

УДК 004.65:336.221

ИСПОЛЬЗОВАНИЕ BIG DATA В НАЛОГООБЛОЖЕНИИ



Т.В. Казак

Заведующий кафедрой инженерной психологии и эргономики, доктор психологических наук Республики Беларусь, доктор психологических наук Российской Федерации, профессор, БГУИР
kazak@bsuir.by



Е.Д. Ракицкий

Студент БГУИР, кафедра инженерной психологии и эргономики
egor.rakitskiy@gmail.com



А.Н. Василькова

Старший преподаватель кафедры инженерной психологии и эргономики, БГУИР
a.vasilkova@bsuir.by

Т.В. Казак

Заведующий кафедрой инженерной психологии и эргономики, доктор психологических наук Республики Беларусь, доктор психологических наук Российской Федерации, член-корреспондент Международной академии психологических наук, профессор.

Образование: Высшее; Аспирантура, специальность: 19.00.03 - психология труда, инженерная психология, эргономика; Докторантура, специальность: 19.00.03 - психология труда, инженерная психология, эргономика.

Область профессиональных интересов / исследований: Психология труда. Инженерная психология. Эргономика. Психология управления. Юридическая психология. Социальная и организационная психология. Клиническая психология.

Е.Д.Ракицкий

Студент кафедры инженерной психологии и эргономики БГУИР.

Область профессиональных интересов / исследований: языки программирования, искусственный интеллект, технологии виртуальной реальности.

А.Н. Василькова

Старший преподаватель кафедры инженерной психологии и эргономики.

Образование: 2007 - МГВРК по специальности «Программное обеспечение информационных технологий»,

2022 - магистратура БГУИР по специальности «Охрана труда и эргономика».

Область профессиональных интересов / исследований: языки программирования, искусственный интеллект, технологии виртуальной реальности.

Аннотация. Целью данного проекта является исследование потенциала *Big Data* в улучшении эффективности налогообложения и предотвращении налоговых уклонений. В проекте будет рассмотрен фреймворк *Keras*, как один из инструментов для обработки и анализа больших объемов данных.

Ключевые слова: налогообложение, *Big Data*, технологии, *Keras*, уплата налогов.

Введение. В современном мире использование *Big Data* и искусственного интеллекта (ИИ) открывает новые возможности во многих областях, включая

налогообложение. В рамках данного проекта предусмотрено сосредоточение на использовании этих технологий для анализа контента в социальных сетях с целью выявления лиц с более высоким экономическим статусом, чем указано в их налоговых декларациях.

Рассматриваются 3 основных налога на имущество: транспортный, земельный и налог на недвижимость.

В этом году оплата 3 платежей производится более чем 3,5 миллиона физических лиц. На сегодняшний день поступление всех 3 платежей обеспечены уже более чем на 90%, однако остаётся около 10% неплательщиков

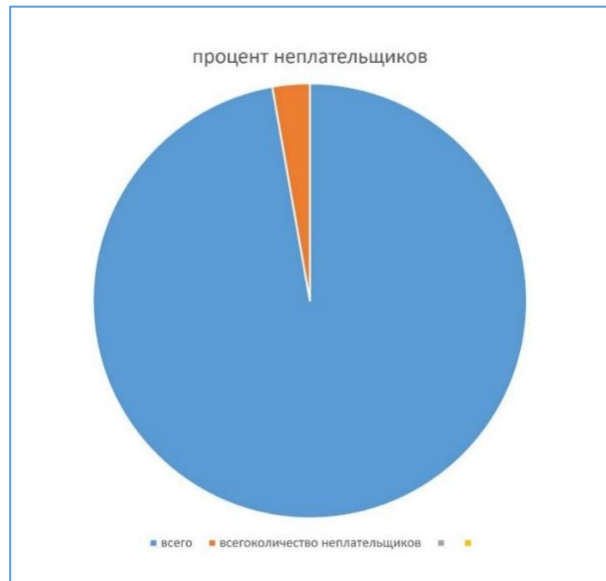


Рисунок 1. Диаграмма, показывающая количество неплательщиков

Ежедневно за неуплату налогов начисляется пеня в размере 0,033% от неуплаченной суммы. Тем самым в год государство теряет более 20 млн. рублей. К примеру, за 2023 год была рекордная сумма по неуплате налогов у интернет-магазина в размере 9,5 млн рублей.

Социальные сети являются богатым источником данных, которые могут быть использованы для получения ценной информации о финансовом положении пользователей. Однако, эти данные часто не структурированы и требуют сложного анализа. С помощью алгоритмов машинного обучения и анализа больших данных существует возможность обработать и проанализировать большие объемы информации из социальных сетей, выявить закономерности и тенденции, а также определить лиц, чей образ жизни не соответствует их заявленным доходам.

Основная часть. Понятия *Big Data*, аналитика данных и искусственный интеллект не являются чем-то новым. Тем не менее, некоторые технологические достижения сделали возможным их интенсивное использование в бизнесе и государственном управлении за последние несколько лет. Основные достижения касаются: колоссального расширения вычислительной мощности и хранения данных на компьютерах, связанного со снижением их стоимости; растущая доступность сетей связи и широкополосного интернета; разработка эффективных моделей сбора, хранения и обработки больших массивов данных и продвинутых когнитивных алгоритмов; появление новых источников данных (например, датчиков, *GPS*, социальных сетей и т.д.), в том числе электронных счетов-фактур, а также обмен фискальной информацией между странами.

Обработка данных начинается с *Big Data*, которая основана на концепции пяти «V»: Объем: относится к большому объему генерируемых данных; Разнообразие: источники

данных очень разнообразны, что повышает сложность анализа; Скорость: из-за большого объема и разнообразия данных вся обработка должна быть гибкой для генерации необходимой информации; Достоверность: напрямую связана с достоверностью данных; Значение: относится к «полезной информации», которую можно получить из данных.

Большинство решений для работы с большими данными включают в себя некоторые из следующих компонентов: источники данных, хранилища данных, обработка лотов, потребление данных в режиме реального времени, потоковая обработка, аналитическое хранилище данных, отчеты и оркестратор.

Сложно определить точную границу между *Big Data* и аналитикой данных или между аналитикой данных и искусственным интеллектом.

Большие данные и аналитика предоставляют метод поиска аномалий, который может привести к предсказанию того, где эти аномалии могут произойти снова.

Покупки в Интернете, электронные и карточные транзакции, оплаченные офлайн-покупки, а также реакция на налоговые уведомления – все это точки данных в налоговом путешествии физического лица, которые можно использовать для поведенческой аналитики налогоплательщика. Большие данные и аналитика в сочетании с искусственным интеллектом могут работать еще лучше, чтобы помочь регулирующим органам значительно улучшить соблюдение налогового законодательства.

Например, исторические данные о реакции налогоплательщиков на получение налоговых уведомлений могут быть использованы для прогнозирования их следующего шага. Предиктивная аналитика также может разрабатывать сложные профили рисков, анализировать тенденции, выявлять потенциальные проблемы аудита и выявлять случаи с более высоким риском для более глубокого расследования, потенциально отсекая пути для мошенничества еще до того, как оно произойдет.

Актуальность – процесс оптимизации различных процессов в сфере налогообложения, следовательно, помогает наладить экономику и стабильность государства. Следует выделить 4 основных пункта, связанных с актуальностью программного продукта:

1 Борьба с налоговыми мошенничествами: Эта программа позволяет выявлять случаи скрытого дохода, не декларированной прибыли и других налоговых преступлений. Автоматический анализ данных из различных источников позволяет выявить несоответствия между заявленными доходами и фактическим финансовым положением. Это помогает налоговым службам активно бороться с налоговыми мошенничествами и улучшить собираемость налогов.

2 Улучшение бюджетной дисциплины: программа позволяет контролировать доходы и расходы большого количества физических и юридических лиц. Программное средство помогает выявить и предотвратить случаи неправильного учета доходов или занижения налоговых обязательств. Это способствует более справедливому и эффективному распределению налоговых средств, что, в свою очередь, улучшает бюджетную дисциплину и общую финансовую стабильность.

3 Оптимизация работы налоговых служб: программа с использованием *Big Data* значительно сокращает ручную работу и повышает эффективность налоговых служб. Благодаря автоматизации процессов анализа больших объемов данных, налоговые службы могут быстрее и точнее выявлять налоговые нарушения. Это позволяет сосредоточить ресурсы налоговых служб на более сложных случаях и скорее принять меры по пресечению нарушений.

4 Повышение финансовой прозрачности: программа позволяет государству и обществу иметь более полную картину о финансовом положении граждан и организаций. Это способствует повышению финансовой прозрачности и уменьшению возможностей для расчётливого уклонения от налогообложения. Более прозрачная налоговая система

способствует развитию бизнеса, привлечению инвестиций и повышению уровня доверия в обществе.

Технологии *Big Data* в налогообложении. Технологии *big data* играют важную роль в сфере налогообложения, позволяя эффективно обрабатывать и анализировать большие объемы данных для выявления налоговых нарушений и оптимизации процессов. Одна из ключевых технологий, применяемых в программе *big data* для налогообложения, это машинное обучение. Машинное обучение позволяет автоматически обрабатывать и анализировать огромные объемы данных, выявлять паттерны и тренды, а также строить модели для прогнозирования и выявления аномалий. На основе этих моделей и алгоритмов, программы *big data* могут автоматически выявлять случаи несоответствия между заявленными доходами и реальным финансовым положением, а также определять потенциальные налоговые преступления. Другая важная технология *big data* для налогообложения – это облачные вычисления. Облачные платформы позволяют хранить и обрабатывать огромные объемы данных, предоставляя высокую производительность и масштабируемость. Они также обеспечивают гибкость и доступность для пользователей, что позволяет налоговым службам быстро и эффективно обрабатывать большие объемы данных. Технологии обработки естественного языка (*Natural Language Processing, NLP*) также широко применяются в программе *big data* для налогообложения. *NLP* позволяет программам анализировать и интерпретировать текстовую информацию, например, декларации о доходах или сообщения в социальных сетях. Автоматический анализ текста позволяет находить ключевые слова, фразы и смысловые связи, что помогает налоговым службам более точно определить налоговые обязательства и выявить возможные налоговые нарушения. Технологии визуализации данных также имеют важное значение в программе *Big Data* для налогообложения. Они позволяют представлять сложные и многомерные данные в наглядной и понятной форме. Визуализация данных помогает налоговым службам обнаруживать связи и тренды, а также принимать информированные решения на основе визуальных представлений данных.

Заключение. В заключении проекта по использованию *Keras, Big Data* и искусственного интеллекта в налогообложении, можно отметить следующее:

1 Проект успешно демонстрирует, как современные технологии могут быть использованы для повышения эффективности налогового администрирования. Использование *Keras* и ИИ позволяет обрабатывать большие объемы данных, собранных из социальных сетей, для выявления лиц, уклоняющихся от уплаты налогов.

2 Алгоритмы машинного обучения, основанные на *Keras*, способны обрабатывать и анализировать данные, что позволяет выявлять потенциальных налоговых уклонистов. Это, в свою очередь, помогает налоговым органам оптимизировать процесс сбора налогов и повысить уровень налоговой дисциплины.

Однако, необходимо учесть, что использование таких технологий требует строгого соблюдения норм законодательства в области защиты персональных данных. Все действия должны быть направлены на обеспечение конфиденциальности и безопасности информации.

В целом, проект показывает большой потенциал использования *Keras, Big Data* и ИС в налогообложении, открывая новые возможности для повышения эффективности работы налоговых органов.

Список литературы

- [1] Kluwer International Tax Blog [Electronic resource]. – <https://kluwertaxblog.com/2021/07/16/big-data-in-tax-administrations/>
- [2] Artificial intelligence in tax administrations: Benefits and risks of its use [Electronic resource], – <https://www.gccfintax.com/articles/artificial-intelligence-in-tax-administrations-benefits-and-risks-of-its-use-1685.aspx>
- [3] Revolutionizing Taxation: How India Uses AI & ML to Improve the Tax Process [Electronic resource]. – <https://www.analyticsvidhya.com/blog/2023/04/how-india-uses-data-analytics-ai-ml-to-improve-tax-process/>
- [4] Белорусский информационный портал Sputnik [Electronic resource]. – <https://sputnik.by/20221207/mns-poschitalo-skolko-belorusov-zaplatili-nalogi-na-imuschestvo-1069872912.html>

Авторский вклад

Казак Тамара Владимировна – руководство и постановка задачи исследования BIG DATA в налогообложении РБ.

Василькова Анастасия Николаевна – постановка задачи исследования, описание принципа работы Big Data в улучшении эффективности налогообложения и предотвращении налоговых уклонений, анализ полученных результатов, формирование структуры статьи.

Ракицкий Егор Дмитриевич – тестирование программного средства, описание принципов использования Keras, Big Data и искусственного интеллекта в налогообложении, формирование структуры статьи.

USING BIG DATA IN TAXATION

T.V. Kazak

Head of the Department of Engineering Psychology and Ergonomics, doctor of psychological sciences of the Republic of Belarus, doctor of psychological sciences of the Russian Federation, Professor, Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus.

E.D. Rakitsky

BSUIR student, Department of Engineering Psychology and Ergonomics

A.N. Vasilkova

Senior Lecturer, Department of Engineering Psychology and Ergonomics, BSUIR

Abstract. The goal of this project is to explore the potential of Big Data to improve tax efficiency and prevent tax evasion. The project will consider the Keras framework as one of the tools for processing and analyzing large amounts of data.

Keywords: taxation, Big Data, technology, Keras, tax payment.