

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра математического моделирования и анализа данных

Аннотация к дипломной работе

**«МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ
ПОВЕДЕНИЯ КЛИЕНТОВ В СТРАХОВАНИИ»**

Мехович Мария Александровна

Научный руководитель – кандидат физ.-мат.
наук, доцент
В. И. Лобач

Минск, 2025

АННОТАЦИЯ

Дипломная работа, 52 страницы, 7 рисунков, 6 таблиц, 1 приложение, 11 источников.

Ключевые слова: МАШИННОЕ ОБУЧЕНИЕ, СТРАХОВАНИЕ, АНАЛИЗ ДАННЫХ, ПРОГНОЗИРОВАНИЕ, СРАВНИТЕЛЬНЫЙ АНАЛИЗ, PYTHON.

Объект исследования: сфера страхования, а именно методы прогнозирования потребительского выбора при покупке страховых продуктов.

Предмет исследования: алгоритмы машинного обучения для предсказания вероятности приобретения страховых полисов на основе данных и поведенческих характеристик клиентов.

Цель исследования: сравнить эффективность различных подходов машинного обучения и выявить оптимальный метод прогнозирования страховых продаж.

Методы исследования: разработка программных решений на Python, сравнительный анализ моделей машинного обучения.

Полученные результаты:

1. Проведено всестороннее тестирование алгоритмов (логистическая регрессия, наивный Байесовский классификатор, k-ближайших соседей, случайный лес, XGBoost) на реальных страховых данных.
2. В качестве основных ключевых параметров использовались демографические показатели (возраст, пол), истории вождения (аварии, стаж), характеристики имущества (возраст автомобиля).
3. Оценка эффективности по следующим метрикам: Accuracy, Precision, Recall, F1-score и AUC-ROC.
4. Наилучшие результаты показал алгоритм XGBoost, который может быть рекомендован как основной инструмент для прогнозирования страховых продаж, персонализации предложений, оптимизации маркетингового бюджета.

Область возможного практического применения: данная дипломная работа имеет широкий спектр применения в различных областях, требующих прогнозирования поведения клиентов на основе данных. Результаты работы могут быть использованы для оптимизации бизнес-процессов, улучшения качества обслуживания клиентов и как следствие - повышения конкурентоспособности и улучшения финансовых показателей компании.

АНАТАЦЫЯ

Дыпломная праца, 52 старонкі, 7 малюнкаў, 6 табліц, 1 дадатак, 11 крыніц.

Ключавыя словы: МАШЫННАЕ НАВУЧАННЕ, СТРАХАВАННЕ, АНАЛІЗ ДАДЗЕННЫХ, ПРАГНАЗАВАННЕ, ПАРАЎНАЛЬНЫ АНАЛІЗ, PYTHON.

Аб'ект даследавання: сфера страхавання, а менавіта метады прагназавання спажывецкага выбару пры набыванні страхавых прадуктаў.

Прадмет даследавання: алгарытмы машыннага навучання для прадказання верагоднасці набывання страхавых полісаў на аснове даных і паводніцкіх характарыстык кліентаў.

Мэта даследавання: параўнаць эфектыўнасць розных падыходаў машыннага навучання і выявіць аптымальны метады прагназавання страхавых продажаў.

Метады даследавання: распрацоўка праграмных рашэнняў на Python, параўнальны аналіз мадэляў машыннага навучання.

Вынік:

1. Праведзена ўсебаковае тэсціраванне алгарытмаў (лагістычная рэгрэсія, найўны Баесаўскі класіфікатар, k-бліжэйшых суседзяў, выпадковы лес, XGBoost) на рэальных страхавых дадзеных.
2. У якасці асноўных ключавых параметраў выкарыстоўваліся дэмаграфічныя паказчыкі (узрост, пол), гісторыя ваджэння (аварыі, стаж), характарыстыкі маёмасці (узрост аўтамабіля).
3. Ацэнка эфектыўнасці па наступных метрыках: Accuracy, Precision, Recall, F1-score і AUC-ROC.
4. Найлепшыя вынікі паказаў алгарытм XGBoost, які можа быць рэкамендаваны як асноўны інструмент для прагназавання страхавых продажаў, персаналізацыі прапановаў, аптымізацыі маркетынгавага бюджэту.

Вобласць магчымага практычнага прымянення: дадзеная дыпломная праца мае шырокі спектр прымянення ў розных галінах, якія патрабуюць прагназавання паводзінаў кліентаў на аснове дадзеных. Вынікі працы могуць быць скарыстаны для аптымізацыі бізнэс-працэсаў, паляпшэнні якасці абслугоўвання кліентаў і як следства - павышэнні канкурэнтаздольнасці і паляпшэнні фінансавых паказчыкаў кампаніі.

ANNOTATION

Diploma thesis, 52 pages, 7 illustrations, 6 spreadsheets, 1 application, 11 sources.

Keywords: MACHINE LEARNING, INSURANCE, DATA ANALYSIS, FORECASTING, COMPARATIVE ANALYSIS, PYTHON.

The object of the research: the insurance industry, namely methods for predicting consumer choices when purchasing insurance products.

The subject of the research: machine learning algorithms for predicting the likelihood of purchasing insurance policies based on personal data and behavioral characteristics of clients. various customer characteristics and their behavior.

The purpose of the research: compare the effectiveness of different machine learning approaches and identify the optimal method for forecasting insurance sales.

Methods of research: development of software solutions in Python, comparative analysis of machine learning models.

The results of the work:

1. Comprehensive testing of algorithms (logistic regression, naive Bayes classifier, k-nearest neighbors, random forest, XGBoost) was conducted on real insurance data.

2. Demographic indicators (age, gender), driving history (accidents, experience), property characteristics (vehicle age) were used as the main key parameters.

3. Evaluation of effectiveness by the following metrics: Accuracy, Precision, Recall, F1-score and AUC-ROC.

4. The best results were shown by the XGBoost algorithm, which can be recommended as the main tool for forecasting insurance sales, personalizing offers, and optimizing the marketing budget.

Recommendations on the usage: This thesis has a wide range of applications in various fields that require forecasting customer behavior based on data. The results of the work can be used to optimize business processes, improve the quality of customer service and, as a result, increase competitiveness and improve the financial performance of the company.