

ПРИМЕНЕНИЕ МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДЛЯ ПОВЫШЕНИЯ ПРИБЫЛЬНОСТИ СЕТЕВОГО БИЗНЕСА¹

Н. П. Савина, Н. А. Галстян, О. В. Литвишко, Е. А. Закревская

Российский экономический университет имени Г. В. Плеханова,
Москва, Россия

В статье рассмотрено применение кластеризации аптечных сетей и определены стратегии работы с каждым из полученных кластеров, что вызвано необходимостью перепроверять основные показатели работы аптечных сетей (скорость продаж, количество брендов исследуемой компании в сети, состояние складов, наличие сильных конкурентов в продажах сети), а также сложностью распределения ресурсов компании для выстраивания коммуникации с аптечными сетями (визитов представителей) и самой продукции между аптечными сетями. Для решения задачи кластеризации аптечных сетей в исследовании выбран метод *k-means*, основанный на учете доли компании и доли конкретного бренда в продажах аптечной сети. Метрикой для определения качества полученных результатов был выбран «силуэт», т. е. форма отображения набора кластеров. Его показатель равен 0,514, что свидетельствует о достаточно высокой точности результатов и возможности внедрения данного алгоритма в реальную бизнес-практику. Итогом кластерного анализа множества аптечных сетей является набор из трех кластеров. Используя полученные результаты кластеризации и имеющиеся бизнес-требования для каждого кластера, в ходе исследования был предложен ряд рекомендаций по взаимодействию фармацевтической компании с аптечными сетями в части оценки конкурентной среды, анализа скорости продаж и пополняемости складских запасов сети, частоты визитов фармацевтических представителей в аптечные сети. По результатам проведенного исследования сделан вывод, что любой предлагаемый метод кластеризации аптечных данных должен обновляться с точки зрения качества интерпретации результатов, актуальности изначально выбранных критериев кластеризации и корректироваться исходя из бизнес-требований, которые поступают со стороны команд маркетологов, менеджеров по категориям лекарственных средств и отдельным брендам.

Ключевые слова: кластеризация, фармацевтический бизнес, интеллектуальный анализ, стратегия, *k-means*-алгоритм.

USING METHODS OF INTELLECTUAL ANALYSIS TO STEP UP PROFITABILITY OF NETWORK BUSINESS

Natalya P. Savina, Narek A. Galstyan,

Oleg V. Litvishko, Ekaterina A. Zakrevskaya

Plekhanov Russian University of Economics, Moscow, Russia

The article studies the use of drug-store chain clusterization and determines strategy of work with each of the clusters, which is necessary because key work parameters of drug-store chains, such as selling speed, the number of brands of the company being investigated in the chain, the condition of stores, availability of strong competitors in the chain selling must be verified. Another reason is complexity of company resource distribution to build communications with drug-store chains (representatives' calls) and products themselves among drug-store chains. To resolve the problem of drug-store clusterization *k-means* method based on the assessment of the company share

¹ Статья подготовлена по результатам исследования, проведенного при финансовой поддержке Российского экономического университета имени Г. В. Плеханова.

and the share of a concrete brand in the drug-store chain was chosen for the research. As metrics for identifying the quality of obtained results 'silhouette' was chosen, i.e. the form of a cluster set representation. Its factor is equal to 0.514, which testifies to rather high accuracy of results and possibility to introduce this algorithm into real business practice. By cluster analyzing the multitude of drug-store chains a set of three clusters was identified. On the basis of these results of clusterization and current business requirements for each cluster a number of recommendations were put forward aimed at interaction between the pharmaceutical company and drug-store chains in the aspect of assessing the competition environment, analyzing selling speed and re-filling stocks of the chain, frequency of pharmaceutical representatives' calls to drug-store chains. The findings of the research allowed us to draw a conclusion that any method of clusterization of drug-store data should be renewed in view of the quality of result interpretation, topicality of initial criteria of clusterization and should be corrected proceeding from business requirements, which arrive from marketer teams and managers on medicine categories and separate brands.

Keywords: clusterization, pharmaceutical business, intellectual analysis, strategy, *k-means* algorithm.

Введение

Одним из ключевых факторов обеспечения высоких продаж и повышения доли рынка для фармацевтической компании является четкое обоснование выбранной стратегии взаимодействия с аптечными сетями, которые в свою очередь являются проводниками между фармацевтической компанией и пациентами, нуждающимися в тех или иных лекарственных средствах. В зарубежной и отечественной бизнес-практиках компании зачастую выстраивают индивидуальную стратегию взаимодействия с каждой аптечной сетью или группой аптечных сетей. Другими словами, планированию работы с аптечными сетями предшествует обязательный процесс распределения каждой сети по группам, к которым привязаны установленные правила взаимодействия. Таким образом, перед бизнес-аналитиками фармацевтической компании ставится важная цель – кластеризация аптечных сетей и определение стратегии работы с каждым из полученных кластеров.

Вместе с тем бизнес-аналитики зачастую не обновляют методологию кластеризации аптечных сетей, не пересматривают актуальность и значимость факторов, выбранных для группировки аптечных сетей в определенное количество кластеров, используют в работе устаревшие инструменты с небольшими возможностями обработки и анализа данных. В условиях постоянной трансформации фармацевтического рынка, появления большого количества данных, в том числе из-за внедрения закона о передаче данных с кассовых аппара-

тов в налоговую службу через операторов фискальных данных, появления новых каналов сбыта лекарственных средств посредством электронной торговли, а также изменения спроса на те или иные лекарственные средства существующие методы кластеризации аптечных сетей могут утратить свою актуальность и эффективность.

Полученные кластеры аптечных сетей должны находиться под наблюдением бизнес-аналитиков. Необходимы с определенной частотой обновления факторов группировки аптечных сетей и пересмотр стратегии с уже существующими кластерами. Для осуществления такой инициативы требуется разработка нового автоматизированного алгоритма кластеризации аптечных сетей, который будет способен аккумулировать, обрабатывать и анализировать часто изменяющиеся данные об аптечных сетях, обогащать текущую модель кластеризации новыми данными. Этим и обусловливается актуальность темы исследования, рассматриваемой в данной статье.

Методология

Для решения задачи кластеризации аптечных сетей за основу выбран метод *k-means*.

Цель любого алгоритма кластеризации состоит в получении такого результата, чтобы расстояние между точками в одном кластере было небольшим по сравнению с расстоянием между двумя соседними кластерами. Другими словами, члены одной группы очень похожи, а члены разных групп крайне непохожи.

Метод *k-means* – это итерационный алгоритм кластеризации, в котором число кластеров k заранее определено и алгоритм итеративно присваивает каждую точку данных одному из k кластеров на основе сходства признаков [13].

Опишем детально технический алгоритм *k-means*, по которому разработан предложенный метод кластеризации:

1. Определение количества кластеров k , базируясь на исходных данных об объектах кластерного анализа и бизнес-требованиях.
2. Определение k -центроидов, которые рассчитываются исходя из набора данных.
3. Назначение каждой точки из датасета на ближайший центроид; в качестве проверки расстояния между точками и центроидами используем евклидово расстояние по формуле (1).
4. Переопределение центроидов и присвоение точек новым кластерам.
5. Если точки присвоились другим кластерам (в отличие от предыдущей итерации), повторяем пункт 4; в противном случае переходим к следующему пункту алгоритма.
6. Сохранение финальных кластеров.

Математическая модель, заложенная в основу предлагаемого автоматизированного метода кластеризации, построена на минимизации суммы квадратов расстояний между центроидом кластера и связанными с ним точками набора данных [11]:

$$\text{Minimize } \sum_{j=1}^k \sum_{i=1}^n (x_{ij} - c_j)^2, \quad (1)$$

где k – количество кластеров;

n – количество точек в датасете (объектов кластеризации);

c – центроид кластера j ;

$(x_{ij} - c_j)$ – расстояние между точкой и центроидом, которому данная точка принадлежит.

Важнейшим шагом в предложенном методе кластеризации аптечных сетей является процесс расчета количества кластеров. Другими словами, основным входным сигналом для кластерного анализа множества объектов с помощью метода *k-средних* яв-

ляется количество кластеров k , которое можно получить с помощью концепции минимизации суммы квадратов внутри кластера (Minimizing Within Cluster Sum of Square – WCSS).

Сумма квадратов внутри кластера является мерой изменчивости наблюдений внутри каждого кластера. Эта изменчивость выражается в виде среднего расстояния от каждого наблюдения до центроида [10]. Кластеры с меньшей суммой квадратов являются более компактными, чем кластеры с большей суммой квадратов.

По мере увеличения числа кластеров показатель WCSS продолжает уменьшаться: вначале WCSS резко снижается, затем скорость снижения замедляется, что приводит к локтевому участку (elbow plot). Количество кластеров в локтевом образовании обычно дает представление об оптимальном количестве кластеров.

При использовании концепции минимизации WCSS нужно учитывать возможные бизнес-требования, поскольку задача кластеризации часто решается для ее внедрения в бизнес-процессы [8]. Поведение показателя WCSS при изменении количества кластеров представлено на рис. 1, где выделенный прямоугольник – локтевое образование.

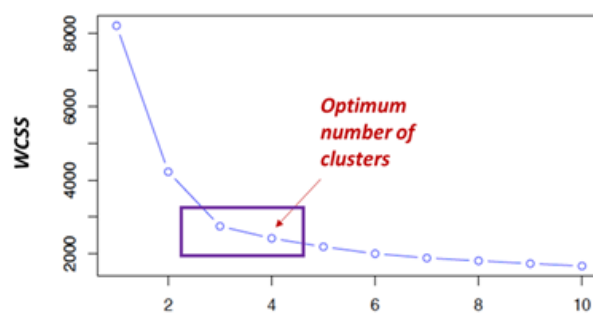


Рис. 1. График зависимости показателя WCSS от количества кластеров

Цель кластерного анализа – разделение текущего набора аптечных сетей на сегменты с определенными сходствами, при этом полученные группы аптечных сетей должны различаться друг от друга.

Результат

Для решения задачи кластеризации был собран набор данных по 250 различным аптечным сетям из 82 городов России.

Исходя из особенностей и требований бизнес-процессов в качестве переменных для кластеризации аптечных сетей были выбраны следующие факторы:

1. Доля продаж всех брендов компании в аптечной сети (Company Share in Chain – CSC).

2. Доля продаж промотируемого бренда в данной сети (Brand Share in Chain – BSC).

На рис. 2 представлены основные показатели исходного набора данных для выборки аптечных сетей.

	count	mean	std	min	25%	50%	75%	max
CSC	246.0	0.038226	0.032030	-0.017771	0.017059	0.031569	0.047370	0.227235
BSC	246.0	0.008415	0.007482	0.000024	0.003676	0.006832	0.010224	0.053672

Рис. 2. Статистические показатели по переменным CSC и BSC

В первую очередь необходимо определить оптимальное количество кластеров. Для этого используем метод elbow. На рис. 3 изображена зависимость WCSS от возможного количества кластеров для исследуемой выборки. Согласно локтевому методу выбираем количество кластеров $k = 3$, так как скорость снижения после этой точки сбавляется и становится равномерной.

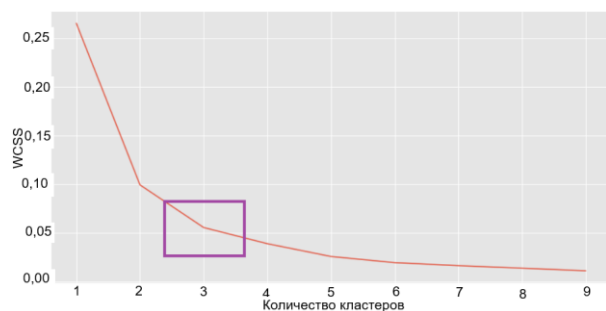


Рис. 3. График зависимости WCSS от количества кластеров

Определенное с помощью локтевого метода потенциальное количество кластеров аптечных сетей позволяет начать запуск алгоритма *k-means*. Автоматизированный программный код на основе предложенного метода кластеризации позволит не только быстро рассчитывать кластеры и соотносить каждую точку данных (аптечную сеть) с одним из кластеров, но и за-

гружать данные, которые будут регулярно обновляться.

Далее алгоритм кластеризации был выполнен с помощью использования специального пакета Python для машинного обучения (*sklearn.cluster – k-means*).

Также с помощью библиотеки *matplotlib* был построен график для визуализации результатов кластерного анализа аптечных сетей. Результат визуализации кластеризации представлен на рис. 4.

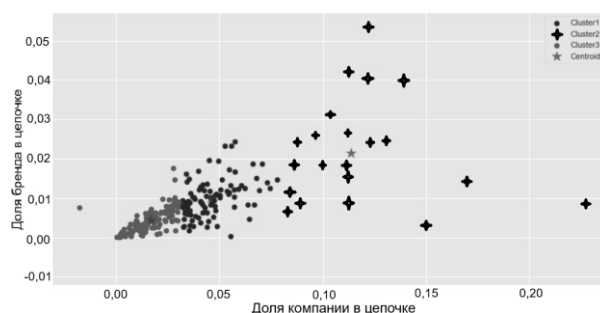


Рис. 4. Визуализация результатов кластерного анализа аптечных сетей

Следующим этапом разработки метода кластеризации аптечных сетей является оценка качества решения. Для этого была выбрана метрика силуэта внутреннего типа, для расчета которого не нужны начальные (первичные) результаты кластеризации (по текущей методологии). Метрика для предложенного метода рассчитывается на основе пакета SPSS и равна 0,514,

что отражает наличие плотных, т. е. четко выделенных, кластеров и свидетельствует о высоком качестве предложенного метода кластеризации и о возможности его внедрения в реальную бизнес-практику.

Последним этапом кластеризации аптечных сетей является интерпретация полученных кластеров с точки зрения выбранных факторов. Результатом кластерного анализа на основе предложенного метода являются три различных сегмента аптечных сетей:

1. *Кластер 1* с низкой долей продаж всех брендов компании и низкой долей промотируемого бренда в сетях.

2. *Кластер 2* со средней долей продаж всех брендов компании и низкой/средней долей промотируемого бренда в сетях.

3. *Кластер 3* с высокой долей продаж всех брендов компании и средней/высокой долей промотируемого бренда в сетях.

Можно заметить, что кластеры явно отличаются на основе доли совокупных продаж по всем брендам компании, в то время как доля отдельного бренда является варьируемым фактором. Таким образом, можно сделать предварительный вывод, что при выстраивании стратегии взаимодействия с каждой аптечной сетью нужно в первую очередь смотреть на долю всех брендов компании в продажах данной сети.

Кластеризация была проведена для одного конкретного бренда лекарственного средства, однако предложенный метод будет работать и при использовании других ключевых брендов в промотируемом портфеле компании.

Заключение

Итогом кластерного анализа множества аптечных сетей является набор из трех кластеров. Используя полученные результаты кластеризации и имеющиеся бизнес-требования для каждого кластера, в ходе исследования был составлен список рекомендаций по взаимодействию. Учитывая возможные изменения описания кластеров во времени, нужно сделать допущения об устойчивости набора факторов каждого

кластера только на текущий момент. Таким образом, с сетями каждого из кластеров предлагается взаимодействовать по рекомендациям, представленным ниже.

Рекомендации для аптечных сетей кластера 1:

- в аптечные сети кластера 1 можно поставлять другие бренды, для которых данные сети относятся к кластеру 3;

- необходимо выяснять текущие объемы поставок анализируемого бренда в эти сети, так как есть вероятность, что продажи низкие, но быстрые (при высоком спросе) из-за неверных решений, основанных на кластеризации по утратившей свою актуальность методологии;

- следует учитывать, что низкая доля всей компании в продажах сети свидетельствует о большем количестве конкурентных компаний и брендов в данных аптечных сетях;

- количество визитов фармацевтических представителей в аптечные сети данного кластера необходимо увеличить с целью изучения конкурентной среды и причин преобладания конкурентов, а также с целью анализа скорости продаж и пополняемости складских запасов сети продукцией компании.

Рекомендации для аптечных сетей кластера 2:

- в аптеки данной категории можно направлять перераспределенные доли поставок из других кластеров для поддержания и увеличения доли компании в продажах данных сетей;

- количество визитов фармацевтических представителей в аптечные сети данного кластера можно оставить на стабильном уровне в связи с тем, что основные критерии кластеризации находятся на среднем значении по всему рынку сети.

Рекомендации для аптечных сетей кластера 3:

- данный кластер может быть использован в качестве территории для экспериментов: во-первых, в аптечных сетях кластера можно тестировать и продвигать запуски новых препаратов, так как эти ап-

течные сети очень лояльны к компании, имея высокую долю их брендов в своих продажах; во-вторых, поставки из этого кластера могут быть перераспределены при необходимости в другие сети;

- целесообразно продолжать поставлять анализируемый бренд в том же объеме либо выше, так как высокие показатели по критериям кластеризации свидетельствуют о лояльности аптечной сети к продукции компании;

- необходимо проверить, как долго каждая из сетей данного кластера держит на высоком уровне показатели по целевым критериям.

При обнаружении аптечных сетей, которые повысили свой кластер, необходимо выяснить причины, из-за которых такие аптечные сети стали более привлекательными для взаимодействия с компанией. Выясненные факторы могут быть применены к сетям из других кластеров, что позволит предпринять меры для повышения их показателей в будущем.

В дополнение к индивидуальным рекомендациям для каждого кластера предлагаются общие инициативы, которые целесообразно внедрить при использовании нового подхода к кластеризации аптечных сетей:

1. Ежемесячное либо ежеквартальное обновление алгоритма кластеризации, корректировки полученных критериев с точки зрения бизнес-процессов, изменения его условий и заранее оформленных договоренностей с аптечными сетями. Любой предлагаемый метод кластеризации аптечных данных должен обновляться с точки зрения качества интерпретации результатов, актуальности изначально выбранных критериев кластеризации и корректироваться исходя из бизнес-требований, которые поступают со стороны команд маркетологов, менеджеров по категориям лекарственных средств и отдельным брендам.

2. Для быстрого обновления использовать продвинутые инструменты бизнес-аналитики, аккумулирующие большие

объемы данных как по аптечным сетям, так и по всему фармацевтическому рынку. Грамотно настроенный процесс загрузки, валидации новых данных и хранения старых позволит быстрее обновлять алгоритм кластеризации и проверять актуальность выбранных критериев группировки сетей за прошлый период.

3. Исходя из возможностей компании использование первичной информации от операторов фискальных данных (ОФД), которые аккумулируют все данные о продажах с кассовых аппаратов аптечных точек всех регионов. Поставщики такого рода данных не поставляют данные в разрезе конкретных аптечных сетей (такие данные можно получить только из отчетов сетей, с которыми взаимодействует напрямую компания), однако с помощью данных ОФД можно следить за динамикой продаж своей продукции и конкурентов на еженедельной основе, в региональном разрезе и на уровне релевантных для каждого бренда рынках.

Качество решения задачи кластеризации аптечных сетей сильно зависит от инструментов и процессов, используемых и построенных для обработки большого количества данных, от частоты обновления данных, их оперативности, точности и количества разрезов, в которых они могут быть представлены и обработаны бизнес-аналитиками.

В условиях изменения спроса на те или иные препараты в различных регионах доли отдельных брендов в продажах аптечных сетей могут быть восприняты в кластеризации по-другому, тем самым определяя потенциально новые кластеры. В кластерном анализе используется доля продаж из аптек, а не доля отгрузок (поставок) во множество аптечных сетей от компании либо ее дистрибьюторов. Так может возникнуть группа аптечных сетей, у которых низкая доля конкретного бренда в продажах, но при этом остальные бренды компании имеются в больших объемах (в упаковках и в денежной форме).

Таким образом, в аптечных точках данной сети может возникнуть ситуация, когда конкретный бренд находится в застое на складах аптек из-за отсутствия спроса на препарат среди покупателей, выбирающих аптеки данной сети. Следовательно, нужно принять решение о перераспределении поставок данного бренда из анализируемой аптечной сети в другие.

Также нужно учитывать, что поведение аптек в каждой из сетей и выполнение их планов по продажам зависят от множества различных регионов, в которых они находятся. Есть вероятность образования другого кластера аптечных сетей, где в продажах доля компании низкая по сравнению с другими сетями, а доля одного бренда намного выше данного показателя в других сетях. В таком кластере аптечных сетей выбранный для анализа бренд вносит основной вклад в продажи компании в этих аптеках.

С созданием кластеров аптечных сетей может измениться стратегия поставок, например, с помощью увеличения поставок других препаратов в эти сети, так как есть вероятность, что эти аптечные сети хорошо выполняют планы для любых препаратов от различных компаний и необходимо перераспределение части отгрузок анализируемого бренда из текущего кластера в другие.

Нужно также учитывать, что любые результаты, полученные на основе того или иного метода кластеризации, не всегда могут быть интегрированы в реальную бизнес-практику с точно таким же набором факторов, какой был на выходе кластерного анализа. В таких случаях описание кластеров и основные факторы их определения должны быть скорректированы на основании бизнес-требований.

Также количество кластеров аптечных сетей и их критерии позволяют определить распределение визитов фармацевтических представителей компании в те или иные аптечные сети. Данные визиты служат инструментом коммуникации с аптечными сетями на первичном этапе переговоров с руководителями аптечных точек и/или сетей, в ходе которых определяются планы поставок, правила выставления препаратов на полки (маркетинговые инструменты), реклама определенных брендов (это касается лекарственных средств безрецептурного рынка), что важно в построении стратегии взаимодействия компании с аптечными сетями.

Следовательно, неправильная интерпретация результатов кластеризации аптечных сетей с точки зрения бизнеса повлечет за собой некорректное распределение визитов между аптечными сетями, что в свою очередь может привести к неэффективному использованию имеющихся ресурсов компании. Таким образом, бизнес-процессы, происходящие внутри фармацевтической компании, и особенности фармацевтического рынка являются источниками ряда трудностей, с которыми сталкиваются бизнес-аналитики при интеграции кластерного анализа в текущую систему работы.

Вышеперечисленные особенности интерпретации результатов кластеризации любого реального множества объектов (в том числе и аптечных сетей) свидетельствуют о том, что необходимо рассматривать различные сценарии бизнес-процессов и корректировать описания кластеров, прежде чем использовать их в принятии бизнес-решений.

Список литературы

1. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с англ. А. А. Слинкина. – М. : ДМК Пресс, 2018.
2. Литвишко О. В. Финансовые аспекты функционирования спортивных клубов // Теория и практика физической культуры. – 2011. – № 5. – С. 77.

3. Обучение без учителя: 4 метода кластеризации данных на Python. 2018. – URL: <https://proglab.io/p/unsupervised-ml-with-python/> (дата обращения: 01.11.2021).
4. Соколов Е. А. Лекция 12. Обучение без учителя. – URL: <https://github.com/esokolov/ml-course-hse/blob/master/2016-fall/lecture-notes/lecture12-unsupervised.pdf> (дата обращения: 01.11.2021).
5. Чубукова И. А. Курс лекций «Data Mining». Национальный открытый интернет-университет информационных технологий «ИНТУИТ». 2018. – URL: <http://www.intuit.ru/departament/database/datamining> (дата обращения: 01.11.2021).
6. Agglomerative Hierarchical Clustering. International online learning resource for Data Mining and Statistics «Data Nova». 2019. – URL: <https://www.datanovia.com/en/lessons/agglomerative-hierarchical-clustering/> (дата обращения: 01.11.2021).
7. Dabbura I. K-means Clustering: Algorithm, Applications, Evaluation Methods, and Drawbacks. Online publishing platform Medium. 2018. – URL: <https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a> (дата обращения: 01.11.2021).
8. Das A. Segmentation using Unsupervised Learning Technique – Clustering. Online publishing platform Medium. 2018. – URL: <https://towardsdatascience.com/segmentation-using-unsupervised-learning-technique-clustering-bb4089c6392e> (дата обращения: 01.11.2021).
9. Girgin S. K-Means Clustering Model in 6 Steps with Python. Online publishing platform Medium. 2019. – URL: <https://medium.com/pursuitnotes/k-means-clustering-model-in-6-steps-with-python-35b532cfa8ad> (дата обращения: 01.11.2021).
10. Korolev S. Open Machine Learning Course. Topic 7. Unsupervised Learning: PCA and Clustering. Online publishing platform Medium. 2018. – URL: <https://medium.com/open-machine-learning-course/open-machine-learning-course-topic-7-unsupervised-learning-pca-and-clustering-db7879568417> (дата обращения: 01.11.2021).
11. Litoishko O., Beketova K., Akimova B., Azhmukhamedova A., Islyam G. Impact of the Digital Economy on the Banking Sector // E3S Web of Conferences. 1st International Conference on Business Technology for a Sustainable Environmental System. BTSES-2020. – Almaty, 2020. – С. 04033.
12. Nakagomi K. Cluster Analysis of Pharmacists' Work Attitudes // Journal of General and Family Medicine, Japan. – 2017. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5729359/> (дата обращения: 01.11.2021).
13. Selecting the number of clusters with silhouette analysis on KMeans clustering. Online platform of machine learning in Python. 2019. – URL: https://scikit-learn.org/stable/auto_examples/cluster/plot_kmeans_silhouette_analysis.html (дата обращения: 01.11.2021).
14. Todd A., Thomson K., Kasim A., Bambra C. Cutting Care Clusters: the Creation of an Inverse Pharmacy Care Law? An area-level analysis exploring the clustering of community pharmacies in England // BMJ Open Journal, England. – 2018. – URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6074640/> (дата обращения: 01.11.2021).
15. VanderPlas J. Python Data Science Handbook: Essential Tools for Working with Data. – O'Reilly Media, 2016.
16. Vivek S. Clustering algorithms for customer segmentation. Online publishing platform Medium. 2018. – URL: <https://towardsdatascience.com/clustering-algorithms-for-customer-segmentation-af637c6830ac> (дата обращения: 01.11.2021).
17. Wijaya C. Y. Breaking down the agglomerative clustering process. Online publishing platform Medium. 2019. – URL: <https://towardsdatascience.com/breaking-down-the-agglomerative-clustering-process-1c367f74c7c2> (дата обращения: 01.11.2021).

References

1. Gudfellou Ya., Bendzhio I., Kurvill A. *Glubokoe obuchenie* [Deep Learning], translated from English by A. A. Slinkin. Moscow, DMK Press, 2018. (In Russ.).
2. Litvishko O. V. *Finansovye aspekty funktsionirovaniya sportivnykh klubov* [Financial Aspects of the Functioning of Sports Clubs]. *Teoriya i praktika fizicheskoy kultury* [Theory and Practice of Physical Culture], 2011, No. 5, pp. 77. (In Russ.).
3. *Obuchenie bez uchitelya: 4 metoda klasterizatsii dannykh na Python. 2018* [Teaching without a teacher: 4 methods of data clustering in Python. 2018]. (In Russ.). Available at: <https://proglib.io/p/unsupervised-ml-with-python/> (accessed 01.11.2021).
4. Sokolov E. A. Lektsiya 12. *Obuchenie bez uchitelya* [Lecture 12. Teaching without a teacher]. (In Russ.). Available at: <https://github.com/esokolov/ml-course-hse/blob/master/2016-fall/lecture-notes/lecture12-unsupervised.pdf> (accessed 01.11.2021).
5. Chubukova I. A. Kurs lektsiy «Data Mining». *Natsionalnyy otkrytyy internet-universitet informatsionnykh tekhnologiy «INTUIT»*. 2018 [Course of lectures "Data Mining". National Open Internet University of Information Technologies "INTUIT". 2018]. (In Russ.). Available at: <http://www.intuit.ru/departament/database/datamining> (accessed 01.11.2021).
6. Agglomerative Hierarchical Clustering. International online learning resource for Data Mining and Statistics «Data Nova». 2019. Available at: <https://www.datanovia.com/en/lessons/agglomerative-hierarchical-clustering/> (accessed 01.11.2021).
7. Dabbura I. K-means Clustering: Algorithm, Applications, Evaluation Methods, and Drawbacks. Online publishing platform Medium. 2018. Available at: <https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a> (accessed 01.11.2021).
8. Das A. Segmentation using Unsupervised Learning Technique – Clustering. Online publishing platform Medium. 2018. Available at: <https://towardsdatascience.com/segmentation-using-unsupervised-learning-technique-clustering-bb4089c6392e> (accessed 01.11.2021).
9. Girgin S. K-Means Clustering Model in 6 Steps with Python. Online publishing platform Medium. 2019. Available at: <https://medium.com/pursuitnotes/k-means-clustering-model-in-6-steps-with-python-35b532cfa8ad> (accessed 01.11.2021).
10. Korolev S. Open Machine Learning Course. Topic 7. Unsupervised Learning: PCA and Clustering. Online publishing platform Medium. 2018. Available at: <https://medium.com/open-machine-learning-course/open-machine-learning-course-topic-7-unsupervised-learning-pca-and-clustering-db7879568417> (accessed 01.11.2021).
11. Litvishko O., Beketova K., Akimova B., Azhmukhamedova A., Islyam G. Impact of the Digital Economy on the Banking Sector. *E3S Web of Conferences. 1st International Conference on Business Technology for a Sustainable Environmental System. BTSES-2020*. Almaty, 2020. C. 04033.
12. Nakagomi K. Cluster Analysis of Pharmacists' Work Attitudes. *Journal of General and Family Medicine, Japan*, 2017. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5729359/> (accessed 01.11.2021).
13. Selecting the number of clusters with silhouette analysis on KMeans clustering. Online platform of machine learning in Python. 2019. Available at: https://scikit-learn.org/stable/auto_examples/cluster/plot_kmeans_silhouette_analysis.html (accessed 01.11.2021).
14. Todd A., Thomson K., Kasim A., Bambra C. Cutting Care Clusters: the Creation of an Inverse Pharmacy Care Law? An area-level analysis exploring the clustering of community pharmacies in England. *BMJ Open Journal, England*, 2018. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6074640/> (accessed 01.11.2021).

15. VanderPlas J. Python Data Science Handbook: Essential Tools for Working with Data. O'Reilly Media, 2016.

16. Vivek S. Clustering algorithms for customer segmentation. Online publishing platform Medium. 2018. Available at: <https://towardsdatascience.com/clustering-algorithms-for-customer-segmentation-af637c6830ac> (accessed 01.11.2021).

17. Wijaya C. Y. Breaking down the agglomerative clustering process. Online publishing platform Medium. 2019. Available at: <https://towardsdatascience.com/breaking-down-the-agglomerative-clustering-process-1c367f74c7c2> (accessed 01.11.2021).

Сведения об авторах

Наталья Павловна Савина

кандидат экономических наук,
доцент кафедры мировой экономики
РЭУ им. Г. В. Плеханова.
Адрес: ФГБОУ ВО «Российский
экономический университет имени
Г. В. Плеханова», 117997, Москва,
Стремянный пер., д. 36.
E-mail: Savina.NP@rea.ru

Нарек Андраникович Галстян

аналитик данных
Novartis International AG.
Адрес: Novartis International AG,
125315, Москва,
Ленинградский проспект, д. 70.
E-mail: Galstyan.NA@novartis.com

Олег Валерьевич Литвишко

кандидат экономических наук,
доцент кафедры финансового
менеджмента РЭУ им. Г. В. Плеханова.
Адрес: ФГБОУ ВО «Российский
экономический университет имени
Г. В. Плеханова», 117997, Москва,
Стремянный пер., д. 36.
E-mail: Litvishko.OV@rea.ru

Екатерина Андреевна Закревская

кандидат экономических наук,
доцент кафедры математических
методов в экономике РЭУ им. Г. В. Плеханова.
Адрес: ФГБОУ ВО «Российский экономический
университет имени Г. В. Плеханова», 117997,
Москва, Стремянный пер., д. 36.
E-mail: Zakrevskaya.EA@rea.ru

Information about the authors

Natalya P. Savina

PhD, Assistant Professor
of the Department for World Economy
of the PRUE.
Address: Plekhanov Russian
University of Economics,
36 Stremyanny Lane, Moscow,
117997, Russian Federation.
E-mail: Savina.NP@rea.ru

Narek A. Galstyan

Data analyst of the Novartis International AG.
Address: Novartis International AG,
70 Leningradsky Avenue,
Moscow, 125315,
Russian Federation.
E-mail: Galstyan.NA@novartis.com

Oleg V. Litvishko

PhD, Assistant Professor
of the Department for Financial
Management of the PRUE.
Address: Plekhanov Russian
University of Economics,
36 Stremyanny Lane, Moscow,
117997, Russian Federation.
E-mail: Litvishko.OV@rea.ru

Ekaterina A. Zakrevskaya

PhD, Assistant Professor of the Department
for Mathematical Methods in Economics
of the PRUE.
Address: Plekhanov Russian University
of Economics, 36 Stremyanny Lane,
Moscow, 117997, Russian Federation.
E-mail: Zakrevskaya.EA@rea.ru