Секция «Цифровые технологии в нефтегазовой отрасли»

**Сегментация пачек баженовской высокоуглеродистой формации
Среднеобской нефтегазоносной области Западно-Сибирского
нефтегазоносного бассейна**

Научный руководитель – Большакова Мария Александровна

Богданов Александр Александрович

Студент (магистр)

Московский государственный университет имени М.В.Ломоносова, Геологический факультет, Кафедра геологии и геохимии горючих ископаемых, Москва, Россия

E-mail: a.bogdanov2512@mail.ru

Расчленение разреза баженовской высокоуглеродистой формации (БВУФ) на пачки по показаниям кривых геофизических исследований скважин (ГИС) в настоящее время выполняется вручную и требует значительных временных ресурсов как на проведение самой интерпретации, так и на дополнительные исследования.

Цель работы – автоматизировать выделение пачек в разрезе БВУФ по данным ГИС с применением методов глубокого обучения. Для решения задачи использовался подход *sequence labeling* (маркирование последовательностей) – на выходе модель проставляет метку определенной пачки каждому интервалу записи методов ГИС.

Для обучения и инференса модели использовались данные по 946 скважинам: 757 в обучающей выборке (108855 интервалов - строк) и 189 (27069 интервалов) в тестовой из Салымской, Сургутской, Самотлорской, Нижневартовской и Малобалыкской структурно-фациальных зон. В качестве признаков использовались: НК (нейтронный каротаж), DGR (нормированный гамма-каротаж), ВК (боковой каротаж), структурно-фациальная зона (метка), нормированная на скважину глубина, дифференцированные один и два раза кривые ГИС, бинарные лейблы для пропущенных в скважине методов ГИС. По показаниям ГИС исследователи вручную выделяют до 4-х пачек в строении БВУФ – LB, VI, V, VI. В работе целевой переменной выступали метки пачек по каждому интервалу глубины, IV и V пачки БВУФ были объединены в одну в связи со схожей конфигурацией кривых и сложностью их разделения. Итого использовались три уникальные метки: 0 – для VI пачки, 1 – для IV и V пачек, 2 – для LB пачек (I-III).

Архитектура нейронной сети, используемая для решения данной задачи, состояла из: а) четырех одномерных сверточных слоев (размер фильтра = 3) без понижения размерности входной последовательности; б) далее выход из сверточного блока подавался на вход двунаправленной рекуррентной сети (BiLSTM) с 3 слоями; в) результаты из прошлого пункта шли на вход в линейный слой, логиты откуда далее переходили в CRF слой, осуществляющий сегментацию последовательности (пачек в скважине). В процессе обучения модели был использован оптимизатор AdamW с параметрами $lr = 0.001$ и $weight_decay = 0.001$. Оптимизировалась лосс-функция - *negative log-likelihood*. Для оценки обобщающей способности модели использовалась метрика *f1-score* взвешенная.

В результате по всем 189 скважинам из тестовой выборки была получена метрика *f1-score = 0.874*. Модель хорошо сегментирует пачки по Салымской, Сургутской, Нижневартовской и Малобалыкской структурно-фациальных зонам (по всем зонам *f1* метрика > 0.85) и заметно ошибается на Самотлорской зоне в связи со специфической конфигурацией кривых ГИС, сокращенной мощностью БВУФ и выпадением/сокращением из разреза скважин нижних пачек ($f1 = 0.63$).

Данный подход к выделению пачек в разрезе БВУФ показал высокую точность и впоследствии позволит автоматически сегментировать пачки, затрачивая значительно меньше временных ресурсов исследователей.