Iran. J. Insur. Res., 12(4): 265-282, 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

Predicting term life insurance surrender using deep neural networks

A. Khandan[1,*], L. Niakan[2], Z. Fakharinezhad[3]

[1] *Department of Economics of Public Affairs, Faculty of Economics, Kharazmi University, Tehran, Iran*

[2] *Department of General Studies of Insurance, Insurance Research Center, Tehran, Iran*

[3] *Department of Life Insurance, Dey Insurance Company, Tehran, Iran*

ARTICLE INFO

Article History:
Received 23 January 2023
Revised 18 April 2023
Accepted 14 May 2023

Keywords:
Neural network
Prediction
Surrender
Term life insurance

*Corresponding Author:
Email: *Khandan.abbas@khu.ac.ir*
Phone: +9821 22522762
ORCID: 0000-0002-4558-6653

ABSTRACT

BACKGROUND AND OBJECTIVES: Life insurance has a very low adoption rate in Iran, mainly due to policy surrender. This research aims to analyze the individual characteristics and insurance contract features that influence the surrendering of term life insurance policies.

METHODS: The study utilizes a pilot database of 35,171 policy-holders and pensioners registered by an Iranian insurance company in 2021. Data mining, deep learning, and neural network algorithms are used for analysis due to their high accuracy in prediction:

FINDINGS: The model demonstrates desirable performance based on evaluation metrics with a 74 percent accuracy in predicting both types of surrendered and non-surrendered insurance policies. The model performs better in predicting non-surrendered insurance policies more attention is given to interpreting those results. Despite imbalanced data, the model still performs well. In the dataset, surrendered policies make up only 3 percent of the total, leading to bias towards predicting the majority class. Nonetheless, the model accurately predicts and categorizes most surrendered policies, covering 59 percent of the total 244 cases.

CONCLUSION: The results indicate that certain demographic characteristics, such as age, female gender, health surcharge, and accident risk rate, as well as specific contract characteristics, including policy term, time since start date, longer premium payment methods, higher annual increase in capital and premium, fewer covered risks, and lower benefits, are negatively correlated with policy surrender. Furthermore, the results suggest that if the insured person is the policy surrender themselves, the probability of surrender is minimized. On the other hand, if the insured person is someone else, especially distant relatives, the probability of surrender increases.

DOI: 10.22056/ijir.2023.04.02

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

نشریه علمی پژوهشنامه بیمه

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

مقاله علمی

پیش‌بینی بازخرید بیمه‌نامه‌های زندگی به شرط فوت با استفاده از شبکه‌های عصبی عمیق

عباس خندان[۱،*]، لیلی نیاکان[۲]، زهرا فخاری‌نژاد[۳]

[۱] گروه اقتصاد امور عمومی، دانشکده اقتصاد، دانشگاه خوارزمی، تهران، ایران

[۲] گروه مطالعات عمومی بیمه، پژوهشکدهٔ بیمهٔ ایران، تهران، ایران

[۳] گروه بیمه‌های زندگی، شرکت بیمه دی، تهران، ایران

اطلاعات مقاله

تاریخ های مقاله:

تاریخ دریافت: ۰۳ بهمن ۱۴۰۱

تاریخ داوری: ۲۹ فروردین ۱۴۰۲

تاریخ پذیرش: ۲۴ اردیبهشت ۱۴۰۲

کلمات کلیدی:

بازخرید

بیمهٔ عمر به شرط فوت

پیش‌بینی

مدل شبکهٔ عصبی

*نویسنده مسئول:

ایمیل: khandan.abbas@khu.ac.ir

تلفن: ۲۲۵۲۲۷۶۲ ۹۸۲۱+

ORCID: 0000-0002-4558-6653

چکیده:

پیشینه و اهداف: ضریب نفوذ بیمهٔ عمر به‌عنوان یک محصول مهم بیمه‌ای و برنامه‌ریزی مالی در ایران بسیار پایین است و یکی از دلایل آن بازخرید بیمه‌نامه‌هاست. هدف این مقاله بررسی تأثیر مشخصه‌های فردی و قراردادی بیمه‌نامه‌هاست که بر بازخرید بیمه‌نامه‌های عمر به شرط فوت اثر می‌گذارند.

روش‌شناسی: برای این منظور از داده‌های آماری و اطلاعات ثبتی ۳۵۱۷۱ خریدار بیمه‌نامه‌های عمر و مستمری یک شرکت بیمه‌ای در مقطع سال ۱۴۰۰ به‌عنوان پایلوت استفاده شد. برای تجزیه‌وتحلیل نیز از داده‌کاوی و الگوریتم‌های یادگیری عمیق و شبکهٔ عصبی استفاده شد که دقت بسیار بالایی در پیش‌بینی دارند.

یافته‌ها: مدل از دقت مطلوب ۷۴ درصد در پیش‌بینی هر دو نوع بیمه‌نامه‌های عدم بازخرید و بازخریدشده برخوردار است. البته عملکرد در پیش‌بینی عدم بازخرید بیمه‌نامه‌ها بسیار بهتر بوده، اما چون موضوع اصلی مقاله پیش‌بینی بیمه‌نامه‌های بازخریدشده است، در تفسیر نتایج بیشتر به آن توجه شد. نتایج به‌دست‌آمده با وجود مشکل نامتوازن بودن داده‌ها مطلوب است. در داده‌های مورد بررسی نسبت بیمه‌نامه‌های بازخریدی به عدم بازخرید ۳ به ۱۰۰ است که این عدم توازن موجب می‌شود فرایند یادگیری به سمت پیش‌بینی طبقه با بیشترین فراوانی سوگیری پیدا کند. باوجوداین، شاخص پوشش ۵۹ درصدی به‌دست‌آمده نشان داد که از مجموع ۲۴۴ بیمه‌نامهٔ بازخریدشده در مجموعه دادهٔ تست، شبکه توانسته اغلب آن‌ها، یعنی ۱۴۵ مورد را به‌درستی در طبقهٔ بیمه‌نامه‌های بازخریدی پیش‌بینی و طبقه‌بندی کند.

نتیجه‌گیری: نشان داده شد که از مشخصه‌های جمعیت‌شناختی متغیرهای سن، جنسیت زن، اضافه نرخ پزشکی، نرخ خطر حادثی و از مشخصه‌های قرارداد نیز مدت بیمه‌نامه، زمان سپری‌شده از شروع بیمه‌نامه، شیوهٔ پرداخت حق‌بیمه با اقساط بلندمدت‌تر، بالاتر بودن ضرایب افزایش سالانهٔ سرمایه و حق‌بیمه و کمتر بودن تعداد موارد پوشش و سرمایهٔ فوت با بازخرید اثر عکس دارند و احتمال آن را کاهش می‌دهند. با بازخرید بیمه‌نامه به‌صورت عکس مرتبط‌اند. نسبت بیمه‌گذار و بیمه‌شده نیز تأثیرگذار است و نشان داده شد که وقتی بیمه‌گذار بیمه‌نامهٔ عمر را برای خود بخرد بازخرید در حداقل است و با دور شدن نسبت خویشاوندی احتمال بازخرید افزایش می‌یابد.

DOI: 10.22056/ijir.2023.04.02

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

مقدمه

بیمهٔ زندگی اگرچه ریشه‌های تاریخی طولانی دارد، اما هنوز هم به‌عنوان یک محصول مهم بیمه‌ای و برنامه‌ریزی مالی نفوذ کمی در بازار دارد. در ایالات متحده فقط ۵۰/۵ درصد از مصرف‌کنندگان ۲۵ تا ۶۴ ساله بیمهٔ زندگی انفرادی دارند و ۴۸ درصد از جمعیت بزرگسال معتقدند به بیمهٔ زندگی نیاز دارند (Life Insurance Marketing and Research Association, 2019).

طبق گزارش سالنامهٔ آماری بیمهٔ مرکزی (۱۴۰۰)، متوسط جهانی ضریب نفوذ بیمه‌های زندگی در سال ۲۰۲۱ برابر ۳ درصد بوده که از متوسط جهانی ضریب نفوذ بیمه‌های غیرزندگی ۳/۹ در همان سال کمتر است. البته مشکل ضریب نفوذ پایین بیمهٔ زندگی در ایران پررنگ‌تر است، به‌گونه‌ای که بر اساس همان گزارش، ضریب نفوذ بیمه‌های زندگی در ایران در سال ۱۴۰۰ فقط ۰/۲۸ درصد بوده که هم نسبت به متوسط جهانی ضریب نفوذ بیمه‌های زندگی و هم در مقایسه با ضریب نفوذ بیمه‌های غیرزندگی در ایران که برای آن سال ۱/۶ درصد بوده، بسیار پایین‌تر است.

دلایل اقتصادی، ساختاری، اجتماعی، فرهنگی و آموزشی متعددی را می‌توان برای عدم توسعهٔ بیمه‌های زندگی در ایران و موانع پیش روی آن برشمرد و مطالعات مختلفی هستند که به بررسی چرایی پایین بودن تقاضا برای محصولات بیمه‌های زندگی پرداخته‌اند. باوجوداین، این مسئله روی دیگری هم دارد و آن بازخرید محصولات بیمه‌های زندگی است. بازخرید بیمه‌نامه به‌طور کلی به فسخ بیمه‌نامه از سوی بیمه‌گذار گفته می‌شود. توجه به نرخ بازخرید بیمه‌نامه‌ها از آن جهت مهم است که در شرایط پایین بودن تقاضای کل برای بیمه‌های زندگی و رقابتی بودن بازارها، این حفظ و نگهداشت مشتریان است که برای شرکت‌ها سودآوری دارد. بازخرید بیمه‌نامه‌ها خود می‌تواند دلایل مختلفی داشته باشد و شناخت دلایل آن وابسته به تحلیل آمار و اطلاعات مشتریان و داده‌کاوی است. تجزیه‌وتحلیل اطلاعات و شناخت الگوی‌های رفتاری مشتریان به شرکت‌ها کمک خواهد کرد تا محصولات خود را با نیاز و خواستهٔ مشتریان منطبق سازند و سازوکارهایی در جهت افزایش رضایت مشتریان و سودآوری بیشتر ایجاد کنند.

درک الگوهای رفتاری از حجم انبوه داده‌ها امروزه با استفاده از روش‌های مبتنی بر هوش مصنوعی و یادگیری ماشینی امکان‌پذیر شده است. این پژوهش قصد دارد تا با استفاده از الگوریتم‌های یادگیری عمیق و شبکهٔ عصبی به پیش‌بینی بازخرید بیمه‌نامه‌های زندگی بپردازد و تأثیر مشخصه‌های فردی و ویژگی‌های بیمه‌نامه‌ها بر بازخرید آن‌ها را تعیین کند. در همین راستا، از اطلاعات مشتریان بیمه‌نامه‌های یکی از محصولات بیمهٔ زندگی در یکی از شرکت‌های بیمه برای سال ۱۴۰۰ استفاده شده است. فرض کلی مطالعه این است که داده‌ها و اطلاعات مشتریان این شرکت به کل بازار بیمه‌های زندگی تعمیم‌پذیر است.

ساختار مقاله در ادامه به این شکل است که ابتدا در بخش دوم مقاله، به ادبیات نظری این حوزه و مرور پیشینهٔ پژوهش پرداخته خواهد شد. بخش سوم به ارائهٔ آمارهای توصیفی و معرفی متغیرهای پژوهش اختصاص دارد. بخش چهارم به معرفی الگوریتم یادگیری عمیق و فرایند پیاده‌سازی شبکه‌های عصبی می‌پردازد. در بخش پنجم یافته‌های پژوهش و در پایان، در بخش ششم نتیجه‌گیری و توصیه‌های سیاستی ارائه خواهد شد.

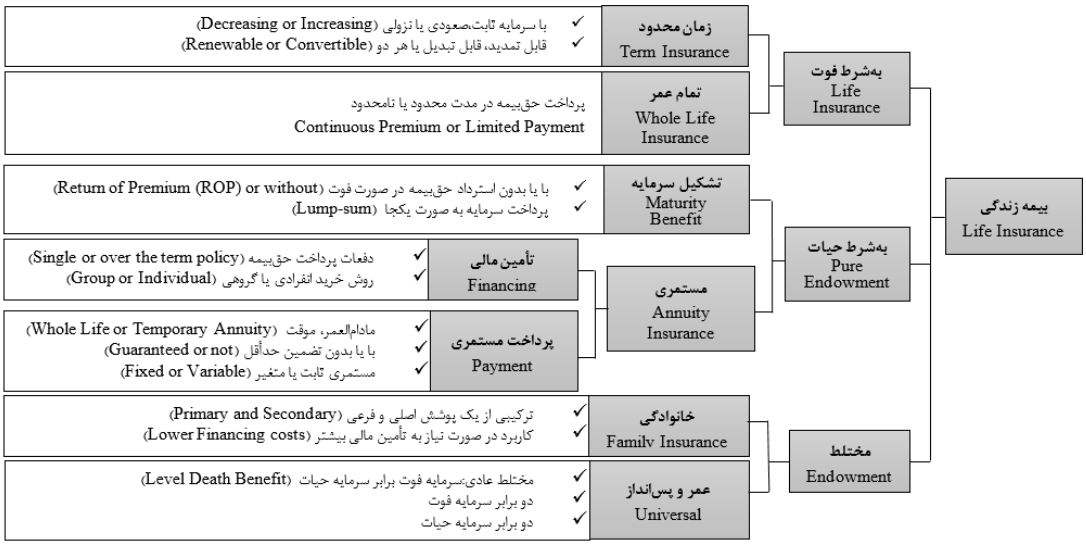
مبانی نظری پژوهش

در این قسمت به‌طور خلاصه ابتدا به معرفی بیمه‌های زندگی و بازار بیمه‌های زندگی در ایران پرداخته خواهد شد. سپس ادبیات نظری این حوزه و مرور پیشینهٔ پژوهش ارائه خواهد شد که درک بهتری از عوامل اثرگذار بر تقاضا و بازخرید بیمه‌های زندگی به‌دست خواهند داد.

بیمه‌های زندگی در ایران

به‌طور کلی بیمه‌های زندگی به سه دستهٔ کلی ۱) به شرط فوت، ۲) به شرط حیات و ۳) مختلط (Endowment Insurance) تقسیم می‌شوند. دو مورد اول نقش پوشش ریسک دارند، یعنی بیمه‌گر در ازای دریافت حق‌بیمه ریسک مشخصی (ریسک فوت یا ریسک طولانی‌شدن عمر) را پوشش می‌دهد. بیمه‌های زندگی مختلط علاوه‌بر کارکرد بیمه‌ای بیمه‌نامه‌های به شرط فوت یا به شرط حیات، نقش پس‌اندازی و سرمایه‌گذاری نیز دارند. تمامی موارد بالا ممکن است مدت‌دار باشند یا اینکه مدت بیمه‌نامه نامحدود باشد که در این صورت به آن بیمهٔ زندگی تمام‌عمر گفته می‌شود. شکل ۱ انواع اصلی بیمه‌های زندگی را نمایش داده است.

در ایران نظارت بر صنعت بیمهٔ زندگی بر عهدهٔ بیمهٔ مرکزی جمهوری اسلامی ایران است و در این خصوص شورای‌عالی بیمه آیین‌نامه‌هایی را تصویب کرده که آیین‌نامهٔ شمارهٔ ۶۸ و اصلاحات بعدی آن از مهم‌ترین آن‌هاست. شرکت‌های بیمه علاوه‌بر محصولات بیمه‌ای دیگر، محصولات بیمهٔ زندگی مختلفی را نیز ارائه می‌دهند و در ایران تنها دو شرکت بیمه خاورمیانه و باران هستند که به‌صورت تخصصی فقط محصولات بیمهٔ زندگی را در پورتفوی خود دارند. در دههٔ گذشته بازار بیمه‌های زندگی رشد فزاینده‌ای داشتند، به‌گونه‌ای که سهم بیمه‌های زندگی از کل حق‌بیمهٔ تولیدی از ۸ درصد در سال ۱۳۹۰ به ۱۵/۷۹ درصد در سال ۱۴۰۰ رسیده است (Bimeh Markazi Centeral Insurance of Iran, 2021). باوجوداین همچنان ضریب نفوذ بیمه‌های زندگی در ایران بسیار پایین است. طبق گزارش (Bimeh Markazi Centeral Insurance of Iran 2021) مجموع حق‌بیمهٔ تولیدی بیمه‌های زندگی در سال ۱۴۰۰ برابر ۱۸۲۱۷ میلیارد تومان بوده که در مقایسه با هزینهٔ ناخالص داخلی جاری (گزارش بانک مرکزی از حساب‌های ملی فصلی ایران) آن سال ۶،۶۷۷،۴۵۰ میلیارد تومان، از ضریب نفوذی حدود ۰/۲۸ درصد حکایت دارد. این در حالی است که ضریب نفوذ بیمه در سال ۱۳۹۹ به‌طور کلی در ایران برابر ۱/۹ و متوسط جهانی آن ۷ درصد بوده است. میانگین ضریب نفوذ بیمه‌های زندگی در جهان نیز در این سال برابر ۳ بوده است.



شکل ۱: انواع کلاسیک بیمه‌های زندگی
Fig. 1: The classic types of life insurance products
(Khandan, 2022)

Deacon and Firebaugh (1988) رفتار افراد در بازخرید بیمه‌نامه‌ها را نوعی تقاضای منفی تلقی می‌کنند و تلاش می‌کنند تقاضای بیمهٔ زندگی اعم از خرید یا بازخرید را که یکی از مهم‌ترین جنبه‌ها و تصمیمات مالی خانوار است، در یک چارچوب تکاملی تبیین کنند. این چارچوب رویکردی سیستمی برای درک رفتارهای فردی و خانوار است که عوامل مختلف تأثیرگذار بر رفتارهای مدیریتی منابع خانوار را طبق شکل ۲ در یکی از سه سیستم فردی، سیستم خانوادگی و سیستم محیطی طبقه‌بندی می‌کند. در سیستم فردی عوامل بسیاری هستند از جمله تقاضای بیرونی (ارزش‌های خانوادگی، اهداف خانوادگی، مطالبات و تعهدات خانوار، هنجارهای اجتماعی، مطالبات اجتماعی، رویدادهای اجتماعی)، تقاضای داخلی (اهداف شخصی)، منابع بیرونی (حمایت‌های خانواده، حمایت‌های اجتماعی) و منابع داخلی (قابلیت‌های شخصی، کیفیت‌های شخصی، تجربیات زندگی و روابط فردی) که بر رفتار فرد در برنامه‌ریزی تأثیر می‌گذارند. علاوه‌براین، ویژگی‌های شناختی، عاطفی، اجتماعی و فیزیکی افراد که ذاتی و اکتسابی هستند نیز در تصمیم‌گیری فرد نقش دارند. بعد از تصمیم‌گیری، فرد می‌تواند راهبردهای متنوعی اتخاذ و با دریافت بازخورد آن‌ها را اصلاح کند. خروجی تصمیم‌گیری سطح فردی در نظام خانوادگی تحت تأثیر ارتباطات بین‌فردی و منابع انسانی و مادی درون خانواده قرار می‌گیرد. دو نظام محیطی خرد (محیط سکونت یا اشتغال، همسایگان و دوستان) و کلان (فرهنگ، سیاست، اقتصاد، فنّاوری) نیز وجود دارند که بر تصمیمات نهایی خانواده و فرد اثر می‌گذارند.

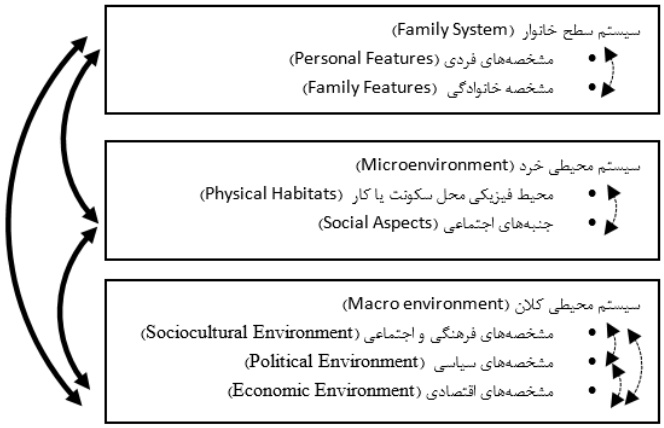
Deacon and Firebaugh (1988) معتقدند هر دو نظام محیطی کلان و خرد وابسته به یکدیگرند و همچنین هر دو با نظام خانوادگی در تعامل‌اند و در نهایت بر تصمیمات مالی خانوار از جمله خرید بیمه‌نامه‌های زندگی تأثیر می‌گذارند. به دلیل تعامل بین نظام‌ها، خانوار بازخوردهای مختلفی دریافت می‌کند و بر اساس این بازخوردها ممکن است راهبردهای مدیریتی خود را حفظ یا تغییر دهد که به معنی بازخرید است.

مروری بر پیشینهٔ پژوهش

شروع مطالعات در زمینهٔ تقاضای بیمهٔ عمر را می‌توان به Yaari (1965) نسبت داد که بر اساس فرض حداکثرسازی مطلوبیت انتظاری سرپرست خانوار در یک مدل پیوستهٔ زمانی، مسئلهٔ طول عمر نامطمئن و تقاضای بیمهٔ عمر را بررسی کرد. از آن پس مطالعات متعدد داخلی و خارجی به بررسی تقاضای بیمهٔ زندگی، خرید و بازخرید و تأثیر عوامل مختلف بر آن پرداخته‌اند. در ادامه جدیدترین مطالعاتی که در پیش‌بینی بازخرید بیمه‌های زندگی بر مشخصه‌های فردی مشتریان تمرکز داشته‌اند، مرور خواهند شد.

Ghorbani *et al.* (2022) با استفاده از روش‌های داده‌کاوی، از جمله جنگل تصادفی، درخت تصمیم، رگرسیون لجستیک و شبکهٔ عصبی به طبقه‌بندی مشتریان بیمه‌های زندگی برحسب ریزش یا عدم ریزش می‌پردازند. در این پژوهش از اطلاعات بیمه‌نامه‌های زندگی یک شرکت بیمهٔ پایلوت در سال ۱۳۹۸ برای استان تهران استفاده شده است و نتایج به‌دست‌آمده نشان می‌دهند که احتمال بازخرید بیمه‌نامه‌های عمر در سنین بالاتر، در میان زنان و افراد دارای مشاغل پرریسک بیشتر است. از بررسی مشخصه‌های قراردادی نیز یافته‌های به‌دست‌آمده حاکی از آن است که در میان بیمه‌نامه‌های با اقساط سالانه، حق‌بیمه کمتر و درصد ضریب تغییر سرمایه بیشتر، بازخرید به احتمال کمتری اتفاق افتاده است.

Bakhtiar Nasrabadi *et al.* (2020A) نشان دادند رفتار بازخرید بیمهٔ زندگی برآیندی از کنش و برهم‌کنش‌های شناختی،



شکل ۲: نظام‌های محیطی و سطح خانوار مؤثر بر تقاضای بیمه‌های زندگی

Fig. 2: The environmental and family level systems determining the demand for life insurance (Deacon and Firebaugh, 1988)

روان‌شناختی و اجتماعی بیمه‌گذار در زمینه‌ای از شرایط خرد (فردی) و کلان با مضمون تردید (عدم قطعیت) است. این مطالعه با رویکرد کیفی و انجام مصاحبه‌های نیمه‌ساختاریافته از ۲۹ فرد (بیمه‌گذار، نمایندگان، فروشندگان و کارشناسان)، در نهایت نتیجه می‌گیرد که ماهیت ناملموس و بلندمدت بیمهٔ زندگی، شرایط محیطی بی‌ثبات اقتصادی (به‌ویژه تورم، نرخ بهره)، تردید در ارتباط با کارآمدی محصول و ایفای تعهدات بیمه‌گر از مهم‌ترین دلایلی است که فرد را به بازخرید بیمه‌نامه (بازخرید یا فسخ به دلیل عدم پرداخت) وامی‌دارد. آن‌ها معتقدند ارتباطات و تعاملات بیمه‌گر در این مرحله و در ارزیابی و تغییر نگرش بیمه‌گذار نقش مهمی دارد و فقدان آن می‌تواند منتج به احساس رهاشدگی، بی‌توجهی، نارضایتی و ... شود و رفتارهای احساسی و تکانشی نیز به دلایل گفته‌شده به بازخرید بیمه‌نامه منجر شود. نسخه‌ای مشابه همین مطالعه توسط Bakhtiar Nasrabadi *et al.* (2020B) در ارتباط با پیش‌بینی خرید بیمه‌نامهٔ عمر به چاپ رسیده است.

در مطالعه‌ای مشابه، Helmzadeh *et al.* (2020) نیز معتقدند که تاکنون بیشتر تحقیقات صورت‌گرفته در زمینهٔ بازخرید بیمه‌های عمر بر عوامل اقتصادی متمرکز بوده‌اند و به همین دلیل، هدف اصلی خود را نگرش جامع و تبیین همه‌جانبهٔ علل بازخرید بیمه‌نامه‌های عمر در ایران معرفی می‌کنند. در این مطالعه نیز از روش کیفی و انجام مصاحبه‌های نیمه‌ساختاریافته با کارکنان عالی‌رتبه و کارشناسان بازخرید بیمه‌های عمر برخی شرکت‌های بیمهٔ فعال در کشور استفاده شده است. این مطالعه نتیجه می‌گیرد که شرایط کلان مانند بی‌ثباتی اقتصادی یا سیاسی و شرایط خرد مانند وضعیت شغلی و درآمد بیمه‌گذار در کنار ویژگی‌های ماهوی بیمهٔ عمر ازجمله پیچیدگی، بلندمدت و ناملموس بودن در بازخرید آن تأثیرگذارند؛ اما آنچه در نهایت موجب بازخرید می‌شود عدم تعامل، عدم ارائهٔ مشاورهٔ مجدد و عدم ارائهٔ تسهیلات است که احساس نارضایتی و سلب اعتماد از بیمه‌گر را بر می‌انگیزانند.

Hozarmoghaddam *et al.* (2020) به بررسی تأثیر شیوع ویروس کرونا بر بیمه‌های زندگی پرداخته‌اند. در این مقاله با توجه به افزایش نرخ مرگ‌ومیر و مواجه شدن افراد با ریسک فوت چنین استدلال می‌شود که احتمالاً بیمه‌نامه‌های عمر به شرط فوت با افزایش فروش مواجه شود. در مقابل، همچنین ممکن است که به دلیل کاهش اشتغال و درآمد، افزایش تورم و کاهش ارزش سرمایهٔ فوت، از بین رفتن شبکهٔ فروش سنتی و نمایندگی‌ها و ... فروش بیمه‌نامه‌ها کاهش یافته و بنابراین اثر نهایی همه‌گیری ویروس کرونا بر فروش بیمه‌نامه‌های عمر مشخص نیست. اما این مقاله به‌طور مشخص با ارائهٔ آمار و ارقام نشان می‌دهد که بازخرید بیمه‌نامه‌های عمر در دوران همه‌گیری کرونا با افزایش روبه‌رو بوده است. از جمله مهم‌ترین دلایل برای افزایش بازخرید نیز می‌توان به عدم آگاهی و ابهام در پرداخت سرمایهٔ فوت به دلیل ابتلا به کرونا، بیکاری و کاهش درآمد خانوارها و افزایش نیاز خانوارها با نقدینگی اشاره کرد.

Bash Afshar *et al.* (2018) با جمع‌آوری اطلاعات ۱۰۰۰ نفر از خریداران بیمه‌های زندگی یک شرکت بیمه در سال ۱۳۹۲ به بررسی عوامل مؤثر بر نگهداشت مشتریان، و الگوی خوشه‌بندی برای ارائهٔ خدمات به آنان می‌پردازند. با استفاده از متغیرهای مربوط به اطلاعات بیمه‌نامه، اطلاعات جمعیتی و تکمیلی بیمه‌شدگان مانند سوابق بیماری، در این مطالعه عوامل مؤثر بر ریزش مشتری شناسایی و یک الگوی خوشه‌بندی برای طبقه‌بندی مشتریان در دو گروه اصلی مشتریان سودبخش و مشتریان دارای ریسک پیشنهاد می‌شود. طبق نتایج به‌دست‌آمده، متغیرهای جمعیتی همچون «جنسیت» و «سن» و متغیرهای بیمه‌ای همچون «حق‌بیمهٔ سالیانه» و «ضریب فوت در اثر حادثه» از جمله مهم‌ترین عوامل تأثیرگذار در شناسایی گروه‌های مشتریان هستند.

Sazgar and Abed (2018) با استفاده از داده‌های ۱۰ سالهٔ بیمهٔ عمر و پس‌انداز یکی از شرکت‌های بیمهٔ فعال کشور به بررسی رفتار بازخرید مشتریان می‌پردازند. مدل برازش‌شده احتمال بازخریدی را ۸/۳ درصد برآورد می‌کند و طبق نتایج به‌دست‌آمده

افرادی که پوشش فوت در اثر حادثه را خریداری کرده‌اند، تمایل کمتری به بازخرید داشته‌اند. پوشش معافیت از پرداخت حق‌بیمه، مدت بیمه‌نامه‌ها، نحوهٔ پرداخت حق‌بیمه به‌صورت ماهیانه، میزان حق‌بیمهٔ پرداختی و سال‌های صدور نیز از عوامل مؤثر بر بازخرید این نوع بیمه‌ها بوده‌اند. Habibi Marand (2016) در پایان‌نامهٔ کارشناسی ارشد خود با مطالعهٔ موردی ۳۶۸ بیمه‌نامهٔ بازخریدشده در بیمهٔ پارسیان مشهد به بررسی عوامل مؤثر بر بازخرید بیمهٔ عمر و سرمایه‌گذاری می‌پردازد. نتایج این پژوهش که با مطالعات اقتصادسنجی انجام شده، نشان داد که متغیرهای سن، درآمد، تأهل و بُعد خانوار بر بازخرید بیمهٔ عمر و سرمایه‌گذاری تأثیر معناداری دارند، اما تأثیر معناداری از تحصیلات مشاهده نشد.

Mahdavi *et al.* (2015) با استفاده از اطلاعات و داده‌های ترکیبی بیمه‌نامه‌های دو شرکت طی سال‌های ۱۳۸۷ تا ۱۳۹۲ و روش رگرسیون لجستیک به پیش‌بینی بازخرید بیمه‌های عمر و بررسی عوامل مؤثر بر آن پرداخته‌اند. طبق نتایج به‌دست‌آمده از مدل با صحت ۹۷/۵ درصد، احتمال بازخریدی در حدود ۰/۰۴۲ درصد برآورد شد. طبق نتایج به‌دست‌آمده سن، اضافه نرخ پزشکی (نرخی که به دلایل بیماری به نرخ حق‌بیمه اضافه می‌شود)، سرمایهٔ اولیهٔ فوت و میزان حق‌بیمه بر بازخریدی بیمه‌نامه تأثیر مثبت و در مقابل تحصیلات، وضعیت تأهل، مدت بیمه‌نامه و تعداد اقساط تأثیر منفی بر احتمال بازخریدی بیمهٔ عمر داشته‌اند. طبق نتایج به‌دست‌آمده، مردان مجرد با اضافه نرخ پزشکی بالا احتمال بازخرید بالاتری دارند که این احتمال با افزایش حق‌بیمه و سرمایهٔ فوت و در مقابل کاهش مدت بیمه و کاهش تعداد اقساط بیشتر نیز می‌شود.

Abbasi and Sazgar (2005) به بررسی علل بازخرید بیمه‌نامه‌های عمر و پس‌انداز در شرکت سهامی بیمهٔ ایران پرداخته‌اند. این مطالعه از طریق توزیع پرسشنامه بین ۶۰ نفر از کارشناسان بیمهٔ عمر انجام و نظرات آن‌ها در خصوص علل بازخرید مشتریان جمع‌آوری شد. نتایج و یافته‌های این پژوهش نشان می‌دهند که احتمال بازخرید بیمه‌نامه‌ها برای افراد در مشاغل کم‌خطر (طبقات شغلی اول و دوم از طبقه‌بندی پنجگانهٔ مشاغل بر اساس خطرات شغلی)، و افرادی که در بیمه‌نامهٔ خود شیوهٔ پرداخت حق‌بیمهٔ ماهانه و سرمایهٔ فوت کمتر را انتخاب کرده‌اند، بیشتر است.

از مطالعات خارجی در زمینهٔ بازخرید بیمه‌نامه‌های عمر و عوامل مؤثر نیز می‌توان به چند مورد از جدیدترین مطالعات اشاره کرد. Azzone *et al.* (2022) با استفاده از الگوریتم جنگل تصادفی و رگرسیون لجستیک به پیش‌بینی مدت‌زمان سپری‌شده تا بازخرید بیمه‌نامه‌های عمر می‌پردازند. نتایج به‌دست‌آمده نشان می‌دهند که عوامل غیراقتصادی (مدت‌زمان مانده تا سررسید، شرکت بیمه و رویکرد بازاریابی آن) در تعیین مدت‌زمان سپری‌شده تا بازخرید نقش مهمی دارند و در مقابل عوامل اقتصادی و مالی بی‌تأثیر بوده‌اند (البته به غیر از نرخ رشد درآمد که تأثیر زیادی داشت). این مطالعه همچنین نشان می‌دهد که مدل‌های خطی مانند رگرسیون لجستیک در مدل‌سازی تصمیمات مالی پیچیده مانند بازخرید بیمه‌نامه‌ها عملکرد مناسبی ندارند. Hu *et al.* (2021) به بررسی این موضوع می‌پردازد که آیا تلفیق داده‌های مناطق جغرافیایی مختلف یک شرکت بیمه در ایرلند برای پیش‌بینی مدت‌زمان سپری‌شده تا بازخرید (توقف پرداخت حق‌بیمه) بیمه‌نامه‌های عمر این شرکت مفید است؟ برای این منظور از یک روش ترکیبی استفاده شد که در آن مدل‌های معمول آماری در خوشه‌های فضایی نشان‌دهندهٔ ساختار فضایی و جغرافیایی بیمه‌شدگان ایرلندی به کار گرفته شده‌اند.

Heo (2020) خود با اتخاذ رویکرد سیستمی و روش‌های یادگیری ماشین و شبکه‌های عصبی مصنوعی به بررسی عوامل مؤثر بر تقاضای بیمهٔ زندگی در ایالات متحده می‌پردازد. یافته‌ها و نتایج به‌دست‌آمده در این مطالعه نشان داد که متغیرهای تأثیرگذار مهم مرتبط با تقاضای بیمهٔ زندگی به هر چهار نظام فردی، خانوادگی، محیطی خرد و محیطی کلان تعلق دارند. این مطالعه همچنین از مقایسهٔ تأثیر متغیرها در روش شبکهٔ عصبی و روش‌های اقتصادسنجی نتیجه می‌گیرد که رویکرد سیستمی و روش یادگیری ماشینی با در نظر گرفتن همهٔ این روابط مستقیم و غیرمستقیم از دقت بالاتری در پیش‌بینی رفتار خریداران و تقاضای بیمهٔ زندگی برخوردار است. Zietz (2003) نیز با همین رویکرد به بررسی عوامل مؤثر بر تقاضای بیمه‌های زندگی می‌پردازد و تناقض و تفاوت‌ها در نتایج برآوردی مطالعات قبلی از تأثیر عوامل مؤثر را ناشی از به‌کارگیری سلیقه‌ای این عوامل در مدل‌های اقتصاد سنجی دانسته و بر به‌کارگیری روش‌های یادگیری ماشین به جهت پیش‌بینی تأکید دارد.

Milhaud *et al.* (2011) با استفاده از مدل‌های درخت تصمیم، رگرسیون لجستیک و مدل‌های طبقه‌بندی به پیش‌بینی احتمال بازخرید بیمه‌نامه‌های عمر و عوامل مؤثر بر آن می‌پردازند. جامعهٔ آماری ۲۸۵۰۶ بیمه‌نامه یک شرکت اسپانیایی از ۱۹۹۹ تا ۲۰۰۷ را دربرگرفته که به‌صورت مقطعی برای سال ۲۰۰۷ بررسی شده‌اند. نتایج به‌دست‌آمده نشان می‌دهند که احتمال بازخرید در بیمه‌نامه‌های با مدت کمتر، شیوهٔ پرداخت با اقساط ماهانه و دوماهه، سرمایهٔ فوت بیشتر، اضافه نرخ پزشکی کمتر و برای افراد در سنین بالاتر بیشتر است.

Dash (2018) با در دست داشتن یک نمونهٔ ۴۰۰ نفری از خریداران به بررسی عوامل مؤثر بر تقاضای بیمهٔ زندگی در هند می‌پردازد و طیف وسیعی از متغیرها را مطالعه می‌کند. نتایج به‌دست‌آمده نشان دادند که درآمد افراد، اشتغال آن‌ها، تحصیلات و سن بر خرید بیمهٔ زندگی اثر مثبتی دارند و در مقابل اثر نرخ حق‌بیمه یا قیمت این بیمه‌ها منفی ارزیابی شد. همچنین نشانه‌ای از تأثیر دیگر متغیرهای جنسیت، تأهل و بار تکفل مشاهده نشد. Sulaiman *et al.* (2015) با استفاده از داده‌های سری زمانی ۱۹۷۹ تا ۲۰۰۸ به بررسی عوامل مؤثر بر تقاضای بیمهٔ عمر در کشور اتیوپی پرداخته‌اند. در این مقاله پنج عامل تأثیرگذار اقتصادی شامل تولید ناخالص داخلی، پس‌انداز، توسعهٔ مالی، نرخ بهره، و تورم و شش عامل تأثیرگذار جمعیتی شامل جمعیت، نسبت وابستگی‌جوان، نسبت وابستگی‌پیری، قیمت بیمه، نرخ مرگ‌ومیر و امید به زندگی بررسی می‌شوند. نتایج به‌دست‌آمده نشان می‌دهد که تورم در این میان

جدول ۱: فراوانی بیمه‌نامه‌های عمر باز و بسته در سال ۱۴۰۰
Table 1: The statistical ferequency of open and closed policies in 2021

انواع بیمه‌نامه Policies	عدم بازخرید Non-Surrendered				بسته در اثر بازخرید Surrendered
	باز Open	بسته در اثر فوت Death	بسته در اثر سررسید Lapsed	بسته در اثر عدم پرداخت Unpaid Premium	
تعداد Number	33782	15	7	2626	839
درصد Percentage	90.6	0.04	0.02	7.05	2.25

دارای اثر قابل توجه و منفی بر تقاضای بیمهٔ عمر است. علاوه‌براین، نتایج از یک اثر منفی نسبت وابستگی جوانان بر تقاضای بیمهٔ عمر حکایت دارد. سایر متغیرها اثر معنی‌داری نداشته‌اند.

Balaji and Serivatsa (2012) از تکنیک طبقه‌بندی بیزین ساده (Naïve Bayesian) به‌منظور پیش‌بینی و دسته‌بندی مشتریان بیمه‌های زندگی استفاده می‌کند. برای این منظور نیز از اطلاعات ۱۰ هزار مشتری شرکت بیمهٔ زندگی هند (IRDA) برای دورهٔ زمانی ژانویه تا دسامبر ۲۰۱۱ استفاده شده است. این مقاله برای بررسی عوامل مؤثر بر حفظ و نگهداشت مشتریان به بررسی تأثیر هفت متغیر مانند سن، جنسیت، تأهل، تعداد فرزندان، نوع محصول، طبقهٔ دریافت خدمات (فرعی یا اصلی)، نوع طرح بیمه (پس‌اندازی/ پوشش ریسک/ بازنشستگی/ فرزندان) پرداختند. نتایج به‌دست‌آمده حاکی از آن بود که روش بیزین ساده قابلیت طبقه‌بندی را به‌خوبی درخت تصمیم و شبکهٔ عصبی دارد.

روش‌شناسی پژوهش

جامعهٔ آماری مورد پژوهش در این مطالعه شامل اطلاعات ۳۷۲۶۹ بیمه‌نامهٔ عمر و مستمری یکی از شرکت‌های بیمه کشور برای سال ۱۴۰۰ است. این افراد کسانی هستند که از فروردین ۱۳۹۵ این بیمه‌نامه را خریده و تا سال ۱۴۰۰ همچنان مشتری این بیمه‌نامه شرکت باقی مانده‌اند. به‌عبارت‌دیگر، این جامعهٔ آماری به‌دلیل محدودیت داده، اطلاعات افرادی که قبل از ۱۳۹۵ این بیمه‌نامه را خریداری کرده‌اند، ندارد و همچنین فقط افرادی را دربرمی‌گیرد که در سال ۱۴۰۰ قرارداد باز و فعال داشته‌اند.

این اطلاعات از دفتر بیمه‌های زندگی شرکت بیمهٔ پایلوت گرفته شده است. داده‌های دریافتی اطلاعات مختلفی از مشخصه‌های افراد و نوع قرارداد آن‌ها را دربرمی‌گیرد. یکی از اطلاعات دریافتی تاریخ و نوع الحاقیه‌ای است که در خصوص بعضی بیمه‌نامه‌ها وجود داشته است. بیمه‌نامه‌های غیرابطالی یا (۱) از طرف بیمه‌گر فسخ شده‌اند که عمدتاً در اثر عدم پرداخت حق بیمه بوده، یا (۲) بیمه‌نامه در اثر فوت بیمه‌شدهٔ اصلی بسته شده، یا (۳) بیمه‌نامه در اثر سررسید بسته شده، یا (۴) بیمه‌نامه از طرف بیمه‌گذار فسخ شده که به آن بازخرید گفته می‌شود، و در نهایت یا (۵) بیمه‌نامهٔ باز (بدون تغییر یا با الحاقیهٔ تغییر در پارامترهای حق‌بیمه، نحوهٔ پرداخت، ذی‌نفع، انواع موارد پوشش، تغییر نشانی و ...) مانده است. از میان این پنج دسته، این پژوهش به‌دنبال توضیح و پیش‌بینی بیمه‌نامه‌هایی است که از طرف بیمه‌گذار فسخ یا بازخرید شده است. جدول ۱ اطلاعات مربوط به این متغیر هدف یا وابسته را نمایش می‌دهد. به‌این‌ترتیب، در این پژوهش متغیر هدف دودویی تعریف شده، به این صورت که بیمه‌نامه‌های بازخریدی مقدار ۱ و عدم بازخرید (سایر مشاهدات اعم از باز، بسته در اثر فوت یا سررسید و بسته در اثر عدم پرداخت) مقدار ۰ به خود می‌گیرند.

متغیرها و اطلاعات مختلفی را که برای توضیح و پیش‌بینی متغیر هدف به کار گرفته می‌شود می‌توان در دو دسته اطلاعات جمعیت‌شناختی و اجتماعی بیمه‌شدگان و یا اطلاعات قراردادی بیمه‌نامه طبقه‌بندی کرد. در مجموعه داده‌های بیمه‌شدگان، اطلاعات جمعیت‌شناختی و اجتماعی اقتصادی متعددی وجود دارند که عبارت‌اند از:

- سن بیمه‌شده
- جنسیت بیمه‌شده: متغیر دودویی که برای مردان مقدار ۱ و برای زنان مقدار ۰ می‌گیرد.
- نسبت بیمه‌شده با بیمه‌گذار: این مشخصه را می‌توان در قالب دو متغیر بررسی کرد. نخست یک متغیر دودویی با نام «یکسان بودن بیمه‌شده و بیمه‌گذار» که در مواردی که بیمه‌شده همان بیمه‌گذار است مقدار ۱ و در مواردی که بیمه‌شده متفاوت با بیمه‌گذار است مقدار ۰ می‌گیرد. در حالت دوم، به نزدیکی رابطهٔ بیمه‌شده با بیمه‌گذار نیز توجه شده و یک متغیر ترتیبی با نام «نسبت با بیمه‌گذار» تعریف خواهد شد که وقتی بیمه‌شده خود بیمه‌گذار باشد مقدار ۱، بیمه‌شده جزء خانواده بیمه‌گذار باشد (همسر یا فرزند) مقدار ۲، بیمه‌شده جزء خانوادهٔ نسبی بیمه‌گذار باشد (والدین یا بردار و خواهر) مقدار ۳، و در صورت داشتن نسبت خویشاوندی دورتر هم مقدار ۴ به خود می‌گیرد.
- طبقهٔ خطر شغلی: اطلاعات شغلی ثبت‌شده از بیمه‌شدگان متأسفانه طبقه‌بندی مشخص و قابلیت استفاده ندارد. باوجوداین، شرکت بیمه یک طبقه‌بندی از مشاغل تحت عنوان «طبقهٔ خطر شغلی» ثبت می‌کند. این متغیر در ۵ دسته (۱= کم خطر تا ۵= بسیار پرخطر)، ریسک و مخاطرات مشاغل مختلف را طبقه‌بندی می‌کند.
- اضافه نرخ پزشکی: وضعیت سلامت بیمه‌شده بر اساس اطلاعات دریافتی از افراد و نظر کارشناس شرکت ارزیابی شده و به‌صورت

اضافه نرخ پزشکی ثبت و در محاسبهٔ حق‌بیمه اعمال می‌شود.

● مجموع نرخ خطر حادثی: کارشناسان شرکت بیمه همچنین بر اساس اطلاعات دریافتی از افراد یک نرخ خطر وقوع حادثه برای بیمه‌شده ارزیابی و در حق‌بیمه اعمال می‌کنند.

● استان محل سکونت: از دیگر اطلاعات ثبت‌شده و قابل استفاده در مدل، استان محل سکونت افراد است. برای هر استان یک متغیر دودویی ساخته خواهد شد که در صورت سکونت فرد بیمه‌شده در آن استان مقدار ۱ و در غیر این صورت مقدار ۰ به خود می‌گیرد.

متأسفانه اطلاعات بیشتری در مورد درآمد افراد، شغل سایر اعضای خانوار، تحصیلات، تأهل و ... در بیمه‌نامه وجود ندارد و بر این اساس تحلیل‌های این پژوهش با این محدودیت‌ها روبه‌رو است. علاوه‌بر مشخصه‌های گفته‌شده، هر قرارداد و بیمه‌نامه‌ای مشخصه‌هایی دارد که توسط افراد هنگام خریداری انتخاب می‌شوند. این مشخصه‌ها و اطلاعات قراردادی ازآنجاکه نشان‌دهندهٔ انتخاب افراد هستند، می‌تواند با رفتار آن‌ها در تقاضای بیمه‌نامه یا بازخرید آن مرتبط باشند و به این منظور برای پیش‌بینی و توضیح متغیر هدف استفاده خواهند شد. این متغیرها عبارت‌اند از:

● مدت بیمه‌نامه: مدت بیمه‌نامه‌ای که افراد هنگام خرید انتخاب می‌کنند، می‌تواند بیانگر چشم‌انداز آن‌ها از خرید بیمه‌نامه یا برنامه‌ریزی مالی باشد.

● مدت‌زمان سپری‌شده از شروع بیمه‌نامه: اختلاف بین تاریخ بازخرید و تاریخ شروع، نشان‌دهندهٔ مدت سپری‌شده از بیمه‌نامه است که طبیعتاً با احتمال بازخرید مرتبط است. در بیمه‌نامه‌های باز نیز مدت سپری شده تا پایان اسفند ۱۴۰۰ که مقطع مورد بررسی است لحاظ خواهد شد.

● تعداد موارد پوشش بیمه‌نامه: افراد در زمان خرید می‌توانند علاوه‌بر پوشش ریسک فوت، متقاضی پوشش‌های اضافه‌تر، از جمله پوشش فوت ناشی از حادثه، پوشش نقص عضو یا ازکارافتادگی ناشی از حادثه، پوشش هزینه‌های پزشکی ناشی از حادثه، و پوشش امراض خاص نیز باشند. این متغیر حداقل مقدار ۱ دارد (چون همهٔ بیمه‌نامه‌ها ریسک فوت به هر علت را پوشش می‌دهند) و حداکثر برابر ۵ است.

● کل سرمایه و سرمایهٔ فوت: در هریک از موارد پوشش‌داده‌شده در بیمه‌نامه، مبلغی به‌عنوان سرمایه مشخص شده که در صورت وقوع یا بروز آن حالت به ذی‌نفع پرداخت می‌شود. مجموع سرمایه‌های ثبت‌شده برای پوشش‌های مختلف تحت عنوان متغیر کل سرمایه در پیش‌بینی استفاده خواهد شد. باوجوداین، به دلیل جایگاه اصلی پوشش فوت در بیمه‌نامهٔ عمر، سرمایهٔ فوت بیمه‌نامه نیز به‌صورت جداگانه در توضیح مدل استفاده خواهد شد. هر دو متغیر به‌صورت حقیقی‌شده به قیمت‌های سال ۱۳۹۵ درآمده و برای این منظور از شاخص کل سالانهٔ قیمت مصرف‌کنندهٔ مرکز آمار ایران استفاده شده است.

● شیوهٔ پرداخت حق‌بیمه: بر اساس موارد پوشش‌داده‌شده، سرمایه در هر مورد پوشش، مدت بیمه‌نامه و نرخ‌های حادثه و پزشکی و ... یک میزان حق‌بیمه برای افراد تعیین می‌شود. حق‌بیمه می‌تواند در تعداد اقساط مختلفی پرداخت شود و این متغیر در صورت انتخاب شیوهٔ پرداخت یکجا مقدار ۰، ماهانه مقدار ۱، دوماهه مقدار ۲، سه‌ماهه مقدار ۳، شش‌ماهه مقدار ۴ و سالانه مقدار ۵ می‌گیرد. ترتیب این متغیر نشان‌دهندهٔ این است که فرد تمایل دارد اقساط پرداختی را به زمانی دورتر بیندازد.

● ضریب افزایش سالانهٔ سرمایه و حق‌بیمه: افراد می‌توانند برای سرمایهٔ فوت و میزان حق‌بیمه ضریب افزایش سالانه‌ای در نظر گیرند که نشان‌دهندهٔ تورم انتظاری در ذهن آن‌هاست.

در این پژوهش از مبلغ حق‌بیمه به‌عنوان متغیر مستقل استفاده نشد، چون میزان حق‌بیمه متغیر مستقل نیست و بر اساس دیگر متغیرهای انواع پوشش، سرمایه، مدت بیمه‌نامه، شیوهٔ پرداخت، نرخ‌های پزشکی و حادثه، طبقهٔ خطر شغلی و ضرایب افزایش سالانهٔ سرمایه و حق‌بیمه مشخص می‌شود و وقتی آن متغیرها در مدل وارد شده، دیگر گنجاندن میزان حق‌بیمه بی‌معنی خواهد بود. به‌عبارت‌دیگر و در اصطلاح اقتصادسنجی، حق‌بیمه تابعی ریاضی از سایر متغیرهاست و گنجاندن آن در مدل در کنار سایر متغیرها مشکل هم‌خطی ایجاد می‌کند. جدول ۲ آمارهایی مقدماتی از داده‌های مورد استفاده را نمایش می‌دهد.

معرفی و پیاده‌سازی شبکهٔ عصبی مصنوعی

گام نخست در داده‌کاوی، آماده‌سازی داده‌ها شامل سه مرحلهٔ پیش‌پردازش داده‌ها، انتخاب مؤثرترین ویژگی‌ها و هم‌مقیاس‌سازی است. پیش‌پردازش خود شامل چهار اقدام کنترل داده‌های طبقه‌بندی‌شده، کنترل داده‌های گم‌شده، کنترل داده‌های پرت و در آخر کنترل داده‌های تکراری می‌شود. آن‌چنان‌که در قسمت قبل گفته شد، متغیرهای مورد بررسی همه به‌صورت عددی (پیوسته، دودویی یا ترتیبی) تعریف شده‌اند. در خصوص برخی متغیرهای کیفی مثل استان محل سکونت از روش وان‌هات و برای نسبت بیمه‌شده با بیمه‌گذار و شیوهٔ پرداخت از روش ترتیبی استفاده شد. در روش وان‌هات به‌طور مثال برای هر استان یک متغیر دودویی جداگانه ساخته شده و در نهایت همهٔ متغیرهای دودویی مربوط به استان‌ها، به‌جز استان تهران (به‌عنوان پایه مقایسه) وارد مدل می‌شود. این روش برای متغیرهای کیفی که فاقد ترتیب باشند مناسب است. برای متغیرهایی کیفی که در آن‌ها ترتیب مشخصی وجود دارد بهتر است از کدگذاری ترتیبی استفاده شود تا مدل بتواند از اطلاعات ترتیبی آن‌ها برای پیش‌بینی استفاده کند، وگرنه این اطلاعات از دست می‌رود. به‌طور مثال در شیوهٔ پرداخت حق‌بیمه می‌توان آن‌ها را برحسب تعداد اقساط مرتب کرد یا در متغیر نسبت بیمه‌گذار می‌توان آن‌ها را به ترتیب خود فرد تا خانوادهٔ درجهٔ اول و دورتر مرتب کرد تا مدل از این اطلاعات استفاده کند.

تعداد انگشت‌شمار مقادیر گم‌شده نیز با میانهٔ متغیرها جایگزین شدند. با رصد مشاهدات تعداد ۱۵۹۸ مورد تکراری شناخته شد که حذف گردید. برای شناسایی مقادیر پرت نیز از بازهٔ میانگین به‌علاوه و منهای ۵ انحراف معیار استفاده شد که در نهایت با حذف آن‌ها ۳۵۱۷۱ مشاهده باقی ماند. همهٔ متغیرها نیز در نهایت استاندارد شده

جدول ۲: آمار توصیفی داده‌های متغیرهای پژوهش
Table 2: Descriptive statistics of data and research variables

حداکثر Max	چارک ۳ Q3	چارک ۲ Q2	چارک ۱ Q1	حداقل Min	انحراف معیار St. deviation	میانگین Mean	نوع متغیر Variable Type		
1	0	0	0	0	0.148	0.0225	دودویی (۰ و ۱) Binary	بازخرید بیمه‌نامه Surrenderd Policies	متغیر هدف Dependent Variable
66	35	24	7	0	16.02	22.68	پیوسته Continous	سن (Age)	متغیر مستقل Indipendent Variables
1	1	1	0	0	0.496	0.56	دودویی (۰ و ۱) Binary	جنسیت (مرد) Sex (Men)	
1	1	0	0	0	0.5	0.48	دودویی (۰ و ۱) Binary	یکسان بودن بیمه‌شده و بیمه‌گذار Insurrer is the Insured	
4	2	2	1	1	0.84	1.689	ترتیبی (۱ تا ۴) Ordinal (1 to 4)	نسبت بیمه‌شده با بیمه‌گذار Relation between Insured and the Insurer	
5	2	2	2	1	0.64	2.01	ترتیبی (۱ تا ۵) Ordinal (1 to 5)	طبقهٔ خطر شغلی Job Risk Class	
1.8	0	0	0	0	0.109	0.039	پیوسته Continous	اضافه نرخ پزشکی Health Extra Premium	
1.15	0	0	0	0	0.043	0.004	پیوسته Continous	نرخ خطر حادثی Accident Rate	
30	30	30	20	5	6.54	24.55	پیوسته (سال) Continous (Years)	مدت بیمه‌نامه Term of Insurance	
6.5	4	2.1	1.1	0	1.6	2.4	پیوسته (سال) Continous (Years)	مدت سپری‌شده Elapsed Time	
5	5	5	2	1	1.65	3.685	ترتیبی (۱ تا ۵) Ordinal (1 to 5)	تعداد موارد پوشش Number of Coverage Cases	
639.1	41.82	21.03	9.22	0.252	42.4	34.7	پیوسته Continous	کل سرمایه (میلیون تومان) Total Benefit	
176.7	12.29	7.67	4.61	0.252	8.41	9.64	پیوسته Continous	سرمایهٔ فوت (میلیون تومان) Death Benefit	
5	4	1	1	0	1.69	2.51	ترتیبی (۰ تا ۵) Ordinal (0 to 5)	شیوهٔ پرداخت حق‌بیمه Payment Method	
20	5	5	5	0	4.06	5.67	ترتیبی (۰ تا ۲۰) Ordinal (0 to 20)	ضریب افزایش سالانه سرمایه Increasing Benefits (Annual rate)	
20	15	10	10	0	5.34	11.67	ترتیبی (۰ تا ۲۰) Ordinal (0 to 20)	ضریب افزایش سالانهٔ حق‌بیمه Increasing Premiums (Annual rate)	
1	0	0	0	0	0.1	0.137	دودویی (۰ و ۱) Binary	استان محل سکونت (تهران) Province (Tehran)	

و به شکل توزیع‌های با میانگین صفر و واریانس ۱ استفاده خواهند شد. در یادگیری عمیق و شبکهٔ عصبی مصنوعی، چند لایه انتخاب ویژگی توسط شبکه به‌صورت خودکار انجام می‌شود. هم‌مقیاس‌سازی ویژگی‌ها نیز یک مرحله از آماده‌سازی است که توصیه می‌شود هنگام کار با بسیاری از الگوریتم‌های یادگیری ماشین انجام شود تا نتایج تحت تأثیر مقیاس و واحد اندازه‌گیری متغیرها قرار نگیرد. در نهایت، مجموعه داده‌ها به دو دسته ۲۴۶۲۰ دادهٔ آموزش و ۱۰۵۵۱ دادهٔ آزمایش تقسیم شد. داده‌های آموزش به‌منظور یادگیری مدل به کار گرفته خواهند شد و داده‌های آزمایش که در مراحل قبل توسط مدل دیده نشده، در پایان برای ارزیابی عملکرد مدل استفاده می‌شود.

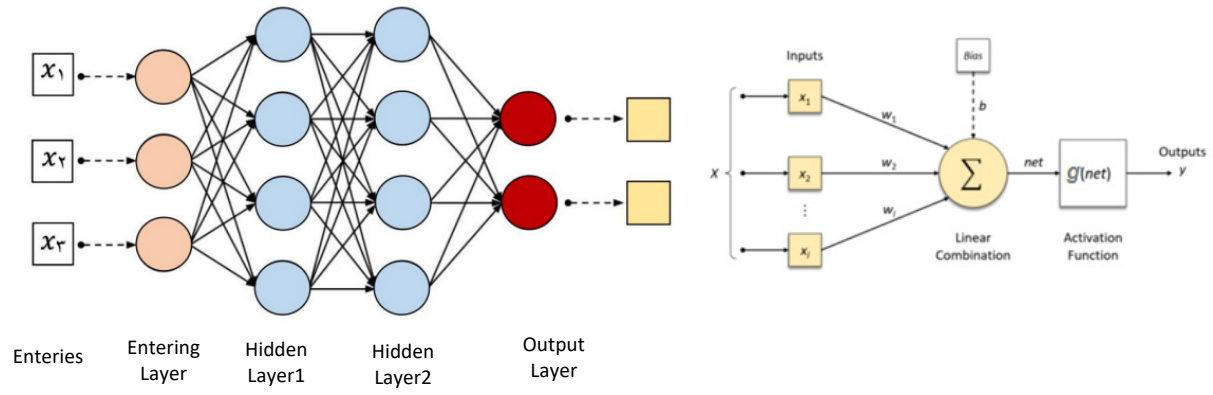
پیاده‌سازی الگوریتم یادگیری عمیق و مدل شبکهٔ عصبی

شبکه‌های عصبی مصنوعی، مدل‌های محاسباتی هستند که سازوکار یادگیری در آن‌ها مشابه مغز انسان پایه‌ریزی شده است و به‌واسطهٔ همین شباهت مدل مفیدی برای طبقه‌بندی داده‌ها و پیش‌بینی الگوی آینده تلقی می‌شود (Berson *et al*. 1999, Kudyba and Kwatinetz, 2014). در این شبکه هر لایه از کنار هم قرار دادن چندین نورون تشکیل شده است. اطلاعات توسط لایهٔ ورودی دریافت، در لایه‌های پنهان پردازش و در لایهٔ خروجی نتایج نشان داده می‌شود. به الگوی اتصال بین نورون‌ها، تعداد نورون‌ها و تعداد لایه‌ها در اصطلاح معماری شبکهٔ عصبی گفته می‌شود. ساده‌ترین ساختار شبکهٔ عصبی تک‌لایه یا پرسپترون است که فقط از یک لایهٔ ورودی و خروجی تشکیل شده است. این ساختار در قسمت سمت راست شکل ۳ نمایش داده شده است.

بردار X ورودی‌ها، محرک‌ها یا همان داده‌های آموزشی هستند. مقادیر ورودی و وزن آن‌ها که با بردار وزن W نشان داده شده مشخص است. ترکیبی خطی از مقادیر ورودی با استفاده از وزن آن‌ها محاسبه شده که خالص اطلاعات (net) نامیده می‌شود. خالص اطلاعات به یک تابع فعال‌ساز g داده می‌شود که در نهایت خروجی یا نتیجه را به‌دست می‌دهد. در شبکه‌های پیچیده‌تر با چند لایهٔ پنهان، خروجی هر لایه در واقع ورودی لایهٔ بالاتر است و به لایهٔ دیگر منتقل می‌شوند تا در نهایت به لایهٔ خروجی برسند. قسمت سمت چپ شکل ۳ این شبکهٔ عصبی عمیق را نمایش می‌دهد. واژهٔ عمیق نیز به داشتن چندین لایه اشاره دارد و در این صورت در اصطلاح یادگیری عمیق نامیده می‌شود.

می‌توان نتایج را با داده‌های آموزشی با برچسب‌های معین مقایسه کرد و هرگونه خطای پیش‌بینی را با محاسبهٔ تابع زیان به شبکه بازخورد داد. شبکه با دریافت بازخوردها وزن‌های به‌کارگرفته‌شده در محاسبات را اصلاح می‌کند تا در تکرارهای بعدی خطا در پیش‌بینی کمتر و تابع زیان به حداقل برسد. به روش تعیین و اصلاح وزن‌ها، الگوریتم یادگیری گفته می‌شود؛ در این فرایند یادگیری، شبکه به‌تدریج می‌آموزد که کدام ورودی‌ها اهمیت و وزن بیشتری در پیش‌بینی و نتیجه‌گیری دارد. عملکرد شبکه به ابرپارامترهایی چون تعداد لایه‌ها، تعداد نورون هر لایه، فرم تابع فعال‌ساز، فرم تابع زیان، الگوریتم و نرخ یادگیری بستگی دارد که تنظیم مقدار مناسب آن‌ها معمولاً با سعی و خطا و به روش اعتبارسنجی متقاطع بر اساس معیارهای ارزیابی مانند ماتریس اغتشاش، معیارهای دقت، صحت و پوشش و امتیاز-F صورت می‌گیرد.

ماتریس اغتشاش یا سردرگمی شامل تعداد عناصری است که به‌درستی (قطر اصلی) یا نادرستی (خارج از قطر اصلی) برای هر کلاس طبقه‌بندی شده‌اند. دقت یکی از معیارهای ارزیابی مدل است که درصد مشاهداتی که به‌درستی طبقه‌بندی شده‌اند (مجموع مقادیر قطر اصلی ماتریس اغتشاش) از کل مشاهدات را اندازه‌گیری می‌کند. معیار پوشش بر مشاهدات واقعی هر کلاس (یک سطر از ماتریس اغتشاش) تمرکز می‌کند و درصد مشاهداتی را می‌سنجد که درست طبقه‌بندی شده‌اند. معیار صحت بر مشاهداتی تمرکز می‌کند که توسط مدل در یک کلاس طبقه‌بندی شده (یک ستون از ماتریس اغتشاش) و درصد مشاهداتی را می‌سنجد که با واقعیت منطبق‌اند. معیار امتیاز-F نیز میانگین هارمونیک دو معیار صحت و پوشش است. برای درک بهتر فرض کنید دو کلاس مثبت (+) و منفی (-) داریم و قرار است مدل مشاهدات را در این دو کلاس طبقه‌بندی و پیش‌بینی کند. ماتریس اغتشاش این مثال در جدول ۳ و فرمول معیارهای یادشده نیز در رابطهٔ (۱) مشخص شده است.



شکل ۳: نمایی از یک پرسپترون ساده و عمیق
Fig. 3: Illustration of simple and deep perceptrons (Vazan, 1992)

جدول ۳: ماتریس اغتشاش
Table 3: Confusion Matrix

	پیش‌بینی مقدار مثبت (+) Predicted Positive	پیش‌بینی مقدار منفی(-) Predicted Negative
مقدار واقعی مثبت (+) Actual Positive	مثبت واقعی (TP) True Positive	منفی کاذب (FN) False Negative
مقدار واقعی منفی (-) Actual Negative	مثبت کاذب (FP) False Positive	منفی واقعی (TN) True Negative

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$
$$Precision = \frac{TP}{TP+FP}$$
$$Recall = \frac{TP}{TP+FN}$$
$$F1 = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (۱)$$

ابرپارامترهای شبکه با سعی و خطا و به روش اعتبارسنجی متقاطع بر اساس معیارهای ارزیابی گفته‌شده تنظیم و در نهایت، مدل نهایی شبکهٔ عصبی با مشخصات جدول ۴ انتخاب شد.

همان‌طور که در اطلاعات جدول نیز می‌توان دید، از یک مدل شبکهٔ عصبی با ۳ لایهٔ پنهان استفاده خواهد شد. تعداد نورون‌ها در این لایه‌های پنهان به‌ترتیب ۳۲، ۱۶ و ۸ نورون است که مقادیری بین تعداد نورون لایهٔ ورودی (تعداد مشخصه‌ها ۴۵) و تعداد نورون لایهٔ خروجی (یک خروجی ۱) گرفته‌اند. تعداد لایه‌ها و تعداد نورون‌ها با اعتبارسنجی متقاطع مشخص شده‌اند. در لایه‌های پنهان از یکسوساز خطی به‌عنوان تابع فعال‌ساز استفاده شده که سرعت همگرایی بالایی دارد و در تابع خروجی نیز تابع فعال‌ساز سیگموید به کار رفته تا مقدار خروجی را به‌صورت مقادیر احتمال بین ۰ و ۱ به‌دست بدهد. برای شبکه یک تابع زیان به شکل آنتروپی متقاطع دودویی وزن‌دهی‌شده در نظر گرفته شد. این تابع مقادیر پیش‌بینی‌شده از بازخرید یا عدم بازخرید بیمه‌نامه‌ها را با مقادیر واقعی مقایسه و پس از محاسبهٔ خطا برای آن یک هزینه یا زیان مشخص می‌کند. به دو دلیل در تابع زیان وزن بالاتری به بیمه‌نامه‌های بازخریدشده داده شد. نخست به دلیل تمرکز موضوعی پژوهش بر بیمه‌نامه‌های بازخریدشده و اینکه می‌خواهیم مدل در پیش‌بینی این دسته از بیمه‌نامه‌ها عملکرد مناسبی داشته باشد. دلیل دوم تلاش برای رفع مشکل توزیع نامتوازن داده‌هاست. در داده‌های مورد بررسی نسبت بیمه‌نامه‌های بازخریدی به عدم بازخرید ۳ به ۱۰۰ است که این عدم توازن موجب می‌شود فرایند یادگیری به سمت پیش‌بینی طبقه با بیشترین فراوانی سوگیری پیدا کند و مدل در پیش‌بینی طبقهٔ اقلیت که اتفاقاً موضوع اصلی این پژوهش است عملکرد موفقی نداشته باشد. برای رفع این مشکل، در تابع زیان وزن بالاتری به بیمه‌نامه‌های بازخریدشده داده شد تا الگوریتم در یادگیری و شناسایی عوامل مؤثر بر آن بیشتر توجه کند.

در شبکهٔ عصبی مقادیر ورودی با وزن‌های مشخص‌شده در هر لایه پس از گذر از توابع فعال‌ساز به یک خروجی تبدیل می‌شود. به‌منظور یادگیری الگوریتم، مقدار خطا و زیان محاسبه‌شده به شبکه بازخورد داده می‌شود تا وزن‌ها و تأثیر ورودی‌ها با هدف حداقل‌سازی خطا و زیان اصلاح شوند. فرایند یادگیری و اصلاح وزن‌ها مبتنی بر گرادیان کاهشی و با نرخ یادگیری تطبیقی به روش آدام (ADAM) انجام خواهد شد.

نتایج و بحث

با تنظیم این پارامترها در شبکه، مدل مورد نظر در محیط پایتون کدنویسی و اجرا شد که دربارهٔ نتایج به‌دست‌آمده در این قسمت بحث خواهد شد. نخست عملکرد مدل در یادگیری با داده‌های آموزش؛ طبق نتایج به‌دست‌آمده می‌توان دید که فرایند یادگیری شبکه مطلوب بوده و مدل توانسته طی دوره‌های تکرار به مقادیر دقت بالای حدود ۹۰ درصد برسد. باوجوداین، آنچه اهمیت دارد عملکرد مدل در پیش‌بینی با مجموعه داده‌های آزمایش است که تاکنون آن‌ها را ندیده و در حافظه خود ندارد.

نتایج پیش‌بینی شبکهٔ عصبی برآزش‌شده با مجموعه داده‌های آزمایش نیز در جدول ۵ نمایش داده شده است. به‌منظور درک و ارزیابی بهتر همچنین می‌توان عملکرد مدل در پیش‌بینی را با مدل رگرسیون لجستیک مقایسه کرد. رگرسیون لجستیک در واقع همان شبکهٔ عصبی است با این تفاوت که هیچ لایهٔ پنهانی نداشته و فقط از دو لایهٔ ورودی و خروجی تشکیل شده است. نتایج به‌دست‌آمده از پیش‌بینی با استفاده از رگرسیون لجستیک در جدول ۶ نمایش داده شده است. همان‌طور که می‌توان دید، مدل نهایی مورد استفاده در این مطالعه در همهٔ شاخص‌های ارزیابی عملکرد بهتری نسبت به رگرسیون لجستیک داشته است.

شاخص دقت ۸۳ درصدی نشان می‌دهد که مدل به‌طور کل و در پیش‌بینی هر دو گروه از بیمه‌نامه‌ها و مجموعه داده‌ها عملکرد مناسبی داشته است. البته عملکرد مدل در پیش‌بینی بیمه‌نامه‌های عدم بازخرید به‌طور ویژه بسیار خوب بوده است. برای طبقهٔ عدم بازخرید، شاخص پوشش ۸۴ درصدی نشان می‌دهد که شبکه از مجموع ۱۰۳۶۶ بیمه‌نامهٔ عدم بازخرید توانسته به‌درستی ۸۶۲۰ مورد را پیش‌بینی و در این دسته طبقه‌بندی کند. البته در شرایط نامتوازن بودن داده‌ها نباید عملکرد مدل صرفاً براساس پیش‌بینی طبقهٔ اکثریت، به‌ویژه شاخص صحت و امتیاز-F ارزیابی شود، چون در این

جدول ۴: ابرپارامترهای مدل نهایی شبکهٔ عصبی
Table 4: Hyper-parameters of the final neural network model

لایه‌های پنهان Hidden Layers	3
تعداد نورون Neurons	تعداد نورون در لایهٔ ورودی برابر تعداد مشخصه‌ها (۴۵) Entering Layer Neurons (45) تعداد نورون در لایه‌های پنهان به‌ترتیب (۳۲ – ۱۶ – ۸) Hidden Layers Neurons (8-16-32) تعداد نورون در لایهٔ خروجی (۱) Output Layer Neurons (1)
تابع فعال‌ساز Activation Function	لایه‌های پنهان میانی: یکسوساز خطی Hidden Layers (Relu) لایه خروجی: سیگموید Output Layer (Sigmoid)
تابع زیان Loss Function	آنتروپی متقاطع دودویی وزن‌دهی‌شده Weighted Binary Cross Entropy وزن بیمه‌نامهٔ بازخرید به عدم بازخرید ۲۰۰ به ۱ Surrender to Non-surrendered Weight (200 to 1)
الگوریتم یادگیری Learning Algorithm	الگوریتم یادگیری پس‌انتشار با مقداردهی تصادفی وزن‌های اولیه Backpropagation with Random Initializing Weights
الگوریتم بهینه‌سازی Optimization Algorithm	گرادیان کاهشی با نرخ یادگیری تطبیقی به روش آدام (ADAM) Gradient Descent (ADAM)
اندازهٔ دسته‌های داده Size of Batches	32
تعداد دورهٔ تکرار Number of Epoches	150
اعتبارسنجی Validation	اعتبارسنجی متقاطع با لحاظ شاخص‌های مختلف Cross-Validation Using Various Criteria

جدول ۵: معیارهای اعتبارسنجی مدل نهایی
Table 5: Cross validation criteria of the final model

ماتریس اغتشاش Confusion Matrix				
	پیش‌بینی بازخرید (۱) Predicted as Surrender (1)	پیش‌بینی عدم بازخرید (۰) Predicted as Non-Surrendered (0)	دقت Accuracy	
بازخرید (۱) Surrender (1)	128	118	0.83	
عدم بازخرید (۰) Non-surrendered (0)	1686	8620		
	صحت Precision	پوشش Recall	F1	F2
پیش‌بینی بازخرید (۱) Prediction in Surrender Class	0.07	0.52	0.12	0.23
پیش‌بینی عدم بازخرید (۰) Prediction in Non-Surrendered Class	0.99	0.84	0.91	0.87

شاخص‌ها موارد پیش‌بینی درست یا اشتباه طبقهٔ اکثریت (سطر دوم) با مشاهدات اقلیت بیمه‌نامه‌های بازخریدشده (سطر اول) مقایسه می‌شود که طبیعتاً داده‌های بسیار کمتری دارد. لذا، در خصوص داده‌های نامتوازن تأکید می‌شود که باید بر طبقهٔ اقلیت، یعنی صرفاً موارد بازخریدی و به‌ویژه شاخص پوشش تمرکز شود. از آن جهت بر این نکته تأکید می‌شود که معمولاً این طبقهٔ اقلیت و در این پژوهش طبقهٔ بازخرید بیمه‌نامه‌هاست که هدف اصلی است. بر این اساس می‌توان دید که عملکرد مدل تاحدودی با توجه به مشکل نامتوازن بودن توزیع داده‌ها مطلوب است. در طبقهٔ اقلیت بازخرید، شاخص پوشش ۵۲ درصدی نشان می‌دهد که از مجموع ۲۴۶ بیمه‌نامهٔ

جدول ۶: معیارهای اعتبارسنجی مدل رگرسیون لجستیک
Table 6: Cross validation criteria of the logestic regression

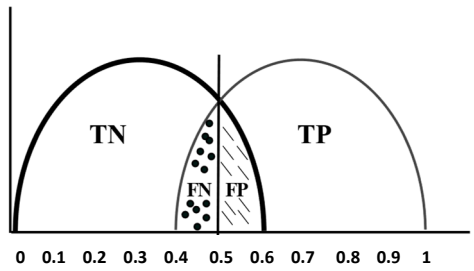
ماتریس اغتشاش Confusion Matrix				
	پیش‌بینی بازخرید (۱) Predicted as Surrender (1)	پیش‌بینی عدم بازخرید (۰) Predicted as Non-Surrendered (0)	دقت Accuracy	
بازخرید (۱) Surrender (1)	202	44	0.3	
عدم بازخرید (۰) Non-surrendered (0)	7394	2912		
	صحت Precision	پوشش Recall	F1	F2
پیش‌بینی بازخرید (۱) Prediction in Surrender Class	0.03	0.82	0.05	0.13
پیش‌بینی عدم بازخرید (۰) Prediction in Non-Surrendered Class	0.99	0.28	0.44	0.33
متوسط Average	0.51	0.55	0.24	0.23

بازخریدشده در مجموعهٔ داده‌های آزمایش، شبکه توانسته اغلب آن‌ها، یعنی ۱۲۸ مورد را به‌درستی در طبقهٔ بیمه‌نامه‌های بازخریدی پیش‌بینی و طبقه‌بندی کند. شاخص صحت چون آن را با سطر بالایی مشاهدات بیمه‌نامه‌های عدم بازخرید که تعداد مشاهدهٔ زیادی دارد مقایسه می‌کند، نمی‌تواند معیار خوبی در اینجا باشد و همچنین شاخص امتیاز-F که متوسط معیار صحت و پوشش است. در موارد داده‌های نامتوازن می‌توان به یک فرم خاص از شاخص یا امتیاز F تحت عنوان امتیاز Fbeta توجه کرد که طبق رابطهٔ (۲) محاسبه می‌شود. در شرایطی که هزینهٔ پیش‌بینی و طبقه‌بندی اشتباه موارد بازخریدی به‌عنوان عدم بازخرید (مثبت کاذب FP) بیشتر باشد، بهتر است ضریب بتا برابر ۲ قرار داده شود و شاخص تحت عنوان امتیاز F2 شناخته می‌شود. در شرایطی که هزینهٔ پیش‌بینی و طبقه‌بندی اشتباه بازخرید به‌عنوان عدم بازخرید و همچنین عدم بازخرید به‌عنوان بازخرید یکسان باشد، همچنان از شاخص F1 استفاده خواهد شد. هر دوی این حالت‌ها، یعنی شاخص F1 و F2 در جدول نمایش داده شده است.

$$Fbeta = \frac{(1+Beta^2)\times Precision\times Recall}{(Beta^2\times Precision)+Recall} => \quad (۲)$$
$$F2 = \frac{5\times Precision\times Recall}{(4\times Precision)+Recall}$$

برای داده‌های نامتوازن شاخص‌های دیگر و بیشتری نیز پیشنهاد شده است که به نظر مناسب می‌رسند. به‌طور مثال متوسط هندسی دو شاخص حساسیت (شاخص پوشش برای طبقهٔ اکثریت عدم بازخرید) و اختصاصی (شاخص پوشش برای طبقهٔ اقلیت بازخرید) که برای مدل برآورد شده برابر ۰/۶۶ به‌دست می‌آید. شاخص حساسیت و شاخص اختصاصی به‌ترتیب نرخ مثبت واقعی و نرخ منفی واقعی نیز شناخته می‌شوند. این دو شاخص برخلاف یکدیگرند. برای درک این موضوع به شکل ۴ نگاه کنید. مدل به هر بیمه‌نامه احتمال بازخریدی بین ۰ و ۱ اختصاص می‌دهد. منحنی نازک احتمال بازخرید، و منحنی تیره احتمال عدم بازخرید فرد را نشان می‌دهند. خروجی‌های نزدیک صفر قطعاً به معنی عدم بازخرید و خروجی‌های نزدیک یک بازخرید شدن بیمه‌نامه است، اما اگر عددی بین این دو، مثلاً ۰/۵، به‌دست آمد، با قطعیت نمی‌توان در مورد بازخریدی بیمه‌نامه نظر داد. حد آستانه‌ای باید تعیین شود که از آن آستانه به بالا، بیمه‌نامه‌ها را بازخرید و از آن آستانه به پایین بیمه‌نامه‌ها را عدم بازخرید طبقه‌بندی کرد. اگر عدد ۰/۵ به‌عنوان آستانه انتخاب شود، ناحیهٔ هاشورخرده بیانگر بیمه‌نامه‌هایی است که به‌اشتباه بازخرید و ناحیهٔ نقطه‌چین هم بیمه‌نامه‌هایی هستند که به‌اشتباه عدم بازخرید تشخیص داده خواهند شد. اگر حد آستانه پایین، مثلاً روی ۰/۴، آورده شود، تمام بیمه‌نامه‌های بازخریدی (منحنی نارک‌تر و روشن‌تر) را تشخیص خواهیم داد یا به‌عبارتی شاخص حساسیت یا پوشش بالا خواهد بود، اما میزان زیادی از بیمه‌نامه‌های عدم بازخرید را هم به‌اشتباه بازخرید اعلام خواهیم کرد که به معنی شاخص اختصاصی پایین است. این رابطه را به شیوهٔ دیگری نیز می‌توان بیان کرد. اگر حد آستانه پایین، مثلاً روی ۰/۴، آورده شود، تمام بیمه‌نامه‌های بازخریدی را تشخیص خواهیم داد یا به‌عبارتی نرخ مثبت واقعی بالا خواهد بود، اما میزان زیادی از بیمه‌نامه‌های عدم بازخرید را هم به‌اشتباه بازخرید اعلام خواهیم کرد و نرخ مثبت کاذب نیز بالا خواهد بود. نرخ مثبت کاذب در اینجا برابر (1-specificity) است.

اگر نرخ مثبت واقعی را در مقابل نرخ مثبت کاذب در یک نمودار و برای سطوح مختلف از حد آستانه رسم کنیم، به یک منحنی تحت عنوان ROC دست می‌یابیم که اتفاقاً یکی از بهترین شاخص‌های ارزیابی عملکرد طبقه‌بندی، به‌ویژه برای داده‌های نامتوازن است. منحنی ROC



شکل ۴: توزیع احتمال بازخرید یا عدم بازخرید بیمه‌نامه‌ها به همراه حد آستانه
Fig. 4: Distribution of surrender and non-surrender policies with their thresholds

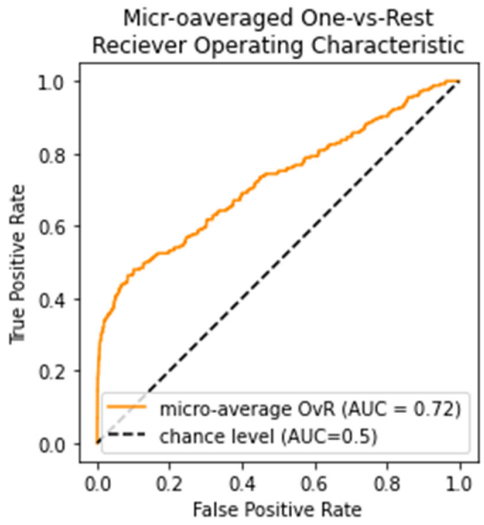
حاصل از طبقه‌بندی مدل نهایی شبکهٔ عصبی در شکل ۵ نشان داده شده است. بهترین عملکرد زمانی است که نرخ مثبت واقعی برابر ۱ و نرخ مثبت کاذب صفر باشد که به آن طبقه‌بندی کامل گفته می‌شود. در مقابل یک طبقه‌بندی کاملاً تصادفی به ما نرخ مثبت واقعی ۰/۵ و نرخ مثبت کاذب ۰/۵ به‌دست خواهد داد. این خط نشان‌دهندهٔ عدم تمایز همان خط نیمساز از گوشهٔ چپ و پایین به راست و بالاست که با خط‌چین نشان داده شده است. یک منحنی ROC که بالای این خط قرار بگیرد در طبقه‌بندی موفق است و هرچه به محور عمودی نزدیک‌تر باشد، عملکرد آن در طبقه‌بندی نیز بهتر بوده است. می‌توان همچنین از سطح زیر این منحنی AUC برای ارزیابی استفاده کرد. اگر مساحت زیر منحنی AUC بالاتر از ۰/۵ و به ۱ نزدیک‌تر باشد، عملکرد مدل در طبقه‌بندی بهتر است. شاخص AUC به‌دست‌آمده در این پژوهش همان‌طور که در شکل ۵ نشان داده شده، برابر ۰/۷۲ است که عملکردی تاحدودی مناسب از خود نشان می‌دهد.

با توجه به عملکرد مناسب شبکهٔ عصبی طراحی‌شده و نتایج جدول ۵، به نظر می‌رسد اکنون می‌توان به بررسی متغیرهای تأثیرگذار و عواملی که توسط مدل مؤثر شناخته شده پرداخت.

تفسیر نتایج و شناسایی عوامل اثرگذار بر بازخرید

روش شبکهٔ عصبی در دهه‌های اخیر طرفداران زیادی پیدا کرده، چون به دلیل قابلیت برآورد روابط غیرخطی عملکرد بهتری در پیش‌بینی و طبقه‌بندی دارند. اما شبکه‌های عصبی اگرچه در پیش‌بینی بسیار خوب عمل می‌کنند، در تفسیر و توضیح نتایج و روابط ضعیف‌ترند. بسیاری از جمله Azodi *et al.* (2020)، Guidotti *et al.* (2018)، Buhrmester *et al.* (2021)، Olden and Jackson (2002) و Liang *et al.* (2021) شبکهٔ عصبی را به یک «جعبهٔ سیاه» تشبیه کرده‌اند که پژوهشگر نمی‌داند در داخل آن چه اتفاقی می‌افتد. باوجوداین این مقاله برای شناسایی عوامل مؤثر بر بازخرید بیمه‌نامه‌ها، وزن‌های اثرگذاری متغیرها را که توسط شبکهٔ عصبی تعیین شده، بررسی خواهد کرد. وزن تأثیرگذاری متغیرها در جدول ۷ ارائه شده است.

نخست می‌توان به بررسی متغیرهای جمعیت‌شناختی و اجتماعی پرداخت. می‌توان دید که سن بیمه‌شده با بازخرید بیمه‌نامه رابطه عکس



شکل ۵: منحنی ROC و شاخص AUC حاصل از طبقه‌بندی مدل نهایی
Fig. 5: ROC curve and AUC of the final model classification

دارد. این تأثیر اگرچه صرفاً یک همبستگی است و از جنس کنترل سایر متغیرها به مانند تحلیل‌های آماری نیست، اما به‌خوبی با انتظارات منطبق است. افزایش سن ارتباط نزدیکی با ریسک فوت بیمه‌شده و سایر ریسک‌ها از جمله ازکارافتادگی و ... دارد. همچنین می‌توان دید که بین مرد بودن (جنسیت) و بازخرید بیمهٔ عمر رابطهٔ مستقیمی وجود دارد. در مجموعهٔ داده‌ها حدود نیمی از بیمه‌شدگان جزء خانواده و خویشاوندان بیمه‌گذار بوده‌اند. از این جهت می‌توان گفت که بیمه‌گذاران نیاز و ضرورت کمتری برای خرید بیمه‌نامهٔ عمر برای اعضای خانواده و خویشاوندان مرد خود می‌دیده‌اند. استان محل سکونت نیز یک متغیر اثرگذار بود. نتایج پژوهش نشان دادند که بازخرید بیمه‌نامه‌ها در استان تهران بیشتر از سایر استان‌هاست. وزن استان محل سکونت در بازخرید در شکل ۶ نمایش داده شده است و می‌توان دید که در مقایسه با تهران، به‌ترتیب استان‌های اصفهان، کهکیلویه و بویراحمد، آذربایجان غربی، لرستان، مرکزی، خراسان رضوی، فارس، سمنان، کردستان و چهارمحال و بختیاری به‌طور ویژه کمترین بازخرید را داشته‌اند.

نرخ خطر شغلی نیز رابطه و همبستگی مثبتی با بازخرید بیمه‌نامه دارد که اگرچه خلاف انتظار است، اما نسبت به دیگر متغیرها پایین‌ترین ضریب اثرگذاری را دارد و به‌نوعی معنی‌دار نیست. دلیل معنی‌دار نبودن این متغیر را باید فراوانی پایین و تعداد کم مشاهدات با نرخ خطر شغلی بالا دانست، به‌گونه‌ای که در جامعهٔ آماری کمتر از ۲ درصد افراد نرخ خطر شغلی بالا (۴ و ۵) داشته‌اند. در مقابل، دو متغیر دیگر، یعنی اضافه نرخ پزشکی و متغیر نرخ حادثی اثرگذاری بالایی داشته که علامت آن‌ها مطابق انتظار نیز هست. نتایج به‌دست‌آمده نشان می‌دهند افرادی که در شغل خود با حوادث مختلف روبه‌رویند یا وضعیت سلامت آن‌ها مناسب نیست با احتمال کمتری بیمه‌نامهٔ خود را بازخرید خواهند کرد که تأییدکنندهٔ رابطهٔ بیمه‌نامهٔ عمر و ریسک فوت (به‌دلیل حادثه یا وضعیت سلامت) است.

یکی دیگر از مشخصه‌های جالب توجه نسبت بیمه‌گذار و

جدول ۷: وزن اثرگذاری متغیرها و مشخصه‌ها در پیش‌بینی بازخرید بیمه‌نامه
Table 7: The influencing weights of research variables on the prediction of policy-holders surrender

اطلاعات جمعیت‌شناختی و اجتماعی بیمه‌شدگان Demographics and Social Characteristics		اطلاعات قراردادی بیمه‌نامه‌های عمر The Insurance Policy Features	
متغیر Variable	وزن اثرگذاری Weight	متغیر Variable	وزن اثرگذاری Weight
سن Age	-26.4	مدت بیمه‌نامه Term of Insurance	-11.8
جنسیت (مرد) Sex (Men)	37.7	مدت سپری‌شده Elapsed Time	-68.7
نرخ خطر شغلی Job Risk Class	1.94	نحوهٔ پرداخت حق‌بیمه Payment Method	-56.9
مجموع نرخ حادثی Accident Rate	-53.2	ضریب افزایش سالانهٔ حق‌بیمه Increasing Premiums (Annual rate)	-6.04
اضافه نرخ پزشکی Health Extra Premium	-60.1	ضریب افزایش سالانهٔ سرمایه Increasing Benefits (Annual rate)	-8.7
یکسان بودن بیمه‌شده و بیمه‌گذار Insurrer is the Insured	-7.9	موارد پوشش Number of Coverage Cases	5.0
نسبت بیمه‌شده با بیمه‌گذار Relation between Insured and the Insurer	25.3	سرمایهٔ فوت Death Benefit	34.5
استان محل سکونت (پایه تهران) Province (Tehran)	Negative	کل سرمایه Total Benefit	-2.9

بیمه‌شده است. نخستین متغیر درصورتی‌که بیمه‌شده همان بیمه‌گذار باشد عدد ۱ و در غیر این صورت عدد صفر به خود می‌گیرد. از نتایج به‌دست‌آمده می‌توان دید که این متغیر رابطه‌ای معکوس با بازخرید بیمه‌نامه دارد، به این معنی که بازخرید کمتر است، وقتی که بیمه‌گذار بیمه‌نامهٔ عمر را برای خود خریده باشد. در شرایطی که بیمه‌نامه برای افراد دیگر خریداری شده باشد، بازخرید بیشتر اتفاق خواهد افتاد و متغیر دوم نشان می‌دهد با دورتر شدن نسبت خویشاوندی این احتمال افزایش می‌یابد.

در ارتباط با اطلاعات قراردادی بیمه‌نامه‌ها، نخست می‌توان دید که مدت بیمه‌نامه ارتباط یا همبستگی منفی با بازخرید بیمه‌نامه دارد. این نتیجه مطابق انتظار است، چون به‌طور معمول مدت بیمه‌نامه نشان‌دهندهٔ چشم‌انداز و رویکرد بلندمدت افراد در برنامه‌ریزی مالی آن‌هاست. مدت‌زمان سپری‌شده نیز یک رابطه یا همبستگی منفی با بازخرید بیمه‌نامهٔ عمر دارد. افرادی که مدت‌زمان بیشتری از زمان قرارداد آن‌ها گذشته باشد، کمتر بیمه‌نامهٔ عمر خود را بازخرید خواهند کرد که می‌توان آن را به تأثیر هزینهٔ مرده یا تأثیر وابستگی به مسیر طی‌شده نسبت داد. متغیر دیگر نحوهٔ پرداخت حق‌بیمه است. طبق نتایج به‌دست‌آمده هرچه شیوهٔ پرداخت حق‌بیمه با اقساط طولانی‌تر همراه باشد (اقساط سالانه یا شش‌ماهه) بازخرید کمتر و در مقابل هرچه اقساط کوتاه‌تر باشد (ماهانه یا دوماهه) بازخرید بیشتر خواهد بود. اثر این متغیر را می‌توان در کنار تورم دید، چراکه بالا بودن تورم در ایران در واقع منجر شده تا اقساط طولانی‌مدت به‌صرفه‌تر به نظر برسد. این تأثیر را می‌توان در اثرگذاری ضرایب افزایش سالانه حق‌بیمه و سرمایه نیز دید. بیمه‌نامه‌هایی که در آن‌ها ضریب افزایش سرمایه بالاتر است، به معنی تعدیل سرمایهٔ بیمه‌نامه با تورم بوده و به‌این‌ترتیب این بیمه‌نامه‌ها با وجود تورم بالا همچنان برای افراد به‌صرفه است و احتمال بازخرید را برای آن‌ها پایین‌تر آورده است.

همچنین طبق نتایج به‌دست‌آمده می‌توان دید که بین تعدد موارد پوشش و همچنین مقدار حقیقی‌شده سرمایهٔ فوت و بازخرید بیمه‌نامهٔ عمر یک رابطه یا همبستگی مثبت وجود دارد. در نگاه اول این شاید غیرقابل درک باشد، اما باید توجه داشت که در ساختار شبکهٔ عصبی نمی‌توان روابط علّی یا حتی روابط مستقیم با ثابت گرفتن سایر متغیرها را مشاهده کرد و وزن یا تأثیر برآوردشده یک تأثیر نهایی از بی‌نهایت متغیرهایی است که از مسیر این متغیرها بر بازخرید اثر گذاشته‌اند. باوجوداین، در توضیح این نتیجه می‌توان گفت که می‌دانیم که حق‌بیمه بر اساس سرمایهٔ فوت و همچنین موارد پوشش و ... مشخص می‌شود. هرچه تعداد موارد پوشش بیشتر باشد یا هرچه مبلغ حقیقی‌شدهٔ سرمایهٔ فوت بالاتر باشد، حق‌بیمهٔ پرداختی نیز مبالغ بالاتر و سنگین‌تری خواهد بود و بر این اساس به افزایش بازخرید بیمهٔ عمر منجر خواهد شد.

جمع‌بندی و پیشنهادها

با وجود سبقهٔ چند صدسالهٔ بیمهٔ عمر در جهان، هنوز هم بیمهٔ عمر به‌عنوان یک محصول مهم بیمه‌ای و برنامه‌ریزی مالی نفوذ کمی در بازار و در میان افراد جامعه، به‌ویژه در ایران دارد. یکی از دلایل

شکل ۶: وزن تأثیر استان محل سکونت در بازخرید در مقایسه با تهران

Fig. 6: The influencing weights of residence province on the policy-holders surrender in comparison with Tehran

ضریب نفوذ پایین، بازخرید بیمه‌نامه‌هاست که به دلایل متعددی می‌تواند اتفاق بیفتد. این مقاله به‌دنبال بررسی تأثیر مشخصه‌های فردی و قراردادی بیمه‌نامه‌هاست که بر بازخرید بیمه‌نامه‌های عمر به شرط فوت اثر می‌گذارند. با اتخاذ هدف مذکور، این مقاله از داده‌های آماری و اطلاعات ثبتی ۳۵۱۷۱ خریدار بیمه‌نامه‌های عمر و مستمری یکی از شرکت‌های بیمه به‌عنوان پایلوت بهره گرفت. در این راستا، از داده‌کاوی و استفاده از الگوریتم‌های یادگیری عمیق و شبکهٔ عصبی استفاده شد که دقت بسیار بالاتری در پیش‌بینی دارند.

پس از اعتبارسنجی مدل‌های مختلف، در نهایت یک مدل شبکهٔ عصبی با سه لایهٔ پنهان به‌ترتیب با تعداد ۳۲، ۱۶ و ۸ نورون استفاده شد که مقادیری بین تعداد نورون لایهٔ ورودی (تعداد مشخصه‌ها ۴۵) و تعداد نورون لایهٔ خروجی (یک خروجی ۱) گرفته‌اند. این مدل با داده‌های آموزش (۷۰ درصد مجموعهٔ داده‌ها) فرایند یادگیری را طی کرد و سپس با داده‌های تست آزمون شد. شاخص دقت ۷۴ درصدی مدل در پیش‌بینی هر دو نوع بیمه‌نامه‌های عدم بازخرید و بازخریدشده عملکرد مطلوبی داشت. البته در پیش‌بینی عدم بازخرید بیمه‌نامه‌ها عملکرد بسیار بهتر بوده، اما چون موضوع اصلی مقاله پیش‌بینی بیمه‌نامه‌های بازخریدشده است، در تفسیر نتایج بیشتر به آن توجه شد. نتایج به‌دست‌آمده با وجود مشکل نامتوازن بودن داده‌ها مطلوب است. در داده‌های مورد بررسی نسبت بیمه‌نامه‌های بازخریدی به عدم بازخرید ۳ به ۱۰۰ است که این عدم توازن موجب می‌شود فرایند یادگیری به سمت پیش‌بینی طبقه با بیشترین فراوانی سوگیری پیدا کند. باوجوداین، شاخص پوشش ۵۹ درصدی به‌دست‌آمده برای طبقهٔ اقلیت بازخریدی نشان داد که از مجموع ۲۴۴ بیمه‌نامهٔ بازخریدشده در مجموعه دادهٔ تست، شبکه توانسته اغلب آن‌ها، یعنی ۱۴۵ مورد را به‌درستی در طبقهٔ بیمه‌نامه‌های بازخریدی پیش‌بینی و طبقه‌بندی کند.

با نگاه به درون این شبکهٔ پیچیده که به «جعبهٔ سیاه» شهرت دارد می‌توان تاحدی به تأثیر عوامل و مشخصه‌های مختلف در بازخرید بیمه‌نامه‌ها پی برد. نتایج به‌دست‌آمده نشان دادند که از مجموع مشخصه‌های جمعیت‌شناختی متغیرهای سن، جنسیت زن، اضافه نرخ پزشکی (نشان‌دهندهٔ سلامت فرد)، نرخ خطر حادثی (نشان‌دهندهٔ حوادث شغلی فرد) با بازخرید بیمه‌نامه به‌صورت عکس مرتبط‌اند. نسبت بیمه‌گذار و بیمه‌شده نیز یکی از عوامل اجتماعی تأثیرگذار است و نتایج نشان دادند بازخرید وقتی بیمه‌گذار بیمه‌نامهٔ عمر را برای خود بخرد در حداقل است، اما با دور شدن نسبت خویشاوندی احتمال موارد بازخرید افزایش می‌یابد. از میان مشخصه‌های قراردادی نیز مدت بیمه‌نامه، مدت‌زمان سپری‌شده از شروع بیمه‌نامه، شیوهٔ پرداخت حق‌بیمه با اقساط بلندمدت‌تر، بالاتر بودن ضرایب افزایش سالانهٔ سرمایه و حق‌بیمه (تعدیل مناسب با تورم) و کمتر بودن تعداد موارد پوشش و سرمایهٔ فوت (و در نتیجه حق‌بیمهٔ کمتر) با بازخرید اثر عکس دارند و آن را کاهش می‌دهند.

از یافته‌های این پژوهش نمی‌توان روابط علّت و معلولی برداشت کرد، اما شرکت‌های بیمه قطعاً می‌توانند از پیش‌بینی‌های آن در ارتباط با بازخرید بیمه‌نامه‌ها استفاده کنند. به‌طور مثال، چون مدت بیمه‌نامه نشان‌دهندهٔ چشم‌انداز افراد از برنامه‌ریزی مالی خود برای خریداری بیمه‌نامهٔ عمر است و هرچه طولانی‌تر باشد بازخرید کمتر است، بنابراین بهتر است شرکت‌ها با تبلیغات یا سازوکارهای تشویقی

افراد را به سمت خریداری بیمه‌نامه‌های با مدت طولانی‌تر ترغیب کنند. یا چون بازخرید در شروع بیمه‌نامه بیشتر است و با گذر زمان و سپری شدن مدت به‌تدریج بازخرید کاهش می‌یابد، شرکت‌ها بهتر است در طول مدت بیمه‌نامه با بیمه‌گذار تعامل داشته باشند و با ارسال پیام‌هایی به آنان مزایا و منافع بیمهٔ عمر را یادآوری کنند. شرکت‌ها می‌توانند در خصوص حفظ و نگهداشت مشتریان سیاست‌های مشخصی را به‌کار گیرند و خدمات جدیدی به‌مرور به آن‌ها پیشنهاد دهند. همچنین، شرکت می‌تواند با گذشت زمان کافی از شروع بیمه‌نامه، تخفیف‌هایی در پرداخت حق‌بیمه به‌منظور حفظ مشتری داشته باشد یا سیاست‌های تخفیفی در تمدید و به‌روزرسانی بیمه‌نامه‌ها داشته باشد. حتی در خصوص عوامل دیگر نیز شرکت‌ها باید سیاست‌هایی را به‌منظور کاهش بازخرید در نظر داشته باشند. به‌طور مثال دیده شد که انگیزهٔ بیمه‌گذار برای پوشش بیمه‌شدگان مرد و خویشاوندان دورتر به‌مراتب کمتر و در نتیجه بازخرید بیشتر خواهد بود و لذا شرکت‌ها می‌توانند برنامه‌های وفاداری مخصوص این گروه‌ها در نظر گیرند تا ماندگاری آن‌ها بیشتر شود یا با توجه به کاهش بازخرید با افزایش سن، شرکت‌ها باید در کنار ارائهٔ پاداش‌های وفاداری، تخفیفاتی برای نگه داشتن مشتریان جوان در نظر گیرند.

مشارکت نویسندگان

عباس خندان: جمع‌آوری مطالعات مرتبط و تدوین مدل، روش پژوهش و متدولوژی، مروری بر ادبیات پژوهش، و نتیجه‌گیری؛ لیلی نیاکان: کنترل چهارچوب تدوین و استانداردهای پژوهشی؛ زهرا فخاری‌نژاد: جمع‌آوری داده‌ها.

تشکر و قدردانی

این مقاله از یک طرح پژوهشی فرصت مطالعاتی در پژوهشکدهٔ بیمه انجام شده است. از پژوهشکدهٔ بیمه به‌سبب فراهم آوردن امکانات و لوازم تحقیق و از دانشگاه خوارزمی برای تأمین بودجه قدردانی می‌شود.

تعارض منافع

نویسندگان اعلام می‌دارند که در مورد 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 به درج نحوهٔ دقیق دسترسی به مجوز CC منوط به ذکر تغییرات احتمالی بر روی مقاله می‌داند. لذا 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

Abbasi, E.; Sazgar, A., (2005). Understanding the reasons for life insurance surrender in Iran insurance company. Iran. J. Insur. Res., 74(2): 87-119 **(33 Pages)**. [In Persian]

Azodi, C.B.; Tang, J.; Shiu, S.H., (2020). Opening the black box: Interpretable machine learning for geneticists. Trends. Genet., 36(6): 442-455 **(14 Pages)**.

Azzone, M.; Barucci, E.; Mancayo, G.G.; Marazzina, D., (2022). A machine learning model for lapse prediction in life insurance contracts. Expert. Syst. Appl., 191(1): 116261.

Bakhtiar Nasrabadi, H.A.; Hassangholipoor, T.; Mira, S.A.; Vedadhir, A.A., (2020A). Developing a model of policyholder's surrender behavior: A study based on the grounded theory. New. Mark. Res. J., 9(4): 33-54 **(22 Pages)**. [In Persian]

Bakhtiar Nasrabadi, H.A.; Hassangholipoor, T.; Vedadhir, A.A.; Badin, M.; Mira, S.A., (2020B). Explanation of skepticism and trust dialectic in buying behavior of life insurance consumers: A grounded theory. Iran. J. Insur. Res., 9(4): 43-88 **(46 Pages)**. [In Persian]

Balaji, S.; Srivatsa, S.K., (2012). Naive bayes classification approach for mining life insurance databases for effective prediction of customer preferences over life insurance products. Int. J. Comput. Appl., 51(3): 22-26 **(5 Pages)**.

Bash Afshar, M.; SaeedPanah, M.; Tireh Eidouzhi, F., (2018). Clustering model of life insurance customers (Case study: An insurance company). Iran. J. Insur. Res., 7(2): 45-64 **(20 Pages)**. [In Persian]

Berson, A.; Smith, S.; Thearling, K., (1999). Building data mining applications for CRM. New York, NY: McGraw-Hill.

Bimeh Markazi Centeral Insurance of Iran (2021). Statistical yearbook of insurance. Tehran: Bimeh Markazi Centeral Insurance of Iran.

Buhrmester, V.; Munch, D.; Arens, M., (2021). Analysis of explainers of black box deep neural networks for computer vision: A survey. Mach. Learn. Knowl. Extr., 3(4): 966-989 **(24 Pages)**.

Dash, G., (2018). Determinants of life insurance demand: Evidences from India. Asia. Pac. Inst. Adv. Res., 4(2): 86-99 **(14 Pages)**.

Deacon, R.E.; Firebaugh, F.M., (1988). Family resource management: Principles and applications. Allyn & Bacon.

Ghorbani, H.; Ghanbarzadeh, M.; Ofoghi, R., (2022). Investigating the churn of life insurance customers using data mining methods (A case study: One of the Iran's insurance companies). Iran. J. Insur. Res., 11(4): 305-320 **(16 Pages)**. [In Persian]

Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D., (2018). A survey of methods for explaining black box models. ACM. Comput. Surv., 51(5): 1-42 **(42 Pages)**.

Habibi Marand, N., (2016). Factors affecting redemption of life insurance and investments with an emphasis on microeconomic factors (The case study of Parsian insurance company). [In Persian]

Helmzadeh, A.; Hamidi, K.; Heidarzadeh Hanzaei, K., (2020). Investigating factors affecting the tendency to life insurance surrender. Iran. J. Insur. Res., 9(2): 67-98 **(32 Pages)**. [In Persian]

Heo, W., (2020). The demand for life insurance: Dynamic ecological systemic theory using machine learning techniques. Palgrave Macmillan.

Hozarmoghaddam, N.; Ghanbarzadeh, M.; Hamze, A.; Ghafoor Boroujerdi, M., (2020). Investigating the effects of the coronavirus outbreak on life, health and travel insurance. Iran. J. Health. Insur., 3(3): 148-161 **(14 Pages)**. [In Persian]

Hu, S.; O'Hagan, A.; Sweeney, J.; Ghahramani, M.H., (2021). A spatial machine learning model for analysing customers' lapse behaviour in life insurance. Anal. Actuarial. Sci., 15(2): 367-393 **(27 Pages)**.

Khandan, A., (2022). Prediction and investigation of various factors' effect on the surrender of life insurance constracts. Insur. Res. Center. Res. Project. [In Persian]

Kudyba, S.; Kwatinetz, M., (2014). Introduction to the big data era. In Big Data, Mining, and Analytics. Taylor & Francis Group.

Liang, Y.; Li, S.; Yan, C.; Li, M.; Jiang, C., (2021). Explaining the black-box model: A survey of local interpretation methods for deep neural networks. Neuro. Comput., 419: 168-182 **(15 Pages)**.

Mahdavi, G.; Ofoghi, R.; Abed, M., (2015). The impact of risk aversion on surrender of life insurance policies case study (Iranian life insurance market). Iran. J. Insur. Res., 4(3): 61-75 **(15 Pages)**. [In Persian]

Milhaud, X., Loisel, S.; Maume-Deschamps, V., (2011). Surrender triggers in life insurance: What main features affect the surrender behavior in a classical economic context? Bull. Fr. d'Actuariat., 11(22): 5-48 **(44 Pages)**.

Olden, J.D.; Jackson, D.A., (2002). Illuminating the "black box": A randomization approach for understanding variable contributions in artificial neural networks. Ecol. Model., 154(1/2): 135-150 **(16 Pages)**.

Sazgar, A.; Abed, M., (2018). Explaining factors affecting life insurance surrender. Life insurance development seminar. [In Persian]

Sulaiman, L.A.; Migiro, S.; Yeshihareg, T., (2015). Investigating the factors influencing the life insurance market in Ethiopia. Probl. Perspect. Manage., 13(2): 152-160 **(9 Pages)**.

Vazan, M., (1992). Deep learning: Bases, concepts, and approches. Miad Andisheh publication. [In Persian]

Yaari, M.E., (1965). Uncertain lifetime, life insurance, and the theory of the consumer. Rev. Econ. Stud., 32(2): 137-150 **(14 Pages)**.

Zietz, E.N., (2003). An examination of the demand for life insurance. Risk Manag. Insur. Rev., 6(2): 159-191 **(33 Pages)**.

AUTHOR(S) BIOSKETCHES **معرفی نویسندگان**

عباس خندان، استادیار اقتصاد، گروه اقتصاد امور عمومی، دانشکده اقتصاد، دانشگاه خوارزمی، تهران، ایران

- Email: *khandan.abbas@khu.ac.ir*
- ORCID: 0000-0002-4558-6653
- Homepage: https://khu.ac.ir/cv/4535/english

لیلی نیاکان، استادیار، گروه پژوهشی عمومی بیمه، پژوهشکدهٔ بیمه، تهران، ایران

- Email: *niakan@irc.ac.ir*
- ORCID: 0000-0002-9821-8512
- Homepage: https://www.irc.ac.ir/niakan

فخاری‌نژاد، مدیر بخش بیمه‌های زندگی، شرکت بیمه دی، تهران، ایران

- Email: *z_fakharinejad@dayins.com*
- ORCID: 0000-0001-6850-721X
- Homepage: https://www.dayins.com

HOW TO CITE THIS ARTICLE

Khandan, A.; Niakan, L.; Fakharinezhad, Z., (2023). Predicting term life insurance surrender using deep neural networks. Iran. J. Insur. Res., 12(4): 265-282.

DOI: 10.22056/ijir.2023.04.02

URL: https://ijir.irc.ac.ir/article_160295.html?lang=en