

**0.1. Черных В.Ю., Степнов А.А., Сорокин А.А.
Автоматизация рутинной обработки
сейсмических данных методами машинного обучения**

Основной вид инструментальных данных в сейсмологии – это непрерывные записи движения грунта (смещения, скорости, ускорения), регистрируемые сейсмическими станциями, как правило, объединёнными в сети наблюдений на определенной территории. Получаемая таким образом информация является основой для детектирования и комплексного исследования землетрясений. С активным развитием сетей наблюдений и возможностей приборной базы происходит кратный рост объема данных, что создает большие сложности в обработке информации, основанной, как правило, на ручном труде специалистов. В связи с этим, существует актуальная задача, связанная с разработкой современных масштабируемых методов и технологий обработки инструментальных данных, способных оперативно производить анализ сейсмологических данных по различным направлениям. Наиболее перспективным и востребованным подходом в решении рассматриваемых задач являются методы машинного обучения, обеспечивающие баланс между точностью и скоростью работы алгоритмов.

В рамках исследования был разработан метод машинного обучения, классифицирующий небольшие фрагменты данных на участки, содержащие вступления сейсмических волн, и на прочие (шум). Созданная нейросеть Seismo-Performer [1] анализирует спектрограммы сейсмограмм и не имеет в своем составе сверточных слоев. Нейросеть состоит из небольшого, по сравнению с аналогами, количества обучаемых параметров – 57 тыс. (популярный аналог – GPD – включает в себя 1,742 млн. параметров). Небольшое число параметров позволяет достичь высокой производительности: Сканирование 24-часовой трехканальной сейсмограммы созданная нейросеть выполняет в среднем на 47,3% быстрее самого быстрого аналога и на 119,9% быстрее самого медленного (56,2 с. $\pm 0,382, 8. \pm 0,3123, 6. \pm 0,4$).

Seismo-Performer демонстрирует лучшую способность к обобщению. Оценка была проведена с использованием массивов сейсмограмм из двух разных источников: Сахалин, Дагестан и Северная Калифорния, в ходе которой Seismo-Performer продемонстрировал стабильное преимущество при распознавании сейсмических волн на данных из источника, отличного от того, на котором была обучена нейросеть. Среднее преимущество в точности оценки (над GPD) в таких условиях составило 4,58%.

Проведенные обучение и оценка эффективности нейросетей требовали сбора большого массива сейсмических данных. Требования к подготовке и составу данных зависят от особенностей их дальнейшего применения и могут определять методы пред-

варительной обработки, типы разрешенных сейсмических инструментов и т.д. Для автоматизации процесса сбора таких массивов был разработан настраиваемый алгоритм, развертываемый на содержащих архивы сейсмических наблюдений машинах, и автоматизирующий решение задачи выборки данных для обучения и оценки методов машинного обучения. Полученный комплекс программ затем использовался для сбора массива данных с о. Сахалин (9,8 тыс. сейсмограмм) и с р. Дагестан (28,1 тыс. сейсмограмм). В ходе выработки алгоритма также был проведен анализ различных методов предварительной обработки сейсмических данных и их влияние на эффективность применения нейросетей.

Использование методов машинного обучения для пособия рутинной ручной обработке сейсмических данных предполагает разработку автоматизирующего алгоритма. Алгоритм, выработанный для решения данной задачи, исследует поступающие в реальном времени (или собранные архивы) данных выбранным методом машинного обучения, используя процедуру скользящего окна. Разработанное решение уже находится в ежедневной эксплуатации на сети сейсмических наблюдений о. Сахалин.

Дальнейшим направлением развития является интеграция методов машинного обучения в сейсмический блок автоматизированной информационной системы «Сигнал».

Список литературы

- [1] STEPNOV A., CHERNYKH V., KONOVALOV A. The Seismo-Performer: generally based and efficient machine learning approach for recognition seismic phases from local earthquakes in real time // Sensors, 2021, 21(18), 6290.