

**Ларченко Ю. Г., Котляров В. П.**  
**Y. G. Larchenko, V. P. Kotlyarov**

## **КЛАСТЕРИЗАЦИЯ КАК МЕТОД ГРУППИРОВКИ И АНАЛИЗА МНОГОМЕРНЫХ ОБЪЕКТОВ В ПРОМЫШЛЕННОСТИ**

## **CLUSTERING AS METHOD OF GROUPING AND ANALYSIS OF MULTIDIMENSIONAL OBJECTS IN INDUSTRY**

**Ларченко Юлия Геннадьевна** – кандидат экономических наук, доцент кафедры проектирования, управления и разработки информационных систем Комсомольского-на-Амуре государственного университета (Россия, Комсомольск-на-Амуре). E-mail: l\_uliya\_g@mail.ru.

**Yulia G. Larchenko** – PhD in Economics, Associate Professor at the Department of Design, Management and Development of Information Systems, Komsomolsk-na-Amure State University (Russia, Komsomolsk-on-Amur). E-mail: l\_uliya\_g@mail.ru.

**Котляров Валерий Петрович** – кандидат технических наук, доцент кафедры проектирования, управления и разработки информационных систем Комсомольского-на-Амуре государственного университета (Россия, Комсомольск-на-Амуре). E-mail: kvp@knastu.ru.

**Valery P. Kotlyarov** – PhD in Technical Sciences, Associate Professor at the Department of Design, Management and Development of Information Systems, Komsomolsk-na-Amure State University (Russia, Komsomolsk-on-Amur). E-mail: kvp@knastu.ru.

**Аннотация.** Работа посвящена кластеризации действующих предприятий крупного промышленного холдинга АО «Полиметалл» на основе наиболее значимых показателей, позволяющих оценить степень достижения стратегических целей хозяйствующих субъектов, направленных, в свою очередь, на прирост совокупного капитала компании согласно Международному стандарту интегрированной отчётности. При этом состав показателей, характеризующий стратегические цели компании, дифференцируется в зависимости от вида используемого капитала и определяется в годовых отчётах холдинга и в отчётах об устойчивом его развитии. В качестве основного метода кластеризации объектов исследования применяется алгоритм К-средних, выполненный в аналитической платформе Deductor. С помощью метода К-средних, предполагающего расчёт средневзвешенного расстояния предприятий до центра кластера, получены три кластера. В первый кластер входят предприятия Майское, Омолон, Светлое, второй кластер представлен предприятиями Кызыл и Воронцовское, третий кластер включает Дукат, Албазино-Амурск, Варваринское. В первом кластере определяющими показателями для формирования стали индикаторы, ориентированные на стратегию корпоративного управления и устойчивого развития, реализация которой, в свою очередь, обеспечивает прирост человеческого и природного капиталов. Второй кластер холдинга состоит из предприятий, нацеленных в первую очередь на стабильное производство и, соответственно, прирост финансового капитала. Третий кластер представлен предприятиями, реализующими в настоящий момент времени стратегии стабилизации производства и устойчивого развития, направленные на увеличение финансового и природного капиталов.

**Summary.** The work is devoted to the clustering of operating enterprises of the large industrial holding JSC «Polymetal» based on the most significant indicators that allow assessing the degree of achievement of the strategic goals of economic entities, aimed in turn at increasing the total capital of the company in accordance with the International Integrated Reporting Standard. At the same time, the composition of indicators characterizing the company's strategic goals is differentiated depending on the type of capital used and is determined in the holding's annual reports and reports on its sustainable development. The K-means algorithm implemented in the Deductor analytical platform is used as the main method for clustering research objects. Using the K-means method, which involves calculating the weighted average distance of enterprises to the center of the cluster, 3 clusters were obtained. The first cluster includes the following enterprises: Maiskoye, Omolon, and Svetloye, the second cluster is represented by enterprises: Kyzyl and Vorontsovskoye, the third cluster includes: Dukat, Albazino-Amursk, and Varvarinskoye. In the first cluster, indicators focused on the strategy of corporate governance and sustainable development, their implementation, in turn, ensures the growth of human and natural capital, became the determining indicators for their formation. The second cluster of the holding consists of enterprises aimed primarily at stable pro-

duction and, accordingly, the growth of financial capital. The third cluster is represented by enterprises currently implementing production stabilization and sustainable development strategies aimed at increasing financial capital and natural capital.

**Ключевые слова:** кластер, алгоритм К-средних, индикаторы, совокупный капитал, предприятие, стратегические цели.

**Key words:** cluster, K-means algorithm, indicators, total capital, company, strategic goals.

УДК 65.011

В современных условиях хозяйствования, характеризующихся усилением уровня конкурентной борьбы, перед коммерческими предприятиями возникает задача стабилизации и усиления своей рыночной позиции. В качестве одного из способов решения поставленной задачи предприятия отдельных отраслей промышленности рассматривают увеличение масштабов производства и реализации путём открытия новых объектов. В этом случае предприятия сталкиваются с необходимостью детализации общей стратегии развития применительно к своим хозяйствующим единицам. Поэтому при появлении большого количества многомерных объектов требуется их объединение в однородные группы с целью выработки внутригрупповой стратегической линии движения. Многие учёные и исследователи для классификации предприятий и объединения их в однородные группы предлагают кластерный анализ, который получил активное развитие во второй половине прошлого столетия. В развитие кластерного анализа как статистического метода классификации многомерных объектов в однородные группы по мере близости их расположения к центру внесли существенный вклад следующие зарубежные и отечественные учёные: Н. Джордайн, Дж. Эверит, М. Андерберг, М. Жамбю и Э. Диде, С. А. Айвазян, З. И. Бежаева, О. В. Староверова, Б. Г. Миркин, И. И. Елисеева, В. О. Рукавишников, И. Д. Мандель и др. [10; 13; 15]. Многообразие научных публикаций в области кластеризации, а также широкая область практического применения данного метода анализа определяют актуальность поставленной задачи исследования.

В настоящее время существует много подходов к проведению кластеризации объектов [11; 12; 16]. Первый подход основан на процедуре иерархической алгоритмизации. В соответствии с ним каждый объект рассматривается как отдельный кластер (группа, класс). Первоначально объединяются два ближних кластера, к которым далее присоединяется один или более схожих объектов, образующих новый кластер. Процедура укрупнения продолжается до формирования кластеров, в которых расстояние между объектами одной группы значительно меньше расстояния до объектов другой группы. Следующий подход подразумевает использование иерархического дивизимного метода (делимого метода). Этот алгоритм – обратный алгоритму агломерации. Согласно данному методу, все объекты изначально входят в один кластер. Затем берутся два наиболее удалённых объекта, образующих два кластера. Последующее разбиение на более мелкие классы продолжается до создания групп с максимально близкими по расстоянию объектами. В экономических исследованиях распространённым является итерационный алгоритм, реализующий метод К-средних. Этот метод основан на выполнении ряда операций:

1. Задаётся количество кластеров.
2. Осуществляется произвольное распределение объектов по кластерам.
3. Определяются координаты центра каждого кластера  $(x, y)$  как среднеарифметическое значение свойств (характеристик, показателей) объектов  $(x_i, y_i)$ :

$$(x, y) = \left( \frac{\sum_{i=1}^n x_i}{n}, \frac{\sum_{i=1}^n y_i}{n} \right).$$

4. Вычисляются расстояния от каждого объекта до центра соответствующего кластера (центра) как Евклидово расстояние:

$$\begin{cases} d(x_i, x) = \sqrt{\sum_{i=1}^n (x_i - x)^2} \\ d(y_i, y) = \sqrt{\sum_{i=1}^n (y_i - y)^2} \end{cases}.$$

5. Анализируется расстояние от объекта до центров выбранного кластера и других выделенных кластеров.

6. Проводится перегруппировка кластеров на основании результатов анализа расстояний объектов до центров тяжести образованных кластеров. Каждый объект присоединяется к тому кластеру, расстояние до центроида которого является наименьшим.

7. Определяется функция штрафов как сумма квадратов всех расстояний. Уменьшение функции штрафов свидетельствует об улучшении результата группировки объектов.

8. Операции перегруппировки объектов и последующего расчёта расстояний до центров тяжести новых кластеров повторяются многократно, пока функция штрафов не примет минимально возможное значение. В последней итерации координаты центров кластеров останутся без изменений.

Кластерный анализ требует проведения многочисленных и трудоёмких расчётов. Поэтому применение программных продуктов, позволяющих снизить временные, трудовые и финансовые затраты аналитика, является необходимой мерой. Одним из популярных программных инструментов проведения кластерного анализа является аналитическая платформа Deductor. Она позволяет провести весь цикл операций по анализу данных, начиная с импорта исходной информации, подбора необходимой модели её обработки и заканчивая получением результатов в табличной и/или графической форме.

Процесс кластеризации объектов в аналитической платформе Deductor включает в себя следующие этапы:

1. Выбор и обоснование показателей, позволяющих классифицировать и объединить многомерные объекты в однородные группы (кластеры), а также принять взвешенное управленческое решение относительно траектории их развития.

2. Составление матрицы значений показателей в разрезе исследуемых объектов. В матрице столбцами являются показатели объектов, а строки – это непосредственно названия единиц множества.

3. Преобразование табличной формы исходной информации в текстовый файл и импорт его в программу.

4. Настройка параметров текстового файла: назначение строковых данных (название объектов) – информационное, тип вещественных данных (показатели) – непрерывный, их назначение – входное.

5. Запуск процесса импорта исходных данных.

6. Запуск мастера обработки данных: Data Mining, кластеризация алгоритмом k-means.

7. Настройка назначений столбцов (информационное или входное); разбиение исходных данных на обучающее подмножество (100%-е обучение); установка фиксированного количества кластеров (k-means); запуск процесса кластеризации; выбор способов отображения результатов визуализации (например, профили кластеров, OLAP-куб в виде кросс-таблицы и кросс-диаграммы); настройка назначений полей куба (объект – информационное, показатели – факт, номер кластера – измерение); настройка размещения измерений (строки – это номера кластеров); настройка фактов и вариантов агрегации (расчёт среднего значения по показателям); завершение визуализации.

Как ранее было указано, одним из основных способов отображения результатов исследования является графическое представление кластеров (профили кластеров), которое показывает значимость показателей в процессе группировки объектов. Значимость показателей выражается в процентах и демонстрирует отклонение среднего группового значения параметра от среднего его значения по всей выборке. Чем больше полученное отклонение, тем выше значимость конкретного показателя для соответствующего кластера. Рассмотрим алгоритм кластеризации многомерных объектов методом К-средних применительно к крупному промышленному холдингу АО «Полиметалл».

АО «Полиметалл» – это компания по добыче драгоценных металлов в России и Казахстане. В структуру холдинга входят следующие действующие предприятия: Кызыл, Дукат, Албазино-Амурск, Майское, Омолон, Воронцовское, Варваринское и Светлое. В отношении них в холдинге сформулированы стратегические цели, связанные с обеспечением роста отдельных видов капиталов, согласно положениям Международного стандарта интегрированной отчетности (см. табл. 1) [3; 6; 14].

Таблица 1

Стратегические цели АО «Полиметалл» в 2020-2021 годах

Стратегическая цель	Ключевой показатель эффективности	Капитал
Стабильное производство	Выработка	Финансовый капитал
	Удельные совокупные затраты	
	Рентабельность инвестированного капитала	
Обеспечение роста	Строительство новых объектов	Производственный капитал
	Инвестиции в поиск новых месторождений	
Корпоративное управление и устойчивое развитие	Коэффициент травматизма с летальным исходом	Человеческий капитал
	Коэффициент несчастных случаев	
	Доля женщин-специалистов	
	Удельное технологическое потребление воды	Природный капитал
	Удельные выбросы парниковых газов	
	Удельный вес вторично используемых отходов	
Инвестиции в будущее	Энергоёмкость	Интеллектуальный капитал
	Инвестиции в обучение персонала	
	Социальные инвестиции	
	Инвестиции в охрану окружающей среды	Социальный капитал

Из табл. 1 видно, что оценка степени достижения стратегических целей холдинга, направленных на прирост всех видов капиталов, осуществляется посредством ключевых показателей эффективности [1; 5; 7]. При этом совокупность оценочных индикаторов указывается в годовом отчёте и отчёте об устойчивом развитии компании за анализируемый период времени.

Следует отметить, что стратегические цели компании сформулированы обобщённо безотносительно к её бизнес-единицам. Поэтому для конкретизации и приоритизации стратегических целей действующих предприятий АО «Полиметалл» необходимо провести их кластеризацию. В качестве группировочных критериев рассматривались следующие показатели: доля женщин, коэффициент травматизма на 1000 человек, энергоёмкость, удельное потребление свежей воды на технологические нужды, доля утилизированных отходов, удельные выбросы парниковых газов, удельные совокупные затраты (по всем действующим предприятиям, кроме Дуката, в расчёте на унцию золотого эквивалента, а по Дукату – на унцию серебряного эквивалента), рентабельность продаж (по прибыли до вычета процентов, налогов и амортизации). Необходимо отметить, что состав показателей ограничивается параметрами, характеризующими реализацию стратегических целей в направлениях прироста финансового, производственного, человеческого и природного ка-

питалов [2; 4; 8; 9]. В процессе кластерного анализа не рассматриваются индикаторы интеллектуального и социального капиталов, т. к. характеристика стратегических целей по их увеличению не представлена в разрезе действующих предприятий АО «Полиметалл». Числовые значения показателей представлены в табл. 2.

Таблица 2

Исходная информация для проведения кластерного анализа

Действующее предприятие	Доля женщин	Коэффициент травматизма на 1000 чел.	Энергоёмкость, ГДж/тыс. унций золотого эквивалента	Удельное потребление свежей воды на технологические нужды, м <sup>3</sup> /тыс. тонн переработанной руды	Доля утилизированных отходов	Удельные выбросы парниковых газов, тонн СО <sup>2</sup> -эквивалент / тыс. тонн переработанной руды	Удельные совокупные затраты, долл. / унцию	Рентабельность продаж (по прибыли до вычета процентов, налогов и амортизации)
Кызыл	0,180	1,54	3783,2	241	0,020	60,0	554,0	0,781
Дукат	0,190	1,19	4960,6	153	0,300	58,7	11,5	0,494
Албазино-Амурск	0,164	1,57	4361,9	236	0,374	58,2	946,0	0,569
Майское	0,120	1,02	5078,8	46	0,450	50,8	1020,0	0,572
Омолон	0,110	2,78	2330,4	65	0,470	26,0	773,0	0,648
Воронцовское	0,290	0,00	4239,8	43	0,790	45,8	679,0	0,669
Варваринское	0,170	0,00	6917,0	215	0,380	55,6	1179,0	0,440
Светлое	0,139	0,00	2777,5	68	0,950	19,3	484,0	0,770

На основе данных табл. 2 получены результаты кластеризации действующих предприятий АО «Полиметалл» в платформе Deductor, которые показаны на рис. 1 и 2:

- «0» кластер включает предприятия Майское, Омолон и Светлое;
- «1» кластер включает предприятия Кызыл и Воронцовское;
- «2» кластер включает предприятия Дукат, Албазино-Амурск и Варваринское.

В «0» кластере наиболее значимыми показателями для формирующих его предприятий стали: энергоёмкость, удельные выбросы парниковых газов, доля женщин и удельное потребление свежей воды на технологические нужды. Высокая значимость данных показателей (от 62,5 до 80,6 %) внутри полученной группы показателей свидетельствует об ориентации предприятий (Майское, Омолон и Светлое) на стратегические цели, связанные с корпоративным управлением и устойчивым развитием и направленные в первую очередь на обеспечение прироста человеческого и природного капиталов.

Предприятие	Доля женщин	Коэффициент травматизма на 1000 чел.	Энергоемкость, ГДж / тыс. унций золотого эквивалента	Удельное потребление свежей воды на технологические нужды, м3 / тыс. тонн переработанной руды	Доля утилизированных отходов	Удельные выбросы парниковых газов, тонн CO2-эквивалент / тыс. тонн переработанной руды	Удельные совокупные затраты, долл. / унцию	Рентабельность по EBITDA	Номер кластера /	Расстояние до центра кластера
Дукат	0,190	1,19	4960,6	153	0,300	58,7	11,5	0,494	2	0,671
Албазино - Амурск	0,164	1,57	4361,9	236	0,374	58,2	946,0	0,569	2	0,471
Варваринское	0,170	0,00	6917,0	215	0,380	55,6	1179,0	0,440	2	0,646
Кызыл	0,190	1,54	3783,2	241	0,020	60,0	554,0	0,781	1	0,909
Воронцовское	0,290	0,00	4239,8	43	0,790	45,8	679,0	0,669	1	0,809
Майское	0,120	1,02	5078,8	46	0,450	50,8	1020,0	0,572	0	0,719
Омолон	0,110	2,78	2330,4	65	0,470	26,0	773,0	0,648	0	0,638
Светлое	0,139	0,00	2777,5	68	0,950	19,3	484,0	0,770	0	0,781

Рис. 1. Таблица показателей кластеров АО «Полиметалл»

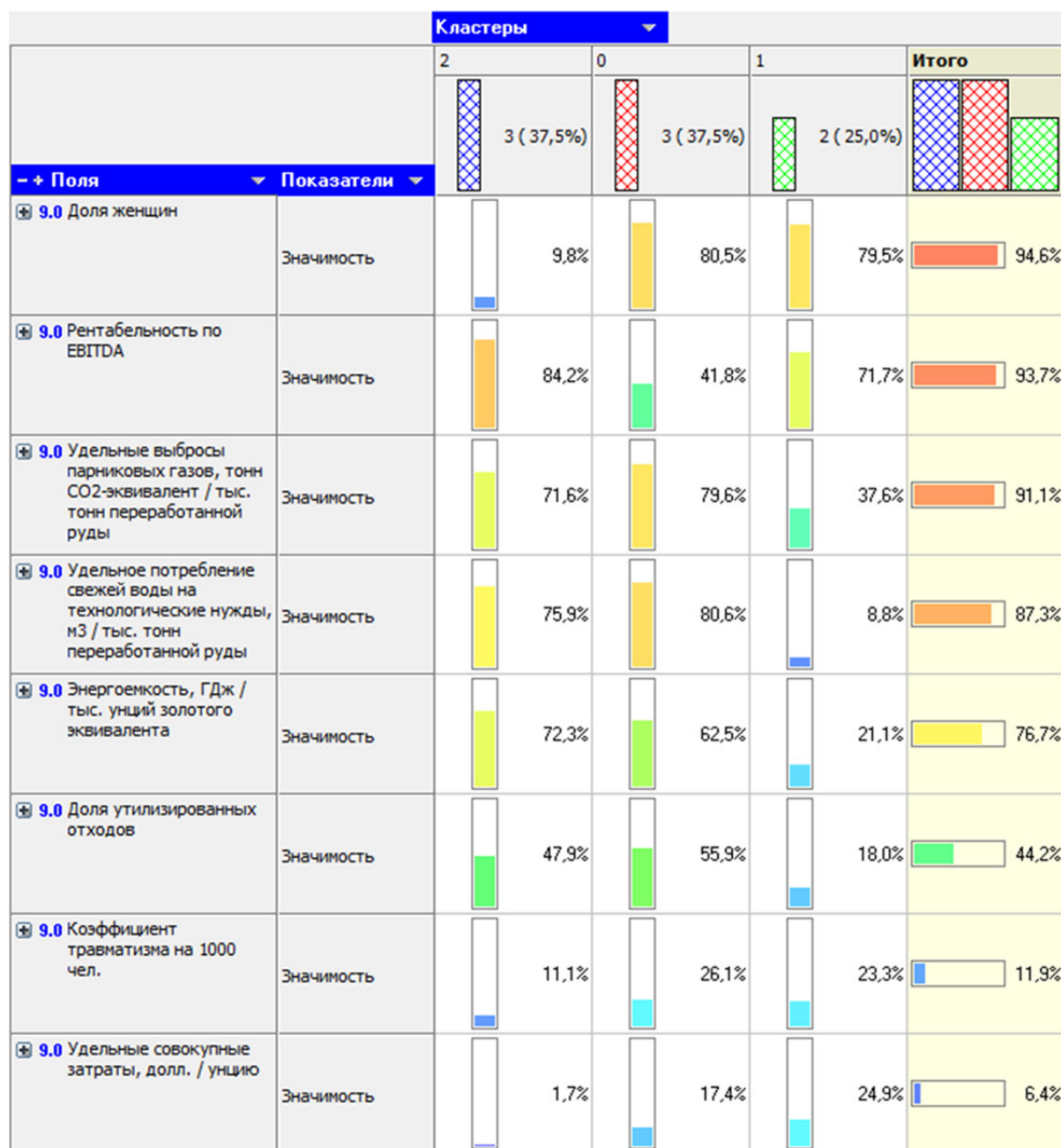


Рис. 2. Профили кластеров АО «Полиметалл»

Второй кластер («2») также состоит из предприятий, нацеленных на стратегию корпоративного управления и устойчивого развития. Индикаторами достижения стратегической цели развития являются: удельные выбросы парниковых газов, энергоёмкость и удельное потребление свежей воды на технологические нужды (важность факторов от 71,6 до 75,9 %). Однако вместе с показателями человеческого и природного капиталов для объектов, вошедших во второй кластер, существенным стратегическим параметром является значение показателя рентабельности продаж, характеризующего эффективность использования финансового капитала предприятий. Следовательно, для предприятий второго кластера актуальной на данный момент времени является смешанная стратегия стабилизации производства, корпоративного управления и устойчивого развития.

Заключительный кластер («1») включает предприятия, сгруппированные по двум важным факторам: доля женщин и рентабельность продаж (значимость 79,5 % и 71,7 % соответственно). Значит, в этом классе стратегический приоритет предприятий также смещается в сторону корпоративного управления и устойчивого развития в совокупности со стабилизацией производства.

Кластеризация многомерных объектов на примере действующих предприятий АО «Полиметалл» позволила сгруппировать бизнес-единицы и установить приоритет стратегических целей их развития в настоящее время. По результатам проведённого кластерного анализа бизнес-единицы холдинга разделены на три группы. Для предприятий всех кластеров выявлена необходимость сосредоточения усилий на реализации стратегии корпоративного управления и устойчивого развития. Однако предприятиям первого и второго кластеров также рекомендуется акцентировать внимание на разработке мероприятий по достижению стратегии стабилизации производства. Вместе с тем отметим, что несмотря на одинаковую стратегическую направленность предприятий разных кластеров на корпоративное управление и устойчивое развитие, существуют расхождения в планах их текущего и оперативного управления. Это объясняется вариативностью значимости группировочных показателей. Поэтому предприятиям внутри сформированных классов следует планировать свою текущую деятельность на основе декомпозиции наиболее значимых внутригрупповых показателей, что в свою очередь позволит сформировать конкретные оперативные мероприятия по достижению стратегических целей.

## ЛИТЕРАТУРА

1. Дубинина, Н. А. Показатели оценки бизнес-процессов предприятия / Н. А. Дубинина // Вестник Пермского университета. Сер. «Экономика». – 2016. – № 2 (29). – С. 179-191.
2. Кизиль, Е. В. Регулирование развития региональных социально-экономических систем: методологический аспект / Е. В. Кизиль // Учёные записки Комсомольского-на-Амуре государственного технического университета. Науки о человеке, обществе и культуре. – 2018. – № III-2 (35). – С. 96-99.
3. Кириченко, Л. П. Экономическая эффективность формирования и использования человеческого капитала корпорации / Л. П. Кириченко, А. С. Мотырева // Учёные записки Комсомольского-на-Амуре государственного технического университета. Науки о человеке, обществе и культуре. – 2021. – № IV-2 (52). – С. 115-120.
4. Ларченко, Ю. Г. Рентабельность собственного капитала как интегральный показатель экономической эффективности предприятия / Ю. Г. Ларченко, В. П. Котляров // Учёные записки Комсомольского-на-Амуре государственного технического университета. Науки о человеке, обществе и культуре. – 2022. – № II-2 (58). – С. 96-102.
5. Мерзликина, Г. С. Эффективное управление по изменениям: методическое и инструментальное обеспечение / Г. С. Мерзликина, Т. Е. Кожанова // Вестник Воронежского государственного университета инженерных технологий. – 2017. – Т. 79. – № 2. – С. 248-256.
6. Международный стандарт интегрированной отчётности. – URL: [https://www.integratedreporting.org/wp-content/uploads/2015/03/13-12-08-THE-INTERNATIONAL-IR-FRAMEWORK.docx\\_en-US\\_ru-RU.pdf](https://www.integratedreporting.org/wp-content/uploads/2015/03/13-12-08-THE-INTERNATIONAL-IR-FRAMEWORK.docx_en-US_ru-RU.pdf) (дата обращения: 05.01.2023). – Текст: электронный.
7. Растова, Ю. И. Стратегическое управление современной организацией: эффект синергии концепций / Ю. И. Растова, М. А. Растов // Управленческие науки. – 2018. – № 8 (3). – С. 20-31.
8. Усанов, И. Г. Трансформационный менеджмент: формы преобразования бизнеса / И. Г. Усанов, Г. И. Усанов // Учёные записки Комсомольского-на-Амуре государственного технического университета. Науки о человеке, обществе и культуре. – 2022. – № VI-2 (62). – С. 124-130.

9. Яковлева, Т. А. Повышение экономической эффективности деятельности организации на основе аутсорсинга / Т. А. Яковлева, Т. М. Пляскина // Учёные записки Комсомольского-на-Амуре государственного технического университета. Науки о человеке, обществе и культуре. – 2020. – № IV-2 (44). – С. 94-98.
10. Alexandru, V. A., Bolisani, E., Andrei A. G., Cegarra-Navarro, J. G., Martínez, A. M., Paiola, M., Scarso, E., Vatamanescu, E.-M., Zieba, M. Knowledge management approaches of small and medium-sized firms: a cluster analysis. *Kybernetes*, 2020, no. 49, p. 73-87. DOI: 10.1108/K-03-2019-0211.
11. Burdakova, G., Byankin, A. The Research of Cluster Initiatives of a Higher Education Institution in a Priority Development Area. *Proceedings of the international scientific conference far east con (2018) (ISCFEC 2018)*, 2018, (Atlantis Press, 2019), vol. 47, p. 367-370. DOI: 10.2991/iscfec-18.2019.94.
12. Filipova, A. G., Inzartsev, A. V., Vysockaya, A. V. Educational inequality in Russian regions: Mathematical modeling. *Current Problems and Ways of Industry Development: Equipment and Technologies. Lecture Notes in Networks and Systems*, Springer, Cham, 2021, 200 p., p. 948-958. DOI: 10.1007/978-3-030-69421-0\_105.
13. Kliestik, T., Valaskova, K., Lazaroiu, G., Kovacova, M., Vrbka, J. Remaining Financially Healthy and Competitive: The Role of Financial Predictors. *Journal of Competitiveness*, 2020, no. 12, P. 74-92. DOI: 10.7441/joc.2020.01.05.
14. АО «Полиметалл»: сайт. – Санкт-Петербург, 2021. – URL: <https://www.polymetalinternational.com> (дата обращения: 05.01.2023). – Текст: электронный.
15. Radeva, I. Multicriteria Fuzzy Sets Application in Economic Clustering Problems. *Cybernetics and information technologies*, 2017, no. 17, p. 1311-9702. DOI: 10.1515/cait-2017-0028.
16. Zharikova, E. P., Grigoryev, Y. Y., Grigoryeva, A. L. Application of neural networks for water area analysis. *Marine Intellectual Technologies. Issue 2*, 2021, no. 2, p. 129-133. Part 2. DOI: 10.37220/MIT.2021.52.2.063.