

DOI: 10.12731/2070-7568-2023-12-2-172-190

УДК 334:316:004.8



Научная статья |

Математические, статистические и инструментальные методы в экономике

ИССЛЕДОВАНИЕ ВЗАИМОСВЯЗИ ОБРАЩЕНИЙ ГРАЖДАН НА ОТКРЫТЫЙ ПОРТАЛ ПРАВИТЕЛЬСТВА МОСКВЫ С СОЦИАЛЬНО-ЭКОНОМИЧЕСКИМИ ПОКАЗАТЕЛЯМИ АДМИНИСТРАТИВНЫХ ОКРУГОВ

П.А. Сахнюк, Т.И. Сахнюк

В статье рассматриваются возможности изучения состояния социальной сферы по данным репозитория портала открытых данных Правительства Москвы по административным округам и городским округам с использованием платформ бизнес-аналитики и интеллектуальных технологий платформ Data Science и машинного обучения. Представлены возможности использования технологий машинного обучения для платформ бизнес-аналитики для выявления скрытых закономерностей с целью принятия обоснованных управленческих решений.

***Цель** – эмпирическая проверка гипотезы о положительной взаимосвязи между количеством семей, получающих субсидию в городе Москве, в различных разрезах, и числом обращений граждан на открытый портал Правительства Москвы*

***Метод или методология проведения работы:** в статье использовались метод кластеризации и статистические методы исследования.*

***Область применения результатов:** полученная в результате мониторинга объективная информация может быть использована для выработки стратегии и принятия управленческих решений по развитию округов и повышению качества жизни г. Москвы*

***Ключевые слова:** платформы науки о данных и машинного обучения; платформы бизнес-аналитики; искусственный интеллект; технологии дополненной аналитики*

***Для цитирования.** Сахнюк П.А., Сахнюк Т.И. Исследование взаимосвязи обращений граждан на открытый портал Правительства Москвы с социально-экономическими показателями административных округов // Наука Красноярья. 2023. Т. 12, №2. С. 172-190. DOI: 10.12731/2070-7568-2023-12-2-172-190*

Original article | Mathematical, Statistical and Instrumental Methods in Economics

STUDY OF THE RELATIONSHIP OF CITIZENS' APPEALS TO THE OPEN PORTAL OF THE GOVERNMENT OF MOSCOW WITH THE SOCIO-ECONOMIC INDICATORS OF THE ADMINISTRATIVE DISTRICT

P.A. Sakhnyuk, T.I. Sakhnyuk

The article discusses the possibilities of studying the state of the social sphere according to the repository of the open data portal of the Government of Moscow for administrative districts and urban districts using business intelligence platforms and intelligent technologies of Data Science and machine learning platforms. The possibilities of using machine learning technologies for business intelligence platforms to identify hidden patterns in order to make informed management decisions are presented.

Purpose – is to familiarize with business intelligence platforms; the use of intelligent BI-platform technologies for monitoring socio-economic indicators.

Method or methodology of the work: the clustering method and statistical research methods were used in the article.

Results: to confirm the hypothesis of a positive relationship between the number of families receiving subsidies in the city of Moscow, in various sections, and the number of appeals of citizens to the open portal of the Government of Moscow

Practical implications: the objective information obtained as a result of monitoring can be used to develop a strategy and make management decisions on the development of districts and improving the quality of life in Moscow

Keywords: data science and machine learning platforms; business intelligence platforms; machine learning; artificial intelligence; augmented analytics technologies

For citation. Sakhnyuk P.A., Sakhnyuk T.I. Study of the Relationship of Citizens' Appeals to the Open Portal of the Government of Moscow with the Socio-Economic Indicators of the Administrative District. *Krasnoyarsk Science*, 2023, vol. 12, no. 2, pp. 172-190. DOI: 10.12731/2070-7568-2023-12-2-172-190

Введение

Цифровая трансформация государственного и муниципального управления привела к тому, что в распоряжении властей оказывается

большой объем данных, эффективное использование которых, может существенно повысить качество принимаемых решений [17-21]. Однако в настоящее время при сборе и анализе данных, получаемых из баз служб государственной статистики, порталов открытых данных и других источников, существует проблема получения качественной и наглядной информации.

В результате информация, которой располагают органы исполнительной власти, и отдельные граждане носит во многом экспертный, эвристический характер, в том числе и в такой важной сфере, как мониторинг социально-экономических показателей административных округов и районов Москвы. Поэтому возникает объективная необходимость в применении интеллектуальных технологий VI-платформ для решения этой задачи [1, 2].

Технологии дополненной аналитики, основанные на машинном обучении, дают возможность анализировать большие объем данных специалистам предметной области, без необходимости глубоких компетенций в области компьютерных наук. Эти технологии включает в себя обработку естественного языка, как способа запроса данных и создания повествований для объяснения основных статистик, драйверов роста (снижения) и визуализаций. Уже сейчас до половины всех аналитических запросов в VI-платформах формируются с помощью поиска, обработки на естественном языке, или генерируются автоматически. Обработка естественного языка и разговорная аналитика увеличивают распространение бизнес-аналитики среди “гражданских датасаентистов” (“Citizen Data Science”), включая новые классы пользователей, особенно сотрудников фронт-офисов и муниципалитетов города.

Цель исследования: выявление взаимосвязей между количеством обращений граждан на открытый портал правительства г. Москва и социально-экономическими показателями развития административных округов. Нами выдвинута гипотеза о том, что между количеством обращений и значениями социально-экономических показателей существует положительная связь.

Материалы и методы исследования

В работе использовались статистические методы. Для интеллектуального анализа были выбран набор данных по количеству семей, получающих субсидию в городе Москве, в различных разрезах и данные по обращению граждан на портал открытых данных (рис. 1) [3].

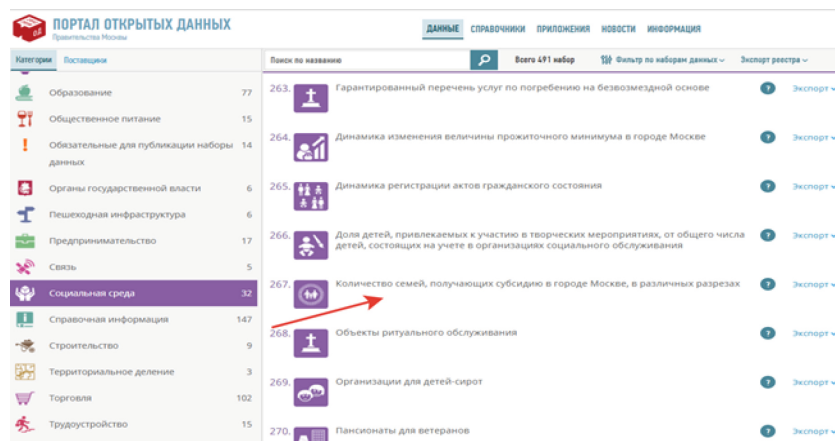


Рис. 1. Категория: «Социальная среда», данные: «Количество семей, получающих субсидию в городе Москве, в различных разрезах»

Результаты: на рисунке 1 показана коллекция различных визуализаций, созданных в службе Power BI.

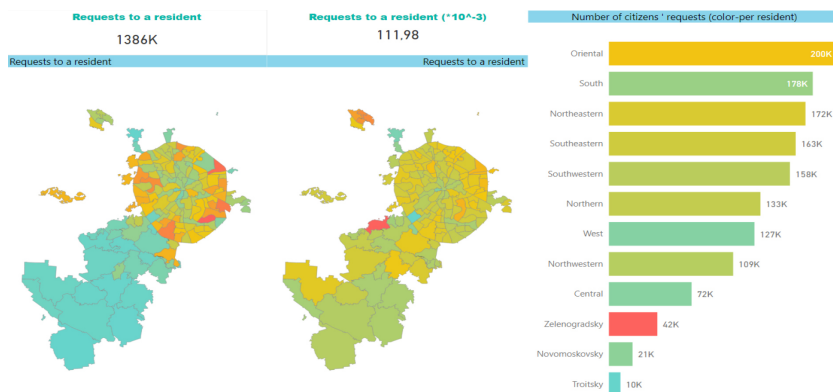


Рис. 2. Дашборд числа обращений граждан на портал открытых данных (в абсолютных единицах и приходящиеся на одного жителя административных округов и районов города)

Найдем “похожие” объекты среди 146 районов Москвы. Поиск схожих объектов – одна из наиболее часто встречающаяся задача в анализе данных. Сходство определим на основе признаков набора данных «Количество семей, получающих субсидию в городе Москве, в различных разрезах»: об-

щее количество семей, малообеспеченных семей, многодетных семей, неполных семей, семей пенсионеров, семей студентов, семей безработных и количество обращений на портал открытых данных. В Power BI поддержка кластеризации предоставляет мощные аналитические возможности, Power BI поддерживает алгоритм k-means, наиболее популярный метод кластеризации. Одной из самых сложных задач в кластеризации является определение количества кластеров (рис. 3). Для облегчения этой задачи Power BI предоставляет как автоматические, так и ручные опции для контроля.



Рис. 3. Результаты кластеризации районов Москвы алгоритмом k-means Power BI

Для проверки полученных результатов проведем кластеризацию районов Москвы с помощью других алгоритмов кластеризации. В качестве инструмента кластерного анализа будем использовать нейронную сеть Кохонена, реализуемой в аналитической платформе Deductor Studio, достоинством которой по сравнению с другими алгоритмами является возможность визуального анализа многомерных данных: схожие объекты попадают в соседние ячейки карты (рис. 4).

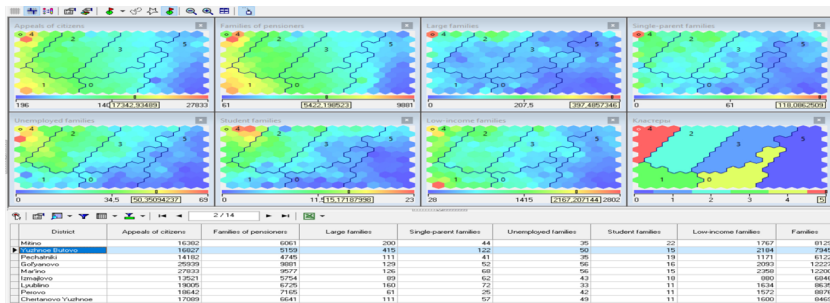


Рис. 4. Карты Кохонена по набору

«Количество семей, получающих субсидию в Москве, в различных разрезах»
(в кластере 4 показаны значения для района Южное Бутово)

Проверку результатов кластеризации методом карт Кохонена и k-means осуществим в KNIME Analytics Platform (рис. 5), в которой легко выполнить нечеткую *c-means* (FCM) кластеризацию, особенностью которой является отнесение каждой точки данных к кластеру с функцией принадлежности, изменяемой в диапазоне от 0 до 1 включительно [4-8]. Кластеры представляются нечеткими множествами, и, кроме того, границы между кластерами также являются нечеткими. Степень принадлежности определяется расстоянием от объекта до соответствующих центроидов.

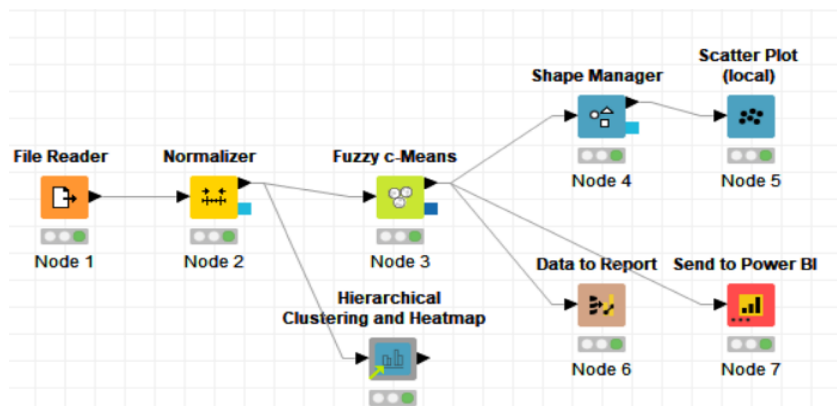


Рис. 5. Workflow FCM кластеризации in the KNIME Analytics Platform

S District	S Winner Cluster	D cluster_4	D cluster_0	D cluster_1	D cluster_2	D cluster_3	D cluster_5
Yasenevo	cluster_4	█	█				
Gol'yanovo	cluster_4	█	█				
Xoroshyovo-Mnevniki	cluster_4	█	█				
Mar'ino	cluster_4	█	█				
Ivanovskoe	cluster_4	█	█				
Otradnoe	cluster_4	█	█				
Mitino	cluster_4	█	█				
Chertanovo Yuzhnoe	cluster_4	█	█				
Lyubino	cluster_4	█	█				
Yuzhnoe Butovo	cluster_4	█	█				
Perovo	cluster_0		█				
Vykhino-Zhulebino	cluster_0		█				
Bibirevo	cluster_2			█			
Izmailovo	cluster_2			█			
Tekstil'shiki	cluster_2			█			
Pechatniki	cluster_2			█			

Рис. 6. Результаты FCM кластеризации административных округов Москвы в KNIME Analytics Platform

Результаты нечеткой кластеризации представлены на рисунке 6. Сравнивая их с результатами кластеризации методом карт Кохонена построенных в Deductor Studio и k-means Power BI (рис. 3, 4), убеждаемся в их объективности. Отличие в результатах наблюдается только для районов Перово, Измайлово, Текстильщики и Печатники: алгоритм FCM, отнес

их ко второму кластеру при незначительной разнице в значениях функции принадлежности.

Взаимосвязи между признаками нашего набора данных, отражающего социальное положение жителей районов города исследуем с помощью дашбордов Power BI включающих матрицу корреляции исследуемых признаков, точечную диаграмму численности населения районов кластера и числа обращений граждан на портал открытых данных, карту районов города (рис. 7).

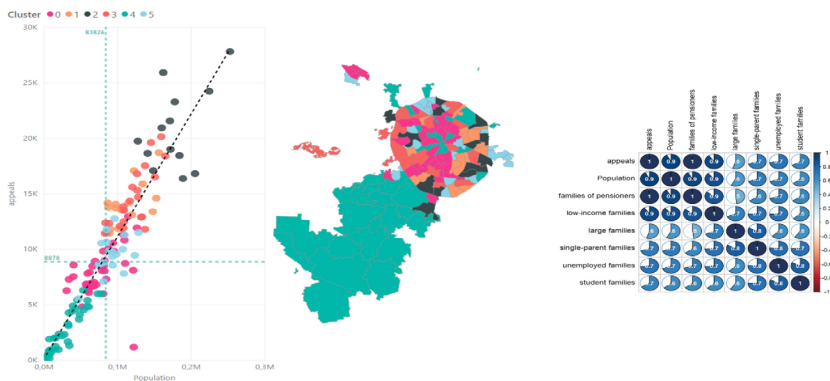


Рис. 7. Дашборд Power BI по районам Москвы в разрезе найденных кластеров

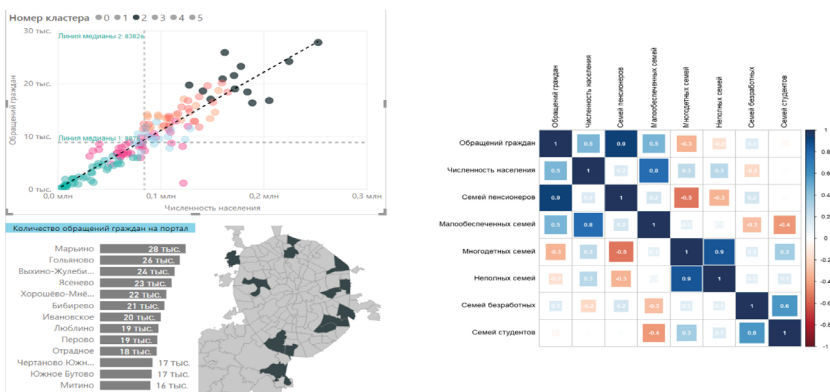


Рис. 8. Дашборд Power BI по районам Москвы 2-го кластера

Для более глубокого исследования данных используем платформу Qlik Sense, которая комбинирует ассоциативные исследования в свободной форме, предоставляет контекстно-зависимые предложения и автоматическим

построенные визуализации, основанные на искусственном интеллекте. Платформа использует Qlik Cognitive Engine для предоставления предложений по автоматическому анализу, которые помогают по-новому взглянуть на данные, автоматически генерируя и определяя приоритеты аналитики на основе общего набора данных и критериев поиска пользователя. Augmented intelligence платформы предлагает понимание и автоматизацию визуальной аналитики, понимая вопросы и отвечая на естественном языке.

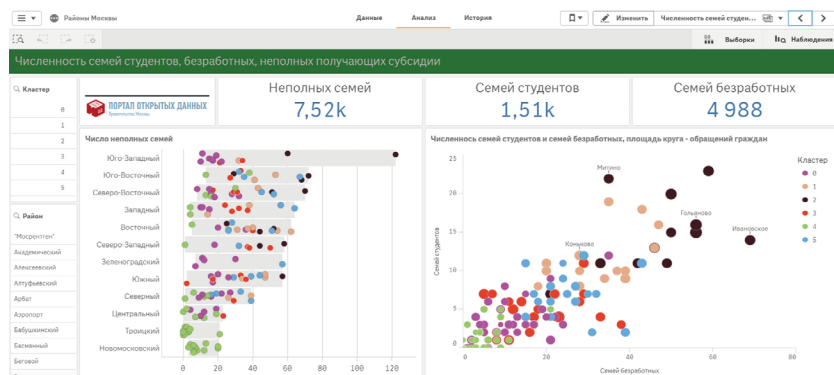


Рис. 9. Дашборд Qlik Sense: распределение районов Москвы в зависимости от взаимосвязи численности семей студентов, семей безработных и неполных семей, получающих субсидию и числа обращений на портал открытых данных в разрезе найденных кластеров

Анализ дашборда (рис. 9) показывает, что по мере увеличения количества семей безработных, неполных семей и семей студентов, получающих субсидию, увеличивается и число обращений граждан на портал открытых данных (характерно для районов второго кластера).

Дополнительную информацию о формировании кластеров можно извлечь, используя инструмент “дерево решений” в Deductor Studio (figure 9). Так например, если получающих субсидии семей пенсионеров ≥ 4488 и неполных семей ≥ 14 и малообеспеченных семей ≥ 1434 , тогда районы относятся ко второму кластеру.

Для выявления ключевых факторов влияния будем использовать соответствующий визуализатор в Power BI который поможет понять, какие факторы влияют на исследуемую метрику (в нашем случае номер кластера, к которому отнесен район Москвы). Применение в Power BI машинного обучения помогает исследовать данные, позволяя выполнять их глубокий анализ для автоматического поиска шаблонов, понятной интерпретации и

прогнозирования результатов. С помощью машинного обучения (модель регрессионной модели) факторы влияния ранжируются от наиболее к наименее значимому, а для объяснения влияния предоставляется показатель вероятности и текстовое описание. Так для районов кластера 4 (рис. 10) наиболее значимым фактором является число семей студентов, получающих субсидию.

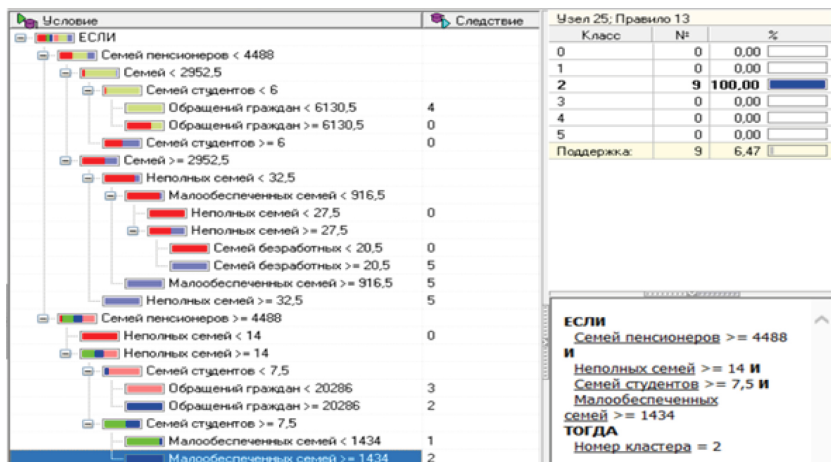


Рис. 10. Дерево решений формирования кластеров и интерпретация правил, сформированных алгоритмом C4.5 для районов кластера 2

Для районов 0-го кластера наиболее значимым фактором является количество семей безработных, получающих субсидию. Для районов 1-го кластера наиболее значимым фактором является количество неполных семей, получающих субсидию. Для районов 3-го кластера наиболее значимым фактором является количество семей студентов, получающих субсидию. Для районов 4-го кластера наиболее значимым фактором является количество семей, получающих субсидию. Для районов 5-го кластера наиболее значимым фактором является количество семей безработных, получающих субсидию.

Значимость числа малообеспеченных семей, получающих субсидии для районов 2-го кластера подтверждается визуализацией профилей кластеров, полученных с помощью алгоритмов k-means и карт Кохонена (рисунок 11). Здесь значимость атрибутов показывает их степень влияния на образования кластера. Атрибуты, по которым проведен кластерный анализ ранжированы по убыванию значимости, показаны их статистические характеристики.

Key influencers Top segments



What influences Clusters to be Cluster4 ?

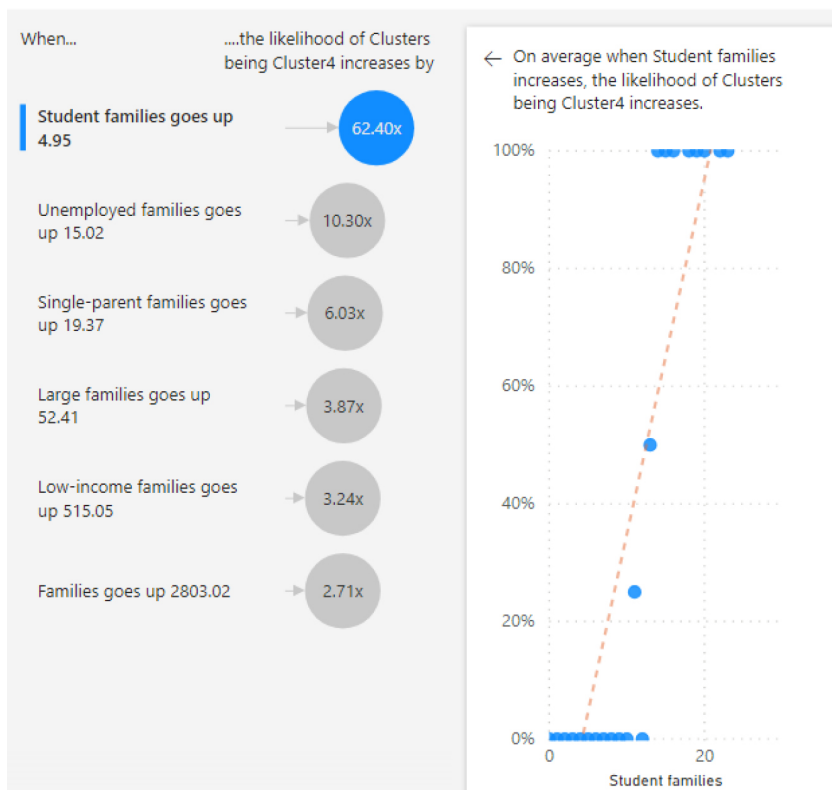


Рис. 10. Визуализатор поиска ключевых факторов влияния Power BI районов 4-го кластера

Одним из самых быстрых способов получить ответ из данных – это задать вопрос на естественном языке. Эту возможность предоставляет функция вопросов и ответов в Power BI. Q&A – это интерактивно, часто один вопрос приводит к другим, поскольку визуализация открывает интересные пути для достижения цели. В службе Power BI панель мониторинга содержит листы, закрепленные в одном или нескольких наборах данных, поэтому можно задавать вопросы о любых данных, содержащихся в наборе данных. Инструмент Q&A распознает введенные слова и выясняет, где

(в каком наборе данных) найти ответ. Q&A также поможет сформировать вопрос с автозаполнением, пересчетом и другими текстовыми и наглядными подсказками. Ответ на вопрос отображается в виде интерактивной визуализации и обновляется по мере изменения вопроса.

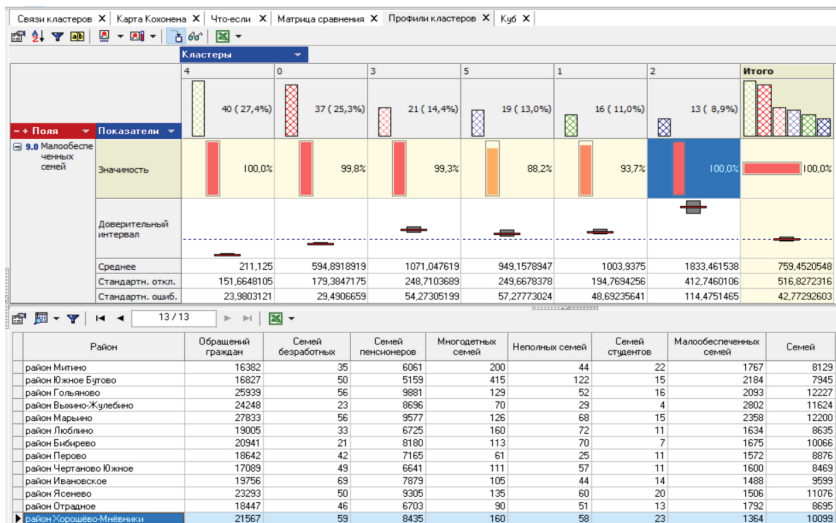


Рис. 11. Визуализатор профили кластеров платформы Deductor Studio (кластер 2)

Подтверждается высокая корреляционная связь между количеством семей, получающих субсидию и числом обращений граждан на портал открытых данных (рис. 12).

Для выявления скрытых закономерностей в данных и автоматизации выполнения основных задач машинного обучения используем настраиваемую платформу искусственного интеллекта H2O Driverless AI. Платформа предоставляет автоматический инжиниринг функций, проверку и настройку моделей, выбор и развертывание моделей, интерпретацию машинного обучения, создание собственных сценариев в построении модели, обработку временных рядов и текстов, автоматическую генерацию конвейеров для скоринга моделей.

H2O Driverless AI выполняет функцию проектирования входных признаков и выходной переменной, чтобы определить их оптимальное представление для построения модели машинного обучения. Различные этапы преобразований функций появляются на протяжении всех итераций.

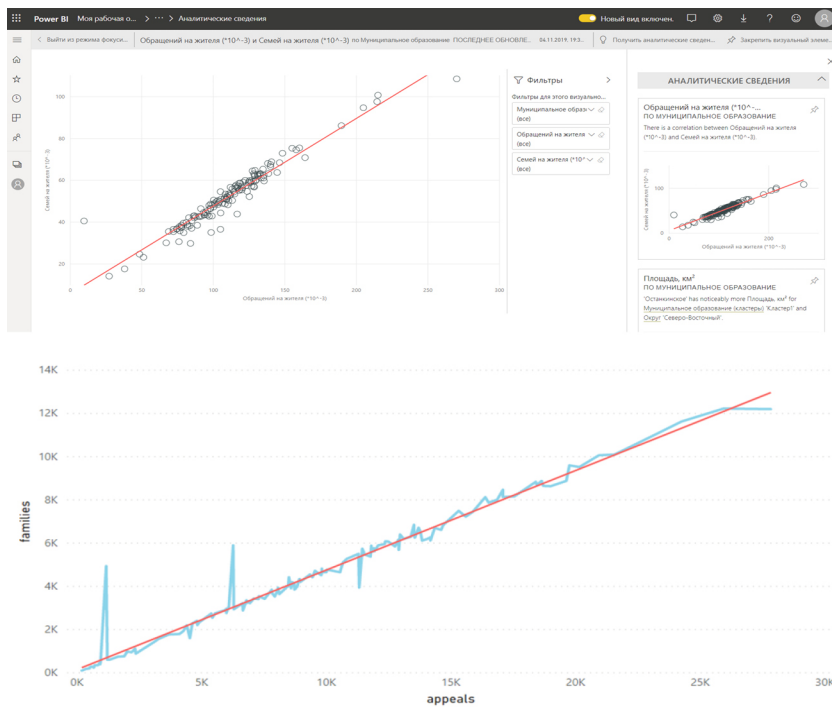


Рис. 12. Визуализатор краткой аналитики для поиска тенденций в наборе данных: точечная диаграмма, демонстрирующая связь между количеством семей, получающих субсидию, и количеством граждан, обратившихся на портал открытых данных

Их можно просмотреть, наведя указатель мыши на точки итерационных данных, в разделе Variable Importance (рис. 13). Преобразования в H2O Driverless AI применяются к столбцам набора данных. Числовые, категориальные, временные и текстовые преобразователи осуществляют проектирование функций, обеспечивающих работу алгоритмов машинного обучения. Проектирование функций является основополагающим, наиболее сложным, и дорогостоящим процессом в машинном обучении.

Наилучшей в нашем эксперименте признана модель LightGBM - фреймворк градиентного бустинга, разработанная Microsoft и использующий алгоритмы обучения на основе деревьев решений. Он был специально разработан для уменьшения использования памяти, повышения скорости обучения и повышения эффективности.

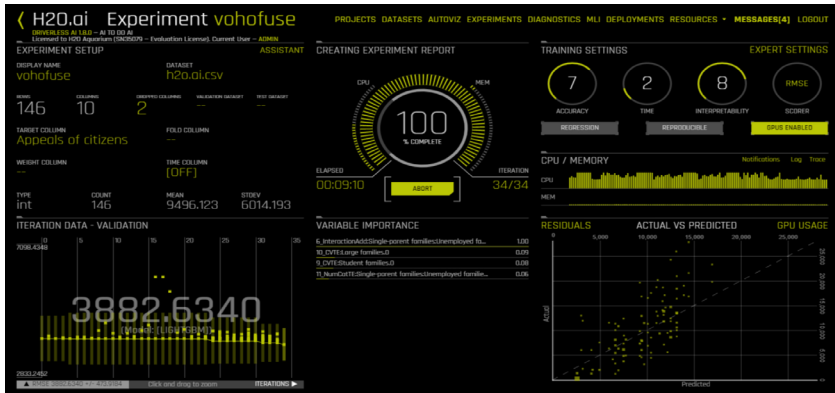


Рис. 13. Дашборд эксперимента H2O Driverless AI

Подобно XGBoost, это одна из лучших доступных реализаций градиентного бустинга [9-13]. Он также используется для подгонки моделей Random Forest внутри Driverless AI. Результаты находятся в сводке эксперимента в правой нижней части страницы эксперимента. После того, как найдена лучшая прогнозная модель, мы можем её интерпретировать. MLI Dashboard представляет различные типы объяснения, касающиеся модели и ее результатов [14-16]. Все графики на дашборде являются интерактивными (рис. 14).



Рис. 14. Дашборд интерпретации модели машинного обучения

Представленное на дашборде дерево решений отображает приближительную блок-схему модели принятия решений сложной моделью H2O Driverless AI. Более высокие и частые функции более важны. Предикторы

находящиеся выше или ниже друг друга могут указывать на взаимосвязь между ними, самые толстые ребра – это наиболее распространенные пути принятия решений в дереве, которые приводят к предсказанному числовому результату. Наибольшее значение для формирования модели предсказания числа обращений граждан на портал открытых данных имеют: количество неполных семей, семей безработных, семей студентов и многодетных семей, получающие субсидии. Причем, вклад этих признаков в модель различный. В силу высокой корреляционной зависимости с выходной переменной признаки: количество семей, семей пенсионеров и малообеспеченных семей, получающих субсидию, не участвуют в формировании модели. Также H2O Driverless AI позволяет загружать автоматически сгенерированные документы, такие как «Скачать отчет об эксперименте» и «Отчет о MLI», одним нажатием кнопки. Разработанную модель можно применять для прогноза числа обращений граждан на портал Правительства Москвы.

Выводы

Полученные результаты подтверждают гипотезу о положительной взаимосвязи между количеством семей, получающих субсидию в городе Москве, в различных разрезах, и числом обращений граждан на открытый портал Правительства Москвы. Это свидетельствует о возможности влиять на механизм участия активных москвичей в жизни города. Полученная объемная информация может быть использована для выработки стратегии и принятия решений по развитию города.

Список литературы

1. Фролов Ю.В., Сахнюк П.А., Сахнюк Т.И. Использование инструментов геоаналитики для подготовки бакалавров и магистров по направлению «бизнес-информатика» // Вестник Российского университета дружбы народов. Серия: Информатизация образования. 2018. Т. 15. №2. С. 151-164.
2. Мамаев И.И., Сахнюк Т.И., Сахнюк П.А. Анализ основных социально-экономических показателей районов Ставропольского края за 2011-2015 годы средствами систем бизнес-аналитики // Научный журнал КубГАУ [Электронный ресурс]. Краснодар: КубГАУ, 2016. №123(09). <http://ej.kubagro.ru/2016/09/pdf/26.pdf>
3. Количество семей, получающих субсидию в городе Москве, в различных разрезах. <https://data.mos.ru/opendata/7701236617-kolichestvo-semey-poluchayushchih-subsidiyu-v-gorode-moskve>

4. Siddiqui T., Ausaf A. Data mining tools and techniques for mining software repositories: A systematic review // *Big Data Analytics*. Springer, Singapore, 2018, pp. 717-726.
5. Verma K., Bhardwaj S., Arya R., UL Islam M.S., Bhushan M., Kumar A., Samant P. Latest Tools for Data Mining and Machine Learning // *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*. 2019. Vol. 8, Issue 9S. <https://doi.org/10.35940/ijitee.I1003.0789S19>
6. Alcalá R., Gacto MJ., Alcalá-Fdez J. Evolutionary data mining and applications: A revision on the most cited papers from the last 10 years (2007–2017) // *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2018. Vol. 8, e1239.
7. Jean N., Burke M., Xie M., Davis W.M., Lobell D.B., Ermon S. Combining satellite imagery and machine learning to predict poverty // *Science*. 2016. Vol. 353(6301). P. 790–794.
8. Shi S., Wang Q., Xu P., Chu X. Benchmarking state-of-the-art deep learning software tools // 2016 7th international conference on cloud computing and big data (CCBD). IEEE, 2016. P. 99–104.
9. Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system // *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016.P. 785–794.
10. Patrick Hall, Navdeep Gill, Nicholas Schmidt. Proposed Guidelines for the Responsible Use of Explainable Machine Learning. arXiv:1906.03533v3 [stat. ML] 29 Nov 2019.
11. Sorelle A. Friedler, Chitradeep Dutta Roy, Carlos Scheidegger, and Dylan Slack. Assessing the Local Interpretability of Machine Learning Models. arXiv preprint arXiv:1902.03501, 2019. <https://arxiv.org/pdf/1902.03501.pdf>.
12. Nguyen G., Dlugolinsky S., Bobák M. et al. Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey // *Artif Intell Rev*. 2019. Vol. 52. P. 77–124. <https://doi.org/10.1007/s10462-018-09679-z>
13. Hall, Patrick. On the Art and Science of Explainable Machine Learning: Techniques, Recommendations, and Responsibilities. <https://arxiv.org/pdf/1810.02909.pdf>
14. Wojciech Froelich. Towards improving the efficiency of the fuzzy cognitive map classifier // *Neurocomputing*. 2017. Vol. 232. P. 83-93, <https://doi.org/10.1016/j.neucom.2016.11.059>
15. Rory Mitchell. Gradient Boosting, Decision Trees and XGBoost with CUDA. <https://devblogs.nvidia.com/parallelforall/gradient-boosting-decision-trees-xgboost-cuda/>

16. Mitchell R, Frank E. Accelerating the XGBoost algorithm using GPU computing // PeerJ Computer Science. 2017. Vol. 3, e127. <https://doi.org/10.7717/peerj-cs.127>
17. Катрашова Ю. В. Использование «сквозных» цифровых технологий в сфере государственного управления / Ю. В. Катрашова, Г. Ю. Митяшин // Наука Красноярья. 2020. Т. 9, № 4. С. 85-102. <https://doi.org/10.12731/2070-7568-2020-4-85-102>
18. Катрашова Ю. В. Система социального рейтинга как форма государственного контроля над обществом: перспективы внедрения и развития, угрозы реализации / Ю. В. Катрашова, Г. Ю. Митяшин, В. А. Плотников // Управленческое консультирование. 2021. № 2(146). С. 100-109. <https://doi.org/10.22394/1726-1139-2021-2-100-109>
19. Плотников В. А. Цифровая революция, Covid-19 и государственное управление / В. А. Плотников, А. В. Пролубников // Вызовы цифровой экономики: тренды развития в условиях последствий пандемии COVID-19 : Сборник статей IV Всероссийской научно-практической конференции, приуроченной к Году науки и технологий в России, Брянск, 25 мая 2021 года. Брянск: Федеральное государственное бюджетное образовательное учреждение высшего образования «Брянский государственный инженерно-технологический университет», 2021. С. 234-237.
20. Катрашова Ю. В. Преимущества и угрозы внедрения технологии больших данных в систему государственного управления / Ю. В. Катрашова, Г. Ю. Митяшин // Актуальные проблемы аграрной науки: прикладные и исследовательские аспекты : Сборник научных трудов Всероссийской (национальной) научно-практической конференции, Нальчик, 04–05 февраля 2021 года. Том I. Нальчик: Федеральное государственное бюджетное образовательное учреждение высшего образования «Кабардино-Балкарский государственный аграрный университет имени В.М. Кокова», 2021. С. 367-370.
21. Апатова Н. В. Цифровые трансформации бизнеса и социума // Вестник Тверского государственного университета. Серия: Экономика и управление. 2022. № 4(60). С. 60-71. <https://doi.org/10.26456/2219-1453/2022.4.060-071>

References

1. Frolov Yu.V., Sakhnyuk P.A., Sakhnyuk T.I. Using geoanalytics tools for training bachelors and masters in business informatics. *Bulletin of the Russian Peoples' Friendship University. Series: Informatization of education*. 2018. Т. 15. No. 2. pp. 151-164.
2. Mamaev I.I., Sakhnyuk T.I., Sakhnyuk P.A. Analysis of the main socio-economic indicators of the regions of the Stavropol Territory for 2011-2015 using business

- analytics systems. *Scientific journal of KubSAU* [Electronic resource]. Krasnodar: KubGAU, 2016. No. 123(09). <http://ej.kubagro.ru/2016/09/pdf/26.pdf>
3. The number of families receiving subsidies in the city of Moscow, in various sections. <https://data.mos.ru/opendata/7701236617-kolichestvo-semey-poluchayushchih-subsidiyu-v-gorode-moskve>
 4. Siddiqui T., Ausaf A. Data mining tools and techniques for mining software repositories: A systematic review. *Big Data Analytics*. Springer, Singapore, 2018, pp. 717-726.
 5. Verma K., Bhardwaj S., Arya R., UL Islam M.S., Bhushan M., Kumar A., Samant P. Latest Tools for Data Mining and Machine Learning. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*. 2019. Vol. 8, Issue 9S. <https://doi.org/10.35940/ijitee.I1003.0789S19>
 6. Alcalá R., Gacto MJ., Alcalá-Fdez J. Evolutionary data mining and applications: A revision on the most cited papers from the last 10 years (2007–2017). *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2018. Vol. 8, e1239.
 7. Jean N., Burke M., Xie M., Davis W.M., Lobell D.B., Ermon S. Combining satellite imagery and machine learning to predict poverty. *Science*. 2016. Vol. 353(6301). P. 790–794.
 8. Shi S., Wang Q., Xu P., Chu X. Benchmarking state-of-the-art deep learning software tools. *2016 7th international conference on cloud computing and big data (CCBD)*. IEEE, 2016, pp. 99–104.
 9. Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016. P. 785–794.
 10. Patrick Hall, Navdeep Gill, Nicholas Schmidt. Proposed Guidelines for the Responsible Use of Explainable Machine Learning. arXiv:1906.03533v3 [stat. ML] 29 Nov 2019.
 11. Sorelle A. Friedler, Chitradeep Dutta Roy, Carlos Scheidegger, and Dylan Slack. Assessing the Local Interpretability of Machine Learning Models. arXiv preprint arXiv:1902.03501, 2019. <https://arxiv.org/pdf/1902.03501.pdf>.
 12. Nguyen G., Dlugolinsky S., Bobák M. et al. Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey. *Artif Intell Rev*. 2019. Vol. 52. P. 77–124. <https://doi.org/10.1007/s10462-018-09679-z>
 13. Hall, Patrick. On the Art and Science of Explainable Machine Learning: Techniques, Recommendations, and Responsibilities. <https://arxiv.org/pdf/1810.02909.pdf>
 14. Wojciech Froelich. Towards improving the efficiency of the fuzzy cognitive map classifier. *Neurocomputing*. 2017. Vol. 232. P. 83-93, <https://doi.org/10.1016/j.neucom.2016.11.059>

15. Rory Mitchell. Gradient Boosting, Decision Trees and XGBoost with CUDA. <https://devblogs.nvidia.com/parallelforall/gradient-boosting-decision-trees-xg-boost-cuda/>
16. Mitchell R, Frank E. Accelerating the XGBoost algorithm using GPU computing. *PeerJ Computer Science*. 2017. Vol. 3, e127. <https://doi.org/10.7717/peerj-cs.127>
17. Katrashova Yu. V. The use of “end-to-end” digital technologies in the field of public administration / Yu. V. Katrashova, G. Yu. Mityashin. *Science of Krasnoyarsk*. 2020. Vol. 9, No. 4. pp. 85-102. <https://doi.org/10.12731/2070-7568-2020-4-85-102>
18. Katrashova Yu. V. Social rating system as a form of state control over society: prospects for implementation and development, threats to implementation / Yu. V. Katrashova, G. Yu. Mityashin, V. A. Plotnikov. *Management consulting*. 2021. No. 2(146). pp. 100-109. <https://doi.org/10.22394/1726-1139-2021-2-100-109>
19. Plotnikov V. A. Digital revolution, Covid-19 and public administration / V. A. Plotnikov, A. V. Prolubnikov. *Challenges of the digital economy: development trends in the context of the consequences of the COVID-19 pandemic: Collection of articles of the IV All-Russian scientific-practical conference dedicated to the Year of Science and Technology in Russia, Bryansk, May 25, 2021*. Bryansk: Bryansk State Engineering and Technology University, 2021. P. 234-237.
20. Katrashova Yu. V. Advantages and threats of introducing big data technology into the public administration system / Yu. V. Katrashova, G. Yu. Mityashin. *Current problems of agricultural science: applied and research aspects: Collection of scientific works of the All-Russian (national) scientific-practical conference, Nalchik, February 04–05, 2021*. Volume I. Nalchik: Kabardino-Balkarian State Agrarian University named after V.M. Kokov, 2021. P. 367-370.
21. Apatova N.V. Digital transformations of business and society. *Bulletin of Tver State University. Series: Economics and management*. 2022. No. 4(60). pp. 60-71. <https://doi.org/10.26456/2219-1453/2022.4.060-071>

ДАнные ОБ АВТОРАХ

Сахнюк Павел Анатольевич, доцент, департамента бизнес-информатики,
кандидат технических наук
Финансовый университет при правительстве Российской Федерации
Ленинградский проспект, 49, г. Москва, 125993, Российская Феде-
рация
pav-sahnyuk@yandex.ru

Сахнюк Татьяна Ивановна, доцент департамента информатики, управления и технологий, кандидат экономических наук
Институт цифрового образования Московского городского педагогического университета
ул. Шереметьевская, 28, г. Москва, 129594, Российская Федерация
Tatiana-sakhnyuk@yandex.ru

DATA ABOUT THE AUTHORS

Pavel A. Sakhnyuk, Associate Professor, Department of Business Informatics, Candidate of Technical Sciences
Financial University under the government of the Russian Federation
49, Leningradsky Ave., Moscow, 125993, Russian Federation
pav-sahnyuk@yandex.ru
SPIN-code: 7192-6869
ORCID: <https://orcid.org/0000-0002-1457-0640>
ResearcherID: GRR-6444-2022
Scopus Author ID: 57191248583

Tatyana I. Sakhnyuk, Associate Professor, Department of Informatics, Management and Technology
Institute of Digital Education, Moscow City Pedagogical University
28, Sheremetevskaya Str., Moscow, 129594, Russian Federation
Tatiana-sakhnyuk@yandex.ru
SPIN-code: 5339-2920
ORCID: <https://orcid.org/0000-0003-2929-0112>
Scopus Author ID: 57191247960

Поступила 07.04.2023

После рецензирования 29.04.2023

Принята 20.05.2023

Received 07.04.2023

Revised 29.04.2023

Accepted 20.05.2023