

## ارائه یک الگوی پیش بینی فرار مالیاتی مشاغل مبتنی بر تکنیک های داده کاوی

محمد قاسمی<sup>۱</sup>، صادق عابدی<sup>۲\*</sup>، علی محتشمی<sup>۳</sup>

<sup>۱</sup> دانشجوی دکتری، گروه مدیریت صنعتی، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران  
<sup>۲</sup> استادیار، گروه مدیریت صنعتی، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران (عهده دار مکاتبات)  
<sup>۳</sup> استادیار گروه مدیریت صنعتی، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران  
تاریخ دریافت: مهر ۱۴۰۱، اصلاحیه: آذر ۱۴۰۱، پذیرش: دی ۱۴۰۱

### چکیده

در پژوهش پیش رو با توجه به اهمیت موضوع مالیات در تهیه بودجه کشور و خلاء تحقیقاتی داخلی و خارجی، یک مدل پیش بینی فرار مالیاتی مشاغل با استفاده از روش های داده کاوی ارائه می گردد. داده های مورد بررسی شامل ۶۴۰۰ پرونده مالیاتی مشاغل دارای کد مالیاتی در استان قزوین طی سال های ۱۳۹۳ تا ۱۳۹۸ می باشد. پرونده های شغلی مرتبط با مشاغل در پنج گروه مالیاتی شامل گروه صنفی وکلا و کارشناسان رسمی دادگستری و مشاورین حقوقی و دفاتر ازدواج و طلاق و مشاغل وابسته، گروه صنفی بساز بفروشان، مصالح ساختمانی، مهندسین ساختمان و سایر مشاغل وابسته، گروه صنفی فست فود ها، هتل ها و متل ها، رستوران ها و چلو کبابی ها و مشاغل وابسته، گروه صنفی بساز بفروشان، مصالح ساختمانی، مهندسین ساختمان و سایر مشاغل وابسته، گروه صنفی فست همراه و سیم کارت و مشاغل وابسته، گروه صنفی فروشندگان لوازم یدکی خودرو سبک و سنگین، تعمیرگاه خودرو و نمایندگی مجاز فروش خودرو و مشاغل وابسته دسته بندی می شوند. جهت مدل سازی با استفاده از نمودار ROC نسبت به بررسی الگوریتم درخت تصمیم، قوانین انجمنی و قانون بیز استفاده گردید که در نهایت مدل درخت تصمیم در این پژوهش مبنا قرار گرفته است. نتایج نشان می دهد معیار صحت برابر ۰/۶۷ و معیار پوشش برابر ۰/۶۸ بدست آمده است که نشان دهنده عملکرد مناسب و بهینه مدل ساز در این پژوهش می باشد. با استفاده از تکنیک Cross Validation صحت اعتبار مدل پیش بینی مورد آزمون قرار گرفت تا با اطمینان بالا درصد عملکرد مدل سازی تخمین زده شود. معیار کاپا برابر ۰/۶۱۷ نشان از قابلیت اطمینان مناسب در جهت مدل پیش بینی می باشد. نتایج حاصل می تواند در تدوین سیاست گذاری راهبردهای اخذ مالیات در استان ها مورد استفاده قرار گیرد.

**واژه های اصلی:** اجتناب مالیاتی، فرار مالیاتی، مشاغل، پیش بینی، داده کاوی، درخت تصمیم

### ۱- مقدمه

کاهش و به تعویق انداختن مالیات بر درآمد خود دارند. [۳۸] 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 درآمدهای خود می گریزند و بخش قابل توجهی از درآمد دولت به این ترتیب از بین می رود [۲۳]. این رفتار مؤدیان مالیاتی در پرداخت نکردن مالیات را می توان به دو طبقه اجتناب مالیاتی و فرار مالیاتی تقسیم کرد. براساس این تفکیک در صورتیکه عدم پرداخت مالیات از طرق قانونی انجام شود، "اجتناب از پرداخت مالیات" و در صورتی که از روش های غیرقانونی از پرداخت مالیات اجتناب شود، "فرار مالیاتی" نامیده میشود. [۳۳]

درآمدهای مالیاتی یکی از پایدارترین منابع جهت تأمین هزینه های حاکمیتی دولت ها می باشد. [۸] در این خصوص یکی از اهداف اقتصادی برنامه های توسعه اقتصادی - اجتماعی و سند چشم انداز کاهش اتکا به ذخایر نفتی و درآمدهای حاصل از آن است؛ بنابراین افزایش سهم درآمدهای مالیاتی در بودجه دولت بهترین شیوه برای کاهش اتکای کشور به درآمدهای نفتی است. [۲۴] در این راستا کاهش وابستگی هزینه های عمومی به منبع نفت و انجام اصلاحات لازم برای افزایش منابع درآمدی مالیاتی دولت و کاهش حجم فرار مالیاتی را می توان یکی از اولویت های مهم و ضروری در نیل به اهداف نظام اقتصادی کشور تلقی نمود. [۲] مبانی نظری و شواهد تجربی نشان میدهند مؤدیان سعی در

\*abedi.sadegh@gmail.com

با توجه به موضوع فرار مالیاتی در مشاغل و اصناف می بایست برای سهولت در روند، اصناف را در گروه های صنفی طبقه بندی کرد با توجه به این موضوع، ۱۵ گروه صنفی در کشور عبارتند از وکلا، مشاوران حقوقی و خانواده، دفتر اسناد رسمی، مشاورین املاک، تالارهای پذیرایی، رستوران ها و کبابی ها، اغذیه فروشی ها و فست فودها، مراکز اقامتی اعم از هتل، هتل آپارتمان ها، مسافرخانه ها، نمایشگاه ها و فروشگاه های وسایل نقلیه، آرایشگاه های زنانه و مردانه، آجیل و خشکبار، قنادی، میوه و تره بار، مجموعه های فرهنگی ورزشی، لوازم التحریر، بازی های رایانه ای، دفاتر خدمات ارتباطی و دیگر اماکن در کل کشور که باید تحت نظارت امور مالیاتی قرار بگیرد [۲۱]. استفاده از ابزارها و روش های پیش بینی سنتی خطای بالایی دارد و به عملکرد ضعیف تری می انجامد [۲۹]. با پیشرفت فناوری و علم تحلیل داده، ابزارهای مدرنی به منظور یافتن موارد عدم رعایت و عدم تطبیق درآمدهای گزارش شده با سایر منابع به وجود آمده است [۷۱].

رویکرد خاص این مقاله بررسی فرار مالیاتی با توجه به رویکرد داده کاوی است و با جمع آوری داده و سپس انجام مراحل پیش پردازش داده، داده هایی که مورد نظر انتخاب شده برای مدل استفاده میشود. با توجه به اینکه علم داده پیشرفت های بسیاری داشته و نرم افزار های زیادی می تواند تحلیل علم داده را انجام دهند در این مقاله سعی بر این است که با توجه به ۶۴۰۰ پرونده مالیاتی مشاغل در استان قزوین و انجام فرآیند داده کاوی بر روی این دیتاها مدل پیش بینی فرار مالیاتی مشاغل ارائه شود که می توان با بررسی گروه های صنفی که در این مقاله بحث شده و جمع آوری دیتا های سایر گروه های صنفی مدل بسیار کارآمد و دقیق برای پیش بینی رفتار مؤدیان را ارائه نمود. این مقاله در حوزه فرار مالیاتی مشاغل و رفتار مالیاتی این دسته برای نخستین بار در استان قزوین ارائه می گردد.

## ۲- ادبیات و پیشینه پژوهش

### ۲-۱- تعاریف

- فرار مالیاتی و اجتناب مالیاتی:

وقتی که یک مؤدی مالیاتی از ارائه گزارش درست در مورد درآمدهای حاصل از کار با سرمایه خود که مشمول پرداخت مالیات می شود، امتناع می کند، یک نوع عمل غیر رسمی انجام می دهد که او را از چشم مقامات دولتی و مالیاتی کشور دور نگه می دارد [۵۵]. در فرار مالیاتی، به

رغم تطبیق فعالیت با قوانین، در اجرای فعالیت یک یا چند ضابطه مصوب مراجع ذی ربط تعدا نادیده گرفته می شود [۸].

فرار مالیاتی و اجتناب از مالیات، هر دو به منظور گریز از پرداخت مالیات صورت می گیرند. اما تفاوت اساسی این دو آن است که اجتناب از پرداخت مالیات، رفتاری قانونی است، درحالی که فرار مالیاتی، رفتاری غیرقانونی است. استفاده مؤدی از راه های مختلف قانونی، مانند محاسبه مالیات براساس نرخ های گوناگون به منظور کاهش از میزان مالیات خود را اجتناب از مالیات می خوانند. در واقع عوامل اقتصادی با استفاده از روزه های قانون مالیات و بازبینی در تصمیمات اقتصادی خود تلاش می کنند تا بدهی مالیاتی خود را کاهش دهند [۸۱].

اجتناب مالیاتی به این معنی که مؤدی سعی میکند با استفاده از مواد قانونی و مفرهای موجود در آن مالیات کمتری را بپردازد بدون اینکه از صورت ظاهر قانون عدول کند اما در صورتی که هر گونه تلاش غیر قانونی برای پرداخت نکردن مالیات، عدم ارائه اطلاعات لازم در مورد عایدات و منافع مشمول مالیات به مقامات مسئول، فرار مالیاتی خوانده می شود.

وقتی مؤدی از گزارش درآمدهای مشمول مالیات خویش عدول کند یا به شیوه های عامدانه ای نظیر گزارش کمتر از واقع درآمد و فروش، گزارش بیش از واقع کسورات و یا تنظیم اظهارنامه غیرواقعی از تعهدات مالیاتی خود طفره رود، فرار مالیاتی محسوب می شود [۲۵]. اشنایدر و اینست فرار مالیاتی را به عنوان کاهش غیرقانونی مالیات پرداختی با عدم گزارش درآمد با اعلام کردن نرخ های تفریق بالاتر تعریف می کند [۴۴].

در تعریف دیگر هرگونه تلاش غیرقانونی به منظور عدم پرداخت مالیات، مانند ندادن اطلاعات لازم در مورد عواید و منافع اموال مشمول مالیات به مقامات مسئول، فرار مالیاتی خوانده میشود [۸۹].

تمایز مفهومی بین فرار مالیاتی و اجتناب از پرداخت مالیات به قانونی یا غیر قانونی بودن رفتار مؤدیان مربوط می شود [۴۲]. اما در اجتناب از مالیات، فرد نگران نیست که عمل او افشا شود. اجتناب از مالیات، از خلاءهای قانونی در قانون مالیات ها نشأت می گیرد. در اینجا فرد به منظور کاهش قابلیت پرداخت مالیات، خود دنبال راه های گریز می گردد [۴۲].

جمع آوری آمار های فرار مالیاتی از طریق داده های بزرگ یکی از عوامل تعیین کننده بی نظمی اقتصادی در یک کشور می باشد [۴۶].

زمینه های پیدایش فرار مالیاتی در بررسی های مختلف نشان می دهد شایع ترین زمینه های پیدایش فرار مالیاتی این چنین است که:

- ۱- عدم شناخت مؤدیان و مستند نبودن میزان درآمد آنها [۷۱].
- ۲- عدم مبادله کامل اطلاعات و نبودن سیستم نظارت و پیگیری در اخذ مالیات [۴۲].
- ۳- تأخیر در وصول مالیات [۴۱].

۴. وجود معافیت های وسیع و متنوع: دولت ها گاه با اتخاذ برخی سیاست های حمایتی مالیاتی و معافیت های مالیاتی، راه فرار مالیاتی را باز می کنند و در صورتی که یک سیستم اطلاعات کارآمد در نظام مالیاتی موجود نباشد، فرار مالیاتی گسترش می یابد [۷۸]. همچنین با اعطای معافیت های مالیاتی، گروهی از پرداخت مالیات معاف می شوند و این امر به کاهش درآمدهای مالیاتی دولت منجر می شود. بنابراین، اگر اعطای معافیت ها با دقت لازم صورت نگیرد، دولت برای کسب درآمد مالیاتی مورد نظر، فشار سنگینی را بر گروه های دیگر جامعه وارد می کند. یعنی برای جبران بخش خالی کاسه درآمد مالیاتی، میزان نرخ های مالیاتی بیشتری، بر گروه هایی که از معافیت های مالیاتی بهره مند نیستند، تحمیل خواهد شد. بی تردید اعطای معافیت های مالیاتی گسترده، به ناکارایی نظام مالیاتی و ضعف نظام اجرایی مالیاتی منجر می شود [۸۷].

۵- تشخیص علی الرأس و وجود ضعف در اجرای آن [۷۸].  
۶- عدم گسترش فرهنگ مالیاتی در جامعه فرهنگ مالیاتی، مجموعه ای از طرز تلقی، بینش و عکس العمل افراد در قبال نظام مالیاتی است. به عبارت دیگر طرز تلقی، بینش، برداشت، آرمان ها، ارزش های اجتماعی، قوانین جاری و میزان تحصیلات و آگاهی، از جمله عواملی هستند که فرهنگ مالیاتی را شکل می دهند [۵۷]. نبود فرهنگ مالیاتی و آموزش آن در سطوح مختلف یکی از موارد مورد اشاره به عنوان فرار مالیاتی ذکر میشود [۷۷].

۷- عدم استقبال از تسلیم اظهارنامه های مالیاتی و ضعیف بودن ضمانت های اجرایی: ارسال نکردن اظهارنامه، دلایل گوناگونی دارد. رایج ترین دلیل عدم ارسال، می تواند این باشد که در صورت به روز نبودن ثبت مؤدی، بهترین امکان برای ارسال نکردن اظهارنامه و پس از آن فرار از پرداخت مالیات فراهم شده است. دلیل دیگر عدم برخورد قاطع ماموران مالیاتی است که آن هم می تواند به دلیل عدم آشنایی و عدم تسلط کافی ماموران مالیاتی به قوانین مالیاتی و نداشتن مهارت کافی در اجرای آنها باشد [۴۰].

-داده کاوی:

به مفهوم استخراج اطلاعات نهان یا الگوها و روابط مشخص در حجم زیادی از داده ها در یک یا چند بانک اطلاعاتی بزرگ گفته می شود. بسیاری از مردم داده کاوی را مترادف واژه های رایج کشف دانش از داده ها (KDD) می دانند. داده کاوی، پایگاه ها و مجموعه حجیم داده ها را در پی کشف و استخراج، مورد تحلیل قرار می دهد. این گونه مطالعات و کاوش ها را به واقع می توان همان امتداد و استمرار دانش کهن و همه جا گیر آمار دانست. تفاوت عمده در مقیاس، وسعت و گوناگونی زمینه ها و کاربردها، و نیز ابعاد و اندازه های داده های امروزی است که شیوه های ماشینی مربوط به یادگیری، مدل سازی، و آموزش را طلب می نماید [۴۸].

داده کاوی فرآیند کشف الگوها و روابط جالب از مقادیر زیاد داده است. منابع داده می تواند شامل پایگاه های داده، انبارهای داده، وب، سایر مخازن اطلاعات یا داده هایی باشد که به صورت پویا به سیستم منتقل می شوند [۷۷].

#### ۲-۲- پیشینه پژوهش های داخلی و خارجی

با توجه به گستردگی پژوهش های حوزه مالیات و همچنین اهمیت موضوع فرار مالیاتی در این زمینه پژوهش هایی صورت گرفته که از پژوهش های داخلی می توان به حمیدی و همکاران (۱۳۹۴) [۱۲] به بررسی جایگاه جرایم مالیاتی در جلوگیری از فرار در نظام مالیات بر ارزش افزوده (مطالعه موردی استان قزوین) پرداختند. رضاقلی زاده و همکاران (۱۳۹۸) [۲۰] به بررسی تحلیل فرار مالیاتی در ایران به روش شاخص چندگانه -علل چندگانه پرداختند، امیری (۱۳۹۷) [۳] بررسی رابطه بین اجتناب از پرداخت مالیات، ریسک مالیاتی و سررسید بدهی در شرکت های پذیرفته شده در بورس اوراق بهادار تهران، امیدی پور و همکاران (۱۳۹۶) [۱] به بررسی فرار مالیاتی در پایه مالیات بر درآمد اشخاص حقوقی در ایران (برآوردهای سالانه ۱۳۹۲-۱۳۵۲) پرداختند. دهقانی و همکاران (۱۳۹۷) [۱۷] به بررسی تئوری چشم انداز؛ رهیافتی نوین در توضیح پدیده فرار مالیاتی پرداختند. پناهی (۱۳۹۵) [۸] تشخیص فرار مالیاتی مؤدیان مالیات بر ارزش افزوده، امیری (۱۳۹۶) [۴] به بررسی اقتصاد رفتاری و فرار مالیاتی پرداخت. اغلب سیاست های مالیاتی مبتنی بر نحوه تصمیم گیری مالیات دهندگان بر اساس مدل های اقتصادی کلاسیک بنا نهاده شده است. اما بررسی ها نشان می دهد مدل های تصمیم گیری متعارف که فارغ از بنیان های روانشناسی - اجتماعی و تنها بر اساس مؤلفه های اقتصادی طراحی شده اند، نمی تواند تحولات و نحوه دقیق عملکرد تصمیم گیران را تبیین کنند. در زمینه راه کارهای جلوگیری از فرار مالیاتی می توان به پژوهش های رضایی سیابیدی (۱۳۹۶) [۲۱] که به بررسی راه های مقابله و جلوگیری از فرار مالیاتی پرداختند، تقوی فرد، رئیسی وانانی و پناهی (۱۳۹۶) [۱۰] به بررسی تحلیل آینده نگر تشخیص فرار مالیاتی مؤدیان مالیات بر ارزش افزوده با استفاده از الگوریتم های طبقه بندی و خوشه بندی، در زمینه مورد مطالعه در استان های داخلی می توان به پژوهش های کرمخانی و همکاران (۱۳۹۵) [۳۱] بررسی میزان کارایی جرایم مالیاتی در جلوگیری از فرار مالیاتی در نظام مالیات بر ارزش افزوده در بین مؤدیان مالیاتی استان ایلام، لاری دشت بیاض، قائم مقامی و کهرمی (۱۳۹۵) [۳۲] به بررسی عوامل مؤثر بر فرار مالیاتی در استان خراسان جنوبی با تأکید بر مؤلفه های فرهنگی پرداختند، فطرس و همکاران (۱۳۹۵) [۲۸] به بررسی اقتصاد زیرزمینی و فرار مالیاتی در چارچوب مدل های تعادل عمومی پویای تصادفی پرداختند. مداح و همکاران (۱۳۹۴) [۳۵] به بررسی عوامل مؤثر بر فرار مالیاتی (مطالعه موردی استان آذربایجان شرقی) پرداختند. نتایج این

پرداختند. فو-سنگ هونگ (۲۰۱۷) [۶۷] در مقاله خود به بررسی توضیح غیرخطی بودن تورم و رشد اقتصادی و نقش فرار مالیاتی پرداختند. این مقاله شامل فرار مالیاتی در یک مدل رشد درون زای ساده با پول در عملکرد اقتصادی است. وو و همکاران (۲۰۱۲) [۸۵] نیز با استفاده از داده کاوی درصد افزایش عملکرد کشف فرار مالیاتی برآمدند، آنها با استفاده از داده کاوی یک چارچوب نظارتی ارائه کردند تا مغایرت های بین گزارش های ارزش افزوده را که نیازمند حسابرسی بیشتری بود، مشخص نمایند.

مطالعات مختلف انجام شده در زمینه های مختلف فرار مالیاتی اکثراً به بررسی شرکت ها و سازمان های بوری پذیرفته شده می پردازند اما خلاء تحقیقاتی در زمینه مشاغل و طریقه محاسبه مالیاتی این گروه می باشد. هرساله مؤدیان این گروه باید اظهارنامه مالیاتی خود را ارسال کنند اما محاسبه مالیات مشاغل یکی از سوالات بسیار مهمی است که سازمان امور مالیاتی می بایست به آن پاسخ دهد تا به گونه ای محاسبات انجام پذیرد که فرار مالیاتی این گروه از مؤدیان به کمترین نرخ خود برسد. تاکنون پژوهش های اندکی در زمینه فرار مالیاتی در حوزه بررسی مشاغل داخل کشور انجام شده است و اغلب پژوهش های انجام شده در این زمینه با توجه به رویکردهای اقتصادی به پیش بینی فرار مالیاتی در شرکت های بوری پرداخته اند. با وجود این، حتی در پژوهش هایی که از منظر ارقام اقتصادی به مسئله فرار مالیاتی پرداخته اند، جای خالی استفاده از رویکردهای نوینی چون داده کاوی مشهود است. این مسئله بدان علت مهم است که در پژوهش های بسیاری سودمندی، کارایی و عملکرد بهتر تکنیک های داده کاوی در مقابل روش های خطی به وضوح مورد تأیید قرار گرفته است. امروز با توجه به پیشرفت تکنولوژی و فناوری، تولید انبوه داده ها، اطلاعات و الگوریتم ها در شبکه های مختلف اطلاعاتی و ارتباطی شکل گرفته است و لازمه پیشرفت و موفقیت در حال حاضر در سازمان ها و شرکت های مختلف و کسب و کارهای کوچک استفاده از این حجم انبوه دیتا و اطلاعات می باشد تا از طریق آنها بتوانند مشکلات سیستماتیک خود را متوجه شوند و با بهره گیری از آن ها نقاط ضعف خود را بهبود بخشند.

### ۳- روش پژوهش

این پژوهش از لحاظ هدف از نوع توصیفی-پیمایشی بوده و در آن جهت جمع آوری اطلاعات و داده های لازم، از پایگاه داده امور مالیاتی استان قزوین به عنوان یک مؤلفه ی بازبایی اطلاعات استفاده شده است. همچنین داده ها از چند منبع مختلف در ادارات مختلف استانی ترکیب شده و با استفاده از روند استخراج، پالایش، بارگذاری (ETL) و یکپارچه شده اند. بدین منظور این پژوهش در دو فاز اساسی صورت گرفت. در فاز اول این پژوهش ابتدا ویژگی های مختلف را برای فرار مالیاتی مشاغل تعریف نموده تا از طریق آن و با استفاده از روش های داده کاوی به یک

پژوهش نشان داد مالیات از اهم درآمدهای دولت است که پذیرش آن از طرف جامعه هدف و نحوه اجرای آن از طرف دولت، به ویژگی های فرهنگی هر کشور بستگی دارد و هر یک از عوامل مثل بار مالیاتی، اندازه دولت و نیز شدت قوانین حوزه مالیات و نرخ بیکاری و تورم نیز از جمله موارد مهم در زمینه پذیرش مالیات و عدم ورود به اقتصاد زیر زمینی می باشند که هر کدام تأثیر خاص خود را بر روی فرار مالیاتی می گذارند.

در زمینه پژوهش های خارج کشور که در سالیان اخیر به چاپ رسیده است می توان به مطالعه لیوهونگ (۲۰۲۲) به بررسی مدیریت و جمع آوری مالیات و بررسی مسائل حقوقی در تجزیه و تحلیل کلان داده های مالیاتی می پردازد [۷۳]، در مطالعه دیگر هوآنگ و همکاران (۲۰۲۱) به بررسی تأثیر اجتناب مالیاتی و بار اضافی مالیات بر درآمد پرداختند [۶۷]. مطالعات ساندر خلیل و همکاران (۲۰۲۰) [۶۹]، بررسی تأثیر دین داری بر نگرش های فرار مالیاتی در کشور لبنان، حوزه اورازو (۲۰۲۰) [۴۷]، بررسی رویکرد کسب و کارهای مرتبط با جایگزین واردات و فرار مالیاتی، جیانفئی روان و همکاران (۲۰۱۹) [۷۸] در مقاله خود به بررسی شناسایی گروه های مشکوک فرار مالیاتی مبتنی بر معاملات وابسته در داده های بزرگ پرداختند، دبور دی جیواچینو و همکاران (۲۰۲۰) [۵۸] در بررسی فرار مالیاتی و روحیه مالیاتی پرداختند. اثرات اخلاقی مالیات و هنجارهای اجتماعی بر فرار مالیاتی را در هنگام تعامل افراد در یک شبکه بررسی شده است. الگویی ارائه شده است که شامل انگیزه هایی برای رعایت مالیات به صورت مجازات و جریمه، روحیه مالیاتی و شهرت برای رفتار اجتماعی است. همچنین بانو دمیر و همکاران (۲۰۲۰) [۵۵] در پژوهش های خود به بررسی تغییرات سیاست تجاری، فرار مالیاتی و قانون بنفورد، والتر دیدیمو و همکاران (۲۰۱۸) [۵۷] بررسی یک سیستم تجزیه و تحلیل بصری برای پشتیبانی از کشف فرار مالیاتی پرداختند. این پژوهش، یک سیستم پشتیبانی تصمیم برای کشف فرار مالیاتی، مبتنی بر یک زبان بصری قدرتمند و تکنیک های پیشرفته تجسم شبکه را توصیف می کند. با همکاری آژانس درآمد ایتالیا که در حال حاضر از آن استفاده می شود، توسعه یافته است. احمد علی و همکاران (۲۰۱۸) [۴۱]. آستادستر و همکاران (۲۰۱۹) در مقاله ای تحت عنوان فرار مالیاتی و عدالت به برآورد مدل های متعارف پیش بینی فرار مالیاتی افراد ثروت مند می پردازد که هر چقدر ثروت مندان از پرداخت مالیات خود طفره برونند عدالت در آن جامعه تضعیف می گردد [۴۵]. در پژوهش خود تحت عنوان تأثیر کیفیت محیط سازمانی در فرار از پرداخت مالیات ها با هدف مقایسه تأثیر کیفیت محیط سازمانی (IEQ) بر فرار مالیاتی در کشورهای عضو قدیمی (قبل از ۲۰۰۴) و جدید (پس از ۲۰۰۴) اتحادیه اروپا توسط شاخص های مدیریت جهانی بانک جهانی اندازه گیری صورت پذیرفت. لومیر عبدیشیکو (۲۰۱۷) [۴۰] در مقاله خود به بررسی عوامل تعیین کننده سطح فرار مالیاتی در اقتصادهای در حال گذار

- مالیات ابرازی (میزان مالیات اعلامی توسط مؤدی بر اساس ریتال (
- مالیات قطعی (محاسبه مالیات واقعی بر اساس ریتال)
- مبلغ فرار مالیاتی (بر حسب ریتال)
- طبقه بندی فرار مالیاتی (براساس دو کلاس پلی و خیر)

در این پژوهش پرونده های مالیاتی مرتبط با اصناف و مشاغل در پنج گروه مالیاتی به عنوان داده های نمونه انتخاب گردید. این پنج گروه عبارت است از :

- G1: گروه صنفی وکلا و کارشناسان رسمی دادگستری و مشاورین حقوقی و دفاتر ازدواج و طلاق و مشاغل وابسته
- G2: گروه صنفی بساز بفروشان، مصالح ساختمانی، مهندسی ساختمان و سایر مشاغل وابسته
- G3: گروه صنفی فست فود ها، هتل ها و متل ها، رستوران ها و چلو کبابی ها و مشاغل وابسته
- G4: گروه صنفی دفاتر پلیس+۱۰، دفاتر خدماتی ارتباطی دولت، ITC، فروشندگان تلفن همراه و سیم کارت و مشاغل وابسته
- G5: گروه صنفی فروشندگان لوازم یدکی خودرو سبک و سنگین، تعمیرگاه خودرو و نمایندگی مجاز فروش خودرو و مشاغل وابسته

-پاک سازی داده ها :

برای انجام یک داده کاوی مطلوب لازم است مقادیر گم شده جایگزین شوند. داده های مزاحم شناسایی و به نحوی مناسب با آنها برخورد و ناپایداری ها اصلاح شوند. جهت آماده سازی داده ها جهت مدل سازی نیاز است پاک سازی داده ها صورت پذیرد.

-انتخاب پایگاه داده مناسب:

بر اساس مراحل قبل ۶۴۰۰ پرونده مالیاتی مرتبط با مشاغل در قالب یک دیتاست یکپارچه و تجمیع شده و انتخاب گردید.

-انتخاب زیر مجموعه ای از ویژگی ها:

ویژگی های غیر مرتبط، ویژگی هایی که شامل هیچ گونه اطلاعات مفیدی برای انجام عملیات داده کاوی نیستند، ویژگی های زائد، حاوی اطلاعاتی هستند که همه یا بخش زیادی از آن در یک یا چند ویژگی دیگر تکرار شده باشد. در این گام از بین ویژگی های شناسایی شده در مرحله تجمیع داده ها ویژگی انتخاب گردید. لازم به ذکر است کد مالیاتی، کد ملی،

تقسیم بندی کلی در زمینه نمونه های آماری بدست آوریم. در فاز دوم پژوهش برای تقسیم بندی مشاغل به دسته های مختلف از بانک اطلاعات سازمان امور مالیاتی استان قزوین استفاده گردید و سپس با توجه به آن تحلیل های مختلف صورت گرفت.

این پژوهش شامل بررسی کلیه اصناف و مشاغلی می باشند که دارای کد مالیاتی در استان قزوین هستند که بالغ بر ۶۴۰۰ پرونده مالیاتی بین سال های ۹۳ تا ۹۸ مورد تحلیل قرار گرفت. با توجه به این که از روش داده کاوی استفاده می نماییم و امکان تجزیه و تحلیل حجم زیادی از داده ها وجود دارد، نمونه گیری صورت نخواهد گرفت و تمام داده های موجود در بانک اطلاعات جهت تحلیل بکار خواهد رفت. چهار مرحله در پیش پردازش داده در داده کاوی عبارت است از

- یکپارچه سازی داده (Data Integration)
- تجمیع داده (Data Aggregation)
- نمونه گیری (Data Sampling)
- پاک سازی داده (Data Cleaning)

- انتخاب زیر مجموعه ای از ویژگی ها (Feature Subset selection)

-تغییر شکل متغیرها (Variable Transformation)

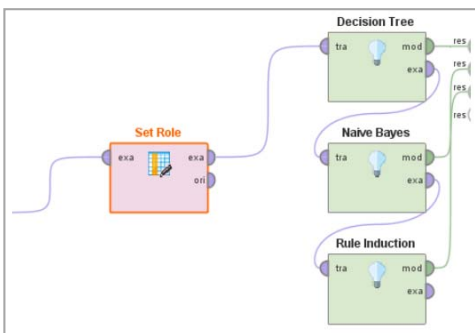
لازم به ذکر است مرحله ۵ و ۶ را می توان جزء فرآیندهای مرحله پاک سازی داده ها نیز در نظر گرفت. در این پژوهش از روش اتصال محکم<sup>۱</sup> جهت یکپارچه سازی داده ها استفاده شده است. بدین صورت که از پایگاه داده امور مالیاتی استان قزوین به عنوان یک مولفه ی بازایی اطلاعات استفاده شده است. همچنین داده ها از چند منبع مختلف در ادارات مختلف استانی ترکیب شده و با استفاده از روند استخراج، پالایش، بارگذاری (ETL) و یکپارچه شده اند.

جهت هم مقیاس سازی داده به دلیل آنکه از پایگاه های داده مختلف انتخاب شده بودند ابتدا متغیرهای تحلیلی را انتخاب سپس آنها را هم مقیاس نمودیم. تعریف ویژگی های اولیه پژوهش بر اساس پایگاه داده در سازمان عبارت است از:

- واحد مالیاتی ( بر اساس تعداد پرونده مالیاتی)
- سابقه فعالیت (سال)
- جنسیت (زن و مرد)
- کد ملی (بر اساس کد ۱۰ رقمی)
- تسلیم اظهارنامه ( بر اساس صفر و یک می باشد)
- نوع اتحادیه ( بر اساس موضوع فعالیت به ۱۵ نوع تقسیم می شود)

<sup>1</sup> Tight Coupling

پس از مرحله آماده سازی و پیش پردازش داده های فرار مالیاتی مشاغل، در این تحقیق از مدل کلاس بندی شامل درخت تصمیم، قوانین انجمنی و قوانین بیز استفاده شده است. در شکل زیر نحوه اتصال دیتاست آماده شده به مدل مشاهده می گردد.



شکل (۱): نحوه پردازش دیتاست فرار مالیاتی مشاغل

در جدول ۱ و ۲ پارامترهای این تحقیق که بر از روش کلاسه بندی قوانین انجمنی و درخت تصمیم انجام شده را مشاهده می کنید.

مالیات تشخیصی ویژگی هایی بودن که در فرآیند مدل سازی تأثیر گذار نمی باشند، لذا حذف گردیدند.

- جایگزین کردن داده های گم شده:

مطابق تحلیل انجام شده، درآمد ابرازی شامل ۶۷ داده گم شده ، مالیات ابرازی ۲ داده گم شده و فرار مالیاتی ۵۱ داده گم شده می باشند. لذا در این گام داده های گم شده را با میانگین داده ها جایگزین گردید.

- فیلتر کردن داده ها :

در این گام داده های مورد تحلیل را بر اساس گروه های صنفی منتخب پایش می نماییم . به عنوان نمونه در تحلیل داده های گروه صنفی صاحبان دفاتر رسمی و حقوقی و مشاغل وابسته مشاهده می شود. گروه صنفی G1 که سابقه فعالیت بالای ۱۴ سال را دارند فرار مالیاتی ثبت نشده است.

- مشخص کردن برجسب/لیبل :

در این گام می بایست یکی از ویژگی های موجود در دیتاست را به عنوان متغیر پاسخ انتخاب نماییم. لذا ویژگی فرار مالیاتی را بر اساس دو کلاس فرار مالیاتی دارد (Yes)، فرار مالیاتی ندارد (No) انتخاب می نماییم.

#### ۴-مدل سازی و تحلیل داده ها

جدول (۱): پارامترهای قوانین انجمنی فرار مالیاتی مشاغل

پارامتر لاتین	پارامتر	مقدار	توضیحات
Sample ratio	نسبت نمونه	۰,۹	این پارامتر نسبت نمونه داده های آموزشی مورد استفاده برای رشد و هرس را مشخص می کند.
Pureness	خلوص	۰,۹	این پارامتر خلوص مورد نظر را مشخص می کند ، یعنی حداقل نسبت کلاس اصلی در یک زیرمجموعه تحت پوشش به منظور خالص بودن زیر مجموعه.
Minimal prune benefit	حداقل سود هرس	۰,۲۵	این پارامتر حداقل میزان سود را مشخص می کند که برای هرس شدن باید بیش از سود منقضی نشده باشد.

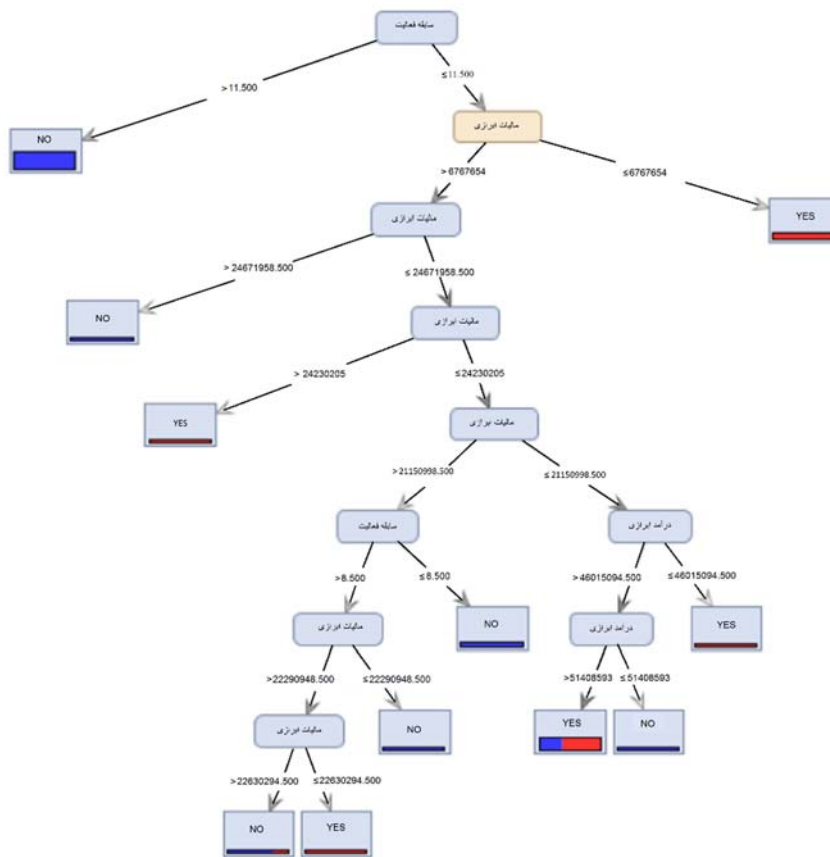
جدول (۲): پارامترهای درخت تصمیم فرار مالیاتی مشاغل

پارامتر	مقدار	توضیحات
حداکثر عمق	۲۰	عمق یک درخت بسته به اندازه و ماهیت مجموعه نمونه متفاوت است. این پارامتر برای محدود کردن اندازه درخت تصمیم استفاده می شود.
میزان اطمینان	۰,۲۵	این پارامتر میزان اطمینان مورد استفاده برای محاسبه خطای بد بینانه هرس را مشخص می کند.
حداقل سود	۰,۱	سود یک گره قبل از تقسیم آن محاسبه می شود.
حداقل اندازه برگ	۲	درخت به گونه ای تولید می شود که هر زیر مجموعه برگ حداقل تعداد نمونه برگ حداقل را داشته باشد.

حدافل اندازه برای تقسیم	۴	فقط آن گره ها تقسیم می شوند که اندازه آنها بزرگتر یا مساوی حدافل اندازه پارامتر split باشد.
تعداد گزینه های پیش از هرس	۳	این پارامتر تعداد گره های جایگزین را که برای تقسیم انتخاب می شود را تنظیم می کند.

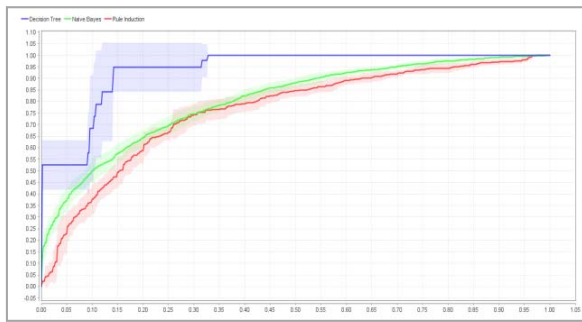
ساختار درخت تصمیم، پیش بینی به دست آمده از درخت در قالب یک سری قواعد توضیح داده میشود. هر مسیر از ریشه تا یک برگ درخت تصمیم، یک قانون را بیان میکند و در نهایت برگ با کلاسی که بیشترین مقدار رکورد در آن تعلق گرفته برچسب می خورد. در شکل زیر درخت تصمیم فرار مالیاتی مشاغل مشاهده می شود. در شکل زیر درخت تصمیم استخراج شده مشاهده می شود:

یکی از پرکاربردترین الگوریتم های داده کاوی، الگوریتم درخت تصمیم است. در داده کاوی، درخت تصمیم یک مدل پیش بینی کننده است به طوری که می تواند برای هر دو مدل رگرسیون و طبقه ای مورد استفاده قرار گیرد. زمانی که درخت برای کارهای طبقه بندی استفاده می شود، به عنوان درخت طبقه بندی شناخته می شود و هنگامی که برای فعالیت های رگرسیونی به کار می رود درخت رگرسیون نامیده می شود. در



شکل (۲): درخت تصمیم فرار مالیاتی مشاغل

۴-۱- انتخاب مدل مناسب جهت داده کاوی

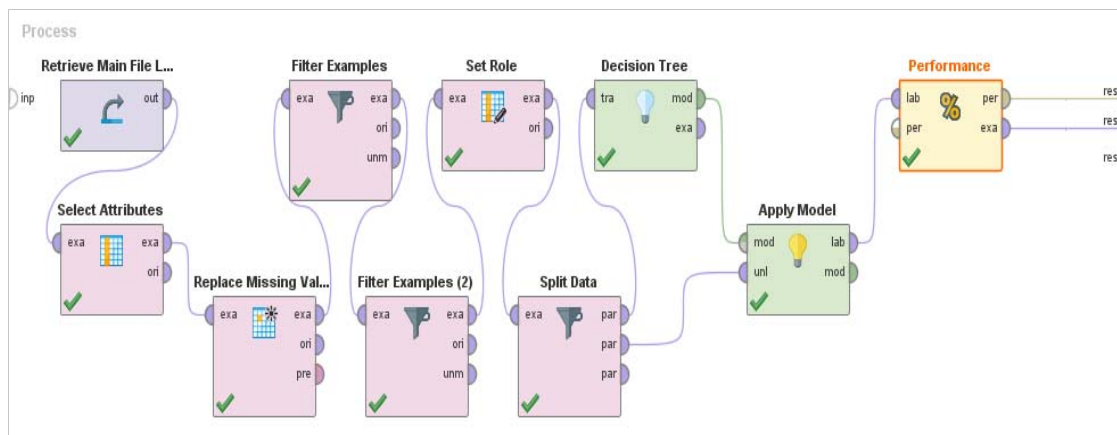


شکل (۳) : رتبه بندی مدل های طبقه بندی داده کاوی فرار مالیاتی مشاغل

۴-۲- بررسی عملکرد مدل پیش بینی فرار مالیاتی مشاغل

در این قسمت با توجه به انتخاب الگوریتم مدل ساز درخت، مدل پیش بینی را طراحی می نمایم. مدل پیش بینی بر اساس داده های آموزشی، دیتاست جدید را برای پیش بینی بدست می آوریم. هدف این می باشد بر اساس داده های آموزشی و یادگیری ماشین، مدل بتواند رفتار داده های جدید را پیش بینی نماید. پس از تفکیک داده ها به داده های آموزشی و داده های تست می توانیم، صحت و دقت عملکرد مدل پیش بینی را اندازه گیری نماییم. بر اساس پیش مدل سازی انجام شده، نسبت به پاک سازی داده ها اقدام نموده و سپس داده ها را با نسبت معمول ۷۰ به ۳۰ به داده های آموزشی و داده های تست تقسیم بندی می نماییم. این بدین معنا می باشد که ۷۰ درصد داده ها، یادگیری الگوریتم را انجام می دهند و ۳۰ درصد از آنها الگوریتم ارائه شده را تست نمایند. در شکل زیر طراحی مدل ساز فرار مالیاتی مشاغل را که در نرم افزار Rapid Miner دیتا ها وارد شده که بر مبنای آن مدل سازی و تحلیل داده انجام می شود.

جهت بررسی عملکرد مدل های داده کاوی از نمودار های ROC استفاده می شود. این نمودارها برای اندازه گیری کردن دیداری مدل ها بکار برده می شود. منحنی های ROC میزان عملکرد خوب مدل های باینری یادگیری ماشین را نشان می دهند. که این مقدار را با دو نرخ مثبت صحیح و نرخ مثبت غلط با بازه های اطمینان مختلف بدست می آورد. در واقع میزان True Positive Rate را در مقابل False Positive Rate واقع می گیریم و ترسیم می کند. در نهایت خط مستقیم رسم می شود که هرچه به سمت بالا و گوشه سمت چپ نزدیک تر باشد مدل مناسب تری محسوب می شود. با این نمودار مشخص می شود کدام مدل عملکرد مناسب تری را دارد. برای بررسی رفتار داده های جدید از مدل های پیش بینی استفاده می نماییم و سپس اعتبارسنجی مدل مورد نظر را اندازه گیری نماییم. هدف در یادگیری ماشین بررسی داده های بزرگ جهت کشف الگوهای پیش بینی رفتار داده ها می باشد. پیش بینی به دو صورت قیاسی و عددی انجام می شود. که پیش بینی قیاسی شامل دسته بندی و پیش بینی عددی شامل رگرسیون می باشد. در اینجا از مدل طبقه بندی داده ها شامل درخت تصمیم، قوانین انجمنی و قانون بیز استفاده شده است. بر اساس تحلیل شکل فوق و ترسیم خط مستقیم مشاهده می شود مدل درخت تصمیم بر اساس داده های موجود مدل مناسب تری محسوب می شود و عملکرد مناسب تری جهت پیش بینی رفتار الگوی فرار مالیاتی را دارد. نتایج نشان می دهد مدل درخت تصمیم عملکرد مناسب تری نسبت به دو مدل قوانین انجمنی و قانون بیز را دارد.



شکل (۴) : طراحی مدل ساز فرار مالیاتی مشاغل



می‌شود. به معیار پوشش، معیار حساسیت<sup>۲</sup> نیز گفته می‌شود.

• معیار کاپا (Kappa) به می‌تواند عملکرد الگوریتم‌های چند کلاسه را ارزیابی کند. این معیار قادر است که تعداد نمونه‌ها را در یک مجموعه‌ی داده‌ی نامتوازن نیز در نظر بگیرد و کیفیت الگوریتم را با توجه به تعداد نمونه‌ها بیان کند. پس برای مسائلی با طبقه‌های مختلف و همچنین مسائلی که داده‌های نامتوازن دارند، می‌توانید از این معیار استفاده نماییم. معیار کاپا مدل برابر ۰,۶۱۲ می‌باشد که عملکرد خوب (بین ۰,۶ تا ۰,۸) مدل ساز را نشان می‌دهد.

در جدول زیر مقادیر پارامترهای تحلیل حساسیت مدل مشاهده می‌شود.

جدول (۳): پارامترهای تحلیل حساسیت مدل

پارامتر لاتین	پارامتر	مقدار	توضیحات
Accuracy	دقت	66.83%	درصد پیش بینی های صحیح
Kappa	ضریب کاپا	0.612	ارزیابی کیفیت الگوریتم‌های طبقه بندی چند کلاسه
Weighted Mean Recall	ضریب پوشش میانگین وزنی	68.00%	محاسبه‌ی پوشش بر روی کل داده ها

۴-۳/ اعتبارسنجی مدل بر اساس داده های تست (آزمون)

برای اعتبارسنجی مدل، داده‌ها را به دو بخش داده‌های آموزش و تست به نسبت ۷۰ به ۳۰ تقسیم تقسیم بندی شد. اما نمی‌دانیم داده‌هایی که برای تست استفاده شده است چقدر شبیه داده‌ی می‌باشند که برای آموزش و ساخت مدل استفاده شده‌اند. برای این منظور از تکنیک Cross Validation، صحت اعتبارسنجی را مورد آزمون قرار دهیم تا با اطمینان بیشتری بتوانیم درصد عملکرد مدل سازی را تخمین بزنیم و مطمئن شویم داده‌هایی که برای تست استفاده شده است مشابه داده‌های آموزشی می‌باشند. در جدول زیر صحت اعتبارسنجی مدل پیش بینی را با استفاده از تکنیک Cross Validation، مورد آزمون قرار داده شده است. تا با اطمینان بیشتری بتوانیم درصد عملکرد مدل سازی را تخمین بزنیم.

باتوجه به مدل سازی صورت گرفته هدف بعدی دستیابی به بالاترین دقت و صحت ممکن و خروجی مدل سازی می‌باشد. در برخی از مسائل، تشخیص صحیح نمونه‌های مربوط به یکی از دسته‌ها برای ما اهمیت بیشتری دارد. در تحقیق مذکور هدف شناسایی افرادی است که فرار مالیاتی دارند. برای درستی عملکرد مدل از آزمون‌های مختلف عملگری استفاده می‌نماییم. یکی از این آزمون‌ها جهت صحت مدل در حوزه هوش مصنوعی، ماتریس آشفتگی<sup>۱</sup> می‌باشد که در آن عملکرد مدل را نشان می‌دهند. بر اساس این ماتریس مشاهده می‌شود، مدل ساز بر اساس دو کلاس فرار مالیاتی دارد (YES) و فرار مالیاتی ندارد (NO)، ۸۰۰ نمونه که در کلاس فرار مالیاتی ندارد (NO) را درست پیش بینی کرده است. لذا ۶۰ درصد پیش بینی در کلاس فرار مالیاتی ندارد (NO) را مدل ساز به درستی پیش بینی کرده است. که نشان از صحت و دقت بالا در پیش بینی دارد. همچنین ۴۷۸ نمونه که در کلاس فرار مالیاتی دارد (YES) را درست پیش بینی کرده است. لذا ۸۳ درصد پیش بینی در کلاس فرار مالیاتی دارد (YES) را مدل ساز به درستی پیش بینی کرده است که نشان از صحت و دقت بالا در پیش بینی دارد.

۴-۳ آنالیز حساسیت عملکرد مدل پیش بینی: ارزیابی عملکرد الگوریتم‌های شرح داده شده در بالا، با استفاده از معیارهای متفاوتی بر مبنای دیدگاه حساسیت و تشخیص صورت گرفته شده است. حساسیت و تشخیص دو شاخص برای ارزیابی نتیجه یک دسته‌بندی دودویی (دو حالته) هستند. زمانی که بتوان داده‌ها را به دو گروه مثبت و منفی تقسیم کرد، دقت نتایج یک آزمایش که اطلاعات را به این دو دسته تقسیم می‌کند با استفاده از شاخص‌های حساسیت و ویژگی قابل اندازه‌گیری و توصیف است.

• معیار صحت (Accuracy) نسبت مقداری موارد صحیح طبقه‌بندی شده توسط الگوریتم از یک کلاس مشخص، به کل تعداد مواردی که الگوریتم چه به صورت صحیح و چه به صورت غلط، در آن کلاس طبقه‌بندی کرده است را محاسبه می‌کند. که در این پژوهش معیار صحت برابر ۶۷ درصد می‌باشد.

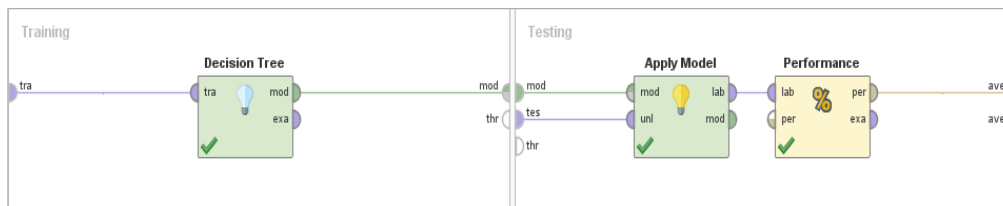
• معیار کاربردی دیگری به نام پوشش (Recall) نیز وجود دارد و همان‌طور که از نامش پیداست، به دنبال محاسبه‌ی پوشش بر روی کل داده‌هاست. همان‌طور که مشاهده می‌کنید، تمرکز اصلی معیار پوشش بر خلاف معیار صحت بر روی داده‌هایی است که واقعاً «YES» بوده‌اند. که در این پژوهش معیار پوشش برابر ۶۸ درصد می‌باشد. در برخی از مسائل، معیار پوشش مهم‌تر از معیار صحت

<sup>2</sup> Sensitivity

<sup>1</sup> Confusion Matrix

جدول (۴): اعتبارسنجی براساس داده های آزمون

Accuracy: 67.79% +/- 1.99%			
	پیش بینی درست (NO)	پیش بینی درست (Yes)	پیش بینی کلاس
کلاس فرار مالیاتی ندارد (NO)	2568	1596	61.67%
کلاس فرار مالیاتی دارد (Yes)	464	1768	79.20%
Recall	84.69%	52.56%	

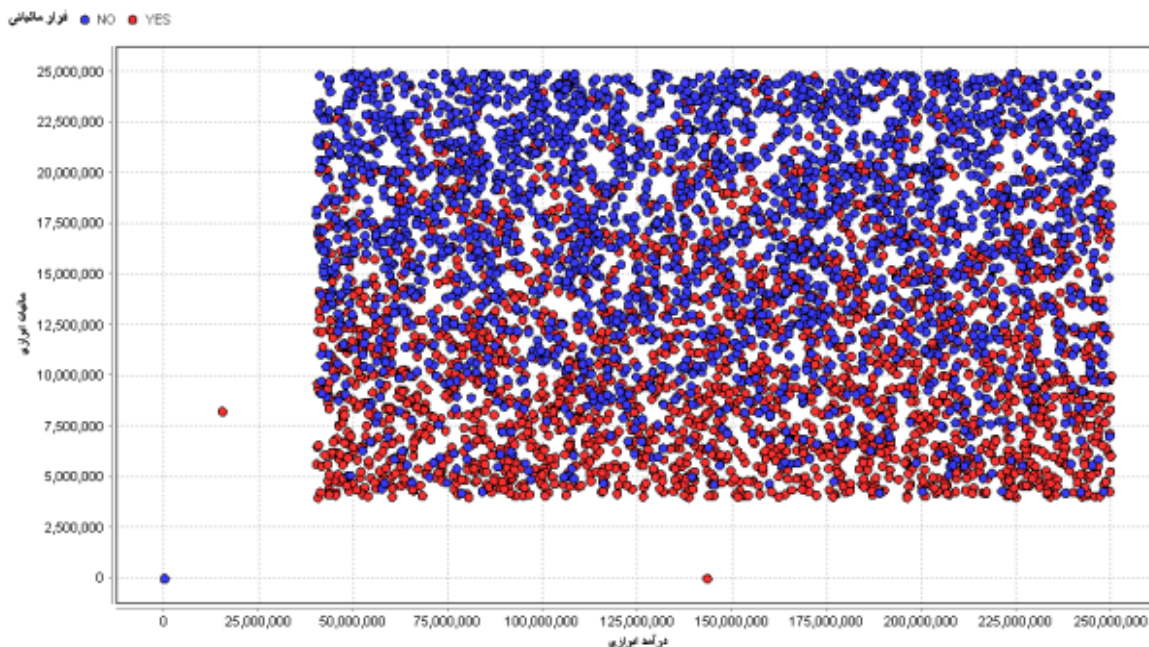


شکل (۵): طراحی مدل ساز بر اساس اعتبارسنجی Cross Validation

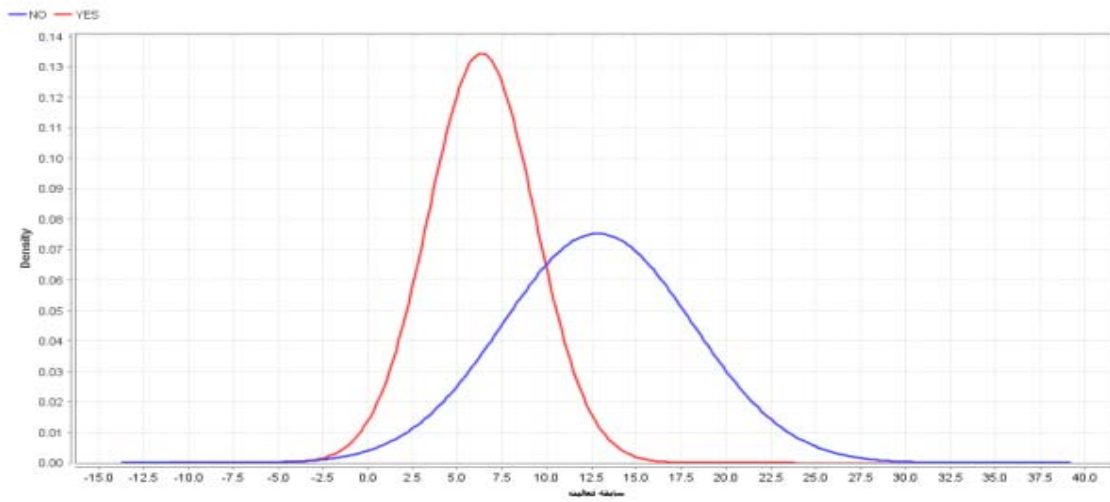
۴-۴-تحلیل نتایج

بر اساس خروجی داده کاوی پرونده های مالیاتی مشاغل مطابق شکل زیر مشاهده می شود میانگین مالیات ابرازی مشاغلی که فرار مالیاتی داشته اند از میانگین مالیات ابرازی مشاغلی که فرار مالیاتی نداشته اند کمتر می باشد.

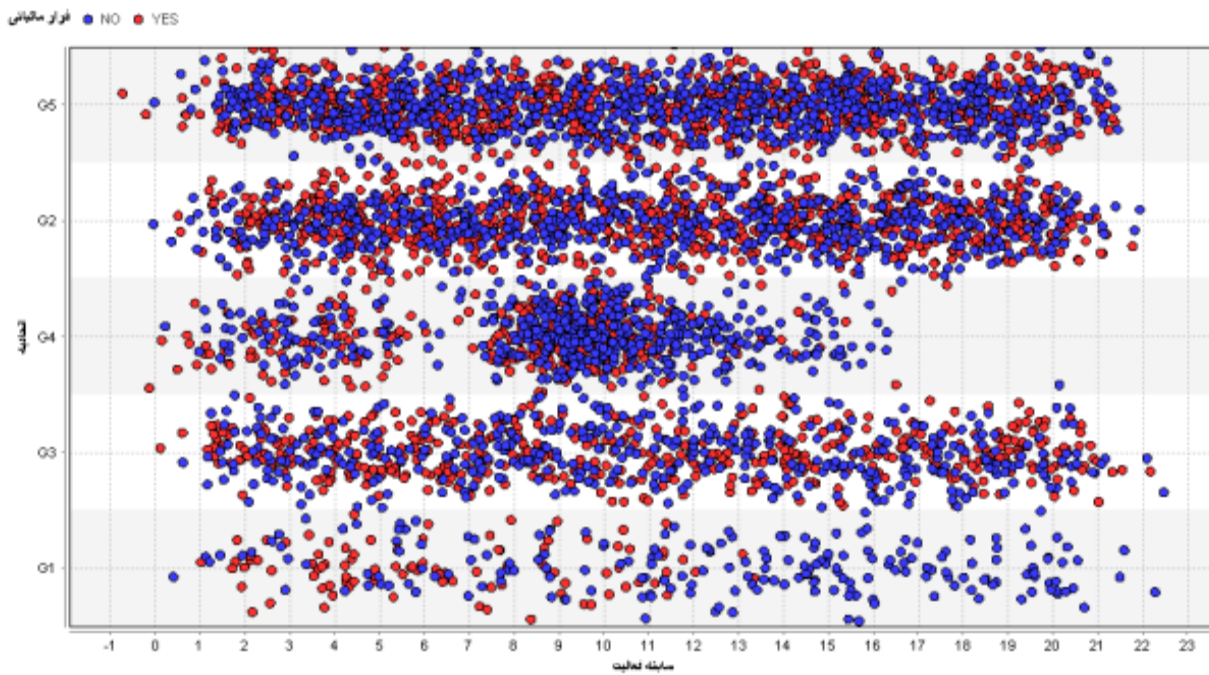
در بحث معیار صحت، معیار صحت (Accuracy) برابر  $67.79\% \pm 1.99\%$  نشان از قابلیت اطمینان و اعتبار مناسب جهت مدل پیش بینی می باشد. به دلیل تکرارهای مختلف جهت انتخاب داده های آموزشی  $67.79\% \pm 1.99\%$  انحراف معیار محاسبه شده است.



شکل (۶): بررسی درآمد ابرازی و مالیات ابرازی مشاغل دارای فرار مالیاتی



شکل (۷): بررسی سابقه فعالیت اصناف دارای فرار مالیاتی



شکل (۸): خوشه بندی فرار مالیاتی مشاغل بر اساس ویژگی سابقه فعالیت

مالیاتی داشته اند (۶ سال) از میانگین مالیات ابرازی مشاغل که فرار مالیاتی نداشته اند (۹ سال) کمتر می باشد. همچنین بر اساس خوشه بندی مشاغل (شکل ۸) مورد مطالعه مشاهده می شود:

یکپارچه سازی داده، تجمیع داده، نمونه گیری، پاک سازی داده ها، انتخاب زیر مجموعه ای از ویژگی ها و تغییر شکل متغیرها صورت پذیرفت. بر اساس تجزیه و تحلیل داده ها نتایج نشان می دهد، مدل درخت تصمیم بر اساس داده های موجود مدل مناسبت محسوب می شود و عملکرد مناسبت جهت پیش بینی رفتار الگوی فرار مالیاتی را دارد. بر اساس ماتریس آشفتگی مدل ساز بر اساس دو کلاس فرار مالیاتی دارد، ۸۰۰ نمونه که در کلاس فرار مالیاتی ندارد (NO) را درست پیش بینی نمود. لذا ۶۰٪ پیش بینی در کلاس فرار مالیاتی ندارد (NO) را مدل ساز به درستی پیش بینی کرده است که نشان از صحت و دقت بالا در پیش بینی دارد. همچنین ۴۷۸ نمونه که در کلاس فرار مالیاتی دارد (YES) را درست پیش بینی کرده است. لذا ۸۳٪ درصد پیش بینی در کلاس فرار مالیاتی دارد (YES) را مدل ساز به درستی پیش بینی کرده است که نشان از صحت و دقت بالا در پیش بینی دارد. نتایج این پژوهش در بخش تجزیه و تحلیل داده نشان می دهد مدل درخت تصمیم عملکرد مناسبت را دارد. لذا به مدیران امور مالیاتی استان پیشنهاد می شود، کمیته داده کاوی در ادارات مختلف استان تشکیل و الگوی پیشنهادی را در دستور کار خود قرار نمایند. نتایج داده کاوی نشان می دهد معیار صحت، معیار پوشش و معیار کاپای، عملکرد خوب مدل پیشنهادی را مورد تایید قرار داد، لذا نتایج این پژوهش اعتبار لازم جهت تصمیم گیری مدیران امور مالیاتی استان در مورد ایجاد مکانیزم اجرایی کنترل فرار مالیاتی پنج گروه مشاغل مورد مطالعه را دارد. نتایج داده کاوی نشان می دهد میانگین مالیات ابرازی مشاغلی که فرار مالیاتی داشته اند از میانگین مالیات ابرازی مشاغلی که فرار مالیاتی نداشته اند کمتر می باشد. لذا این امر می تواند به تصمیم گیری ارزیابان امور مالیاتی این کمک را نماید که در مورد مشاغلی که مالیات ابرازی کمتر از میانگین محاسبه شده را اعلام می نمایند، حساسیت و کنترل های بیشتری نسبت به فرار مالیاتی این گروه از مشاغل داشته باشند. نتایج داده کاوی نشان می دهد میانگین سابقه فعالیت مشاغلی که فرار مالیاتی داشته اند از میانگین مالیات ابرازی مشاغلی که فرار مالیاتی نداشته اند کمتر می باشد. این امر به تصمیم گیری ارزیابان امور مالیاتی این پیشنهاد را می دهد که مشاغلی که کمتر از ۶ سال از سابقه فعالیت آنها نمی گذرد، حساسیت بیشتری نسبت به فرار مالیاتی این گروه از مشاغل را داشته باشند. نتایج داده کاوی نشان می دهد بیش از ۳۲ درصد از مشاغلی که مالیات ابرازی خود را اعلام نموده اند فرار مالیاتی داشته اند. پیشنهاد می گردد که از تمامی پایگاه های اطلاعاتی در اختیار دولت از قبیل حساب بانکی، درگاه های پرداخت الکترونیک (پوز و کش ایس) برای کشف درآمد واقعی و جلوگیری از فرار مشاغل استفاده نمود. همچنین پیشنهاد می گردد با استفاده از ابزار هوش مصنوعی و علم داده سازمان امور مالیاتی تشخیص مالیات مشاغل را بر مبنای واقعیت با استفاده از مشاوره

در تمامی خوشه ها فرار مالیاتی وجود دارد ولی نسبت فرار مالیاتی به کل ۳۲٫۵ درصد از کل می باشد.

- بر اساس داده های مورد تحلیل مشاهده می شود، گروه صنفی وکلا و کارشناسان رسمی دادگستری و مشاورین حقوقی و دفاتر ازدواج و طلاق و مشاغل وابسته (G1) فرار مالیاتی کمتری نسبت به سایر گروه های صنفی دیگر دارد. همچنین گروه صنفی صنفی فروشندگان لوازم یدکی خودرو سبک و سنگین، تعمیرگاه خودرو و نمایندگی مجاز فروش خودرو و مشاغل وابسته (G5) فرار مالیاتی بیشتری نسبت به سایر خوشه های صنفی دیگر را دارد.

- بیشتر نمونه های مورد تحلیل در گروه صنفی فروشندگان لوازم یدکی خودرو سبک و سنگین، تعمیرگاه خودرو و نمایندگی مجاز فروش خودرو و مشاغل وابسته (G5) (۹۶۱ تعداد نمونه) می باشند و کمترین گروه صنفی وکلا و کارشناسان رسمی دادگستری و مشاورین حقوقی و دفاتر ازدواج و طلاق و مشاغل وابسته (G1) (۳۴۸ نمونه) می باشد.

- در گروه صنفی وکلا و کارشناسان رسمی دادگستری و مشاورین حقوقی و دفاتر ازدواج و طلاق و مشاغل وابسته (G1) بیشتر فرار مالیاتی مربوط به صنوفی می باشد که سابقه فعالیت کمتر ۱۱ سال داشته اند.

#### ۵- نتیجه گیری

در این پژوهش یک الگوی پیش بینی فرار مالیاتی مشاغل مبتنی بر تکنیک های داده کاوی ارائه گردید، که بر اساس جستجوی انجام شده و خلاء های پژوهشات داخل و خارج کشور تا کنون پژوهشی در این زمینه انجام نشده بود و این پژوهش برای اولین بار در حوزه امور مالیاتی استان قزوین اجرا گردید. این پژوهش شامل بررسی کلیه صنوف و مشاغلی که دارای کد مالیاتی در استان قزوین بین سال های ۹۳ تا ۹۸ مورد تحلیل قرار گرفت. در این پژوهش پرونده مالیاتی مرتبط با مشاغل در پنج گروه مالیاتی به عنوان داده های نمونه انتخاب گردید این پنج گروه عبارت است از گروه صنفی وکلا و کارشناسان رسمی دادگستری و مشاورین حقوقی و دفاتر ازدواج و طلاق و مشاغل وابسته، گروه صنفی بساز بفروشان، مصالح ساختمانی، مهندسین ساختمان و سایر مشاغل وابسته، گروه صنفی فست فود ها، هتل ها و متل ها، رستوران ها و چلو کبابی ها و مشاغل وابسته، گروه صنفی دفاتر پلیس +۱۰، دفاتر خدماتی ارتباطی دولت، ITC، فروشندگان تلفن همراه و سیم کارت و مشاغل وابسته، گروه صنفی فروشندگان لوازم یدکی خودرو سبک و سنگین، تعمیرگاه خودرو و نمایندگی مجاز فروش خودرو و مشاغل وابسته می باشند. شش مرحله در پیش پردازش داده در فرآیند داده کاوی شامل

داده سازمان امور مالیاتی تشخیص مالیات اصناف را بر مبنی واقعیت با استفاده از مشاوره شرکت های دانش بنیان انجام دهد. پیشنهاد می گردد نسبت به تفکیک حساب های تجاری و غیر تجاری فعالان اقتصادی اقدام گردیده تا از این طریق میزان اخذ مالیات از تراکنش های تجاری شفاف تر گردد. همچنین پیشنهاد می شود که جهت انتخاب متغیرهای ورودی و خروجی الگوریتم های حل مسئله از روش دلفی و توافق خبرگان استفاده شود و همچنین به سایر پژوهشگران پیشنهاد می شود از روش های تصمیم گروهی جهت انتخاب این متغیرها استفاده نمایند.

### منابع و مأخذ

- [۱] امید پور، رضا. پژوهش، رضایان، جمشید. (۱۳۹۶). فرار مالیاتی در پایه مالیات بر درآمد اشخاص حقوقی در ایران (برآوردهای سالانه ۱۳۹۲-۱۳۵۲). اقتصاد مالی (اقتصاد مالی و توسعه)، تابستان ۱۳۹۶، دوره ۱۱، شماره ۳۹، ۲۷-۵۶.
- [۲] امید پور، رضا. پژوهش، رضایان، جمشید. محمدی، تیمور. معمارنژاد، عباس. (۱۳۹۴). برآورد حجم اقتصاد زیرزمینی و فرار مالیاتی: تحلیل تجربی در ایران. پژوهشنامه مالیات، زمستان ۱۳۹۴، دوره ۲۳، شماره ۲۸ (مسلسل ۷۶)، ۶۹-۹۴.
- [۳] امیری، رسول. (۱۳۹۷). اجتناب از پرداخت مالیات، ریسک مالیاتی و سررسید بدهی در شرکت های پذیرفته شده در بورس اوراق بهادار تهران. دانشگاه آزاد اسلامی واحد مرودشت. پایان نامه کارشناسی ارشد، ۱۳۹۷.
- [۴] امیری، میثم. (۱۳۹۶). اقتصاد رفتاری و فرار مالیاتی. پژوهشنامه اقتصادی. بهار ۱۳۹۶، دوره ۱۷، شماره ۶۴، ۹۵-۱۳۰.
- [۵] باقری، بهروز. عموری، مهدی. (۱۳۸۸). شناسایی و اولویت بندی نسبت های مالی اثرگذار بر درآمد مشمول مالیات. پژوهشنامه مالیات، بهار ۱۳۸۸، دوره جدید - ۱۷، شماره ۴ (مسلسل ۵۲)، ۷۳-۸۰.
- [۶] باقرپور ولشانی، محمد علی. باقری، مصطفی. خادم، حمید. حسینی پور، رضا. (۱۳۹۰). بررسی عوامل مالی و غیرمالی موثر بر گریز مالیاتی با استفاده از تکنیک های داده کاوی: صنعت خودرو و ساخت قطعات. مطالعات تجربی حسابداری مالی، دوره ۹، شماره ۳۴، تابستان ۱۳۹۰، ۱۰۳-۱۲۸.
- [۷] پژوهش، رضایان، جمشید. درویشی، باقر. (۱۳۸۹). اصلاحات ساختاری در نظام مالیاتی ایران، پژوهشنامه مالیات. بهار و تابستان ۱۳۸۹، دوره جدید ۱۸، شماره ۸ (مسلسل ۵۶)، ۹-۴۷.
- [۸] پناهی، ریحانه. (۱۳۹۵). تشخیص فرار مالیاتی مؤدیان مالیات بر ارزش افزوده با استفاده از روش های داده کاوی (مطالعه موردی: مؤدیان مالیات بر ارزش افزوده شهر تهران). دانشگاه علامه طباطبائی، پایان نامه کارشناسی ارشد، ۱۳۹۵.
- [۹] پورحیدری، امید. فدوی، محمدحسن. امینی نیا، میثم. (۱۳۹۳). بررسی تاثیر اجتناب از پرداخت مالیات بر شفافیت گزارشگری مالی شرکت های پذیرفته شده در بورس اوراق بهادار تهران. پژوهشنامه اقتصادی، بهار ۱۳۹۳، دوره ۱۴، شماره ۵۲، ۶-۸۵.

شرکت های دانش بنیان انجام گردد. پیشنهاد می شود فرار مالیاتی با استفاده از ظرفیت قوانین مجازات اسلامی جرم انگاری شده تا همگان از هزینه وقوع جرم مطلع گردند. همچنین پیشنهاد می گردد قوانینی در جهت کاهش معافیت های مالیاتی زائد و هدمند نمودن آن برای افراد خوش حساب مالیاتی در نظام مالیاتی تهیه و تدوین شود. نتایج داده کاوی نشان می دهد گروه صنفی فروشندگان لوازم یدکی خودرو سبک و سنگین، تعمیرگاه خودرو و نمایندگی مجاز فروش خودرو و مشاغل وابسته فرار مالیاتی بیشتری نسبت به سایر خوشه های صنفی دیگر را دارد. لذا پیشنهاد می گردد ارزیابان امور مالیاتی حساسیت بیشتری نسبت به فرار مالیاتی این گروه از مشاغل را داشته باشند. نتایج داده کاوی نشان می دهد در گروه صنفی وکلا و کارشناسان رسمی دادگستری و مشاورین حقوقی و دفاتر ازدواج و طلاق و مشاغل وابسته بیشتر فرار مالیاتی مربوط به صنفی می باشد که سابقه فعالیت کمتر ۱۱ سال دارا می باشند لذا پیشنهاد می گردد ارزیابان امور مالیاتی حساسیت بیشتری نسبت به فرار مالیاتی این گروه از مشاغل را داشته باشند. جهت داده کاوی بررسی فرار مالیاتی مشاغل از پایگاه داده امور مالیاتی استان قزوین به عنوان مورد مطالعه در یک سری زمانی مشخص استفاده شده است. به سایر پژوهشگران پیشنهاد می گردد با توجه به کلاس بندی منتخب در این پژوهش، داده های سایر استان های کشور را مورد تجزیه و تحلیل قرار داده و نتایج را جهت بررسی جامعیت نتایج مقایسه نمایند. در این پژوهش به دلیل اهمیت موضوع، داده های مالیاتی مرتبط با مشاغل در پنج گروه مالیاتی شامل گروه صنفی وکلا و کارشناسان رسمی دادگستری و مشاورین حقوقی و دفاتر ازدواج و طلاق و مشاغل وابسته، گروه صنفی بساز بفروشان، مصالح ساختمانی، مهندسین ساختمان و سایر مشاغل وابسته، گروه صنفی فست فود ها، هتل ها و متل ها، رستوران ها و چلو کبابی ها و مشاغل وابسته، گروه صنفی دفاتر پلیس +۱۰، دفاتر خدماتی ارتباطی دولت، ITC، فروشندگان تلفن همراه و سیم کارت و مشاغل وابسته، گروه صنفی فروشندگان لوازم یدکی خودرو سبک و سنگین، تعمیرگاه خودرو و نمایندگی مجاز فروش خودرو و مشاغل وابسته مورد تجزیه و تحلیل قرار گرفت. به سایر پژوهشگران پیشنهاد می گردد سایر گروه های مشاغلی که در این پژوهش به آنها پرداخته نشده است را انتخاب و مورد مطالعه خود قرار نمایند. در این پژوهش به دلیل محدودیت های اطلاعاتی موجود در پایگاه داده، کمترین داده های مورد تحلیل در گروه صنفی وکلا و کارشناسان رسمی دادگستری و مشاورین حقوقی و دفاتر ازدواج و طلاق و مشاغل وابسته می باشد. به سایر پژوهشگران پیشنهاد می گردد که در صورت دسترسی به داده های مالیاتی مشاغل و شرکت ها که تواما در این پژوهش انجام شده و بررسی نقش آنها در فرار مالیاتی بررسی گردد و پیشنهاد می گردد با استفاده از ابزار هوش مصنوعی و علم

- [۱۰] تقوی فرد، سیدمحمدتقی. رئیسی وانانی، ایمان. پناهی، ریحانه. (۱۳۹۶). تحلیل آینده نگر تشخیص فرار مالیاتی مودیان مالیات بر ارزش افزوده با استفاده از الگوریتم های طبقه بندی و خوشه بندی. پژوهشنامه مالیات، پاییز ۱۳۹۶، دوره ۲۵، شماره ۳۵ (مسلسل ۸۳)، ۱۱-۳۵.
- [۱۱] حسنی، محمد. شبان مختاری، مهدی. مودی، مسیانی. (۱۳۹۱). بررسی عوامل تاثیر گذار بر ظرفیت مالیاتی و پیش بینی درآمدهای مالیاتی استان خراسان جنوبی. پایان نامه کارشناسی ارشد
- [۱۲] حمیدی، ناصر. محمدزاده، امیر. محمدی، فاطمه. (۱۳۹۴). بررسی جایگاه جرایم مالیاتی در جلوگیری از فرار در نظام مالیات بر ارزش افزوده (مطالعه موردی استان قزوین). پژوهشنامه مالیات. پاییز ۱۳۹۴، دوره ۲۳، شماره ۲۷ (مسلسل ۷۵)، ۱۴۷-۱۶۶.
- [۱۳] خالقی، ابوالفتح. سیفی قره یثاق، داوود. (۱۳۹۴). رویکرد نظام کیفری ایران به جرم فرار مالیاتی. پژوهشنامه مالیات، زمستان ۱۳۹۴، دوره ۲۳، شماره ۲۸ (مسلسل ۷۶)، ۱۸۱-۲۰۷.
- [۱۴] خداپرست، محبوبه. سلطان حسینی، محمد. سلیمی، مهدی. (۱۳۹۷). برآورد سهم نسبی عوامل مدیریتی بر فرار مالیاتی ورزشکاران و مربیان حرفه ای باشگاه های ورزشی استان اصفهان. پژوهش های فیزیولوژی و مدیریت در ورزش، تابستان ۱۳۹۷، دوره ۱۰، شماره ۱۰۱، ۲-۱۱۳.
- [۱۵] پیرایی، خسرو. سلطانی شیرازی، الیزابت. (۱۳۸۷). برآورد ظرفیت مالیاتی در استان فارس. پژوهشنامه مالیات، فصلنامه علمی سازمان امور مالیاتی کشور، ۱۳۸۷، دوره ۱۶، شماره ۱
- [۱۶] خسروی، تانیا. پژویان، جمشید. (۱۳۹۲). تاثیر مالیات بر شرکت ها بر سرمایه گذاری بخش خصوصی با استفاده از رهیافت کرانه ها. اقتصاد مالی (اقتصاد مالی و توسعه). زمستان ۱۳۹۲، دوره ۷، شماره ۲۵، ۹۵-۱۲۱.
- [۱۷] دهقانی، سحر. موسوی جهرمی، یگانه. عبدلی، قهرمان. (۱۳۹۷). تئوری چشم انداز. رهیافتی نوین در توضیح پدیده فرار مالیاتی. تحقیقات اقتصادی. بهار ۱۳۹۷، دوره ۵۳، شماره ۱، ۱-۲۳.
- [۱۸] دستگیر، محسن. غریبی، مریم. (۱۳۹۴). کاربست روش های داده کاوی به منظور ارتقای عملکرد تشخیص فرار مالیاتی. پژوهشنامه مالیات. زمستان ۱۳۹۴، دوره ۲۳، شماره ۲۸ (مسلسل ۷۶)، صفحه ۹۵-۱۱۶.
- [۱۹] رحیمی کیا، اقبال. محمدی، شاپور. غضنفری، مهدی. (۱۳۹۴). تشخیص فرار مالیاتی با استفاده از سیستم هوشمند ترکیبی. پژوهشنامه مالیات. تابستان ۱۳۹۴، دوره ۲۳، شماره ۲۶ (مسلسل ۷۴)، ۱۳۵-۱۶۳.
- [۲۰] رضاقلی زاده، مهدیه. آقایی، مجید. عالمی، امیرحسین. (۱۳۹۸). تحلیل فرار مالیاتی در ایران به روش شاخص چندگانه- علل چندگانه (MIMIC). مجلس و راهبرد، بهار ۱۳۹۸، دوره ۲۶، شماره ۹۷، ۱۹۱-۲۲۶.
- [۲۱] رضایی سیابیدی، محمد. (۱۳۹۶). راه های مقابله و جلوگیری از فرار مالیاتی. قانون پار، زمستان ۱۳۹۶، دوره ۴، شماره ۴، ۱۴۵-۱۵۹.
- [۲۲] زهی، نقی. محمدخانلی، شهرزاد. (۱۳۸۹). بررسی عوامل موثر بر فرار مالیاتی (مطالعه موردی استان آذربایجان شرقی). پژوهشنامه مالیات، پاییز و زمستان ۱۳۸۹، دوره دوره جدید ۱۸، شماره ۹ (مسلسل ۵۷)، ۲۵-۶۰.
- [۲۳] شیرخانی، فریبا. (۱۳۸۴). بررسی مالیات بر درآمد بخش کشاورزی در ایران، مجله اقتصادی ۱۴-۲۱، ۱۱-۱۸.
- [۲۴] عباسیان، عزت ا... محمودی، وحید. شاکر، ایمان. (۱۳۹۰). تحلیل خطای پیشبینی درآمدهای مالیاتی دولت در اقتصاد ایران. تحقیقات مالی دانشکده مدیریت دانشگاه تهران، دوره ۱۳، شماره ۳۲، پاییز و زمستان ۱۳۹۰، ۱۰۹-۱۳۲
- [۲۵] عبدالله میلانی، مهنوش. اکبرپور روشن، نرگس. (۱۳۹۱). فرار مالیاتی ناشی از اقتصاد غیررسمی در ایران.
- [۲۶] عبدلی، قهرمان. ابریشمی، حمید. حسینی فرد، سید محمد. (۱۳۹۴). تحلیل نظری و تجربی حسابرسی مالیاتی مطلوب در مالیات بر درآمد جهت کاهش فرار مالیاتی. پژوهشنامه مالیات، زمستان ۱۳۹۴، دوره ۲۳، شماره ۲۸ (مسلسل ۷۶)، ۴۱-۶۷.
- [۲۷] عرب مازار، علی اکبر. دهقانی، علی. (۱۳۸۸). برآورد کارایی مالیات بر درآمد مشاغل و اشخاص حقوقی. پژوهشنامه مالیات، زمستان ۱۳۸۸، دوره دوره جدید ۱۷، شماره ۷ (مسلسل ۵۵)، ۴۵-۶۴.
- [۲۸] فطرس، محمدحسین. دلایی میلان، علی. (۱۳۹۵). بررسی اقتصاد زیرزمینی و فرار مالیاتی در چارچوب مدل های تعادل عمومی پویای تصادفی (DSGE). پژوهش های رشد و توسعه اقتصادی، زمستان ۱۳۹۵، دوره ۷، شماره ۲۵، ۶۵-۸۴
- [۲۹] فلاح پور، سعید. گل ارضی، غلامحسین. فتوره چیان، ناصر. (۱۳۹۲). پیش بینی روند حرکتی قیمت سهام با استفاده از ماشین بردار پشتیبان بر پایه الگوریتم ژنتیک در بورس اوراق بهادار تهران. تحقیقات مالی، پاییز و زمستان ۱۳۹۲، دوره ۱۵، شماره ۲، ۲۶۹-۲۸۸.
- [۳۰] کریمی پتانلار، سعید. گیلک حکیم آبادی، محمدتقی. صابرنوچمنی، فاضل. (۱۳۹۴). بررسی تاثیر اثربخشی دولت بر کاهش فرار مالیاتی در کشورهای منتخب. پژوهشنامه مالیات، پاییز ۱۳۹۴، دوره ۲۳، شماره ۲۷ (مسلسل ۷۵)، ۶۳-۹۰.
- [۳۱] کرمانخی، جواد. ویسمرادی، اکبر. علی مدد، زهرا. (۱۳۹۵). بررسی میزان کارایی جرایم مالیاتی در جلوگیری از فرار مالیاتی در نظام مالیات بر ارزش افزوده در بین مودیان مالیاتی استان ایلام، حسابداری دولتی، بهار و تابستان ۱۳۹۵، دوره ۲، شماره ۴، ۲۵-۳۶.
- [۳۲] لاری دشت بیاض، محمود. قائم مقامی، کامران. کهرمی، قاسم. (۱۳۹۵). بررسی عوامل موثر بر فرار مالیاتی در استان خراسان جنوبی با تاکید بر مولفه های فرهنگی. حسابداری ارزشی و رفتاری، بهار و تابستان ۱۳۹۵، دوره ۱، شماره ۱، ۱۳۹-۱۶۴.
- [۳۳] مجدآبادی فراهانی، فرزانه. (۱۴۰۱). سیاست جنایی ایران در قبال جرم مالیاتی. پژوهشنامه مالیات فصلنامه علمی سازمان امور مالیاتی کشور، دوره ۳۰، شماره ۵۶.
- [۳۴] محمدی خیاره، محسن. (۱۳۹۷). بررسی تاثیر فرار مالیاتی و فساد بر کارآفرینی: مورد مطالعه کشورهای OECD. توسعه کارآفرینی، پاییز ۱۳۹۷، دوره ۱۱، شماره ۳، ۵۰۱-۵۲۰.

- [54] DeBarr, D., Eyster-Walker, Z. (2006). **Closing the Gap: Automated Screening of Tax Returns to Identify Egregious Tax Shelters**, *ACM SIGKDD Explor. Newslett.* 8 (1), 11–16.
- [55] Demir, B., Javorcik, B. (2020). **Trade Policy Changes, Tax Evasion and Benford's law**, *Journal of Development Economics*, Volume 144, 102456
- [56] Devereux, M.P., Griffith, R., Klemm, A. (2002). **Can International Tax Competition Explain Corporate Income Tax Reforms**, *Econ. Pol.* 35.
- [57] Didimo, W., Giamminonni, G., Liotta, G., Montecchiani, F., Pagliuca, D. (2018). **A Visual Analytics System to Support tax Evasion Discovery**. *Decision Support Systems* Volume 110, pp.71-83.
- [58] Di Gioacchino, D., Fichera, D.(2020). **Tax Evasion and Tax Morale: A Social Network Analysis European**. *Journal of Political Economy*, Volume 65, 101922.
- [59] Drezewski, R., Sepielak, J., Filipkowski, W. (2015). **The Application of Social Network Analysis Algorithms in a System Supporting Money Laundering Detection**, *Inf. Sci.* 295,18–32
- [60] Ferrantino, M.J., Liu, X., Wang, Z. (2012). **Evasion Behaviors of Exporters and Importers: Evidence from the US–China Trade Data Discrepancy**. *J. Int. Econ.* 86 (1) , 141–157.
- [61] González, P.C., Velásquez, J.D. (2017). **Characterization and Detection of Taxpayers With False Invoices Using Data Mining Techniques**. *Expert Syst. Appl.* 40 (5).
- [62] Goumagias, N.D., Hristu-Varsakelis, D., Saraidaris, A. (2012). **A Decision Support Model for Tax Revenue Collection in Greece**, *Decis. Support Syst.* 53 (1), 76–96.
- [63] Han, J., Kamber, M., Pei, J. (2011) **Data Mining Concepts and Techniques**. Third Edition, ISBN 978-0-12-381479-1
- [64] Han, J., Kamber, M., Pei, J. (2012). **Data Mining: Concepts and Techniques A volume in Series in Data Management Systems**. Morgan Kaufmann, Third Edition.
- [65] Hobson, J.L., Mayew, W.J., Peecher, M., Venkatachalam, M. (2015). **Auditors Ability to Detect Financial Deception: the Role of Auditor Experience and Management**, *Cognitive Dissonance*.
- [66] Hsu, K.W., Pathak, N., Srivastava, J., Tschida, G., Bjorklund, E. (2015). **Data Mining Based Tax audit Selection: A Case Study of a Pilot Project at the Minnesota Department of Revenue**, in: **Real World Data Mining Applications**. Springer, pp. 221–245.
- [67] Hung, F.S. (2017). **Explaining the Nonlinearity of Inflation and Economic Growth: The Role of Tax evasion International Review of Economics & Finance**. Volume 52, pp.436-445.
- [68] Hwang, S., Nagac, K. (2021), **Tax Avoidance and Excess Burden of Income Tax**, *Korean Econometric Society*. Volume 32, Issue 3, September 2021, Pages 69-93
- [69] Khalil S., Sidani, Y. (2020). **The Influence of Religiosity on Tax Evasion Attitudes in Lebanon**. *Journal of International Accounting, Auditing and Taxation*, Volume 40.
- [70] Kim, Y.J., Baik, B., Cho, S. (2016). **Detecting Financial Misstatements with Fraud Intention Using Multi-Class Cost-Sensitive Learning**. *Expert Syst. Appl.* 62 .
- [71] Kwak, W., Shi, Y., Lee, C.F. (2010). **The Fuzzy Set and Data Mining Applications in Accounting and Finance**. in: *Handbook of Quantitative Finance and Risk Management*, Springer, (2010), pp. 1307–1331.
- [72] Jia, S. (2016). **Tax Losses in China: Estimates, Analysis and Countermeasures**. *Friends Account.* 13, 39–40.
- [73] Lihong, C. (2022). **Research on Tax Collection and Administration Application and Legal Issues on Big Data Analysis**. *Hindawi,journal of sensors* volume 2022,article id 6578964, page11
- [74] Liu, X., Pan, D., Chen, S. (2010). **Application of Hierarchical Clustering in Tax Inspection Case-Selecting**. in: *Proceedings of the International Conference on Computational Intelligence and Software Engineering (CISE)*, IEEE, pp. 1–4.
- [35] مداح، مجید. خالق پناه، زهرا. (۱۳۹۴). **فرار مالیاتی در واردات ایران رویکرد مدل ترکیبی شبکه عصبی مصنوعی و الگوریتم تدریجی**. برنامه ریزی و بودجه، تابستان ۱۳۹۴، دوره ۲۰، شماره ۲، ۸۵ – ۱۰۲.
- [36] مقیمی نیا، علی. (۱۳۸۷). **طراحی راهکاری مناسب جهت تعیین و محاسبه ضرایب مالیات بر مشاغل**. پژوهشنامه مالیات فصلنامه علمی سازمان امور مالیاتی کشور، ۱۳۸۷، دوره ۱۶، شماره ۳، ۹–۳۸.
- [37] موسوی جهرمی، یگانه. طهماسبی بلداجی، فرهاد. خاکی، نرگس. (۱۳۸۸). **فرار مالیاتی در نظام مالیات بر ارزش افزوده: یک مدل نظری**. پژوهشنامه مالیات، تابستان ۱۳۸۸، دوره دوره جدید ۱۷، شماره ۵ (مسلسل ۵۳)، ۲۷ – ۳۸.
- [38] مهران، ساسان. سیدی، سیدجلال. (۱۳۹۳). **بررسی تاثیر مالیات بر درآمد و حسابداری محافظه کارانه بر اجتناب مالیاتی شرکت ها**. فصلنامه علمی پژوهشی دانش حسابداری و حسابرسی مدیریت سال سوم، شماره دهم
- [39] نظری، علیرضا. فدایی، ایمان. (۱۳۹۲). **آسیب شناسی نظام مالیاتی ایران**. فصلنامه سیاستهای مالی و اقتصادی، سال اول، شماره ۴، زمستان ۱۳۹۲، ۱۱۰–۹۵.
- [40] Abdixhiku, L., Krasniqi, B., Pugh, G., Hashi, I. (2017). **Firm-Level Determinants of Tax Evasion in Transition Economies**. *Economic Systems*, Volume 41, Issue 3, pp.354-366.
- [41] Ahmad al-Zaqeba, M.A., Suhaila A.H., izlawanie, M. (2018). **Tax Compliance of Individual Taxpayers: A Systematic Litreature Review**.
- [42] Agarwal, S., Li, K., Qin, Y., Jing, W., Yan, J. (2020). **Tax Evasion, Capital Gains Taxes, and the Housing Market**. *Journal of Public Economics*, Volume 188, 104222.
- [43] Akoglu, L., Tong, H., Koutra, D. (2015). **Graph Based Anomaly Detection and Description: A Survey**. *Data Min. Knowl. Discov.* 29 (3), pp. 626–688.
- [44] Alaghi, K. (2015). **Operating Profit, Productivity Indexes and Net Earnings**. *Afr. J. Bus. Manag.* 6 (11), 4397.25.
- [45] Alstadsæter, A., Johannesen, N., Zucman, G. (2019). **Tax Evasion and Inequality**. *American Economic Review* 2019, 109(6): 2073–2103
- [46] Assylbekov, Z., Melnykov, I., Bekishev, R., Baltabayeva, A., Bissengaliyeva, D., Mamlin, E. (2016). **Detecting Value-Added Tax Evasion by Business Entities of Kazakhstan**, in: *Intelligent Decision Technologies* Springer, pp. 37–49.
- [47] Aurazo, J. (2020). **Merchant Internalization and Tax Evasion**. *Research in Economics* in press, 2020.
- [48] Barak, S., Modarres, M. (2015). **Developing an Approach to Evaluate Stocks by Forecasting Effective Features With Data Mining Methods**. *Expert Systems with Applications*, 42(3), 1325–1339.
- [49] Barman, S., Pal, U., Sarfaraj, M.A., Biswas, B., Mahata, P. Mandal, P. (2016). **A Complete Literature Review on Financial Fraud Detection Applying Data Mining Techniques**. *Int. J. Trust Manag. Comput. Commun.* 3 (4), 336–359.
- [50] Beutel, L.A., Akoglu, C., Faloutsos, F. (2015). **Detection through Graph-Based User Behavior Modeling**, in: *Proceedings of the Twenty-Second ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2015, 1696–1697.
- [51] Bowen, R.M., Call, A.C., Rajgopal, S. (2010). **Whistle-Blowing: Target Firm Characteristics and Economic Consequences**. *Account. Rev.* 85 (4), 1239–1271.
- [52] Chen, Y.S., Cheng, C.H. (2010). **A Delphi-Based Rough Sets Fusion Model for Extracting Payment Rules of Vehicle License Tax in the Government Sector**. *Expert Syst. Appl.* 37 (3), 2161–2174.
- [53] Davia, H.R., Coggins, P.C., Wideman, J.C., Kastantin, J.T. (2000). **Accountant's Guide to Fraud Detection and Control**. 2nd ed. New York: John Wiley.

- [75] McCue, C. (2007). **Data Mining and Predictive Analysis Intelligence Gathering and Crime Analysis**. Butterworth-Heinemann
- [76] McCue, C. (2015). **Data Mining and Predictive Analysis Intelligence Gathering and Crime Analysis**. Butterworth-Heinemann, Second Edition
- [77] Richardson, G. (2008). **The Relationship Between Culture and Tax Evasion Across Countries: Additional Evidence And Extensions**. Journal of International Accounting, Auditing and Taxation. 17: 67-78.
- [78] Ruan, J., Yan, Z., Dong, B., Zheng, Q., Qian, B. (2019). **Identifying Suspicious Groups of Affiliated-Transaction-Based Tax Evasion in Big Data**, Information Sciences, Volume 477, pp.508-532 .
- [79] Sikka, P. (2010). **Smoke and Mirrors: Corporate Social Responsibility and Tax Avoidance**, in: **Accounting Forum**. 34, Elsevier, 2010, pp. 153-168.
- [80] Stankevicius, E. (2015). **Hybrid Approach Model for Prevention of Tax Evasion and Fraud**, **Procedia - Social and Behavioral Sciences**. Volume 213, pp.383-389.
- [81] Tajmir Riahi, H., Al Mohammad, F. (2019). **Analyzing and Criticizing the Legislative Criminal Policy of Tax Crimes in the Reformed Direct Taxes Law of 2014**. Research Journal of Taxation, 46:29, 37-38.
- [82] Tian, F., Lan, T., Chao, K.M., Godwin, N., Zheng, Q., Shah, N., Zhang, F. (2016). **Mining Suspicious Tax Evasion Groups in Big Data**, **IEEE Trans. Knowl. Data Eng.** 28(10), 2651-2664.
- [83] Tselykh, A., Knyazeva, M., Popkova, E., Durfee, A., Tselykh, A. (2016). **An Attributed Graph Mining Approach to Detect Transfer Pricing Fraud**. in: Proceedings of the Ninth International Conference on Security of Information and Networks, ACM, 2016, pp. 72-75.
- [84] West, J., Bhattacharya, M. (2016). **Intelligent Financial Fraud Detection: A Comprehensive Review**, **Comput. Secur**, 57,47-66
- [85] Wu, R.S., Ou, C.S., Lin, H.Y., Chang, S.I., Yen (2012). **Using Data Mining Technique to Enhance Tax Evasion Detection Performance**. **Expert Systems with Applications** 39(10): 8769-8777.
- [86] Yang, X.S. (2019). **Introduction to Algorithms for Data Mining and Machine Learning**. Academic Press, (2019 )
- [87] Yamen, A. (2018). **Impact of Institutional Environment Quality on Tax Evasion: A Vomparrative Investigation of Old versus New EU Members**. Journal of International Accounting, Auditing and Taxation, Volume 32, (2018), pp.17-29.
- [88] Yousefi, K., Vesal, M., Hanifa, P. (2020). **Import Tax Evasion and Avoidance: Evidence from Iran** .The Quarterly Review of Economics and Finance, Volume 75, pp.31-39.
- [89] Ziai Bigd, E., Gholami, T., Ebi, B., Idaji, B. (2013). **Investigating the Effect of Economic Sanctions on Iran's Trade: An Application of the Gravity Model**. **Economic s Research**, 13(48): 282 -291.