



Международный журнал информационных технологий и энергоэффективности

Сайт журнала:

<http://www.openaccessscience.ru/index.php/ijcse/>



УДК 004.8

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ФИНАНСОВЫХ ПОКАЗАТЕЛЕЙ СТРАХОВОЙ КОМПАНИИ

Карасёв Д.М.

ФГБОУ ВО "МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМЕНИ Н.Э. БАУМАНА (НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)", Москва, Россия (105005, город Москва, 2-Я Бауманская ул, д. 5 стр. 1), e-mail: weasdqw72@gmail.com

В данной статье осуществлено комплексное исследование методов прогнозирования финансовой деятельности страховой компании, что представляет собой ключевую задачу в условиях высокого риска и неопределённости, характерных для конкурентных рынков страховых услуг. В работе подробно анализируются как классические, так и современные подходы к прогнозированию, основанные на обработке сложных данных страховой сферы. В частности, рассмотрены особенности использования статистических методов, машинного обучения и глубокого обучения для эффективного прогнозирования финансовых показателей. Особое внимание уделено методам градиентного бустинга и анализу временных рядов, результаты применения которых демонстрируют значительное улучшение в точности прогнозов финансовой деятельности страховой компании по сравнению с традиционными моделями.

Ключевые слова: Риск, страхование, прогнозирование, страховые услуги, финансовая деятельность.

THE USE OF MACHINE LEARNING METHODS TO PREDICT THE FINANCIAL PERFORMANCE OF AN INSURANCE COMPANY

Karasev D.M.

BAUMAN MOSCOW STATE TECHNICAL UNIVERSITY (NATIONAL RESEARCH UNIVERSITY), Moscow, Russia (105005, Moscow, 2nd Baumanskaya str., 5/1), e-mail: weasdqw72@gmail.com

This article provides a comprehensive study of methods for forecasting the financial performance of an insurance company, which is a key task in conditions of high risk and uncertainty characteristic of competitive insurance markets. The paper analyzes in detail both classical and modern approaches to forecasting based on the processing of complex insurance data. In particular, the features of using statistical methods, machine learning and deep learning for effective forecasting of financial indicators are considered. Special attention is paid to gradient boosting methods and time series analysis, the results of which demonstrate a significant improvement in the accuracy of forecasts of the financial activity of an insurance company compared with traditional models.

Keywords: Risk, insurance, forecasting, insurance services, financial activities.

Введение

Стабильная экономическая деятельность многих промышленных предприятий и организаций, компаний на конкурентном рынке является залогом социально-экономической политики государства. Согласно Тарасовой Ю.А. «Эффективное функционирование различных отраслей экономики, таких как отрасли промышленности, сельское хозяйство, а также предпринимательства, возможно только при наличии хорошо организованной системы государственного управления в страховой области» [1]. Это предполагает взаимодействие четкой иерархии власти, где каждый уровень выполняет свои определенные функции и задачи.

Вертикаль государственного управления позволяет координировать деятельность различных секторов экономики, устанавливать правила и стандарты, а также обеспечивать контроль за их выполнением и соблюдением.

Например, в современных конкурентных рыночных условиях компании должны управлять объемом предоставления определенных страховых продуктов и услуг, и их продаж. Актуальность объема продаж связана напрямую с финансовой деятельностью страховой компании. Она непосредственным образом влияет на доходы страховой компании, так как именно от продажи страховых продуктов и услуг зависит ее прибыль. Она влияет на финансовую устойчивость компании сбалансировав размер ее активов и обязательств.

Поэтому очень важно уметь прогнозировать спрос и управлять предложением таким образом, чтобы финансовые результаты были направлены на минимизацию затрат и избежание потерь потенциальной прибыли.

Рыночная конкурентная среда всегда остается нестабильной и непредсказуемой от субъективных и объективных причин [2]. В таких рыночных условиях на конкурентном рынке, компании вынуждены защитить свои интересы на рынке от различных рисков, связанные с их коммерческой деятельностью [3]. Например, прогнозирование многих показателей с учетом рисков финансово-хозяйственной деятельности компании позволяет им эффективно позиционировать себя на конкурентном рынке, а также стратегически управлять им на долгосрочной перспективе. Стратегическое планирование помогает компании оценить свои текущие возможности и потенциальные угрозы, чтобы разработать эффективную стратегию для достижения поставленных целей. Оно также позволяет компании адаптироваться к изменяющимся требованиям рынка и предпочтениям клиентов, расширяя спектр предоставляемых страховых продуктов и услуг, максимально удовлетворяя все потребности потребителей.

В этом случае страхование компании позволяют снизить финансовые потери организаций, предприятий и физических лиц в случае возникновения непредвиденных ситуаций, таких как аварии, кражи, пожары, природные катаклизмы и другие форс- мажорные обстоятельства. Также и сами страховые компании сталкиваются с трудностями при планировании своей деятельности на длительный период времени из-за быстро меняющейся и непредсказуемой ситуации на страховом конкурентном рынке. Это может быть вызвано различными факторами, основными из которых являются изменения в законодательстве государственной политики, экономические кризисы во внешних и внутренних рынках, внедрением новых технологий в работу конкурентов и другие [4].

Несмотря на значительные успехи в методах прогнозирования финансовой деятельности страховой компании, все еще остаются вопросы, связанные со сложностью проведения процессов прогнозирования временных рядов и они требуют дальнейшего изучения.

Цель работы – исследование и анализ применения методов машинного обучения для прогнозирования финансовых показателей страховой компании.

Процесс прогнозирования включает в себя анализ данных и информации из различных источников для определения будущих тенденций и возможностей спроса на товары или услуги страховой компании. Результаты этого исследования помогают руководству страховой компании принимать обоснованные решения, связанные с планированием мероприятий по продвижению бренда компании, маркетингом, ценообразованием и другими аспектами

бизнеса. Прогнозирование спроса является важной частью процесса принятия управленческих решений, поскольку оно позволяет страховой компании лучше понимать потребности своих клиентов и адаптировать свою стратегию для удовлетворения этих потребностей.

В наши дни главной целью страховых компаний является извлечение прибыли на перспективу [5]. Но страховые компании работают в специфическом секторе экономики, где получение прибыли неразрывно связано с выполнением социальной функции – предоставлением защиты от рисков и неопределенности будущего. Необходимо также определить категорию населения или клиентов страховой компании – это различные физические и юридические лица, адекватно реагирующие на изменения государственной политики, социально-экономических отношений в обществе, а также предугадывающие будущие риски различных событий в будущем. Поэтому, хотя прибыль и является ключевым показателем эффективности работы страховой компании, она должна достигаться в рамках справедливого и прозрачного ценообразования, адекватной оценки рисков и соблюдения интересов клиентов.

Согласно Т.Г.Гурнович и Е.И.Ладыгиной в настоящее время «существует большое количество методов, позволяющих осуществлять прогнозирования объемов страховых услуг. Наиболее часто используются три группы методов: методы экспертных оценок; методы анализа и прогнозирования временных рядов; казуальные (причинно-следственные) методы» [6].

Прогнозирование посредством экспертных оценок основывается на принципах анализа текущего состояния страховой компании и ее потенциала для дальнейшего развития. Этот подход предполагает привлечение специалистов-экспертов, которые анализируют внутренние и внешние факторы, влияющие на компанию, и делают прогнозы относительно ее будущего развития. Конъюнктурные оценки, полученные с помощью экспертных методов полезны в случаях, когда использование других методов затруднено из-за отсутствия полной информации, необходимой для качественного прогноза.

Методы прогнозирования данных страховой компании на основе анализа и прогнозирования временных рядов включают в себя различные статистические и математические модели, которые используются для предсказания будущих событий, таких как количество страховых случаев, размер выплат по страховым полисам и другие. Они позволяют более точно оценивать риски и принимать обоснованные решения по управлению своим объектом, компанией или другим имуществом.

Казуальные методы прогнозирования данных основаны на анализе причинно-следственных связей между различными переменными. Например, при исследовании связи между уровнем доходов населения и спросом на страховые продукты используют корреляционный анализ. Если обнаруживается устойчивая корреляция между ними, то можно сделать прогноз о том, как изменение уровня доходов повлияет на спрос на страховые продукты. Они позволяют понимать динамику конкурентного рынка и позволяют принимать решение о стратегии развития страховой компании.

В настоящее время наиболее успешными методами прогнозирования финансовой деятельности страховой компании являются исследование временных рядов и корреляционно-регрессионные модели.

Прогнозирование временных рядов финансовой деятельности (недельные, квартальные, годовые) страховых компаний включают в себя такие методы и подходы [7]:

1. Анализ временных рядов. Он включает в себя анализ трендов во времени, сезонности, цикличности и других характеристик данных.

2. Прогнозирование на основе ARIMA-моделей. ARIMA (авторегрессионная интеграция скользящего среднего) – это статистическая модель, которая используется для прогнозирования временных рядов. Она учитывает авторегрессию (зависимость текущего значения от предыдущих), интеграцию (убыток тренда) и скользящее среднее (учет сезонности).

3. Прогнозирование на основе нейронных сетей. Нейронные сети используются для прогнозирования временных рядов, если они имеют сложную структуру или нелинейные зависимости.

4. Прогнозирование на основе деревьев решений. Деревья решений – это алгоритмы машинного обучения, которые используются для прогнозирования временных рядов. Успешность их применения обосновано на том, что они обрабатывают данные с большим количеством переменных, а затем результаты интерпретируются.

5. Прогнозирование на основе методов машинного обучения. Например, метод опорных векторов, случайный лес, градиентный бустинг и другие. Градиентный бустинг позволяет улучшить производительность модели путем последовательного добавления деревьев решений к слабому предсказателю, т.е. итерация будет проводиться до тех пор, пока не будет достигнута определенная, заранее заданная точность.

6. Прогнозирование на основе статистических моделей. Статистические модели, такие как линейная регрессия, логическая регрессия, дискриминантный анализ, используются для прогнозирования сложных данных путем построения линейных и нелинейных моделей, а также построения моделей на основе дифференцированных данных.

В настоящее время финансовый план страховой компании формируется из следующих составных частей:

- формирования структуры капитала;
- определение объема привлеченных средств;
- определения направлений размещения привлеченных средств;
- определение лимитов по видам страхования;
- определения порядка формирования и использования страховых резервов;
- определения порядка формирования и использования прибыли;
- формирования показателей развития на предстоящий период;
- определения порядка налогообложения.

Рассмотрим использование градиентного бустинга для прогнозирования финансовой деятельности страховой компании. Он основан на принципе обучения на основе ансамбля слабых моделей, таких как деревья решений. Пусть имеется большой массив данных наблюдаемых параметров, содержащих информацию о финансовых результатах и клиентах страховой компании, таких как страховые резервы, страховые выплаты, денежный поток, бюджет доходов, количество клиентов и прибыль компании (рисунок 1). Обучающая выборка состоит из $m=1002$ результатов наблюдений и $n=6$ входящих переменных: $X = \{X_1, X_2, X_3, X_4, X_5\}$ и результирующей переменной Y .

X1	X2	X3	X4	X5	Y
317,45	234,26	309,12	378,12	527	12673408
321,34	264,71	311,17	365,46	524	11745691
318,16	256,08	308,23	354,18	538	11894367
319,10	237,12	307,22	364,80	544	12367234
318,74	238,23	306,86	356,24	529	12094567
317,03	247,72	307,28	366,17	541	11945274
319,62	218,72	308,15	373,44	536	12178361

Рисунок 1 – Исходный фрагмент обучающей выборки

Необходимо построить отображение $F : X \rightarrow Y$, где X – вектор внутренних переменных, а Y – прогнозируемая величина.

Наша цель – предсказать вероятность того, что клиенты купят новый полис страхования в следующем году. Для этого используем алгоритм градиентного бустинга, а также исследуем возможности снижения размерности задачи без потери информационной емкости исходных данных финансовых результатов страховой компании.

Решение данной задачи проведем с помощью метода градиентного бустинга на базе деревьев решений (*Gradient boosting trees*) [8,9].

Необходимо также оценить качество прогноза с помощью статистических показателей, представленных в таблице 1, где Y_i – известные табличные данные, а \hat{Y}_i – расчетные предсказанные значения.

Таблица 1 – Статистические показатели оценки качества прогноза

Средние оценки	Среднеквадратическая ошибка (Root Mean Square Error)	$RMSE = \sqrt{\frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{m}}$
	Средняя % ошибка (Mean Absolute Percentage Error)	$MAPE = \frac{1}{m} \sum_{i=1}^m \frac{ y_i - \hat{y}_i }{y_i} \cdot 100\%$
Точечная оценка	Ошибка экстраполяции (Relative Extrapolation Error)	$REE = \frac{ y_{m+1} - \hat{y}_{m+1} }{y_{m+1}} \cdot 100\%$

Расчеты проведем в программной среде Wolfram Mathematica, использующейся для численных расчетов, построения графиков, работы с массивами данных, программирования и других задач связанных с математикой и статистикой [9].

На Рисунке 2 представлены результаты работы оператора модели прогнозирования методом градиентного бустинга.



Рисунок 2 – Составление базы данных и построение модели

Характеристика процесса обучения модели включает в себя множество параметров, включая изменение во времени стандартного отклонения. Оно является мерой разброса значений случайной величины относительно ее математического ожидания. В контексте машинного обучения, стандартное отклонение здесь используется для оценки стабильности модели во время обучения.

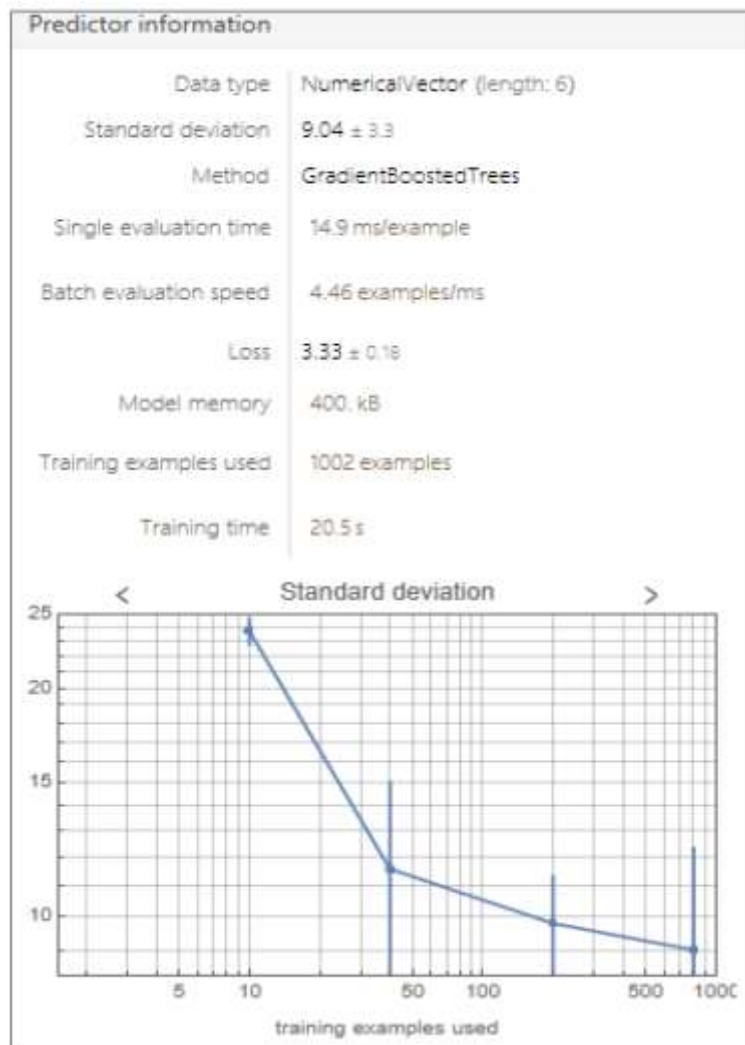


Рисунок 3 – Процесс обучения программы

Расчет качества обучения модели представлен на Рисунке 4.

```

predictresults = Flatten[GBT[x]]; (*прогноз результатов для расчета оценок*)
                                |уплостить
(*Среднеквадратическая ошибка*)
RMSEGBT = Sqrt[Sum[Power[y[[i]] - predictresults[[i]], 2], {i, Length[y]}] / Length[y]]
           |кв... |с... |степень           |длина |длина
(*Средняя процентная ошибка*)
MAPEGBT = PercentForm[Sum[Abs[y[[i]] - predictresults[[i]]] / y[[i]], {i, Length[y]}] / Length[y]]
           |форма проц... |с... |абсолютное значение           |длина |длина
(*Ошибка экстраполирования для завтрашнего дня*)
Error = PercentForm[Abs[GBTResult - answer] / answer]
           |форма проц... |абсолютное значение
    
```

Рисунок 4 – Расчет статистических показателей согласно таблице 1

На основании вычисленных значений две переменные оказались несущественными на основании коэффициента корреляции, поэтому проводилось поэтапное понижение размерности (методом главных компонент) от четырех до единицы, т.е количества клиентов.

Как следует из результатов, уменьшение параметров не привела к существенному изменению оценка качества прогноза, представленная на Рисунке 5.

Ошибка	n				
	6	4	3	2	1
RMSE	3.672	3.886	4.007	3.898	5.031
MAPE	1.03%	1.09%	1.09%	1.02%	1.38%
REE	0.33%	0.55%	1.12%	0.35%	0.45%

Рисунок 5 – Вычисленные значения статистических переменных при понижении числа параметров модели

Таким образом, было проведено обучение модели с помощью обучающей выборки, а с помощью тестовой выборки проведена проверка ее эффективности. В процессе обучения модель добавляла новые деревья решений до тех пор, пока не достигла определенной точности или не превысила заданное количество итераций.

Таким образом, мы можем использовать обученную модель для прогнозирования финансовой деятельности страховой компании. Для этого мы можем ввести новые данные клиентов и предсказать вероятность того, что они купят новый полис страхования в будущем.

Вывод.

Для того чтобы повысить эффективность работы страховой компании в условиях неопределенности и риска, необходимо систематически планировать и прогнозировать финансовую деятельность. Основой для этого служит детальное изучение общей ситуации на конкурентном рынке страховых услуг, включая динамику предпринимательской активности и жизненный уровень населения в регионе присутствия компании. Такой подход позволяет учитывать все возможные факторы, влияющие на успешность бизнеса, и принимать взвешенные решения, направленные на минимизацию рисков и максимизацию прибыли. Градиентный бустинг временных рядов является более успешным методом прогнозирования финансовой деятельности страховой компании.

Список литературы

1. Тарасова Ю. А. Страховое дело. -Москва: Изд-во Юрайт, 2024. - 232 с.
2. Дрюк Т. В. Целевая модель развития страхового рынка России: клиентоцентричная трансформация и лучшие практики зарубежных стран//Финансы и кредит. 2023. Т. 29, № 6(834). – С. 1339-1369.
3. Илюхин А. А., Илюхина С.В. Экономика страхового рынка: моделирование ключевых показателей//Human Progress. 2022. Т. 8, № 3. – С. 6-11.
4. Иванов Д. Ю., Ростова Е.П., Клевина М.В. Страховая компания в системе управления рисками производства//Вестник Самарского университета. Экономика и управление. 2023. Т. 14, № 3. – С. 167-176.
5. Крамчанинова, А. С. Перспективы развития страхования в аспекте цифровизации экономики/А.С.Крамчанинова, А.А.Новикова, Е.И.Выхристюк//Социальные и технические сервисы: проблемы и пути развития : Сборник статей по материалам VIII Всероссийской научно-практической конференции , Нижний Новгород, 18 ноября 2021 года. – Нижний Новгород: ФГБОУ ВО "Нижегородский государственный педагогический университет имени Козьмы Минина", 2021. – С. 177-180.
6. Джалилова Г. А., Градусова В.Н. Современные подходы к прогнозированию деятельности страховых компаний//Научное обозрение. Международный научно-практический журнал. 2020. № 4. – С. 1-6.
7. Савинская Д.Н., Кочкарова П.А., Зейн В.К, Шуняев А.А. Современные методы прогнозирования временных рядов//Современная экономика: проблемы и решения. 2021. Том 11(143). –С.56-62.
8. А. Дьяконов «Введение в анализ данных и машинное обучение». URL: https://alexanderdyakonov.files.wordpress.com/2017/06/book_boosting_pdf.pdf (Дата обращения 17.03.2024).
9. Stephen Wolfram. An Elementary Introduction to the Wolfram Language. URL: <https://www.wolfram.com/language/elementary-introduction/2nd-ed/> (дата обращения 17.03.2024).
10. Джункеев У. Прогнозирование инфляции в России на основе градиентного бустинга и нейронных сетей//деньги и кредит, 2024. Т.83. №1. –С.53-76.

References

1. Tarasova Yu. A. Insurance business. -Moscow: Yurayt Publishing House, 2024. – pp.232
2. Dryuk T.V. Target model for the development of the Russian insurance market: client-centric transformation and best practices of foreign countries // Finance and Credit. 2023. Т. 29, No. 6(834). – pp. 1339-1369.
3. Piyukhin A.A., Piyukhina S.V. Economics of the insurance market: modeling of key indicators // Human Progress. 2022. Т. 8, No. 3. – pp. 6-11.
4. Ivanov D.Yu., Rostova E.P., Klevina M.V. Insurance company in the production risk management system // Bulletin of Samara University. Economics and Management. 2023. Т. 14, No. 3. – pp. 167-176.
5. Kramchaninova, A. S. Prospects for the development of insurance in the aspect of digitalization of the economy / A. S. Kramchaninova, A. A. Novikova, E. I. Vykhristyuk // Social and

- technical services: problems and development paths: Collection of articles based on materials from VIII All-Russian Scientific and Practical Conference, Nizhny Novgorod, November 18, 2021. – Nizhny Novgorod: Federal State Budgetary Educational Institution of Higher Education “Nizhny Novgorod State Pedagogical University named after Kozma Minin”, 2021. – pp. 177-180.
6. Dzhililova G. A., Gradusova V.N. Modern approaches to forecasting the activities of insurance companies // Scientific review. International scientific and practical journal. 2020. No. 4. – pp. 1-6.
 7. Savinskaya D.N., Kochkarova P.A., Zein V.K, Shunyaev A.A. Modern methods of forecasting time series // Modern economics: problems and solutions. 2021. Volume 11(143). –pp.56-62.
 8. A. Dyakonov “Introduction to data analysis and machine learning.” URL: https://alexanderdyakonov.files.wordpress.com/2017/06/book_boosting_pdf.pdf (Accessed 03/17/2024).
 9. Stephen Wolfram. An Elementary Introduction to the Wolfram Language. URL: <https://www.wolfram.com/language/elementary-introduction/2nd-ed/> (access date 03/17/2024).
 10. Dzhunkeev U. Forecasting inflation in Russia based on gradient boosting and neural networks // money and credit, 2024. T.83. No. 1. –pp.53-76.
-