



فصلنامه علمی پژوهشی دانش سرمایه‌گذاری  
سال هشتم / شماره سی‌ودوم / زمستان ۱۳۹۸

## استفاده از تکنیک‌های داده‌کاوی برای سنجش ریسک مالیاتی مؤدیان مالیات بر ارزش افزوده

محمد مسیحی

دانشجوی دکتری تخصصی، گروه مدیریت مالی، واحد امارات، دانشگاه آزاد اسلامی، دبی، امارات متحده عربی  
Mohammad-masahi53@yahoo.com

احمد یعقوب نژاد

دانشیار، گروه حسابداری، واحد تهران مرکز، دانشگاه آزاد اسلامی، تهران، ایران (نویسنده مسئول)  
Yaghoobacc@gmail.com

امیررضا کیقبادی

استادیار، گروه حسابداری، واحد تهران مرکز، دانشگاه آزاد اسلامی، تهران، ایران  
Acc1388@gmail.com

تقی ترابی

دانشیار، گروه اقتصاد، واحد علوم تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران  
tttorabi@gmail.com

تاریخ دریافت: ۹۷/۱۰/۱۲ تاریخ پذیرش: ۹۸/۰۱/۲۵

### چکیده

در این مقاله با استفاده از تکنیک‌های داده‌کاوی ریسک مالیاتی مؤدیان در نظام مالیات بر ارزش افزوده مورد مطالعه قرار می‌گیرد. اهمیت ارزیابی ریسک مالیاتی مؤدیان مالیات بر ارزش افزوده به منظور تدوین طرح مؤثر انتخاب مؤدیان برای حسابرسی مالیاتی با هدف افزایش کارایی و اثر بخشی در نظام مالیات بر ارزش افزوده کشور می‌باشد. مؤدیان مالیاتی در این تحقیق به سه گروه مؤدیان فاقد ریسک، با ریسک پایین و پر ریسک طبقه بندی شده‌اند به منظور ارزیابی ریسک مالیاتی از دو تکنیک داده‌کاوی ماشین بردار پشتیبان و رگرسیون لجستیک استفاده شده است. جامعه آماری پژوهش شامل اشخاص حقوقی بزرگ در شهر تهران می‌باشد که در سال‌های ۱۳۹۰ تا ۱۳۹۳ مورد حسابرسی مالیاتی در نظام مالیات بر ارزش افزوده قرار گرفته‌اند در این تحقیق متغیرها شامل ساز و کارهای حاکمیت شرکتی، ویژگی‌های خاص شرکتی، ماهیت فعالیت مؤدیان سیستم کنترلی مؤدیان و نسبت‌های مالی می‌باشد که به منظور آموزش و آزمون مدل استفاده شده‌اند نتایج تحقیق نشان می‌دهد دو تکنیک Logistic و SVM از توان صحت ارزیابی ۷۰٪ برخوردار هستند و نوعی ادغام در نتایج این دو تکنیک توانسته است با کسب نزدیک به ۸۳٪ صحت ارزیابی از توان بالاتری برخوردار باشد.

**واژه‌های کلیدی:** تکنیک‌های داده‌کاوی، ریسک مالیاتی، مالیات بر ارزش افزوده، حسابرسی مالیاتی.

## ۱- مقدمه

مالیات قدرت ذاتی و مهمترین ابزار تأمین مالی مخارج سرمایه‌گذاری و سرمایه در گردش دولت‌ها است اما یک مشکل بزرگ در اجرای مؤثر مدیریت مالیاتی دولت‌ها فرار از پرداخت مالیات است. در ایران نیز منابع اصلی درآمدهای دولت عمدتاً از دو منبع مالیات و نفت تأمین می‌شود لیکن در این میان یکی از مسائل اساسی اقتصاد ایران، سهم اندک درآمدهای مالیاتی در مجموع درآمدهای دولت می‌باشد که نسبت پایین مالیات به تولید ناخالص داخلی (T/GDP) در کشور که در بهترین سال (۱۳۸۸) معادل ۸/۵٪ بوده است مبین این موضوع می‌باشد به منظور رفع این مشکل در کشورهای مختلف دنیا حسابرسی مالیاتی را به سیستم مالیاتی خود وارد نموده‌اند. (جان اوبی آنیا دوبا ۲۰۱۴) البته باید توجه داشت که در یک سیستم مدرن جهت وصول مالیات: ۱- غیر ممکن است که دولت تک تک مؤدیان را مورد حسابرسی قرار دهد؛ ۲- بررسی مداوم مؤدیان کم خطر و تمکین‌کننده، هدر دادن منابع توسط دولت می‌باشد. لیکن در نظام مالیاتی کشور الزام به حسابرسی همه مؤدیان مالیاتی وجود دارد. که این امر موجب عدم کارایی نظام مالیاتی در مواجهه با فرار مالیاتی می‌گردد و ضرورت دارد که سیستم حسابرسی مبتنی بر ریسک جایگزین حسابرسی همه مؤدیان گردد. این موضوع در نظام مالیات بر ارزش افزوده که مبتنی بر خوداظهاری می‌باشد و بار مالی آن بر مؤدیان مالیاتی تحمیل نمی‌گردد بلکه صرفاً وظیفه وصول آن از مصرف‌کنندگان و پرداخت به سازمان امور مالیاتی کشور را دارند از اهمیت بالایی برخوردار می‌باشد تا با تأکید بر حسابرسی مؤدیان پر ریسک خود اظهاری مؤدیان افزایش یابد. در سیستم حسابرسی مبتنی بر ریسک انتخاب مؤدی جهت حسابرسی بر اساس معیار ریسک خواهد بود. در نتیجه برنامه‌ریزی حسابرسی با توجه به طبقات مختلف ریسک مؤدیان انجام خواهد شد (James Alm 2016). که امر مستلزم شناسایی و سنجش ریسک مالیاتی مؤدیان می‌باشد.

مدیریت ریسک، یک استراتژی کلیدی برای سازمان مالیاتی نوین خواهد بود و تمرکز این استراتژی، بر پایه موارد زیر است:

- الف) ایجاد سیستم خوداظهاری؛
- ب) ارتقاء تمکین داوطلبانه؛
- ج) ساختار سازمانی براساس تمرکز عملکردی نه بر اساس تمرکز جغرافیایی؛
- د) تمرکز بر روی تخصص؛
- ه) ارتقاء انصاف و عدالت؛
- و) خدمات متنوع برای پاسخگویی به نیازهای طبقات مختلف مؤدیان؛

دولت نمی‌تواند تک تک مؤدیان را مورد ممیزی و رسیدگی قرار دهد و همچنین بررسی مداوم مؤدیان کم خطر و تمکین‌کننده، هدر دادن منابع توسط دولت می‌باشد. به نوعی می‌توان گفت هزینه فرصت از دست رفته برای یک چنین بررسی‌هایی بسیار بالا است، بنابراین منابع باید برای مؤدیان صرف شود که پتانسیل ایجاد درآمد و سود بالایی دارند. سازمان مالیاتی باید با هزینه منفعت نمودن روش‌ها و تکنیک‌های خود، تمکین

مالیاتی توسط مؤدیان را به منظور شناسایی و پیشگیری از رفتار مجرمانه مدیریت نماید و از طریق ارائه خدمات و آموزش مناسب، مؤدیان را در تمکین و پرداخت تعهدات مالیاتی شان کمک نماید. چنین سیستم مدرنی بر پایه خود اظهاری و تمکین داوطلبانه توسط مؤدیان پایه ریزی شده است و کارکرد ارزیابی ریسک از کارکرد اجرای حسابرسی تفکیک می‌گردد (کاواجا و همکاران، ۲۰۱۱).

استراتژی‌های شناسایی ریسک، بر مبنای تجزیه و تحلیل روند کسب و کار و الگوهای اجتناب مالیاتی توسعه یافته‌اند. انتخاب مؤدیان جهت حسابرسی بر پایه ارزیابی ریسک و توسعه تکنیک‌های انتخاب مبتنی بر ریسک می‌باشد. برای این منظور، ویژگی‌ها و مشخصات مؤدی در یک سیستم نمره دهی ریسک، جهت شناسایی و ارزیابی ریسک عدم تمکین استفاده می‌شود. این کار باعث اولویت‌بندی و برنامه‌ریزی حجم کار ممیزی و خدمات مؤدیان شده و باعث می‌شود که منابع به گروه‌های با ریسک بالا اختصاص داده شود (شکل زیر).



شکل (۱)

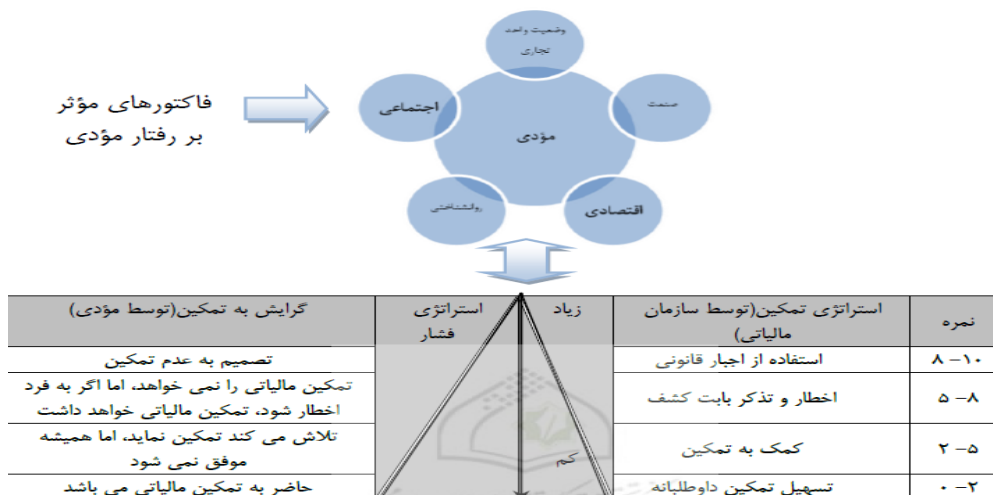
- اثربخشی مدیریت مبتنی بر ریسک تمکین، اساساً به موارد زیر بستگی دارد:
- جمع‌آوری اطلاعات شخص ثالث و مطابقت آن با اظهارنامه مؤدی با استفاده از پایگاه‌های داده قابل اتکاء و شماره شناسایی منحصر به فرد مؤدی مالیاتی.
  - انجام نمونه‌گیری بر پایه تجزیه و تحلیل ریسک.
  - استاندارد سازی فرآیندهای پرداخت، از قبیل پرداخت از طریق بانک و الزامات حسابداری.
  - اطمینان در خصوص به کارگیری یکنواخت قوانین و روش‌ها.
  - ارائه اطلاعات به موقع و مناسب برای تصمیمات مدیریتی در جهت تدوین سیاست مالیاتی.

در چند سال گذشته استفاده از اطلاعات اشخاص ثالث با استفاده از اتوماسیون افزایش یافته که به سازمان مالیاتی این امکان را می‌دهد تا پرو فایل‌های اظهارنامه‌های مالیاتی را ایجاد نموده و به صورت آنلاین به اظهارنامه‌ها دسترسی داشته باشد. این تحولات باید در تنظیم استراتژی‌های حسابرسی مالیاتی خوب و رویکردهای مبتنی بر ریسک در نظر گرفته شود. استراتژی حسابرسی، ابزار مهم و حیاتی در مدیریت تمکین از طریق سازمان مالیاتی می‌باشد. یکی از جنبه‌های طراحی استراتژی حسابرسی، روش مورد استفاده برای انتخاب مؤدی می‌باشد. بسیاری از سازمان‌های مالیاتی استراتژی‌های حسابرسی خود را بر پایه روش‌های نمره دهی بر مبنای ریسک توسعه داده‌اند.

اما در این میان، نقش سازمان مالیاتی در مدل مبتنی بر ریسک، به عنوان یک بازیگر استراتژیک خواهد بود. در این مدل، این واقعیت در نظر گرفته می‌شود که احتمال شناسایی (با فرض اینکه مؤدیان بخواهند به طور طبیعی درآمدها را کمتر اظهار کنند) بستگی به اطلاعات ارائه شده توسط مؤدیان به سازمان مالیاتی در هنگام تکمیل اظهارنامه مالیاتی آنها، خواهد داشت.

## ۲- مبانی نظری و مروری بر پیشینه پژوهش

استراتژی حسابرسی مؤثر و کارآمد بر اساس چندین جنبه حیاتی از قبیل برنامه‌های آموزشی و انگیزشی برای حسابرسان مالیاتی، عملکرد فناوری اطلاعات و کافی بودن چارچوب قانونی برای حسابرسان می‌باشد. نگرش مؤدیان نسبت به تمکین از اهمیت بالایی برخوردار است. در حقیقت، اعمال یک استراتژی تمکین توسط سازمان مالیاتی، عکس‌العمل متناسب با سطح ریسک تمکین هر مؤدی می‌باشد (نظیر مکاتبه، اخطار، اعتماد به خود اظهاری، حسابرسی محدود و حسابرسی جامع).



شکل (۲) منبع: سازمان توسعه و همکاری‌های اقتصادی، ۲۰۰۲

یکی از جنبه‌های طراحی استراتژی حسابرسی، روش‌های مورد استفاده برای انتخاب مؤدیان می‌باشد که در این زمینه، مهمترین روش‌های مورد استفاده توسط سازمان مالیاتی جهت انتخاب مؤدیان برای ممیزی به شرح زیر می‌باشد.

## ۲-۱- انتخاب دستی

این روش هرچند مزایای غیر قابل انکاری از قبیل استفاده از دانش محلی و غیر رسمی داشته و مورد علاقه حسابرسان بوده و نیز نیاز محدود به پردازش داده‌ها دارد، اما بسیار چالش برانگیز است زیرا به وضوح خطر فساد را افزایش می‌دهد.

- نظام دستی در مقابل نظام سیستمی خواهد بود و در این صورت، حسابرسان، برخی جنبه‌های عدم تمکین را فراموش می‌کنند و هیچ تلاشی جهت کشف الگوهای مخفی شده در سابقه عدم تمکین در برخی زمینه‌های مشابه، بخش‌ها یا در ویژگی‌های سایر مؤدیان مالیاتی، انجام نمی‌دهند.
- به طور مشابه، اگر فرایند غربالگری به شدت دستی باشد، حسابرسان اغلب تنها از داده‌های داخل استفاده می‌کنند (داده‌های صادر شده توسط خود سازمان مالیاتی) و احتمالاً از اطلاعات مناسب از طریق داده‌های خارجی غفلت می‌ورزند (از جمله گمرکات و یا سایر ادارات مالیاتی).

## ۲-۲- انتخاب تصادفی

در این روش، مؤدیان به صورت تصادفی جهت حسابرسی انتخاب می‌شوند. استفاده از روش نمونه‌گیری طبقه‌ای در بسیاری از کشورها فراگیر شده است. در این روش، مؤدیان بر اساس معیارهایی از قبیل اندازه، جنسیت، نوع مالیات، طبقه‌بندی شده و سپس یک نمونه از هر طبقه انتخاب می‌گردد.

انتخاب تصادفی از چندین مزیت قابل توجه برخوردار است:

- الف) به سازمان مالیاتی این امکان را می‌دهد که با اطلاعات مربوط در مورد تمکین و جنبه‌های دیگری از رفتار مؤدیان بدون جانبداری در انتخاب ممیزی، حسابرسی را انجام دهد. انتخاب تصادفی روش خوبی برای جمع‌آوری داده‌ها و ارتقاء دانش در سازمان‌های مالیاتی می‌باشد.
- ب) استراتژی انتخاب منصفانه است چرا که همه مؤدیان در طبقه مشابه، احتمال مساوی جهت انتخاب برای ممیزی را دارند. همچنین روش انتخاب تصادفی خطرات ناشی از فساد و انتخاب خودسرانه را در مقایسه با روش غربالگری دستی کاهش می‌دهد. با این حال، انتخاب تصادفی یک مشکل عمده دارد و که آن هزینه فرصت بسیار بالا است. حسابرسی مؤدیان بوسیله روش‌های متمرکز بر مؤدیان با ریسک بالا و یا حتی حسابرسی بوسیله روش انتخاب غربالگری دستی درآمد بالاتری را نسبت به روش انتخاب تصادفی دارند (حتی با طبقه‌بندی)، در نتیجه حسابرسی تصادفی تأثیر کمتری بر درآمدهای مستقیم و احتمالاً بر عوامل بازدارنده عدم تمکین مالیاتی دارد.

## ۲-۳- انتخاب مبتنی بر ریسک

اکثر سازمان‌های مالیاتی استراتژی‌های حسابرسی خود را براساس تمرکز بر ریسک‌های عدم تمکین مؤدیان، گسترش داده‌اند (سازمان توسعه و همکاری‌های اقتصادی، ۲۰۰۶). تجربه نشان داده است که یک استراتژی انتخاب حسابرسی کارآمد، باید مؤدیانی را که به احتمال زیاد تمکین نمی‌کنند (مقررات گریز) می‌باشند را شناسایی کند چرا که احتمال تحصیل مبالغ بزرگ از طریق اصلاحات حسابرسی و جرایم از این گونه افراد بیشتر است. در این راستا بسیاری از سازمان‌های مالیاتی استراتژی‌های حسابرسی خود را بر پایه روش‌های نمره دهی بر مبنای ریسک توسعه داده‌اند. در یک سیستم نمره دهی ریسک، از مشخصات و ویژگی‌های مؤدی، جهت شناسایی و ارزیابی ریسک تمکین استفاده می‌شود. سیستم، ریسک اطلاعات و نحوه رعایت و تمکین مؤدی را مورد تجزیه و تحلیل قرار می‌دهد. چنانچه ریسک از آستانه از پیش تعیین شده پائین تر باشد، یک ارزیابی خودکار انجام پذیرفته و برگ تشخیص صادر می‌گردد و چنانچه بیش از آستانه باشد، مؤدی برای حسابرسی و بازنگری به صورت دستی انتخاب می‌گردد (انتخاب جهت حسابرسی مبتنی بر ریسک) و پس از حسابرسی و بررسی‌های لازم، برگ تشخیص صادر و ابلاغ می‌شود. این روش بر مبنای شناسایی و سنجش عوامل مؤثر بر ریسک مالیاتی انجام می‌شود. سازمان مالیاتی را قادر می‌سازد پرو فایلهایی از مؤدیان بسازند و کسانی که با احتمال زیاد قوانین مالیاتی را تمکین نمی‌کنند، را شناسایی نمایند.

این استراتژی ضمن ارائه اطلاعات آگاهی بخش برای سازمان مالیاتی و روش آماری قوی در جمع‌آوری داده، هزینه بر است و به طور معمول در کشورهای توسعه یافته با اتکا به بخش IT سیستم‌ها موارد ذیل اجرا می‌گردد:

- الف) وجود مقدار قابل توجهی از داده‌های با کیفیت (داخلی و خارجی به سازمان مالیاتی) در هر دو مورد حسابرسی گذشته و ویژگی‌های مورد حسابرسی فعلی (حسابرسی شده یا نشده).
- ب) وجود سیستم‌های فناوری اطلاعات (سخت افزار، نرم افزار و آموزش) برای پردازش داده‌ها و ارائه نمره و دادن اطلاعات لازم جهت برنامه‌ریزی حسابرسی.

به دلایل فوق، اجرای انتخاب حسابرسی بر مبنای ریسک می‌تواند با تجهیزات ضعیف کشورهای در حال توسعه همراه با فقر اطلاعاتی، چالش‌زا باشد، البته این بدان معنی نیست که مطلوب و یا امکان‌پذیر نباشد.

حسابرس به عنوان یک شخصیت مستقل و متخصص که هم مورد اعتماد مؤدیان و هم مورد تأیید دولت است، اطلاعات مالی شرکت‌ها را حسابرسی می‌نماید که این امر موجب بهبود کیفیت حسابرسی مالیاتی می‌شود. بهبود کیفیت حسابرسی مالیاتی و تشخیص مالیات به کمک نیروهای ماهر، فضای اعتماد متقابل بین نظام مالیاتی و مؤدیان را ایجاد نموده و رضایت‌مندی آن‌ها را افزایش می‌دهد، لذا مالیات تعیین شده با احتمال بالاتری مورد پذیرش مؤدیان قرار گرفته و تمکین مالیاتی افزایش می‌یابد (Melat, Abera, 2016). از جمله پژوهش‌های مرتبط با این موضوع می‌توان به پژوهش کووی هسو و همکاران (۲۰۱۵) اشاره نمود که در مقاله‌ای با عنوان «انتخاب برای حسابرسی مالیاتی با استفاده از داده کاوی» مطالعه موردی در یک پروژه پایلوت ارائه نموده‌اند که به منظور انتخاب برای حسابرسی در اداره مالیات مینه‌سوتا از روش داده‌کاوی استفاده گردید. در این مقاله با اشاره به این موضوع که برای دولت کاهش شکاف مالیاتی بسیار مهم است در ادامه یک روش مبتنی

بر استخراج داده‌ها ارائه گردید که برای بهبود فرآیند انتخاب برای حسابرسی در اداره مالیات بکار رفته است. طی این مقاله فرآیند انتخاب دستی برای حسابرسی در حوزه مالیات فروش و مصرف شرح داده شده است. درباره داده‌های استخراجی از منابع مختلف بحث گردید و به مسائل مرتبط با انتخاب شاخص‌ها پرداخته می‌شود و تکنیک‌های داده‌کاوی که مورد استفاده قرار می‌گیرند توضیح داده شد و نتایج بدست آمده از حسابرسی‌های میدانی واقعی انجام شده توسط حسابرسان گزارش گردید. این نتایج مؤید مفید بودن روش مبتنی بر استخراج داده‌ها در انتخاب برای حسابرسی هستند. برای استفاده از روش مبتنی بر داده‌کاوی (مدل‌سازی) با هدف بهبود انتخاب برای حسابرسی به کار می‌رود. داده‌ها نقش حیاتی در پروژه داده‌کاوی دارند به همین منظور به فرآیند پردازش، پالایش و فرمت مجدد داده‌ها توجه خاص گردید. سپس مدل‌ها استخراج و با استفاده از داده‌های دنیای واقعی تست شده‌اند. نتایج برای این پروژه نشان می‌دهد که روش مبتنی بر داده‌کاوی یک افزایش ۶۳/۱٪ در کارایی ایجاد نموده است. با اهمیت‌ترین قسمت این پروژه، اعتبار سنجی ناشی از حسابرسی‌های میدانی واقعی است که نشان می‌دهد مفید بودن استخراج داده‌کاوی، برای بهبود انتخاب‌های حسابرسی از نظر دقت و درآمد تولید شده است بنابراین داده‌کاوی راهکاری است برای کمک به افزایش کارایی و وصول درآمد که برای بیشتر فعالیت‌های مبتنی بر تمکین قابل استفاده می‌باشد. همچنین هاپتمن و همکاران (۲۰۱۴) در مقاله‌ای با عنوان "بهبود خدمات ادارات مالیات به عنوان عاملی در تمکین مالیاتی"، از رابطه بین مؤدیان و خدمات حسابرسی مالیاتی به عنوان ابزاری برای ادارات مالیاتی در جلوگیری از فرار مالیاتی مورد بحث واقع می‌شود استفاده کرده‌اند در بخش تئوریک، عوامل تمکین مالیاتی (کاهش ریسک مالیاتی) و ابزارهای ادارات مالیاتی در مبارزه با فرار مالیاتی، سطح نرخ‌های مالیاتی و مسئولیت رعایت قوانین در سیستم مالیاتی اعضای اتحادیه اروپا مورد مقایسه واقع می‌شود. در بخش تجربی مقاله، تجارب شرکت‌های اسلوونیایی در رابطه با خدمات حسابرسی مالیاتی مورد بررسی و تجزیه و تحلیل قرار می‌گیرد و نتیجه‌گیری می‌شود که درک بهتر محرک‌های مؤثر بر رفتار تمکین مؤدیان به ادارات مالیاتی در شناسایی و بکارگیری مؤثر ابزارهای سیاستگذاری کمک می‌کند. در پژوهش دیگری توسط راویسانکار و همکاران (۲۰۱۱)، با عنوان «تشخیص تقلب در صورت‌های مالی و متغیرهای انتخاب شده با استفاده از روش‌های داده‌کاوی» به شناسایی شرکت‌هایی که به تقلب در صورت‌های مالی متوسل می‌شوند پرداختند. آن‌ها روش‌های داده‌کاوی را بر روی مجموعه‌ای از داده‌های شرکت‌ها در دو مرحله آزمون کردند. در مرحله اول با انتخاب ۱۸ ویژگی، شبکه‌های عصبی احتمالاتی از بین تمامی روش‌ها برتر بود و در مرحله دوم نیز با انتخاب ۱۰ ویژگی، شبکه‌های عصبی احتمالاتی و برنامه نویسی ژنتیک با حاشیه دقت برابر نسبت به سایر روش‌ها برتر بودند. در بخش تحقیقات داخلی که توسط دستگیر و غریبی (۱۳۹۴)، با عنوان کاربست روش‌های داده‌کاوی به منظور ارتقای عملکرد تشخیص فرار مالیاتی صورت گرفت نتیجه گرفته شد که روش‌های داده‌کاوی مبتنی بر قواعد وابستگی با ایجاد دو مدل با درصد صحت ۹۱٪ بر روی داده‌های آموزش، با درصد صحت ۸۸٪ بر روی داده‌های اعتبار سنجی و با درصد صحت ۸۶٪ بر روی داده‌های آزمون توانسته است موفق به تشخیص فرار مالیاتی گردد. پیری و جباری (۱۳۹۴)، در تحقیق دیگری با عنوان "عوامل مرتبط با به کارگیری حسابرسی داخلی مبتنی بر ریسک"، به پردازش عوامل شرکتی خاص مرتبط با به کارگیری حسابرسی

داخلی مبتنی بر ریسک و بررسی نقش حسابرسی داخلی در مدیریت ریسک بنگاه پرداختند. نمونه‌هایی این پژوهش دربرگیرنده ۵۲ شرکت است. همچنین به منظور گردآوری اطلاعات از پرسشنامه استفاده شده است. نتایج پژوهش نشان می‌دهد تعداد شرکت‌هایی که روش‌های مبتنی بر ریسک را برای طرح‌ریزی برنامه سالانه حسابرسی بکار می‌برند در صنعت مالی نسبت به بخش‌های غیر مالی، بیشتر است. تعداد شرکت‌های که روش‌های مبتنی بر ریسک را برای طرح‌ریزی هر کار حسابرسی بکار می‌برند در صنعت مالی نسبت به بخش‌های غیر مالی، بیشتر است و ارتباطی معنادار بین نقش فعال حسابرسی داخلی در پیاده‌سازی مدیریت ریسک بنگاه و عضویت شرکت‌ها در صنعت مالی، وجود دارد. رضانی و همکاران (۱۳۹۲)، نیز در تحقیقی با عنوان " حسابرسی مالیاتی مبتنی بر ریسک و معیارهای تعیین‌کننده آن با تأکید بر اندازه مؤدیان مالیاتی و تجربه سایر کشورها" ضمن تبیین حسابرسی مالیاتی مبتنی بر ریسک و معیارهای آن، تفکیک مؤدیان براساس اندازه را اولین گام در تحلیل محیط ریسک و در نتیجه مدیریت ریسک می‌داند. در واقع درک طبقات مختلف مؤدیان اجازه می‌دهد تا مقامات مالیاتی، استراتژی‌های مؤثر و هدفمندی برای تمکین مالیاتی ارائه دهند تا مناسب‌ترین عکس‌العمل برای ریسک‌های درآمدی و بهترین پاسخ به خدمات مورد نیاز مؤدیان فراهم گردد.

### ۳- فرضیه پژوهش

فرضیه پژوهش به این صورت بیان می‌شود: ریسک مالیاتی مؤدیان مالیات بر ارزش افزوده با استفاده از روش‌های داده‌کاوی ماشین بردار پشتیبان و رگرسیون لجستیک قابل پیش‌بینی می‌باشد.

### ۴- روش شناسایی تحقیق پژوهش

پژوهش حاضر با توجه به نتایجی که می‌تواند به همراه داشته باشد یک تحقیق بنیادی است زیرا درصد شناسایی عوامل مختلف تأثیرگذار بر رفتار مؤدیان مالیاتی و مدل‌سازی این رفتارها است و از طرف دیگر با توجه به کاربرد این تحقیق که برای مسائل اجرایی در نظام مالیاتی بکار گرفته می‌شود یک تحقیق کاربردی می‌باشد از نظر روش تحقیق با توجه به ماهیت پژوهش در حوزه علوم مالی از روش پیمایشی استفاده می‌نماید. برای شناسایی عوامل مؤثر بر ریسک مالیاتی در مالیات بر ارزش افزوده ابتدا با انجام تحقیقات میدانی، مطالعات کتابخانه‌ای و تجربیات محقق پرسشنامه مبتنی بر روش دلفی تهیه گردیده و با کسب نظر از ۳۴ نفر از خبرگان مالیاتی عوامل مؤثر بر ریسک مالیاتی نهایی گردید. همچنین جامعه آماری شامل اشخاص حقوقی اداره کل مؤدیان بزرگ مالیاتی که طی سال‌های ۱۳۹۰ تا ۱۳۹۳ حسابرسی و مالیات آن‌ها قطعی گردیده می‌باشد - فرآیند نمونه‌گیری تصادفی و به روش کوکران صورت پذیرفته است که ۲۸۱ مؤدی به عنوان نمونه واجد شرایط مورد پژوهش قرار گرفته است.

فرآیند طی شده برای ارائه مدل بهینه به شرح ذیل می‌باشد.

نسبت به جمع‌آوری داده‌ها از بانک‌های اطلاعاتی موجود اقدام و عوامل مؤثر بر ریسک مالیاتی مؤدیان شناسایی گردیده شاخص مؤدیان فاقد ریسک، با ریسک کم و پر ریسک تعیین و داده‌های نمونه به سه مجموعه داده‌های



آموزشی، آزمون و اعتبار سنجی تقسیم گردیده‌اند. مدل‌های با استفاده از داده‌های آموزشی ساخته شده نسبت به آزمون مدل‌ها با استفاده از داده‌های آزمون اقدام گردید. در نهایت آزمون دقت و سنجش اعتبار مدل‌ها در تعیین وضعیت مؤدیان با استفاده از داده‌های اعتبار سنجی انجام و بهترین مدل ارائه گردید.

#### ۴-۱- تبیین متغیر وابسته

متغیر وابسته عبارت است از ریسک مالیاتی که همه انواع عدم قطعیت در رویدادها و گزارشات مالی یک شرکت می‌باشد و در برگزیده عدم قطعیت بکارگیری قانون و مقررات مالیاتی در گزارشات شرکت و همچنین احتمال حسابرسی شدن و عدم قطعیت در حسابداری مالی مالیات‌های دریافتی و کیفیت اطلاعات حسابداری که تصمیمات مالیات برآن اساس گرفته می‌شود (هاتچنرز و ریگو، ۲۰۱۵). در این پژوهش مؤدیان در سه دسته فاقد ریسک، کم ریسک و پر ریسک طبقه‌بندی می‌گردند که معیار اندازه‌گیری آن درصد اختلاف بین مالیات ابرازی با مالیات قطعی است.

#### ۴-۲- تبیین و اندازه‌گیری متغیرهای مستقل

متغیرهای مستقل: این پژوهش در قالب ۵ طبقه شامل ساز و کارهای حاکمیت شرکتی، ویژگی‌های خاص شرکتی، ویژگی‌های ماهیت و فعالیت مؤدیان، ویژگی‌های کنترلی مؤدیان و نسبت‌های مالی قرار می‌گیرند که اجزاء هر کدام از این طبقات به شرح زیر است.

جدول شماره (۱) خلاصه متغیرهای مستقل

شرح متغیرها	شاخص‌ها	نوع متغیر
ساز و کارهای حاکمیت شرکتی	نقش دوگانه مدیر عامل	مجازی (صفر و یک)
	درصد سهام مدیران	مجازی (صفر و یک)
	درصد سهامداران نهادی	مجازی (صفر و یک)
	درصد مدیران غیر مؤلف	مجازی (صفر و یک)
	گزارش حسابرسی مالی	مجازی (صفر و یک)
ویژگی‌های خاص شرکتی	نوع مالکیت	مجازی (صفر و یک)
	نوع شرکت	مجازی (صفر و یک)
	نوع فعالیت	مجازی (چند ارزشی)
	تمرکز شرکت مؤدی	مجازی (صفر و یک)
	مالیات و عوارض ابرازی (فروش - خرید)	پیوسته (ریال)
	مالیات و عوارض قطعی	پیوسته (ریال)
	تفاوت مالیات ابرازی و قطعی	درصد
	اندازه شرکت مؤدی (فروش)	پیوسته (ریال)

شرح متغیرها	شاخص‌ها	نوع متغیر
	عمر شرکت	گسسته (سال)
	سابقه مالیاتی شرکت مؤدی (مراجعه به هیئت)	دو ارزشی (بله - خیر)
	ترکیب درآمدها (معاف و غیرمعاف)	درصد
	ترکیب درآمدها (صادراتی و داخلی)	درصد
ویژگی‌های ماهیت و فعالیت مؤدیان	حساب ارزی درآمدی و هزینه‌ای	مجازی (صفر و یک)
	میزان اعتبارپذیری فعالیت	درصد
	صدور صورتحساب فاکتور شده	مجازی (صفر و یک)
	حسابرس داخلی و کمیته حسابرسی	مجازی (صفر و یک)
ویژگی‌های کنترلی مؤدیان	عضویت در بورس	مجازی (صفر و یک)
	قیمت‌گذاری مصوب کالا و خدمات	مجازی (صفر و یک)
	نوع اظهار نظر حسابرسان	مجازی (چند ارزشی)
	یکی بودن مالکیت و مدیریت	مجازی (صفر و یک)
نسبت‌های مالی	نسبت‌های جاری	پیوسته (درصد)
	نسبت‌های آنی	پیوسته (درصد)
	نسبت‌های بدهی	پیوسته (درصد)
	نسبت بدهی‌های بلند مدت به حقوق صاحبان سهام	پیوسته (درصد)
	نسبت گردش دارایی ثابت	پیوسته (درصد)
	نسبت گردش جمع دارایی‌ها	پیوسته (درصد)
	نسبت سود عملیاتی	پیوسته (درصد)
	نسبت سود خالص	پیوسته (درصد)
	نسبت بازده جمع دارایی‌ها	پیوسته (درصد)
	نسبت بازده حقوق صاحبان سهام	پیوسته (درصد)

#### ۴-۳- آمارهای عددی متغیرهای پژوهش

جدول زیر میزان دامنه تغییرات نمرات متغیرها، میانگین و انحراف معیار آن‌ها اشاره شده است همان‌طور که مشاهده می‌شود تمامی متغیرهای یادشده در دامنه قابل قبولی برای دو شاخص چگونگی و کشیدگی قرار دارند که در این دامنه قابل قبول برای چگونگی و کشیدگی بین ۳ تا ۳- می‌باشد و در صورتی که متغیر در این دامنه قرار داشته باشد دارای توزیع نرمال است.

جدول شماره (۲) آماره‌های عددی متغیرهای پژوهش

	ساز و کارهای حاکمیت شرکتی	ویژگی‌های خاص شرکتی (مؤدیان)	ویژگی‌ها و ماهیت فعالیت مؤدیان	ویژگی‌های سیستم کنترلی مؤدیان	ویژگی‌های نسبت‌های مالی
Mean	25.50	29.75	29.75	21.25	8.50
Standard Error	9.37	11.85	11.17	9.09	3.23
Median	17.00	18.50	20.00	13.50	5.00
Standard Deviation	26.49	33.53	31.60	25.70	9.13
Sample Variance	701.71	1124.21	998.50	660.50	83.43
Kurtosis	-0.80	-0.68	-1.05	0.04	-1.62
Skewness	0.76	0.89	0.77	1.01	0.62
Range	71.00	86.00	80.00	69.00	22.00
Minimum	0.00	1.00	0.00	0.00	0.00
Maximum	71.00	87.00	80.00	69.00	22.00
Sum	204.00	238.00	238.00	170.00	68.00

#### ۴-۴- تکنیک‌های داده‌کاوی مورد استفاده در پژوهش

در این تحقیق برای سنجش ریسک مؤدیان مالیات بر ارزش افزوده (اشخاص حقوقی بزرگ در تهران) و تفکیک آن به مؤدیان فاقد ریسک، کم ریسک و پر ریسک از تکنیک‌های داده‌کاوی رگرسیون لجستیک و ماشین بردار پشتیبان استفاده می‌شود در میان تکنیک‌های مختلف داده‌کاوی، رگرسیون لجستیک به طور گسترده‌ای در پیش بینی احتمال باینری یا بیشتری برای متغیر گسسته مانند مشتریان خوب یا بدو یا تمکین مالیاتی یا عدم تمکین مالیاتی مؤدیان می‌باشد.

ماشین بردار پشتیبان یکی از روش‌های یادگیری با نظارت است که از آن برای طبقه‌بندی و رگرسیون - استفاده می‌کنند این روش از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای طبقه بندی از جمله شبکه‌های عصبی نشان داده است. مبنای کار این روش دسته بندی خطی داده‌ها است. در تقسیم خطی داده‌ها سعی می‌کنیم خطی را انتخاب کنیم که حاشیه اطمینان بیشتری داشته باشد.

#### ۵- یافته‌های پژوهش

با استفاده از پیشینه پژوهش و استفاده از روش دلفی متغیرهای کیفی و کمی و نسبت‌های مالی انتخاب گردید. با استفاده از نرم افزار ۱۸ Spss Moldeler Verin کارایی این مدل بیشتر در پیش بینی ریسک مالیاتی

مؤدیان بررسی قرار گرفت. به گونه‌ای که ابتدا داده‌های خام را جمع‌آوری و طبقه‌بندی کردیم و سپس در قالب Excel 2016 وارد نرم‌افزار کردیم و با فیلتر کردن داده‌ها عمل پالایش و پیش‌پردازش روی داده‌ها انجام گردید. سپس متغیرهای ورودی و خروجی مشخص و بر اساس ایجاد پارتیشن به مقدار آموزش، آزمون و ارزیابی (اعتبار مدل) شخصی‌گردید. در این تحقیق ۷۰٪ را به عنوان داده‌های آموزش، ۲۰٪ را برای داده‌های آزمون و ۱۰٪ باقی‌مانده را به عنوان داده‌های لازم جهت اعتبار مدل انتخاب نموده‌ایم الگوریتم‌های ماشین بردار پشتیبان و درخت تصمیم را ساخته و در آن هدف و متغیرهای ورودی تعیین گردیده‌اند و با اجراء الگوریتم‌های یاد شده مدل مذکور و خروجی‌های آن مشخص گردید که جداول زیر نمایشگر میزان صحت هر یک از مدل‌ها در تشخیص و خوشه‌بندی داده‌ها در شش مدل به کار رفته (دو مدل ماشین بردار پشتیبان و چهار مدل درخت تصمیم) می‌باشد.

#### ۵-۱- جدول نتایج ماشین بردار پشتیبان

این مدل در سه بخش آموزش، آزمون و اعتبار سنجی مدل بدست آمده به روش (LSVM) به ترتیب دارای صحت ۶۸٪، ۶۰٪ و ۷۰٪ می‌باشد بر این اساس تا ۷۰ درصد از مؤدیان به وسیله این الگوریتم خوشه‌بندی به درستی باتوجه به واکاوی داده‌های موجود تفکیک خواهند شد.

جدول شماره (۳) نتایج مدل LSVM

Partition	آموزش	آزمون	اعتبار سنجی		
Correct	539	137	77	68.4%	70%
Wrong	249	89	33	31.60%	30%
Total	788	226	110		

جدول شماره (۴) نتایج مدل SVM

Partition	آموزش	آزمون	اعتبار سنجی		
Correct	516	135	59	65.48%	53.64%
Wrong	272	91	51	34.52%	46.36%
Total	788	226	110		

این مدل در سه بخش آموزش، آزمون و اعتبار سنجی مدل بدست آمده به روش (SVM) به ترتیب دارای درجه صحت تقریبی ۶۵٪، ۶۰٪ و ۵۴٪ می‌باشد که بر این اساس تا ۵۴ درصد از مؤدیان به وسیله این الگوریتم خوشه‌بندی، به درستی باتوجه به واکاوی داده‌های موجود تفکیک خواهند شد.

### ۵-۲- جداول نتایج رگرسیون لجستیک

این مدل در سه بخش آموزش، آزمون و اعتبار سنجی مدل بدست آمده به روش (Logistic) به ترتیب دارای درجه صحت تقریبی ۶۹، ۶۷ و ۷۰ درصد می‌باشد، بدیهی است که بر این اساس تا ۷۰ درصد از مؤدیان به وسیله این الگوریتم خوشه‌بندی، به درستی با توجه به واکاوی داده‌های موجود تفکیک خواهند شد.

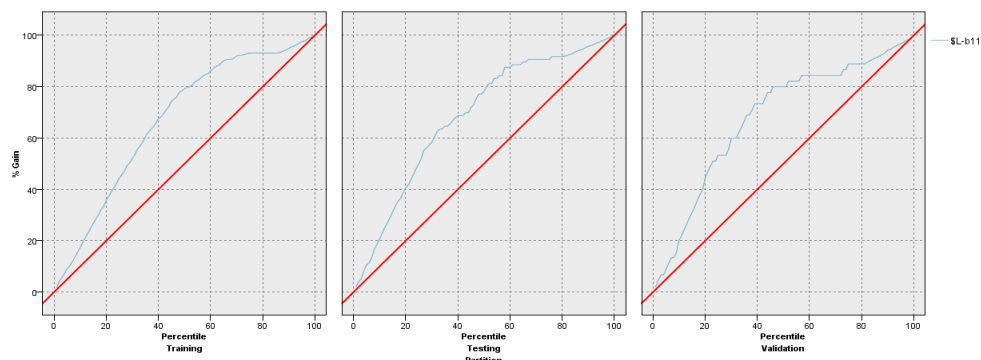
جدول (۵) نتایج مدل Logistic

Partition	آموزش		آزمون		اعتبار سنجی	
Correct	548	69.54%	151	66.81%	77	70%
Wrong	240	30.46%	75	33.19%	33	30%
Total	788		226		110	

### ۵-۳- ادغام نتایج جهت افزایش کارآمدی مدل تحقیق

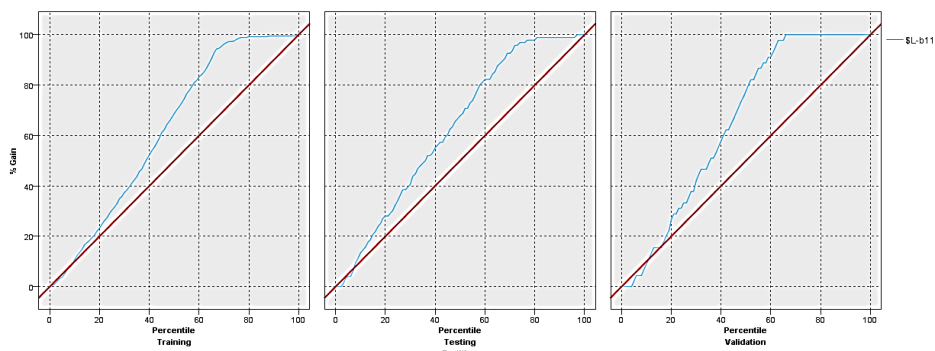
دو مدل Logistic و SVM با بالاترین توان ممکن در حدود ۷۰ درصد صحت می‌توانند به عنوان کاندیدای مناسبی جهت خوشه‌بندی مؤدیان پر ریسک مورد استفاده قرار گیرند.

شکل‌های (۳) و (۴) بهبود ایجاد شده توسط هریک از مدل‌ها را جهت پیش‌بینی خوشه مورد نظر بر اساس ریسک اعتباری مؤدیان در هریک از الگوریتم‌های ماشین بردار پشتیبان خطی و رگرسیون لجستیک نمایش می‌دهد:



Logistic Evaluation Graphs

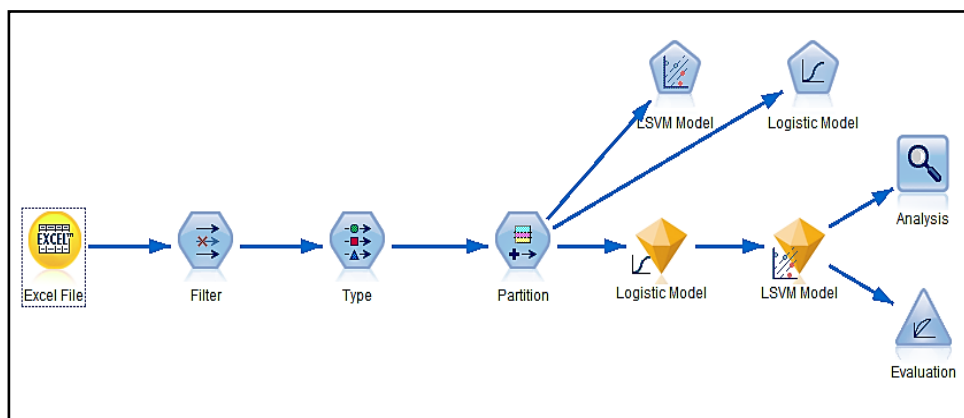
شکل (۳) نمایش ارزیابی نحوه بهبود رگرسیون لجستیک Logistic Regration



LSVM Evaluation Graphs

شکل (۴) نمایش ارزیابی نحوه بهبود ماشین بردار پشتیبان خطی LSVM

از سوی دیگر با ترکیب نتایج در نرم افزار به شکل (۵) می‌توان نتایج مدل‌ها را علاوه بر نحوه توافق و همگرایی آنان در دسته بندی به نوع دیگر نیز بهبود داده و همانگونه که در جداول (۹) و (۱۰) نیز ملاحظه می‌گردد مدل رگرسیون لجستیک با نتایج مناسب تر در بهبود می‌تواند با ادغام با روش ماشین بردار پشتیبان و همگرایی اعتبار نتایج را تا نزدیک ۸۳ درصد بهبود دهد:



Combined Result of Logistic and LSVM Model

شکل (۵) نحوه ادغام نتایج الگوریتم‌ها در نرم افزار SPSS Modeler Ver. 18

همان‌گونه که در نتایج جداول (۹) و (۱۰) نیز قابل مشاهده است، میزان همپوشانی نتایج این دو مدل نزدیک به ۶۸ درصد و همچنین مطابقت همگرایی نتایج آنان در صورت ادغام با نتایج مورد نظر نزدیک به ۸۳ درصد است:

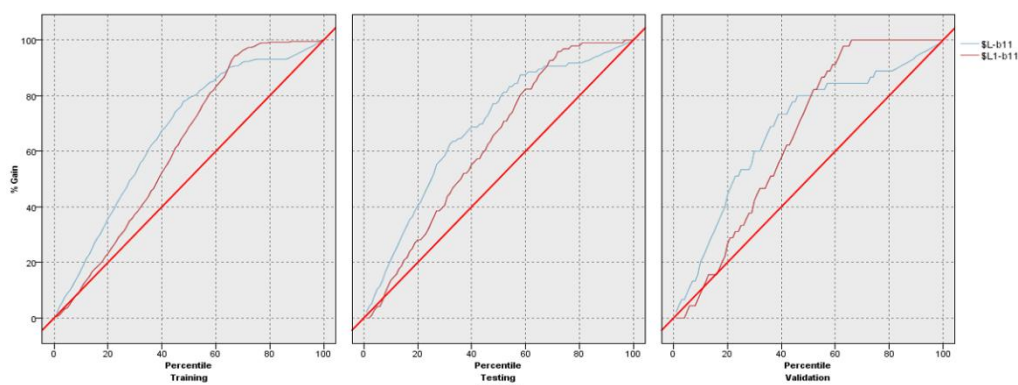
جدول (۹) نتایج مدل همپوشانی بین الگوریتم‌های LSVM & Logistic

Partition	آموزش		آزمون		اعتبار سنجی	
Agree	548	69.54%	152	67.26%	75	68.18%
Disagree	240	30.46%	74	32.74%	35	31.82%
Total	788		226		110	

جدول (۱۰) مطابقت همگرایی نتایج و افزایش اعتبار مدل تحقیق در صورت ادغام

Partition	آموزش		آزمون		اعتبار سنجی	
Correct	446	81.39%	114	75%	62	82.67%
Wrong	102	18.61%	38	25%	13	17.33%
Total	548		152		75	

شکل (۶) نیز نحوه همگرایی این نتایج را بطور همزمان بایکدیگر مقایسه می‌نماید که نشان از ارجحیت مدل رگرسیون لجستیک در این رابطه دارد:



Comparing Evaluation LSVM and Logistic

شکل (۶) مقایسه نتایج ارزیابی نحوه بهبود رگرسیون لجستیک و Logistic Regression و ماشین بردار پشتیبان خطی LSVM

## ۶- نتیجه گیری و بحث

مالیات ابزار تأمین مالی دولت‌ها می‌باشد که پرداخت صحیح و به موقع آن بخش با اهمیتی از مسئولیت اجتماعی بنگاه‌های اقتصادی تلقی می‌گردد. که در این راستا بهبود کارایی تشخیص و وصول مالیات از فعالیت‌های اقتصادی با اهمیت تلقی می‌گردد. یکی از گام‌های اساسی در تحقق این امر بهبود مدیریت تمکین

مالیاتی با شناسایی و ارزیابی عوامل مؤثر بر ریسک مالیاتی می‌باشد. با سنجش عواملی نظیر اندازه شرکت، نوع مالکیت، نوع صنعت و همچنین نسبت‌های مالی و سایر مؤلفه‌های تأثیرگذار مؤدیان با ریسک بالا شناسایی و در اولویت حسابرسی قرار می‌گیرند مشروط سازی انتخاب برای حسابرسی موجب می‌گردد تا سهم بیشتری از تقلب‌های مالی و مالیاتی مؤدیان شناسایی گردد. نتایج حاصل از تحقیقات نشان می‌دهد که استفاده از تکنیک‌های داده‌کاوی امکان تشخیص فرار مالیاتی را فراهم می‌نماید و در نتیجه بهبود کارایی نظام‌های مالیاتی را به دنبال دارد. در این تحقیق از تکنیک‌های داده‌کاوی برای سنجش ریسک مالیاتی مؤدیان در نظام مالیات بر ارزش افزوده کشور استفاده گردید. جامعه آماری پژوهش اشخاص حقوقی بزرگ در تهران که در قلمرو زمانی ۱۳۹۰ تا ۱۳۹۳ مورد حسابرسی قرار گرفته‌اند، می‌باشد از تکنیک‌های داده‌کاوی ماشین بردار پشتیبان و رگرسیون لجستیک استفاده گردیده‌اند. نتایج تحقیق نشان می‌دهد که تکنیک‌های Logistic و L SVM دارای توان ارزیابی معادل ۷۰٪ برای سنجش ریسک مالیاتی مؤدیان مالیات بر ارزش افزوده می‌باشند و ترکیب و ادغام دو مدل Logistic و L SVM توان ارزیابی و سنجش ریسک مالیاتی را افزایش داده به گونه‌ای که میزان همپوشانی نتایج این دو مدل نزدیک به ۶۸٪ و همچنین مطابقت همگرایی نتایج آنان در صورت ادغام با نتایج مورد انتظار نزدیک به ۸۳٪ می‌باشد. بنابراین می‌توان نتیجه گرفت که استفاده از تکنیک‌های داده‌کاوی راهکاری مناسبی برای کمک به بهبود مدیریت تمکین مالیاتی از طریق سنجش ریسک مالیاتی مؤدیان در نظام مالیات بر ارزش افزوده می‌باشد و برای افزایش توان تکنیک‌های داده‌کاوی برای سنجش ریسک مالیاتی پیشنهاد می‌گردد. با استفاده از ظرفیت‌های قانونی مواد ۱۶۹ و ۱۶۹ مکرر موضوع قانون مالیات‌های مستقیم نسبت به تکمیل و تقویت بانک‌های اطلاعاتی سازمان اقدام نمود تا با شناسایی و ارزیابی ریسک مالیاتی مؤدیان در نظام مالیات بر ارزش افزوده زمینه انتخاب مؤدیان برای حسابرسی را با استفاده از قواعد ریسک فراهم نمود. مشروط سازی انتخاب برای حسابرسی بر پایه قواعد ریسک مالیاتی موجب می‌گردد تا سهم بیشتری از تقلب‌های مالیاتی کشف گردد.

#### فهرست و منابع

- \* موسوی جهرمی، یگانه، فرهاد طهماسبی بلداجی و نرگس خاکی (۱۳۸۸)، «فرار مالیاتی در نظام مالیات بر ارزش افزوده: یک مدل نظری»، فصلنامه تخصصی مالیات، دوره جدید، شماره پنجم (مسلسل ۵۳)، ص ۲۷-۳۸.
- \* پیری، صفر و جباری، حسین (۱۳۹۴) "عوامل مرتبط با به کارگیری حسابرسی داخلی مبتنی بر ریسک"، اولین همایش ملی حسابداری و حسابرسی
- \* رحمانی و ابوحمز (۱۳۹۴) "حسابرسی مالیاتی و تمکین مالیات شرکتی"، مجله حسابرسی، سال اول، شماره ۲، پاییز ۹۳
- \* دستگیر، محسن و غریبی، مریم (۱۳۹۴) «کاربست روش‌های داده‌کاوی به منظور ارتقای عملکرد و تشخیص فرار مالیاتی» پژوهشنامه مالیات، ش ۲۰۸ ص ۹۵-۱۱۶



- \* رمضانی، سید مهدی و دستگیر، محسن و عسکری، علی و خانی، عبدالله (۱۳۹۲)، «حسابرسی مالیاتی مبتنی بر ریسک و معیارهای تعیین کننده آن با تأکید بر اندازه مؤدیان مالیاتی و تجربه سایر کشورها». پژوهش نامه مالیات، شماره ۶۶، تابستان ۱۳۹۲.
- \* سازمان مالیاتی ایران (شهریور ۱۳۹۱)، «ظرفیت‌های مالیاتی اقتصاد ایران، چالش‌ها و راه‌کارهای توسعه آن»، دفتر پژوهش و برنامه ریزی.
- \* Alabede J O, Zaimah B t. Zainol Ariffin and Kamil Md Idris.2011.Determinants of Tax compliance Behaviour : A Proposed Model for Nigeria.International Research Journal of Finance and Economics,ISSN1450-2887 Issue78.
- \* Alabede J O, Zaimah Zainol Ariffin and Kamil Md Idris.2011.Individual Taxpayers Attitude and compliance Behaviour in Nigeria: The Moderating Role of Financial Condition and Risk Preference. Journal of Accounting and Taxation Vol.3(5),pp.91-104
- \* Abebaw Kassie, Melat, Abera (2016), Factors Affecting Tax Audit Effectiveness Evidence From Large Tax Payers Office of Ethiopian Revenue And Customs Authority , Thesis - Accounting & Finance
- \* James Alm, Michele Bernasconi , Susan Laury , Daniel J. Lee (2016), Culture, Compliance, and Confidentiality: Taxpayer Behavior in the United States and Italy, University Ca' Foscari of Venice, Dept. of Economics Research Paper Series No. 36
- \* Kuo-Wei Hsu , Nishith Pathak, Jaideep Srivastava, Greg Tschida, Eric Bjorklund (2015), "Data Mining Based Tax Audit Selection: A Case Study of a Pilot Project at the Minnesota Department of Revenue", Real World Data Mining Applications, Volume 17 of the series Annals of Information Systems pp 221-245
- \* 20. Khwaja, Munawer Sultan, Rajul Awasthi, and Jan Loeprick (2011). "Risk-Based Tax Audits: Approaches and Country Experiences", The International Bank for Reconstruction and Development / The World Bank, 1818 H Street NW, Washington DC20433.
- \* LIDIJA HAUPTMAN, MIRJANA HORVAT & ROMANA KOREZ-VIDE (2014) Improving Tax Administration's Services as a Factor of Tax Compliance: The Case of Tax Audit, LEX LOCALIS - JOURNAL OF LOCAL SELF-GOVERNMENT Vol. 12, No. 3, pp. 481-501, July 2014
- \* Lu, S.X. and Wang, X.Z., 2004, August. A comparison among four SVM classification methods:LSVM, NLSVM. SSVM and NSVM. In Machine learning and cybernetics,2004. Proceeding of 2004 international conference on (vol.7,pp.4277-4282). Ieee.
- \* Rayisankar, P., V. Ravi, G. Raghava Rao & Bose. (2011). Detection of Financial Statement Fraud and Feature Selection Using Data Mining Techniques. Decision Support Systems, 50,491-500