Iran. J. Insur. Res., 12(4): 251-264, Autumn 2023, Serial #46

Iranian Journal of Insurance Research
(IJIR)

Homepage: https://ijir.irc.ac.ir/?lang=en

Iranian Insurance Research Center (IIRC)

ORIGINAL RESEARCH PAPER

Designing an intelligent evaluation system for predicting fire insurance claims using deep learning

I. Raeesi Vanani[1,*], M. Taghavifard[1], B. Sohrabi [2], M. Amirhosseini[1]

[1] *Department of Operations and Information Technology Management, Faculty of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran*

[2] *Department of Information Technology Management, Faculty of Management, University of Tehran, Tehran, Iran*

ARTICLE INFO

Article History:
Received 03 May 2023
Revised 01 July 2023
Accepted 14 August 2023

Keywords:
Deep learning
Fire insurance
Loss prediction
Risk evaluation
Risk theory

*Corresponding Author:
Email: *Imanraeesi@atu.ac.ir*
Phone: +9821 44744370
ORCID: 0000-0001-8324-9896

ABSTRACT

BACKGROUND AND OBJECTIVES: Fire insurance is a type of insurance that provides financial protection against property damage caused by events such as fire, theft, and weather-related damage. The use of modern analytics methods in analyzing policies and predicting potential losses is of great interest to insurance companies and those involved in property insurance services. Claim prediction is a crucial measure for predicting future losses in insurance companies, particularly in property insurance, where it helps evaluate the Probable Maximum Loss (PML) based on risk theory.

METHODS: This study focuses on developing claim prediction models for occurrence probability, severity, and time by creating a dataset, training various algorithms, and comparing their performance. Exploratory Data Analysis (EDA) is conducted to select relevant features, resulting in the selection of 44 fields from insurance policy and claim information. Recursive Field Elimination (RFE) is used to reduce the dataset dimensions, and informative fields are selected for each prediction model. The dataset consists of over 780,000 policy records and approximately 70,000 loss payments from the Iran Insurance Company's actual fire application database spanning 10 years (2011-2021). Linear regression, random forest regression (RFR), support vector regression (SVR), and deep neural network algorithms are implemented for each prediction model. The accuracy of the algorithms is evaluated using mean squared error (MSE) and mean absolute error (MAE) values.

FINDINGS: The results of the prediction models show that the deep multi-layer perception (MLP) algorithm performs the best. After hyperparameter tuning and multiple runs, the final MSE values are determined as 0.117 for occurrence probability prediction, 0.042 for loss severity prediction, and 0.106 for claim time prediction. The comparison of the innovative model results with test data demonstrates that intelligent models provide more accurate predictions. Accurate prediction is highly valuable for insurance companies as it allows them to mitigate financial losses. The use of deep learning in loss time prediction, in addition to severity and probability prediction, is a significant innovation in this research.

CONCLUSION: Machine learning techniques, such as deep learning, can help insurance companies optimize their services with higher accuracy, strengthen risk management practices, and provide tools for better decision-making. The application of deep learning in loss prediction can replace complex and inaccurate manual processes in insurance, leading to advancements in risk management, reinsurance management, and improvement of fire insurance pricing.

DOI: 10.22056/ijir.2023.04.01

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

نشریه علمی پژوهشنامه بیمه دوره ۱۲، شماره ۴، پاییز ۱۴۰۲، شماره پیاپی ۴۶، ص ۲۵۱-۲۶۴

نشریه علمی پژوهشنامه بیمه

سایت نشریه: https://ijir.irc.ac.ir/?lang=fa

مقاله علمی

طراحی سیستم ارزیابی هوشمند جهت پیش‌بینی خسارت بیمه‌های آتش‌سوزی با استفاده از یادگیری عمیق

ایمان رئیسی وانانی[۱،*]، محمدتقی تقوی فرد[۱]، بابک سهرابی[۲]، مرتضی امیرحسینی[۱]

[۱] گروه مدیریت عملیات و فناوری اطلاعات، دانشکدهٔ مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران

[۲] گروه مدیریت فناوری اطلاعات، دانشکدهٔ مدیریت، دانشگاه تهران، تهران، ایران

اطلاعات مقاله

تاریخ های مقاله:

تاریخ دریافت: ۱۳ اردیبهشت ۱۴۰۲

تاریخ داوری: ۱۰ تیر ۱۴۰۲

تاریخ پذیرش: ۲۳ مرداد ۱۴۰۲

کلمات کلیدی:

ارزیابی ریسک

بیمهٔ آتش‌سوزی

پیش‌بینی خسارت

نظریهٔ ریسک

یادگیری عمیق

*نویسنده مسئول:

ایمیل: Imanraeesi@atu.ac.ir

تلفن: ۴۴۷۴۴۳۷۰ ۹۸۲۱+

ORCID: 0000-0001-8324-9896

چکیده:

پیشینه و اهداف: بیمهٔ آتش‌سوزی نوعی بیمه است که از خسارت مالی به اموال محافظت می‌کند. معمولاً حوادثی مانند آتش‌سوزی، سرقت و خسارت‌های مربوط به آب‌وهوا را پوشش می‌دهد و می‌تواند به جبران هزینه‌های تعمیر یا جایگزینی اموال آسیب‌دیده کمک کند. شرکت‌های بیمه و علاقه‌مندان به توسعهٔ خدمات بیمهٔ آتش‌سوزی به‌دنبال استفاده از روش‌های تحلیلی مدرن برای تجزیه‌وتحلیل بیمه‌نامه‌ها، ارزیابی و پیش‌بینی خسارت احتمالی آن‌ها برای مدیریت ریسک هستند. پیش‌بینی ادعای خسارت، معیاری حیاتی برای پیش‌بینی خسارت‌های آتی در شرکت‌های بیمه است. براساس نظریهٔ ریسک، پیش‌بینی خسارت عنصری مهم در کسب‌وکار بیمهٔ آتش‌سوزی برای ارزیابی حداکثر خسارت احتمالی است.

روش‌شناسی: در این پژوهش سه معیار پیش‌بینی خسارت (احتمال وقوع، شدت، زمان بروز) با تهیهٔ مجموعه داده، یادگیری و مقایسهٔ الگوریتم‌های مختلف توصیف می‌شوند. در ابتدا، تجزیه‌وتحلیل داده‌های اکتشافی برای انتخاب ویژگی‌های مورد نیاز انجام شد و در نهایت ۴۴ قلم اطلاعاتی از اطلاعات بیمه‌نامه و خسارت پرداختی رشتهٔ آتش‌سوزی انتخاب گردید. ابعاد مجموعه داده‌ها توسط روش حذف بازگشتی ویژگی‌ها کاهش یافته و برای هر الگوریتم، مجموعهٔ مختلفی از فیلدهای اطلاعاتی مؤثر انتخاب شده است. ما بیش از ۷۸۰،۰۰۰ رکورد بیمه‌نامه و حدود ۷۰،۰۰۰ رکورد مرتبط خسارت پرداختی را برای یک بازهٔ ده‌ساله (ابتدای ۱۳۹۰ تا ابتدای ۱۴۰۰) از بانک اطلاعاتی عملیاتی سامانهٔ آتش‌سوزی بیمهٔ ایران انتخاب کرده‌ایم. مدل‌های یادگیری رگرسیونی برتر مانند رگرسیون خطی، رگرسیون جنگل تصادفی، رگرسیون بردار پشتیبان و شبکهٔ عصبی عمیق برای هر سه الگوریتم پیش‌بینی خسارت پیاده‌سازی شد. سپس دقت الگوریتم‌ها با مقدار میانگین مربعات خطا و مقدار میانگین خطای مطلق مقایسه شد.

یافته‌ها: نتایج پیش‌بینی مدل نشان داد که بهترین الگوریتم برای هر سه معیار، یادگیری عمیق و مشخصاً شبکهٔ عصبی چندلایهٔ پرسپترون است. پس از تنظیم فراپارامترها و چندین بار اجرا، بهترین الگوریتم یادگیری عمیق با کمترین خطا با مقادیر ۰،۱۱۷ (احتمال وقوع)، ۰،۰۴۲ (شدت خسارت)، ۰،۱۰۶ (زمان بروز خسارت) حاصل شد. پیش‌بینی نتایج مدل نوآورانهٔ ما در داده‌های آزمایشی، به این نتیجه رسید که مدل هوشمند ارائه‌شده دقت مناسبی دارد. شرکت‌های بیمه به‌شدت علاقه‌مند پیش‌بینی آینده‌اند و پیش‌بینی خسارت فرصتی برای کاهش زیان مالی برای شرکت فراهم می‌کند. به‌کارگیری یادگیری عمیق در پیش‌بینی خسارت آتش‌سوزی و پیش‌بینی زمان بروز خسارت، علاوه‌بر احتمال و شدت، نوآوری‌های مدل هستند.

نتیجه‌گیری: یادگیری ماشین می‌توانند به شرکت‌ها کمک کنند تا خدمات خود را با دقت بیشتری بهینه کنند، مدیریت ریسک را تقویت و در نتیجه ابزارهایی برای تصمیم‌گیری بهتر فراهم نمایند. به‌کارگیری یادگیری عمیق در پیش‌بینی خسارت بیمه می‌تواند به‌صورت کاربردی جایگزین فرایند دستی پیچیده، زمان‌بر و نادقیق موجود در شرکت‌های بیمه شود و سرآغاز توسعهٔ نوین در مدیریت ریسک، مدیریت اتکایی و بهبود نرخ‌گذاری بیمهٔ آتش‌سوزی باشد.

DOI: 10.22056/ijir.2023.04.01

توجه: مدت‌زمان بحث و انتقاد برای این مقاله تا ۱ ژوئیه ۲۰۲۳ در وب‌سایت IJIR در «نمایش مقاله» باز است.

مقدمه

شرکت‌های بیمه، سرمایهٔ خود را در معرض ریسک خسارت قرار می‌دهند و براساس آن کسب درآمد می‌کنند. یکی از مشکلات شرکت‌های بیمه، زمان زیاد و ارزیابی نادقیق پیش‌بینی خسارت و ارزش در معرض ریسک (Value at Risk) است. آن‌ها باید بتوانند براساس تحلیل‌های داده‌محور و به‌کارگیری روش‌های نوین و براساس اطلاعات و داده‌های مربوط به بیمه‌نامه‌ها و سوابق خسارت، ارزیابی ریسک خسارت را انجام دهند (Deelstra and Plantin, 2014).

بیمهٔ آتش‌سوزی، در ایران در گروه بیمه‌های اموال و با نام بیمهٔ آتش‌سوزی شناخته می‌شود. این بیمه، شامل پوشش خطراتی از قبیل آتش‌سوزی، زلزله، سیل، انفجار، طوفان، ترکیدن لوله، سرقت و صاعقه برای ساختمان‌ها، تأسیسات، ماشین‌آلات و اثاثیه است.

یکی از اصول ضروری در مدیریت ریسک شرکت‌های بیمه‌ای، روش ارزیابی مناسب پیش‌بینی خسارت احتمالی و سپس اخذ تصمیمی صحیح براساس راهبردهای مدیریت ریسک شرکت است. برای این منظور، استفاده از راه‌حل‌های هوش مصنوعی، یادگیری ماشین و فنّاوری‌های داده‌محور، می‌تواند کمک شایان توجهی به این مهم داشته باشد (Olarewaju and Msomi, 2022). ما می‌توانیم با تحلیل سطح خطرات، تشخیص مناسبی از تخمین خسارت داشته باشیم و به‌منظور پیش‌بینی بهتر خسارت، می‌توانیم با استفاده از سوابق اطلاعات و اقلام اطلاعاتی مؤثر، تخمین مناسبی از خسارت آتی داشته باشیم (Chen and Yang, 2020).

همچنین باید این ارزیابی و پیش‌بینی خسارت، در زمان کوتاه و معقولی انجام شود. ارزیابی داده‌محور خسارت بیمه‌های آتش‌سوزی، در زمان کوتاه و با دقت مناسب، از اهداف اصلی این پژوهش است. در حال حاضر ارزیابی خسارت آتی در رشتهٔ آتش‌سوزی در شرکت‌های بیمه به‌صورت داده‌محور انجام نمی‌گیرد و معمولاً بسیار پیچیده، زمان‌بر و نادقیق است. همچنین در عمده تحقیقات قبلی، پیش‌بینی خسارت دربارهٔ دو معیار احتمال وقوع و شدت خسارت انجام شده است و در این پژوهش و به‌منظور بهبود روش ارزیابی ریسک خسارت، معیار پیش‌بینی زمان بروز خسارت بعد از صدور بیمه‌نامه نیز در کنار دو معیار دیگر پیشنهاد شده است. در این پژوهش قصد داریم با به‌کارگیری علوم داده، به‌صورت هوشمند پیش‌بینی خسارت بیمه‌نامه‌های آتش‌سوزی را انجام داده، مدل بهبودیافته‌ای برای معیارهای ارزیابی و پیش‌بینی خسارت ارائه دهیم. خروجی این پژوهش می‌تواند برای تحلیل ریسک بیمه‌ای، قیمت‌گذاری و اتکایی موردی استفاده شود.

عناوین ضرورت پژوهش به شرح زیر است:

- به‌کارگیری الگوریتم‌های یادگیری ماشین و به‌صورت داده‌محور (نیاز به پژوهش بیشتر در اینشورتک).
- پژوهشی کاربردی، برای ارزیابی و پیش‌بینی خسارت در بیمهٔ آتش‌سوزی (کاربردی در صنعت و مخصوصاً بیمهٔ آتش‌سوزی).
- حوزهٔ ارزیابی خسارت در اولویت پژوهشی، پژوهشکدهٔ بیمه است (اولویت پژوهش).
- پیش‌بینی زمان بروز خسارت (بعد از صدور بیمه‌نامه) در کنار پیش‌بینی احتمال و شدت خسارت (بهبود روش‌های پیشین).

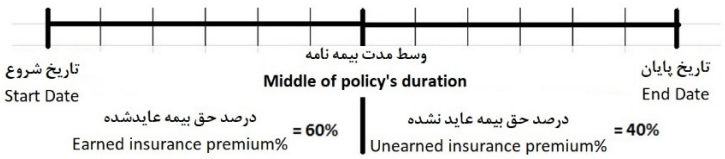
مبانی نظری پژوهش

ارزیابی ریسک بیمه

ریسک دلیل وجودی صنعت بیمه است. در واقع، شرکت‌های بیمه باید بتوانند با ابزارهای مناسب و تا حد ممکن دقیق، ریسک قراردادهای بیمه یا بیمه‌نامه‌ها را ارزیابی و مدیریت کنند. انواع ریسک در شرکت‌های بیمه، با در نظر گرفتن آیین‌نامهٔ شمارهٔ ۶۹ مصوب سال ۱۳۹۰ شورای عالی بیمه عبارت است از ریسک بیمه‌گری (Underwriting risk)، ریسک اعتبار (Credit risk)، ریسک نقدینگی (Liquidity risk)، ریسک بازار (Market risk) و ریسک عملیاتی (Operational risk) که در شکل ۱ این دسته‌بندی نمایش داده شده است. ریسک بیمه‌گری برای ارزیابی خسارت استفاده می‌شود.

ریسک بیمه‌گری اصلی‌ترین ریسک صنعت بیمه است. ریسک بیمه‌گری، پذیرش ریسک و تعهد در مقابل خسارات احتمالی است. خود ریسک‌های بیمه‌گری انواع مختلفی دارد که ریسک حق‌بیمه، ریسک ذخایر ناکافی و ریسک فاجعه‌آمیز را شامل می‌شود. ریسک حق‌بیمه در واقع ریسک بروز خسارات فراتر از خسارات انتظاری در سال جاری است و این ریسک، ریسک معمول صدور نیز نامیده می‌شود. درحالی‌که ریسک ذخایر ناکافی در اثر ناکافی بودن ذخایر

شکل ۱: انواع ریسک بیمه‌ای
Fig. 1: Types of insurance risk



نمودار ۱: حق‌بیمهٔ عایدشده و عایدنشده (وب‌سایت www.agordon.com)
Chart 1: Earned and unearned insurance premium

برای پوشش مطالبات خسارات است و در نهایت ریسک فاجعه‌آمیز ریسک ناشی از خسارات در مقیاس بزرگ است (Shiu, 2020).

تجمیع ارزیابی ریسک برای بیمه‌نامه‌ها و خسارت‌های قبلی بر روی اقلام اطلاعاتی مؤثر، روشی کلیدی برای شرکت‌های بیمه است. مثلاً در بیمهٔ اتومبیل اگر فقط پرداخت خسارت برای یک بیمه‌نامه را بررسی کنیم، قطعاً برای ارزیابی ریسک نامناسب و کاملاً نادقیق است. درحالی‌که اگر این ارزیابی ریسک خسارت بر روی ۱۰۰ هزار بیمه‌نامهٔ مشابه قبلی انجام شود، می‌تواند برای ارزیابی و پیش‌بینی با دقت قابل قبولی استفاده شود (Lentz *et al.*, 2015).

یکی از ریسک‌های مطرح‌شده در ریسک‌های بیمه‌گری، ریسک ذخایر ناکافی است، این ریسک با محاسبه و نگهداری مبلغی نزدیک به واقعیت به‌عنوان ذخیرهٔ خسارت به‌دست می‌آید. در صورت بروز خسارت، مبلغ ذخایر پشتیبان شرکت برای ایفای تعهدات پرداخت خسارت خواهد بود. بر همین اساس، شرکت‌های بیمه بخشی از درآمد را به‌عنوان رزرو نگه می‌دارند و مابقی را در حوزه‌های دیگر سرمایه‌گذاری می‌کنند (Breuer and Staudt, 2022). با توجه به اینکه مبلغی از درآمد بیمه‌نامه به‌عنوان ذخیره نگهداری می‌شود، در صورت داشتن پیش‌بینی مناسب از زمان وقوع خسارت، می‌تواند منبع مالی مناسبی برای سرمایه گذاری باشد.

زمان بروز ریسک یا همان وقوع خسارت در صنعت بیمه بسیار مهم است و بهتر است در ارزیابی‌های ریسک نقش مؤثری داشته باشد. این موضوع وقتی بیشتر مشخص می‌شود که با مفهوم حق‌بیمهٔ عایدشده و عایدنشده آشنا شویم. هر مدت که از زمان بیمه‌نامه سپری می‌شود به همان میزان (به‌اضافهٔ هزینه‌های بیمه‌گری) از حق‌بیمهٔ دریافت‌شده از مشتری عایدشده است. در نمودار ۱ با فرض رسیدن به ماه ششم از یک قرارداد فرضی بیمه‌ای یک‌ساله، مقدار عایدشده از حق‌بیمه به میزان ۶۰ درصد حق‌بیمهٔ دریافتی است (۵۰ درصد بابت سپری شدن نیمی از قرارداد و ده درصد بابت هزینه‌های غیرقابل عودت صدور بیمه‌نامه) و طبیعتاً ۴۰ درصد از حق‌بیمهٔ دریافتی هنوز عایدنشده است و تا پایان مدت بیمه‌نامه، ممکن است با خسارت مواجه شویم.

بنابراین هرچه دیرتر خسارت احتمالی رخ بدهد، مبلغ عایدی از حق‌بیمه بیشتر می‌شود و خود می‌تواند عامل بسیار مهمی در ارزیابی ریسک و تصمیم‌گیری باشد. برای ارزیابی ریسک و پیش‌بینی خسارت بیمه‌های آتش‌سوزی، باید علاوه‌بر دسته‌بندی اطلاعات قبلی فروش و خسارت، به روشی مناسب احتمال وقوع، شدت و زمان بروز خسارت پیش‌بینی و ارزیابی شود.

نظریهٔ ریسک

ریسک موجود در بیمه و خسارت‌های احتمالی پرداختی، با نظریهٔ ریسک قابل بررسی است. این نظریه منطبق بر دیدگاه نوسان احتمال وقوع اتفاق است و به‌عنوان انحراف از میانگین بازدهی، در نظریه‌ها تعریف می‌شود (Roeser *et al.*, 2012).

نظریهٔ ریسک یکی از ابزارهای مهم بیم‌سنجی در حوزهٔ بیمه‌های غیرزندگی (اموال، مهندسی، حوادث، مسئولیت، انرژی، حمل کالا و اتومبیل) است. از قیمت‌گذاری، طراحی و ارزیابی یک محصول بیمهٔ غیرزندگی تا مباحث اتکایی آن‌ها، بدون استفاده از نظریهٔ ریسک تقریباً ناممکن است. متأسفانه در ایران آن‌گونه که شایسته است این بخش از علم بیم‌سنجی، مورد توجه صنعت بیمه و محققان قرار نگرفته و حتی برخی فعالان صنعت بیمه، علم بیم‌سنجی را معادل بیمه‌های زندگی (عمر و درمان) دانسته‌اند. این در حالی است که برخلاف بیمه‌های زندگی، بیمه‌های غیرزندگی بسیار متنوع و پیچیده‌اند (Payandeh Najafabadi, 2020).

از منظر نظریهٔ ریسک و برای شرکت‌های بیمه، می‌خواهیم ریسک آتی را پیش‌بینی کنیم و برای این منظور دسته‌بندی اطلاعات قبلی بیمه‌نامه‌ها و خسارات وارده را براساس عوامل و متغیرهای مؤثرتر آن‌ها مد نظر قرار می‌دهیم. وقتی ارزش در معرض ریسک (جمع سرمایه‌های مورد پوشش در یک بیمه‌نامه) را اندازه‌گیری می‌کنیم، یعنی ما x درصد مطمئنیم که در طول زمان t از مبلغ v دلار، خسارت بیشتری پرداخت نخواهیم کرد و v جمع سرمایهٔ بیمه‌نامه‌هاست که به احتمال x درصد و در زمان t خسارت می‌بینند (Hull, 2015).

با نگاه از منظر نظریهٔ ریسک می‌خواهیم احتمال وقوع خسارت و زمان بروز خسارت را پیش‌بینی کنیم و همچنین با پیش‌بینی درصد میزان خسارت نسبت به سرمایهٔ تحت پوشش، می‌توانیم شدت خسارت را نیز پیش‌بینی کنیم.

مروری بر پیشینهٔ پژوهش

در سال‌های اخیر و با به‌کارگیری بیشتر روش‌های هوشمند و داده‌محور در صنعت و در ارزیابی‌ها و تصمیم‌گیری‌ها، در صنعت بیمه نیز تحولاتی در حال وقوع است و برخی تحقیقات در این بخش معرفی می‌شود.

پیشینهٔ تحقیقات داخلی

Farhang Adib *et al.* (2014) با استفاده از روش‌های داده‌کاوی

و الگوریتم‌های یادگیری درخت تصمیم، یک مدل برای پیش‌بینی خسارت برای بیمه‌نامه‌های شخص ثالث ارائه کرده‌اند. در این روش و براساس خسارت‌های پرداخت‌شدهٔ قبلی، ابتدا ارتباط اقلام اطلاعاتی بیمه‌نامه شناسایی و میزان وابستگی آن‌ها با پرداخت خسارت مشخص شد و سپس عوامل مؤثرتر انتخاب و براساس الگوریتم‌های مختلف یادگیری ماشین ارزیابی انجام می‌شود و در نهایت بهترین نتیجه به الگوریتم درخت تصمیم تعلق گرفت که می‌تواند با خطای قابل قبولی احتمال وقوع خسارت را براساس متغیرهای بیمه‌نامهٔ ثالث شناسایی کند. مشابه تحقیق فوق، Sharifi *et al.* (2018) تحقیق دیگری در رشتهٔ بیمه‌های زندگی (عمر و درمان) انجام داده‌اند که شدت و احتمال وقوع خسارت را براساس یادگیری ماشین پیش‌بینی می‌کند. Hanafizadeh and Rastkhiz (2013)، تحقیقی با عنوان ارزیابی ریسک و پیش‌بینی دسته‌بندی مشتریان انجام داده که در حوزهٔ ارزیابی ریسک و بر دسته‌بندی مشتریان بیمهٔ اتومبیل تمرکز دارد. با پیش‌بینی دسته‌بندی مشتری، سود بیشتر و طبیعتاً زیان کمتری برای شرکت بیمه به همراه خواهد داشت. بیشتر تحقیقات داخلی انجام‌شده مشابه موارد یادشده است و بر پیش‌بینی احتمال و میزان خسارت و براساس الگوریتم‌های یادگیری نظارت‌شده متمرکز شده‌اند. تعداد زیادی از آن‌ها هم بر رشتهٔ اتومبیل، عمر و درمان تمرکز داشته و در دیگر رشته‌ها، مخصوصاً رشته‌های آتش‌سوزی تحقیقات کمی صورت گرفته است.

پیشینهٔ تحقیقات خارجی

Sakthivel and Rajitha (2017) دربارهٔ استفاده و مقایسهٔ مدل‌های مختلف یادگیری ماشین برای تخمین خسارت آتی در بیمه‌نامه‌های غیرزندگی پژوهش کرده‌اند و یا Gamaliel and Murfi (2020)، با استفاده از شبکهٔ عصبی برای یادگیری ماشین پیش‌بینی خسارت در بیمه‌نامه‌های مسافرتی را بررسی کرده‌اند. Kuo (2019) مدل نگهداری ذخیره برای ریسک براساس شبکهٔ عصبی یادگیری عمیق را ارائه کرده که در این پژوهش نیز ابتدا ریسک یا خسارت آتی پیش‌بینی می‌شود و براساس احتمال و شدت آن، ذخیرهٔ قابل نگهداری بابت ریسک‌های آتی پیش‌بینی می‌شود.

Bücher and Rosenstock (2022) مدل نوآورانه‌ای در سطح مایکرو با بهره‌گیری از شبکهٔ عصبی برای پیش‌بینی تعداد خسارت ارائه کرده‌اند که استفاده از سطح مایکرو شامل شبکهٔ عصبی در یادگیری ماشین، نوآوری این تحقیق برای پیش‌بینی تعداد خسارت آتی است. در همین حوزه Selvakumar *et al.* (2021)، Ogunnaike and Si (2017)، Baran and Rola (2022) مدل‌های پیش‌بینی خسارت را براساس روش‌های داده‌محور و با استفاده از الگوریتم‌های یادگیری ماشین انجام داده و نتایج تحقیقات آن‌ها براساس کمترین خطای الگوریتمی منتشر شده است. در سال‌های قبل نیز این قبیل تحقیقات برای پیش‌بینی خسارت رواج داشته و Wüthrich (2018)، بررسی مروری مقالات و تحقیقات پیش‌بینی خسارت بیمه‌ای با روش‌های مختلف رگرسیون را انجام داده است که بیشتر آن‌ها براساس رگرسیون است و در سنوات بعد به‌کارگیری یادگیری ماشین براساس شبکهٔ عصبی و روش‌های نوین هوش مصنوعی رونق گرفته است.

تعدادی از پژوهش‌های مرتبط با موضوع پژوهش، در جدول ۱ ارائه می‌شود. عمدهٔ پژوهش‌ها در ارزیابی ریسک خسارت بر روی دو معیار احتمال و شدت بروز خسارت صورت گرفته و همچنین در حوزهٔ پیش‌بینی خسارت رشته اموال پژوهش‌های چندانی صورت نگرفته است.

روش‌شناسی پژوهش

با توجه به ضرورت‌های مطرح‌شده، روش‌شناسی پژوهش علم طراحی (Design Science Research Methodology) انتخاب شده است. این روش‌شناسی، بیشتر در علوم کامپیوتر و پژوهش‌های سیستم‌های اطلاعاتی استفاده می‌شود و فولر در سال ۱۹۵۷ آن را به‌عنوان یک روش تحقیق مبتنی بر جمع‌آوری دانش ایجاد کرد که به‌عنوان روشی موفق استفاده می‌شود (Bisandu, 2016).

هونر در سال ۲۰۰۴ راهنمایی برای این روش تحقیق پیشنهاد کرده است که خلاصهٔ آن به شرح زیر است و چرخهٔ تولید و آزمون محصولات برای پوشش نیازمندی و مشکلات در شکل ۲ نمایش داده شده است (Hevner *et al.*, 2004).

- باید دستاورد زندهٔ مبتنی بر فنّاوری و قابل ارائه تولید کند.
- راه‌حل مبتنی بر فنّاوری آن باید مرتبط با یک مشکل کسب‌وکار باشد.
- باید کیفیت و اثربخشی دستاورد مربوطه به‌خوبی ارزیابی شود.
- خروجی‌ها علاوه‌بر فنّاوری‌گرا بودن باید به‌خوبی مدیریت‌گرا هم باشند.

پرسش‌های پژوهش

- سیستم ارزیابی هوشمند خسارت بیمه‌های آتش‌سوزی چگونه است؟
- کدام روش‌های ارزیابی ریسک بیمه و یادگیری ماشین، در ارزیابی ریسک پیش‌بینی خسارت بیمه‌نامه‌های آتش‌سوزی، کارایی بهتری دارند؟

جامعه و نمونهٔ پژوهش

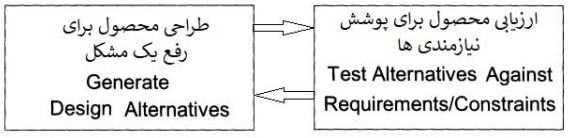
در پژوهش حاضر جامعهٔ پژوهش مورد نظر، بیمهٔ آتش‌سوزی در شرکت‌های بیمه‌ای است و با توجه به سهم بازار بالای شرکت بیمهٔ ایران، مورد مطالعهٔ پژوهش عبارت از داده‌های سیستم اطلاعاتی بیمهٔ آتش‌سوزی شرکت بیمهٔ ایران است. واحد تحلیل، اطلاعات سالیانهٔ فروش و خسارت بیمه‌های آتش‌سوزی است. نمونهٔ پژوهش اطلاعات موجود در بانک اطلاعاتی نرم‌افزار آتش‌سوزی بیمهٔ ایران طی سال‌های ۱۳۹۰ تا ۱۴۰۰ با حدود ۱ میلیون رکورد است.

گام‌های اجرای پژوهش

برای پژوهش حاضر و منطبق با روش تحقیق علم طراحی، ابتدا

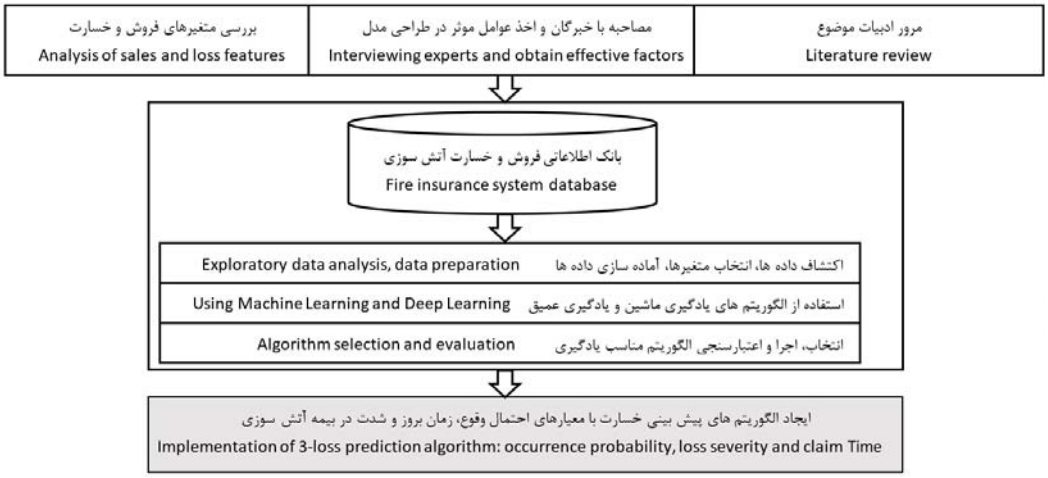
جدول ۱: برخی پژوهش‌های مرتبط به‌عنوان پیشنهٔ پژوهش
Table 1: Some related research as background

عنوان مقاله Article Title	نویسنده(گان) Author(s)	سال Year	روش تحلیل Analysis Method	یافته‌ها و نتایج Results and Contribution
Optimal reinsurance-investment problem under mean-variance criterion with risky assets	(Yang, 2020)	2020	Optimization Mean-Variance	بهینه‌سازی سرمایه‌گذاری اتکایی به‌منظور کاهش ریسک‌های مختلف و در کنار کسب ارزش مورد انتظار
Evaluation methods of flood risk models in the insurance	(Franco and Becker, 2020)	2020	Numerical models Evaluation	ارزیابی مدل‌های ریسک سیلاب در بیمهٔ اتکایی
Research on risk management of big data and machine learning insurance based on internet finance	(Liu, 2019)	2019	Big Data and Machine Learning	مدیریت ریسک بیمه‌ها با استفاده از یادگیری ماشین بر روی فروش بیمهٔ اینترنتی
Machine Learning in Individual Claims Reserving	(Wüthrich, 2018)	2018	claims prediction by regression	بررسی مروری مقالات و تحقیقات پیش‌بینی خسارت بیمه‌ای با روش‌های مختلف رگرسیون
Prediction of motor insurance claims occurrence as an imbalanced Machine Learning problem	(Baran and Rola, 2022)	2022	claim occurrence prediction by Machine learning	مدل پیش‌بینی احتمال وقوع خسارت برای داده‌های نامتوازن با استفاده از یادگیری ماشین
Prediction of insurance claim severity loss using regression models	(Ogunnaike and Si, 2017)	2017	Loss severity prediction Machine Learning	مدل پیش‌بینی شدت خسارت در بیمه‌ها با استفاده از مدل‌های رگرسیون و یادگیری ماشین
Predictive modeling of insurance claims using Machine Learning	(Selvakumar *et al.*, 2021)	2021	ANN in Claim Prediction	یک مدل شبکهٔ عصبی برای پیش‌بینی میزان خسارت در بیمه‌نامه‌های اتومبیل
Micro-level prediction of outstanding claim counts based on mixture models and neural networks	(Bücher and Rosenstock, 2022)	2022	micro-level involving neural networks	مدلی نوآورانه در سطح مایکرو با بهره‌گیری از شبکهٔ عصبی برای پیش‌بینی تعداد خسارت
A deep learning approach to loss reserving	(Kuo, 2019)	2019	loss reserving - deep learning	مدل نگهداری ذخیره برای ریسک براساس شبکهٔ عصبی یادگیری عمیق
The regularization learning network for insurance claim prediction in travel insurance	(Gamaliel and Murfi, 2020)	2020	Deep NN for Claim Prediction	استفاده از شبکهٔ عصبی برای یادگیری ماشین پیش‌بینی خسارت در بیمه‌نامه‌های مسافرتی
AI for estimation of future claim frequency in non-life insurance	(Sakthivel and Rajitha, 2017)	2017	Machine Learning Claim Prediction	مقایسهٔ مدل‌های مختلف یادگیری برای تخمین خسارت آتی در بیمه‌نامه‌های غیرزندگی
کاربرد الگوریتم های یادگیری ماشین در پیش‌بینی ریسک بیمهٔ زندگی	(Sharifi *et al.*, 2018)	۱۳۹۷	تحلیل پیش‌بینی با یادگیری ماشین	رگرسیون خطی چندگانه بهترین الگوریتم برای پیش‌بینی ریسک در بیمهٔ زندگی انتخاب شد
پیش‌بینی مبلغ خسارت برای بیمهٔ شخص ثالث خودرو با استفاده از داده‌کاوی	(Farhang Adib *et al.*, 2014)	۱۳۹۳	داده‌کاوی، خوشه‌بندی و درخت تصمیم	الگوریتم درخت تصمیم بهترین مدل برای پیش‌بینی خسارت در بیمه‌های ثالث انتخاب شد



شکل ۲: چرخهٔ تولید و آزمون
Fig. 2: The generate/test cycle

روند پژوهش معرفی و سپس متغیرهای ورودی، الگوریتم‌های مورد استفاده و خروجی‌ها به‌عنوان فرایند بخش‌های مختلف ارائه می‌شود. در شروع کار و با بررسی دقیق متغیرهای موجود بر پایهٔ مبانی نظری و بررسی پژوهش‌های صورت‌گرفتهٔ قبلی، متغیرهای مورد نیاز برای طراحی مدل ارزیابی را مشخص می‌کنیم. همچنین با مصاحبه با مدیران و خبرگان صنعت، متغیرهای مهم برای ارزیابی خسارت اخذ



شکل ۳: فرایند پیش‌بینی ادعا
Fig. 3: Claim prediction process

می‌شود که روند پژوهش در شکل ۳ مشاهده می‌شود.

براساس مبانی نظری (نظریهٔ ریسک، ارزیابی ریسک بیمه‌ای) مطرح‌شده، باید دسته‌بندی اطلاعات فروش و خسارت براساس متغیرهای مؤثر انجام و الگوریتم‌های یادگیری ماشین، باید الگوریتم‌های ارزیابی پیش‌بینی احتمال، زمان و شدت خسارت طراحی و ایجاد شود.

گردآوری داده‌ها

تمامی داده‌های فروش، الحاقیهٔ بیمه‌نامه و خسارت به‌صورت ریزاطلاعات در بانک‌های اطلاعاتی بیمهٔ ایران نگهداری می‌شود. برخی اقلام اطلاعاتی در بیمه‌نامه‌های آتش‌سوزی به شرح زیر است که برای یادگیری الگوریتم‌های ارزیابی خسارت استفاده می‌شود:

- ریزاطلاعات فروش و خسارت بیمه‌نامه‌های آتش‌سوزی به تفکیک اقلام اطلاعاتی اصلی عبارت‌اند از: نوع بیمه‌نامه، نوع ذی‌نفع، دولتی یا خصوصی، حقیقی یا حقوقی، پوشش‌های بیمه (از قبیل سیل، زلزله، آتش‌سوزی، سرقت) و سقف تعهد آن‌ها، ردهٔ سقف تعهد، حق‌بیمهٔ دریافتی به تفکیک پوشش‌های مختلف (۲۵ پوشش مختلف)، جمع حق‌بیمه، جمع سرمایه و ردهٔ سرمایه، نوع فروش نقدی یا قسطی، استان و شهر فروش و ردهٔ ریسک استان، مرهونه دارد یا خیر، گروه و زیرگروه فعالیت بیمه‌گذار، نوع سازه، خسارت دیده یا خیر، مبلغ خسارت، فاصلهٔ زمانی خسارت بعد از فروش.
- پیش‌بینی احتمال، زمان و شدت ریسک (خسارت). برای محاسبهٔ احتمال و پس از دسته‌بندی اطلاعات بیمه‌نامه‌ها براساس اقلام اطلاعاتی مؤثر، تعداد بیمه‌نامه‌های خسارت‌دیده بر تعداد کل تقسیم و درصد احتمال وقوع خسارت حاصل می‌شود. همچنین برای خسارت‌دیده‌ها، با تقسیم مبلغ خسارت پرداختی بر سقف تعهدات، شدت خسارت مشخص می‌شود و در نهایت برای پیش‌بینی زمان بروز خسارت، مدت‌زمان بین صدور بیمه‌نامه و بروز خسارت براساس روز به‌عنوان این معیار است که برای اطلاعات دسته‌بندی‌شده، میانگین روزهای محاسبه‌شده لحاظ می‌شود.

روش تجزیه‌وتحلیل داده‌ها

پژوهش حاضر، تابع فرایند تحقیق علمی و دارای گام تجزیه‌وتحلیل متغیرهاست. برای این کار، باید با شناخت دقیق مسئله، اقلام اطلاعاتی را به‌درستی شناسایی کنیم که به آن مهندسی ویژگی‌ها (Feature Engineering) می‌گویند و با انتخاب و اجرای الگوریتم‌های مختلف، متغیرهای مؤثر را شناسایی و سپس با استفاده از تجزیه‌وتحلیل اطلاعات، داده‌های مورد نیاز برای مدل‌سازی را به‌دست آوریم. بدین روش اقلام داده‌های مورد نیاز را جمع‌آوری، استانداردسازی، اعتبارسنجی، و نرمال‌سازی می‌کنیم. در نهایت با به‌کارگیری الگوریتم‌های یادگیری ماشین، مدلی مناسب برای یادگیری انتخاب، آزمون و مستقر می‌کنیم. زبان Python و کتابخانه‌های یادگیری ماشین آن، یکی از بهترین ابزارهای این حوزه است.

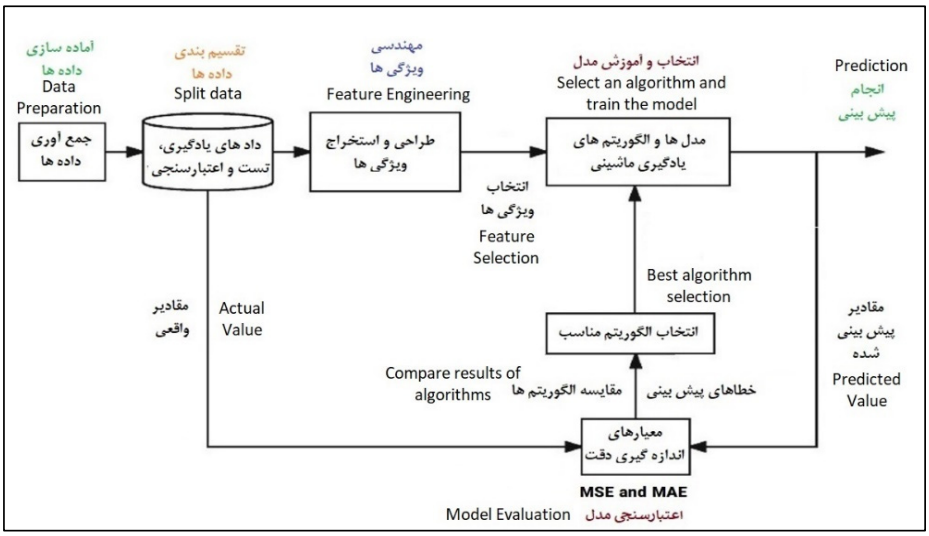
به‌منظور ارزیابی و دستیابی به سه معیار پیش‌بینی خسارت مطرح‌شده و برای هریک، از الگوریتم‌های مختلف یادگیری ماشین استفاده و با تغییر فراپارامترها (Hyper-Parameters) و مقایسهٔ الگوریتم‌ها، بهترین الگوریتم با میزان خطای کمتر انتخاب می‌شود.

نتایج و بحث

مدل پیش‌بینی خسارت بیمه‌نامهٔ آتش‌سوزی

جریان کار پیش‌بینی خسارت براساس یادگیری ماشین

چارچوب پیش‌بینی یادگیری ماشین به هفت مرحلهٔ اصلی تقسیم می‌شود: جمع‌آوری داده، آماده‌سازی داده، انتخاب مدل، آموزش مدل، ارزیابی و اعتبارسنجی مدل، تنظیم پارامتر، و ساخت الگوریتم پیش‌بینی. نمای کلی روش در شکل ۴ نمایش داده شده است. فرایند، سه بار برای سه معیار مختلف پیش‌بینی خسارت (احتمال وقوع خسارت، پیش‌بینی شدت خسارت و پیش‌بینی زمان بروز خسارت) به‌صورت مجزا اجرا می‌شود. دقت مدل با استفاده از دو نوع محاسبهٔ خطاهای مقادیر به‌دست‌آمده با عنوان میانگین خطای مطلق (Mean Absolute Error)



شکل ۴: جریان کار پیش‌بینی خسارت براساس یادگیری ماشین
Fig. 4: Claim prediction workflow according to Machine Learning

یا MAE و میانگین خطای مربعات (Mean Square Error) یا MSE بین مقادیر واقعی و پیش‌بینی‌شده اعتبارسنجی می‌شود.

گردآوری و آماده‌سازی داده‌ها

مجموعهٔ داده‌ها شامل اطلاعات فروش و خسارت بیمه‌نامه‌های آتش‌سوزی شرکت بیمهٔ ایران است که شامل اطلاعات با حدود ۷۰،۰۰۰ حواله خسارت از مجموع ۷۷۸،۲۴۶ بیمه‌نامهٔ صادره برای یک دورهٔ ۱۰ساله از ابتدای سال ۱۳۹۰ تا ابتدای سال ۱۴۰۰ است. برای هر بیمه‌نامه و خسارت اقلام اطلاعاتی مختلفی ثبت شده است. کاوش داده‌ها با استفاده از تجزیه‌وتحلیل داده‌های اکتشافی (Exploratory Data Analysis) یا EDA به محقق این امکان را می‌دهد که توزیع‌های مختلف این ویژگی‌ها را درک کند. در نهایت براساس سؤالات پژوهش و مبانی نظری، تعداد ۴۴ قلم اطلاعات به‌عنوان اقلام اطلاعاتی مرتبط از جداول بانک اطلاعاتی موجود استخراج شده که شامل اقلام اطلاعاتی موجود در بیمه‌نامه و مبلغ و زمان خسارت‌های مرتبط با بیمه‌نامه‌های خسارت‌دیده می‌شود. این اقلام در جدول ۲ نمایش داده شده است.

مهندسی ویژگی‌ها

طراحی ویژگی به‌معنای چگونگی شناسایی ویژگی‌های مهم است، یعنی اقلام اطلاعاتی با همبستگی بالا با متغیر پاسخ را انتخاب کنیم. در این پژوهش از ترکیبی از توابع برای محاسبهٔ اهمیت هر متغیر پیش‌بینی‌کننده استفاده شده است. این فرایند به انتخاب ویژگی‌ها، کاهش ابعاد و استخراج ویژگی منجر می‌شود که در این بخش هریک و راه‌حل خاص آن توضیح داده و مراحل این کار در شکل ۵ نمایش داده شده است.

پیش از نرمال‌سازی و کاهش ابعاد، باید دسته‌بندی لازم برای اقلام اطلاعاتی انجام پذیرد. در واقع مسئلهٔ اصلی در مجموعه داده این است که در داده‌های موجود نمونه‌های زیادی با ابعاد مشابه و نتایج متفاوت داریم، به‌طور مثال در یک شهر، با سازهٔ یکسان، ویژگی‌های یکسان و در بسیاری از موارد هیچ خسارتی رخ نداده و برای برخی بیمه‌نامه‌ها خسارت رخ داده است. بنابراین برای هریک از سه مجموعه داده (مورد استفاده در سه الگوریتم پیش‌بینی ریسک) دسته‌بندی لازم برای ابعاد اطلاعاتی مشترک انجام و قلم اطلاعاتی خروجی نیز سرجمع شد. بنابراین برای داشتن داده‌های تمیز و قابل اعتماد، باید ابعاد را گروه‌بندی کنیم و از توابع جمع‌بندی مانند میانگین استفاده کنیم تا رکوردهایی منحصربه‌فرد از اطلاعات در هر مجموعه داده داشته باشیم. این فرایند در شکل ۵ با عنوان گروه‌بندی داده‌ها نمایش داده شده است. حال سه مجموعه دادهٔ مختلف با گروه‌بندی ابعاد و محاسبهٔ میانگین وقوع خسارت (با تقسیم تعداد خسارت برای همان ابعاد بر تعداد کل آن نوع بیمه‌نامه‌ها)، شدت خسارت (متوسط مبلغ خسارت پرداختی برای همان گروه از ابعاد) و زمان بروز خسارت (میانگین زمان خسارت برای همان گروه از ابعاد) در اختیار داریم که ما را به داشتن داده‌های مناسب برای یادگیری ماشین هدایت می‌کند. هر الگوریتم دارای ابعاد ورودی و یک عدد پیوسته به‌عنوان خروجی است، حال شروع مرحلهٔ بعد در طراحی ویژگی‌ها، نرمال‌سازی، استانداردسازی و مدیریت داده‌های خارج محدوده (Outlier detection) است. منظور از نرمال‌سازی تبدیل داده‌ها به دامنهٔ [۰ و ۱] و یا دامنهٔ [۱- و ۱+] آموزش بهتر مدل است. منظور از استانداردسازی، متقارن کردن توزیع داده‌هاست.

اکنون باید ویژگی‌های موجود را به تعداد بهینه کاهش دهیم. کاهش ابعاد شامل کاهش تعداد متغیرهایی است که برای مدل‌سازی کارآمد استفاده می‌شود. روش‌های انتخاب و استخراج ویژگی‌های زیادی وجود دارد، ما از کاهش بازگشتی ابعاد یا RFE و همچنین از انتخاب تحلیل همبستگی برای انتخاب ویژگی استفاده می‌کنیم. یک انتخاب مبتنی بر تحلیل همبستگی، تحت روش حداکثر ضریب اطلاعات، ویژگی‌ها را با استفاده از بهینه‌سازی غیرخطی انتخاب

جدول ۲: اقلام اطلاعاتی مجموعه دادهٔ موجود پس از تجزیه‌وتحلیل اکتشافی
Table 2: Dataset fields after EDA

ردیف Row	عنوان قلم اطلاعاتی Field title	عنوان انگلیسی در مجموعه دادهٔ نهایی Field English title in dataset	شرح و نمونهٔ اطلاعات Description and sample value
۱	نوع گروه بیمه‌نامه	SUBFIELDGROUP_CODE	صنعتی، مسکونی و تجاری
۲	نوع ذی‌نفع بیمه‌گذار	STAKEHOLDER_TYPECODE	شخصی، شرکتی، بانک، خودبیمه‌گذار
۳	زیررشتهٔ بیمه‌ای	SUBFIELD_CODE	منازل مسکونی، ویلایی، کارخانه‌ها، صنایع کوچک
۴	مرهونه دارد یا خیر؟	IS_MORTGAGE	بلی، خیر
۵	نوع اقتصاد بیمه‌گذار	ECONOMIC_TYPECODE	خصوصی، دولتی
۶	نوع فروش بیمه‌نامه	SALES_TYPE_CODE	نقدی، اقساطی
۷	نوع بیمه‌گذار	CORP_STATUS_CODE	حقیقی، حقوقی
۸	شهر مورد بیمه	CITY_ID	کد شهرها
۹	استان مورد بیمه	PROVINCEID	کد استان‌ها
۱۰	سطح ریسک استان	PROVINCERISKLEVEL	۵ سطح موجود: کم‌ریسک تا پرخطر در بیمهٔ آتش‌سوزی مشخص شده است
۱۱	سرگروه فعالیت بیمه‌گذار	TOP_GROUP_ID	۱۲ گروه تعریف‌شده شامل خانگی، انبار، کارخانه‌ها و ...
۱۲	گروه فعالیت بیمه	GROUP_ID	۳۰ گروه تعریف‌شده
۱۳	نوع سازهٔ ساختمان	STRUCTURE_CODE	آجری، خشتی، آهنی، بتونی و گلی
۱۴	سطح سرمایهٔ مورد بیمه	CAPITAL_LEVEL	۱۰ سطح تعریف‌شده براساس مبلغ سرمایه
۱۵ تا ۳۹	انواع پوشش‌های مورد بیمه	COVERAGES (25 items)	۲۵ پوشش مختلف و هریک با تعهد ریالی مشخص مانند آتش‌سوزی، انفجار، سیل، رانش، سرقت، آشوب و بلوا، ترکیدگی لوله، هزینهٔ پاک‌سازی، زمین‌لرزه، اتشفشان
۴۰	سطح تعهدات	TOTALOBLIGATION_LEVEL	براساس هر پوشش تعهدی مشخص می‌شود و براساس جمع ریالی تعهدات یک سطح تعیین می‌شود و ۱۰ سطح مشخص دارد
۴۱	مبلغ حق‌بیمه	PREMIUM_AMOUNT	مبلغ پرداختی بابت بیمه‌نامه که درآمد ناخالص حاصل از فروش بیمه‌نامه است
۴۲	آیا خسارت دارد یا خیر	IS_LOSS	اگر خسارت دیده باشد بلی، وگرنه خیر
۴۳	زمان وقوع خسارت	LOSS_TIME	اگر خسارت دیده باشد، تعداد روز مابین صدور بیمه‌نامه و بروز خسارت
۴۴	شدت خسارت	DAMAGED_RATE	اگر خسارت داشته باشد، نسبت خسارت(های) پرداختی به جمع تعهدات به‌صورت درصدی

شکل ۵: جریان کار مهندسی ویژگی‌ها
Fig. 5: Feature engineering workflow

می‌کند که به‌دنبال بهینه‌سازی رابطه بین دقت و تنوع خروجی‌های پایه است (Hu *et al.*, 2017). حذف بازگشتی ویژگی‌ها یا RFE، یک روش پوششی و حذف‌کننده در انتخاب ویژگی‌هاست. در این روش ویژگی اضافی و ضعیف را حذف می‌کند به شرطی که حذف آن کمترین تأثیر را بر خطای یادگیری ماشین دارد و از طرفی ویژگی مستقل و قوی را برای بهبود عملکرد مدل حفظ می‌کند (Mathew, 2019). برای هر الگوریتم، همهٔ این روش‌ها و کل فرایند مهندسی ویژگی‌ها را اجرا کرده‌ایم و برای هر معیار پیش‌بینی خسارت، مجموعه‌ای از ویژگی‌های متفاوت را به‌دست آورده‌ایم. ویژگی‌های انتخاب‌شده برای هر الگوریتم در جدول ۳ نشان داده شده است

انتخاب، اجرا و ارزیابی الگوریتم‌های یادگیری ماشین

نوع مسئله برای هر سه معیار پیش‌بینی خسارت، به‌صورت یادگیری ماشین نظارت‌شده و خروجی پیوسته است. در این شرایط

به‌کارگیری مدل‌های یادگیری رگرسیون استفاده می‌شود. براساس پیشینهٔ پژوهش‌ها و بررسی پروژه‌های کاربردی در دنیا، تعدادی از برترین مدل‌های طراحی و پیاده‌سازی نظارت‌شدهٔ رگرسیونی شامل رگرسیون خطی (Linear Regression)، رگرسیون بردار پشتیبان (Support Vector Regression) یا SVR، رگرسیون جنگل تصادفی (Random Forest Regression) و یادگیری عمیق است. برای هر سه معیار پیش‌بینی خسارت (پیش‌بینی احتمال وقوع خسارت، شدت خسارت و زمان بروز خسارت) ، این مدل‌ها را اجرا می‌کنیم. در هر بار اجرا، تنظیم فراپارامترها انجام و خطای پیش‌بینی با معیارهای اندازه‌گیری MSE و MAE محاسبه شد و در نهایت بهترین مدل با بهترین پارامترها برای هر الگوریتم انتخاب می‌شود. به‌منظور آشنایی با هریک از مدل‌های به‌کارگرفته‌شده، توضیحاتی ارائه می‌شود.

رگرسیون خطی یک مدل پایه از تجزیه‌وتحلیل پیش‌بینی‌کننده است. هدف پایه‌ای رگرسیون بررسی این موضوع است: آیا مجموعه‌ای از متغیرهای پیش‌بینی‌کننده، به‌درستی پیش‌بینی یک متغیر نتیجه را انجام می‌دهند؟ (Jadhav *et al.*, 2021) ماشین بردار پشتیبان (Support Vector Machine) یا SVM را کورتس و واپنیک پیشنهاد کرده‌اند. این مدل یک الگوریتم قدرتمند و قوی یادگیری ماشین برای طبقه‌بندی و رگرسیون در نظر گرفته می‌شود و برای پیش‌بینی موارد دسته‌بندی و رگرسیون کاربرد دارد (Malik *et al.*, 2020).

مدل رگرسیونی SVM را با اختصار SVR نام‌گذاری کرده‌اند که یک رویکرد رگرسیون‌محور مبتنی بر SVM است و یک تابع را با به حداکثر رساندن تعداد انحرافات واقعی به‌دست‌آمده در اهداف، در داخل نوار حاشیهٔ نرمال‌شده تخمین می‌زند. الگوریتم SVR یک مسئلهٔ کمینه‌سازی محدب است که بردار نرمال تابع خطی را پیدا می‌کند (Lee *et al.*, 2020). رگرسیون جنگل تصادفی یک الگوریتم مجموعه‌ای است که چندین درخت رگرسیون را ترکیب می‌کند و یک جنگل از درخت‌های رگرسیون تشکیل می‌دهد. هر درخت با استفاده از زیرمجموعه‌ای تصادفی از ویژگی‌ها آموزش داده می‌شود و خروجی میانگین درخت‌هاست (Dewi *et al.*, 2019).

شبکه‌های عصبی روش یادگیری ماشین بسیار محبوبی هستند. آن‌ها ثابت کرده‌اند که در موقعیت‌های مختلف بسیار قدرتمندند و به عبارت ساده، شبکه‌های عصبی را می‌توان به‌عنوان توابع رگرسیون غیرخطی با ابعاد بالا برای پیش‌بینی یا طبقه‌بندی مشاهدات استفاده کرد(Mare *et al.*, 2022). شبکهٔ عصبی‌ای که تعداد زیادی لایهٔ پنهان داشته باشد شبکهٔ عصبی عمیق ((Deep Neural Network گفته می‌شود. مدل MLP با لایه‌های متعدد نیز نوعی از یادگیری عمیق تمام‌متصل چندلایهٔ پیش‌انتشار (Feedforward Fully Connected MLP) است (Géron, 2018).

در پیاده‌سازی مدل‌ها متوجه شدیم که با توجه به تعداد زیاد

جدول ۳: اقلام اطلاعاتی انتخاب‌شده برای هریک از سه معیار پیش‌بینی خسارت
Table 3: Selected fields for each loss prediction measure

الگوریتم پیش‌بینی خسارت Loss prediction algorithm	اقلام اطلاعاتی طراحی و انتخاب‌شده Designed and selected fields
پیش‌بینی احتمال وقوع خسارت: ورودی ۲۴ قلم اطلاعاتی تعداد رکوردها پس از دسته‌بندی: ۱۳۱،۸۰۴ رکورد مجزا خروجی به‌صورت عدد پیوسته: احتمال درصدی وقوع خسارت براساس تعداد خسارت/ تعداد بیمه‌نامه	CITY_ID, CAPITAL_LEVEL, STRUCTURE_CODE, PROVINCEID, PROVINCERISKLEVEL,COVER03_OBLIGATED,COVER02_OBLIGATED, COVER04_OBLIGATED, SALES_TYPE_CODE, COVER05_OBLIGATED, CORP_STATUS_CODE,COVER24_OBLIGATED,GROUP_ID,STAKE_HOLDER_TYPE_CODE, SUBFIELDCODE, TOP_GROUP_ID, COVER14_OBLIGATED, COVER18_OBLIGATED, COVER09_OBLIGATED, COVER22_OBLIGATED, ECONOMIC_TYPE_CODE, SUBFIELDGROUPCODE, COVER06_OBLIGATED,TOTAL_COVER_OBLIGATION_LEVEL Output: CLAIM_OCCUR_RATE
پیش‌بینی شدت خسارت: ورودی ۲۵ قلم اطلاعاتی تعداد رکوردها پس از دسته‌بندی: ۳۰،۷۷۹ رکورد مجزا خروجی به‌صورت عدد پیوسته: درصد شدت خسارت براساس خسارت‌ها/ سقف تعهد	CITY_ID, CAPITAL_LEVEL, STRUCTURE_CODE, PROVINCEID, PROVINCERISKLEVEL, IS_MORGAGE_CODE, SALES_TYPE_CODE, COVER02_OBLIGATED COVER24_OBLIGATED, STAKE_HOLDER_TYPE_CODE, CORP_STATUS_CODE, GROUP_ID, TOP_GROUP_ID, ECONOMIC_TYPE_CODE, COVER05_OBLIGATED, COVER13_OBLIGATED, COVER08_OBLIGATED, SUBFIELDGROUPCODE, COVER35_OBLIGATED, COVER04_OBLIGATED, COVER22_OBLIGATED, SUBFIELDCODE, COVER06_OBLIGATED, COVER03_OBLIGATED,TOTAL_COVER_OBLIGATION_LEVEL Output: LOSS_SEVERITY_RATE
پیش‌بینی زمان بروز خسارت: ورودی ۲۲ قلم اطلاعاتی تعداد رکوردها پس از دسته‌بندی: ۳۱،۵۳۰ رکورد مجزا خروجی به‌صورت عدد پیوسته : تعداد روز بروز خسارت بعد از صدور بیمه‌نامه	SUBFIELDGROUPCODE, CITY_ID, CAPITAL_LEVEL, STRUCTURE_CODE, PROVINCEID, PROVINCERISKLEVEL,SALES_TYPE_CODE, COVER03_OBLIGATED, COVER05_OBLIGATED, COVER02_OBLIGATED, GROUP_ID, COVER35_OBLIGATED, COVER06_OBLIGATED, CORP_STATUS_CODE, COVER04_OBLIGATED, TOP_GROUP_ID, STAKE_HOLDER_TYPE_CODE, SUBFIELDCODE, COVER22_OBLIGATED, COVER24_OBLIGATED, IS_MORGAGE_CODE, TOTAL_COVER_OBLIGATION_LEVEL Output: DAYDIFF_F_SALES

جدول ۴: بهترین پارامترها و بهترین الگوریتم برای معیارهای پیش‌بینی خسارت
Table 4: Best algorithm and tuned hyper-parameter for loss prediction measures

مشخصات الگوریتم انتخاب‌شده برای پیش‌بینی خسارت Selected algorithms results			نتایج اعتبارسنجی سایر الگوریتم‌های انتخاب‌نشده Unselected algorithms evaluation results	
الگوریتم Algorithm	فراپارامترها Hyper-parameter	اعتبارسنجی Evaluation		
MLP پیش‌بینی احتمال وقوع خسارت	batch size = 32, epochs = 50, verbose = 1, units: 49, 23, 7	MSE: 0.11 MAE: 0.23	Regression Linear	MAE: 0.52 MSE: 0.33
			Forest Random	MAE: 0.41 MSE: 0.20
			SVR	MAE: 0.38 MSE: 0.19
MLP پیش‌بینی شدت خسارت	batch size =100, epochs = 32, verbose = 1, units: 61,29,13	MSE: 0.04 MAE: 0.07	Regression Linear	MAE: 0.09 MSE: 0.05
			Forest Random	MAE: 0.25 MSE: 0.12
			SVR	MAE: 0.16 MSE: 0.09
MLP پیش‌بینی زمان بروز خسارت	batch size = 16, epochs = 50, verbose = 1, units: 67, 23, 6	MSE: 0.10 MAE: 0.21	Regression Linear	MAE: 0.35 MSE: 0.18
			Forest Random	MAE: 0.27 MSE: 0.13
			SVR	MAE: 0.52 MSE: 0.30

ورودی‌ها، مدل‌های یادگیری غیرشبکهٔ عصبی خطای زیادی دارند و ازاین‌رو و بر پایهٔ پژوهش‌های پیشین در حوزهٔ یادگیری ماشین، نوعی از یادگیری شبکهٔ عصبی عمیق و به‌صورت تمام‌متصل چندلایه با نام MLP انتخاب و اجرا شد. این مدل روشی کارآمد برای حل مسائل رگرسیون است و می‌تواند روابط غیرخطی پیچیده بین متغیرهای ورودی و خروجی را بیاموزد. یکی از مزایای اصلی آن نسبت به سایر الگوریتم‌ها، توانایی در مدیریت داده‌ها با تعداد ویژگی زیاد است.

اعتبارسنجی مدل‌ها

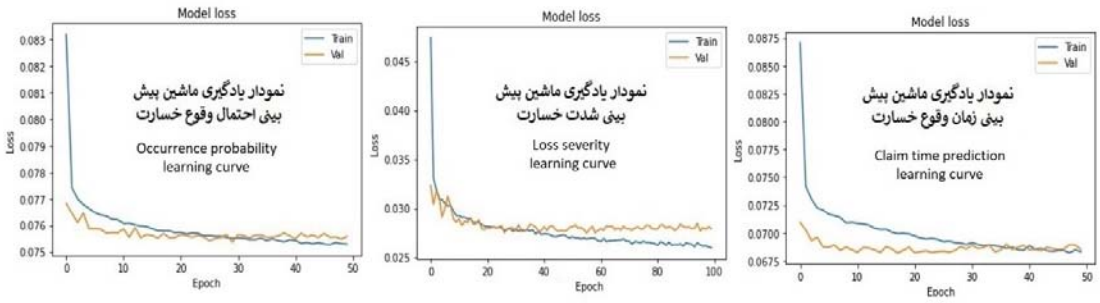
پس از برازش و مقایسهٔ نتایج MSE و MAE متوجه شدیم که در هر سه الگوریتم پیش‌بینی، مدل یادگیری عمیق از نوع شبکهٔ عصبی پرسپترون چندلایه (MLP) بهترین عملکرد را دارد و برای هریک فراپارامترهای بهینه نیز شناسایی شد. در جدول ۴ الگوریتم انتخابی به همراه بهترین تنظیمات فراپارامترها معرفی شده است و برای مقایسهٔ نتایج ارزیابی خطای الگوریتم‌ها، مقدار MSE و MAE الگوریتم مطلوب با سایر الگوریتم‌ها در کنار هم قرار گرفته و قابل مقایسه و مشاهده است.

در انتها باید به مورد دیگری نیز توجه شود که می‌تواند کارایی مدل را زیر سؤال ببرد. این موضوع نحوهٔ مرحلهٔ یادگیری است که براساس شرایط داده‌های ورودی و تنظیمات الگوریتم، ممکن است دچار بیش‌برازش (Overfitting) یا کم‌برازش (Under fitting) شده باشیم. بر همین اساس و پس از اجرای مدل‌ها، منحنی‌های یادگیری برای سه الگوریتم ترسیم و مشاهده شد که هیچ تطابق بیش از حد یا کمتری اتفاق نیفتاده است. این نمودارها در نمودار ۲ نمایش داده شده است. منحنی‌های یادگیری در یادگیری ماشین به‌طور کلی، منحنی یادگیری است که زمان یا تجربه را در محور x و یادگیری یا بهبود را در محور y نشان می‌دهد. یک منحنی یادگیری مقدار بهینه تابع خطای یک مدل را برای یک مجموعهٔ آموزشی در برابر همان تابع بر روی داده‌های اعتبارسنجی ارزیابی می‌کند.

نمونه‌ای از اجرای مدل ارزیابی هوشمند خسارت بیمه‌نامه‌های آتش‌سوزی در شکل ۶ مشاهده می‌شود. در فرم طراحی‌شده، اقلام اطلاعاتی بیمه‌نامهٔ آتش‌سوزی به‌عنوان ورودی در نظر گرفته شده است که کاربر می‌تواند اطلاعات را وارد کند. سپس این اقلام براساس توضیحات ارائه‌شده نرمال‌سازی و استانداردسازی می‌شود. اقلام تمیزکاری‌شده اطلاعاتی مرتبط با هریک از معیارهای پیش‌بینی خسارت، به‌صورت مستقل به الگوریتم‌های آماده‌شده پیش‌بینی خسارت ارسال و در کسری از ثانیه، خروجی هریک ارائه می‌شود. خروجی شامل پیش‌بینی احتمال وقوع خسارت، پیش‌بینی شدت خسارت (درصد خسارت احتمالی براساس مبلغ سقف تعهد) و پیش‌بینی مدت‌زمان بروز خسارت بعد از صدور بیمه‌نامه می‌شود. این خروجی کاربردی یکی از دستاوردهای پژوهش است.

جمع‌بندی و پیشنهادها

مدل ارائه‌شده یک سیستم ارزیابی و پیش‌بینی خسارت بیمهٔ آتش‌سوزی است که براساس مبانی نظری ارزیابی ریسک بیمه‌ای و نظریهٔ ریسک و نظر خبرگان تحلیل و طراحی شد. اقلام اطلاعاتی و داده‌ها از بانک اطلاعاتی سامانهٔ آتش‌سوزی شرکت بیمهٔ ایران برای یک دورهٔ ده‌ساله (ابتدای ۱۳۹۰ تا ابتدای ۱۴۰۰) استفاده شده است. بر این اساس سه الگوریتم پیش‌بینی خسارت برای پیش‌بینی احتمال وقوع خسارت، احتمال شدت خسارت و پیش‌بینی زمان بروز خسارت طراحی شد که پیش‌بینی زمان بروز خسارت در کنار پیش‌بینی احتمال و شدت خسارت به‌عنوان بهبود تحقیقات پیشین ارائه شده است. در نهایت برای هر سه الگوریتم، یادگیری عمیق به‌عنوان بهترین روش انتخاب شد که با انجام تنظیمات لازم به حداقل خطای ممکن در الگوریتم‌ها دست یافتیم. در این راستا اعتبارسنجی الگوریتم‌های انتخاب‌شده‌ها با استفاده از میانگین مربعات خطا به میزان قابل قبول ۰,۱۱۷ (برای احتمال وقوع)، ۰,۰۴۲ (برای پیش‌بینی شدت



نمودار ۲: منحنی‌های یادگیری ماشین معیارهای پیش‌بینی خسارت
Chart 2: Loss prediction measures learning curves

شکل ۶: تصویر اجرای فرم اصلی سیستم به‌عنوان دستاورد پژوهش
Fig. 6: Main form of system as a research contribution

خسارت)، ۰,۱۰۶ (برای پیش‌بینی زمان بروز خسارت) حاصل شد. به‌کارگیری یادگیری عمیق به‌صورت داده‌محور برای داده‌های بیمهٔ آتش‌سوزی نسبت به دیگر روش‌های بررسی‌شده، کارایی بهتری داشته و افزودن معیار پیش‌بینی زمان بروز خسارت موجب بهبود روش‌های رایج پیش‌بینی خسارت شده است. پس از پیاده‌سازی الگوریتم‌ها، اقلام اطلاعاتی یک بیمه‌نامه به سیستم ارزیابی هوشمند پیش‌بینی خسارت ارائه می‌شود و سیستم براساس الگوریتم‌های یادگیری‌شده، بیمه‌نامه را ارزیابی و سه خروجی بابت درصد احتمال وقوع خسارت، میزان شدت خسارت احتمالی و پیش‌بینی تعداد روز بروز خسارت بعد از صدور بیمه‌نامه را ارائه می‌کند. بخش پیش‌بینی زمان بروز خسارت در صورت به‌کارگیری می‌تواند به‌عنوان عامل مهمی در تصمیم‌گیری شرکت‌های بیمه استفاده شود. بدین نحو که در مدت‌زمان رسیدن به خسارت احتمالی می‌توانند از درآمد و ذخایر خسارت درنظرگرفته‌شده برای هر بیمه‌نامه در بازارهای دیگر سرمایه‌گذاری کنند و منبع مالی ذخیرهٔ خسارت می‌تواند کاراتر باشد.

پیشنهادها شامل پیشنهادهای کاربردی و پژوهشی است. در بخش کاربردی، پیشنهاد می‌شود از این مدل در ارزیابی ریسک خسارت بیمه‌نامه‌های آتش‌سوزی استفاده شود و با ارائهٔ هر بیمه‌نامه به سامانه، خروجی‌های پیش‌بینی ریسک در کوتاه‌ترین زمان و با دقت مناسب توسط سیستم ایجاد و از آن‌ها برای تصمیم‌گیری یا حتی ارسال به دیگر واحدهای ذی‌نفع (اتکایی، قیمت‌گذاری) استفاده شود. در بخش پژوهشی و به‌عنوان اولین پیشنهاد، انجام پژوهشی در همین حوزه و بر روی اطلاعات سایر شرکت‌های بیمه‌ای و سایر رشته‌های بیمه است. پیشنهاد دوم پژوهشی، انجام پژوهش‌های تکمیلی در راستای این پژوهش است، به‌نحوی که خروجی این پژوهش به‌عنوان پیش‌بینی خسارت در تعیین سطح ریسک در مدیریت ریسک و یا تصمیم‌گیری در اتکایی موردی آتش‌سوزی به کار گرفته شود. همچنین می‌توان پژوهش در حوزهٔ قیمت‌گذاری و اعمال تخفیف یا تقسیط بیمه‌نامه‌های آتش‌سوزی براساس خروجی این پژوهش و در ادامهٔ آن صورت پذیرد. به‌طور کلی خروجی این

پژوهش به‌عنوان ارزیابی ریسک و پیش‌بینی خسارت بیمه‌نامه‌های آتش‌سوزی می‌تواند در مدیریت ریسک، مدیریت سرمایه‌گذاری، مدیریت اتکایی و قیمت‌گذاری استفاده شود.

مشارکت نویسندگان

این پژوهش مستخرج از رسالهٔ دکتری مرتضی امیرحسینی است و بررسی مفاهیم و ادبیات موضوعی مرتبط با پژوهش، مطالعهٔ پیشینهٔ پژوهش، بررسی و جمع‌آوری داده‌ها، طراحی و پیاده‌سازی مدل، کنترل صحت مدل و بحث و نتیجه‌گیری توسط وی انجام شده است. دکتر ایمان رئیسی وانانی به‌عنوان استاد راهنما و دکتر محمدتقی تقوی فرد و دکتر بابک سهرابی به‌عنوان استادان مشاور در تمام مراحل انجام پژوهش، نظارت و راهبری داشته‌اند.

تشکر و قدردانی

از مدیران و کارشناسان شرکت معظم بیمهٔ ایران که به‌عنوان خبرگان این حوزه دانش و داده‌های مورد نیاز را در اختیار پژوهشگران این پژوهش گذاشته‌اند، کمال تشکر را داریم.

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) ©2023: این مقاله تحت مجوز بین‌المللی Creative Commons Attribution 4.0 اجازهٔ استفاده، اشتراک‌گذاری، اقتباس، توزیع و تکثیر را در هر رسانه یا قالبی مشروط بر درج نحوهٔ دقیق دسترسی به مجوز CC منوط به ذکر تغییرات احتمالی بر روی مقاله می‌داند. ازاین‌رو به استناد مجوز مذکور، درج 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 مذکور و یا استفادهٔ فراتر از مجوز فوق، نویسنده ملزم به دریافت مجوز حق نسخه‌برداری از شخص ثالث است.

به‌منظور مشاهدهٔ مجوز بین‌المللی Creative Commons Attribution 4.0 به نشانی زیر مراجعه شود:

http://creativecommons.org/licenses/by/4.0

یادداشت ناشر

ناشر نشریهٔ پژوهشنامهٔ بیمه با توجه به مرزهای حقوقی در نقشه‌های منتشرشده بی‌طرف باقی می‌ماند.

منابع

Baran, S.; Rola, P., (2022). Prediction of motor insurance claims occurrence as an imbalanced machine learning problem. Department. Math. Cracow. Univ. Econ., 2022(1): 1-12 **(12 Pages)**.

Bisandu, D.B., (2016). Design science research methodology in computer science and information systems. Int. J. Inf. Technol., 1(1): 1-6 **(6 Pages)**.

Breuer, A.; Staudt, Y., (2022). Equalization reserves for reinsurance and non-life undertakings in Switzerland. Risks., 10(3): 1-41 **(41 Pages)**.

Bücher, A.; Rosenstock, A., (2022). Micro-level prediction of outstanding claim counts based on novel mixture models and neural networks. Eur. Actuarial. J., 13(385): 55-90 **(36 Pages)**.

Chen, Z.; Yang, P., (2020). Robust optimal reinsurance–investment strategy with price jumps and correlated claims. Insur. Math. Econ., 92(1): 27-46 **(20 Pages)**.

Deelstra, G.; Plantin, G., (2014). Risk theory and reinsurance. Springer.

Dewi, K.C.; Murfi, H.; Abdullah, S., (2019). Analysis accuracy of random forest model for big data – A case study of claim severity prediction in car insurance. Fifth international conference on science in information technology., 5(46713): 60-65 **(6 Pages)**.

Farhang Adib, S.; Damyar, A.; Minaee, B., (2014). Predicting the amount of damage for third party car insurance policies using data mining algorithms. The second national conference on applied research in computer science and information technology., 2(33): 454-468 **(15 Pages)**. [In Persian]

Franco, G.; Becker, J.F., (2020). Evaluation methods of flood risk models in the (re)insurance industry. Water. Secur., 11(1): 294-302 **(9 Pages)**.

Gamaliel, J.T.; Murfi, H., (2020). Regularization learning network for insurance claim prediction in travel insurance. J. Adv. Res. Dyn. Control. Syst., 12(SP4): 1496-1503 **(8 Pages)**.

Géron, A., (2018). Neural networks and deep learning. O'reilly. Media., 1(1): 52-97 **(46 Pages)**.

Hanafizadeh, P.; Rastkhiz, N., (2013). Comparison of two data mining methods in the segmentation of car body insurance customers based on risk (Case study: Mellat insurance company). Ind. Manage. Stud., 11(30): 563-584 **(22 Pages)**. [In Persian]

Hevner, R.A.; Salvator, T.M.; Park, J.; Ram, S., (2004). Design science in information science. Manage. Inf. Syst. Res. Center. Univ. Minnesota., 28(1): 75-105 **(31 Pages)**.

Hu, X.; Duan, B.; Zhang, L., (2017). De vylder approximation to the optimal retention for a combination of quota-share and excess of loss reinsurance with partial information. Insur. Math. Econ., 76(1): 48-55 **(8 Pages)**.

Hull, J.C., (2015). Risk management and financial institutions. Wiley.

Jadhav, A.; Kulkarni, M.; Abute, P.; Rajarapollu, P., (2021). Design & development of insurance money predictor to claim with insurance company. Global conference for advancement in technology (GCAT)., 2(1): 1-4 **(4 Pages)**.

Kuo, K., (2019). Deep triangle: A deep learning approach to loss reserving. Risks., 7(3): 97-105 **(9 Pages)**.

Lee, H.; Wang, J.; Leblon, B., (2020). Using linear regression,

random forests, and support vector machine with unmanned Aerial vehicle multispectral images to predict canopy Nitrogen weight in corn. Remote. Sens., 12(13): 312-325 **(14 Pages)**.

Lentz, T.J.; Dotson, G.S.; Williams, P.R.; Maier, A.; Gadagbui, B.; Pandalai, S.P.; Lamba, A.; Hearl, F.; Mumtaz, M., (2015). Aggregate exposure and cumulative risk assessment--integrating occupational and non-occupational risk factors. J. Occup. Environ. Hyg., 12(sup1): 112-126 **(15 Pages)**.

Liu, Q., (2019). Research on risk management of big data and machine learning insurance based on internet finance. J. Phys. Conf. Ser., 1345(5): 65-76 **(12 Pages)**.

Malik, A.; Tikhamarine, Y.; Souag-Gamane, D.; Kisi, O.; Pham, Q.B., (2020). Support vector regression optimized by meta-heuristic algorithms for daily streamflow prediction. Stochastics. Environ. Res. Risk. Assess., 34(1): 1755-1773 **(19 Pages)**.

Mare, C.; Manaţe, D.; Mureşan, G.M.; Dragoş, S.L.; Dragoş, C.M.; Purcel A.A., (2022). Machine learning models for predicting Romanian farmers' purchase of crop insurance. Math., 10(19): 1-13 **(13 Pages)**.

Mathew, T.E., (2019). A logistic regression with recursive feature elimination model for breast. Int. J. Emerging. Technol., 10(3): 55-63 **(9 Pages)**.

Ogunnaike, R.M.; Si, D., (2017). Prediction of insurance claim severity loss using regression models. International conference on machine learning and data mining in pattern recognition., 10358(12): 233-247 **(15 Pages)**.

Olarewaju, O.M.; Msomi, T.S., (2022). Factors affecting the profitability of reinsurance companies in Sub-Saharan Africa: Evidence from dynamic panel analysis. Cogent. Bus. Manage., 9(1): 1-18 **(18 Pages)**.

Payandeh Najafabadi, A.T., (2020). Risk theory in non-life insurance. Tehran: Insurance Research Center. [In Persian]

Roeser, S.; Hillerbrand, R.; Sandin, P.; Petersen, M., (2012). Handbook of risk theory: Epistemology, Decision theory, Ethics, and social implications of risk. Springer.

Sakthivel, K.M.; Rajitha, C.S., (2017). Artificial intelligence for estimation of future claim frequency in non-life insurance. Global. J. Pure. Appl. Math., 13(6): 1701-1710 **(10 Pages)**.

Selvakumar, V.; Satpathi D.K.; Praveen Kumar, P.T.V.; Haragopal, V.V., (2021). Predictive modeling of insurance claims using machine learning approach for different types of motor vehicles. Univers. J. Accounting. Finance., 9(1): 1-14 **(14 Pages)**.

Sharifi, M.; Sarbakhshian, A.; Rashidi, A., (2018). Application of machine learning algorithms in life insurance risk prediction. National conference on insurance and development., 1-20 **(20 Pages)**. [In Persian]

Shiu, Y.M., (2020). How does reinsurance and derivatives usage affect financial performance? Evidence from the UK non-life insurance industry. Econ. Model., 88(1): 376-385 **(10 Pages)**.

Wüthrich, M.V., (2018). Machine learning in individual claims reserving. Scand. Actuarial. J., 2018(6): 465-480 **(16 Pages)**.

Yang, P., (2020). Optimal reinsurance-investment problem under mean-variance criterion with risky assets. Discrete. Dyn. Nat. Soc., 2020(6489532): 1-16 **(16 Pages)**.

معرفی نویسندگان

AUTHOR(S) BIOSKETCHES

ایمان رئیسی وانانی، دانشیار گروه مدیریت عملیات و فناوری اطلاعات، دانشکدهٔ مدیریت و حسابداری، دانشگاه علامه طباطبایی، تهران، ایران

- Email: *imanraeesi@atu.ac.ir*
- ORCID: 0000-0001-8324-9896
- Homepage: https://aris.atu.ac.ir/imanraeesi

محمدتقی تقوی فرد، استاد گروه مدیریت عملیات و فناوری اطلاعات، دانشکدهٔ مدیریت و حسابداری، دانشگاه علامه طباطبایی، تهران، ایران

- Email: *taghavifard@atu.ac.ir*
- ORCID: 0000-0002-4212-2079
- Homepage: https://aris.atu.ac.ir/taghavifard

بابک سهرابی، استاد گروه مدیریت فناوری اطلاعات، دانشکدهٔ مدیریت، دانشگاه تهران، تهران، ایران

- Email: *bsohrabi@ut.ac.ir*
- ORCID: 0000-0001-6188-2607
- Homepage: https://profile.ut.ac.ir/~bsohrabi

مرتضی امیرحسینی، دانشجوی دکتری فناوری اطلاعات، گروه مدیریت عملیات و فناوری اطلاعات، دانشکدهٔ مدیریت و حسابداری، دانشگاه علامه طباطبایی، تهران، ایران

- Email: *m.amirhosseini@atu.ac.ir*
- ORCID: 0000-0001-5813-8027
- Homepage: https://sma.atu.ac.ir/fa

HOW TO CITE THIS ARTICLE

Raeesi Vanani, I.; Taghavifard, M.; Sohrabi, M.; Amirhosseini, M., (2023). Designing an intelligent evaluation system for predicting fire insurance claims using deep learning. Iran. J. Insur. Res., 12(4): 251-264.

DOI: 10.22056/ijir.2023.04.01

URL: https://ijir.irc.ac.ir/article_160304.html?lang=en