

УДК 622.276:004.896

## ГЛУБОКИЕ ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ ДЛЯ ПРОГНОЗА ЗНАЧЕНИЙ ДЕБИТОВ ДОБЫВАЮЩИХ СКВАЖИН

Евсюткин Иван Викторович<sup>1</sup>,  
ive1@tpu.ru

Марков Николай Григорьевич<sup>1</sup>,  
markovng@tpu.ru

<sup>1</sup> Национальный исследовательский Томский политехнический университет,  
Россия, 634050, г. Томск, пр. Ленина, 30.

**Актуальность** исследования обусловлена необходимостью поддержки принятия решения специалистами добывающих предприятий нефтегазовой отрасли при управлении производством. Точное прогнозирование значений дебитов добывающих скважин позволяет определить такие технологические режимы работы фонда скважин и технологического оборудования, которые позволяют достичь заданного объёма выпуска продукции. Существующие методы не всегда обеспечивают требуемый уровень точности при прогнозе значений дебитов скважин, что приводит к ошибкам при расчёте экономического эффекта при оценке рентабельности добывающих скважин и последующих поставок углеводородного сырья, а также при учёте ограничений, накладываемых органами надзора за использованием недр.

**Цель:** разработать и предложить наиболее эффективные модели глубоких искусственных нейронных сетей при прогнозе значений компонентов добычи углеводородного сырья – нефти, газа, жидкости (газового конденсата) и воды.

**Объекты:** технологические параметры дебитов добывающих скважин фонда нефтяных, газовых, нефтегазовых и нефтегазоконденсатных месторождений.

**Методы:** методы анализа больших объёмов технологических данных скважин, развиваемые в соответствии с концепцией «Big Data»; модели глубоких искусственных нейронных сетей; объектно-ориентированное программирование; методы оценки и статистического анализа результатов исследований эффективности глубоких искусственных нейронных сетей при прогнозе значений дебитов добывающих скважин.

**Результаты.** Разработана методика подготовки данных по дебитам скважин для обучения и тестирования глубоких искусственных нейронных сетей прямого распространения. Проведены исследования различных архитектур таких нейросетей при решении задач прогноза дебитов нефти, газа, жидкости (газового конденсата) и дебита воды. Выявлены наиболее эффективные архитектуры глубоких нейросетей прямого распространения. Такие нейронные сети позволяют увеличить точность прогноза в два и более раза по сравнению с точностью прогноза, даваемой традиционным методом экстраполяции по скользящей средней.

### Ключевые слова:

Месторождение нефти и газа, добывающая скважина, прогноз значений дебитов скважин, модели глубоких искусственных нейронных сетей, исследование эффективности глубоких нейронных сетей прямого распространения.

### Введение

При добыче углеводородного сырья (УВС) на промыслах происходит измерение большого числа технологических и геологических параметров. Значения этих параметров служат для оценки состояния скважин и продуктивных пластов, а также позволяют понять, какова технологическая и экономическая эффективность проведённых работ на фонде скважин, в том числе и геолого-технических мероприятий (ГТМ).

Задача прогноза значений параметров актуальна не только для расчёта технологических режимов скважин, но и для управления фондом скважин, так как выявленный критический тренд может позволить заранее установить необходимость проведения определённого вида ГТМ на скважине [1]. Однако существующие методы прогноза значений дебитов скважины в лучшем случае позволяют выполнить экстраполяцию данных путём линейной регрессии [2, 3]. Разумеется, столь простые методы не способны учесть множество аспектов, заложенных в данных, и взаимовлияние их друг на друга в рамках получаемых временных рядов для параметров скважин [4]. Соот-

ветственно, точность прогноза не всегда удовлетворяет специалистов нефтегазодобывающих предприятий. Сделаны первые шаги в использовании методов интеллектуального анализа данных при решении таких задач, давшие первые обнадеживающие результаты [5–7].

Всё это указывает на актуальность разработки новых методов интеллектуального анализа данных, позволяющих существенно повысить точность прогноза значений параметров скважин и продуктивных пластов.

В данной работе для прогноза значений дебитов скважин было предложено использовать глубокие искусственные нейронные сети (ИНС) [8], показавшие свою эффективность при решении задач экстраполяции данных в других предметных областях. Проведено исследование эффективности различных архитектур глубоких ИНС прямого распространения при решении таких задач прогноза. Результаты исследований показали, что использование этих ИНС позволяет значительно повысить точность прогноза значений дебитов скважин по сравнению с точностью, достигаемой с помощью известных методов.

### Постановка задачи исследования

Целью планирования работы производства нефтегазодобывающего предприятия является определение таких технологических режимов работы фонда скважин и технологического оборудования, которые позволят достичь заданного объёма выпуска продукции (нефти, газа и газового конденсата). При этом наиболее важно планирование технологических режимов каждой добывающей скважины фонда и технологических режимов установок подготовки продукции из добываемого УВС, а также планирование режимов транспортировки продукции до магистральных нефте- и газопроводов.

Обычно на эксплуатируемых месторождениях прогнозу подлежат довольно большое число измеряемых параметров добывающих скважин и продуктивных пластов, среди них: дебит нефти, дебит газа, дебит жидкости (конденсата), дебит воды, давление забойное, давление пластовое, обводнённость и т. д. Однако наиболее востребованными являются прогнозы значений дебитов скважин.

Точное предсказание дебитов скважин является необходимой частью управления производством любого нефтегазодобывающего предприятия. Важно прогнозировать (планировать) дебиты не просто на конкретный день, а на определённый промежуток времени (например, равный месяцу), чтобы понимать, сколько в среднем УВС будет добываться из скважины в сутки. В этом случае говорят о планировании технологического режима скважины. Значение технологического режима скважины позволяет рассчитать экономический эффект от добычи УВС и понять рентабельность эксплуатации конкретной скважины, построить перспективные планы поставок продукции и учесть ограничения, наложенные органами надзора за пользование недрами [9].

Для обеспечения высокой точности прогноза значений дебитов скважины было предложено использовать глубокие ИНС прямого распространения. Поэтому была поставлена задача исследования эффективности таких ИНС при решении задачи прогноза значений дебитов при использовании реальных данных с эксплуатируемого месторождения.

Для проведения исследований использовался архив данных одного из нефтегазоконденсатных месторождений Томской области, где за 6 лет эксплуатации накоплен значительный объём данных. Фонд добывающих скважин содержит 17 кустов скважин, в каждом кусте от 6 до 21 добывающей скважины, всего 142 добывающих скважины.

Для экспериментов были выбраны исключительно параметры типа дебитов (ежесуточные объёмы добычи компонентов УВС – нефти, воды и т. д.), так как они являются наиболее значимыми интегральными показателями добычи УВС. Для них в базе данных (БД) предприятия обычно имеется наибольший объём данных, не является исключением и предприятие, из архива которого были взяты данные для исследований.

### Методика подготовки данных для обучения и тестирования искусственных нейронных сетей

На предприятии использовалась методика прогноза (планирования) значений дебитов, основанная на методе экстраполяции по скользящей средней [10, 11]. Для каждого из дебитов скважин в анализируемой БД имеются прогнозное и фактическое значения. Прогноз геологической службой предприятия, судя по имеющимся данным, осуществлялся от –8 до 229 дней до измерения фактического значения дебита. То есть в БД имеются явно ошибочные значения параметров, по крайней мере, при планировании значений дебитов, так как нельзя планировать значение параметра уже после получения его фактического значения. Также вызывают сомнения прогнозы на большие временные промежутки (на 2 и более месяцев). Все такие данные отбраковывались и в экспериментах не использовались [12].

Все фактические измеренные значения дебитов на дату планирования и ранее подаются на входы ИНС (признак, англ. feature). Фактические значения дебитов через месяц после этой даты являются эталоном, то есть они служат эталоном выходным значениям ИНС (метка, англ. label) при её обучении и тестировании.

Данные, которые находятся между датой планирования и более поздней датой получения фактического значения дебита-эталона, не используются, несмотря на их наличие в БД, чтобы находиться в тех же самых условиях, что и находилась геологическая служба предприятия, планировавшая технологический режим работы скважины. Это позволит корректно осуществить сравнение полученных при исследовании результатов с имеющимися в БД значениями прогнозных (плановых) дебитов.

При проведении экспериментов было решено брать значения дебитов от даты планирования в течение трёх–пяти недель до получения фактического значения. Это практически важно при формировании технологических режимов скважин. Исходное число записей в БД 17117, после сужения периода планирования до заданного записей осталось 13032. Видно, что такие действия уменьшили объём обучающей выборки. Однако это позволило нам получить более точную оценку результатов прогноза значений дебитов на месяц по существующей на предприятии методике.

Число анализируемых записей и рассчитанные нами значения взвешенных средних абсолютных ошибок (англ. Weighted absolute percent error – WAPE) [13] при прогнозе каждого параметра (вида дебита) по всему месторождению по методике предприятия указаны в табл. 1. При разработке наших моделей ИНС и при оценке их точности необходимо ориентироваться на уменьшение WAPE при прогнозе значений этих дебитов.

Текущий или капитальный ремонт скважин, остановка их по другим причинам или на проведение ГТМ вносят непредсказуемость в изменение параметров скважин. Это должно быть учтено при составлении обучающих примеров, чтобы в вектор признаков на входах ИНС по каждой скважине не попадали зна-

чения дебитов одновременно до и после проведения тех или иных ГТМ или остановок скважин. То есть планирование технологического режима скважины должно осуществляться исключительно с учетом периодов её эксплуатации.

**Таблица 1.** Оценки точности прогноза параметров по методике предприятия

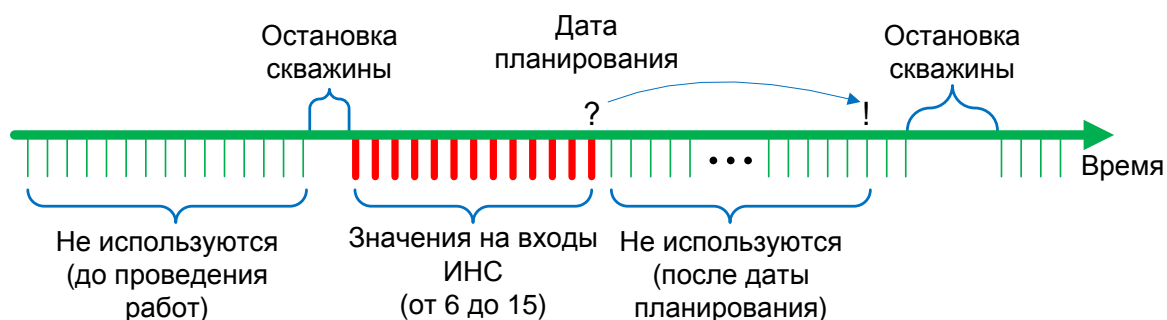
**Table 1.** Estimates of the forecasting precision of parameters by the enterprise method

Название параметра Name of parameter	Обозначение Designation	Единица измерения Measuring unit	Число записей Record count	Взвешенная средняя абсолютная ошибка, % Weighted absolute percent error, %
дебит нефти debit of oil	$Q_n$ $Q_o$	т/сут tons/day	3557	27,347
дебит газа debit of gas	$Q_r$ $Q_g$	тыс. м <sup>3</sup> /сут 1000 m <sup>3</sup> /day	3458	35,133
дебит жидкости debit of fluid	$Q_{ж}$ $Q_f$	м <sup>3</sup> /сут m <sup>3</sup> /day	3459	26,734
дебит воды debit of water	$Q_в$ $Q_w$	т/сут tons/day	2558	67,044

Не ясно, какое оптимальное число значений дебита нужно подавать на входы ИНС, это предстоит определить экспериментально. Также открытым остаётся вопрос, стоит ли подавать на входы ИНС значения только одного какого-либо дебита или добавлять к ним ещё значения других параметров, например,

дебита другой компоненты УВС или пластового давления продуктивного пласта. Так как число таких сочетаний очень велико, подобные эксперименты не выполнялись, и в качестве входных значений для ИНС использовались только фактические значения прогнозируемого дебита.

Суть предлагаемой методики подготовки данных в том, чтобы увеличивать при прогнозе число входных значений дебитов для ИНС за счёт расширения числа известных из БД фактических значений дебитов от 6 (6 дней назад от даты планирования) с постепенным добавлением таких значений за предыдущие дни до 15. Схематично процесс выбора данных изображён на рис. 1, где показано движение скважины во времени. Видно, что скважина была два раза остановлена на разные промежутки времени. Жирными красными штрихами отмечены взятые в эксперимент 13 фактических значений дебита на скважине, получаемые ежедневно. Пусть дата прогноза значения дебита будет отмечена знаком «!» (метка для ИНС), и на эту дату в БД должно быть фактическое значение дебита. Дата, с которой планируется значение дебита, показана знаком «?». С даты планирования и по дату прогноза должен пройти один месяц (период планирования). Все фактические значения дебита (от 6 до 15 значений) до даты планирования становятся признаками, подаваемыми на входы ИНС (отмечены жирными красными штрихами).



**Рис. 1.** Схема выбора значений измеряемых дебитов  
**Fig. 1.** Selection scheme of values for the measured debits

Дате планирования может предшествовать остановка скважины для проведения каких-либо работ (например, ГТМ). Поэтому число признаков для ИНС, большее некоторого имеющегося числа, взять не получится. Например, из рис. 1 следует, что выбрать 14 признаков не удастся. Это значит, что в общем случае обучающие выборки будут несопоставимы по объёму, так как для малого числа признаков на входах ИНС (например, 6) будет всегда больше обучающих примеров, чем для большого числа признаков (например, 15).

Если за какой-либо день в БД отсутствует фактическое значение дебита, и это не связано с проведением работ на скважине, то производится линейная интерполяция значения дебита по соседним известным значениям. Использование линейной интерполяции было продиктовано разработчиками БД.

Если между двумя остановками скважины есть только одно измерение дебита, или оно вообще отсутствует, то в таких условиях провести интерполяцию нельзя, и прогноз на этом интервале становится невозможным [14]. Если на скважине ГТМ не проводились, то берётся весь массив значений фактических дебитов для анализа без каких-либо ограничений. Аналогичным образом для анализа используются все значения параметров после последнего проведённого ГТМ на скважине.

Для извлечения данных о дебитах и остановках скважин, в том числе для ГТМ, из БД было создано специальное программное обеспечение (ПО). Данное ПО позволяет провести всю необходимую предварительную обработку данных, выбрать вид дебита (дебит нефти, газа, жидкости (газового конденсата) или воды), число признаков на входе ИНС (от 6 до 15),

указать число блоков для кросс-валидации данных методом K блоков [15], а также создать файл для последующего обучения или тестирования ИНС. Пользователь видит при загрузке данных информацию о характеристиках выборки в текстовом поле. При формировании файла указывается, сколько было задействовано скважин при построении выборки и общее число обучающих и тестовых примеров.

Для проведения экспериментов с помощью этого ПО было построено 40 пар файлов (для обучения и тестирования ИНС) – по 10 пар на каждый вид дебита. Объёмы пар «признаки–метка» для разного числа входов ИНС и для всех дебитов указаны в табл. 2. При использовании метода K блоков выбрано K=5. Это означает, что все вычисленные в ходе исследований итоговые значения WAPE прогноза дебитов являются результатом усреднения по пяти значениям ошибок, полученным в ходе обучения и тестирования ИНС.

**Таблица 2.** Объёмы выборок для различных дебитов и разного числа входов искусственной нейронной сети

**Table 2.** Volumes of selections for various debits and different number of artificial neural network inputs

Дебиты Debits	Число входов ИНС Number of ANN inputs									
	6	7	8	9	10	11	12	13	14	15
$Q_n$ $Q_w$	1248	1248	1246	1246	1244	1244	1243	1243	1243	1235
$Q_g$ $Q_w$	1199	1199	1197	1197	1195	1195	1194	1194	1194	1188
$Q_w$ $Q_g$	1199	1199	1197	1197	1195	1195	1194	1194	1194	1188
$Q_w$ $Q_w$	880	877	875	875	873	873	872	872	872	868

#### Эксперименты по прогнозу значений дебитов скважин

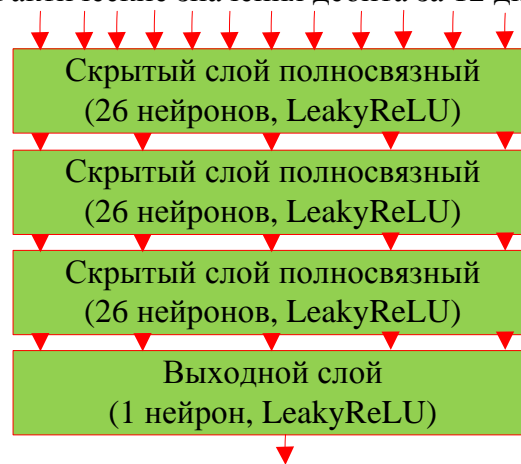
Программная реализация ИНС для проведения экспериментов выполнена на языке C# с использованием библиотеки Microsoft CNTK [16], имеющей определённые преимущества перед другими библиотеками [17, 18]. С её помощью были автоматизированы операции по изменению архитектуры и гиперпараметров ИНС, обучены и протестированы ИНС на вышеуказанных выборках.

Число скрытых слоёв ИНС прямого распространения выбиралось равным 2, 3, 4, 5, 6, 7, 8, 10, 15. В ходе экспериментов было выявлено, что число слоёв больше восьми не даёт прироста точности прогноза. Исследовалось также применение различных функций активации (TanH, LeakyReLU) [19], позволяющих получать связи между нейронами с отрицательными воздействиями. Алгоритм оптимизации при обучении ИНС Adam [20]. Скорость обучения динамически определялась по этому алгоритму. Каждый обучающий пример уникален и должен осуществлять воздействия индивидуально, но скорость обучения тем больше, чем меньше объём минивыборки, поэтому был выбран компромиссный вариант, равный 5. Число эпох выбиралось динамически, в соответствии с достигаемой точностью

обучения, если она несколько эпох существенно не менялась. Число нейронов в каждом скрытом слое выбиралось в соответствии с эвристическим правилом:  $Число\ Нейронов = Число\ Входов * 2 + 2$ .

На входы ИНС подаются фактические значения дебита за определённое число дней (от 6 до 15). Далее идёт ряд скрытых слоёв ИНС – от 2 до 8. В выходном слое ИНС один нейрон (прогнозируемое значение дебита). На рис. 2 в качестве примера приведена архитектура ИНС с 3 скрытыми слоями, 12 фактическими значениями дебита на входах ИНС и 26 нейронами в скрытых слоях, функция активации LeakyReLU.

#### Фактические значения дебита за 12 дней



#### Прогнозное значение дебита

**Рис. 2.** Схема архитектуры искусственной нейронной сети  
**Fig. 2.** Architecture scheme of an artificial neural network

В табл. 3 приведены в качестве примера результаты экспериментов с ИНС различных архитектур при решении на тестовых выборках задачи прогноза значений дебита нефти  $Q_n$ . При этом переменными являются как функции активации и число скрытых слоёв ИНС, так и число входов ИНС. Значение ошибки прогноза вычисляется как среднее по всем ошибкам, полученным по схеме кросс-валидации по 5 блокам, и по всем ошибкам для добывающих скважин в выборках. Жирным отмечается лучшая точность, результаты округлены до тысячных. Аналогичные эксперименты на тестовых выборках были проведены для дебитов газа, жидкости и воды, их результаты сведены в три таблицы.

Такие же четыре таблицы содержат результаты прогноза значений дебитов в случае обучающих выборок. Отметим, что значения WAPE на обучающих выборках меньше до 30 % по сравнению с ошибками, полученными с помощью ИНС на тестовых выборках. Причём чем больше число входов и число слоёв ИНС, тем меньше разница WAPE, полученных на обучающей и тестовой выборках.

Лучшие результаты из всех четырёх таблиц со значениями WAPE при прогнозе дебитов, полученные при исследованиях ИНС с различными архитектурами на тестовых выборках, занесены в табл. 4.

Сравнительный анализ лучших результатов из табл. 4 с WAPE прогноза из табл. 1 показал, что для каждого из дебитов точность прогноза с помощью ИНС выше, чем полученная с помощью метода экстраполяции по скользящей средней. Прирост точности прогноза дебитов с помощью ИНС указан для тестовых выборок

в сравнении с результатами из табл. 1 в последнем столбце табл. 4. Видно, что лучший результат прогноза каждого из дебитов, полученный с использованием ИНС, в два и более раза точнее, чем результат прогноза каждого из дебитов по методике, принятой на предприятии.

**Таблица 3.** Результаты прогноза  $Q_n$  с помощью искусственной нейронной сети на тестовых выборках

**Table 3.** Results of forecasting  $Q_o$  by means of artificial neural network on test selections

Параметр Parameter	Число входов ИНС Number of ANN inputs	Функция активации Activation function	Взвешенная средняя абсолютная ошибка, % Weighted absolute percent error, %							
			Число слоёв Number of layers	2	3	4	5	6	7	8
$Q_n$ $Q_o$	6	TanH	16,705	16,841	17,775	15,163	16,695	17,893	16,705	
		LeakyReLU	20,538	17,978	18,385	16,645	18,146	17,445	20,538	
	7	TanH	18,197	17,723	17,356	17,617	15,420	17,109	18,197	
		LeakyReLU	18,545	17,942	17,563	17,500	17,123	15,777	18,545	
	8	TanH	16,282	17,384	17,141	17,147	20,219	19,081	16,282	
		LeakyReLU	16,993	17,496	17,621	16,837	16,168	17,370	16,993	
	9	TanH	16,551	17,380	17,356	18,837	17,678	15,810	16,551	
		LeakyReLU	17,040	18,290	16,720	18,991	15,858	15,467	17,040	
	10	TanH	14,127	14,674	12,623	13,296	13,985	14,523	13,298	
		LeakyReLU	13,889	13,849	12,477	13,337	13,456	13,351	13,768	
	11	TanH	13,236	12,794	14,579	14,174	13,643	16,728	13,236	
		LeakyReLU	13,718	13,463	13,680	12,685	13,222	13,438	13,718	
	12	TanH	15,568	16,006	15,149	17,024	16,967	16,413	15,568	
		LeakyReLU	15,858	15,660	15,712	15,617	16,456	15,315	15,858	
	13	TanH	16,869	15,179	15,378	16,124	16,000	16,430	16,869	
		LeakyReLU	16,541	15,398	14,602	15,631	14,332	15,240	16,541	
	14	TanH	15,637	15,179	15,962	15,968	16,270	16,099	15,637	
		LeakyReLU	17,494	15,988	14,995	15,611	14,843	15,019	17,494	
	15	TanH	15,360	15,224	14,962	17,200	15,052	15,123	15,360	
		LeakyReLU	15,372	14,731	14,427	14,756	14,050	13,868	15,372	

**Таблица 4.** Лучшие результаты при прогнозе значений дебитов скважин с помощью искусственной нейронной сети

**Table 4.** The best results at forecasting of wells debit values by means of artificial neural network

Параметр Parameter	Взвешенная средняя абсолютная ошибка, % Weighted absolute percent error, %	Особенности архитектуры ИНС Features of ANN architecture	Прирост точности прогноза, % Boost of the forecasting precision, %
$Q_n$ $Q_o$	12,477	10 входов ИНС; 4 скрытых слоя; функция активации LeakyReLU 10 inputs of ANN; 4 hidden layers; activation function LeakyReLU	54,438
$Q_r$ $Q_s$	17,129	11 входов ИНС; 6 скрытых слоёв; функция активации LeakyReLU 11 inputs of ANN; 6 hidden layers; activation function LeakyReLU	51,245
$Q_k$ $Q_f$	13,584	10 входов ИНС; 7 скрытых слоёв; функция активации LeakyReLU 10 inputs of ANN; 7 hidden layers; activation function LeakyReLU	49,188
$Q_b$ $Q_w$	29,513	10 входов ИНС; 7 скрытых слоёв; функция активации LeakyReLU 10 inputs of ANN; 7 hidden layers; activation function LeakyReLU	55,980

#### Заключение

При планировании показателей производства нефтегазодобывающего предприятия важным является прогноз значений дебитов каждой добывающей скважины. Для повышения точности такого прогноза было предложено использовать глубокие искусственные нейронные сети.

Разработана методика подготовки реальных данных по дебитам скважин для обучения и тестирования глубоких искусственных нейронных сетей прямо-

го распространения. Проведены исследования различных архитектур таких искусственных нейронных сетей при решении задач прогноза дебитов нефти, газа, жидкости (газового конденсата) и дебита воды. В результате этих исследований выявлены наиболее эффективные архитектуры искусственных нейронных сетей, позволяющие увеличить точность прогноза в два и более раза по сравнению с точностью прогноза, даваемой традиционным методом экстраполяции по скользящей средней.

## СПИСОК ЛИТЕРАТУРЫ

1. Анкудинов А.А., Полякова Н.С., Радевич Ю.Е. LUKOIL: Мониторинг разработки месторождения с использованием статистических методов анализа на примере перфорационных работ // ROGTEC Field Development, 2019. URL: <https://rogtecmagazine.com/wp-content/uploads/2019/12/LUKOIL-How-Data-Mining-Enhances-Oilfield-Development.pdf> (дата обращения 26.06.2020).
2. Марков Н.Г. Информационно-управляющие системы для газодобывающего производства. – Томск: Изд-во Томского политехнического университета, 2016. – 261 с.
3. Improvement of technological-mathematical model for the medium-term prediction of the work of a gas condensate field / M. Kutia, M. Fyk, O. Kravchenko, S. Palis, I. Fyk // Eastern-European Journal of Enterprise Technologies. – 2016. – V. 83. – P. 40–48.
4. Косков В.Н. Геофизические исследования скважин. – Пермь: Пермский государственный технический университет, 2004. – 122 с.
5. Евсюткин И.В., Марков Н.Г. Управление геологическими мероприятиями на месторождениях нефти и газа с использованием искусственных нейронных сетей // Доклады ТУСУР. – 2020. – Т. 23. – № 1. – С. 62–69.
6. Al-Fattah S.M., Startzman R. Predicting natural gas production using artificial neural network // 2001 SPE Hydrocarbon Economics and Evaluation Symposium, SPE 68593. Dallas, TX, USA. URL: [https://www.researchgate.net/publication/261697475\\_Predicting\\_Natural\\_Gas\\_Production\\_Using\\_Artificial\\_Neural\\_Network](https://www.researchgate.net/publication/261697475_Predicting_Natural_Gas_Production_Using_Artificial_Neural_Network) (дата обращения 26.06.2020).
7. Новый подход к доуточнению прогнозов прокси-моделей пласта с помощью алгоритмов машинного обучения / О.В. Зоткин, М.В. Симонов, А.Е. Осокина, А.М. Андрианова, А.С. Маргарит, В.А. Тимошенко, С.Ф. Лебедев // Нефтяное хозяйство. – 2019. – № 12. – С. 60–63.
8. Khandakar M.R., Louis J. Times-series data augmentation and deep learning for construction equipment activity recognition // Advanced Engineering Informatics. – 2019. – V. 42. – P. 1–12.
9. Раджабов С.Р., Базаров Г.Р. Технологический режим эксплуатации скважин месторождения Зеварды // Молодой ученый. – 2015. – Т. 88. – № 8. – С. 292–294. URL: <https://moluch.ru/archive/88/16556/> (дата обращения: 26.06.2020).
10. Generating moving average trading rules on the oil futures market with genetic algorithms / L. Wang, H. An, X. Xia, X. Liu, X. Sun, X. Huang // Mathematical Modelling and Algorithms in Finance. – 2014. – V. 2014. – 10 p. [Article ID 101808].
11. Вальес-Лосано Дж., Репин-Поляков Н.Н. Применение метода кривой падения Арпса для прогноза дебита скважины // Наука, техника и образование. – 2020. – Т. 69. – № 5. URL: <https://cyberleninka.ru/article/n/primenenie-metoda-krivoy-padeniya-arpsa-dlya-prognoza-debita-skvazhiny> (дата обращения: 26.06.2020).
12. Kofi Sarpong Adu-Manu, John Kingsley Arthur. A review of data cleansing concepts achievable goals and limitations // International Journal of Computer Applications. – 2013. – V. 7. – № 76. – P. 19–22.
13. A survey of forecast error measures / M. Shcherbakov, N.L. Shcherbakova, A. Brebels, A. Tyukov // World Applied Sciences Journal. – 2013. – V. 24. – P. 171–176.
14. Bourke P. Interpolation methods. 1999. URL: <http://paulbourke.net/miscellaneous/interpolation/> (дата обращения: 26.06.2020).
15. Berrar D. Cross-Validation // Encyclopedia of Bioinformatics and Computational Biology. – 2018. – V. 1. – P. 542–545.
16. The Microsoft Cognitive Toolkit. URL: <https://docs.microsoft.com/en-us/cognitive-toolkit/> (дата обращения 26.06.2020).
17. Benchmarking state-of-the-art deep learning software tools / Shaohuai Shi, Qiang Wang, Pengfei Xu, Xiaowen Chu // arXiv:1608.07249v7 [cs. DC] 17 Feb 2017. – URL: <https://arxiv.org/pdf/1608.07249.pdf> (дата обращения: 26.06.2020).
18. Григоренко Е. «Почему CNTK?». URL: <https://habr.com/ru/company/microsoft/blog/336552/> (дата обращения: 26.06.2020).
19. «7 Types of Neural Network Activation Functions: How to Choose?». URL: <https://missinglink.ai/guides/neural-network-concepts/7-types-neural-network-activation-functions-right/> (дата обращения: 26.06.2020).
20. Baptista D., Rodrigues S., Morgado-Dias F. Performance comparison of ANN training algorithms for classification // Conference: Intelligent Signal Processing (WISP). IEEE 8th International Symposium. – Funchal, Portugal, 2013. – P. 115–120.

Поступила 03.11.2020 г.

## Информация об авторах

**Евсюткин И.В.**, ассистент отделения информационных технологий Инженерной школы информационных технологий и робототехники Национального исследовательского Томского политехнического университета.

**Марков Н.Г.**, доктор технических наук, профессор отделения информационных технологий Инженерной школы информационных технологий и робототехники Национального исследовательского Томского политехнического университета.

UDC 622.276:004.896

## DEEP ARTIFICIAL NEURAL NETWORKS FOR FORECASTING DEBIT VALUES FOR PRODUCTION WELLS

Ivan V. Evsyutkin<sup>1</sup>,  
ive1@tpu.ru

Nikolay G. Markov<sup>1</sup>,  
markovng@tpu.ru

<sup>1</sup> National Research Tomsk Polytechnic University,  
30, Lenin avenue, Tomsk, 634050, Russia.

**The relevance** of the research is caused by the need for supporting the decision-making by specialists of the extracting enterprises of the oil and gas industry at production management. Precise forecasting of production wells debits values allows defining such technological operating modes of a well stock and processing equipment which would allow reaching the prescribed production volume. The existing methods do not always provide the demanded precision level at the forecast of wells debits values that leads to mistakes at the calculation of economic effect at a profitability assessment of production wells and the subsequent deliveries of hydrocarbonic raw materials, and also at taking into account the restrictions imposed by environmental supervision of natural resources use.

**The main aim** of the research is to develop and offer the most effective models of deep artificial neural networks at the forecast of production components values for hydrocarbonic raw materials – oil, gas, fluid (gas condensate), and water.

**Objects** of the research are technological parameters of production wells debits of well-stock of oil, gas, oil and gas, and oil-gas condensate fields.

**Methods:** the methods of the analysis of a large amount of technical data of wells developed according to the concept of «Big Data»; models of deep artificial neural networks; object-oriented programming; methods of an assessment and the statistical analysis of research results of deep artificial neural networks efficiency at the forecast of production wells debit values.

**Results.** The technique of data preparation for wells debits is developed for training and testing of feed-forward deep artificial neural networks. Research is carried out for various architectures for such artificial neural networks at the solution of the forecasting task of oil, gas, fluid (gas condensate), and water debits. The most effective architecture of feed-forward deep artificial neural networks is developed. Such neural networks allow increasing the forecasting accuracy in two and more times in comparison with the accuracy of the forecasting received by a traditional method of extrapolation (moving average).

### Key words:

Oil and gas field, production well, forecasting of wells debits values, models of deep artificial neural networks, research of efficiency of feed-forward deep neural networks.

### REFERENCES

- Ankudinov A.A., Polyakova N.S., Radevich Y.E. LUKOIL: How Data Mining Enhances Oilfield Development. *ROGTEC Field Development*, 2019. Available at: <https://rogtecmagazine.com/wp-content/uploads/2019/12/LUKOIL-How-Data-Mining-Enhances-Oilfield-Development.pdf> (assessed 26 June 2020).
- Markov N.G. *Informatsionno-upravlyayushchie sistemy dlya gazodobyvayushchego proizvodstva* [Management information systems for gas production]. Tomsk, Tomsk Polytechnic University Publ. house, 2016. 261 p.
- Kutia M., Fyk M., Kravchenko O., Palis S., Fyk I. Improvement of technological-mathematical model for the medium-term prediction of the work of a gas condensate field. *Eastern-European Journal of Enterprise Technologies*, 2016, vol. 83, pp. 40–48.
- Koskov V.N. *Geofizicheskie issledovaniya skvazhin* [Geophysical surveys of wells]. Perm, Perm State Technical University Publ., 2004. 122 p.
- Evsyutkin I.V., Markov N.G. Upravlenie geologo-tekhnicheskimi meropriyatiyami na mestorozhdeniyakh nefi i gaza s ispolzovaniem iskusstvennykh neyronnykh setey [Management of geological and technical arrangements at oil and gas fields with the use of artificial neural networks]. *Doklady TUSUR*, 2020, vol. 23, no. 1, pp. 62–69.
- Al-Fattah S.M., Startzman R. Predicting natural gas production using artificial neural network. *2001 SPE Hydrocarbon Economics and Evaluation Symposium*, SPE 68593. Dallas, TX, USA. Available at: [https://www.researchgate.net/publication/261697475\\_Predicting\\_Natural\\_Gas\\_Production\\_Using\\_Artificial\\_Neural\\_Network](https://www.researchgate.net/publication/261697475_Predicting_Natural_Gas_Production_Using_Artificial_Neural_Network) (assessed 26 June 2020).
- Zotkin O.V., Simonov M.V., Osokina A.E., Andrianova A.M., Margarit A.S., Timoshenko V.A., Lebedev S.F. Novy podkhod k doutochneniyu prognozov proksi-modeley plasta s pomoshchyu algoritmov mashinnogo obucheniya [New approach to the additional improvement of forecasts of layer proxy models using machine learning algorithms]. *Neftyanoe hoz'yaystvo*, 2019, no. 12, pp. 60–63.
- Khandakar M.R., Louis J. Times-series data augmentation and deep learning for construction equipment activity recognition. *Advanced Engineering Informatics*, 2019, vol. 42, pp. 1–12.
- Radzhabov S.R., Bazarov G.R. Tekhnologicheskii rezhim ekspluatatsii skvazhin mestorozhdeniya Zevardy [Technological mode of the wells' operation of Zevarda field]. *Molodoy Ucheniy*, 2015, vol. 88, no. 8, pp. 292–294. Available at: <https://moluch.ru/archive/88/16556/> (assessed 26 June 2020).
- Wang L., An H., Xia X., Liu X., Sun X., Huang X. Generating moving average trading rules on the oil futures market with genetic algorithms. *Mathematical Modelling and Algorithms in Finance*, 2014, vol. 2014, 10 p. [Article ID 101808].
- Valyes-Losano J., Repin-Polyakov N.N. Primenenie metoda krivoy padeniya Arpsa dlya prognoza debita skvazhiny [Application of the Arps decline curve method for the forecasting of well debit]. *Nauka, tekhnika i obrazovanie*, 2020, vol. 69, no. 5. Available at: <https://cyberleninka.ru/article/n/primenenie-metoda-krivoy-padeniya-arpsa-dlya-prognoza-debita-skvazhiny> (assessed 26 June 2020).
- Kofi Sarpong Adu-Manu, John Kingsley Arthur. A review of data cleansing concepts achievable goals and limitations. *International Journal of Computer Applications*, 2013, vol. 7, no. 76, pp. 19–22.
- Shcherbakov M., Shcherbakova N.L., Brebels A., Tyukov A. A survey of forecast error measures. *World Applied Sciences Journal*, 2013, vol. 24, pp. 171–176.
- Bourke P. *Interpolation methods*. 1999. Available at: <http://paulbourke.net/miscellaneous/interpolation/> (assessed 26 June 2020).

15. Berrar D. Cross-Validation. *Encyclopedia of Bioinformatics and Computational Biology*, 2018, vol. 1, pp. 542–545.
16. *The Microsoft Cognitive Toolkit*. Available at: <https://docs.microsoft.com/en-us/cognitive-toolkit/> (assessed 26 June 2020).
17. Shaohuai Shi, Qiang Wang, Pengfei Xu, Xiaowen Chu. *Benchmarking state-of-the-art deep learning software tools*. arXiv:1608.07249v7 [cs. DC] 17 Feb 2017. Available at: <https://arxiv.org/pdf/1608.07249.pdf> (assessed 26 June 2020).
18. Grigorenko E. «*Pochemu CNTK?*» [«Why CNTK?»]. Available at: <https://habr.com/ru/company/microsoft/blog/336552/> (assessed 26 June 2020).
19. «*7 Types of Neural Network Activation Functions: How to Choose?*». Available at: <https://missinglink.ai/guides/neural-network-concepts/7-types-neural-network-activation-functions-right/> (assessed 26 June 2020).
20. Baptista D., Rodrigues S., Morgado-Dias F. Performance comparison of ANN training algorithms for classification. *Conference: Intelligent Signal Processing (WISP). IEEE 8<sup>th</sup> International Symposium*. Funchal, Portugal, 2013. pp. 115–120.

*Received: 3 November 2020.*

#### **Information about the authors**

**Ivan V. Evsyutkin**, assistant, National Research Tomsk Polytechnic University.

**Nikolay G. Markov**, Dr. Sc., professor, National Research Tomsk Polytechnic University.