



АҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР
ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ
INFORMATION AND COMMUNICATION TECHNOLOGIES

КИБЕРНЕТИКА
КИБЕРНЕТИКА
CYBERNETICS

DOI 10.51885/1561-4212_2024_3_217

MPHTI 28.23.01

З.Т. Хасенова¹, М.А. Рақышева¹, Н.С. Глазырина², А.Н. Исайнова², А.Т. Мукатов²

¹НАО «Восточно-Казахстанский технический университет имени Д. Серикбаева»,

г. Усть-Каменогорск, Казахстан

*E-mail: zkhasenova@edu.ektu.kz**

E-mail: mrakysheva@edu.ektu.kz

²НАО «Евразийский национальный университет имени Л.Н. Гумилева»,

г. Астана, Республика Казахстан

E-mail: glazyrina_ns_1@enu.kz

E-mail: issainova_an@enu.kz

E-mail: abzal_mukatov@mail.ru

СЕНСОРНЫЕ ТЕХНОЛОГИИ И МАШИННОЕ ОБУЧЕНИЕ: НОВЫЕ ВОЗМОЖНОСТИ ДЛЯ ЭКОЛОГИЧЕСКОГО МОНИТОРИНГА

СЕНСОРЛЫҚ ТЕХНОЛОГИЯЛАР МЕН МАШИНАЛЫҚ ОҚЫТУ: ЭКОЛОГИЯЛЫҚ МОНИТОРИНГТІҢ ЖАҢА МҮМКІНДІКТЕРІ

SENSOR TECHNOLOGIES AND MACHINE LEARNING: NEW OPPORTUNITIES FOR ENVIRONMENTAL MONITORING

Аннотация. В данной работе исследуется применение методов статистического машинного обучения для калибровки и анализа данных, полученных от системы искусственного обоняния при экспериментальных исследованиях со стандартными газовыми смесями, включающими ключевые загрязнители атмосферного воздуха (диоксид углерода, оксид углерода, диоксид азота, аммиак и сероводород) в диапазоне концентраций от 5 до 50 ppm. Актуальность исследования обусловлена необходимостью разработки точных и надежных методов мониторинга качества воздуха в условиях низких концентраций загрязняющих веществ. В качестве методов машинного обучения были выбраны многомерная линейная регрессия и кластеризация на основе метода ближайших соседей. Полученные модели продемонстрировали высокую степень адекватности экспериментальным данным, о чем свидетельствуют значения коэффициента детерминации R^2 , близкие к единице. С использованием алгоритма k -means была успешно проведена кластеризация многомерных откликов сенсоров, позволившая выявить четкую зависимость между характеристиками сенсорных сигналов, видом газа и его концентрации в смеси. Результаты исследования могут быть использованы для создания автономных систем мониторинга качества воздуха, способных оперативно выявлять превышения предельно допустимых концентраций вредных веществ.

Ключевые слова: система искусственного обоняния, газовый сенсор, электронный нос, машинное обучение, статистическое машинное обучение, многомерная линейная регрессия, алгоритм k -means.

Аңдатпа. Бұл жұмыста стандартты газ қоспаларымен жүргізілген эксперименттік зерттеулерде атмосфералық ауаның негізгі ластаушылары (көмірқышқыл газы, көміртек оксиді, азот диоксиді, аммиак және күкіртсутек) концентрациясы 5-тен 50 ppm-ге дейін

болатын жасанды иіс сезу жүйесінен алынған деректерді калибрлеу және талдау үшін статистикалық машиналық оқыту әдістерін қолдану зерттеледі. Зерттеудің өзектілігі – ластаушы заттардың төмен концентрациясы жағдайында ауа сапасын бақылаудың дәл және сенімді әдістерін әзірлеу қажеттілігімен байланысты. Машиналық оқыту әдістері ретінде көпөлшемді сызықтық регрессия және ең жақын көршілер әдісі негізіндегі кластерлеу таңдалды. Алынған модельдер тәжірибелік деректерге жоғары сәйкестік көрсетті, бұл R^2 детерминация коэффициентінің мәндерімен дәлелденеді, олар бірлікке жақын. k -means алгоритмін пайдалана отырып, сенсорлардың көпөлшемді жауаптарын сәтті кластерлеу жүргізіліп, сенсорлық сигналдардың сипаттамалары, газ түрі және оның қоспадағы концентрациясы арасындағы айқын тәуелділік анықталды. Зерттеу нәтижелері ауа сапасын бақылаудың автономды жүйелерін құру үшін пайдаланылуы мүмкін, олар зиянды заттардың шекті рұқсат етілген концентрациясынан асып кетуді жедел анықтай алады.

Түйін сөздер: жасанды иіс сезу жүйесі, газ сенсоры, электронды мұрын, машиналық оқыту, статистикалық машиналық оқыту, көпөлшемді сызықтық регрессия, k -means алгоритмі.

Abstract. This paper investigates the application of statistical machine learning methods for the calibration and analysis of data obtained from an artificial olfaction system during experimental studies with standard gas mixtures, which include key air pollutants (carbon dioxide, carbon monoxide, nitrogen dioxide, ammonia, and hydrogen sulfide) in concentrations ranging from 5 to 50 ppm. The relevance of the research is due to the need for accurate and reliable methods for monitoring air quality in conditions of low pollutant concentrations. Multivariate linear regression and clustering based on the nearest neighbors method were chosen as the machine learning methods. The resulting models demonstrated a high degree of agreement with the experimental data, as evidenced by the R^2 coefficient values, which were close to one. Using the k -means algorithm, successful clustering of the multivariate sensor responses was carried out, revealing a clear relationship between sensor signal characteristics, gas type, and its concentration in the mixture. The research results can be used to create autonomous air quality monitoring systems capable of promptly detecting exceedances of the maximum allowable concentrations of harmful substances.

Keywords: artificial olfaction system, gas sensor, electronic nose, machine learning, statistical machine learning, multivariate linear regression, k -means algorithm

Введение. Загрязнение воздуха – одна из наиболее острых экологических проблем современности, особенно в крупных городах и промышленных регионах. Выбросы вредных веществ в атмосферу от промышленных предприятий, транспорта, энергетики и других источников приводят к серьезным последствиям для окружающей среды и здоровья человека: к повышенным рискам развития у населения сердечно-сосудистых заболеваний, рака, хронических обструктивных заболеваний легких и других заболеваний. Кроме того, ухудшение качества воздуха вызывает значительные экономические потери из-за снижения производительности труда, увеличения расходов на здравоохранение, ухудшения качества жизни и миграции населения.

Промышленные города и регионы особенно подвержены проблемам загрязнения воздуха из-за высокой концентрации производственных предприятий. Из-за выбросов от металлургических, химических, энергетических и других предприятий наблюдается повышенный уровень загрязнения воздуха различными вредными веществами. Загрязнение воздуха в промышленных регионах часто носит комплексный характер, что затрудняет оценку его воздействия на здоровье населения.

Таким образом, загрязнение воздуха является серьезной проблемой, требующей комплексного решения. Разработка и внедрение эффективных систем мониторинга качества воздуха является важным шагом в решении этой проблемы.

Традиционные методы анализа качества воздуха (газовая и жидкостная хроматография, масс-спектрометрия, ядерный магнитный резонанс и спектрофотометрия), основанные на физико-химических принципах, хотя и обеспечивают высокую точность, характеризуются высокой стоимостью и сложностью обслуживания (Bogowik et al., 2020).

В качестве перспективной альтернативы предлагается использовать системы искусственного обоняния (для их описания также используется термин «электронный нос»), основанные на использовании матрицы газовых сенсоров и системы распознавания образов (Hurot et al., 2020). «Электронные носы» обладают рядом преимуществ: они компактны, энергоэффективны и могут быть легко интегрированы в различные системы мониторинга. Благодаря способности имитировать человеческое обоняние, «электронные носы» позволяют быстро и точно определять наличие различных летучих органических соединений и других загрязняющих веществ в воздухе (Yan et al., 2015). Их применение охватывает широкий спектр областей от пищевой промышленности (Konduru et al., 2015; Wojnowski et al., 2017; Maimunah Mohd et al., 2020) и медицины (Adiguzel, Kulah, 2015; Sujono, Amin, 2018; Wilson, 2018) до экологического мониторинга (Harrou et al., 2018; Baby, Cabezas & Walsøe de Reca, 2000) и безопасности (Norman et al., 2003; Kanakam, Hussain & Chakravarthy, 2015).

Ранее нами были представлены результаты исследований по созданию мобильного мультисенсорного газоанализатора с функцией беспроводной передачи сенсорных данных (Сатыбалдина, et al., 2023), разработаны программные средства для сбора автоматического выделения признаков, разметки, датасетов и обучения нейросетевого классификатора газов и запахов на языке Python (Мукатов et al., 2024; Сатыбалдина, Сыдыков & Исайнова, 2023), получены оценки точности распознавания исследуемых образцов при варьировании исследуемых газов, паров спиртов, летучих органических соединений (Alseitov et al., 2023; Satybaldina et al., 2023; Сатыбалдина et al., 2024).

В настоящей работе представлены результаты исследований по разработке методов калибровки сенсоров и анализа сенсорных данных, полученных от систем искусственного обоняния, для решения задач мониторинга качества воздуха. Применение методов статистического машинного обучения, таких как линейная регрессия и кластеризация, позволяет создавать точные и надежные модели калибровки сенсоров, что повышает точность определения концентраций различных газов и способствует решению проблемы загрязнения окружающей среды.

Материалы и методы исследования. В табл. 1 представлены данные по образцам газов, использованных в экспериментах. Выбор данных газов обусловлен тем, что именно уровень их содержания в воздухе контролируется на станциях мониторинга качества воздуха в связи с их вредным воздействием на здоровье людей. Стандартные смеси синтетического/сжатого воздуха и целевых газов в известных соотношениях (концентрациях исследуемых газов в ppm) для калибровки датчиков «электронного носа» были предоставлены компанией ООО «Линде Газ Рус», Россия.

Таблица 1. Концентрации целевого газа в смеси с синтетическим воздухом

№	Целевой газ	Концентрации целевого газа в смеси с синтетическим воздухом, ppm
1	Оксид азота NO ₂	5 ppm, 10 ppm, 20 ppm, 35 ppm и 50 ppm
2	Аммиак NH ₃	5 ppm, 10 ppm, 20 ppm, 35 ppm и 50 ppm
3	Сероводород H ₂ S	5 ppm, 10 ppm, 20 ppm, 35 ppm и 50 ppm
4	Монооксид углерода (угарный газ) CO	5 ppm, 10 ppm, 20 ppm, 35 ppm и 50 ppm
5	Диоксид углерода (углекислый газ) CO ₂	5 ppm, 10 ppm, 20 ppm, 35 ppm и 50 ppm

Примечание – составлено авторами

Экспериментальные измерения проведены на экспериментальной установке (рис. 1) в Лаборатории наноматериалов Сколковского института науки и технологий, Сколково, Россия. Использован «электронный нос» собственной разработки, на которую получен патент РК на полезную модель (Патент РК №9539, 2024).

В табл. 2 представлены характеристики и обозначения использованных металлооксидных газовых сенсоров.

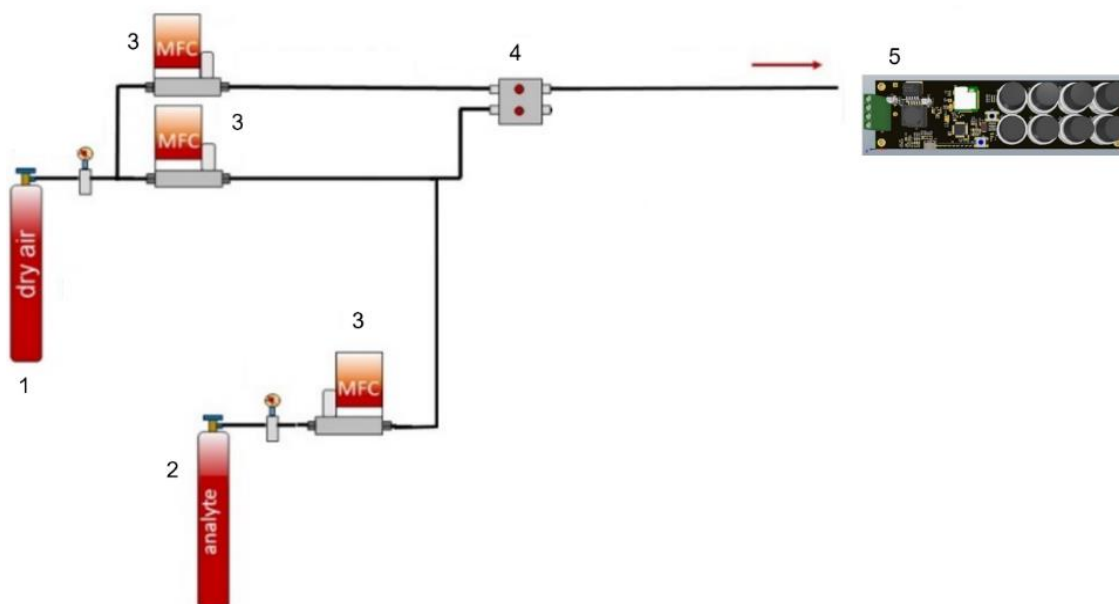


Рисунок 1. Экспериментальная установка с регулируемой подачей газов, объединенная с прибором типа «электронный нос»

Примечание – составлено авторами

Таблица 2. Обозначение сенсоров

Наименование марки сенсора	Перечень газов, к которым сенсор числителен	Обозначение сенсора
MQ-7	CO	R1
MQ-3	Alcohol	R2
MQ-5	LPG	R3
MQ-2	Smoke	R4
MQ-9	CO, CH ₄ , LPG	R5
MQ-8	H ₂	R6
MQ-135	NH ₃ , CO ₂ , NO ₂	R7
MQ-4	CH ₄	R8

Примечание – составлено авторами

Результаты и их обсуждение. Основная цель работы состоит в том, чтобы обеспечить точное калибрование сенсоров «электронного носа» для измерения концентраций целевых газов. Для этого предлагается подход, при котором калибровочные кривые

сенсоров строятся на основе метода наименьших квадратов при анализе поверочной газовой смеси в разных концентрациях. Предложенный подход позволяет улучшить точность и воспроизводимость измерений сенсоров «электронного носа», что имеет ключевое значение для применения «электронного носа» в таких областях, как мониторинг окружающей среды и промышленные процессы.

Экспериментальные данные были получены в соответствии с проведенным протоколом эксперимента, в котором поверочная газовая смесь подавалась в концентрациях 5 ppm, 10 ppm, 20 ppm, 35 ppm и 50 ppm в течение 10 минут. Между подачей аналита на «электронный нос» подавался воздух в течение 15 минут. Эксперимент повторялся для каждого образца по два раза. На рис. 2 показаны графики напряжений всех сенсоров «электронного носа» в течение двух экспериментов.

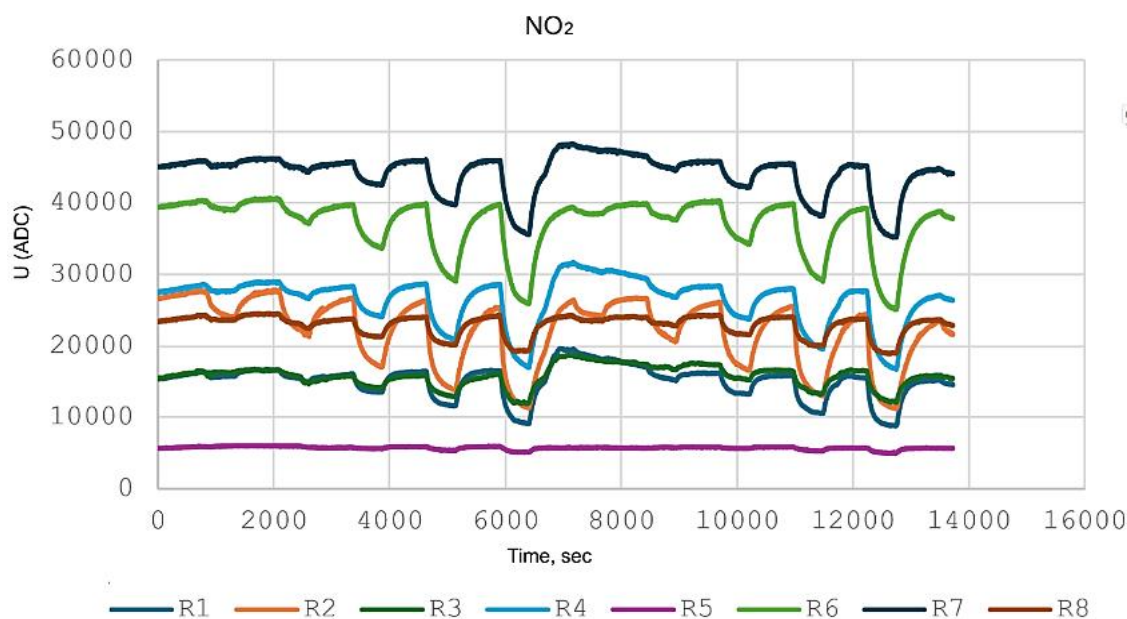


Рисунок 2. Графики временной зависимости электрических характеристик сенсоров R1-R8 «электронного носа» при подаче на него поверочной газовой смеси сухого воздуха и оксида азота (NO₂)

Примечание – составлено авторами

Металлооксидные хемистезисные газовые сенсоры, используемые в «электронном носе», – это устройства, основанные на изменении электрического сопротивления тонкой пленки металлооксидного полупроводника при взаимодействии с газами. При контакте с газовой смесью молекулы газов адсорбируются на поверхности полупроводника. Тип адсорбции зависит от температуры сенсора и природы газа. Адсорбция газов приводит к изменению концентрации носителей заряда в полупроводнике, что в свою очередь вызывает изменение его электрического сопротивления. Изменение сопротивления регистрируется с помощью электродов, нанесенных на поверхность сенсора. Полученный аналоговый сигнал обрабатывается аналогово-цифровым устройством и записывается в файл в виде цифрового сигнала напряжения отдельно для каждого сенсора и далее используется для обработки и анализа.

Рассмотрим подробно алгоритм построения калибровочных кривых сенсоров на примере сенсора R2. Для построения калибровочной кривой сенсора R2 отдельно рассмотрим временную зависимость электрических характеристик R2, представленную на рис. 3.

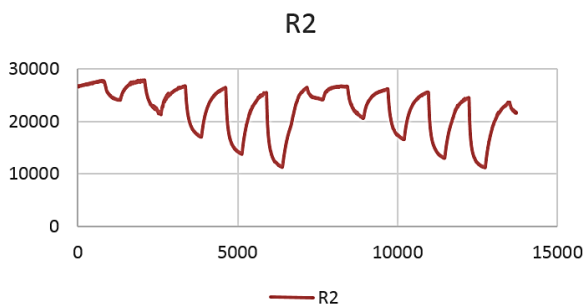


Рисунок 3. Графики временной зависимости электрических характеристик сенсоров R1-R8 «электронного носа» при подаче на него поверочной газовой смеси сухого воздуха и оксида азота (NO₂)

Примечание – составлено авторами

Рассчитаем среднее значение нулевого смещения (U_{air}) сенсора R2. Для этого на графике напряжения сенсора R2 были выбраны две точки на базовых линиях сенсора – точка 2112 со значением в этой точке 27218 в первом эксперименте и точка 8452 со значением 26116 во втором эксперименте (рис. 4).

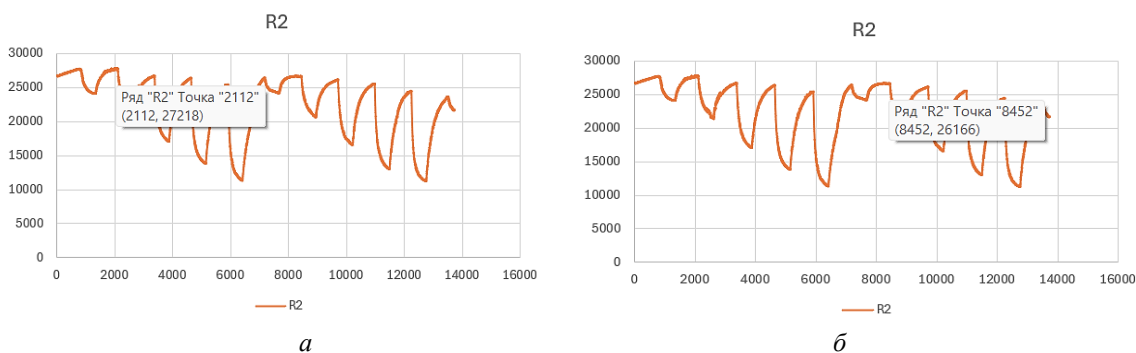


Рисунок 4. Выбранные точки на базовых линиях: а – точка на базовой линии первого эксперимента; б – точка на базовой линии второго эксперимента

Примечание – составлено авторами

В наборе полученных экспериментальных данных к выбранным точкам было добавлено по 5 точек от выбранной точки, из которых было вычислено среднее значение нулевого смещения (табл. 3).

Таблица 3. Данные для расчета U_{air}

Time	R1	R2	R3	R4	R5	R6	R7	R8
1	15452	26759	15482	27571	5733	39391	45027	23452
2	15367	26620	15405	27607	5753	39370	45031	23477
...								

2106	16529	27481	16514	28846	5982	40452	46129	24320
2107	16526	27483	16613	28860	6027	40484	46046	24368
2108	16524	27335	16483	28806	5902	40357	46050	24344
2109	16490	27359	16470	28835	6008	40405	46106	24371

Окончание таблицы 3

Time	R1	R2	R3	R4	R5	R6	R7	R8
2110	16460	27308	16560	28730	6011	40407	46115	24339
2111	16467	27316	16523	28788	6004	40400	46080	24359
2112	16461	27218	16563	28803	6020	40313	46041	24244
2113	16430	27213	16501	28710	5990	40325	46082	24290
2114	16449	27129	16570	28708	6057	40293	45946	24304
2115	16448	27096	16505	28772	6033	40230	46056	24249
2116	16445	27033	16421	28665	6024	40246	46036	24194
2117	16433	26959	16563	28661	5952	40195	45936	24187
2118	16378	26931	16488	28599	5979	40256	45961	24174
...								
8446	16991	26487	17353	29244	5823	39839	46496	24052
8447	16957	26445	17424	29291	5743	39851	46572	23876
8448	16983	26414	17335	29308	5689	39856	46484	23983
8449	16999	26330	17461	29276	5742	39773	46540	23927
8450	16905	26341	17375	29292	5782	39860	46542	23914
8451	16919	26294	17434	29316	5793	39831	46474	23979
8452	16882	26166	17364	29243	5723	39681	46449	23980
8453	16891	26178	17327	29317	5762	39704	46410	23906
8454	16885	26144	17344	29231	5713	39641	46427	23865
8455	16935	26121	17269	29243	5801	39677	46295	23923
8456	16855	26089	17248	29219	5705	39713	46371	23879
8457	16842	25991	17330	29070	5755	39703	46365	23903
8458	16868	25941	17290	29173	5765	39609	46350	23854
...								

Примечание – составлено авторами

Аналогично были рассчитаны средние значения нулевых смещений для всех сенсоров «электронного носа». Средние значения U_{air} каждого сенсора приведены в табл. 4.

Таблица 4. Данные для расчета U_{air}

	R1	R2	R3	R4	R5	R6	R7	R8
U_{air}	16690,27	26725,55	16940,14	29006,55	5874,364	40042,95	46246,5	24108,36

Примечание – составлено авторами

Далее были рассчитаны средние значения максимальных напряжений (U_{max}) сенсоров при подаче на «электронный нос» поверочной газовой смеси NO_2 в

концентрациях 5 ppm, 10 ppm, 20 ppm, 35 ppm и 50 ppm. Для этого были выбраны точки на графике, соответствующие максимальным значениям напряжения при различных концентрациях аналита. Выбранные точки и значения напряжения сенсора R2 в этих точках при подаче поверочной газовой смеси NO₂ в концентрации 5 ppm для каждого эксперимента показаны на рис. 5.

В наборе полученных экспериментальных данных к выбранным точкам было добавлено по 5 точек от выбранной точки, из которых было вычислено среднее значение максимального напряжения сенсора R2 (табл. 5).

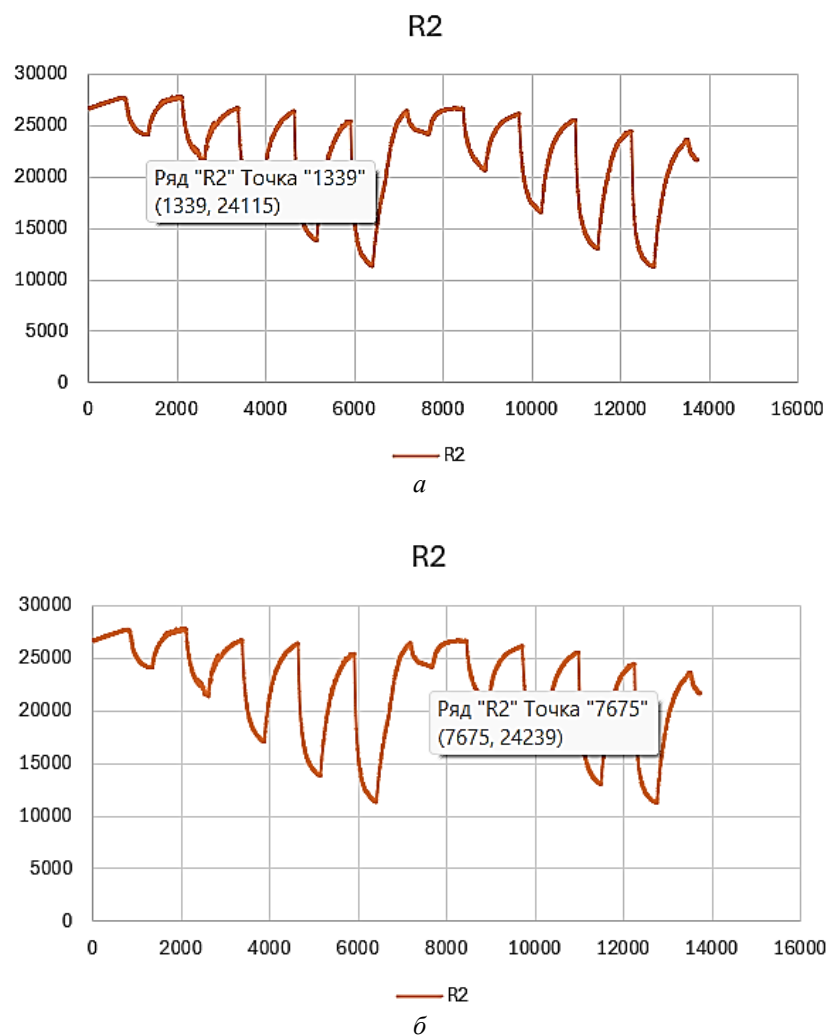


Рисунок 5. Выбранные точки и значения напряжения сенсора R2 в этих точках при концентрации 5 ppm: *a* – точка U_{\max} при концентрации 5 ppm NO₂ в первом эксперименте; *б* – точка U_{\max} для концентрации 5 ppm NO₂ во втором эксперименте
Примечание – составлено авторами

Таблица 5. Данные для расчета U_{air}

Time	R1	R2	R3	R4	R5	R6	R7	R8
1	15452	26759	15482	27571	5733	39391	45027	23452

2	15367	26620	15405	27607	5753	39370	45031	23477
...								
1333	15871	24122	16365	27787	6003	39095	45416	23713
1334	15878	24149	16316	27751	6076	39106	45382	23636

Окончание таблицы 5

Time	R1	R2	R3	R4	R5	R6	R7	R8
1335	15904	24118	16377	27831	6004	39174	45390	23725
1336	15856	24113	16390	27777	6050	39109	45418	23657
1337	15908	24051	16384	27786	6016	39108	45464	23772
1338	15852	24066	16373	27802	6064	39062	45281	23722
1339	15784	24115	16268	27762	6015	39055	45372	23685
1340	15848	24094	16415	27841	6052	39178	45454	23719
1341	15892	24129	16327	27840	6027	39126	45559	23797
1342	15894	24153	16403	27761	5985	39143	45450	23746
1343	15906	24142	16339	27867	6031	39176	45379	23703
1344	15907	24109	16297	27807	6017	39150	45373	23765
1345	15904	24183	16393	27824	5999	39247	45430	23765
...								
7670	18111	24150	17949	30379	5878	38561	47308	23506
7671	18122	24193	17984	30357	5838	38541	47376	23587
7672	18127	24185	17916	30314	5791	38618	47269	23553
7673	18084	24207	17950	30350	5776	38675	47331	23435
7674	18066	24124	17960	30329	5846	38571	47337	23603
7675	18119	24239	17907	30355	5802	38538	47222	23575
7676	18113	24234	17941	30311	5817	38469	47304	23603
7677	18143	24208	17975	30380	5843	38541	47404	23527
7678	18109	24179	17920	30294	5803	38533	47272	23592
7679	18091	24187	17945	30354	5760	38499	47275	23612
7680	18152	24182	17946	30327	5836	38560	47215	23548
...								

Примечание – составлено авторами

Аналогично были рассчитаны средние максимальные значения напряжений (U_{max}) для всех сенсоров «электронного носа» при подаче на него поверочной газовой смеси NO_2 в концентрациях 5 ppm, 10 ppm, 20 ppm, 35 ppm и 50 ppm. Значения U_{max} всех сенсоров «электронного носа» приведены в табл. 6.

Для каждого сенсора были построены графики зависимости его отклика от концентрации поверочной газовой смеси NO_2 , представленные на рис. 5 и 6.

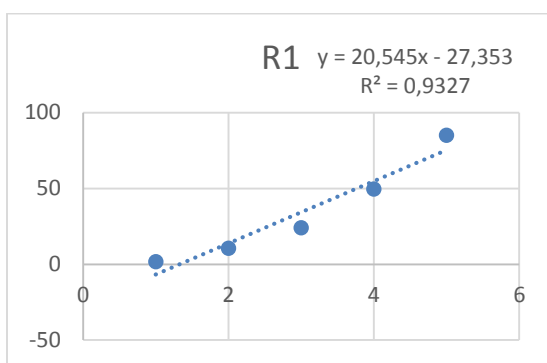
Для аппроксимации зависимости отклика сенсора от концентрации аналита была

использована модель линейной регрессии. На каждом графике представлены уравнения прямых, аппроксимирующих экспериментальные данные, и коэффициенты детерминации (R^2), которые показывают степень соответствия аппроксимирующей модели исходным данным.

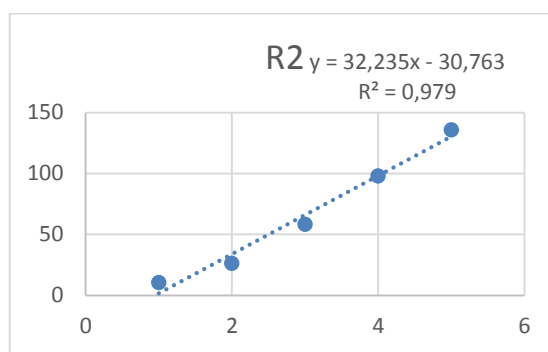
Таблица 6. Значения $U_{\text{тах}}$ для каждого сенсора

Концентрация, ppm	R1	R2	R3	R4	R5	R6	R7	R8
5	16994,52	24152,03	17149,11	29072,76	5926,068	38840,55	46355,09	23638,23
10	15077,14	21136,91	15769,68	26791,77	5815,318	37436,59	44566,77	22692,77
20	13435	16855,41	14690,23	23956,27	5669,045	34036,95	42438,32	21490,77
35	11152,27	13496,73	13178,91	20297,32	5384,727	29254,77	39067,86	20179,05
50	9019,955	11322,23	12113,14	16963,86	5088,727	25569,82	35451,64	19220,27

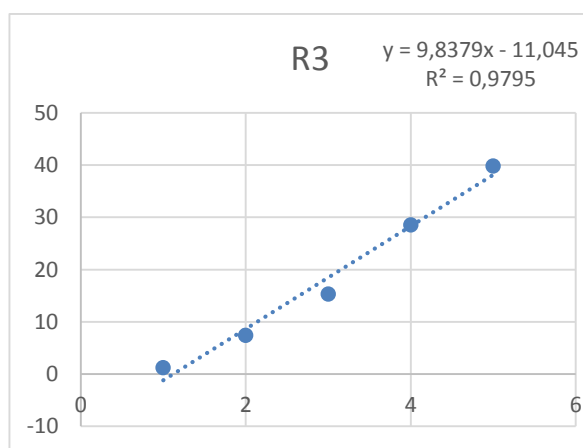
Примечание – составлено авторами



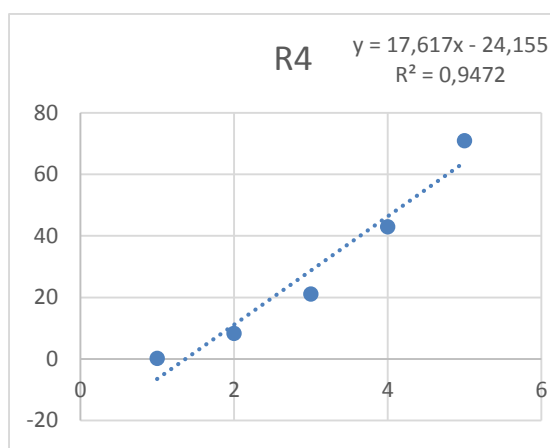
а) калибровочная кривая сенсора R1



б) калибровочная кривая сенсора R2



в) калибровочная кривая сенсора R3



г) калибровочная кривая сенсора R4

Рисунок 5. Графики зависимости откликов сенсоров (R1-R4) «электронного носа» от концентрации поверочной газовой смеси NO₂

Примечание – составлено авторами

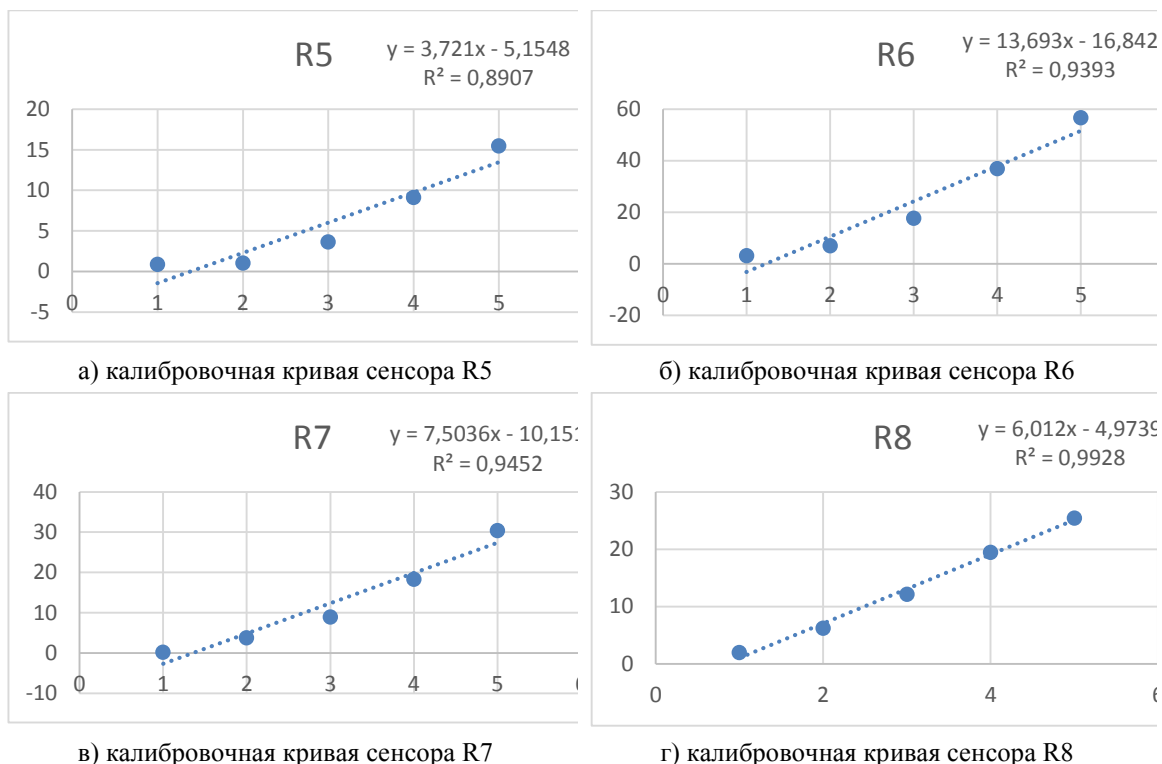


Рисунок 6. Графики зависимости откликов сенсоров (R5-R8) «электронного носа» от концентрации поверочной газовой смеси NO₂

Примечание – составлено авторами

По формуле (1) были вычислены сенсорные отклики для каждой концентрации подаваемой поверочной газовой смеси:

$$R = \left| \left(\frac{U_{air}}{U_{max}} - 1 \right) * 100 \right| \quad (1)$$

Значения сенсорных откликов приведены в табл. 7.

Таблица 7. Значения сенсорных откликов

Концентрация, ppm	R1	R2	R3	R4	R5	R6	R7	R8
5	1,790239	10,6555	1,218546	0,227759	0,872493	3,095757	0,234251	1,988849
10	10,69922	26,44018	7,422182	8,266615	1,015343	6,962075	3,769013	6,238069
20	24,22979	58,55768	15,31569	21,08121	3,621742	17,64553	8,973451	12,18007

35	49,65804	98,01501	28,53975	42,90827	9,093058	36,87666	18,37479	19,47227
50	85,03722	136,0449	39,8493	70,99021	15,43876	56,60242	30,44955	25,43195

Примечание – составлено авторами

Значения коэффициента достоверности аппроксимации R^2 могут лежать в диапазоне от 0 до 1. Чем ближе значение R^2 к 1, тем точнее модель описывает имеющиеся данные. Коэффициенты R^2 для каждой корреляционной кривой сенсоров приведены в табл. 8.

Таблица 8. Коэффициенты R^2

	R1	R2	R3	R4	R5	R6	R7	R8
R^2	0,9327	0,979	0,9795	0,9472	0,8907	0,9393	0,9452	0,9928

Примечание – составлено авторами

Представленный в статье алгоритм калибровки сенсоров «электронного носа» при подаче на него поверочной газовой смеси NO_2 был также применен для калибровки сенсоров при подаче на «электронный нос» других образцов газов.

Далее мы использовали алгоритм кластеризации k-means для анализа многомерных хемирезистивных ответов сенсорной матрицы. Все статистические расчеты и алгоритмы машинного обучения были реализованы и выполнены с использованием языка программирования R. Интегрированная среда разработки RStudio (Stander, Dalla, 2017) использовалась в качестве основной платформы для разработки и выполнения кода.

Результаты анализа показали четкую тенденцию кластеризации в данных, эффективно разделяя сенсорные ответы на основе различных концентраций целевых газов и сухого воздуха. Как показано на рис. 7, алгоритм k-means разделил экспериментальный набор данных, включающий сенсорные отклики на диоксид азота и сухой воздух, на шесть отдельных кластеров, продемонстрировав свою способность к выделению паттернов (цифровых отпечатков газов), позволяющих классифицировать (различать) образцы разной концентрации, учитывая статистические закономерности в сенсорных данных.

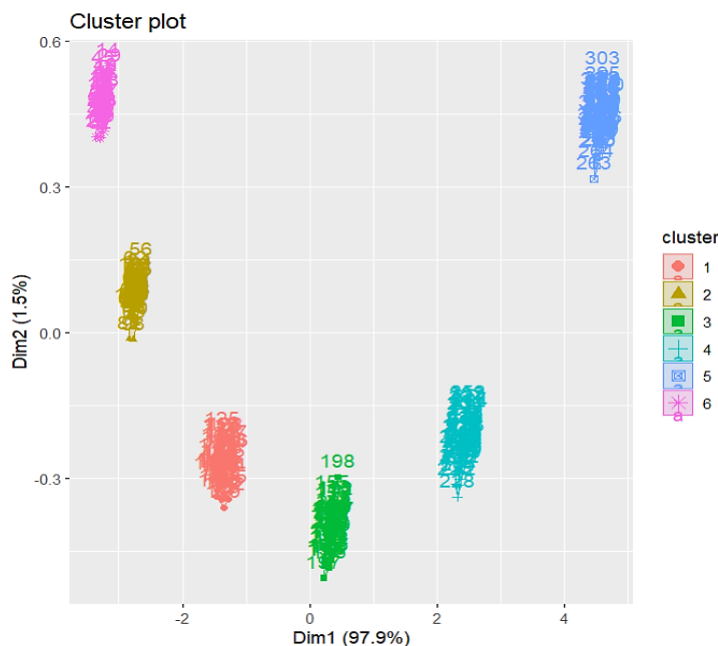


Рисунок 7. Результаты кластеризации для набора данных NO_2 с использованием алгоритма k-means
Примечание – составлено авторами

Каждый кластер 1-5 охватывал 50 векторов сенсорной матрицы, соответствующих введению в многосенсорный газоанализатор газовых смесей калибровки NO_2 в концентрациях 10 ppm, 5 ppm, 20 ppm, 35 ppm и 50 ppm соответственно. Шестой кластер включал 50 измерений, связанных с подачей сухого синтетического воздуха.

Заключение. Проблема загрязнения воздуха является одной из наиболее актуальных экологических проблем современности. В данной работе было продемонстрировано успешное применение металлооксидных хемирезистивных газовых сенсоров в сочетании с методами математической статистики и алгоритмов машинного обучения как важных инструментов для решения этой проблемы.

Разработанная система способна с высокой точностью определять концентрации различных загрязняющих веществ, таких как оксиды азота, аммиак, сероводород, углекислый и угарный газы. Полученные результаты свидетельствуют о перспективности использования данного подхода для создания компактных и экономичных систем мониторинга качества воздуха, которые могут быть широко внедрены как в помещениях, так и на открытых пространствах.

Дальнейшие исследования в данной области будут направлены на улучшение селективности сенсоров, разработку новых алгоритмов обработки данных и интеграцию систем мониторинга в единую сеть устройств Интернета вещей для создания "умных городов".

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Благодарности. Данная работа выполнена при финансовой поддержке Комитета науки Министерства науки и высшего образования Республики Казахстан (грант №AP14872171 – «Разработка моделей и алгоритмов обработки и переноса сенсорных данных при производстве мультизадачных газоаналитических систем»).

Список литературы

- Borowik P., Adamowicz L., Tarakowski R., Siwek K., Grzywacz T. (2020). Odor Detection Using an E-Nose with a Reduced Sensor Array. *Sensors*, 20 (12), 3542; <https://doi.org/10.3390/s20123542>.
- Hurot C., Scaramozzino N., Buhot A., Hou Y. (2020). Bio-Inspired Strategies for Improving the Selectivity and Sensitivity of Artificial Noses: A Review. *Sensors*, 20, 1803.
- Yan J., Guo X.Z., Duan S.K., Jia P.F., Wang L.D., Peng C., Zhang S.L. (2015). Electronic Nose Feature Extraction Methods: A Review. *Sensors*, 15, 27804–27831.
- Konduru T., Rains G.C., Li C. (2015). Detecting sour skin infected onions using a customized gas sensor array. *J. Food Eng*, 160, 19–27.
- Wojnowski W., Majchrzak T., Dymerski T., Gębicki J., Namieśnik J. (2017). Electronic noses: Powerful tools in meat quality assessment. *Meat Science*, Vol. 131, 119-131. <https://doi.org/10.1016/j.meatsci.2017.04.240>.
- Maimunah Mohd Ali, Norhashila Hashim, Samsuzana Abd Aziz, Ola Lasekan. (2020). Principles and recent advances in electronic nose for quality inspection of agricultural and food products *Trends in Food Science & Technology* Volume 99, May, 1-10.
- Adiguzel Y., Kulah H. (2015). Breath sensors for lung cancer diagnosis. *Biosens. Bioelectron.* 65, 121–138.
- Sujono H.A.R., Amin M. (2018). Asthma Identification Using Gas Sensors and Support Vector Machine. *Telkomnika*, 6, 1468–1480.
- Wilson A. (2018). Application of electronic-nose technologies and VOC-biomarkers for the noninvasive early diagnosis of gastrointestinal diseases. *Sensors*, 18, 2613.
- Harrou F., Dairi A., Sun Y., Kadri F. (2018). Detecting Abnormal Ozone Measurements with a Deep Learning-Based Strategy. *IEEE Sens. J.*, 18, 7222–7232.
- Baby R.E., Cabezas M., Walsøe de Reça E.N. (2000). Electronic nose: a useful tool for monitoring environmental contamination *Sensors and Actuators B: Chemical* Volume 69, Issue 3, 25 October, 214-218.
- Norman A., Stam F., Morrissey A., Hirschfelder M., Enderlein D. (2003). Packaging effects of a novel explosion-proof gas sensor. *Sens. Actuators B Chem*, 95, 287–290.
- Kanakam P., Hussain S.M., Chakravarthy A. (2015). Electronic noses: Forestalling fire disasters: A technique to prevent false fire alarms and fatal casualties. In *Proceedings of the IEEE International Conference on Computational Intelligence and Computing Research (ICIC)*, Madurai, India, 10–14, 1–6.
- Сатыбалдина Д., Исайнова А., Альсеитов О., Байдельдинов М., Сеилов Ш. (2023). Создание газоаналитической мультисенсорной системы типа «электронный нос» и исследование ее производительности для задач экологического мониторинга. *Вестник КазАТК – №125(2)*, 315–323. // Satibaldina D., Isainova A., Alseitov O., Baideldinov M, Seilov Sh. (2023). Sozdanie gazoanaliticheskoi multisensornoj sistemi tipa «elektronnii nos» i issledovanie ee proizvoditelnosti dlya zadach ekologicheskogo monitoringa. *Vestnik KazATK, – №125(2)*, 315–323.
- Мукатов А.Т., Исайнова А.Н., Глазырина Н.С., Сатыбалдина Д.Ж. (2024). Программное обеспечение подготовки обучающих наборов данных (датасетов) для распознавания основных загрязнителей атмосферного воздуха. Свидетельство о внесении сведений в государственный реестр прав, охраняемых авторским правом №45747 от 14 мая. // Mukatov A.T., Isainova A.N., Glazirina N.S, Satibaldina D.J. (2024). Programmnoe obespechenie podgotovki obuchayuschih naborov dannih (datasetov) dlya raspoznavaniya osnovnih zagryaznitelei atmosfernogo vozduha. Svidetelstvo o vnesenii svedenii v gosudarstvennii reestr prav, ohranyaemie avtorskim pravom №45747 ot 14 maya.
- Сатыбалдина Д.Ж., Сыдыков Д.М., Исайнова А.Н. (2023). Нейросетевой классификатор газов для газоаналитической мультисенсорной системы. Свидетельство о внесении сведений в государственный реестр прав, охраняемых авторским правом №37464 от 23 июня. // Satibaldina D.J., Sidikov D.M., Isainova A.N. (2023). Neirosetevoi klassifikator gazov dlya gazoanaliticheskoi multisensornoj sistemi. Svidetelstvo o vnesenii svedenii v gosudarstvennii reestr prav, ohranyaemie avtorskim pravom №37464 ot 23 iyunya.
- Alseitov O., Issainova A., Satyaldina D., Konyrkhanova A. (2023). Detection and Classification of the Volatile Organic Compounds from the Petroleum and Petroleum-Derived Products Using an Electronic Nose and Machine Learning Algorithms // *Proceedings of International Conference on Electrical Facilities and informational technologies 2023 (ICEIC 2023)*, Almaty, Kazakhstan, August 22nd to 25th, 218-223.
- Satyaldina D., Baydeldinov M., Issainova A., Alseitov O., Konyrkhanova A., Akhmetova Z., Seilov S. (2023). Artificial Olfactory System for Distinguishing Oil-Contaminated Soils // *WSEAS Transactions on Environment and Development*, 19, 951–960.

- Сатыбалдина Д., Исайнова А., Глазырина Н., Шагабудинов А. (2024). Статистические методы извлечения признаков для обучения мультисенсорной газоаналитической системы. Вестник КазАТК, №131(2), 250–259. // Satibaldina D., Isainova A., Glazirina N., Shagabudinov A. (2024). Statisticheskie metodi izvlecheniya priznakov dlya obucheniya multisensornoj gazoanaliticheskoi sistemi. Vestnik KazATK, №13 (2), 250–259.
- Мобильный газоанализатор с функцией беспроводной передачи данных для системы мониторинга качества воздуха, Патент РК №9539 на полезную модель, 28.06.2024, по заявке №2024/0827.2 // Mobilnii gazoanalizator s funkciei besprovodnoi peredachi danih dlya sistemi monitoringa kachestva vozduha. Patent RK №9539 na poleznuyu model, 28.06.2024 po zayavke №2024/0827.2
- Stander J., Dalla L.V. (2017). On Enthusing Students About Big Data and Social Media Visualization and Analysis Using R, RStudio, and RMarkdown, Journal Statistics Education, VOLUME 25, 60–67.

Information about authors

Khasenova Zarina Toleubekovna – PhD, dean, D. Serikbayev East Kazakhstan technical university, Ust-Kamenogorsk, Kazakhstan, E-mail: zkhasenova@edu.ektu.kz, ORCID: 0009-0001-1417-7793, +77715492087

Rakysheva Madina Adilkhanovna – master of Technical Sciences, D. Serikbayev East Kazakhstan Technical University, Ust-Kamenogorsk, Kazakhstan, E-mail: mrakysheva@edu.ektu.kz, ORCID: 0000-0002-5050-7390, +77029118914

Issainova Aliya Nasipovna – master of Technical Sciences, Senior Lecturer of the Department of Information Security, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan, E-mail: issainova_an@enu.kz, ORCID: 0000-0001-8008-4096, +77017361017

Glazyrina Natalya Sergeevna – PhD, Acting Associate Professor of the Department of Computer and Software Engineering, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan, E-mail: glazyrina_ns_1@enu.kz, ORCID: 0000-0002-0259-774X, +77015263168

Mukatov Abzal Tynyshtybayevich – student of the Department of Computer and Software Engineering, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan, E-mail: abzal_mukatov@mail.ru, +77777414761