

doi:10.25689/NP.2019.4.157-171

УДК 004.896:553.98

**НЕЙРОСЕТЕВОЕ ДЕТЕКТИРОВАНИЕ ПРОДУКТИВНЫХ
ИНТЕРВАЛОВ НА ПРИМЕРЕ ОБЪЕКТА БВ10 САМОТЛОРСКОГО
НЕФТЕГАЗОКОНДЕНСАТНОГО МЕСТОРОЖДЕНИЯ**

Канаев И.С.

ООО «Тюменский нефтяной научный центр»

НИ «Томский политехнический университет»

**PAY INTERVALS DETECTION BY NEURAL NETWORK ON THE
EXAMPLE OF THE BV10 RESERVOIR OF THE SAMOTLOR OIL FIELD**

Kanaev I.S.

LLC «Tyumen Petroleum Research Center»

«National research Tomsk polytechnic university»

E-mail: Kanaevis@tnc.rosneft.ru

Аннотация. Данная работа посвящена анализу применимости нейросетевого подхода при решении задачи автоматического выделения продуктивных интервалов. Методы машинного обучения позволяют наиболее быстрым способом обрабатывать большие массивы данных, а также выявлять необходимые признаки и связи. Проблематика данной работы заключается в создании нейронной сети, которая позволит наиболее точно определять продуктивные интервалы, используя данные ГИС. Для получения точного результата одним из наиболее значительных аспектов является подготовка данных для исследования. Предобработка данных является обязательным условием для любого метода машинного обучения. Полученные результаты были соотнесены с результатами интерпретации специалиста геофизика. Разработанная нейросетевая модель позволяет автоматизировать процесс выделения продуктивных интервалов.

Ключевые слова: машинное обучение, нейронные сети, выделение продуктивных интервалов, анализ последовательностей, предобработка данных.

Abstract. This paper is devoted to the applicability analysis of the neural network usage for automatic pay intervals detection. Machine learning methods allow the fastest way to process large data arrays, as well as to identify the necessary signs and relationships. The

problem of this work is to find the optimal neural network, which will most accurately determine the pay intervals using well logs data. To obtain an accurate result, one of the most significant aspects is the preparation of data for the study. Preprocessing of data is a prerequisite for any method of machine learning. The results obtained were compared with the results of geoscientist's interpretation. The selected algorithm allows automating the process of pay zones detection.

***Key words:** Machine learning, neural network, pay intervals detection, sequence analysis, data preprocessing*

Современный этап развития нефтегазовой отрасли в мире и в Российской Федерации в частности, тесно связан с внедрением широкого спектра технологий автоматизированного анализа данных. Так в Тюменском нефтяном научном центре (входит в корпоративный научно-проектный комплекс ПАО «Роснефть») ведется инновационная работа по цифровизации согласно принятой стратегии «Роснефть-2022». Актуальной задачей в настоящее время является автоматизация процесса обработки разносторонней промысловой информации, объемы которой постоянно растут. Для зрелых месторождений особенно остро стоит вопрос поддержания уровней добычи, решение которого, в том числе, может заключаться в повторном анализе существующей петрофизической информации и выделении ранее недренируемых нефтенасыщенных интервалов. Такая задача переинтерпретации данных является многовариантной, трудозатратной и нетривиальной. При использовании аппарата машинного обучения, а в частности нейронных сетей, процесс детектирования продуктивных интервалов можно в значительной степени автоматизировать.

Обобщенно область автоматизации интеллектуальных задач, для решения которых обычно необходим человек, называют искусственным интеллектом [1]. Термин «искусственный интеллект» в свою очередь объединяет ряд таких понятий, как машинное обучение и глубокое обучение. Нейронные сети – один из методов машинного обучения. Для

методов машинного обучения характерно не прямое решение задачи, а процесс нахождения статистической структуры в многочисленных анализируемых примерах, которая затем позволяет создать правила для получения автоматического решения задачи [1]. Сам процесс нахождения такой структуры и называется обучением.

С позиции детектирования продуктивных интервалов при анализе каротажных данных с применением нейронных сетей интерес представляет ряд исследовательских работ. Так в своей работе Gharib Namada и др. [2] использовали нейронные сети для задачи выделения нефтенасыщенных толщин. Данную задачу авторы разделили на несколько подзадач: определение пористости и водонасыщенности, для решения каждой задачи использовались нейронные сети. В целом авторы анализировали следующие каротажные кривые: гамма-каротаж (ГК), боковой каротаж (БК), плотностной гамма-гамма каротаж (ГГК-П), нейтронный каротаж (НК), селективный гамма-гамма каротаж (ГГК-С) и акустический каротаж (АК). Для определения пористости авторы использовали полносвязную сеть прямого распространения с двумя скрытыми слоями с 25-ю и 12-ю нейронами соответственно, на скрытых слоях использовалась такая функция активации, как гиперболический тангенс, а на выходном слое применена логистическая. Для решения задачи определения водонасыщенности использовалась схожая сеть, однако на скрытых слоях она содержала 16 и 5 нейронов соответственно. Полученные результаты точности определения параметров для одной приведенной в исследовании скважины составляют 99%.

Аналогичную задачу решали и хорватские специалисты Marko Svetkovic и др. [3]. В своем исследовании специалисты применили нейронные сети для определения нефтенасыщенных интервалов по данным месторождения Клоштар (Хорватия). Данную задачу авторы представили в виде задачи бинарной классификации, где

нефтенасыщенные интервалы были объединены в класс «1», а интервалы неколлекторы и водонасыщенные участки в класс «0». В работе рассматриваются различные варианты реализации сети, но наилучшим оказался вариант реализации полносвязной сети с двумя скрытыми слоями с 6-ю и 8-ю нейронами соответственно, точность результатов составила 91 %.

Следует упомянуть работу специалистов из Иранского нефтяного университета. В своей статье [4] Mohammad Mardi и др. исследуют возможность определения водонасыщенности на скважинах крупного нефтяного месторождения Азадеган (Иран) с помощью аппарата машинного обучения. Авторами была принята следующая конфигурация сети: полносвязная трехслойная сеть прямого распространения с одним скрытым слоем, содержащим 25 нейронов. На вход сети подавались следующие типы каротажных данных: БК, НК, ГГК-П, АК и ГК. Исследовались различные варианты реализации сетей с позиции изменения функций активации нейронов. Так наилучший коэффициент корреляции между керновой водонасыщенностью и значением, полученным с помощью нейросети, составил 0,97 при использовании логистической функции активации на скрытом слое.

В последнее время для решения широкого спектра петрофизических задач применяются сверточные нейронные сети. Сверточные сети относятся к методам глубокого обучения и направлены на анализ изображений.

В обзорной статье Chicheng Xu и др. отмечается использование сверточных сетей при анализе каротажных данных [5]. На основе кривых геофизических исследований скважин с применением сверточных нейронных сетей производится классификация отложений по литологии, также в работе упоминается анализ седиментационных структур по фотографиям.

Сравнительный анализ применения сверточной нейронной сети и персептрона для решения задачи определения литологии по данным геофизических исследований скважин проводился специалистами Пекинского нефтяного университета Zhu Liping и др. [6]. Используемый набор каротажных данных скважин месторождения Дацин состоял из: ГК, АК, ГГК-П, БК и микроградиентное зондирование (МГЗ). Точность определения литологии вертикального разреза скважин для сверточных нейронных сетей составила 92,3%, а для обычной трехслойной полносвязной нейронной сети прямого распространения с одним скрытым слоем процент точности определения литологии составил 82,6 %.

В настоящей работе исходными данными для анализа являлся набор каротажных исследований скважин, вскрывающих отложения объекта БВ10 на Самотлорском нефтегазоконденсатном месторождении. Было проанализировано 130 скважин, расположение которых представлено на рис.1. Зачастую не все скважины обладают одинаковым набором каротажных исследований, но чем больше уникальных признаков, описывающих систему, имеется, тем точнее можно решать задачу классификации. В результате анализа представительности различных совокупностей каротажных данных, была сформирована следующая выборка: каротаж потенциалов собственной поляризации (ПС), ГК, индукционный каротаж (ИК), НК и четыре разноглубинных измерения бокового каротажного зондирования (БКЗ).

Данные одинаковых методов каротажных измерений в соседних скважинах могут иметь значительно отличающиеся средние значения ввиду различных условий проведения измерений, содержать в себе пропуски, а также могут быть прописаны в различных диапазонах. Ввиду перечисленных проблем перед обучением нейронной сети проводилась предобработка данных.

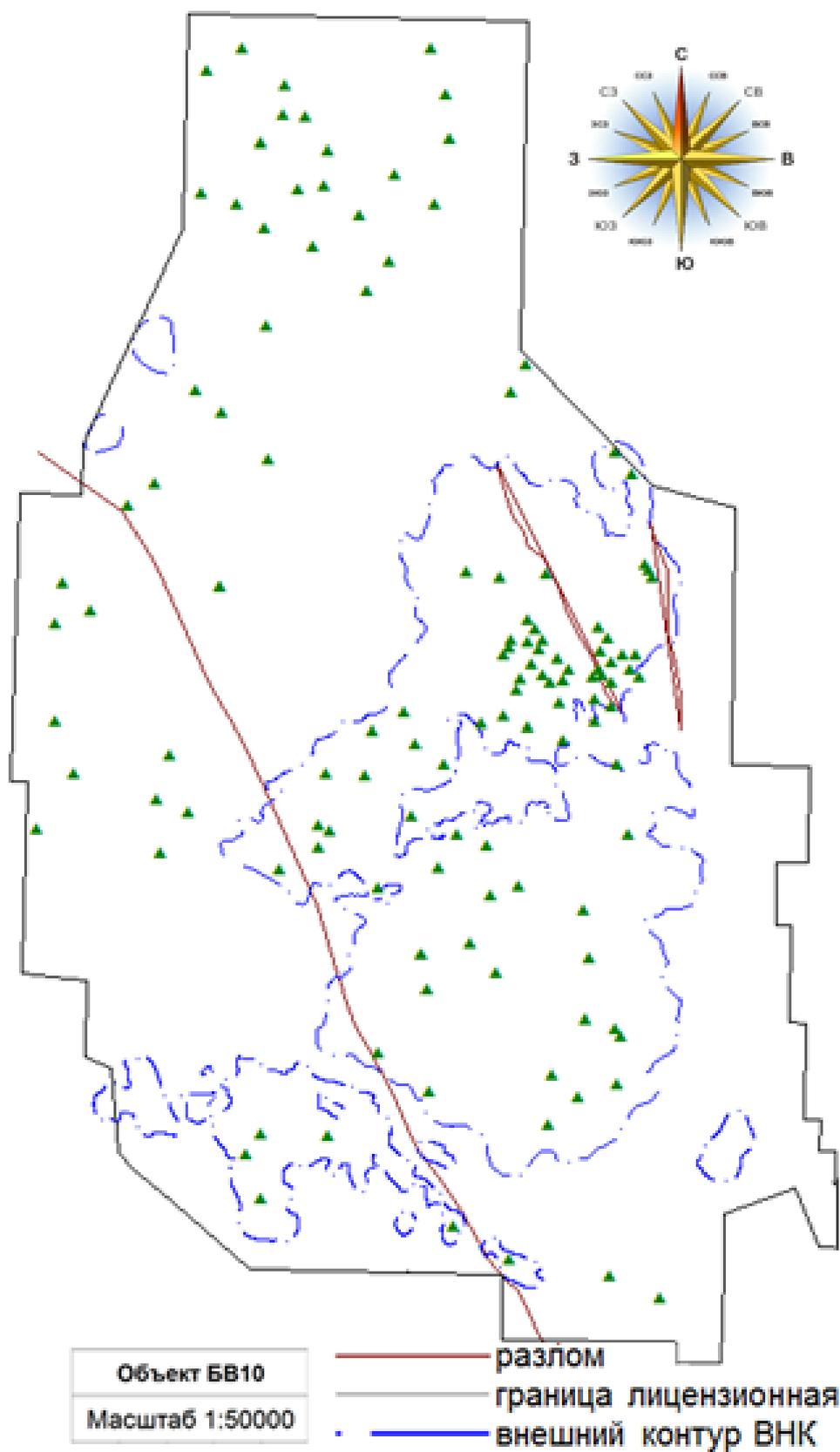


Рис. 1. Схема расположения исследуемых скважин объекта БВ10 Самотлорского месторождения

На первом этапе подготовки данных по каждой скважине отдельно составляются исходные каротажные выборки и при этом отбрасываются пропущенные или неверные значения.

Затем для решения проблемы различающихся условий измерений проводится нормировка каротажей путем вычисления стандартизированных оценок (z -оценок) для устранения негативных эффектов от разномасштабных наборов данных с разных скважин:

$$z_i = \frac{x_i - \bar{X}}{S_x}, \quad (1)$$

где $X = \{x_1, x_2, \dots, x_n\}$ – множество данных каротажа со средним значением \bar{X} и стандартным отклонением S_x , z_i – нормированные значения.

Ввиду различного шага измерений в разных скважинах при составлении обучающей выборки каротажные данные разбиваются по абсолютной глубине с шагом дискретизации в 0,1 м.

Задачу определения продуктивных зон можно представить в виде бинарной классификации, как и в работе [3], объединив неколлекторы и коллекторы, насыщенные водой в класс «0». Класс «1» при этом представляет собой нефтенасыщенные коллекторы, которые имели расчетную нефтенасыщенность выше значения остаточной нефтенасыщенности, а при интерпретации специалистом-геофизиком получили значение «нефть» или «нефть с водой».

Важным вопросом является определение входного вектора для подачи в нейронную сеть. Ряд работ основывается на поточечном составлении обучающей выборки, т.е. на вход сети подается n -мерный вектор признаков – значения n -различных каротажей для данной глубины скважины. Таким образом, чем больше будет значимых параметров в обучающем векторе, тем точнее будет обучаться классификатор. Однако, ввиду специфики каротажных данных, для ряда скважин с одного месторождения сложно составить репрезентативную выборку, т.к. скважин

с одинаковыми наборами каротажей немного. В данной работе задача классификации и выделения продуктивных интервалов будет рассматриваться в качестве задачи анализа временных рядов.

Последовательность значений некоторой величины $X = \{x_1, x_2, \dots, x_n\}$, которая меняется во времени $\{t_1, t_2, \dots, t_n\}$ называется временным рядом. Задача классификации временных рядов состоит в поиске как можно более точного отображения $f: X \rightarrow \{1, 2, \dots, N\}$, которое принимает на вход временной ряд и ставит ему в соответствие один из N классов, а процесс поиска такого отображения называется обучением. Обучение осуществляется по обучающей выборке – набору $[(X_1, y_1), (X_2, y_2), \dots, (X_m, y_m)]$ объектов с известными ответами искомой функции. Результатом обучения можно записать в следующем виде:

$$Acc = \frac{1}{K} \sum_i^K [f(X_i) = y_i], \quad (2)$$

где Acc – процент верно классифицированных объектов, $f(X_i)$ – функция, которая максимизирует качество классификации на тестовой выборке (X_i, y_i) размером K , y_i – метка класса.

Для составления входной выборки использовался метод «скользящего окна». По относительной глубине проходят скользящим окном размером в n измерений. Затем в i -й образец данных последовательно записываются попавшие в окно данные: сначала n точек из гамма каротажа, затем n точек из нейтронного и т.д., что в результате формирует $(n \cdot 8)$ -мерный входной вектор.

В работе используется ширина окна $n = 5$, что формирует входной вектор из 40 значений, только если все данные окна относятся к одному классу, т.е. выходной вектор является числом 0 или 1. Шаг прохода скользящим окном определяется смещением вниз на единицу данных, что способствует увеличению входной выборки.

Еще одним немаловажным недостатком исходных данных является несбалансированное количество обучающих векторов по классам – примеры данных класса «1», или иными словами «нефтенасыщенный коллектор», составляют порядка 10 % от всей составленной выборки. Этот факт может привести к «ленивому обучению», т.е. даже если классификатор для любых обучающих примеров будет выдавать результат «0», то точность классификации составит 90 %. Для решения данной проблемы составленная выборка урезалась следующим образом: оставались все положительные (нефтенасыщенный коллектор или класс «1») примеры, оставались отрицательные примеры, находящиеся на границе с положительными, и далее оставались в случайном порядке отрицательные примеры, пока соотношение классов не составит примерно 60/40. Также для обучения нейронной сети используется перекрестная энтропия в качестве критерия оптимизации, которая уменьшает меру неопределенности классификатора.

В ходе реализации проекта была создана полносвязная нейронная сеть прямого распространения с тремя скрытыми слоями, представленная на рис. 2. Как уже говорилось выше, входной слой состоит из 40 нейронов и имеет функцию активации гиперболический тангенс. Первый скрытый слой состоит из 20 нейронов, второй из 12 и третий из 10 нейронов, на скрытых слоях используется функция активации ReLU. Данная функция активации характеризуется выражением $f(x) = \max(0, x)$. Использование данной функции активации обусловлено быстрым обучением и тем, что она менее требовательна к вычислительным ресурсам, так как производит более простые математические операции. На выходном слое, состоящем из одного нейрона, используется сигмоидальная функция активации, возвращающая значение от 0 до 1. Для бинарной классификации вещественный выход сети сравнивается со значением 0,5, если выход меньше среднего значения 0,5, то класс интерпретируется как 0, иначе как 1.

Следует также выделить другие аспекты реализации сети: в качестве метода оптимизации применялся Adam. При обучении сети использовался метод обучения с помощью мини-пакетов (mini-batch). Для этого весь массив обучающих данных случайным образом разбивается на пакеты размером по 1024 экземпляра. На каждой эпохе обучения на вход сети подаются пакеты данных, затем производится подсчет ошибки и корреляция весов в ходе обратного прохода по сети. Лучшие результаты по точности дало обучение с использованием L2-регуляризация на скрытых слоях с весовым значением 0,002, а также использовалась перекрестная энтропия в качестве критерия оптимизации. Нейронная сеть была создана с помощью библиотеки для машинного обучения TensorFlow на языке программирования Python.

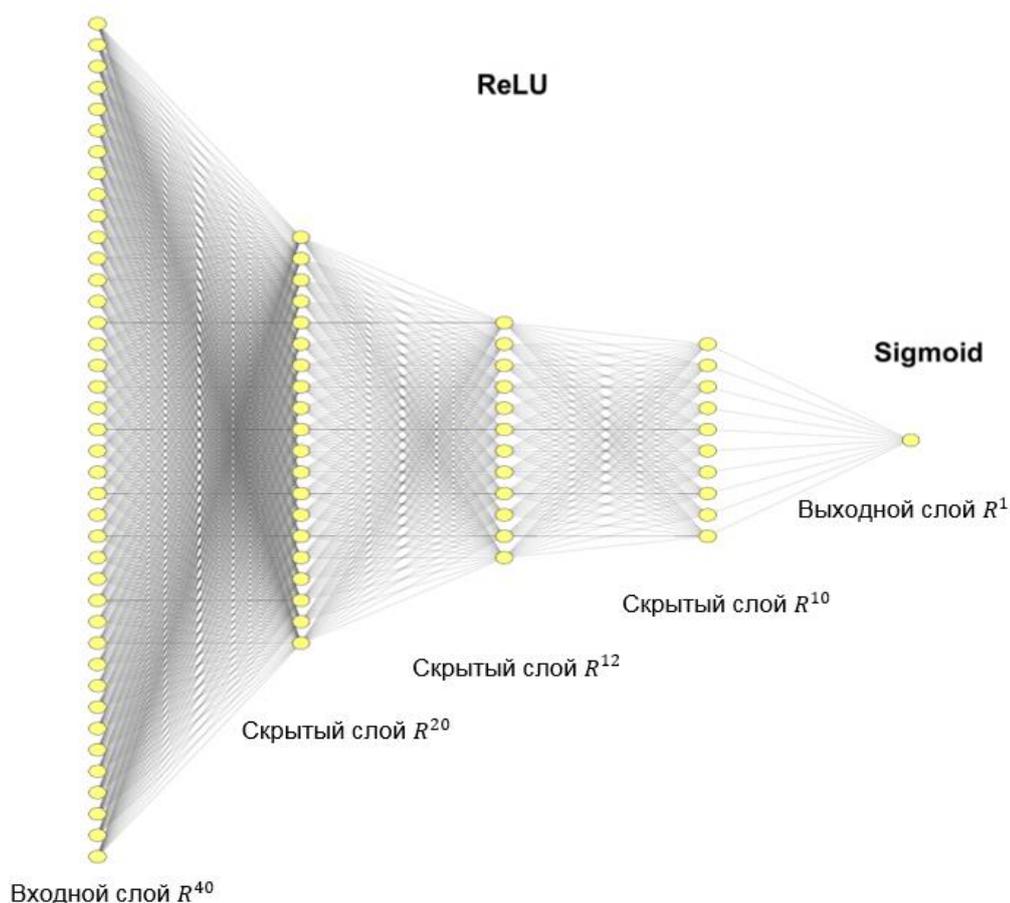


Рис. 2. Архитектура используемой нейронной сети

Для контроля результатов обучения входная выборка была разбита на тренировочную, тестовую и валидационную. Также отдельно

были исключены данные по 3 скважинам для визуализации результатов и оценки качества детектирования продуктивных интервалов относительно конкретных скважин. Достигнутая точность нейронной сети составила на обучающей выборке 97 %, а на тестовой 92 % и 86 % на валидационной. График изменения ошибки предсказания для тестовой и валидационной выборок представлен на рис. 3. Процесс обучения был остановлен на 102 эпохе.

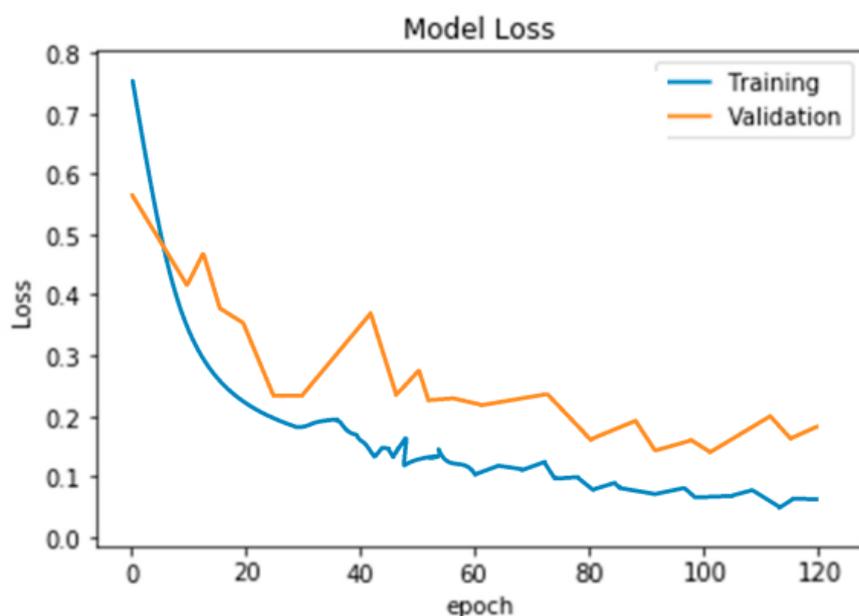


Рис. 3. График изменения ошибки на тестовом множестве и на валидационной выборке.

Результаты детектирования продуктивных интервалов по трем исключенным скважинам представлены на рис. 4. Так для скважины 11-Х (Рис. 4а) результат классификации полностью совпал с интерпретацией эксперта. Для скважины 10-У (Рис. 4б) нейронная сеть подтвердила выделенный экспертом интервал, дополнительно выделив несколько маломощных пропластков мощностью менее 0,4 метра. Для скважины 12-З (Рис. 4в) прогноз разработанного классификатора также в целом совпал с экспертной интерпретацией, дополнительно был выделен один маломощный интервал, нефтенасыщенная толщина которого меньше 0,4 метра.

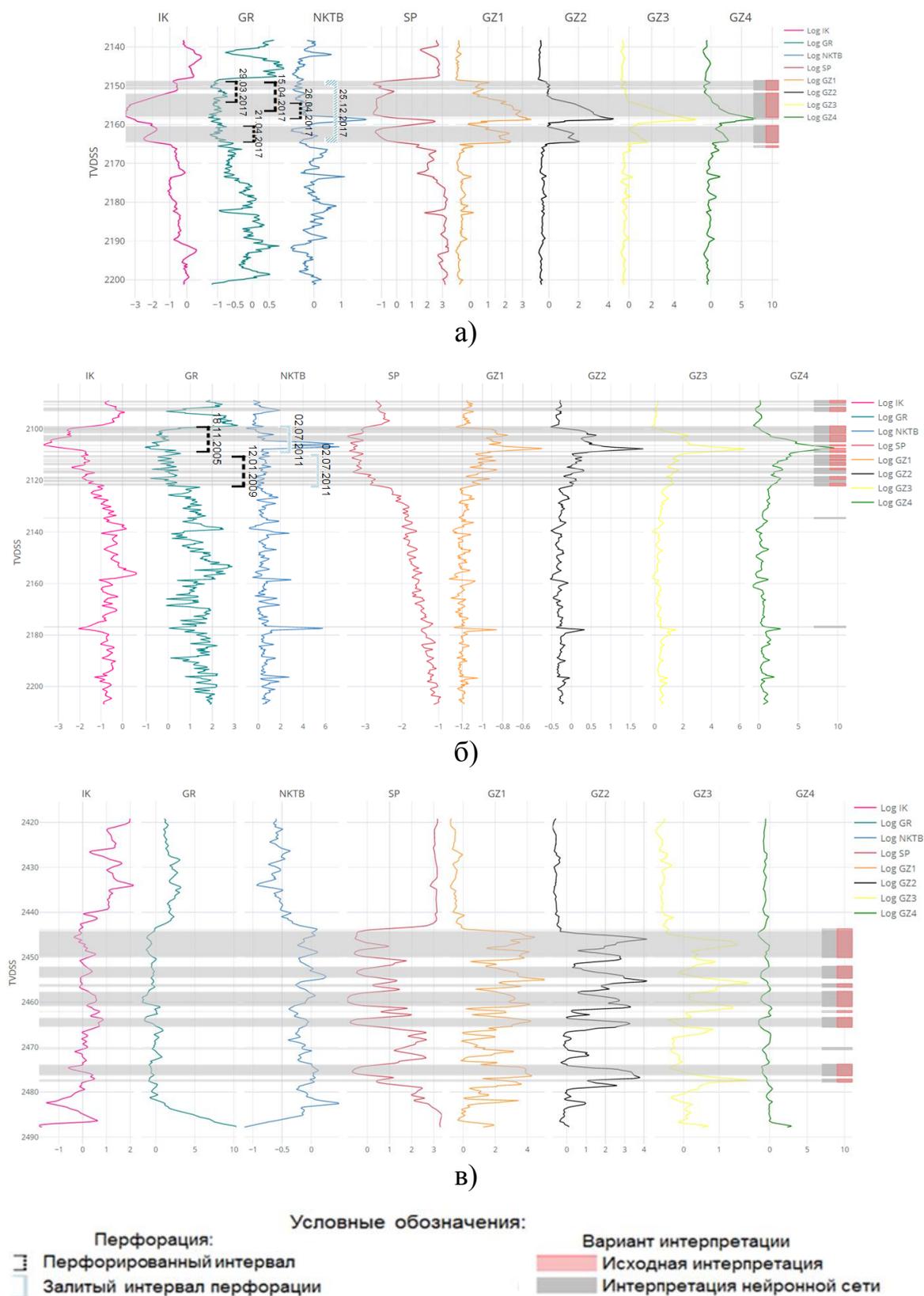


Рис. 4. Сравнение полученных результатов выделения продуктивных интервалов для полностью исключенных из обучения скважин:
а) 11-X; б) 10-Y; в) 12-Z (где IK–ИК, GR–ГК, NKTБ–НК, SP–ПС, GZ1, GZ2, GZ3, GZ4 – зонды различной глубинности БКЗ)

Выводы:

1. Применение нейронных сетей позволяет автоматизировать процесс выделения продуктивных интервалов.
2. В результате проведения исследования нейросетевое детектирование показало высокую точность. Достигнуто значение точности при обучении в 97 % и 92 % при проверке на тестовой выборке, а также 86 % на валидационном наборе данных.
3. Использование полносвязной нейронной сети прямого распространения с представленной архитектурой позволило достичь высокой сходимости нейросетевой и экспертной интерпретаций.
4. Полученный аппарат детектирования в дальнейшем может быть применен для автоматизированной переинтерпретации большего фонда скважин с целью выделения пропущенных нефтенасыщенных интервалов.

Список литературы

1. Шолле, Ф. Глубокое обучение на Python [Текст]: монография-СПб: Питер, 2018. – 400 с.
2. Хамада, Г. Нейросетевой расчет пористости и водонасыщенности для песчано-глинистых коллекторов [Текст]/ Г. Хамада, Э. Ахмед, Н. Чао// Достижения в области прикладных научных исследований. - 2018. - № 8. С. 26-31.
3. Цветкович, М. Применение нейронных сетей для определения литологии коллекторских горных пород и расчета насыщенности [Текст]/ М. Цветкович, Й. Велик, Т. Малвик// Геология Хорватии. - 2009. - №62. С. 115-121.
4. Марди, М. Расчет водонасыщенности с применением нейронных сетей и исследование влияния изменения коэффициента цементации и показателя насыщенности на Иранской нефтяной скважине [Текст]/ М. Марди, Х. Нуроzi, С. Эдалатха// Нефтегазовая наука и технологии. – 2012. -№30. С. 42-434.
5. Чиченг, Х. Когда петрофизика сталкивается с огромным массивом данных: Что может машинное обучение? [Текст]/ Х. Чиченг, С. Мисра, П. Сринивасан, М. Шушианг// Ближневосточная нефтегазовая конференция SPE. - 2019. - 24с.
6. Чжу Л. Улучшенная литологическая интерпретация каротажных данных с применением сверточных нейронных сетей [Текст]/ Чжу Л., Ли Х., Ян Ч., Ли Ч., Ао И. // Петрофизика. - 2018. -№ 6. С. 799-810.

References

1. François Chollet *Glubokoe obuchenie na Python* [Deep Learning with Python]. St. Petersburg: Piter Publ., 2018 (translated into Russian). 400 p.
2. Hamada G., Ahmed E., Chao N. *Neirosetevoi raschet poristosti i vodonasyshchennosti dlya peschano-glinistyh kollektorov* [Neural network prediction of porosity and water saturation for sand shale reservoirs] *Dostizheniya v oblasti prikladnyh nauchnyh issledovaniy*, No.8, 2018. pp. 26-31.
3. Tsvetkovich M. Velik J., Malvik T. *Primenenie nejronnyh setej dlya opredeleniya litologii kollektorskih gornyh porod i rascheta nasyshchennosti* [Neural network for determination of reservoir rock lithology and prediction of saturation]. *Geologiya Horvatii*, No. 62, 2009. pp. 115-121.
4. M. Mardi, H. Nurozi, S. Edalatkhah *Raschet vodonasyshchennosti s primeneniem nejronnyh setej i issledovanie vliyaniya izmeneniya koefficienta cementacii i pokazatelya nasyshchennosti na Iranskoj neftyanoj skvazhine* [A water saturation prediction using artificial neural networks and an investigation on cementation factors and saturation exponent variations in an Iranian oil well]. *Neftegazovaya nauka i tekhnologii*, No. 30, 2012. pp. 42-434 (translated from English)
5. H. Chicheng, S. Misra, P. Srinivasan, M. Shuhian *Kogda petrofizika stalkivaetsya s ogromnym massivom dannyh: Chto mozhet mashinnoe obuchenie?* [When petrophysics meets big data: what can machine do?]. *SPE Middle East Oil and Gas Conference*, 2019. 24 p.
6. Zhu L., Li H., Yang Zh., Li Ch., Ao Yi. *Uluchshennaya litologicheskaya interpretaciya karotazhnyh dannyh s primeneniem svertochnykh neironnykh setei* [Intelligent logging lithological interpretation with convolution neural networks]. *Petrophysics*, Vol. 59, Iss. 06, December 2018. pp. 799-810.

Сведения об авторах

Канаев Илья Сергеевич, специалист управления по ГРМ Самотлорнефтегаз,
ООО «Тюменский нефтяной научный центр», г. Тюмень, Российская федерация
E-mail: iskanaev@tnnc.rosneft.ru

Authors

Kanaev I.S., LLC «Tyumen Petroleum Research Center», Tyumen, Russian Federation
E-mail: iskanaev@tnnc.rosneft.ru

Канаев Илья Сергеевич
625002, Российская Федерация
г. Тюмень, ул. Осипенко 79/1
Тел.: +7 (3452) 52-90-90 вн. 7981
E-mail: iskanaev@tnnc.rosneft.ru