

АННОТАЦИЯ

**диссертационной работы Турарбек Әсем Турарбекқызы на тему
«Исследование и разработка модели определения сейсмической активности
Казахстана на основе методов машинного обучения», представленной
на соискание степени доктора философии (PhD)
по специальности 6D070300 – Информационные системы**

Актуальность работы. Во второй половине XX века выросло количество и сила природных катаклизмов. Только за последние десятилетия во всем мире произошло значительное количество сильных разрушительных землетрясений (в Чили, Перу, Мексике, США, Японии, Армении, на Сахалине, в Афганистане, Турции и др.), которые вызвали огромный ущерб и немалые человеческие жертвы. Землетрясения занимают одно из первых мест среди других видов природных катастроф. Землетрясения приводят к массовым жертвам и разрушениям. Землетрясения могут наносить существенный экономический ущерб странам и регионам, особенно тем, что экономически зависимы от конкретных географических областей. Современный мир тесно связан, и сейсмическое событие в одной стране может оказать влияние на международные цепочки поставок и экономику других стран.

Стихийные явления, и землетрясения в том числе, неизбежны. Их нельзя предотвратить, но уменьшить их разрушительное влияние можно и необходимо. Для этого нужно знать причины возникновения землетрясений, проводить оценку сейсмической опасности, изучать процессы, связанные с их подготовкой и возникновением, разрабатывать методы прогноза этих явлений, выявлять места возможных источников сейсмических колебаний. Землетрясения составляют 13% от общего числа природных катастроф. Сейсмические явления, такие как землетрясения, являются результатом сложных геологических и геофизических процессов, которые постоянно происходят внутри земной коры. Изучение сейсмической активности и ее последствий имеет большое значение для обеспечения безопасности труда и занятости. В Казахстане, как и в других странах Центральной Азии, сейсмические явления представляют собой серьезную угрозу, и их понимание играет решающую роль в подготовке и снижении риска. Ряд регионов Казахстана подвержен сейсмическим рискам из-за их геологической структуры и расположения. В Казахстане на 2023 год в наиболее сейсмоопасных регионах проживает более 7 млн человек, сосредоточено свыше 50% промышленного потенциала, расположено более 400 городов и населенных пунктов.

Сейсмическая активность, регулярно наблюдаемая на территории Казахстана, обуславливает серьезную потребность в разработке эффективных методов прогнозирования сейсмических явлений.

Сейсмическая активность может представлять угрозу для населенных пунктов и ключевых отраслей экономики, таких как добыча полезных ископаемых, промышленность и туризм. В истории страны зафиксированы

случаи разрушительных землетрясений, повлекших за собой значительные человеческие и материальные убытки. Прогнозирование сейсмической активности путем применения методов машинного обучения может стать мощным инструментом предупреждения и минимизации последствий природных катаклизмов. Прогнозирование сейсмической активности может помочь в разработке строительных кодексов и норм, чтобы обеспечить безопасность зданий и инфраструктуры. Изучение сейсмической активности также может способствовать развитию науки и технологий в стране, а также подготовке специалистов в области геофизики и сейсмологии.

В настоящее время существует целый ряд систем и мониторов. Среди них особо выделяются USGS, EMSC, IRIS, они выводят информацию о сейсмической активности всего мира, а также National Data Center of Kazakhstan (NDC), Institute of Seismology in Kazakhstan, Darmen предоставляют информацию по сейсмическим событиям Казахстана.

Эти системы имеют схожую функцию, они визуализируют эпицентры землетрясений, показывают полное описание, то есть дату, время, магнитуду, глубину и указывают на карте прошедшее событие. В моделях машинного обучения для определения сейсмической активности была использована собранная база сейсмологических данных в размере 89629 сейсмических событий из данных Института сейсмологии и данных USGS. Данные прошли этапы предобработки. На этапе классификации использовался ряд наиболее популярных алгоритмов машинного обучения (Support vector machine – SVM, Logistic regression – LR, Decision tree – DT, Random forest – RF, k-nearest neighbors – k-NN, and XGBoost) и нейронных сетей (Deep neural networks – DNN, Convolutional neural networks – CNN). Результаты классификации представлены в виде сводных таблиц метрик оценки эффективности алгоритмов: правильности (accuracy), точности (precision), полноты (recall) и F-меры (F1-score), графиков кривых (Area under curve – Receiver operating characteristics – AUC-ROC). Для анализа сейсмической активности разработана модель с применением исторических данных сейсмических событий Казахстана.

Машинное обучение быстро развивалось в последнее десятилетие, обещая значительно изменить и улучшить функции анализа больших данных в самых разных областях. По сравнению с традиционными методами машинное обучение обеспечивает значительные преимущества в решении сложных задач, производительности вычислений, распространении неопределенности и обработке ее, а также поддержке принятия решений. Все используемые методы были обучены и протестированы на реальных данных, собранных в Казахстане за последние 98 лет, с 1906 по 2023 год. Предлагаемая модель превзошла другие модели по показателям precision, recall, accuracy, F-score и AUC-ROC на 63%, 82,4%, 87,4%, 62,7% и 83% соответственно. Основываясь на результатах, можно сделать вывод, что предложенная модель Conv2D эффективна для определения сейсмической активности.

Цель диссертационной работы – создание модели и информационной системы определения сейсмической активности Республики Казахстан на основе сейсмологических данных с использованием методов машинного обучения.

Задачи исследования:

1. Исследование существующих информационных технологий и систем определения сейсмической активности и обработки данных;
2. Исследование и анализ алгоритмов, методов машинного обучения, используемых для обнаружения сейсмической активности;
3. Создание алгоритма определения сейсмической активности территории Казахстана;
4. Создание и обучение модели глубокого обучения обнаружения сейсмической активности на основе признаков (feature);
5. Разработка информационной системы обнаружения сейсмической активности.

Объект исследования: Сейсмическая активности территории Казахстана.

Предмет исследования: Алгоритмы и методы машинного обучения определения сейсмической активности.

Методы исследования: Методы машинного обучения, нейронные сети, глубокие нейронные сети, статистические методы, алгоритмы машинного обучения.

Научная новизна проведенных исследований и полученных результатов:

1. Разработан алгоритм метода определения сейсмической активности территории Казахстана;
2. Разработана модель глубокого обучения для определения сейсмической активности на основе признаков;
3. Разработана информационная система определения сейсмической активности на основе созданной модели глубокого обучения.

Теоретическая и практическая значимость исследования: Исследование позволяет расширить понимание о возможностях и ограничениях применения алгоритмов машинного обучения в определении сейсмической активности. Разработанная модель и методология могут быть интегрированы в существующие системы мониторинга и прогнозирования сейсмической активности, повышая их точность и своевременность.

Объем и структура работы. Диссертационная работа состоит из введения, 4 глав и заключения. Общий объем диссертации составляет 102 страницы, 36 рисунков, 15 таблиц. Список литературы состоит из 79 наименований.

Во введении дано описание актуальности, цели, задач, объектов и методов исследования, теоретической и практической значимости, а также новизны диссертационной работы.

Первый раздел описывает сейсмический режим Казахстана, подробно рассматривает методы анализа и прогноза землетрясений. Представлен

обзор и анализ систем контроля и прогнозирования землетрясений. Проведен эксперимент определения ущерба землетрясений на ГИС ITRIS.

Во втором разделе приведено описание основных типов моделей машинного обучения и нейронных сетей. Представлен анализ и обзор результатов других авторов использования машинного обучения и нейронных сетей в сейсмологии и определении сейсмической активности.

В третьем разделе подробно представлена разработка модели на основе методов машинного обучения и нейронных сетей, что является важнейшей частью диссертационной работы. Представлен алгоритм разработки модели, определено целевое событие и признаки, показан этап сбора и предобработки сейсмологических данных. Выполнена оценка эффективности использования предложенной модели для определения сейсмической активности территории Казахстана.

В четвертом разделе представлена система, разработанная на фреймворке Django Python. Система выполняет следующие основные функции: создание основных категорий сейсмических событий, извлечение данных по регионам и годам, визуальное представление полученных результатов в виде эпицентров на карте Казахстана и сводных таблиц.

В заключении обобщены теоретические и практические результаты данной диссертационной работы, приведены ее наиболее значимые аспекты в анализе сейсмической активности с применением моделей машинного обучения, нейронных сетей и показателей сейсмических событий.

Основное положение, выносимое на защиту. Новая модель и информационная система определения сейсмической активности, основанная на машинном обучении. Результаты экспериментов определения сейсмической активности, подтверждающих эффективность предложенной модели.

Личный вклад исследователя. Диссертантом был выполнен анализ существующих систем и приложений мониторинга землетрясений; разработан алгоритм метода определения сейсмической активности территории Казахстана; разработана модель глубокого обучения для определения сейсмической активности на основе признаков; разработана информационная система определения сейсмической активности на основе созданной модели глубокого обучения.

Степень обоснованности и достоверности научных результатов. Результаты диссертации были представлены в 13 научных работах, из них 2 статьи опубликованы в журнале рецензируемые в базе Scopus, 3 статьи – в журналах, рекомендуемых Комитетом по обеспечению качества в сфере образования и науки Министерства образования и науки Республики Казахстан, и 8 статей – в материалах международных конференций.

Статьи, опубликованные в издании, представленном комитетом по обеспечению качества в области науки и высшего образования Министерства науки и высшего образования Республики Казахстан:

- 1) Турарбек А.Т. Геоинформационные системы в сейсмологии // Вестник КазННТУ, Серия «Наука о Земле». –2016. –№5. –С.51-53.

- 2) Турарбек А.Т. Использование геоинформационной системы ITRIS для оценки последствий землетрясений Казахстана // Вестник КазНУ, Серия «Технические науки». – 2017. – №3. – С.394-399.
- 3) Турарбек А.Т., Садыкова А.Б. Современное состояние анализа и прогноза землетрясений в Казахстане // Вестник КазНУ, Серия «Технические науки». – 2018. – №2. – С.184-191.

Научные статьи в журналах, индексируемых в базе Скопус:

1. Turarbek Assem, Adetbekov Yeldos, Bektemesov Maktagali 2-D Deep Convolutional Neural Network for Predicting the Intensity of Seismic Events // International Journal of Advanced Computer Science and Applications vol.14(1), pp.788-796, 2023 (IF=0.675, CiteScore rank =44, Q3 Scopus).
2. Turarbek Assem, Bektemesov Maktagali et al. Deep Convolutional Neural Network for Accurate Prediction of Seismic Events // International Journal of Advanced Computer Science and Applications vol.14(10), pp.604-613, 2023 (Q3, Scopus indexed, процентиль 47)

Статьи в материалах международной конференции:

- 1) Криворотько О.И., Кабанихин С.И., Бектемесов М.А., Садыкова А.Б., Турарбек А.Т. Геоинформационная система Казахстана. Математические модели геоинформационной системы // Труды международной научной конференции «Марчуковские научные чтения 2017», Россия, Новосибирск, 2017. – С.455-462.
- 2) Турарбек А.Т. Применение геоинформационной системы ITRIS для моделирования и прогноза сейсмического режима Казахстана // Труды международной научной конференции «Марчуковские научные чтения 2017», Россия, Новосибирск, 2017. – С.98-99.
- 3) Турарбек А.Т. О методах качества изображений при дистанционном мониторинге землетрясений // Материалы Международной научно-методической конференции «Математика в Казахстане - прошлое и перспективы», посвященной 100-летию Ибрашева Хасана Ибрашевича Алматы, 2016 г. –С.115-117.
- 4) Турарбек А.Т. Применения данных дистанционного зондирования при работе с геоинформационными системами для прогнозирования землетрясений // Материалы международной научной конференции студентов и молодых ученых «Мир Фараби», Алматы, 2016 г. –С.237.
- 5) Кенес А., Турарбек А.Т. Составные части геоинформационной системы землетрясений // Материалы международной научной конференции студентов и молодых ученых «Мир Фараби», Казахстан, Алматы, 2017. –С.193.
- 6) Турарбек А.Т., Адетбеков Е.Н. Қазақстанның жер сілкінісі мониторингін жүргізу үшін Python тілін қолдану // Материалы VIII международной научно-методической конференции «Математическое моделирование и информационные технологии в образовании и науке», посвященная 90-летию юбилею КазНПУ имени Абая. Алматы, 2018. – С.109-112.

- 7) Турарбек А.Т., Адетбеков Е.Н Машинное обучение в прогнозировании землетрясений //Материалы международной научной конференции «Обратные и некорректные задачи в естествознании», Алматы,2023 г. – С.44-45.
- 8) Турарбек А.Т. Глубокое машинное обучение в определении сейсмической активности//Материалы международной научной конференции «Обратные и некорректные задачи в естествознании», Алматы,2024 г. –С.56-57.