

ПРОГНОЗИРОВАНИЕ ДЕБИТА НЕФТИ И ОБВОДНЕННОСТИ СКВАЖИННОЙ ПРОДУКЦИИ НА ОСНОВЕ ИСТОРИЧЕСКИХ ДАННЫХ О ЗАВОДНЕНИИ С ПРИМЕНЕНИЕМ ТЕХНОЛОГИЙ ГЛУБОКОГО ОБУЧЕНИЯ

Сальников М.А.¹, Крамойкин И.А.²

¹ НИ ТПУ, ИШИТР, А1-36, e-mail: mas55@tpu.ru

² НИ ТПУ, ИШИТР, 8ПМ1И, e-mail: iak68@tpu.ru

Введение

В условиях современной нефтедобывающей отрасли, прогнозирование дебита нефти и обводненности является критически важной задачей для эффективной эксплуатации нефтяных месторождений. Ранее использовавшиеся традиционные методы прогнозирования, основанные на физических моделях, уже не всегда удовлетворяют потребностям рынка, так как они часто не способны давать качественный прогноз при высокой степени неопределенности, которая является ключевой особенностью данной предметной области. В последние годы технологии глубокого обучения (Deep Learning) получили широкое распространение в нефтегазовой отрасли и стали использоваться для анализа больших объемов данных и прогнозирования параметров добычи.

Прогнозирование дебита нефти и воды с применением технологий глубокого обучения

Задачей данной сети выступает предсказание дебита нефти и воды на 120 дней вперед опираясь на данные работы системы ППД. Осложняющим фактором является учет разномасштабных данных в предсказании. Для этого была разработана следующая архитектура (рисунок 1), включающая две ветви. Каждая ветвь, по задумке, эмбежит входящие в неё данные в вектор. Эти вектора конкатенируются и далее два линейных слоя выполняют непосредственно предсказание [1].

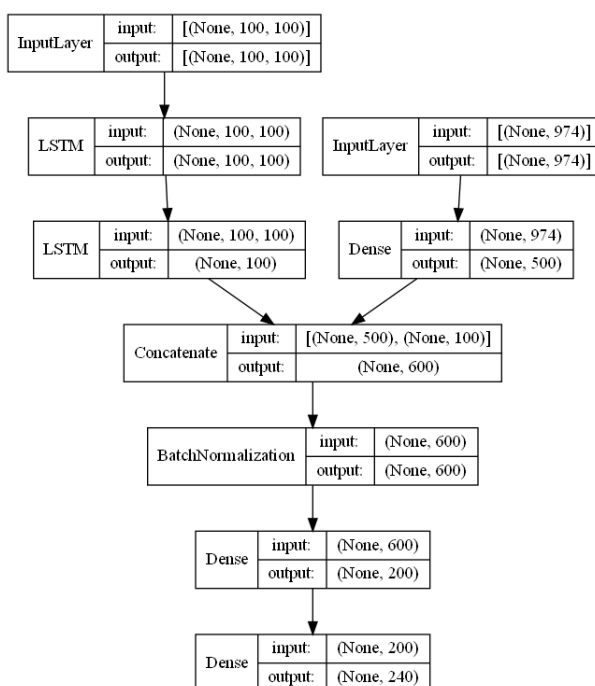


Рис. 1. Архитектура нейронной сети

Первым типом данных выступают временные ряды (двумерный массив). К ним относятся история работы скважины, то есть дебит нефти, дебит воды и забойные давления. Так же сюда добавляются метки ГТМ (геолого-технические мероприятия) подвергшиеся процедуре one-hot-encoding. Последним типом данных, зависящих от времени, являются истории нагнетания жидкости в нагнетательных скважинах всего месторождения. Все эти данные можно считать динамическими. Подготовка этих данных для обучения производилась методом скользящего окна. То есть предсказание должно формиро-

ваться на основе истории работы скважин за определенный период (в нашем случае 100 дней). Лидерами в работе с последовательностями в глубоком обучении на текущий момент являются рекуррентные нейронные сети (RNN), а в работе непосредственно с временными рядами хорошо себя показывают LSTM. Поэтому ветка, отвечающая за обработку динамических фичей включает в себя два слоя таких нейронов.

Вторым типом данных выступают статические фичи. Первый тип статических фичей это вектор расстояний от добывающей скважины, для которой делается предсказание до всех нагнетательных. Таким образом мы пытаемся передать модели пространственное представление о скважинах. Второй статической фичей выступает длина пересечения пласта и скважины, с расчетом на то, что дебит прямо пропорционален этому значению. Для учета геологической составляющей в модель подается набор коэффициентов, описывающих ГИСы. Коэффициенты представляют собой веса аппроксимирующей базисной функции кубического сплайна для каждой кривой. Все три составляющие объединяются в один вектор и подаются во вторую ветку нейронной сети. Статические фичи обрабатываются обычным линейным слоем [2].

Эвристически было доказано, что важной частью такой архитектуры является слой BatchNormalization после объединения выходов из веток сети. Без него качество обученной модели оказывалось в разы хуже. Объясняется это разной природой статических и динамических данных. Слой BatchNormalization приводит эмбединги этих данных к одному распределению [3].

Так же стоит упомянуть небольшую хитрость, примененную при обучении модели, называемую методом взятия первой разности. Модель предсказывает не вектор дебитов напрямую, а вектор приращений этих дебитов относительно последнего известного модели дня работы скважины. Таким образом удастся существенно улучшить качество обученной модели. Объясняется это тем, что такой трюк снижает дисперсию предсказываемых моделью значений, что положительно влияет на её способность к генерализации.

В качестве метрики качества использовалась SMAPE (Symmetric Mean Absolute Percentage Error), одна из метрик, используемых для оценки точности прогнозирования. Она позволяет оценить относительную ошибку прогноза и измеряется в процентах.

Существует несколько преимуществ использования SMAPE. Во-первых, она является симметричной метрикой, что означает, что она не зависит от порядка сравнения фактических значений и прогнозов. Во-вторых, SMAPE позволяет измерять точность прогнозов в процентах, что удобно при сравнении точности различных прогнозных моделей. Однако, как и любая метрика, SMAPE имеет свои недостатки. Например, она не учитывает различия в значимости отдельных ошибок, что может привести к несправедливым оценкам прогнозов, особенно в случае больших выбросов. Также, SMAPE может давать бесконечные значения, если фактический или прогнозируемый ноль, что требует дополнительной обработки.

SMAPE итоговой модели на тестовой выборке составил 9,64.

Заключение

Полученные результаты могут быть использованы для оптимизации процессов добычи нефти и газа, улучшения управления скважинами и повышения эффективности работы нефтегазовых компаний, а самое главное, для оптимизации процесса заводнения пласта. Таким образом, прогнозирование дебита нефти и обводненности скважинной продукции с использованием технологий глубокого обучения представляет собой перспективное направление в области нефтегазовой промышленности, которое может привести к значительному улучшению эффективности и рентабельности производства.

Список использованных источников

1. Al-Qutami T.A., Ibrahim R., Ismail I., Ishak M.A. Virtual multiphase flow metering using diverse neural network ensemble and adaptive simulated annealing // *Expert Systems with Applications*. – 2018. 93 (2018), pp. 72–85.
2. Amin, Evaluation of commercially available virtual flow meters (VFM), Offshore Technology Conference, 2015, P. 1–26. 25764-MS.
3. F. Chollet et al., Keras. <https://github.com/fchollet/keras>, 2015.