



Promotional-Science Article

## A comprehensive analysis of keywords co-occurrence network and the most cited journals on data mining techniques in insurance industry using scientometrics approach

H. Hajian<sup>1\*</sup>, A. Zarjini<sup>2</sup>

<sup>1</sup> Department of Computer Science, Faculty of Science, Ontario University of Technology, Oshawa, Ontario, Canada

<sup>2</sup> Department of Technology Management, Faculty of Management and Economics, Islamic Azad University, Science and Research Unit, Tehran, Iran

### ARTICLE INFO

#### Article History:

Received: 23 March 2023

Revised: 29 July 2023

Accepted: 26 September 2023

#### Keywords:

Data mining

Fraud detection

Insurance industry

Risk management

Scientometrics

### ABSTRACT

**BACKGROUND AND OBJECTIVES:** Data mining is known as a process of discovering patterns in large datasets through a combination of statistical tools and techniques. In recent years, data mining and its applications in different businesses have increasingly grown. Insurance industry is one of the data-driven businesses whose survival is so dependent on satisfying customers besides achieving the highest benefit. Information or data is a vital asset of the insurance industry; accordingly, using data mining techniques to discover patterns behind large datasets is a need. Having seen the increasingly high rate of information technology and recorded data in data-driven businesses, lots of industries like the insurance industry have been urged to use state-of-the-art data mining techniques to turn raw data into useful information using Big Data Analytics.

**METHODS:** Looking at the current research on data mining applications in the insurance industry proves the fact that we should recognize the state-of-the-art techniques in data mining and set new strategies to focus on Big Data Analytics more. Big Data Analytics consists of the algorithms which are more efficient and less time-consuming so it can help to identify patterns and rules in complex datasets. For this purpose, this paper presents a comprehensive literature review regarding the usage of data mining techniques in the insurance industry by the scientometrics approach. For this purpose, first we searched and gathered bibliometrics files of recent researches from Web of Science and Scopus into four different scenarios. In each scenario, we looked up for different keywords regarding "Data Mining", "Insurance Industry", and "Risk Management" to make sure that all the results would be specifically focused on the research topic. Then, we used R programming software to analyze the results of each scenario based on keywords co-occurrence in the given research.

**FINDINGS:** The results of keywords co-occurrence and a word cloud of recent research confirm that insurance companies should focus on Big Data Analytics instead of traditional data processing to get information systematically from too large or complex datasets. Big Data Analytics has been used for several years, but in recent years many data-driven businesses, like the insurance industry, have used its techniques associated with risk and risk factor identification. Risk management in the insurance industry has been widely considered in recent researches. Therefore, in this paper, some high-ranked journals and the most significant researches have been identified and recommended in order to pave the way for future researches in this field.

**CONCLUSION:** We hope that the comprehensive literature review provided in this paper can help the researchers to focus on the relative journals and researches published then get into more details. For this purpose, the lists of all journals and conferences besides the most cited researches are provided in the experimental section of this paper. Also, the ranking list of different countries from all around the world related to data mining and Big Data Analytics in the insurance industry is presented. The results show that Iran is the 15th country that uses data mining techniques and it is the 17th country in the world focusing on risk management in the insurance industry.

\*Corresponding Author:

Email: [hooria\\_hajiyani@ontariotechu.net](mailto:hooria_hajiyani@ontariotechu.net)

Phone: +9821 44865154

ORCID: [0000-0003-1098-4552](https://orcid.org/0000-0003-1098-4552)

DOI: [10.22056/ijir.2024.01.06](https://doi.org/10.22056/ijir.2024.01.06)

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).





## مقاله علمی ترویجی

# تحلیلی بر کاربردهای داده‌کاوی در صنعت بیمه براساس شبکه هم‌رخدادی واژگان‌ها و شناسایی معتبرترین مجلات با شاخص استناد به پژوهش‌های علمی با استفاده از رویکرد علم‌سنجی

حورا حاجیان<sup>۱\*</sup>، اعظم زرچینی<sup>۲</sup>

<sup>۱</sup> گروه علوم کامپیوتر، دانشکده علوم، دانشگاه فنی انتاریو، اوشاوا، انتاریو، کانادا

<sup>۲</sup> گروه مدیریت تکنولوژی، دانشکده مدیریت و اقتصاد، دانشگاه آزاد اسلامی واحد علوم و تحقیقات، تهران، ایران

## اطلاعات مقاله

### تاریخ‌های مقاله:

تاریخ دریافت: ۰۳ فروردین ۱۴۰۲

تاریخ داوری: ۰۷ مرداد ۱۴۰۲

تاریخ پذیرش: ۰۴ مهر ۱۴۰۲

### کلمات کلیدی:

داده‌کاوی

شناسایی ریسک

صنعت بیمه

علم‌سنجی

کشف تخلفات

## چکیده:

**پیشینه و اهداف:** با توجه به رشد روزافزون فناوری اطلاعات و کسب‌وکارهای مبتنی بر داده و روند رو به رشد اطلاعات در پایگاه‌های داده، صنعت بیمه نیز به‌عنوان یکی از کسب‌وکارهای یادشده با حجم انبوهی از اطلاعات ذخیره‌شده روبه‌رو است. این حجم انبوه داده‌ها شرکت‌های بیمه‌ای را ملزم می‌سازد تا با شیوه‌های داده‌کاوی، دانش موجود در پایگاه‌های داده را کشف کنند. از سوی دیگر توسعه روش‌های نوین داده‌کاوی و کشف الگوهای پنهان در پایگاه‌های داده به‌سرعت رو به افزایش است.

**روش‌شناسی:** در این مقاله با هدف بررسی جامع بر کاربرد داده‌کاوی در صنعت بیمه دنیا، جدیدترین مستندات علمی مربوطه در چهار سناریوی مختلف شامل تحقیقات صنعت بیمه، داده‌کاوی در بیمه، کشف تخلفات و ریسک در بیمه از وبسایت علمی Web of Science دریافت و با رویکرد علم‌سنجی تحلیل و بررسی شده است. نتایج به‌دست‌آمده نشان می‌دهد رشد داده‌های موجود در شرکت‌های بیمه و شناسایی ریسک و عامل‌های آن، شرکت‌های بیمه و پژوهشگران را به بررسی روش‌های نوین مدیریت پایگاه‌های داده و داده‌های بزرگ و همچنین شناسایی ریسک و فاکتورهای آن با استفاده از شیوه‌های داده‌کاوی ملزم می‌سازد.

**یافته‌ها:** بدین منظور برای تسهیل در تحقیقات پژوهشگران آینده و استفاده از مطالعات پیشین داده‌کاوی در صنعت بیمه، در این مقاله شبکه هم‌استنادی کشورهای مختلف و معتبرترین مجلات علمی براساس شاخص ارزیابی استناد به مقالات معرفی شده است.

**نتیجه‌گیری:** در نتایج این مقاله فهرست مجلات معتبر علمی و موضوعاتی را که در هریک به‌طور تخصصی به آن پرداخته شده، ارائه کرده‌ایم تا مرور ادبیات جامع در این مقادیر مبروری مسیر مطالعات آینده صنعت بیمه را هموارتر سازد. همچنین رتبه‌بندی برترین کشورهای جهان که متمرکز بر مطالعات داده‌کاوی در صنعت بیمه بوده است، آورده شده و رتبه ایران در میان سایر کشورها نیز مشخص شده است. استفاده از مطالعات و مجلات علمی معتبر می‌تواند راه پژوهش‌های آینده را در پیشرفت این مطالعات در صنعت بیمه ایران را هدفمندتر کند.

### نویسنده مسئول:

ایمیل: [hooria\\_hajijan@ontariotechu.net](mailto:hooria_hajijan@ontariotechu.net)

تلفن: ۰۹۸۲۱ ۴۴۸۶۵۱۵۴

ORCID: 0000-0003-1098-4552

DOI: 10.22056/ijir.2024.01.06

توجه: مدت‌زمان بحث و انتقاد برای این مقاله تا ۱ آوریل ۲۰۲۴ در وبسایت IJIR در «نمایش مقاله» باز است.

علم‌سنجی (*Journal of scientometrics*) این حوزه گسترش بیشتری یافت. علم‌سنجی را بررسی جنبه‌های کمی علوم، پژوهش‌ها و سیاست‌های علمی گویند که هدف از آن ارزشیابی آخرین مستندات علمی و پژوهش‌های انجام‌شده در حوزه علمی خاص و عوامل مؤثر در رشد و تغییرات آن در طول زمان است (Chen et al., 2018). در ادامه این مقاله ضمن بررسی جامع روش‌های مختلف داده‌کاوی و کاربرد آن در صنعت بیمه با رویکرد علم‌سنجی، به بررسی روند هم‌رخدادی واژگان‌ها، هم‌استنادی کشورهای مختلف به‌منظور شناخت و استفاده از مستندات علمی کشورهایی که شرایط اقتصادی و سیاسی مشابه در صنعت بیمه دارند و همچنین مجلات علمی دارای بیشترین استناد در چهار سناریوی مختلف پرداخته شده است. مقاله پیش رو در پنج بخش اصلی تهیه شده که بخش اول آن شامل چکیده و مقدمه است و در آن توضیحی درباره مفاهیم علم داده‌کاوی و مزایای کاربرد آن در صنایع ارائه شده است. بخش دوم نگاهی بر داده‌کاوی و کاربردهای آن در صنعت بیمه و همچنین پیشینه علم‌سنجی به‌عنوان روشی علمی برای سنجش و ارزیابی مستندات یک حوزه خاص دارد. در بخش سوم با بیان روش‌شناسی پژوهش و آمار داده‌های به‌دست‌آمده از مستندات علمی گذشته، به بررسی پژوهش‌های علمی مرتبط با صنعت بیمه، کاربرد داده‌کاوی در صنعت بیمه، مدیریت ریسک و کشف تخلفات در بیمه پرداخته‌ایم. در آخر بخش چهارم نتیجه‌گیری مقاله و بخش پنجم نیز شامل منابع استفاده‌شده در این پژوهش است.

### مروری بر پیشینه پژوهش

در این بخش ابتدا پرکاربردترین روش‌های داده‌کاوی در پژوهش‌های متمرکز در صنعت بیمه که طی سال‌های اخیر به چاپ رسیده، به تفصیل بررسی شده است. سپس پیشینه علم‌سنجی و روش‌های مختلف آن برای بررسی روند پژوهش‌های یک حوزه علمی خاص به‌منظور هدفمندسازی مطالعات آینده شرح داده شده است.

#### داده‌کاوی در صنعت بیمه

با مروری بر مطالعات انجام‌شده در صنعت بیمه در طول دهه گذشته، مدیریت ریسک و کشف تخلفات از پرکاربردترین پژوهش‌های مبتنی بر داده‌کاوی در صنعت بیمه بوده است (Boodhun and Jayabalan, 2018; Hessami, 2018; Kara et al., 2020; Wuyu and Cerna, 2019) از کاربردهای داده‌کاوی در صنایع مختلف می‌توان به کشف تخلفات اشاره کرد. (Phua et al., 2010) به بررسی مطالعات انجام‌شده در حوزه کشف تخلفات صنایع مختلف با رویکرد داده‌کاوی پرداخته است. این مطالعه پوشش مناسبی از روش‌های حل و رویارویی با داده‌ها در صنایع مختلف ارائه می‌دهد. کشف تخلفات در صنعت بیمه‌گری در حوزه‌های مختلف نظیر اتومبیل و درمان نیز یک از رایج‌ترین روش‌های بهره‌گیری از داده‌کاوی است (Ghorbani and Farzai, 2018; Liu et al., 2018; Wang and Xu, 2018)

علم داده‌کاوی فرایند انتخاب، کشف و مدل‌سازی حجم زیادی از داده‌ها به‌منظور کشف الگوهای ناشناخته است. با توجه به مقادیر زیاد داده‌ها در صنعت بیمه، می‌توان گفت فرایند داده‌کاوی در این صنعت امری اجتناب‌ناپذیر است. داده‌کاوی فناوری جدید و قدرتمندی است که با ظرفیت بسیار بالای خود به یاری شرکت‌های بیمه آمده است. به‌عبارت‌دیگر، داده‌کاوی ترکیبی از روش‌های یادگیری ماشین، تشخیص الگو، آمار، نظریه پایگاه داده و خلاصه کردن و یافتن ارتباط بین مفاهیم و الگوها به‌صورت خودکار از پایگاه داده‌های انبوه است (Liao et al., 2012). هدف اصلی داده‌کاوی را می‌توان کمک به فرایند تصمیم‌گیری مدیران سازمان‌ها از طریق کشف دانش نهفته در پایگاه‌های داده دانست. دو کاربرد اصلی داده‌کاوی عبارت‌اند از پیش‌بینی و همچنین یافتن روابط و الگوهای ناشناخته و توصیفات که شرح یک پایگاه داده بزرگ را ارائه می‌دهد. از آنجاکه در صنعت بیمه با حجم زیادی از داده‌ها سروکار داریم، داده‌کاوی با تمرکز بر کشف مهم‌ترین اطلاعات و دانش موجود در داده‌ها کمک شایان توجهی به تحلیل‌گران شرکت‌های بیمه، به‌خصوص در زمینه پیش‌بینی خسارت‌های تقلبی، پوشش‌های درمانی واهی، شناسایی فاکتورهای ریسک برای پیش‌بینی میزان سود و خسارت (Kaščelan et al., 2016). تحلیل سطح مشتری (Boroufar et al., 2017). تحلیل فروش و بازاریابی، گسترش محصولات و خدمات جدید، بیمه اتکایی، تحلیل مالی و محاسبه ذخایر فنی شرکت‌های بیمه کرده است (Jahangiri et al., 2014). به‌عبارت‌دیگر، شرکت‌های بیمه‌ای با حجم انبوهی از اطلاعات ذخیره‌شده روبه‌رو هستند و پایگاه‌های داده به‌سرعت در حال رشدند که این حجم زیاد از داده‌ها فرصت‌های بسیاری را برای یافتن اطلاعات ارزشمند پنهان در پایگاه‌های داده فراهم آورده است (Qarakhani and Abolqasmi, 2019). با پیشرفت روزافزون کاربرد پایگاه داده‌ها، حجم داده‌های ثبت‌شده به‌طور متوسط هر پنج سال دو برابر شده که در این میان سازمان‌هایی موفق‌اند که بتوانند حداقل ۷ درصد از داده‌هایشان را تحلیل کنند (Rahmani and Haji Zain al-Abidini, 2014). رشد سریع پایگاه‌های داده و نیاز به استخراج اطلاعات و دانش از آن‌ها مستلزم استفاده از روش‌های داده‌کاوی و بهبود آن است که باعث شده داده‌کاوی در سال‌های اخیر به حوزه پژوهشی مهم و رو به رشدی تبدیل شود (Liao et al., 2012). اگرچه داده‌کاوی به‌صورت گسترده در صنعت بیمه به کار برده شده، اما با نگاه به روش‌های متنوع و متدهای بهبودیافته آن، تنها شرکت‌هایی از مزایای رقابتی آن بهره‌مند می‌شوند که روش‌های داده‌کاوی را به صورت هدفمند و با نگاه به روند پژوهش‌های روز دنیا انتخاب کنند. از طرفی، به‌منظور بررسی پژوهش‌های انجام‌شده در یک حوزه علمی خاص، از روشی به نام علم‌سنجی استفاده می‌شود. اصطلاح علم‌سنجی را نخستین بار در سال ۱۹۶۹ به زبان روسی و نالیموف (Nalimov) و مولچنکو (Mulchenko) به کار گرفتند و در سال ۱۹۷۸ با تأسیس مجله

### پیشینه علم‌سنجی

در سال‌های اخیر، علم‌سنجی نقشی اساسی در ارزیابی پژوهش‌های علمی ایفا کرده است (Mingers and Leydesdorff, 2015). با دریافت فایل داده‌های منبع‌شناسی برای هر رشته علمی از وب‌سایت‌هایی علمی چون Web of Science یا Scopus می‌توان به ارزیابی پژوهش‌های یک حوزه علمی خاص پرداخت. از مهم‌ترین مزایای استفاده از مفهوم منبع‌شناسی برای بررسی پژوهش‌های علمی می‌توان به تحلیل داده‌هایی که ذات علمی یک پژوهش را نشان می‌دهند، مانند استناد به مقالات دیگر پژوهشگران و هم‌استنادی مقالات به‌عنوان شاخص ارزشمندی و میزان تأثیر یک پژوهش، اشاره کرد (Chen et al., 2018). علم‌سنجی روشی کارآمد برای شناسایی الگوها و روندها در یک حوزه علمی در طول زمان است که شامل شیوه‌های مختلفی برای بررسی پژوهش‌های انجام‌شده در یک رشته علمی خاص است. از این میان می‌توان به شبکه هم‌رخدادی واژگان- (Keywords co-occurrence)، هم‌استنادی مقالات (Citation)، خوداستنادی مقالات (Co-citation)، ضریب تأثیر مجله (Impact factor)، روند پژوهش‌های آینده و مجلاتی که دارای بیشترین مقالات چاپ‌شده با موضوع مشابه‌اند، اشاره کرد (Mingers and Leydesdorff, 2015). روش تحلیل هم‌رخدادی واژگان‌ها نخستین بار برای ترسیم پویایی علم پیشنهاد شد و سپس مورد توجه بسیاری از پژوهشگران در حوزه‌های مختلف علمی در سراسر دنیا قرار گرفته است (Abbas et al., 2019). مرور ادبیات موضوع نشان می‌دهد که تاکنون مطالعه‌ای در خصوص بررسی روند هم‌رخدادی واژگان و شناسایی و استناد به مقالات چاپ‌شده در مجلات معتبر در رابطه با کاربرد داده‌کاوی در صنعت بیمه انجام نشده است. از این رو این مقاله با هدف نگرشی تحلیلی در پژوهش‌های حوزه داده‌کاوی در صنعت بیمه در جهان، نقشه هم‌رخدادی واژگان‌های پرکاربرد و هم‌استنادی برترین مطالعات کشورهای مختلف در جهان را ترسیم کرده تا گامی بزرگ در راستای هدف‌مندی پژوهش‌های آینده در صنعت بیمه ایران باشد. همچنین شناسایی کشورهای پیشرو در مطالعات صنعت بیمه به‌خصوص پژوهش‌های مبتنی بر روش‌های داده‌کاوی، می‌تواند راه را برای دسترسی به پژوهش‌های مشابه در این حوزه برای پژوهشگران آینده هموار سازد.

### روش‌شناسی پژوهش

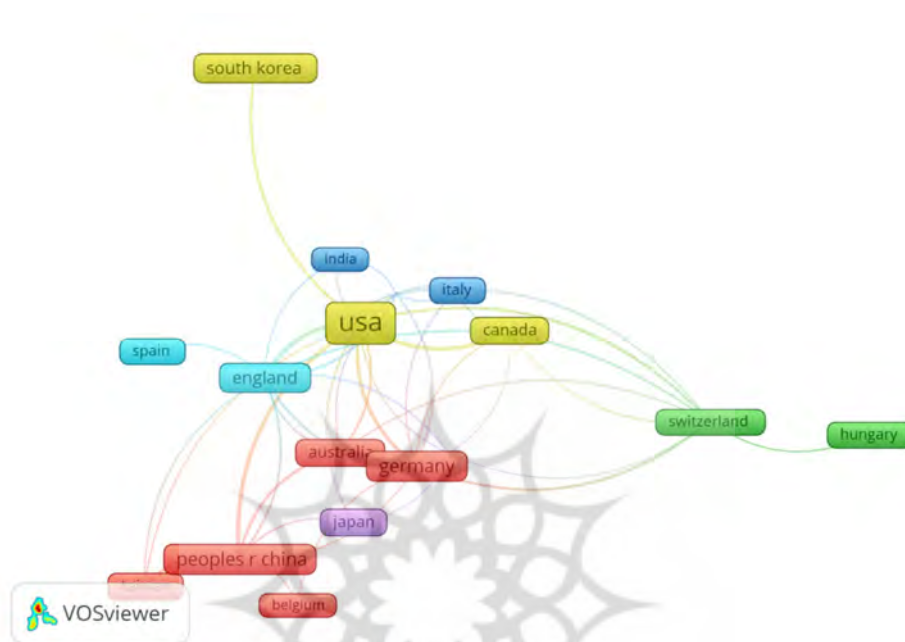
این پژوهش با هدف سوق دادن مطالعات مبتنی بر کاربرد داده‌کاوی در صنعت بیمه انجام شده تا پژوهشگران، دانشگاهیان، مؤسسات پژوهشی و شرکت‌های بیمه با نگاهی بر روند پژوهش‌های مشابه در سال‌های اخیر، پژوهش‌های آینده را هدفمندتر انتخاب کنند. با توجه به پیشرفت روزافزون فناوری اطلاعات در صنعت بیمه و نیاز به روش‌های نو در راستای بهبود عملکرد خدمت‌رسانی شرکت‌های بیمه‌ای به مشتریان، این مقاله می‌تواند تغییر چشمگیری در نگرش سازمان‌ها و مدیران صنعت بیمه ایران به کاربرد تحلیلی

کشف تخلفات صنعت بیمه به‌خصوص در حوزه درمان با روش‌های سنتی بسیار زمان‌بر است، بنابراین ترکیب روش‌های آماری با مدل‌های هوشمند به ظهور یک علم بین‌رشته‌ای به نام داده‌کاوی و کشف دانش منجر شد (Joudaki et al., 2015). بسیاری از پژوهش‌های انجام‌شده با رویکرد داده‌کاوی در صنعت بیمه به‌منظور شناسایی تخلفات حوزه درمان و رسیدگی به اظهارات مشتریان است که یکی از شایع‌ترین تخلفات بیمه‌ای است (Kumar et al., 2010; Koh and Tan, 2011; Kirlidog and Asuk, 2012; Rawte and Anuradha, 2015; Pawar, 2016; Ghuse et al., 2017; Pandey et al., 2016).

از پرکاربردترین روش‌های داده‌کاوی به‌منظور شناسایی کشف تخلفات بیمه‌ای در حوزه درمان می‌توان روش‌های یادگیری ماشین (Waghade and Karandikar, 2018; Obodoekwe and Van der Haar, 2019) روش‌های با نظارت (Sheshasaayee and Thomas, 2017; Herland et al., 2019) و بدون نظارت را نام برد (Bauder et al., 2018; Massi et al., 2020). از طرفی داده‌های بزرگ و روش‌های رویارویی با آن‌ها به‌سرعت در صناعی چون بانکداری، بیمه و تجارت الکترونیک رو به رشد است که بررسی‌ها نشان می‌دهد با رشد روزافزون فناوری و کسب‌وکارهای مالی، داده‌های بزرگ و روش‌ها و شیوه‌های کار با آن‌ها یکی از موضوعات رو به رشد در پژوهش‌های روز دنیا به‌خصوص صنعت بیمه است (Cui et al., 2016). داده‌کاوی همچنین می‌تواند در کسب‌وکارهای مالی مانند بیمه با استفاده از الگوریتم‌های بلاکچین به افزایش امنیت اطلاعات پایگاه‌های داده کمک کند (Raikwar et al., 2018). این رویکرد براساس فلسفه کسب‌وکار مورد نظر، مدل کسب‌وکار، مدیریت داخلی و فروش و امثال آن به تغییرات ساختار صنایع مالی با ارائه چارچوب‌های جدید و نو در معماری مدیریت داده‌های بزرگ منجر شده است (Liu et al., 2018). با رشد روزافزون صناعی که بر پایه داده‌های ثبت‌شده بنا شده است، و توزیع نامتقارن این داده‌ها در انبارهای داده، روش‌های مختلفی برای بررسی داده‌های بزرگ و مدل‌های توسعه‌یافته این روش‌ها مانند رگرسیون لجستیک و مدل‌های درخت تصمیم به‌شدت مورد توجه پژوهشگران در کسب‌وکارهای مختلف از جمله صنعت بیمه قرار گرفته است (Lin et al., 2017). در رابطه با مطالعاتی که از روش‌های رویارویی با داده‌های بزرگ برای کاهش هزینه‌های شرکت‌های بیمه، افزایش کارایی و کشف تخلفات درمان استفاده شده است، می‌توان به Cui et al. (2016); Duman and Sağiroğlu (2017); Herland et al. (2018); Kenyon (2018) اشاره کرد. همچنین در مطالعات جدید، الگوریتم‌ها و روش‌های بهینه‌سازی، الگوریتم‌های ترکیبی، روش‌های جست‌وجوی همسایگی و شبکه‌های عصبی نیز با هدف بهبود روش‌های داده‌کاوی در کشف تخلفات بیمه استفاده شده است (Luan et al., 2019; Li et al., 2018; Yao et al., 2018).

جدول ۱: جست‌وجوی کلید واژه‌ها در مستندات علمی مرتبط با صنعت بیمه  
Table 1: Keyword search in scientific documents related to the insurance industry

تعداد مستندات بررسی شده	کلیدواژه‌های جست‌وجوشده	سناریو
500 مقاله اخیر	Insurance	اول
252	Insurance and data mining	دوم
830 مقاله اخیر	Insurance and risk	سوم
63	Insurance, data mining and fraud detection	چهارم



شکل ۱: شبکه هم‌استنادی کشورهای مختلف در مطالعات صنعت بیمه  
Fig. 1: Reference network of different countries in insurance industry studies

تا شبکه ابری کلیدواژه‌های پر کاربرد، شبکه هم‌استنادی کشورهای مختلف و مجلات معتبر علمی برای پژوهش‌های انجام شده در صنعت بیمه در سال‌های اخیر ترسیم شود. شکل ۱ شبکه هم‌استنادی کشورهای مختلف در پژوهش‌های صنعت بیمه را نشان می‌دهد. هم‌استنادی را می‌توان شاخصی برای ارتباط بین مستندات علمی دانست، به طوری که هم‌استنادی مقالات نشان از ارتباط محتوا و موضوع میان پژوهش‌های علمی انجام شده در آن کشورها دارد. ازین رو می‌توان با توجه به شرایط اقتصادی، سیاسی و وضعیت صنعت بیمه در کشورهای مختلف و شباهت میان آن‌ها، از مستندات علمی سایر کشورها بهره برد. شکل ۱ نشان می‌دهد بیشترین هم‌استنادی پژوهش‌های انجام شده در صنعت بیمه در کشورهای آمریکا\_کانادا\_ کره جنوبی، چین\_ استرالیا\_ تایوان\_ بلژیک\_ آلمان، هند\_ ایتالیا، هلند\_ اسپانیا\_ انگلستان، سوئیس و مجارستان رخ داده است. رنگ‌های مشابه و اتصالات میان کشورهای نام‌برده، نشان از هم‌استنادی پژوهش‌های انجام شده در این کشورها دارد.

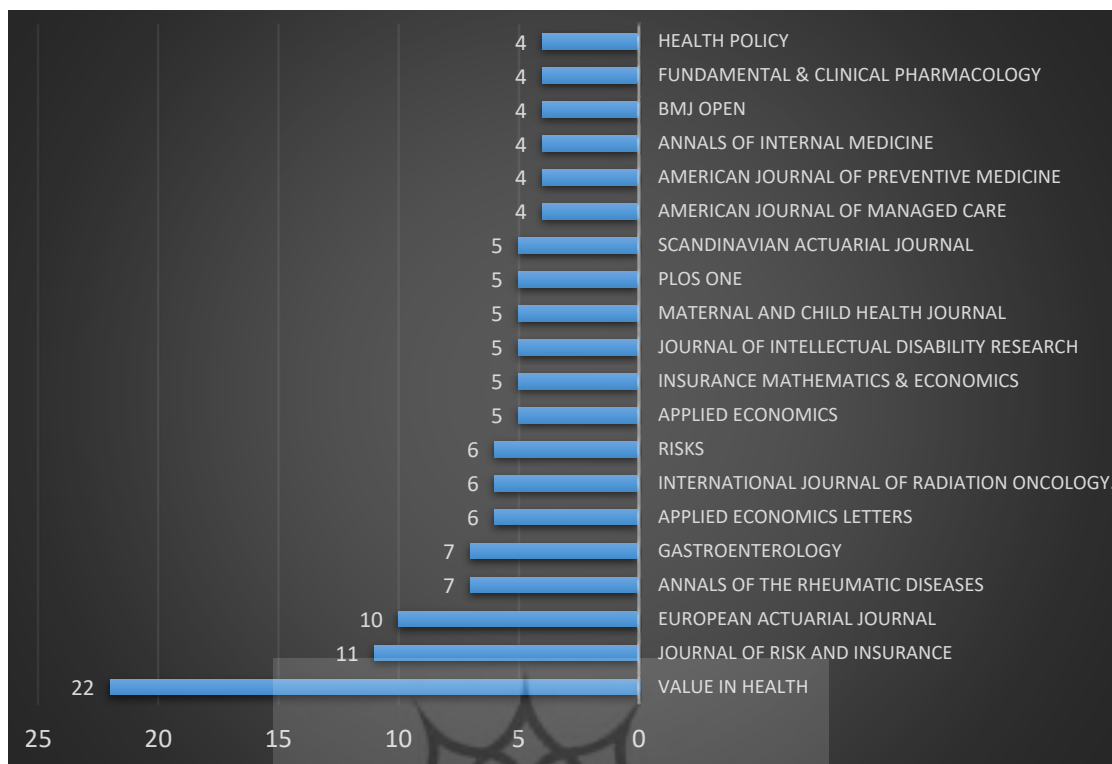
با تحلیل داده‌های سناریوی اول در نرم‌افزار vosviewer و با در نظر گرفتن حداقل ۴ مقاله به چاپ رسیده در هر مجله، مجلات معتبر

روش‌های داده کاوی در این صنعت ایجاد کند. به منظور پوشش جامع تمامی مستندات مرتبط با صنعت بیمه، با استفاده از جست‌وجوی پیشرفته، واژگان‌های منتخب در چهار سناریو در پایگاه داده Web Of Science که با نام پایگاه WOS نیز شناخته می‌شود، جست‌وجو شده که جدول ۱ واژگان‌های منتخب مورد نظر و تعداد مقالات دریافتی را نشان می‌دهد. در ادامه مقاله، فایل داده‌های مستندات هر سناریو در دو نرم‌افزار تحلیلی vosviewer و R وارد شده و تحلیل هم‌رخدادی واژگان‌ها، هم‌استنادی بین کشورهای مختلف و معتبرترین مجلات در هر چهار سناریو به تفکیک بررسی شده است.

شایان ذکر است در هر سناریو با حذف واژگانی که از جنس مفاهیم مورد نظر نبوده‌اند، واژگان اصلی تحقیق مبنای تجزیه و تحلیل‌های بعدی قرار گرفته است.

#### جست‌وجوی کلیدواژه Insurance

در سناریوی اول کلیدواژه insurance در پایگاه داده WoS جست‌وجو شده و ۵۰۰ مقاله اخیر در صنعت بیمه بررسی شده است



شکل ۲: تعداد مستندات علمی مرتبط با صنعت بیمه در مجلات معتبر  
 Fig. 2: The number of scientific documents related to the insurance industry in reputable journals

جدول ۲: اطلاعات مربوط به تعداد مستندات، نوع پژوهش، کلیدواژه‌ها و نویسندگان  
 Table 2: Information related to the number of documents, type of research, keywords and authors

تعداد	شرح
252	مستندات
212	منابع (Journals, Books, etc.)
374	کلیدواژه‌ها
1996 - 2019	بازه زمانی مورد بررسی
4.702	متوسط استناد به هر مقاله
741	نویسندگان
884	تعداد دفعات تکرار نویسندگان در مستندات مورد نظر
35	تعداد نویسندگان مستقل که مستندات تک‌نویسنده‌ای دارند
706	تعداد نویسگانی که در مستندات چندنویسنده‌ای حضور داشته‌اند
37	تعداد مستندات تک‌نویسنده‌ای
3.28	شاخص همکاری نویسندگان

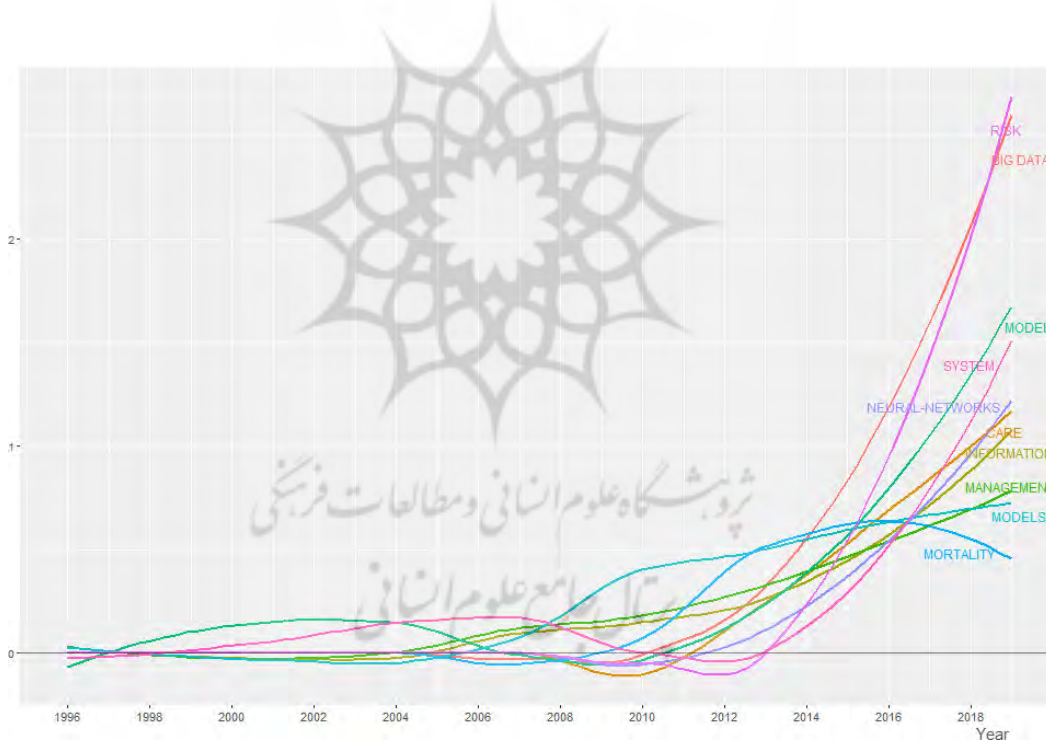
واژه Insurance و Data mining به‌طور هم‌زمان در واژگان کلیدی استفاده شده است. بنابراین با جست‌وجوی هم‌زمان این کلیدواژه در WOS، تعداد ۲۵۲ مستند علمی به‌دست آمده است که اطلاعات مربوط به تعداد مستندات، نوع پژوهش از قبیل کتاب یا مقاله، بازه زمانی مورد بررسی و همچنین نویسندگان و همکاری بین آن‌ها در پژوهش‌های مختلف در جدول ۲ آورده شده است. همان‌طور که در ابتدای این بخش نیز ذکر شد، پس از حذف واژگان‌های غیرمرتبط،

علمی متمرکز در پژوهش‌های مرتبط با صنعت بیمه در شکل ۲ نشان داده شده است. استفاده از این مجلات که حاوی مستندات معتبر علمی صنعت بیمه در سال‌های اخیرند، می‌تواند راه‌گشای پژوهشگران آینده برای یافتن برترین مطالعات بیمه باشد.

جست‌وجوی کلیدواژه‌های Insurance و Data mining در این بخش، هدف بررسی تعداد مقالاتی است که در آن‌ها از دو



شکل ۳: شبکه ابری کلیدواژه‌های پر کاربرد در مطالعات مرتبط با داده کاوی در صنعت بیمه  
 Fig. 3: Cloud network of frequently used keywords in studies related to data mining in the insurance industry



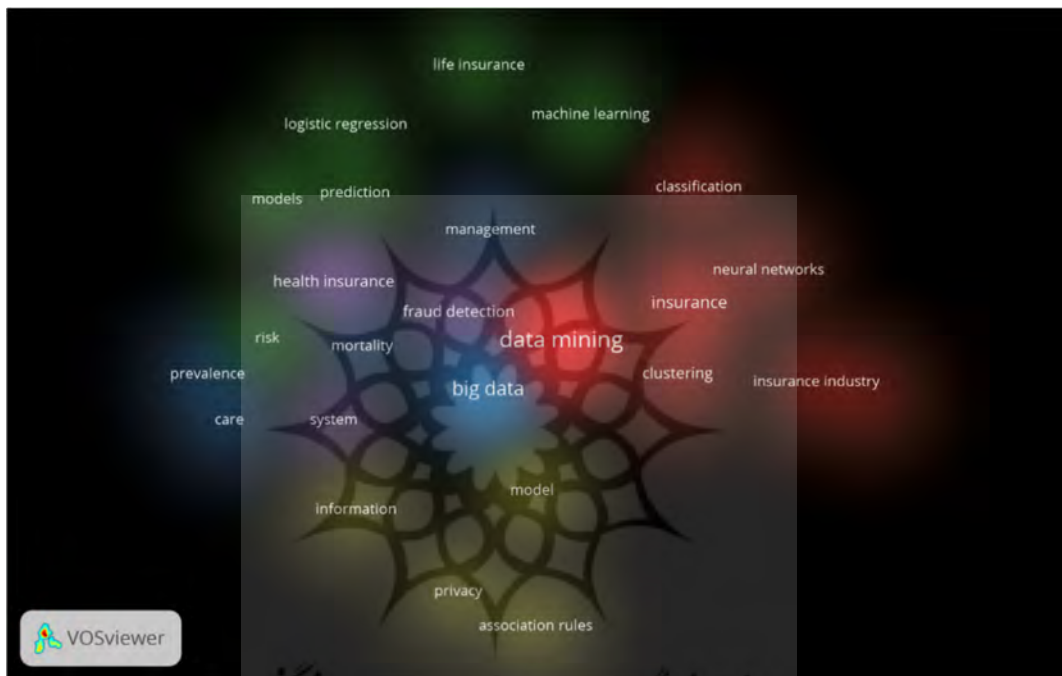
شکل ۴: روند رو به رشد کلیدواژه‌های مرتبط با داده کاوی در صنعت بیمه  
 Fig. 4: The growing trend of keywords related to data mining in the insurance industry

شکل ۳ نشان می‌دهد که صنعت بیمه به‌عنوان یکی از کسب‌وکارهای مبتنی بر داده، به‌سرعت در حال رشد داده‌های موجود در پایگاه‌های داده است که این حقیقت، شرکت‌های بیمه و مؤسسات پژوهشی را ملزم می‌سازد تا به روش‌های مدیریت پایگاه‌های داده و داده‌های انبوه روی آورند. در شکل ۴ نیز روند تغییرات کلیدواژه‌های پر کاربرد قابل مشاهده است. شکل ۴ نشان می‌دهد در میان واژگان پر کاربرد

با استفاده از نرم‌افزار vosviewer، کلیدواژه‌های اصلی و پرتکرار در مستندات مرتبط با کاربرد داده کاوی در صنعت بیمه برای تحلیل‌های بعدی مد نظر قرار گرفته است. همان‌طور که در شکل ۳ نیز مشاهده می‌شود، واژگان پر کاربرد شامل big data، health insurance، machine learning، fraud detection و clustering است. همان‌گونه که در مرور ادبیات نیز بیان شد، شبکه ابری کلیدواژه‌ها در

جدول ۳: خوشه‌بندی براساس شبکه هم‌رخدادی کلیدواژه‌ها در مطالعات داده‌کاوی صنعت بیمه  
Table 3: Clustering based on co-occurrence network of keywords in data mining studies of insurance industry

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Classification	Life Insurance	Big data	Association Rules	Fraud Detection
Clustering	Logistic Regression	Care	Information	Healthy Insurance
Data Mining	Machine learning	Management	Model	system
Data Warehouse	Models	Mortality	Privacy	
Insurance	Prediction	prevalence		
Insurance Industry	Risk			
Neural Network				



شکل ۵: شبکه هم‌رخدادی کلیدواژه‌ها در مطالعات داده‌کاوی در صنعت بیمه  
Fig. 5: Co-occurrence network of keywords in data mining studies in the insurance industry

اول روش‌های خوشه‌بندی، دسته‌بندی و شبکه عصبی از روش‌های پرکاربرد داده‌کاوی در صنعت بیمه است. در خوشه دوم نیز مشاهده می‌شود که روش‌های پیش‌بینی نظیر رگرسیون لجستیک و یادگیری ماشین از روش‌های پیش‌بینی روندهای آینده و ریسک در صنعت بیمه بوده است. به عبارت دیگر می‌توان گفت مدل‌های یادگیری ماشین و رگرسیون لجستیک از روش‌های کاربردی به‌منظور پیش‌بینی ریسک در صنعت بیمه بوده است. به همین ترتیب در خوشه پنجم نیز می‌توان کشف تخلفات، به خصوص در حوزه درمان را از موضوعات پرکاربرد در پژوهش‌های صنعت بیمه دانست. کلیدواژه‌های نام‌برده همچنین در شکل ۵ نیز نشان داده شده است که رنگ مربوط به هر کلیدواژه به‌عبارتی باهم‌آیی واژگان‌های خوشه‌بندی شده در جدول ۳ نشان می‌دهد. همچنین جایگاه کلیدواژه‌ها نشان از ارتباط بین

در طول زمان، داده‌های بزرگ و مدیریت ریسک در صنعت بیمه رشد چشمگیری در مطالعات سال‌های اخیر داشته است که این ضرورت مطالعه و استفاده از روش‌های مختلف برای مدیریت پایگاه‌های داده و رویارویی با داده‌های بزرگ و مدیریت و کاهش ریسک صنعت بیمه را نشان می‌دهد.

از آنجاکه کلیدواژه‌های هر پژوهش علمی بهترین راه برای دسترسی پژوهشگران به مستندات علمی گذشته است، به همین منظور واژگان به‌دست‌آمده از فایل مستندات علمی مبتنی بر کاربرد داده‌کاوی در صنعت بیمه از نظر تکرار یا با هم‌آیی واژگان و تعداد دفعات استفاده‌شده در مستندات علمی با نرم‌افزار vosviewer خوشه‌بندی شده‌اند که ۵ خوشه به‌دست‌آمده در جدول ۳ نشان داده شده است. همان‌طور که جدول ۳ نشان می‌دهد، در خوشه



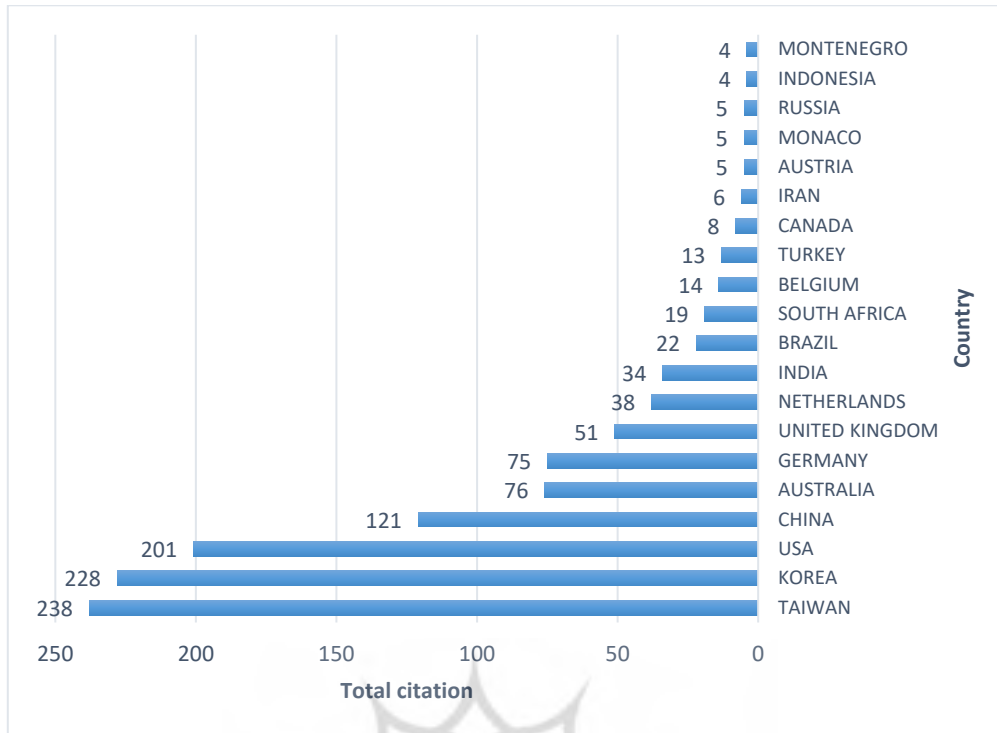
جدول ۴: مجموع و متوسط استناد به پژوهش‌های داده‌کاوی در صنعت بیمه در کشورهای مختلف  
Table 4: Total and average citations to data mining researches in the insurance industry in different countries

متوسط استناد به هر پژوهش	مجموع استناد به پژوهش‌ها	کشورها
11.90	238	Taiwan
9.50	228	South Korea
8.74	201	USA
2.12	121	China
5.85	76	Australia
7.50	75	Germany
7.29	51	United Kingdom
38.00	38	Netherlands
2.62	34	India
5.50	22	Brazil
4.75	19	South Africa
7.00	14	Belgium
4.33	13	Turkey
2.00	8	Canada
1.00	6	Iran
2.50	5	Austria
5.00	5	Monaco
5.00	5	Russia
0.80	4	Indonesia
4.00	4	Montenegro

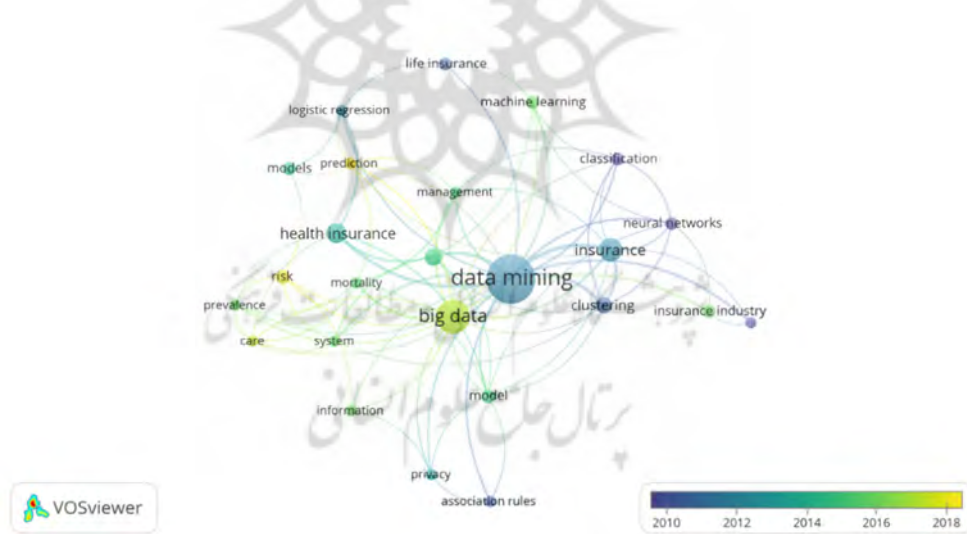
این زمینۀ پژوهشی در جایگاه ۱۵ام قرار دارد. شبکه هم‌رخدادی کلیدواژه‌های مورد بررسی طی سال‌های اخیر در شکل ۷ ترسیم شده که همان‌گونه که در شکل قابل مشاهده است، به‌دلیل روند رو به رشد داده‌ها و حجم انبوه داده‌های ثبت‌شده در شرکت‌های بیمه‌ای، مطالعات سال‌های ۲۰۱۶ تا ۲۰۱۸ به‌سمت تمرکز بر داده‌های بزرگ، ریسک و مدل‌های پیش‌بینی در صنعت بیمه بوده است. درحالی‌که روش‌های خوشه‌بندی، دسته‌بندی و شبکه‌های عصبی که در خوشه‌بندی واژگان کلیدی در شکل ۳ از روش‌های پرکاربرد داده‌کاوی در صنعت بیمه شناخته شد، در مطالعات قدیمی‌تر مشاهده شده که می‌توان گفت از روش‌های اشباع‌شده در مطالعات صنعت بیمه است.

همچنین شکل ۷ نشان می‌دهد الگوریتم‌های یادگیری ماشین (Machine learning) در مستندات دهۀ اخیر بیشتر به چشم می‌خورد که می‌توان آن را جایگزین روش‌های سنتی خوشه‌بندی و دسته‌بندی دانست. روش‌های جدید یادگیری ماشین را می‌توان برای هوشمندسازی پایگاه‌های داده با استفاده از روش‌های یادگیری عمیق، به‌خصوص با روش‌های تنظیم پارامترهای (Parameter tuning) این الگوریتم‌ها استفاده کرد.

آن‌ها در مستندات علمی مورد نظر دارد. کاربرد داده‌کاوی در صنعت بیمه و پژوهش‌های مرتبط با آن در کشورهای مختلف در سال‌های اخیر متفاوت است که در جدول ۴ نیز نشان داده شده است. شاخص استناد به مطالعات یک کشور یا مجلات معتبر علمی به‌عنوان شاخص ارزیابی کیفیت و محتوای پژوهش‌های مذکور در حوزه علمی مورد نظر است. ازین‌رو، کشور تایوان به‌عنوان کشوری پیشرو در پژوهش‌های داده‌کاوی در صنعت بیمه شناخته شده که پس از آن بیشترین تعداد استناد به مقالات را کرۀ جنوبی، آمریکا و چین در سال‌های اخیر در حوزه مورد بررسی داشته‌اند. همچنین داده‌های مورد نظر برای بررسی تعداد پژوهش‌ها و استناد به آن‌ها در نرم‌افزار R وارد شده و حداقل تعداد استناد به مقالات برابر ۴ در نظر گرفته شده است. شکل ۶ مجموع استناد به پژوهش‌های داده‌کاوی در بیمه در کشورهای مختلف را نشان می‌دهد که با توجه به تعداد پژوهش‌های چاپ‌شده در هر کشور و مجموع استناد به این پژوهش‌ها، متوسط استناد به هر مقاله به‌دست آمده است. شایان ذکر است با مجموع ۶ استناد به پژوهش‌های مرتبط با داده‌کاوی در صنعت بیمه که در ایران انجام شده و در وبسایت علمی WOS به چاپ رسیده است، ایران در میان کشورهای پیشرو در



شکل ۶: مجموع استناد به پژوهش‌های انجام‌شده در کشورهای مختلف مبنی بر کاربرد داده‌کاوی در صنعت بیمه  
 Fig. 6: The sum of references to the researchers conducted in different countries based on the application of data mining in the insurance industry



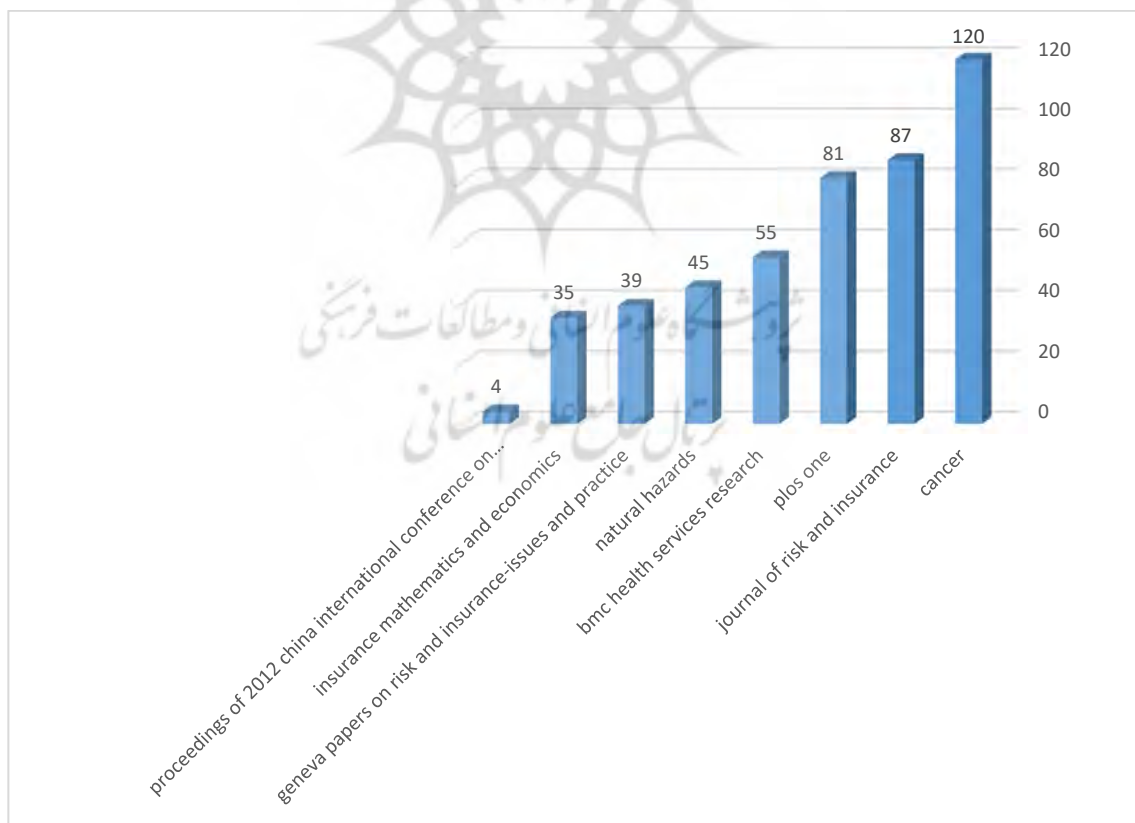
شکل ۷: شبکه هم‌رخدادی واژگان در طول زمان در مطالعات داده‌کاوی در صنعت بیمه  
 Fig. 7: Co-occurrence network of words over time in data mining studies in the insurance industry

همچنین شکل ۸ نیز شامل تعداد استنادهای سایر پژوهش‌های علمی به مقالات چاپ‌شده در این مجلات است. ازین‌رو مقالات چاپ‌شده در مجلات cancer، natural hazard و plos one دارای بیشترین استنادهای علمی در سال‌های اخیر هستند. بدیهی است که شاخص استناد به پژوهش‌های علمی به‌عنوان معیار ارزیابی مجلات و شناسایی معتبرترین مجلات علمی، می‌تواند برای مطالعات آینده

جست‌وجوی کلیدواژه‌های Risk و Insurance با توجه به افزایش قابل توجه پژوهش‌های مرتبط با ریسک در صنعت بیمه و ضرورت استفاده از آن بطور مداوم در مطالعات صنعت بیمه، در این بخش دو واژه ریسک و بیمه به‌عنوان واژگان کلیدی در WOS جست‌وجو شده و مجلات معتبر علمی با حداقل ۵ پژوهش چاپ‌شده در هر مجله، در جدول ۵ نشان داده شده است.

جدول ۵: تعداد پژوهش و مجموع استناد به مقالات چاپ شده در مجلات معتبر در حوزه بررسی ریسک در صنعت بیمه  
 Table 5: The number of researches and total citations of articles published in reliable journals in the field of risk assessment in the insurance industry

تعداد استناد به پژوهش‌ها	تعداد پژوهش‌ها	مجلات
120	6	<i>Cancer</i>
87	5	<i>Journal of Risk and Insurance</i>
81	19	<i>Plos one</i>
55	6	<i>Bmc Health Services Research</i>
45	5	<i>Natural Hazards</i>
39	13	<i>Jeneva Papers on Risk and Insurance-Issues and Practice</i>
35	5	<i>Insurance Mathematics and Economics</i>
4	6	Proceedings of 2012 China international conference on insurance and risk
1	5	Proceedings of 2013 China international conference on insurance and risk
0	9	Proceeding of 2017 International conference on insurance and risk
0	7	Proceedings of 2016 China international conference on insurance and risk
0	6	Proceedings of 2014 China international conference on insurance and risk
0	5	Proceedings of 2015 China international conference on insurance and risk



شکل ۸: مجموع استناد به پژوهش‌های چاپ شده در مجلات معتبر در حوزه بررسی ریسک در صنعت بیمه  
 Fig. 8: Total citations to researches published in reliable journals in the field of risk assessment in the insurance industry

جدول ۶: مجموع و متوسط استناد به پژوهش‌های مرتبط با ریسک و مدیریت آن در صنعت بیمه  
Table 6: Total and average citations to research related to risk and its management in the insurance industry

متوسط استناد به هر پژوهش	مجموع استناد به پژوهش‌ها	کشورها
3.36	329	USA
2.94	147	South Korea
3.64	40	France
0.86	38	China
1.95	37	Germany
1.47	25	Taiwan
2.40	12	Ghana
3.33	10	Vietnam
1.29	9	Canada
7.00	7	Austria
0.64	7	Japan
0.75	6	India
6.00	6	Lebanon
5.00	5	Kenya
0.56	5	United kingdom
4.00	4	Azerbaijan
0.80	4	Iran
1.33	4	Italy
1.33	4	Sweden
0.60	3	Australia

شد، یکی از پرکاربردترین کلیدواژه‌ها در رابطه با کاربرد داده‌کاوی در صنعت بیمه، کشف تخلفات در بیمه، به‌خصوص در حوزه سلامت و درمان است. ازین‌رو در **جدول ۷** مجموع استناد به پژوهش‌های علمی مربوط به کشف تخلفات در کشورهای مختلف و متوسط استناد به هر یک نشان داده شده است. کشور آمریکا همچنان در مقالات معتبر حوزه کشف تخلفات بیمه نیز دارای معتبرترین مستندات علمی و پس از آن کشورهای اسپانیا، بلژیک، انگلستان، اسلوانی، هند و سپس آلمان بیشترین مستندات معتبر با شاخص ارزیابی علمی بالا را داشته‌اند. با هدف شناسایی مجلاتی که بیشترین مقالات پژوهشی مرتبط با کشف تخلفات را به چاپ رسانده‌اند، **جدول ۸** منتخبی از مجلات معتبر علمی را با حداقل تعداد ۱ مقاله چاپ‌شده نشان می‌دهد که این مجلات به‌ترتیب بیشترین استناد علمی صورت‌گرفته به مستندات مورد نظر را نشان می‌دهد. شناسایی این مجلات می‌تواند راه مطالعات آینده پژوهشگران را برای تمرکز بر کشف تخلفات در صنعت بیمه روشن سازد.

#### نتایج و بحث

با نگاه به رشد روزافزون فناوری اطلاعات و پایگاه‌های داده

مورد توجه پژوهشگران قرار گیرد. با نظر به روند رو به رشد پژوهش‌های ریسک در صنعت بیمه در دنیا، ضروری است تا با نگاهی اجمالی به مستندات کشورهای مختلف و تعداد استناد به آن‌ها به‌عنوان شاخص ارزیابی کیفیت مستندات علمی، رشد روزافزون این مطالعات را در دنیا بررسی کنیم و با استناد به پژوهش‌های معتبر ریسک در بیمه در کشورهای توسعه‌یافته، جدیدترین روش‌های علمی را شناسایی کنیم و به کار بندیم. ازین‌رو در **جدول ۶** مجموع استناد به پژوهش‌های علمی چاپ‌شده در زمینه ریسک و مدیریت آن در صنعت بیمه در کشورهای مختلف و متوسط استناد به هر یک نشان داده شده است. همان‌گونه که در **جدول ۶** نیز دیده می‌شود، کشورهای آمریکا، کره جنوبی، فرانسه، چین، آلمان و تایوان دارای معتبرترین مستندات علمی مرتبط با ریسک در صنعت بیمه هستند. با توجه به **جدول ۶** از مستندات علمی چاپ‌شده در WOS، ایران در رده ۱۱۷ام جهان در مطالعات ریسک در صنعت بیمه قرار دارد.

#### جست‌وجوی کلیدواژه‌های *Fraud, data mining, insurance detection*

همان‌گونه که در شبکه ابری کلیدواژه‌های **شکل ۱** نشان داده

جدول ۷: مجموع و متوسط استناد به پژوهش‌های مرتبط با کشف تخلفات در صنعت بیمه  
Table 7: Total and average citation of researches related to the detection of violations in the insurance industry

کشورها	مجموع استناد به پژوهش‌ها	متوسط استناد به هر پژوهش
USA	136	10.46
Spain	93	23.25
Belgium	74	74.00
United Kingdom	74	24.67
Slovenia	43	21.50
India	30	3.00
Germany	21	10.50
South Korea	18	18.00
South Africa	16	16.00
China	14	1.75
Canada	13	6.50
France	13	6.50
Turkey	13	13.00
Sudan	7	7.00
Sweden	6	6.00
Taiwan	5	2.50
Australia	3	1.50
Italy	1	1.00
Indonesia	0	0.00
Malaysia	0	0.00

جدول ۸: منتخبی از مجلات علمی معتبر در چاپ مقالات مرتبط با کشف تخلفات در صنعت بیمه  
Table 8: A selection of authoritative scientific journals in publishing articles related to the discovery of violations in the insurance industry

مجلات	تعداد پژوهش‌ها	تعداد استناد به پژوهش‌ها
<i>Journal of Risk and Insurance</i>	7	255
Expert Systems with Applications	2	53
Legal and Criminological Psychology	1	45
European Journal of Operational Research	1	30
Psychology Crime and Law	2	29
Engineering Applications of Artificial Intelligence	1	18
World Conference on Business, Economics and Management	1	16
Applied Soft Computing	1	13
Decision Support Systems	1	11

آینده در صنعت بیمه ایران و تغییر در نگرش سازمان‌ها و مدیران صنعت بیمه ایران در راستای تهیه و تدوین برنامه‌های راهبردی سازمان‌ها باشد. ازین‌رو در این مقاله، مستندات علمی در رابطه با صنعت بیمه در جهان در سال‌های گذشته در چهار سناریوی مختلف تحلیل شده است. در سناریوی نخست، تحلیل شاخص هم‌استنادی پژوهش‌های متمرکز بر صنعت بیمه میان کشورهای مختلف که نشان از شباهت محتوا و موضوع مستندات مربوطه دارد، می‌تواند با تحلیل وضعیت سیاسی و اقتصادی صنعت بیمه و

در صنایع مختلف، به‌خصوص کسب‌وکارهای مالی مبتنی بر داده، مانند صنعت بیمه، کاربرد علم داده‌کاوی و روش‌های آن به‌عنوان ابزاری تحلیلی برای کشف دانش موجود در پایگاه‌های داده بزرگ بیش‌ازپیش مورد توجه پژوهشگران قرار گرفته است. در این مقاله با هدف بررسی پژوهش‌های انجام‌شده با روش‌های داده‌کاوی در صنعت بیمه در جهان طی سال‌های گذشته، نقشه هم‌واژگانی و هم‌استنادی برترین مطالعات کشورهای مختلف در جهان ترسیم شده تا گامی در راستای هدفمندسازی پژوهش‌های

تجارت الکترونیک است، پیشنهاد می‌شود در پژوهش‌های آینده صنعت بیمه، ترکیب روش‌های داده‌کاوی و بلاکچین برای افزایش امنیت تراکنش‌های مالی و سهولت در پیگیری پرداخت‌ها، وصول به‌موقع حق بیمه‌ها، کاهش تخلفات در پوشش‌های بیمه‌ای و فروش بیمه‌نامه‌ها مد نظر قرار گیرد. همچنین ترکیبی از روش‌های داده‌کاوی، به‌خصوص روش‌های یادگیری ماشین با شیوه‌های بهینه‌سازی مانند ساختارهای جست‌وجوی همسایگی، روش‌های انطباقی (Adaptive) و خودانطباقی (Self-adaptive) و الگوریتم جست‌وجوی ممنوعه (Tabu search) می‌تواند موجب بهبود مدل‌های موجود و یا ارائه مدل‌های نو در قابلیت ردیابی داده‌ها (Data traceability)، کشف الگوهای (Discover patterns) موجود و پیگیری تخلفات مشتریان در پایگاه داده‌های بزرگ شود. همچنین پیشنهاد می‌شود شناسایی ریسک و فاکتورهای آن که از موضوعات تحقیقاتی رو به رشد دنیا در صنعت بیمه است، در شرایط عدم قطعیت سنجدیده شود. بدین‌منظور روش‌های مدل‌سازی در شرایط عدم قطعیت و ترکیب آن با الگوریتم‌های یادگیری عمیق می‌تواند از کاربردی‌ترین مطالعات آینده صنعت بیمه باشد.

#### مشارکت نویسندگان

حوریا حاجیان: جمع‌آوری مطالعات مرتبط و تدوین مدل، تجزیه و تحلیل و تفسیر داده‌های مقاله، چکیده، مقدمه، نتیجه‌گیری، مروری بر ادبیات پژوهش، روش پژوهش و متدولوژی. اعظم زرچینی: جست‌وجوی کلیدواژه‌ها، کنترل چهارچوب تدوین و استانداردهای پژوهشی، طراحی مطالعه، بازبینی و اصلاح.

#### تشکر و قدردانی

پژوهشگران حاضر، از پیشنهادهای داوران محترم که به غنای علمی مقاله کمک کردند، مراتب سپاس خود را به جای می‌آورند.

#### تعارض منافع

نویسندگان اعلام می‌دارند که در خصوص انتشار این مقاله تضاد منافع وجود ندارد. علاوه‌براین، موضوعات اخلاقی شامل سرقت ادبی، رضایت آگاهانه، سوءرفتار، جعل داده‌ها، انتشار و ارسال مجدد و مکرر توسط نویسندگان رعایت شده است.

#### دسترسی آزاد

کپی‌رایت نویسنده (ها) ©2024: این مقاله تحت مجوز بین‌المللی Creative Commons Attribution 4.0، اجازه استفاده، اشتراک‌گذاری، اقتباس، توزیع و تکثیر را در هر رسانه یا قالبی مشروط بر درج نحوه دقیق دسترسی به مجوز CC، منوط به ذکر تغییرات احتمالی بر روی مقاله می‌داند. لذا به استناد مجوز مذکور، در هرگونه تغییرات در تصاویر، منابع و ارجاعات یا سایر مطالب از اشخاص ثالث در این مقاله باید در این مجوز گنجانده شود، مگر اینکه در راستای اعتبار مقاله به اشکال دیگری مشخص شده باشد. در صورت عدم درج مطالب

شبهات وضعیت موجود در ایران با کشورهای مشابه، راه مطالعات آینده را هدفمند سازد. به همین ترتیب مجلات معتبر علمی با در نظر گرفتن حداقل ۴ مستند علمی چاپ‌شده در رابطه با صنعت بیمه نیز شناسایی و برای تسهیل در مطالعات آینده معرفی شده است. در سناریوی دوم، شبکه ابری کلیدواژه‌ها و روند رو به رشد آن‌ها در مطالعات سال‌های اخیر نشان می‌دهد رشد پایگاه‌های داده در صنعت بیمه به استفاده از روش‌های مدیریت پایگاه‌های داده و داده‌های بزرگ منجر شده است. با نظر به روند رشد کلیدواژه‌ها در سال‌های اخیر، مشاهده می‌شود شناسایی فاکتورهای ریسک و مدیریت آن در صنعت بیمه به‌سرعت رو به رشد است که با توجه به ماهیت ناپایدار فاکتورهای ریسک و ماهیت رقابتی بازار در صنعت بیمه، تمرکز بر مطالعات ریسک به‌صورت فزاینده‌ای در حال افزایش است. همچنین با تحلیل هم‌واژگانی مستندات داده‌کاوی در صنعت بیمه، مشخص شد که روش‌های خوشه‌بندی، دسته‌بندی و شبکه عصبی از روش‌های رایج داده‌کاوی در بیمه و رگرسیون لجستیک و یادگیری ماشین از روش‌های پیش‌بینی ریسک در این صنعت بوده است. تحلیل مستندات کشورهای مختلف نشان می‌دهد کشور تایوان به‌عنوان کشوری پیشرو در پژوهش‌های داده‌کاوی در صنعت بیمه شناخته‌شده که پس از آن کره جنوبی، آمریکا و چین بیشترین تعداد استناد به مقالات را در سال‌های اخیر در این حوزه داشته‌اند. همچنین کشور ایران ۱۵امین کشور در میان کشورهای پیشرو در مطالعات جدید و به‌روز دنیا مبنی بر کاربردهای داده‌کاوی در صنعت بیمه است. در ادامه با توجه به روند رو به رشد پژوهش‌های ریسک در صنعت بیمه در دنیا، در سناریوی سوم به بررسی ریسک در صنعت بیمه پرداخته‌ایم که با تحلیل تعداد مستندات و مجموع استناد به پژوهش‌های چاپ‌شده، برترین مجلات علمی متمرکز بر ریسک در بیمه معرفی شده است. کشورهای آمریکا، کره جنوبی، فرانسه، چین، آلمان و تایوان دارای معتبرترین مستندات علمی حوزه ریسک در صنعت بیمه هستند. در انتها با توجه به تعداد پژوهش‌های مربوط به کشف تخلفات در صنعت بیمه که در مرور ادبیات موضوع نیز به‌تفصیل شرح داده شد، در سناریوی چهارم مستندات علمی متمرکز بر کشف تخلفات در بیمه، به‌خصوص در حوزه سلامت و درمان بررسی شده است. ازین‌رو مجموع استناد به این پژوهش‌های علمی در کشورهای مختلف و متوسط استناد به هریک نشان می‌دهد کشور آمریکا همچنان در مقالات معتبر حوزه کشف تخلفات بیمه نیز دارای معتبرترین مستندات علمی و پس از آن کشورهای اسپانیا، بلژیک، انگلستان، اسلوانی، هند و آلمان بیشترین مستندات معتبر با شاخص ارزیابی علمی بالا را داشته‌اند.

#### جمع‌بندی و پیشنهادها

با توجه به اینکه صنعت بیمه در دنیا به‌عنوان یکی از کسب‌وکارهای مستعد برای بهره‌گیری بیشتر از فناوری اطلاعات و ورود به دنیای

یادداشت ناشر


ناشر نشریه پژوهشنامه بیمه با توجه به مرزهای حقوقی در نقشه‌های منتشر شده بی طرف باقی می‌ماند.

منابع

- Abbas, N.N.; Ahmed, T.; Shah, S.H.U.; Omar, M.; Park, H.W., (2019). Investigating the applications of artificial intelligence in cyber security. *Scientometrics.*,
- Bauder, R.; Da Rosa, R.; Khoshgoftaar, T., (2018). identifying medicare provider fraud with unsupervised machine learning. 2018 IEEE International conference on information reuse and integration (IRI), 158(18): 5-22 **(18 Pages)**.
- Boodhun, N.; Jayabalan, M., (2018). Risk prediction in life insurance industry using supervised learning algorithms. *Complex. Intell. Syst.*, 3(2).
- Boroufar, A.; Rezaeian, A.; Shokohyar, S., (2017). Identifying the customer behavior model in life insurance sector using data mining. *Sci. Res. Q. Manage. Res. Iran.*, 121: 1189-1211 **(23 Pages)**. [In Persian]
- Chen, J.; Su, Y.; Si, H.; Chen, J., (2018). Managerial areas of construction and demolition waste: A scientometric review. *Int. J. Environ. Res. Public Health.*, 285-292 **(8 Pages)**.
- Chen, J.; Su, Y.; Si, H.; Chen, J., (2018). Managerial areas of construction and demolition waste: A scientometric review. *Int. J. Environ. Res. Public Health.*, 4: 145-154 **(10 Pages)**.
- Cui, H.; Li, Q.; Li, H.; Yan, Z., (2016). Healthcare fraud detection based on trustworthiness of doctors. 2016 IEEE Trustcom/BigDataSE/ISPA., 20(4): 65-94 **(30 Pages)**.
- Duman, E.A.; Sağiroğlu, Ş., (2017). Health care fraud detection methods and new approaches. *International conference on computer science and engineering (UBMK).*, 15(11).
- Ghorbani, A.; Farzai, S., (2018). Fraud detection in automobile insurance using a data mining based approach. *Int. J. Mechatron. Electr. Comput. Technol.*, 15.
- Ghuse, N.; Pawar, P.; Potgantwar, A., (2017). An improved approach for fraud detection in health insurance using data mining techniques. *Int. J. Sci. Res. Netw. Secur. Commun.*
- Herland, M.; Bauder, R.A.; Khoshgoftaar, T.M., (2019). The effects of class rarity on the evaluation of supervised healthcare fraud detection models. *J. Big. Data.*
- Herland, M.; Khoshgoftaar, T.M.; Bauder, R.A., (2018). Big data fraud detection using multiple medicare data sources. *J. Big. Data.*, 8(27): 3764-3771 **(8 Pages)**.
- Hessami, F., (2018). Business risk evaluation and management of Iranian commercial insurance companies. *Manage. Sci. Lett.*, 5: 27-32 **(6 Pages)**.
- Jahangiri, M.; Ahmadi Varzaneh, M.; Naderi Dehkordi, M., (2014). Application of data mining in insurance industry and customer classification. *National conference of computer engineering and information technology management.*, 6: 1-33 **(33 Pages)**. [In Persian]
- Joudaki, H.; Rashidian, A.; Minaei-Bidgoli, B.; Mahmoodi, M.; Geraili, B.; Nasiri, M.; Arab, M., (2015). Using data mining to detect health care fraud and abuse: A review of literature. *Global. J. Health. Sci.*, 5(1): 1-21 **(21 Pages)**.
- Kara, M.E.; Firat, S.Ü.O.; Ghadge, A., (2020). A data mining-based framework for supply chain risk management. *Comput. Ind. Eng.*, 8(2): 91-102 **(12 Pages)**.
- Kaščelan, V.; Kaščelan, L.; Novović Burić, M., (2016). A nonparametric data mining approach for risk prediction in car insurance: A case study from the Montenegrin market. *Econ. Res. Ekonomska. Istraživanja.*, 7(1): 194-202 **(9 Pages)**.
- Kenyon, D.L., (2018). An intelligent method of predicting insurance claims fraud. *Univ. Pretoria.*, 139.
- Kirlidog, M.; Asuk, C., (2012). fraud detection approach with data mining in health insurance. *Procedia. Social. Behav. Sci.*, 29(1): 545-558 **(14 Pages)**.
- Koh, H.C.; Tan, G., (2011). Data mining applications in healthcare. *J. Healthcare. Inf. Manage.*
- Kumar, M.; Ghani, R.; Mei, Z.S., (2010). Data mining to predict and prevent errors in health insurance claims processing. *Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining, ACM.*, 62: 989-994 **(6 Pages)**.
- Li, Y.; Yan, C.; Liu, W.; Li, M., (2018). A principle component analysis-based random forest with the potential nearest neighbor method for automobile insurance fraud identification. *Appl. Soft. Comput.*, 19(2): 64-72 **(9 Pages)**.
- Liao, S.H.; Chu, P.H.; Hsiao, P.Y., (2012). Data mining techniques and applications—A decade review from 2000 to 2011. *Expert. Syst. Appl.*, 65-74 **(10 Pages)**.
- Lin, W.; Wu, Z.; Lin, L.; Wen, A.; Li, J., (2017). An ensemble random forest algorithm for insurance big data analysis. *IEEE.*, 70: 1000-1009 **(10 Pages)**.
- Liu, Y.; Peng, J.; Yu, Z., (2018). Big data platform architecture under the background of financial technology: In the insurance industry as an example. *Proceedings of the 2018 international conference on big data engineering and technology, ACM.*, 39(12): 11303-11311 **(9 Pages)**.
- Luan, T.; Yan, Z.; Zhang, S., (2019). Fraudster detection based on modularity optimization algorithm. 2019 IEEE 23rd International conference on computer supported cooperative work in design (CSCWD), *IEEE.*, 5: 16568-16575 **(8 Pages)**.
- Massi, M.C.; Ieva, F.; Lettieri, E., (2020). Data mining application to healthcare fraud detection: Two-step unsupervised clustering method for outlier detection with administrative databases. *BMC. Med. Inf. Decis. Making.*, 31-35 **(5 Pages)**.
- Mingers, J.; Leydesdorff, L., (2015). A review of theory and practice in scientometrics. *Eur. J. Oper. Res.*, 422-427 **(6 Pages)**.
- Obodoekwe, N.; Van der Haar, D.T., (2019). A comparison of machine learning methods applicable to healthcare claims fraud detection. *International conference on information technology & systems, Springer.*, 20: 1-11 **(11 Pages)**.
- Pandey, P.; Saroliya, A.; Kumar, R., (2016). Analyses and detection of health insurance fraud using data mining and predictive modeling techniques. *Soft. Comput. Theor. Appl.*, 246(1): 1-19 **(19 Pages)**.
- Pawar, M.P., (2016). Review on data mining techniques for

- fraud detection in health insurance. *IJETT*, 548-557 (10 Pages).
- Phua, C.; Lee, V.; Smith, K.; Gayler, R., (2010). Comprehensive survey of data mining-based fraud detection research. *arXiv. Prepr. arXiv*, 2: 41-49 (9 Pages).
- Qarakhani, M.; Abolqasmi, M., (2019). News of the insurance world. *NoorMags.*, 3(2). [In Persian]
- Rahmani, M.; Haji Zain al-Abidini, M., (2014). Applications of data mining in information science and epistemology. *Q. J. Inf. Manage. Epistemology.*, 1009-6119 (5111 Pages). [In Persian]
- Raikwar, M.; Mazumdar, S.; Ruj, S.; Gupta, S.S.; Chattopadhyay, A.; Lam, K.Y., (2018). A blockchain framework for insurance processes. 2018 9th IFIP international conference on new technologies, mobility and security (NTMS), IEEE., 1-4 (4 Pages).
- Rawte, V.; Anuradha, G., (2015). Fraud detection in health insurance using data mining techniques. 2015 International conference on communication, information & computing technology (ICCICT), IEEE.
- Sheshasaayee, A.; Thomas, S.S., (2017). A purview of the impact of supervised learning methodologies on health insurance fraud detection. *Inf. Syst. Des. Intell. Appl.*, Springer: 978-984 (7 Pages).
- Waghade, S.S.; Karandikar, A.M., (2018). A comprehensive study of healthcare fraud detection based on machine learning. *Int. J. Appl. Eng. Res.*, 13(6): 4175-4178 (4 Pages).
- Wang, Y.; Xu, W., (2018). Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. *Decis. Support. Syst.*, 105: 87-95 (9 Pages).
- Wuyu, S.; Cerna, P., (2019). Risk assessment predictive modelling in Ethiopian insurance industry using data mining. *Software. Eng.*, 6(4).
- Yao, J.; Zhang, J.; Wang, L., (2018). A financial statement fraud detection model based on hybrid data mining methods. 2018 International conference on artificial intelligence and big data (ICAIBD), IEEE., 57-61 (5 Pages).

AUTHOR(S) BIOSKETCHES	معرفی نویسندگان
<p data-bbox="742 853 1311 882">حوریا حاجیان، گروه علوم کامپیوتر، دانشکده علوم، دانشگاه فنی انتاریو، اوشاوا، انتاریو، کانادا</p> <ul style="list-style-type: none"> <li data-bbox="296 904 671 932">▪ Email: <a href="mailto:hooria_hajiyani@ontariotechu.net">hooria_hajiyani@ontariotechu.net</a></li> <li data-bbox="296 932 576 960">▪ ORCID: 0000-0003-1098-4552</li> <li data-bbox="296 960 635 989">▪ Homepage: <a href="https://ontariotechu.ca/">https://ontariotechu.ca/</a></li> </ul> <p data-bbox="536 1011 1311 1039">اعظم زرچینی، دگروه مدیریت تکنولوژی، دانشکده مدیریت و اقتصاد، دانشگاه آزاد اسلامی واحد علوم و تحقیقات، تهران، ایران</p> <ul style="list-style-type: none"> <li data-bbox="296 1061 576 1089">▪ Email: <a href="mailto:a_zarjini@taavon-ins.ir">a_zarjini@taavon-ins.ir</a></li> <li data-bbox="296 1089 576 1118">▪ ORCID: 0000-0002-8638-7220</li> <li data-bbox="296 1118 788 1146">▪ Homepage: <a href="https://srb.iau.ir/faculty/m-malkhalifeh/fa">https://srb.iau.ir/faculty/m-malkhalifeh/fa</a></li> </ul>	

HOW TO CITE THIS ARTICLE	
<p data-bbox="296 1253 1107 1319"><i>Hajian, H.; Zarjini, A., (2024). A comprehensive analysis of keywords co-occurrence network and the most cited journals on data mining techniques in insurance industry using scientometrics approach. Iran. J. Insur. Res., 13(1): 71-86.</i></p> <p data-bbox="296 1340 552 1369">DOI: 10.22056/ijir.2024.01.06</p> <p data-bbox="296 1369 707 1397">URL: <a href="https://ijir.irc.ac.ir/article_139314.html?lang=en">https://ijir.irc.ac.ir/article_139314.html?lang=en</a></p>	