Development of the Approach for Automatic Well Logging Interpretation for Big Number of Wells with the Use of Machine Learning

M.F. Validov* (Kazan federal university), A.R. Ismagilov (Kazan federal university), D.S. Voloskov (Kazan federal university), M.S. Magdeev (PJSC "Tatneft"), A.A. Nazarov (PJSC "Tatneft")

SUMMARY

The increased interest for geological and reservoir simulation model construction for the old fields raises the issue of reinterpretation of well logging data of the old well stock, taking into account the concepts of geology, specified during the period of field production. This work shows the results of mathematical approach development for automatic interpretation of well logging data of the old well stock. The work is aimed to solve the problem of fast reinterpretation of a standard logging set using machine learning algorithms. The solutions obtained for the determination of stratigraphic boundaries with the use of logistic regression and the detection of lithotypes basing on the support vector machine are presented. Mathematical algorithms and approaches to use them are presented and described.

Разработка подхода к автоматической интерпретации ГИС в терригенном разрезе для большого массива скважин с применением машинного обучения

М.Ф. Валидов* (КФУ), А.Р. Исмагилов (КФУ), Д.С. Волосков (КФУ), М.Ш. Магдеев (ПАО «Татнефть»), А.А. Назаров (ПАО «Татнефть»)

Введение

Возросшая актуальность построения геолого-гидродинамических моделей на месторождениях последней стадии эксплуатации остро ставит вопрос о переинтерпретации ГИС старого фонда скважин, с учетом представлений о геологии, скорректированных за время эксплуатации. Ввиду значительного объема данных разрабатывается техническое решение, позволяющее сократить время затраты на переинтерпретацию ГИС. В данной работе представлены результаты создания алгоритма автоматической интерпретации стандартного комплекса ГИС (литология, коэффициент пористости, глинистости и нефтенасыщенности, тип насыщения) в терригенных отложениях тульских и бобриковских горизонтов Татарстана при помощи машинного обучения.

Основной целью данной работы является разработка подходов и математических алгоритмов автоматизации процессов интерпретации материалов ГИС.

Метод

В качестве исходных данных для проведения анализа были использованы результаты исследования кернового материала по ряду площадей изучаемого месторождения, но по причине малой изученности парных керновых значений тульско-бобриковских отложений в качестве эталонной выборки были использованы скважины, проинтерпретированные согласно стандарту интерпретации, ГИС компании заказчика [1].

Первоочередной задачей было автоматическое определение зоны интереса – стратиграфических отбивок при помощи алгоритма логистической регрессии, обученного на результатах ручной корреляции пластов.

Разработанный алгоритм позволяет автоматически выделять стратиграфические отбивки для кровли тульско-алексинских отложений, подошвы тульского репера, кровли и подошвы бобриковского горизонта и подошвы черепетского горизонта турнейского яруса.

В результате многочисленных вариантов опробования нейросетевых алгоритмов было решено применить метод логистической регрессии.

Преимуществом данного подхода, применительно к стратиграфическим отбивкам, является успешное применение граничного значения вероятности верного ответа для отклонения выдаваемого ответа, в случае его низкой правдоподобности. Соответственно, при данном методе получена наилучшая статистика верных ответов при ~10% отклонений вероятно ложных результатов.

Стратиграфические отбивки выделяются по кривым ГК, НГК, КС (PZ), для этого проводится обучение на репрезентативном наборе скважин (20-25 скважин). Обучающие данные должны содержать:

- Отбивку подошвы тульского репера;
- Отбивку кровли турнейских отложений.

Для определения стратиграфических границ на основе кривых ГК и НГК проводится нормализация кривых в интересующем интервале (бобриковские отложения).

Первоначально, по пересечениям нормализованных кривых автоматически выделяются потенциальные стратиграфические границы, после этого при помощи логистической регрессии среди них выделяются предварительные «истинные» границы, представляющие собой кровлю бобриковских отложений и кровлю отложений турнейского яруса.

Далее предварительная «истинная» граница корректируется для более точного позиционирования с учетом кривой КС. «Истинные» границы кровли и подошвы бобриковских отложений находятся схожим образом.

Автоматизация процесса выделения стратиграфических отбивок по предварительным оценкам составляет 90%. На наборе из 3000 скважин затрачиваемое время обработки машиной составляет около 15 минут. Как правило, около 10% скважин даже после проверки качества имеют те или иные отклонения, в результате чего такие скважины необходимо обработать вручную.

Литологическое расчленение разреза является одной из важных задач при интерпретации скважин. Для качественного выделения литологических типов пород вручную необходимы высоко квалифицированные специалисты. В связи с этим, решение данной задачи в дальнейшем позволит значительно ускорить процесс интерпретации скважины и увеличить качество интерпретации.

Были выделены следующие литологические типы пород:

- Глина
- Алевролит уплотненный
- Песчаник глинистый
- Песчаник

В качестве обучения использовалось 10% от общего количества скважин, но не менее 50 скважин. Математический алгоритм обучается на выделение литотипов на основе обучающей выборки. Далее, для проверки качества обучения проводится тестовая интерпретация скважин, на которых имеются результаты ручного выделения литотипов и проводится проверка эффективности работы алгоритма. После этого, в полуавтоматическом режиме проходит несколько циклов обучение-проверка, с увеличением количества обучающих скважин.

Репрезентативная обучающая выборка по скважинам должна соответствовать следующим критериям:

- Должны быть представлены все типы разреза на исследуемом объекте (типовой, врезовой, заглинизированный и т.д.)
- Разрезы скважин должны в достаточной мере содержать предполагаемые к выделению литотипы.

После обучения производится запуск алгоритма на остальных (интерпретируемых) скважинах. В качестве входных данных используется LAS-файл, который содержит стандартный набор:

DS – каверномер (необязательный параметр);

GK – гамма-каротаж;

NGK – нейтронный каротаж;

PZ – метод кажущегося сопротивления;

PS – метод потенциалов самопроизвольной поляризации.

Процесс обработки данных перед интерпретацией начинается с расчленения используемых кривых ГИС на потенциальные пласты посредством сглаживания данных и выделения межэкстремальных участков, как показано на рисунке 1. В выделенных отрезках данным присваивается постоянная величина, более приближенная к экстремальному значению в отрезке, чем к среднему арифметическому.

Рисунок 1 Построение сегментации кривой двойного разностного параметра AG

Принципиальный алгоритм выделения литологии и на Рисунок 2.

Рисунок 2 Принципиальный алгоритм выделения литологии

Для определения литотипов на основе набора кривых ГИС используется машина опорных векторов с полиномиальными ядрами. По количественным показаниям методов ГИС генерируется оптимальное разделение пространства значений методов, по положению которого в дальнейшем определяется тот или иной литотип.

Преимущества метода опорных векторов:

- Эффективен в случае, когда размерность пространства признаков выше, чем размерность обучающей выборки;
- Использует лишь некоторое подмножество обучающих данных (называемых опорными векторами), поэтому он также эффективен в отношении памяти;
- Универсальность: для процесса построения нейросетевой функции могут быть заданы различные функции ядра.

Для оценки качества выделения литологических типов пород, разработан алгоритм на основе вероятностного анализа, позволяющий на качественном уровне определить наиболее вероятный литологический тип в конкретном интервале пласта.

Вероятностная оценка реализуется на основе применения метода опорных векторов (SVM) и обучающей выборке, использованной в SVM [2]. Её существенное отличие от прямой интерпретации по разделяющей классы плоскости, получаемое в SVM, заключается в оценке положения показаний методов ГИС относительно не только полученной в обучении гиперплоскости, но и главным образом по положению относительно расположения обучающей выборки относительно гиперплоскости.

Данная реализация вероятности позволяет дополнительно оценивать показания методов ГИС в сильном соответствии с обучающей выборкой, тогда как используемое разделение классов строго по плоскости сильно обобщает и расширяет положение показаний методов ГИС, ослабляя привязанность обучения к частным случаям, что является как положительным эффектом, так и негативным в близлежащей локации к обучающей выборке.

Ниже представлена таблица поточечного сравнения выделенных литотипов при запуске алгоритма на 450 скважинах (Рисунок 3). Следует отметить, что выделение интервалов, представленных коллектором, в том числе и с ухудшенными ФЕС (алевролит), показывает высокую сходимость (нейросетевой алгоритм верно выделяет коллектор в 91% случаев).

На данный момент при разбиении разреза бобриковских отложений на 4 литотипа: 2 класса относятся к породам коллекторам и 2 класса – к породам неколлекторам. Данный подход позволяет разделять переходные фации на потенциально-коллектор (песчаник глинистый) и неколлектор (алевролит уплотненный).

		Нейронная сеть			
		Песчаник	Песчаник глинистый	Алевролит уплотненный	Глина
Интерпретация вручную	Песчаник	85.29	17.15	0.12	0.06
	Песчаник глинистый	10.95	57.06	3.41	0.34
	Алевролит уплотненный	0.59	6.22	77.17	4.06
	Глина	3.17	19.58	19.31	94.74
		Коллектор	Неколлектор	Коллектор по Нейронным сетям соответствует коллектору по Интерпретации вручную на 91%	
		Коллектор	Неколлектор		
		91.18	1.18		
		8.82	98.82		

Рисунок 3 Сравнение результатов выделения литотипов вручную и при помощи нейронных сетей

Примеры

Испытания разработанного подхода были проведены на нескольких залежах в терригенных отложениях тульского, бобриковского и турнейского возраста, разрабатываемых ПАО «Татнефть» с общим количеством скважин 3000 шт. В обрабатываемом разрезе присутствовали как скважины с выдержанными песчаными прослоями, так и скважины с частым переслаиванием различных литологических типов (Рисунок 4). Сравнение представлено с ручной интерпретацией по стандарту компании [1].

Рисунок 4 Типовой планшет окончательной интерпретации (1 – литология по автоматической интерпретации, 2 – тип насыщения, 3 – SVM анализ, 4 – автоматические стратиграфические отбивки, 5 – пористость, 6 – нефтенасыщенность, 7 – проницаемость).

Выводы

Разработанный алгоритм адаптирован к использованию информации по старому фонду скважин. Удовлетворительная сходимость полученных результатов на различных залежах нефти в бобриковских терригенных отложениях демонстрирует закономерность выбранных критериев и актуальность разработанного подхода для использования с целью оперативной интерпретации ГИС при построении геолого-гидродинамических моделей.

Благодарности

Данная работа выполнена в рамках государственной программы повышения конкурентоспособности Казанского (Приволжского) федерального университета среди ведущих мировых научно-образовательных центров

Авторы выражают благодарность ПАО «Татнефть» им. В.Д.Шашина, в особенности – коллективу Центра Моделирования ПАО «Татнефть» им. В.Д.Шашина за тесное сотрудничество и возможность опубликовать полученные результаты.

Библиография

1. Интерпретация ГИС, алгоритмы определения параметров продуктивных пластов нефтяных месторождений Республики Татарстан. Стандарт организации ОАО «Татнефть» им. В.Д. Шашина СТО ТН – 2013.

2. Wu, Lin and Weng, “Probability estimates for multi-class classification by pairwise coupling”, JMLR 5:975-1005, 2004.

References

1. Interpretatsiya GIS, algoritmyi opredeleniya parametrov produktivnyih plastov neftyanyih mestorozhdeniy Respubliki Tatarstan. Standart organizatsii PJSC TATNEFT STO TN – 2013.
2. Wu, Lin and Weng, “Probability estimates for multi-class classification by pairwise coupling”, JMLR 5:975-1005, 2004.