جدول (۳). نتایج ماتریس درهم‌ریختگی الگوریتم‌های پژوهش

	کاهش خطای هرس		درخت تصادفی		جنگل تصادفی		LMT		J48		ریشه تصمیم	
	+	−	+	−	+	−	+	−	+	−	+	−
+	٤١٥	١٧٢	٣٧٢	٢١٥	٤٢٢	١٦٥	٤٠٢	١٨٥	٤١٠	١٧٧	٣٩٢	١٩٥
−	١٧٩	٣١٥	٢١١	٢٨٣	١٥٧	٣٣٧	١٧٢	٣٢٢	١٦٥	٣٢٩	١٧٧	٣١٧

جدول شماره ۴ معروف‌ترین معیارهای ارزیابی عملکرد الگوریتم‌های درخت تصمیم و نتایج آن را برای هر کدام از الگوریتم‌های درخت تصمیم استفاده شده در این پژوهش نشان می‌دهد. این معیارها شامل مثبت‌های واقعی[۱]، مثبت‌های کاذب[۲]، دقت[۳]، فراخوانی[۴]، F-Measure، ناحیه ROC[۵] و ناحیه منحنی دقت فراخوانی[۶] می‌باشند.

مهم‌ترین معیار برای تعیین کارایی یک الگوریتم طبقه‌بندی معیار دقت یا نرخ صحیح طبقه‌بندی است. این معیار نشان‌دهنده این حقیقت می‌باشد که چند درصد از کل مجموعه رکوردهای آزمایشی به درستی طبقه‌بندی شده است. معیار فراخوانی، کارایی الگوریتم را با توجه به تعداد رخداد آن طبقه نشان می‌دهد. حال آنکه معیار دقت اساساً مبتنی بر دقت پیش‌بینی الگوریتم می‌باشد و مبین این موضوع است که به چه میزان می‌توان به خروجی‌های الگوریتم اعتماد کرد. معیار F-measure ترکیب معیارهای فراخوانی و دقت را نشان می‌دهد و در مواردی به کار می‌رود که نتوان اهمیت ویژه‌ای را برای هر یک از دو معیار فراخوانی و دقت نسبت به یکدیگر قائل شد. این معیار از طریق فرمول زیر محاسبه می‌شود:

$$F - measure = \frac{2 * recall * precsion}{recall + precision} \qquad (۱)$$

معیار ناحیه زیر منحنی (AUC) دیگر معیار مهمی است که برای تعیین میزان کارایی یک الگوریتم بسیار مؤثر می‌باشد. این معیار نشان‌دهنده سطح زیر نمودار ROC است. نمودار ROC روشی برای بررسی کارایی الگوریتم‌ها بوده و هر چه این مقدار بزرگ‌تر باشد کارایی نهایی الگوریتم مطلوب‌تر ارزیابی می‌شود.

1 True Positive Rate
2 False Positive Rate
3 Precision
4 Recall
5 Receiver Operating Characteristics
6 Precision-Recall Curves

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

جدول (٤). نتایج معیارهای مختلف الگوریتم‌های پژوهش

	مثبت‌های واقعی	مثبت‌های کاذب	دقت	فراخوانی	F-Measure	ROC Area	ناحیه منحنی دقت فراخوانی	طبقه
ریشه تصمیم	۰/٦٦٨	۰/۳٥٨	۰/٦٨٩	۰/٦٦٨	۰/٦٧٨	۰/٦۳٦	۰/٦۳۰	دارای فرار مالیاتی
	۰/٦٤۲	۰/۳۳۲	۰/٦۱۹	۰/٦٤۲	۰/٦۳۰	۰/٦۳٦	۰/٥۷۰	بدون فرار مالیاتی
	۰/٦٥٦	۰/۳٤٦	۰/٦٥۷	۰/٦٥٦	۰/٦٥٦	۰/٦۳٦	۰/٦۰۳	میانگین موزون
J48	۰/٦۹۸	۰/۳۳٤	۰/۷۱۳	۰/٦۹۸	۰/۷۰٦	۰/٦۹٤	۰/٦۹۰	دارای فرار مالیاتی
	۰/٦٦٦	۰/۳۰۲	۰/٦٥۰	۰/٦٦٦	۰/٦٥۸	۰/٦۹٤	۰/٦۲۰	بدون فرار مالیاتی
	۰/٦۸٤	۰/۳۱۹	۰/٦۸٤	۰/٦۸٤	۰/٦۸٤	۰/٦۹٤	۰/٦٥۸	میانگین موزون
LMT	۰/٦۸٥	۰/۳٤۸	۰/۷۰۰	۰/٦۸٥	۰/٦۹۳	۰/۷۱٤	۰/۷۱۸	دارای فرار مالیاتی
	۰/٦٥۲	۰/۳۱٥	۰/٦۳٥	۰/٦٥۲	۰/٦٤۳	۰/۷۱٤	۰/٦٤۱	بدون فرار مالیاتی
	۰/٦۷۰	۰/۳۳۳	۰/٦۷۱	۰/٦۷۰	۰/٦۷۰	۰/۷۱٤	۰/٦۸۳	میانگین موزون
جنگل تصادفی	۰/۷۱۹	۰/۳۱۸	۰/۷۲۹	۰/۷۱۹	۰/۷۲٤	۰/۷٦۹	۰/۷۸۳	دارای فرار مالیاتی
	۰/٦۸۲	۰/۲۸۱	۰/٦۷۱	۰/٦۸۲	۰/٦۷۷	۰/۷٦۹	۰/۷۲۸	بدون فرار مالیاتی
	۰/۷۰۲	۰/۳۰۱	۰/۷۰۳	۰/۷۰۲	۰/۷۰۲	۰/۷٦۹	۰/۷٥۸	میانگین موزون
درخت تصادفی	۰/٦۳٤	۰/٤۲۷	۰/٦۳۸	۰/٦۳٤	۰/٦۳٦	۰/٦۲۲	۰/٦۲٦	دارای فرار مالیاتی
	۰/٥۷۳	۰/۳٦٦	۰/٥٦۸	۰/٥۷۳	۰/٥۷۱	۰/٦۲٦	۰/٥۳۹	بدون فرار مالیاتی
	۰/٦۰٦	۰/۳۹۹	۰/٦۰٦	۰/٦۰٦	۰/٦۰٦	۰/٦۲٤	۰/٥۸٦	میانگین موزون
کاهش خطای هرس	۰/۷۰۷	۰/۳٦۲	۰/٦۹۹	۰/۷۰۷	۰/۷۰۳	۰/۷۱۱	۰/۷۲٤	دارای فرار مالیاتی
	۰/٦۳۸	۰/۲۹۳	۰/٦٤۷	۰/٦۳۸	۰/٦٤۲	۰/۷۱۱	۰/٦۳۲	بدون فرار مالیاتی
	۰/٦۷٥	۰/۳۳۱	۰/٦۷٥	۰/٦۷٥	۰/٦۷٥	۰/۷۱۱	۰/٦۸۲	میانگین موزون

منحنی‌های ROC رفتار یک الگوریتم را بدون توجه به توزیع دسته یا هزینه خطا نشان می‌دهند؛ بنابراین کارایی الگوریتم را از این عوامل جدا می‌کنند. متأسفانه در حالی که منحنی

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

ROC یک تکنیک با ارزش مصورسازی است اما در انتخاب الگوریتم مناسب کمک کمی می‌کند؛ به عبارت دیگر تنها زمانی می‌توان گفت یک الگوریتم نسبت به الگوریتم دیگر دارای کارایی بهتری است که در کل فضای کارایی به وضوح بر الگوریتم دیگر تسلط یابد. به همین دلیل معیار AUC که سطح زیر نمودار ROC را نشان می‌دهد می‌تواند نقش تعیین‌کننده‌ای را در معرفی الگوریتم برتر معرفی کند. معیار AUC برای یک الگوریتم که به صورت تصادفی، طبقه نمونه مورد بررسی را تعیین می‌کند برابر ۰/۵ است. همچنین بیشترین میزان این معیار برابر با یک بوده و برای وضعیتی رخ می‌دهد که در آن الگوریتم کلیه نمونه‌ها را بدون هرگونه اشتباهی تعیین کند.

با توجه به جدول شماره ۴، در تمام معیارهای ارزیابی عملکرد، الگوریتم جدول تصادفی دارای بهترین عملکرد و درخت تصادفی دارای ضعیف‌ترین عملکرد بوده است. در ادامه، به منظور بررسی تفاوت بین عملکرد الگوریتم‌های مختلف درخت تصمیم از تحلیل واریانس استفاده شده است. در این راستا، از نتایج معیار F-measure موزون حاصل از ۱۰ بار تکرارِ روایی متقابل ۱۰ بخشی (روایی متقابل ۱۰ بخشی با ۱۰ بار تکرار) استفاده شد که منجر به ایجاد ۱۰۰ داده در مورد هر الگوریتم می‌شود (در مجموع ۶۰۰ دقت).

نتایج حاصل از بررسی نرمال بودن متغیر عملکرد مدل در جدول شماره ۵ نشان داده شده است. مطابق با اطلاعات این جدول، با توجه به آماره آزمون شاپیرو ویلک مشاهده می‌شود که متغیر وابسته در تمام روش‌ها از توزیع نرمال برخوردار است.

جدول (۵). آزمون نرمال بودن

	آزمون شاپیرو-ویلک		
	آماره	درجه آزادی	معناداری
ریشه تصمیم	۰/۹۸۲	۱۰۰	۰/۱۹۹
J48	۰/۹۸۳	۱۰۰	۰/۲۱۷
LMT	۰/۹۷۹	۱۰۰	۰/۱۱۷
جنگل تصادفی	۰/۹۸۵	۱۰۰	۰/۳۱۱
درخت تصادفی	۰/۹۸۲	۱۰۰	۰/۱۹۲
کاهش خطای هرس	۰/۹۸۳	۱۰۰	۰/۲۴۵

جدول شماره ۶ نتایج مرتبط با بررسی همگنی واریانس‌ها با استفاده از آزمون لون را نشان می‌دهد؛ مطابق با اطلاعات مندرج در این جدول که حاکی از بیشتر بودن سطح معناداری آماره لون از ۰/۰۵ است، می‌توان گفت که مشکلی در خصوص همگنی واریانس‌ها وجود ندارد.

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

جدول (٦). آزمون همگنی واریانس F- measure

آماره لون	درجه آزادی ۱	درجه آزادی ۲	معناداری
۰/۷۰٦	٥	٥٩٤	۰/٦۱۹

برای بررسی تفاوت بین عملکرد الگوریتم‌های مختلف درخت تصمیم، نتایج آزمون تحلیل واریانس در جدول ۷ نشان داده شده است. مطابق با آماره F مندرج در این جدول، تفاوت میانگین گروه‌ها تائید و در مقابل فرض یکسان بودن آن‌ها رد می‌شود. به عبارت دقیق‌تر، دست کم یکی از گروه‌ها از نظر میانگین نمره مورد نظر با سایر گروه‌ها متفاوت است.

جدول (۷). نتایج تحلیل واریانس F- measure

	مجموع مربعات	درجه آزادی	میانگین مجموع مربعات	آماره F	سطح معناداری
بین گروه‌ها	۰/۳٦۲	٥	۰/۰۷۲		
درون گروه‌ها	۱/۰۰۸	٥٩٤	۰/۰۰۲	٤۲/٦۷٦	۰/۰۰۰
کل	۱/۳۷۰	٥٩۹	–		

جدول (۸). نتایج آزمون توکی

الگوریتم	فراوانی	زیرمجموعه‌های همگون (در سطح خطای ۵٪)			
		۱	۲	۳	٤
درخت تصادفی	۱۰۰	۰/٦۲			
ریشه تصمیم	۱۰۰		۰/٦٥		
J48	۱۰۰			۰/٦۷	
LMT	۱۰۰			۰/٦۷	
کاهش خطای هرس				۰/٦۸	
جنگل تصادفی	۱۰۰				۰/۷۰
معناداری	–	۱/۰۰	۱/۰۰	۰/۹۹٥	۱/۰۰

مطابق جدول شماره ۸، نتایج آزمون توکی نیز نشان داد که می‌توان شاخص F-measure موزون را بر حسب میانگین‌ها در سطح هر یک از الگوریتم‌های درخت تصمیم به چهار زیرگروه همگن تقسیم کرد. به‌عبارت دیگر، آزمون توکی درخت تصادفی را در گروه اول، ریشه تصمیم را در گروه دوم؛ J48، کاهش خطای هرس و LMT را در گروه سوم و در نهایت جنگل تصادفی را در گروه چهارم طبقه‌بندی کرده است.

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

۸- بحث و نتیجه‌گیری

در چند دهه اخیر فرار مالیاتی تبدیل به یکی از دغدغه‌های دولت‌ها شـــده اســـت. با این حال، پژوهش‌های انجام شده در رابطه با فرار مالیاتی، معمولاً از دیدگاه اقتصادی به بررسی موضوع پرداخته و کمتر پژوهشی به بررسی فرار مالیاتی با استفاده از اطلاعات حسابداری پرداخته است. از سوی دیگر، عمدتاً پژوهش‌های انجام شده داخلی نیز تنها یک صنعت خاص و یک محدوده زمانی کوتاه در بورس اوراق بهادار تهران را مدنظر قرار داده‌اند. افزون بر این، اســتفاده از رویکرد داده‌کاوی نیز به عنوان یکی از روش‌های جدید کمتر مورد اســتفاده پژوهشــگران در این زمینه قرار گرفته اســـت. از این رو، این پژوهش با معرفی کاربرد رویکردهای داده‌کاوی در بحث فرار مالیاتی، به شـــناســـایی روش‌ها و الگوریتم‌های کاراتر برای بررســـی فرار مالیاتی در شـــرکت‌ها پرداخته است؛ بنابراین، این پژوهش به دنبال پاسخگویی به این پرسش است که آیا می‌توان با استفاده از روش‌های داده‌کاوی به پیش‌بینی فرار مالیاتی در بورس اوراق بهادار تهران پرداخت و آیا میان الگوریتم‌های مختلف این روش‌ها، تفاوتی در کارایی آن‌ها وجود دارد یا خیر؟

نتایج پژوهش بیانگر وجود شـــواهدی دال بر عملکرد مناســـب الگوهای پیشـــنهادی برای پیش‌بینی فرار مالیاتی اســـت. به بیان دقیق‌تر تحلیل داده‌های پژوهش نشـــان داد که بهترین روش‌های الگوریتم درخت تصـــمیم برای پیش‌بینی فرار مالیاتی که از عملکرد بالاتری برخوردار هستند، عبارت‌اند از:

- روش جنگل تصادفی (برابر با ۰/۷۲۴)؛
- روش کاهش خطای هرس (برابر با ۰/۷۰۳)؛
- روش J48 (برابر با ۰/۷۰۶)؛
- روش LMT (برابر با ۰/۶۹۳)؛
- روش ریشه تصمیم (برابر با ۰/۶۷۸)؛ و
- روش درخت تصادفی (برابر با ۰/۶۳۶).

این موضوع نشان می‌دهد که در وضعیت کشور ایران روش جنگل تصادفی بهترین روش برای پیش‌بینی فرار مالیاتی است. افزون بر این نتایج حاصل از آزمون تحلیل واریانس نشان داد که بین میزان دقت پیش‌بینی الگوهای مختلف درخت تصـــمیم تفاوت معناداری وجود دارد. نتایج این پژوهش بیانگر آن است که با استفاده از اطلاعات حسابداری در سطح خرد (شرکت‌ها و داده‌های حسابداری) می‌توان اقدام به پیش‌بینی فرار مالیاتی از سوی شرکت‌ها کرد. این در حالی است که رویکردهای اقتصادی نگاهی کلی به فرار مالیاتی داشته و در سطح خرد اقدام به بررسی فرار مالیاتی نمی‌کنند اما با استفاده از اطلاعات حسابداری و فن‌های داده‌کاوی می‌توان

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

به پیش‌بینی فرار مالیاتی در سطح خرد پرداخت. نتایج این پژوهش نیز همانند پژوهش‌های باقرپور ولاشانی و همکاران (۱۳۹۱)، دستگیر و غریبی (۱۳۹۴)، رحیمی‌کیا و همکاران (۱۳۹۴)، سامعی راد و شاه‌بهرامی (۱۳۹۵)، تقوی‌فرد و همکاران (۱۳۹۶)، وو و همکاران (۲۰۱۲)، رحیمی کیا و همکاران (۲۰۱۷) و دیدیمو و همکاران (۲۰۱۸) نشان داد که استفاده از رویکردهای داده‌کاوی می‌تواند منجر به بهبود روش‌های کشف فرار مالیاتی شود.

۹- محدودیت‌های پژوهش

لازم به ذکر است که علی‌رغم اهمیت نوع صنعت در پیش‌بینی فرار مالیاتی، به دلیل حجم محدود جامعه، پژوهشگران در این مطالعه در خصوص تطبیق شرکت‌های دارای فرار مالیاتی و بدون فرار مالیاتی از نظر نوع صنعت با محدودیت مواجه بودند. با وجود این، استفاده از صنایع مختلف در این پژوهش، باعث افزایش تعمیم‌پذیری یافته‌های پژوهش به کلیه شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران می‌شود. همچنین، با وجود این که برخی از متغیرهای بکار رفته در پژوهش‌های خارجی برای پیش‌بینی فرار مالیاتی سودمند به نظر می‌رسید، به دلیل عدم امکان محاسبه این متغیرها از طریق پایگاه‌های اطلاعاتی ذکر شده در پژوهش، از محاسبه و استفاده از آن‌ها در پیش‌بینی فرار مالیاتی صرف‌نظر شد.

۱۰- پیشنهادهای پژوهش

بر اساس این یافته‌ها به مسئولان اقتصادی و مالیاتی کشور پیشنهاد می شود که از الگوهای مورد اشاره در پژوهش حاضر جهت پیش‌بینی فرار مالیاتی در شرکت‌ها و جلوگیری از این اقدام ناپسند اجتماعی و اقتصادی بهره برند. همچنین، به پژوهشگران آینده پیشنهاد می‌شود که با استفاده از سایر تکنیک‌های داده‌کاوی (به‌عنوان نمونه، ماشین بردار پشتیبان و شبکه‌های عصبی مصنوعی) به مطالعه در زمینه پیش‌بینی فرار مالیاتی بپردازند.

۱۱- منابع و مأخذ

اسماعیلی، مهدی (برگرداننده). (۱۳۹۳). **داده‌کاوی: مفاهیم و تکنیک‌ها**، چاپ اول، تهران، ایران: انتشارات نیاز دانش.

افلاطونی، عباس. (۱۳۹۴). بررسی عوامل تعیین‌کننده سطح نگهداشت موجودی‌های مواد و کالا و رتبه‌بندی آن‌ها با استفاده از الگوریتم‌های درخت تصمیم و شبکه عصبی. **فصلنامه حسابداری مالی** ۲۷: ۵۲-۳۰.

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

باقرپور ولاشانی، محمدعلی؛ مصطفی باقری؛ حمید خادم و رضا حسینی پور. (۱۳۹۱). بررسی عوامل مالی و غیرمالی مؤثر بر گریز مالیاتی با استفاده از تکنیک‌های داده کاوی: صنعت خودرو و ساخت قطعات. **مطالعات تجربی حسابداری مالی** ۳۴: ۱۲۸-۱۰۳.

باقری، بهروز و مهدی عموری. (۱۳۸۸). شناسایی و اولویت‌بندی نسبت‌های مالی اثرگذار بر درآمد مشمول مالیات. **فصلنامه تخصصی مالیات** ۴: ۸۰-۷۳.

پورحیدری، امید و امیر سروستانی. (۱۳۹۱). بررسی تأثیر ویژگی‌های شرکت، نوع صنعت و مالکیت نهادی بر اختلاف مالیات ابرازی و قطعی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. **پژوهشنامه مالیات** ۲۰(۱۴): ۷۷-۶۱.

تقوی‌فرد، سید محمدتقی؛ ایمان رئیسی وانانی و ریحانه پناهی. (۱۳۹۶). تحلیل آینده‌نگر تشخیص فرار مالیاتی مؤدیان مالیات بر ارزش افزوده با استفاده از الگوریتم‌های طبقه‌بندی و خوشه‌بندی. **پژوهشنامه مالیات** ۳۵: ۳۵-۱۱.

جعفری، بهزاد و عادل آذر. (۱۳۹۲). درخت تصمیم فازی؛ رویکردی نوین در تدوین استراتژی. **پژوهش‌های مدیریت عمومی** ۶(۱۹): ۳۹-۲۵.

چالاکی، پری و مرتضی یوسفی. (۱۳۹۱). پیش‌بینی مدیریت سود با استفاده از درخت تصمیم‌گیری. **مطالعات حسابداری و حسابرسی** ۱: ۱۲۳-۱۱۰.

دستگیر، محسن و مریم غریبی. (۱۳۹۴). کاربست روش‌های داده‌کاوی به منظور ارتقای عملکرد تشخیص فرار مالیاتی. **پژوهشنامه مالیات** ۲۸: ۱۱۶-۹۵.

رحیمی‌کیا، اقبال؛ شاپور محمدی و مهدی غضنفری. (۱۳۹۴). تشخیص فرار مالیاتی با استفاده از سیستم هوشمند ترکیبی. **پژوهشنامه مالیات** ۲۳(۲۶): ۱۶۴-۱۳۶.

رضایی، فرزین و روح‌اله جعفری نیارکی. (۱۳۹۴). رابطهٔ بین اجتناب مالیاتی و تقلب در حسابداری شرکت‌ها. **پژوهشنامه مالیات** ۲ (۲۶): ۱۳۴-۱۰۹.

سامعی‌راد، مهدی و اسداله شاه‌بهرامی. (۱۳۹۵). بهبود کارایی الگوریتم‌های تشخیص تقلب مالیاتی با استفاده از الگوهای پردازش موازی، **پژوهشنامه مالیات** ۲۹: ۳۲-۱۱.

شیرخانی، فریبا. (۱۳۸۵). بررسی مالیات بر درآمد بخش کشاورزی در ایران، **مجله اقتصادی** ۵۳ و ۵۴: ۶۴-۴۱.

صنیعی آباده، محمد و سینا محمودی. (۱۳۹۴). **داده‌کاوی کاربردی**. چاپ دوم، تهران، ایران: انتشارات نیاز دانش.

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

ضیائی بیگدلی، محمدتقی و فرهاد طهماسبی بلداجی. (۱۳۸۴). **مالیات بر ارزش افزوده (مالیات مدرن)**. تهران: پژوهشکده امور اقتصادی.

عبداله‌میلانی، مهنوش و نرگس اکبرپورروشن. (۱۳۹۱). فرار مالیاتی ناشی از اقتصاد غیررسمی در ایران. **پژوهشنامه مالیات** ۱۳(۶۱): ۱۶۷-۱۴۱.

عبدلی، محمدرضا؛ محمدرضا شورورزی و سعیده میرزائی. (۱۳۹۲). بررسی ارتباط بین تفاوت مالیات ابرازی و تشخیصی با شفافیت شرکتی. **فصلنامه حسابداری مالی** ۱۹: ۹۰-۶۹.

فلاحتی، علی؛ مینو نظیفی و سحر عباسپور. (۱۳۹۱). مدل‌سازی اقتصاد سایه‌ای و تخمین فرار مالیاتی در ایران با استفاده از شبکه عصبی مصنوعی. **فصلنامه تحقیقات توسعه اقتصادی** ۲(۶): ۵۸-۳۳.

لاری دشت بیاض، محمود؛ کامران قائم‌مقامی و قاسم کهرمی. (۱۳۹۵). بررسی عوامل مؤثر بر فرار مالیاتی در استان خراسان جنوبی با تأکید بر مؤلفه‌های فرهنگی، **حسابداری ارزشی و رفتاری** ۱: ۱۶۴-۱۳۹.

مهرانی، ساسان و سید جلال سیدی. (۱۳۹۳). بررسی تأثیر مالیات بر درآمد و حسابداری محافظه‌کارانه بر اجتناب مالیاتی شرکت‌ها، **دانش حسابداری و حسابرسی مدیریت**، ۱۰: ۱۳-۳۳.

موسوی‌جهرمی، یگانه؛ فرهاد طهماسبی بلداجی و نرگس خاکی. (۱۳۸۸). فرار مالیاتی در نظام مالیات بر ارزش‌افزوده :یک مدل نظری، **فصلنامه تخصصی مالیات** ۱۷(۵): ۳۸-۲۷.

نویدی قاضیانی، حمیدرضا؛ علی طیب‌نیا و زهرا اسکندری دورباطی. (۱۳۸۹). نظام تشخیص علی‌الراس در یک نظام آرمانی. **پژوهشنامه مالیات** ۱۸(۹): ۲۴-۹.

هادیان، ابراهیم و علی تحویلی. (۱۳۹۲). شناسایی عوامل مؤثر بر فرار مالیاتی در اقتصاد ایران. **فصلنامه برنامه‌ریزی و بودجه** ۱۸(۲): ۵۸-۳۹.

Alfaro, E., N. García, M. Gámez, and D. Elizondo. (2008). Bankruptcy forecasting: An empirical comparison of adaboost and neural networks. **Decision Support Systems** 45: 110–122.

Allingham, M.G., and A. Sandmo. (1972). Income tax evasion: A theoretical analysis. **Journal of Public Economics** 1(3-4): 323-338.

Arlot, S. and A. Celisse. (2010). A Survey of Cross-Validation Procedures for Model Selection, **Statistics Surveys** 4: 40-79.

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

Breiman, L., and P. Spector. (1992). Submodel Selection and Evaluation in Regression: The X-Random Case, **International Statistical Review** 60(3): 291-319.

Didimo, W., L. Giamminonni, G. Liotta, F. Montecchiani, and Pagliuca, D. (2018). A Visual Analytics System to support Tax Evasion Discovery. **Decision Support Systems**.

Gallemore, J., and E. Labro. (2015). The importance of the internal information environment for tax avoidance. **Journal of Accounting and Economics** 60(1): 149-167.

Graham, J.R., M. Hanlon, T. Shevlin, and N. Shroff. (2013). Incentives for tax planning and avoidance: Evidence from the field. **The Accounting Review** 89(3): 991-1023.

Han, J., and Y. Fu. (1999). Mining multiple-level association rules in large databases. **Knowledge and Data Engineering**, **IEEE Transactions on** 11(5): 798-805.

Hu, Y.C. (2010). Analytic Network Process for Pattern Classification Problems Using Genetic Algorithms, **Information Sciences** 180(13): 2528–2539.

Isgiyarta, J. (2014). Tax Avoidance through Thin Capitalization (Evidence from Indonesian Firms). **International Journal of Research in Business and Technology** 5(3): 692-699.

Johnson, S., D. Kaufmann, J. McMillan, and C. Woodruff. (2000). Why do firms hide? Bribes and unofficial activity after communism. **Journal of Public Economics** 76(3): 495-520.

Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection, **IJCAI'95 Proceedings of the 14th international joint conference on Artificial intelligence** pp.1137-1143.

Lee, M-C., and C. To. (2010). Comparison of support vector machine and back propagation neural network in evaluating the enterprise financial distress. **International Journal of Artificial Intelligence & Applications** 1(3): 31-43.

Lisowsky, P. (2010). Seeking shelter: Empirically modeling tax shelters using financial statement information. **The Accounting Review** 85(5): 1693-1720.

Min, J.H., and Y. Lee. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. **Expert Systems with Applications** 28, 603–614.

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]

Mo, P.L.L. (2003). **Tax avoidance and anti-avoidance measures in major developing economies**. Greenwood Publishing Group.

Moore, A., J. Schneider, and K. Deng, (1997). Efficient Locally Weighted Polynomial Regression Predictions, **International Conference on Machine Learning** pp. 236-244.

Mukkamala, S., A.H. Sung, B. Ribeiro, and A. Vieira (2006). Computational intelligent techniques for financial distress detection. **Journal of Computational Intelligence Research** 2(1): 60-65.

Rahimikia, E., S. Mohammadi, T. Rahmani, and M. Ghazanfari. (2017). Detecting corporate tax evasion using a hybrid intelligent system: A case study of Iran. **International Journal of Accounting Information Systems** 25, 1-17.

Sandmo, A. (2005). The theory of tax evasion: A retrospective view. **National Tax Journal** 1, 643-663.

Stankevicius, E., and L. Leonas. (2015). Hybrid approach model for prevention of tax evasion and fraud. **Procedia-Social and Behavioral Sciences** 213, 383-389.

Taylor, G., and G. Richardson. (2013). The determinants of thinly capitalized tax avoidance structures: Evidence from Australian firms. **Journal of International Accounting, Auditing and Taxation** 22(1): 12-25.

Tsai, C. (2005). Feature selection in bankruptcy prediction. **Knowledge-Based Systems** 22, 120–127.

Wu, R.S., C.S. Ou, H.Y. Lin, S.I. Chang, and D.C. Yen. (2012). Using data mining technique to enhance tax evasion detection performance. **Expert Systems with Applications** 39*(*10): 8769-8777.

[Downloaded from qfaj.mobarakeh.iau.ir on 2025-03-19]