

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ КАЗАХСТАН

Казахский национальный исследовательский технический университет
имени К.И.Сатпаева

Институт геологии и нефтегазового дела имени К. Турысова
Кафедра «Нефтяная инженерия»

УДК 004.932.2

На правах рукописи

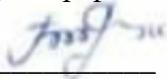
Салиев Нурлан Бакитжанович

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

На соискание академической степени магистра

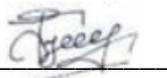
Название диссертации «Использование нейронных сетей при анализе и прогнозировании базовой добычи и ГТМ (геолого-технических мероприятий)»
Направление подготовки 7M07202– Нефтяная инженерия

Научный руководитель
Доктор философии PhD,

 Жаппасбаев Б.Ж.

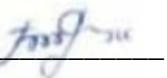
"__" _____ 2021 г.

Рецензент
Доктор философии PhD,
Младший научный сотрудник
Лаборатории

 Гусенов И.Ш.

"__" _____ 2021 г.

Нормоконтроль
PhD, ассистент-профессор кафедры
«Нефтяная инженерия»

 Жаппасбаев Б.Ж.

"__" _____ 2021 г.

ДОПУЩЕН К ЗАЩИТЕ

Заведующий кафедрой
«Нефтяная инженерия»

_____ м.т.н, Дайров Ж.

"__" _____ 2021 г.

Алматы 2021

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ

КАЗАХСТАН

Казахский национальный исследовательский технический
университет имени К.И.Сатпаева

Институт геологии и нефтегазового дела имени К. Турысова
Кафедра «Нефтяная инженерия»

7M07202– Нефтяная инженерия

УТВЕРЖДАЮ
Заведующий кафедрой
«Нефтяная инженерия»

_____ м.т.н, Дайров Ж.К.

“ _ ” _____ 2021 г.

ЗАДАНИЕ

на выполнение магистерской диссертации

Магистранту Салиеву Нурлану Бакитжановичу

Тема: Использование нейронных сетей при анализе и прогнозировании базовой добычи и ГТМ (геолого-технических мероприятий)

Утверждена приказом Ректора Университета №330–м от "11" 11 2021г.

Срок сдачи законченной диссертации "___" июня 2021 г.

Исходные данные к магистерской диссертации: _____

Перечень подлежащих разработке в магистерской диссертации вопросов:

а) Построение архитектуры нейронных сетей для построения моделей работы скважин, оценке влияния различных факторов, а также предсказанию результатов б)

Моделирование физико-химических свойств пластовой нефти горизонтов месторождения

в) Проведен анализ имеющихся наработок мирового научного сообщества применительно к теме работы добывающих и нагнетающих скважин

Перечень графического материала (с точным указанием обязательных чертежей):

Рекомендуемая основная литература:

- 1) Zeng, B., Yang, H., Xu, R., Zhou, W., & Han, X. (2019). Lcf: A local context focus mechanism for aspect-based sentiment classification. Applied Sciences, 9(16), 3389.
- 2) Tseng, B. H., Kreyszig, F., Budzianowski, P., Casanueva, I., Wu, Y. C., Ultes, S., & Gasic, M. (2018). Variational cross-domain natural language generation for spoken dialogue systems. arXiv preprint arXiv:1812.08879.
- 3) Yan, X., Rastogi, A., Villegas, R., Sunkavalli, K., Shechtman, E., Hadap, S., ... & Lee, H. (2018). Mt-vae: Learning motion transformations to generate multimodal human dynamics. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 265-281).

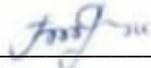
ГРАФИК
подготовки магистерской диссертации

Наименование разделов, перечень разрабатываемых вопросов	Сроки представления научному руководителю	Примечание
Сбор и изучение научной литературы по теме	май 2020	
Анализ данных	декабрь 2020	
Построение архитектуры нейронных сетей для построения моделей работы скважин	май 2021	

Подписи

консультантов и нормоконтролера на законченную магистерскую диссертацию с указанием относящихся к ним разделов диссертации

Наименования разделов	Консультанты, И.О.Ф. (уч. степень, звание)	Дата подписания	Подпись
Экономическая часть			
Охрана труда			
Нормоконтролер			

Научный руководитель  Жаппасбаев Б.Ж.
(подпись)

Задание принял к исполнению обучающийся  Салиев Н.Б.
(подпись)

Дата " _____ " _____ 20__ г.

ОТЗЫВ

на магистерскую диссертацию

Салиева Нурлан Бакитжановича

по специальности 7М07202 – «Нефтяная инженерия»

на тему «Использование нейронных сетей при анализе и прогнозировании базовой добычи и ГТМ (геолого-технических мероприятий)».

В данной работе представлено использование нейронных сетей в нефтегазовой промышленности, в последнее время с увеличением производительности компьютеров актуализировался вопрос обработки больших данных и применение нейронных сетей в научно-исследовательских работах получило широкое распространение. В сфере нефтегазовой отрасли на текущий момент проведено недостаточное количество исследований с использованием машинного обучения и нейронных сетей, что говорит об ее актуальности.

Актуальность применения данного подхода вызвана тем, что в пластовых условиях очень много неопределенностей начиная с геологических обстановок осадконакопления и движения флюидов в данной среде и остро стоит вопрос прогнозирования добычи.

В результате проведенных работ подготовлены начальные модели нейронных сетей для проекта с различными слоями регуляризации и нормализации данных. Произведена попытка добавления слоев «умной» чистки данных от шума, где по результатам можно сказать, что лучше всего себя показали сверточные рекуррентные многослойные модели.

По результатам изучения опыта обработки аналогичных данных в других компаниях, таких, как Газпром, Schlumberger, IBM, Total, Шелл и опыта других коллективов со всего мира в смежных направлениях, выявлена необходимость более точного прогнозирования при добыче УВ сырья.

Данная работа проведена в целях поиска и применимости новейших подходов в анализе данных, где следующими этапами требуется продолжить исследования в нахождении путей решения задач с помощью интеграции новых методов, работающих с различными данными как пример: пороговые и дискретные значения.

Необходимо отметить большой интерес крупных НГ компаний к данному направлению отмечается попытками ускорения развития направления, где зачастую они являются спонсорами и инициаторами при проведении различных конкурсов по всему миру для анализа данных нефтегазового сектора с применением новейших подходов обработки.

В период исследования по теме данной магистерской диссертации:

- опубликована статья в Научно-практическом журнале «Вестник нефтегазовой отрасли Казахстана» №2 (3) 2020, на тему «ПРИМЕНЕНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ В ЛИТОЛОГИЧЕСКОМ ОПИСАНИИ КЕРНА».

В целом работа выполнена на достаточном уровне и является законченной, заслуживает высокую оценку, а ее автор Салиев Н.Б. достоин присвоения магистра по специальности 7М07202 – «Нефтяная инженерия».

Научный руководитель
Доктор философии PhD



Жаппасбаев Б.Ж.

РЕЦЕНЗИЯ

**на диссертацию Салиева Н.Б.
на тему «Использование нейронных сетей при анализе и прогнозировании
базовой добычи и ГТМ (геолого-технических мероприятий)
представленную на соискание степени магистра
по специальности 7М07202 – Нефтяная инженерия**

На магистерскую диссертацию Салиева Н.Б.

Рецензируемая работа посвящена использованию нейронных сетей в нефтегазовой промышленности, в последнее время с увеличением производительности компьютеров актуализировался вопрос обработки больших данных и применение нейронных сетей в научно-исследовательских работах получило широкое распространение. В сфере нефтегазовой отрасли на текущий момент проведено недостаточное количество исследований с использованием машинного обучения и нейронных сетей, что говорит об ее актуальности. Актуальность применения данного подхода вызвана тем, что в пластовых условиях очень много неопределенностей начиная с геологических обстановок осадконакопления и движения флюидов в данной среде и остро стоит вопрос прогнозирования добычи.

Данная охватывает направление Data Science:

В рамках работы в подготовлены и представлены следующие основные результаты:

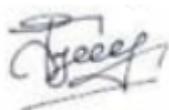
- Проведен анализ имеющихся данных показателей скважин различными статистическими методами с целью выявления зависимостей и влияния различных факторов на работу скважин и друг на друга;
- Выполнено глубокое ознакомление с данными и их особенностями;
- Проведен анализ имеющихся наработок мирового научного сообщества применительно к теме работы добывающих и нагнетающих скважин;
- Построены архитектуры нейронных сетей для построения моделей работы скважин, оценке влияния различных факторов, а также предсказанию результатов;
- Проверены гипотезы о влиянии различных факторов;
- Произведена оценка направлений для будущих исследований.

Проделанная работа выполнена на высоком уровне и заслуживает высокую оценку. Выбранный подход заслуживает одобрения, т.к. соответствует имеющимся мировым практикам и методологиям.

Из недостатков следует отметить усложненный метод подачи материала, требующий разъяснений узких специалистов, а также замечания по оформлению графического материала.

Магистерская диссертация по теме: «Использование нейронных сетей при анализе и прогнозировании базовой добычи и ГТМ (геолого-технических мероприятий)» имеет научно-практическую значимость и рекомендуется к защите.

Рецензент



Гусенов И.Ш.

Доктор PhD, младший научный сотрудник лабораторий

Протокол анализа Отчета подобия

заведующего кафедрой / начальника структурного подразделения

Заведующий кафедрой / начальник структурного подразделения заявляет, что ознакомился(-ась) с Полным отчетом подобия, который был сгенерирован Системой выявления и предотвращения плагиата в отношении работы:

Автор: Салиев Нурлан

Название: Использование нейронных сетей при анализе и прогнозировании базовой добычи и ГТМ (геолого-технических мероприятий)

Координатор: Биржан Жаппасбаев

Коэффициент подобия 1:0.2

Коэффициент подобия 2:0

Замена букв:0

Интервалы:0

Микропробелы:0

Белые знаки:0

После анализа отчета подобия заведующий кафедрой / начальник структурного подразделения констатирует следующее:

- обнаруженные в работе заимствования являются добросовестными и не обладают признаками плагиата. В связи с чем, работа признается самостоятельной и допускается к защите;
- обнаруженные в работе заимствования не обладают признаками плагиата, но их чрезмерное количество вызывает сомнения в отношении ценности работы по существу и отсутствием самостоятельности ее автора. В связи с чем, работа должна быть вновь отредактирована с целью ограничения заимствований;
- обнаруженные в работе заимствования являются недобросовестными и обладают признаками плагиата, или в ней содержатся преднамеренные искажения текста, указывающие на попытки сокрытия недобросовестных заимствований. В связи с чем, работа не допускается к защите.

Обоснование:

.....
.....
.....
.....
.....
.....
.....
.....

Дата

Подпись заведующего кафедрой /

начальника структурного подразделения

Окончательное решение в отношении допуска к защите, включая обоснование:

.....
.....
.....
.....
.....
.....
.....

Дата

Подпись заведующего кафедрой /

начальника структурного подразделения

Протокол анализа Отчета подобия Научным руководителем

Заявляю, что я ознакомился(-ась) с Полным отчетом подобия, который был сгенерирован Системой выявления и предотвращения плагиата в отношении работы:

Автор: Салиев Нурлан

Название: Использование нейронных сетей при анализе и прогнозировании базовой добычи и ГТМ (геолого-технических мероприятий)

Координатор: Биржан Жаппасбаев

Коэффициент подобия 1: 0.2

Коэффициент подобия 2: 0

Замена букв: 0

Интервалы: 0

Микропробелы: 0

Белые знаки: 0

После анализа Отчета подобия констатирую следующее:

- обнаруженные в работе заимствования являются добросовестными и не обладают признаками плагиата. В связи с чем, признаю работу самостоятельной и допускаю ее к защите;
- обнаруженные в работе заимствования не обладают признаками плагиата, но их чрезмерное количество вызывает сомнения в отношении ценности работы по существу и отсутствием самостоятельности ее автора. В связи с чем, работа должна быть вновь отредактирована с целью ограничения заимствований;
- обнаруженные в работе заимствования являются недобросовестными и обладают признаками плагиата, или в ней содержатся преднамеренные искажения текста, указывающие на попытки сокрытия недобросовестных заимствований. В связи с чем, не допускаю работу к защите.

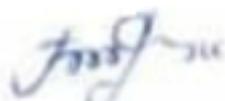
Обоснование:

.....

.....

.....

Дата



Жаппасбаев Б.Ж.

Подпись Научного руководителя

АҢДАТПА

Бүгінгі күні компьютерлердің өнімділігінің артуымен үлкен деректерді өңдеу және ғылыми-зерттеу жұмыстарында нейрондық желілерді қолдану өзекті мәселе болып табылады. Қазіргі уақытта мұнай-газ саласында машиналық оқыту мен нейрондық желілерді пайдалану арқылы зерттеулер жеткіліксіз жүргізілуде. Жұмыста нақты уақыт режимінде көмірсутектерді өндіруді талдау және болжау үшін нейрондық желілердің математикасына негізделген модельді құру қарастырылады.

АННОТАЦИЯ

В последнее время с увеличением производительности компьютеров актуализировался вопрос обработки больших данных и применение нейронных сетей в научно-исследовательских работах. В сфере нефтегазовой отрасли на текущий момент проведено недостаточное количество исследований с использованием машинного обучения и нейронных сетей. В работе рассматривается построение модели, основанной на математике нейронных сетей для анализа и прогнозирования добычи углеводородов в режиме реального времени.

ABSTRACT

Recently, with the increase in computer performance, the issue of big data processing and the use of neural networks in research has become more relevant. At the moment, insufficient number of studies using machine learning and neural networks have been conducted in the oil and gas industry. The paper considers the construction of a model based on the mathematics of neural networks for the analysis and forecasting of hydrocarbon production in real time.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	1
Глава 1. Исходные данные и статистический анализ данных	2
1.1 Исходные данные	2
1.2 Методология и статистический анализ данных	4
Глава 2. Общая концепция и сложности в данных.....	8
2.1 Разработка имитационной модели (цифрового двойника) объекта разработки	8
2.2 Возможные конфигурации вариантов моделей ML (Machine Learning)	19
2.3 Сложности в данных – затруднения качественного обучения моделей	20
Глава 3. Конфигурации и структуры моделей.....	22
3.1 Результаты тестирования моделей и сопоставление полученных результатов по плоским моделям МЛ.....	22
3.2 Результаты тестирования моделей и сопоставление полученных результатов по моделям ИНС.....	25
ЗАКЛЮЧЕНИЕ	27
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ.....	30

ВВЕДЕНИЕ

Актуальность

Актуальность применения данного подхода вызвана тем, что в пластовых условиях очень много неопределенностей начиная с геологических обстановок осадконакопления и движения флюидов в данной среде и остро стоит вопрос прогнозирования добычи. На текущий момент существует достаточно большое количество подходов к прогнозированию добычи как базовой, так и при проведении ГТМ, но суть данной работы будет заключена в том, что построенная модель сможет адаптироваться и учиться с входом новых данных и практически в режиме реального времени сможет корректировать процесс прогнозирования, как на уровне скважин так и на уровне месторождений.

Цель

Построение модели, основанной на математике нейронных сетей для анализа и прогнозирования добычи УВ в режиме реального времени.

Основные задачи работы

1. Создание поэтапного метода построения модели, корректно прогнозирующей добычу УВ.
2. Подбор наилучшего алгоритма нейронных сетей именно для условий месторождения.
3. Подбор индивидуальных подходов при моделировании процесса прогнозирования.

Научная новизна

Большое количество работ о применении ИНС в нефтяной литературе было рассмотрено и обобщено в таблицах. Заявки были разделены на четыре группы: применение ИНС в разведке, бурении, добыче и разработке пластов. Было сведено в таблицу большое количество приложений в литературе по нефтяной инженерии. Кроме того, была представлена формализованная методология применения ИНС для любого нефтяного применения, которая была реализована с помощью блок-схемы, которая может служить практическим руководством для применения ИНС для любого нефтяного применения. Метод был разбит на этапы, которым можно легко следовать. Наличие огромных массивов данных в нефтяной промышленности дает возможность использовать эти данные для принятия более эффективных решений и прогнозирования будущих результатов. В этом документе будет представлен обзор применения ИНС в нефтяной промышленности, а также четкая методология применения ИНС для любого применения в нефтяной промышленности.

Глава 1. Исходные данные и статистический анализ данных

1.1 Исходные данные

Искусственная нейронная сеть (ИНС) - это вычислительная модель, которая имитирует структуру и функциональные аспекты биологических нейронных сетей. Нейронная сеть имеет организацию, аналогичную организации человеческого мозга, и представляет собой сеть, состоящую из обрабатывающих элементов, называемых нейронами.

Нейроны получают данные от окружающих нейронов, выполняют некоторые вычисления, передают результаты другим нейронам. Связи между нейронами имеют вес, связанный с ними. Каждое соединение может передавать сигнал от одного искусственного нейрона к другому. Искусственный нейрон, который получает сигнал может обработать его, а затем подать сигнал дополнительным искусственным нейронам, подключенным к нему. Информация, проходящая через сеть, влияет на структуру ИНС, потому что нейронная сеть изменяется - или, в некотором смысле, обучается - на основе этих входных и выходных данных. ИНС считаются инструментами моделирования нелинейных статистических данных, в которых моделируются сложные взаимосвязи между входными и выходными данными или обнаруживаются закономерности (Хейкин, 1999). Нейронные сети могут использоваться для моделирования взаимосвязей между

входными и выходными данными, поиска закономерностей в данных, прогнозирования данных и т.д.

Нейронная сеть была разработана с использованием метода обратного решения для формулировки данных обучения и тестирования. Нормализация данных упростила работу нейронной сети, повысила ее эффективность и повысила производительность.

Точное предсказание и прогноз добычи нефти всегда были одной из основных областей исследований. Основной целью данной работы является разработка нейросетевой модели для прогнозирования добычи нефти.

Данные представлены

- Для Объекта-1 месторождения А (А-1, А-2)
- 78 добывающих и 12 нагнетательных скважин
- Северный и Южный купола

Структура данных

- Дискретность данных - суточная
 - Охватывает период с 22-05-2017 по 15-05-2020
- Добывающие скважины:
 - Объем добычи жидкости, объем добычи нефти, объем добычи газа
 - WaterCut
 - Bhp_Pump_Press, Bhp_fluid_level
 - Способ эксплуатации

- ГТМ (вид)
- Нагнетательные скважины:
 - Объем закачки воды
 - Тип воды
 - ГТМ

1.2 Методология и статистический анализ данных

ИНС-это мощная техника моделирования, способная решать определенные проблемы. У него есть возможность учиться или проходить обучение о конкретной задаче, ее вычислительных возможностях и способности формулировать абстракции и обобщения. ИНС также обладают способностью изучать и моделировать нелинейные и сложные взаимосвязи. ИНС может извлекать уроки из структуры обучающего набора данных и прогнозировать будущие данные.

Обучение можно классифицировать следующим образом (Хейкин, 2009).:

- a. контролируемое обучение, которое требует наличия целевого или желаемого ответа для реализации конкретного сопоставления ввода–вывода;
- b. неконтролируемое обучение, реализация которого полагается на предоставление независимой от задачи меры качества представления, которую сеть должна изучить самоорганизованным образом;
- c. обучение с подкреплением, при котором сопоставление ввода–вывода выполняется посредством непрерывного взаимодействия обучающей системы с ее средой, чтобы свести к минимуму скалярный показатель производительности.

Было проведено много исследований по разработке соответствующих рекомендаций при выборе сетевых архитектур, определении предварительной обработки данных, алгоритма обучения и других критериев эффективности. Эти параметры выбора будут фактически повлиять на представление нейронной сети и повысить производительность. Важно применять системный подход при разработке моделей ИНС с учетом таких факторов, как предварительная обработка данных, определение адекватных входных данных модели и подходящей сетевой архитектуры, оценка параметров и проверка модели.

Оценка результатов обучения и предсказаний

После обучения на тестовой выборке и достижения хороших показателей предсказаний, в сеть подаются основные (рабочие) данные и оцениваются результаты. Оценка результатов ведется при помощи анализа статистических величин: R^2 (Р-квадрат), Среднеквадратическое отклонение, Ошибки среднеквадратического действующего значения, Средняя абсолютная ошибка и других.

Статистический анализ данных

Стратегия применения LSTM моделей предполагает проверку данных на вариации, ковариации векторов, гетероскедастичность, шумы, выбросы, коллинеарность и другие тесты.

Одной из главных причин такого анализа является уязвимость Deep Learning к разного рода статистическим особенностям. Альтернативные алгоритмы поиска такие как деревья имеют свои преимущества и недостатки.

Изначально требовалось определить принадлежность к распределению затем произвести попытку найти способ перевести распознанное распределение к нормальному виду.

При помощи QQ-тестов проводится анализ принадлежности к нормальному распределению. График на рисунке (см. Рис. 1) не показывает принадлежность выборки к какому-либо статистическому распределению. Красная линия на графике - это теоретическое ожидаемое распределение данных для нормального распределения. Синие точки — это действительные значения величин, которые распределены аномально (не вблизи красной линии). В качестве анализируемой величины использовалась: Liquid_volume (объем добываемой жидкости, множество скважин).

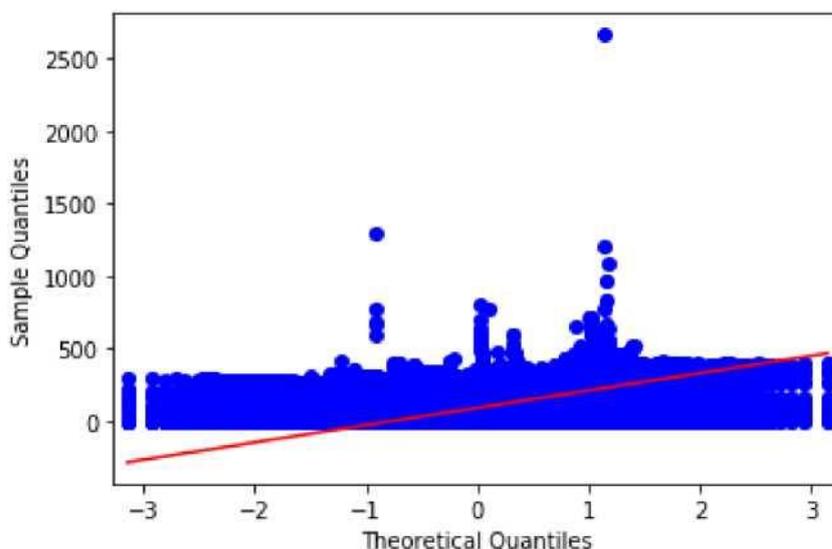


Рисунок 1. QQ тест для Liquid volume, множество скважин

Подобное наблюдалось для всех входных переменных (ВНР_pump_press, Water_volume). В ходе исследований не удалось установить принадлежность к распределениям через ряд тестов. В дальнейшем предполагалось применить тесты

Андерсона-Дарлинга, Колмогорова-Смирнова, графики корреляций: РРСС, РС и тд. Для чего потребуется дальнейший анализ.

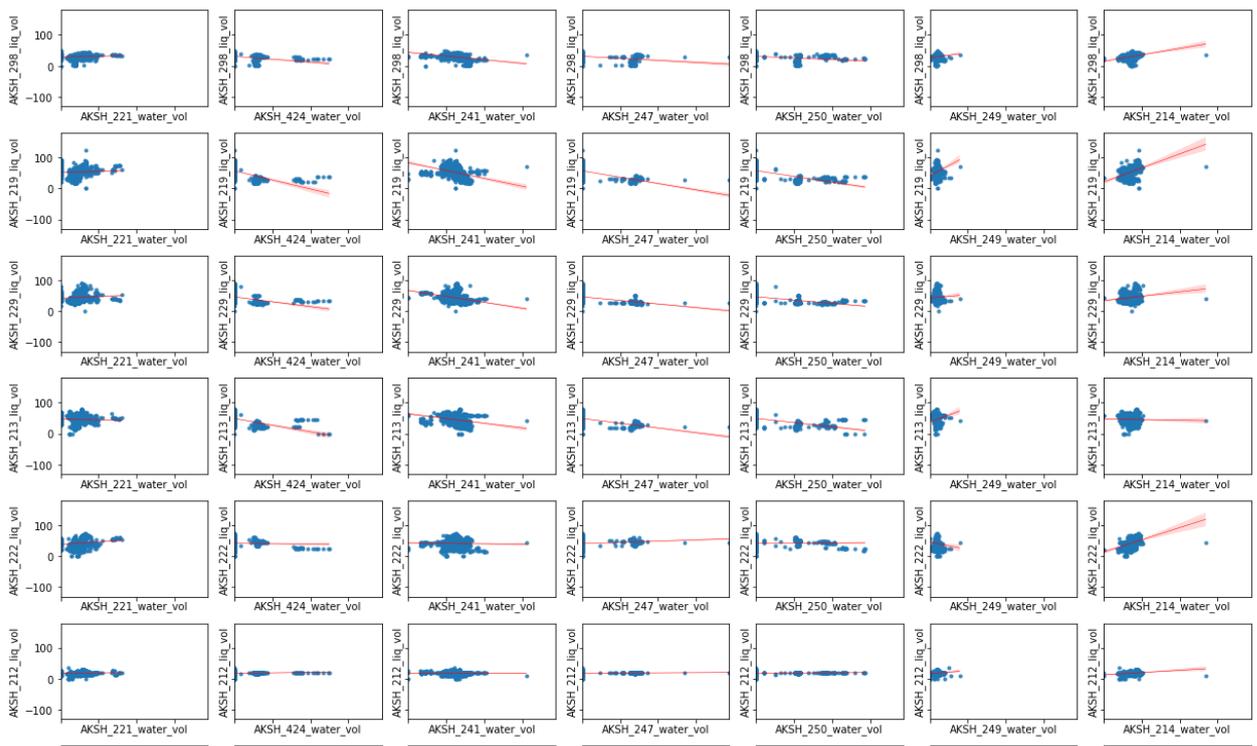


Рисунок 2. QQ тест для Liquid volume, множество скважин, данные подавались с лагом

Анализ парных корреляций так же не выявил зависимостей (см. Рис 2) показывает, что для скважин не выявлено зависимостей между параметрами Liquid_volume и Water_volume (с лагом 15 дней), видны выбросы, данные anomalously распределены вокруг некоторого значения.

Однако оценки точности предсказания Liquid_volume (поданные без лага, без пропусков временного ряда) на различных моделях машинного обучения, в частности деревьях, показали хорошие результаты, представлены в Табл. 1

Таблица 1. Данные по моделям машинного обучения для параметра Liquid volume

Метод	Точность (scores)
Random Forest	91 % (test)
Ada Boost	94 % (test)
XGB	69 % (test)

Простая OLS модель подтвердила гипотезу отрицательного влияния переменных (Water_volume на уровень Liquid_volume) отдельной скважины (liq_vol выбранной как зависимая переменная), что говорит невозможности определения зависимостей между отдельными скважинами. За исключением двух скважин: A_221 и A_225, что видно по параметрам $P > |T|$, что подтверждает гипотезу о взаимовлиянии этих двух скважин.

Низкие значения R2-adjusted (0.063) и R2 (0.074) говорят об отсутствии влияния независимых переменных (water_volume) на зависимую переменную

(liquid_volume), либо же о нелинейной природе связей между переменными.

Omnibus и Jarque-Bera имеют так-же аномально высокие значения, что говорит об отсутствии нормальности в распределении.

Статистика Durbin-Watson подтверждает неоднородность и наличие автокорреляции остатков.

Данные требуют обработки, такой как нивелирования влияния случайных процессов, таких как остановки-пуски скважин, проведение ГТМ, а также аппроксимация к плавным градиентам изменений физических параметров резервуара и скважин.

Глава 2. Общая концепция и сложности в данных

2.1 Разработка имитационной модели (цифрового двойника) объекта разработки

Разработка имитационной модели (цифрового двойника) объекта разработки:

- Загрузка и обработка входных данных
- Обучение ML-модели(-ей)
- Обучение «физической» модели(-ей)
- Создание эффективного ансамбля моделей для генерации прогнозов

В ходе работы были проверены ряд видов моделей нейронных сетей, которые условно можно разделить на группы, представленные в таблице (Табл. 2):

Таблица 2. Модели нейронных сетей и характеристики

Модель	Тип
LSTM-Repeater-LSTM/LSTM-Repeater-LSTM	Многослойная, многошаговая, многофакторная
ConvLstm2D(8-9 слоёв)	Сверточная, рекуррентная, многошаговая
MLP-MLP	Многослойный перцептрон, многошаговая
MLP-LSTM	Комбинированная рекуррентная, многошаговая,
LSTM-Repeater-LSTM	Многослойная, многошаговая, последовательная
ConvLstm2D(8-9)-ConvLstm2D (8-9 слоёв)	Сверточная, рекуррентная, многошаговая, многопрогнозная

Наилучшие результаты точности были получены у LSTM-Repeater, ConvLSTM2D и многоядерных сверточных рекуррентных сетей последовательных архитектур. Архитектура такой сети представлена на рисунке ниже.

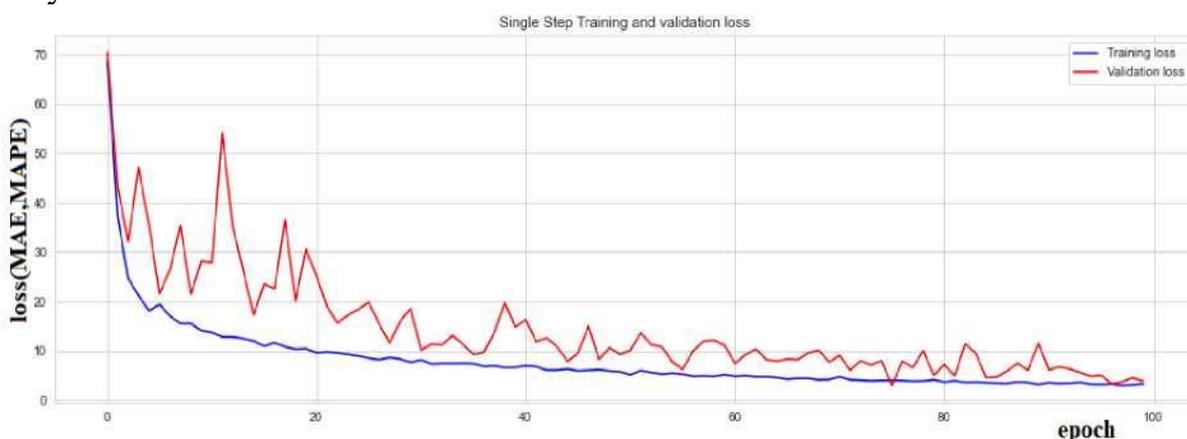


Рисунок 3. Процесс обучения LSTM модели на 40 эпохах

Процесс обучения LSTM сети на разном количестве эпох (интервалов). Видно (см. Рис. 3), что сходимость результатов прогнозных и фактических увеличивается с увеличением числа эпох. Входные данные: ВНР_pump_press, Water_volume для множества скважин.

Результаты предсказаний

Результаты предсказаний моделей представлены в сводной таблице (Табл. 3).

Таблица 3. Сводная таблица результатов предсказаний моделей

Модель	R2	RMSE	MAE	MAPE
LSTM-Repeater-LSTM/LSTM-Repeater-LSTM	-1.24	6.69 (точность 87%)	5.20 (точность 88%)	0.14 (точность 86%)
ConvLstm2D(8-9)-ConvLstm2D(8-9слоёв)	-64.70	36.17 (точность 11 %)	1308.59 (точность 0%)	0.97 (точность 2%)
ConvLstm2D(8-9)	-62.66	35.60 (точность 10%)	1267.98 (точность 0%)	0.96 (точность 2%)
LSTM-Repeater-LSTM	-0.00498	4.47 (точность 90%)	3.43 (точность 92%)	0.097 (точность 99,9%)
MLP-MLP	-8.13	13.49 (точность 76%)	12.42 (точность 78%)	0.328 (точность 68%)
LSTM-MLP	-4.025	10.05 (точность 81%)	9.06 (точность 82%)	0.23 (точность 77%)

Аномальные значения «Коэффициента детерминации (R2)» потребовали проверки результатов моделей при помощи других методов, которые позволили лучше оценить показатели точности предсказаний. Стоит отметить высокую точность моделей LSTM Repeater LSTM, MLP-MLP и LSTM-MLP.

Данные оценки получены для тестовой выборки по данным скважины А_211, за период 2018-2020 года (тренировочная выборка 2018-2019 год, прогнозная - 2020 год), проводилась нормализация данных (приведение шкалы), проводилась регуляризация данных (пенализация шумов), Гауссова регуляризация.

Входные данные: Забойное давление, Объем закачиваемой воды. Данные по объему закачиваемой воды подавались с запаздыванием в 15 дней.

Прогнозируемые: Объем добываемой жидкости.

На рисунке 4 представлена архитектура модели MLP-MLP.

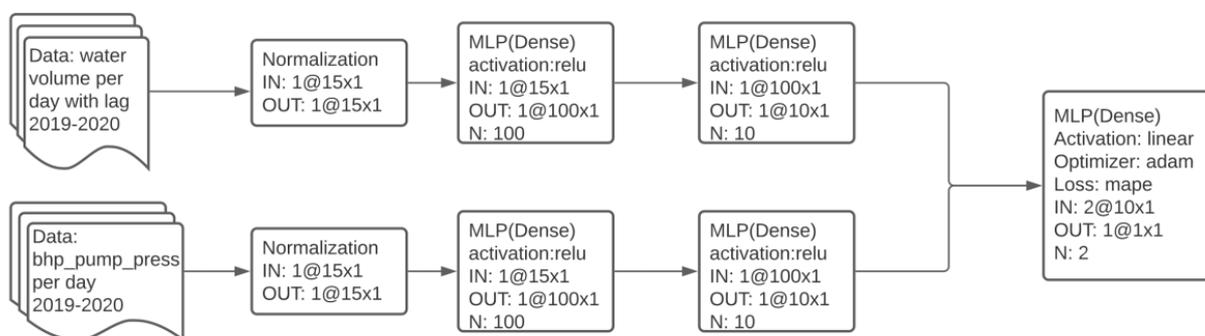


Рисунок 4. архитектура модели MLP-MLP

Описание работы модели MLP-MLP:

- Данные подаются на вход одной ветви с лагом, длина выборок 365 дней (шагов), ширина по 5 векторов), water_volume (с 5-ти скважин) за 2019-2020 годы;
- На вход второй ветви без лага, длина выборок 365 дней (шагов), ширина по 5 векторов bhp_pump_press (5 скважин) за 2019-2020 годы;
- Далее оба слоя работают одинаково со своими наборами данных.

Описание работы одного слоя: данные попадают на слой нормализации (централизация данных) тензором 1 @ 15x 1, затем на слой MLP с активационной функцией Relu в виде тензора 1 @15x1, далее на слой MLP с активацией Relu в виде тензором 1 @100x1. Результирующий тензор 1 @10x1.

После работы каждой ветви данные совместно подаются на слой полносвязной активации Linear с оптимизатором Adam и Loss-функцией MAPE. В результате получается тензор 1@1x1 предсказаний уровня добычи жидкости в день.

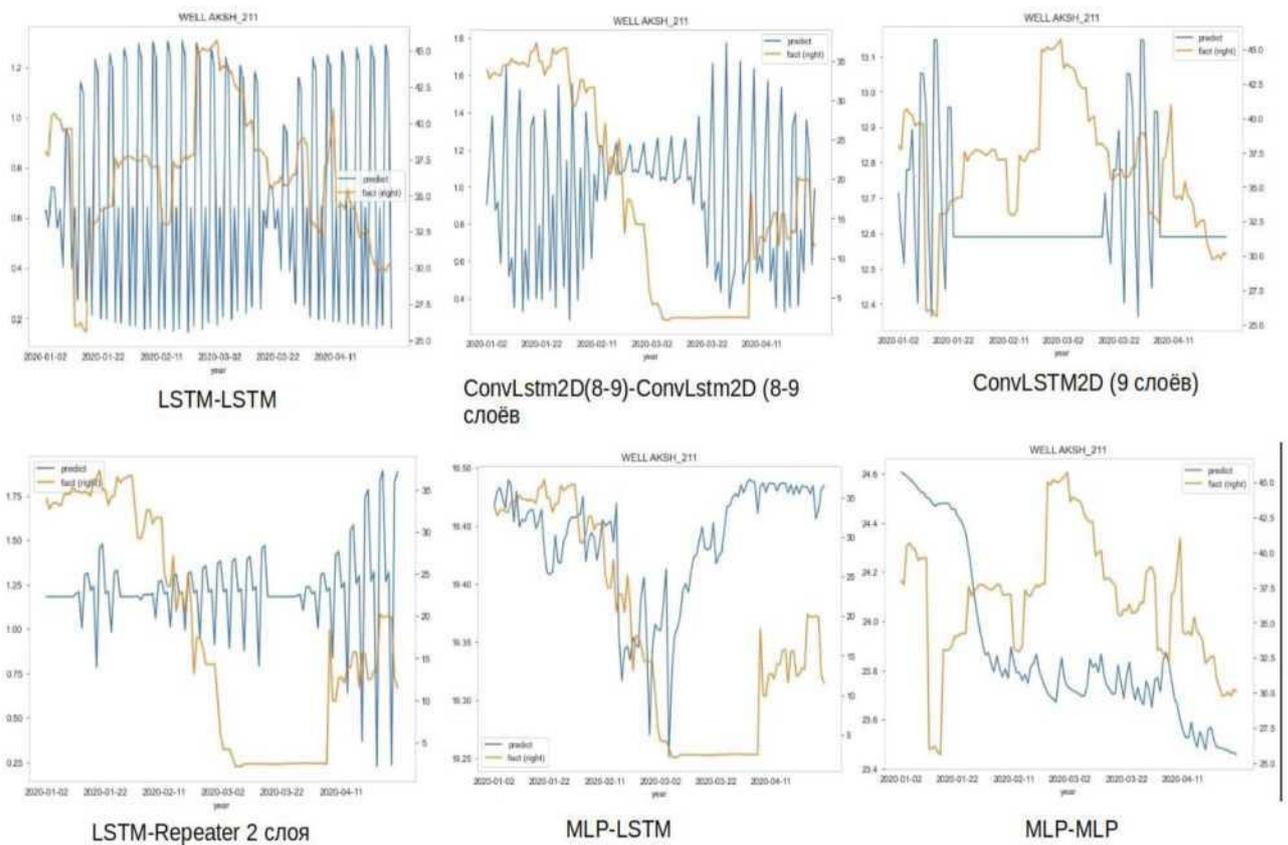


Рисунок 5. Данные прогнозов архитектур

Рисунок 5 показывает графики прогнозов для 6-ти архитектур. Входные данные: Забойное давление, Объем закачиваемой воды. Прогнозируемые: Объем добываемой жидкости (желтая линия — факт, синяя — предсказания). Данные подавались с запаздыванием в 15 дней (Объем закачиваемой воды). Видно, что модели LSTM-Repeater-LSTM, MLP-MLP и MLP-LSTM демонстрируют неплохую точность предсказаний, согласующуюся с показанными ранее результатами. Скважина A_211 была выбрана для анализа из группы скважин Северного Купола, как скважина с наибольшей полнотой данных (меньше всего пропусков) и хорошим показателем R2.

Было выдвинуто предположение о необходимости интеграции Conv-LSTM2D модулей для обработки сложных участков с пропусками и выбросами. Более сложная 4-х уровневая Conv-LSTM2D сеть позволила улучшить прогнозы на некоторых участках. Возможно, благодаря свёртке признаков участков с выбросами, модель лучше обучалась, получая больше точной информации о последовательностях данных. Однако, сложность прогнозирования на длинные многошаговые периоды и необходимость дальнейшего исследования данных на данном этапе делают сложным создание качественного прогноза для произвольных параметров.

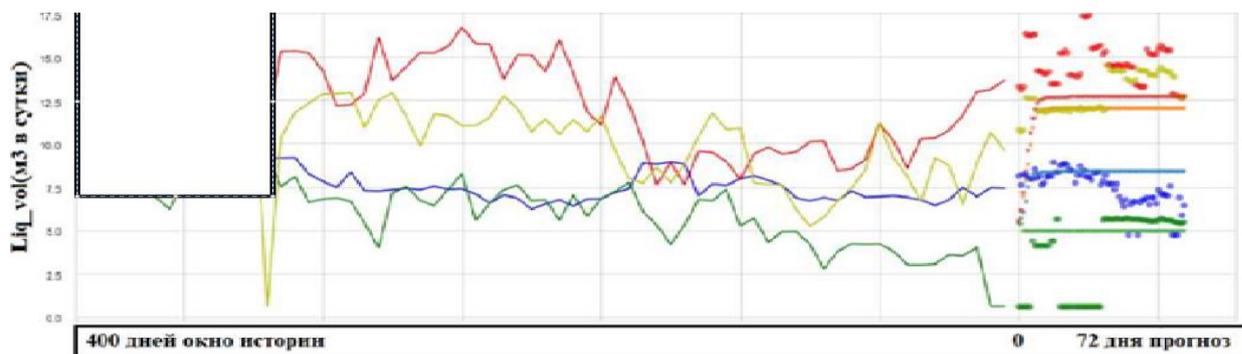


Рисунок 6. Применение сложного 4х уровневго ядра Conv-LSTM2D (multistep) строящего пространственно временное влияние переменных многомерного ряда данных multivariate output. 4 прогнозируемые переменные, 4 water_vol на вход, 4 liq_vol на выход (statefull)

Было замечено, что увеличение количества Conv_LSTM2D слоёв приводило к росту обучаемости модели. То есть гипотеза о необходимости свёртки сложных участков (скачки, резкие градиенты) временных диаграмм подтвердилась. Это представлено на рисунке (см. Рис. 7), где модель Conv_LSTM2D имеет уже 8 слоев и 12 переменных на вход.

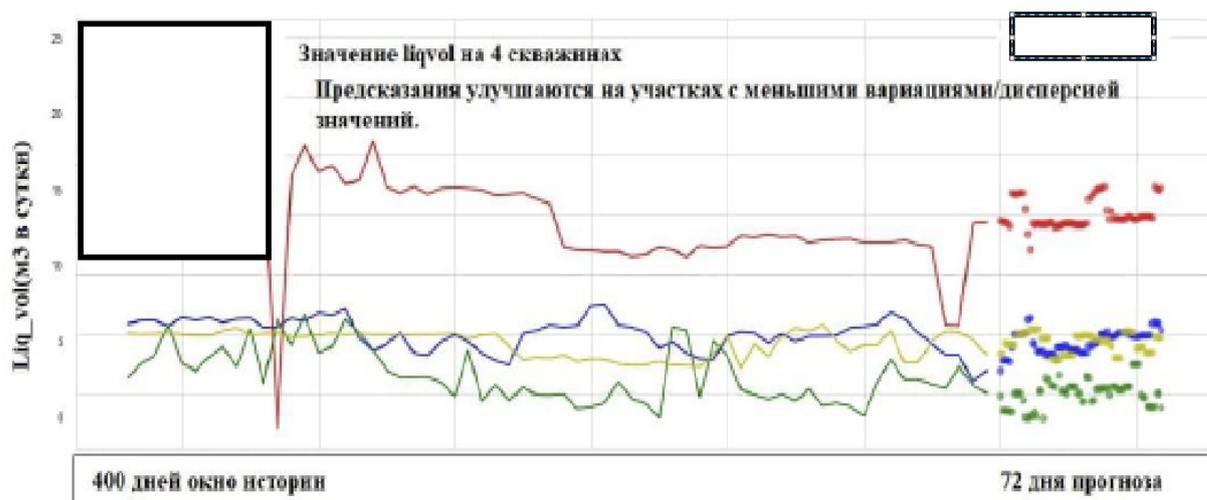
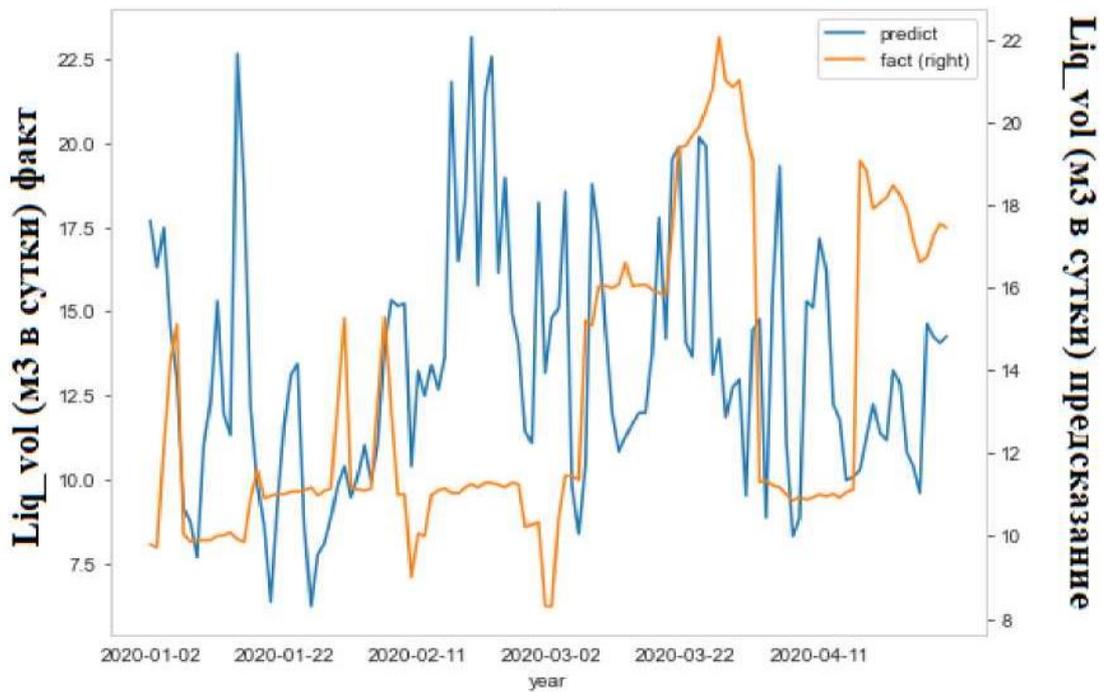


Рисунок 7. Применение сложного 8-ми уровневго ядра Conv-LSTM2D (multistep) строящего пространственно временное влияние переменных многомерного ряда данных multivariate output. 12 прогнозируемые переменные water vol, ВНР_pump_press на вход, 4 переменные liq vol на выход (statefull)

Рисунок 8 показывает работу MLP-MLP с множеством однодневных шагов на период с января 2020 года по конец апреля 2020 года. На вход подаются ВНР_pump_press, water_volume. Прогнозируется liquid_volume. Сильный разброс данных объясняется слабой корреляцией входных параметров сети друг с другом и с выходным параметром.



Тестовые данные - предсказание 1 дневными лагами

Рисунок.8. MLP-MLP ядро сети bhp_pump_press, water volume - liqvolume 1 шаг

Описание работы модели Модели ConvLstm2D(8-9)-ConvLstm2D(8-9слоёв)

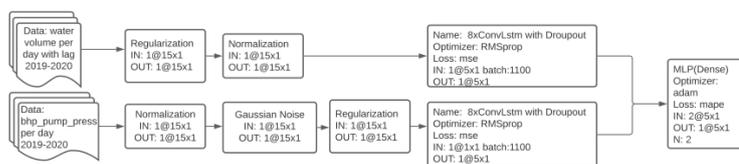
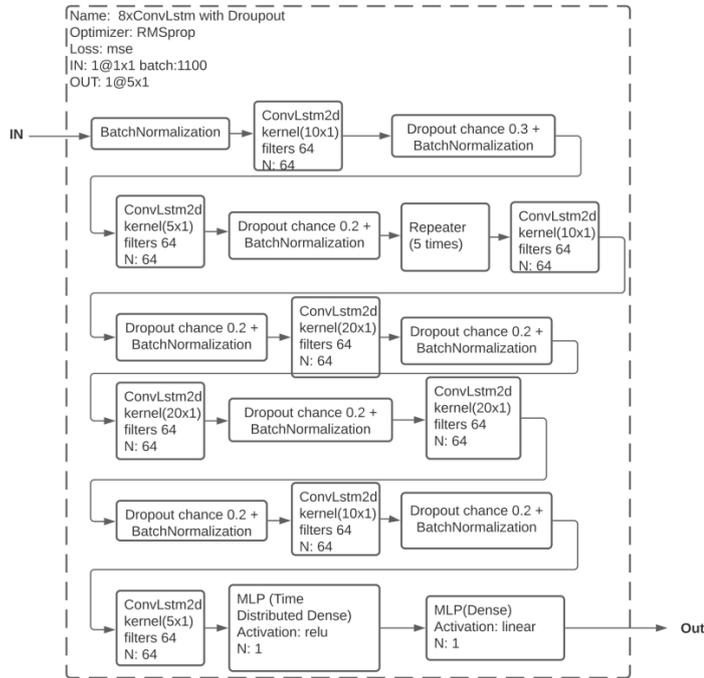


Рисунок 9. Архитектура ConvLstm2D(8-9)-ConvLstm2D(8-9слоёв)

Описание модели

- Данные подаются на вход одной ветви с лагом, длина выборок 365 дней (шагов), ширина по 5 векторов), `water_volume` (с 5-ти скважин) за 2019-2020 годы.
- На вход второй ветви без лага, длина выборок 365 дней (шагов), ширина по 5 векторов `bhp_pump_press` (5 скважин) за 2019-2020 годы.
- Далее каждый слой работают со своими наборами данных.

Описание работы первого слоя

Данные подаются на вход слоя регуляризации в виде тензора $1@15x1$, затем на слой нормализации в виде тензора $1@15x1$, после чего подаются на слой ConvLSTM (описание см. ниже), в результате работы которого получается тензор $1@5x1$.

Описание работы второго слоя

Данные попадают на слой нормализации (централизация данных), Гауссовой пенализации шумов (входной тензор $1@15x1$) и регуляризации, после чего в виде тензора $1@15x1$ попадают на вход ConvLSTM (описание см. ниже), в результате получается тензор $1@5x1$.

После работы каждой ветви данные совместно подаются на слой полносвязной активации с оптимизатором Adam и Loss-функцией MAPE, в виде тензора $2@5x1$. В результате получается тензор $1@5x1$ предсказаний уровня добычи жидкости в день.

Описание работы слоя ConvLSTM with Dropout

Данные на вход подаются в виде тензора $1@1x1$, количеством 1100 пакетов. Выполняется пакетная нормализация данных. Далее данные поступают на слой ConvLSTM2D, где выполняется отбор значимых параметров окном 10, количеством фильтров 64. Слой содержит 64 нейрона. На следующем слое происходит отбрасывание нейронов с вероятностью 30% и пакетная нормализация оставшихся. Затем данные попадают на следующий слой ConvLSTM2D, где происходит фильтрация окном размером 5 на 64 нейронах (32 нейрона в слое), далее отбрасывание с вероятностью 20% и пакетная нормализация. Далее данные десериализуются пятикратно. После этого данные проходят подобный путь несколько раз, до слоя полносвязной активации Relu, после чего подаются на слой полносвязной активации Linear. После чего, в виде тензора $1@5x1$ подаются для оптимизации методом RMSProp с Loss-функцией MSE.

Описание работы модели Модель ConvLstm2D(8-9)

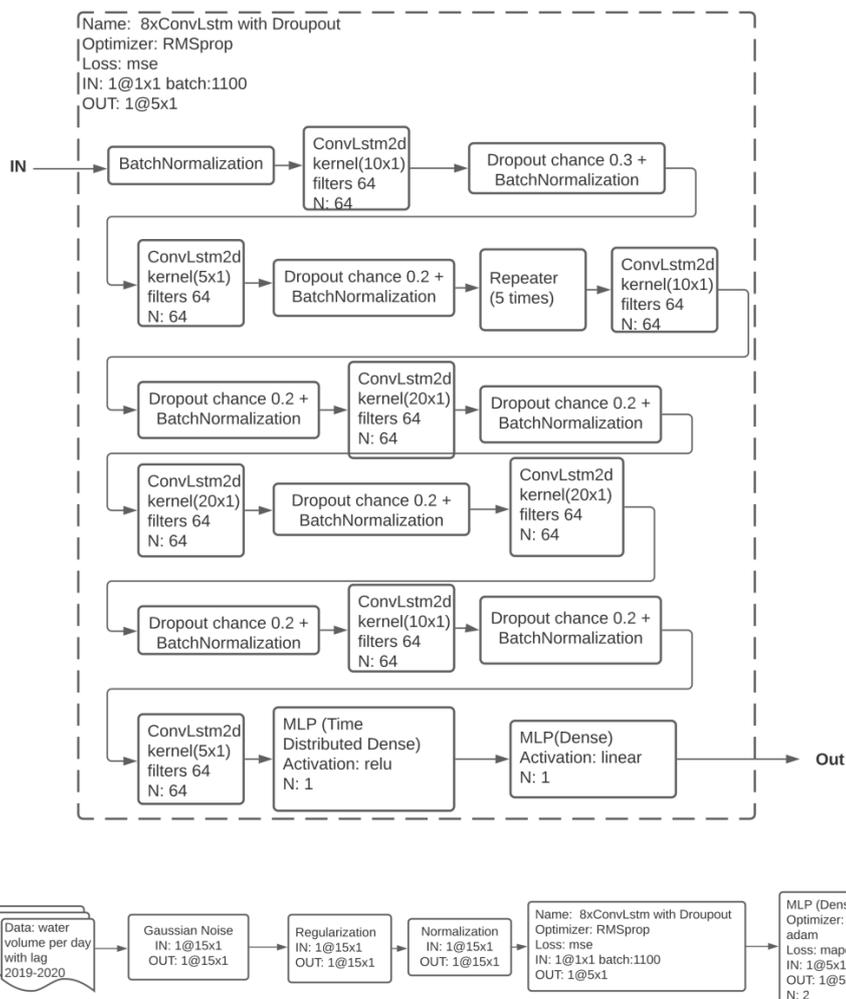


Рисунок 10. Архитектура ConvLstm2D(8-9)

Данные подаются на вход одной ветви с лагом, длина выборки 365 дней (шагов), ширина по 5 векторов), water_volume («объем закачиваемой воды в день») (с 5-ти скважин) за 2019-2020 годы; попадают на слой Гауссовой пенализации шумов (входной тензор $1@15x1$) после чего в виде тензора $1@15x1$ попадают на слой регуляризации и нормализации (централизация данных) тензорами $1@15x1$ и попадают на вход ConvLSTM (описание см. ниже), после чего в виде тензора $1@5x1$ подаются на слой полносвязной активации с оптимизатором Adam и с Loss-функцией MAPE и в результате получается тензор предсказаний $1@5x1$ — «Объем добычи жидкости в день».

Описание работы слоя ConvLSTM with Dropout

- Данные на вход подаются в виде тензора $1@1x1$, количеством 1100 пакетов. Выполняется пакетная нормализация данных. Далее данные поступают на слой ConvLSTM2D, где выполняется отбор значимых параметров окном 10, количеством фильтров 64. Слой содержит 64 нейрона. На следующем слое происходит отбрасывание нейронов с вероятностью 30% и пакетная нормализация оставшихся.
- Затем данные попадают на следующий слой ConvLSTM2D, где происходит фильтрация окном размером 5 на 64 нейронах (32 нейрона в

слое), далее отбрасывание с вероятностью 20% и пакетная нормализация. Далее данные десериализуются пятикратно.

- После этого данные проходят подобный путь несколько раз, до слоя полносвязной активации Relu, после чего подаются на слой полносвязной активации Linear. После чего, в виде тензора $1@5 \times 1$ подаются для оптимизации методом RMSProp с Loss-функцией MSE.

Описание работы Модели LSTM-Repeater-LSTM

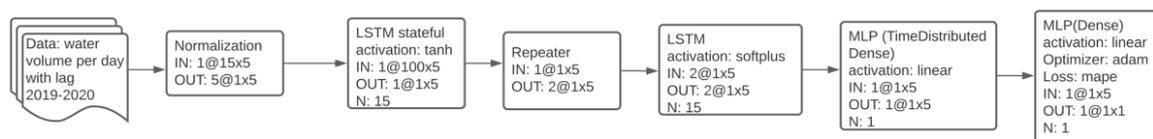


Рисунок 11. Архитектура LSTM-Repeater-LSTM

Описание работы модели LSTM-Repeater -LSTM

- Данные подаются на вход одной ветви с лагом, длина выборок 365 дней (шагов), ширина по 5 векторов), water_volume (с 5-ти скважин) за 2019-2020 годы; Затем попадают на слой нормализации (централизация данных) в виде тензора $1@15 \times 5$.
- Затем данные подаются на слой LSTM, активация Tanh, с опций работы с данными имеющими разные лаги относительно исследуемой зависимой величины в виде тензора $1@100 \times 5$. Далее данные подаются на слой репитера, где происходит двукратная десериализация.
- После чего подаются на LSTM слой активации Softplus, тензором $2@1 \times 5$, затем на слой MLP активации linear тензором $1@1 \times 5$. После чего в виде тензора $1@5 \times 1$ подаются на слой полносвязной активации с оптимизатором Adam и с Loss функцией MAPE и в результате получается тензор предсказаний $1@1 \times 1$ — «Объема добычи жидкости в день».

Описание работы модели MLP-MLP

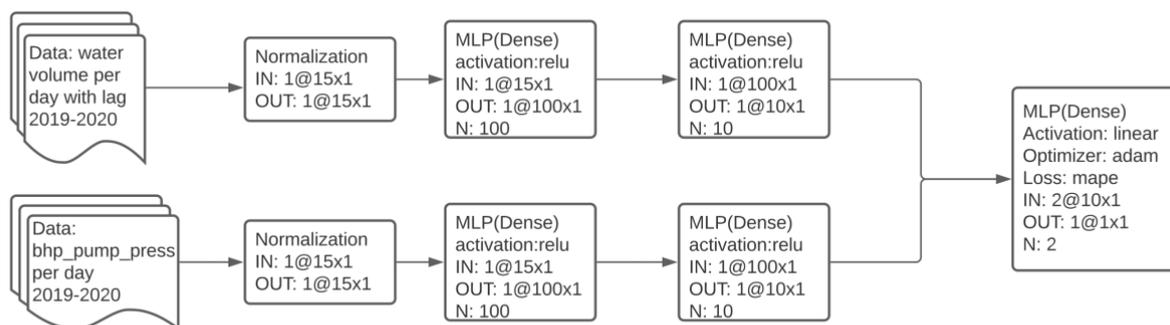


Рисунок 12. Архитектура MLP-MLP

- Данные подаются на вход одной ветви с лагом, длина выборок 365 дней (шагов), ширина по 5 векторов), water_volume (с 5-ти скважин) за 2019-2020 годы;
- На вход второй ветви без лага, длина выборок 365 дней (шагов), ширина по 5 векторов bhp_pump_press (5 скважин) за 2019-2020 годы.
- Далее оба слоя работают одинаково со своими наборами данных.

Описание работы одного слоя: данные попадают на слой нормализации (централизация данных) тензором $1@15x1$, затем на слой MLP с активационной функцией Relu в виде тензора $1@15x1$, далее на слой MLP с активацией Relu в виде тензором $1@100x1$. Результирующий тензор $1@10x1$.

После работы каждой ветви данные совместно подаются на слой полносвязной активации Linear с оптимизатором Adam и Loss-функцией MAPE. В результате получается тензор $1@1x1$ предсказаний уровня добычи жидкости в день.

Описание работы Модели LSTM-MLP

- Данные подаются на вход одной ветви с лагом, длина выборок 365 дней (шагов), ширина по 5 векторов), water_volume (с 5-ти скважин) за 2019-2020 годы.
- На вход второй ветви без лага, длина выборок 365 дней (шагов), ширина по 5 векторов bhp_pump_press (5 скважин) за 2019-2020 годы.
- Далее оба слоя работают одинаково со своими наборами данных.

Описание работы одного слоя:

Данные попадают на слой нормализации (централизация данных) в виде тензора $1@15x5$, затем на слой MLP с активационной функцией Relu в виде тензора $1@15x5$, где преобразуются в тензор $1@100x5$ и подаются на слой LSTM, активация Tanh, с опций работы с данными, имеющими разные лаги относительно исследуемой зависимой величины. Результирующий тензор подается на слой репитера, где данные десериализуются (двукратно) после

чего подаются на LSTM слой активации Softplus, тензором 2@100x5, затем на слой MLP активации linear тензором 1@100x5.

После работы каждой ветви данные совместно подаются на слой полносвязной активации с функцией Linear, с оптимизатором Adam и Loss-функцией MAPE. В результате получается тензор 1@1x1 предсказаний уровня добычи жидкости в день.

Подобные модели были сделаны для скважин месторождения, для которых имелись данные. Результаты получились схожими.

В ходе работы предполагалось усложнение и его интеграция с нелинейными физическими моделями. На данный момент уже внедряется архитектура интеграции физических и нейронных параметров. Пример такой модели полностью на рисунке ниже (см. Рис. 9). Смысл такой интеграции состоит в том, чтобы нейронная сеть учитывала физические процессы взаимодействия скважин. Такие как, например, обводненность и разница давлений между скважинами, где одно давление подаётся на одну сеть, а другое на вторую, а результирующая модель показывает разницу или иные зависимости. Блоки sequential это восьмиуровневые CNN(Conv)LSTM ядра сворачивающие признаки временных рядов. После каждого слоя CNN-LSTM слой batch-normalization необходимые для конвертации данных между ядрами разного размера. И в конце слой time-Distributed для конвертации формата. Дальнейшее изменение возможно в сторону создания принципиально новых компонент сетей и построения новых принципов их работы на базе интеграции с физическими процессами резервуара.

В ходе работ с отделом ML сделаны рекомендации по улучшению работы и применению алгоритмов анализа данных. В настоящее время есть несколько ключевых направлений машинного обучения и методов моделирования нефтяных месторождений.

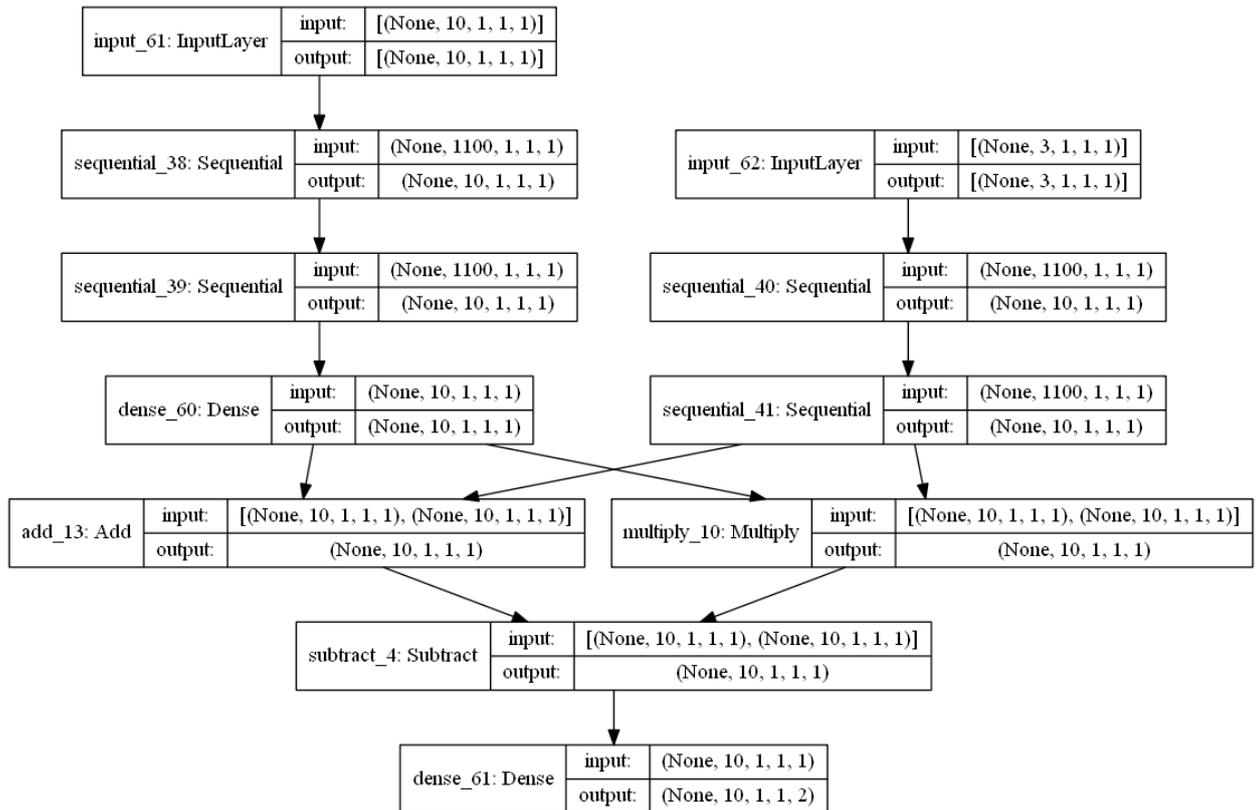


Рисунок 9. Интеграция моделей CNN_LSTM в ансамбль архитектуры для моделирования физических процессов между скважинами

Вышеописанные работы и рассмотренные модели нейронных сетей: MLP-MLP, LSTM-LSTM, LSTM-MLP, ConvLSTM2D-ConvLSTM2D с различными слоями для регуляризации и нормализации данных, где добавлены слои чистки данных от шума через свёртку. По результатам можно сказать, что лучше всего себя показали сверточные рекуррентные многослойные модели.

2.2 Возможные конфигурации вариантов моделей ML (Machine Learning)

Следующим этапом было - разработка и проверка гипотез возможных конфигураций с упрощением комбинаций связки нейронных сетей, в связи с достаточно долгим процессом проведения расчетов на рабочих станциях, имеющих графические процессоры.

На основании полученных результатов были выдвинуты возможные конфигурации и структуры моделей, одним из также требований является возможность интеграции создаваемой модели с физическими моделями (модели построенные на физических законах – существующих симуляторах), ниже приведены предполагаемые варианты моделей:

Количество моделей (на каждый объект разработки и каждый купол)

- 1 доб.скважина – 1 модель – 1 архитектура
- 1 доб.скважина – 1 модель – общая архитектура
- Все добывающие скважины – 1 модель – общая архитектура

Структура имитационной модели

- Модель 1
 - Вход: данные по заводнению и давлению
 - Прогноз: объем жидкости
- Модель 2:
 - Вход: данные по заводнению и давлению
 - Прогноз: обводненность
- Итоговый прогноз: объем нефти (M1-output x M2-output)

Интеграция с физическими моделями:

- Моделируемое пластовое давление – на вход в ML-модели
- Верификация / корректировка / усреднение / сглаживание прогнозов ML-моделей путем сопоставления с прогнозами физ.моделей
- Определение кластеров скважин через физ.модели

2.3 Сложности в данных – затруднения качественного обучения моделей

При проведении расчетов также было отмечено, что множество факторов выявленных в ходе работ влияют на достоверность получения качественных результатов, где следующим пунктом потребовалось провести работу по выявлению и компоновке ряда основных факторов, которые могли быть выявлены только на качественном уровне и визуальном просмотре данных:

Факторы влияющие на качество расчетов модели и предположения к их решению приведены ниже:

- Зашумленность данных
 - В выборке имеются статистические выбросы (скачки) в данных
 - Возможное решение: сглаживание
- Существенные (длительные) пропуски в данных
 - пример: отсутствуют данные по забойному давлению на 2017 год
- Влияние редких событий (ГТМ):
 - Влияние ГТМ вносит как скачки в виде шумов, так и постоянные изменения (сдвиги) в динамику добычи
 - При этом кол-во ГТМ относительно небольшое
- Паузы в добыче
 - Трактуются в моделях как 0
 - Учитывая временную природу данных, исключать данные периоды нельзя
- «Короткие» скважины
 - В некоторых скважинах наблюдается короткая история эксплуатации

Ниже приведены визуальные примеры статистических выбросов, структурных сдвигов, длительные паузы и короткие скважины:

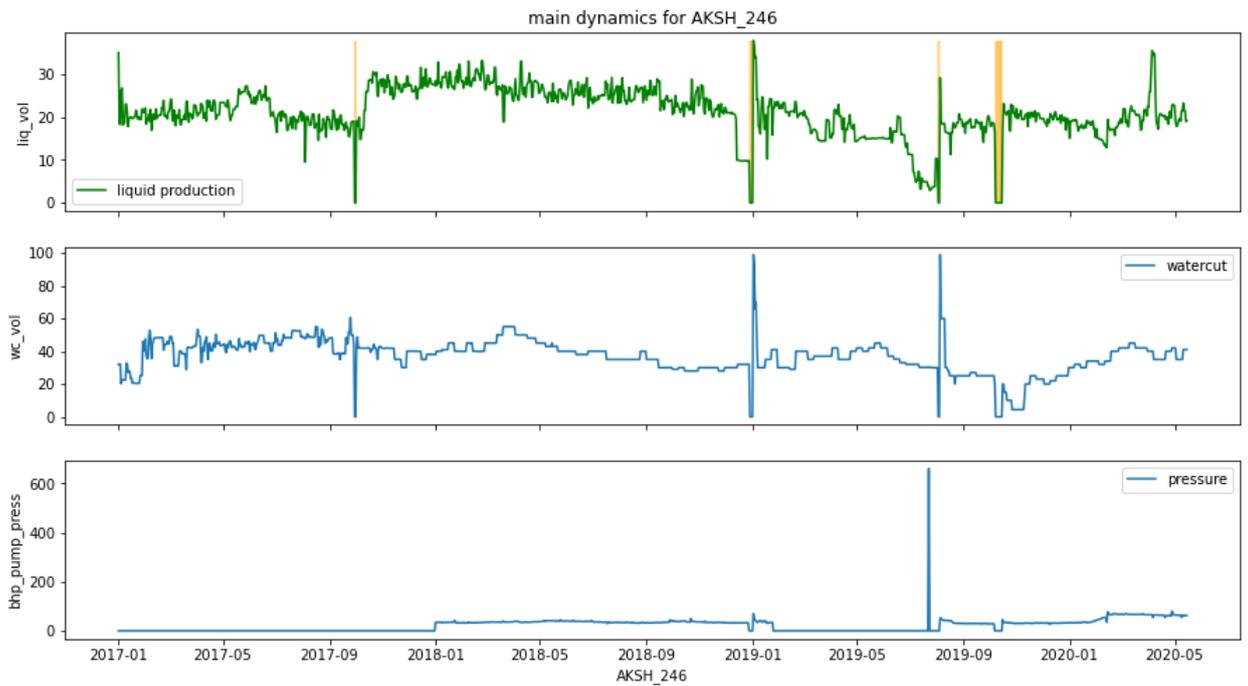


Рисунок 10. Статистические выбросы и структурные сдвиги

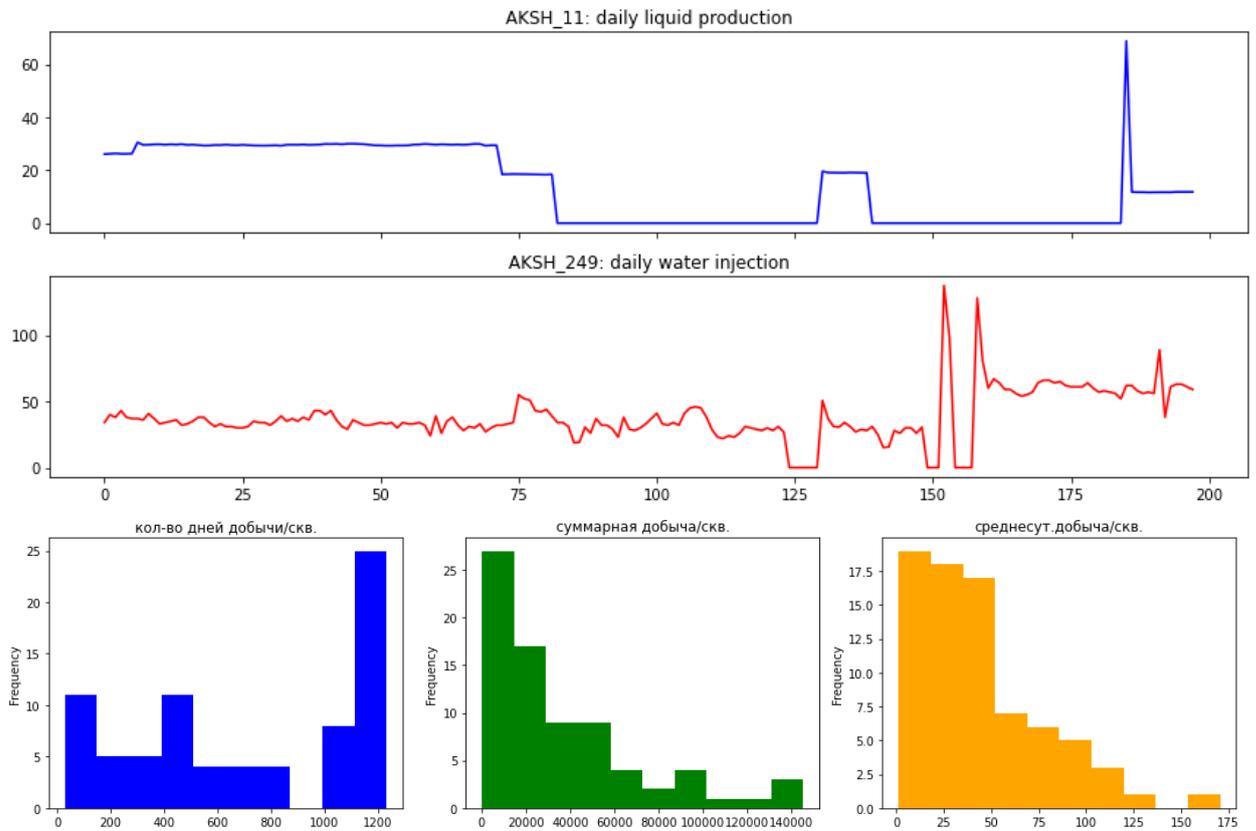


Рисунок 11. Длительные паузы и короткие скважины

Глава 3. Конфигурации и структуры моделей

После проведения вышеописанных работ и проведенного анализа данных выдвинуто решение о принятии в работу по тестированию конфигураций и структур, описаны ниже:

- Модель-1 (регрессор на Liq_Vol)
 - Общие параметры моделей:
 - Таргет: LiqVol
 - Feats: LiqVol (lagged), WaterVol (в пределах купола), ВНР_pump_press (day-to-day basis)
 - Размер лага: 15 суток (1..15)
 - Длина прогноза – 1 сутки
 - Конечная длина (30 суток) будет генерироваться рекурсивным алгоритмом
 - Влияние ГТМ, выбросов и пауз – в процессе ожидания
- Типы моделей
 - «Плоские» модели
 - Разработана модель many-to-many (генерирующая прогноз сразу по нескольким скважинам)
 - Хорошие метрики сходимости (R2: 0,85-0,9)
- ИНС (Искусственные нейронные сети):
 - Разработана модель single-well-output (combined LSTM-MLP)
 - Результаты – умеренные
 - Варианты улучшения:
 - Препроцессинг данных (пластовое давление, сглаживание шумов, ГТМ)
 - Улучшение архитектуры
 - Подбор индивидуальной архитектуры для каждой скважины (трудоемко, но гарантирует улучшение метрик)

3.1 Результаты тестирования моделей и сопоставление полученных результатов по плоским моделям МЛ

Протестированные алгоритмы: линейные, SVM, RandomForest, Бустинговые, Наилучшая модель (на тестовой выборке – 10%, Северный Купол):

- Алгоритм: AdaBoost (kernel – DecisionTree)
- R2 on test (усреднение по суткам): 0,99
- R2 on test (усреднение по скважинам): 0,81
- Не учитывалось отключение скважин на тестовый период -> 25 скважин из 45 – неудовл.результат

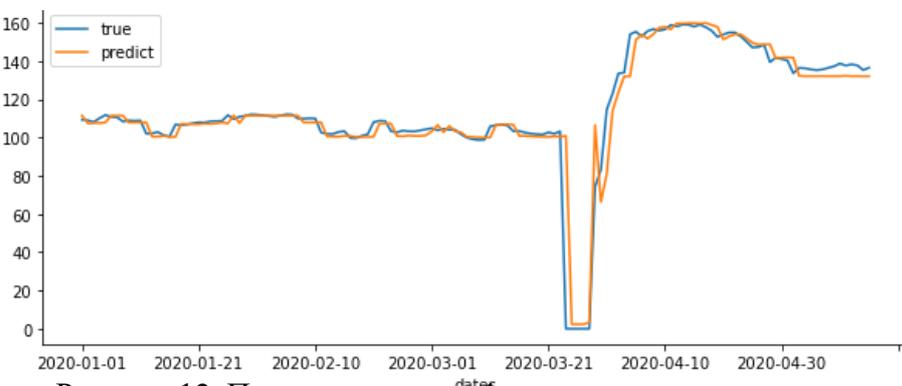
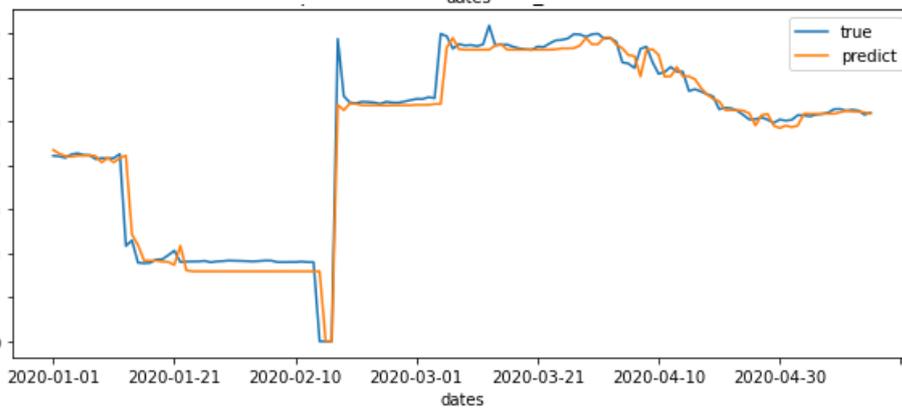
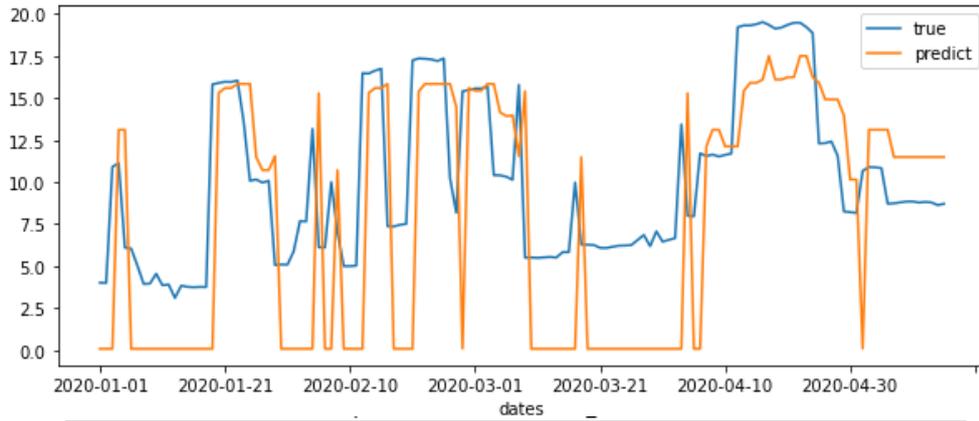
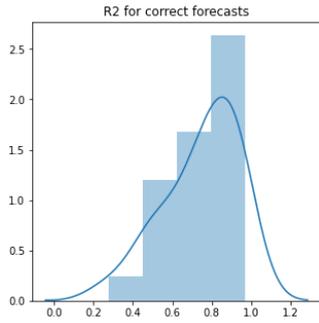


Рисунок 12. Плоские модели с Γ шагом предсказания

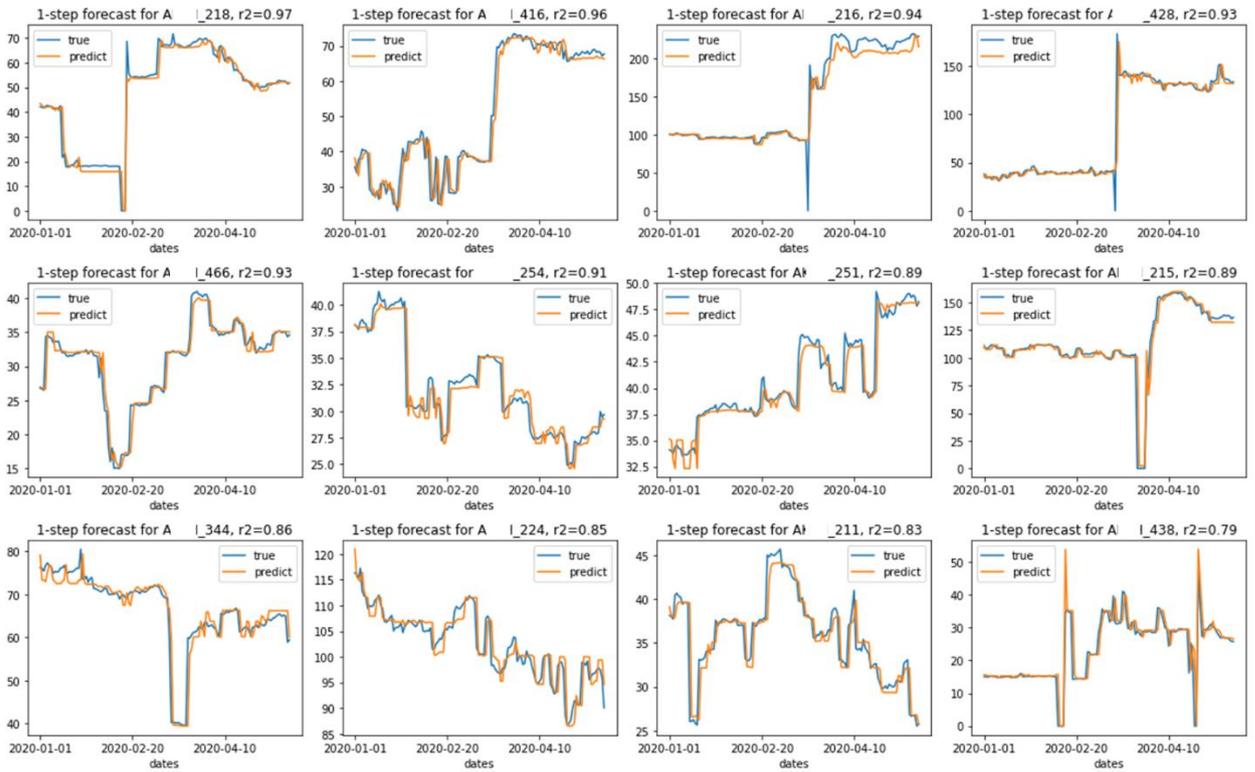


Рисунок 13. Плоские модели с 1 шагом предсказания - Хорошие Прогнозы

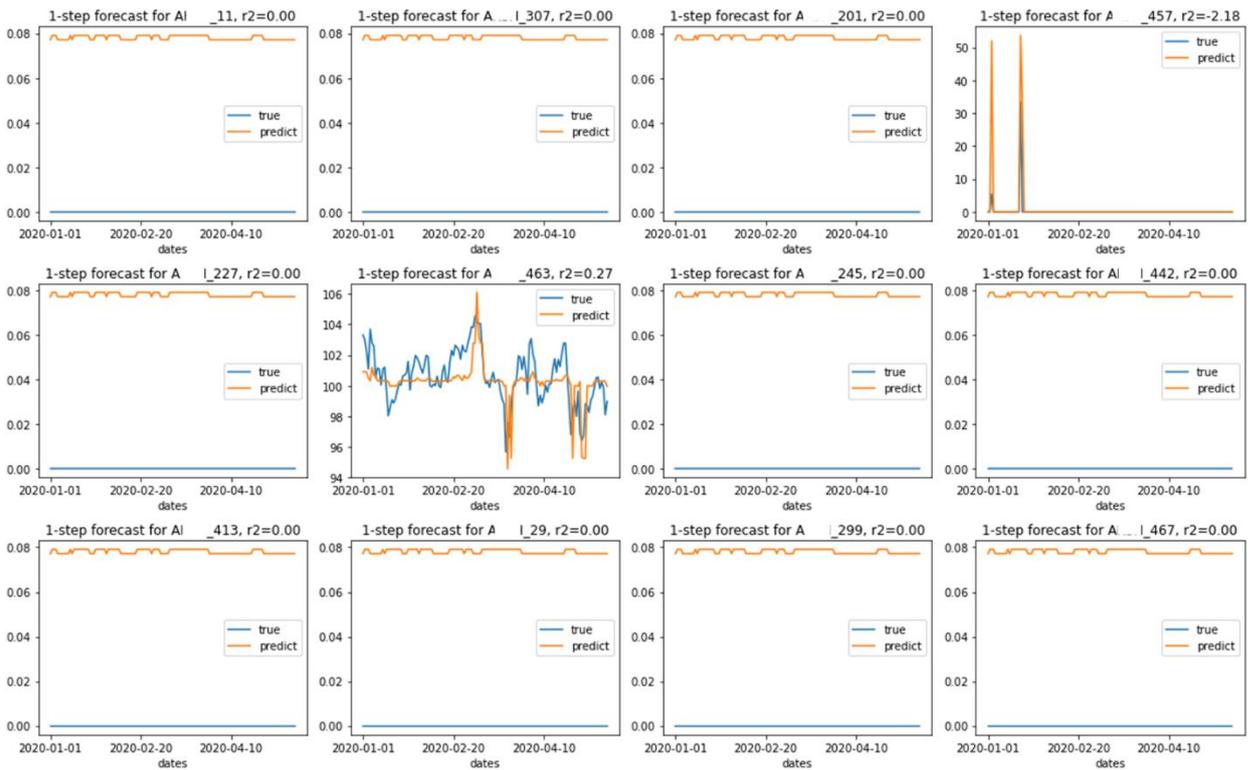


Рисунок 13. Плоские модели с 1 шагом предсказания – Неудовлетворительные результаты прогнозов

3.2 Результаты тестирования моделей и сопоставление полученных результатов по моделям ИНС

Протестированная архитектура:

- LSTM(watervol, liqvol) + MLP (pump_press)
- Модель типа Single Well – 1 step forecast
- Для тестирования выбраны по 5 скважин с каждого купола, имеющих минимальное кол-во пропусков
- Результаты валидации (test, 0,1):
 - R2 mean (with bad wells): negative
 - R2 mean (w/o bad wells): 0.82
 - R2 median (with bad wells): 0,63
 - R2 median (w/o bad wells): 0.82
 - # bad wells (R2<0,5): 5 of 10
 - R2 mean (good wells) North vs South: 0,86-0,80
- Future Steps – Будущие шаги:
 - Mini_Clusters (детализация архитектуры)
 - Pauses and GTM
 - Сглаживание шумов
 - Пластовое давление
 - Multi-well version:
 - 2D -> 3D
 - Long DF with dummies

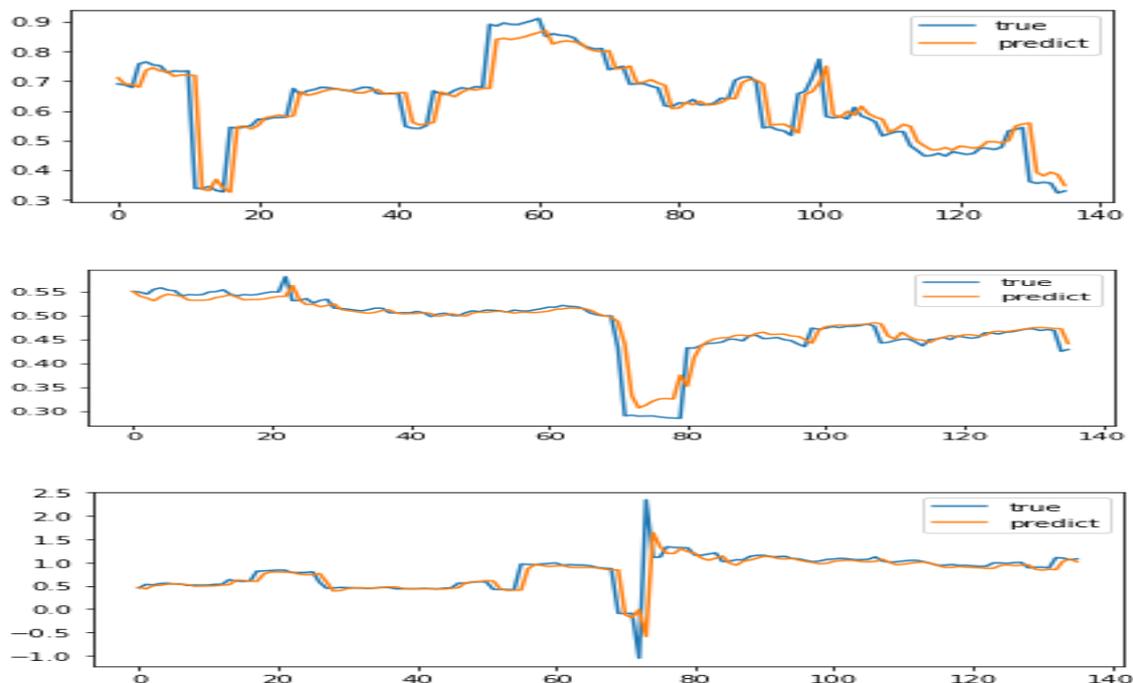


Рисунок 14. модели ИНС

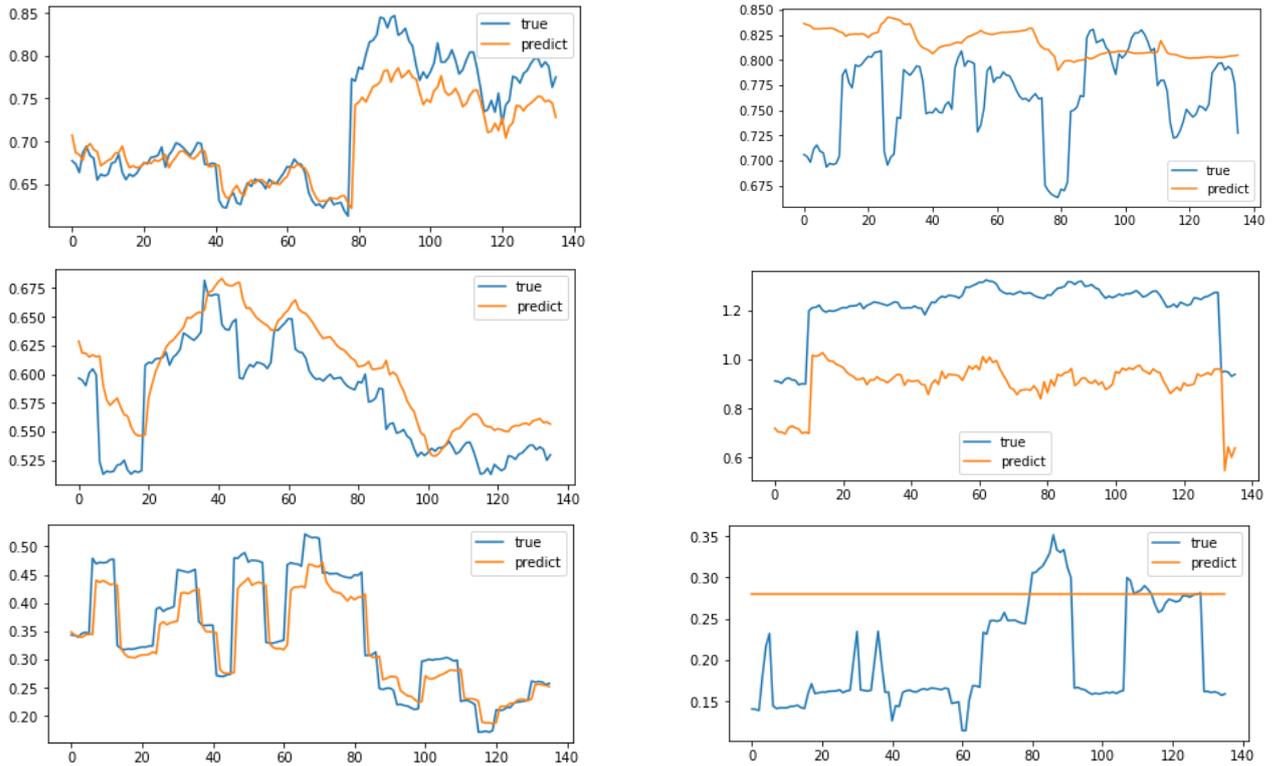


Рисунок 15. модели ИНС – Прогнозы

ИНС. Архитектура

- LSTM Network (15-days window)
 - Input -> LSTM(32, linear) -> Dropout(0,2) -> LayerNormalization(axis=2, 1e-9) -> LSTM(16, softplus) -> Dense(1, None)
- MLP Network (1-feat)
 - Input -> Dense(16, linear) -> Dropout(0,25) -> LayerNormalization(axis=1, 1e-9) -> Dense(32, sigmoid) -> Dropout(0,25) -> LayerNormalization(axis=1, 1e-9) -> Dense(1, softplus)
- Combined Network:
 - Merged Inputs -> {LSTM N, MLP N} -> Dense(2, None) -> Dense(1, selu)
- Optimizer and fit params:
 - Adam(lr=0,001)
 - Batch size: 128
 - Epoches: 500
 - Early Stop Patience: 0,2
 - Loss function: MSE
 - Metrics: MSE
 - Validation subset: TEST subset

ЗАКЛЮЧЕНИЕ

В результате проведенных работ подготовлены начальные модели нейронных сетей для проекта с различными слоями регуляризации и нормализации данных. Произведена попытка добавления слоев «умной» чистки данных от шума через свёртку, где по результатам можно сказать, что лучше всего себя показали сверточные рекуррентные многослойные модели.

Получен опыт к пониманию картины и видение в следующих шагах к продолжению работ:

- Продолжение работ над улучшением параметров нейронных сетей;
- Увязка информации моделей резервуара с моделями нейронных сетей;
- Интегрирование прокси-моделей и ИНС для улучшения результатов;
- Рассмотрение возможностей интеграции данных в единую систему или сеть;
- Продолжение изучения альтернативных методов, таких как физические-ML, методы Reinforcement-learning моделирование, пространственные графы, цепи и др.

Вышеописанные шаги позволят улучшить прогнозирование для «шумных» и «неполных» данных. Возможно, потребуется интеграция для преодоления данных барьеров. На данный момент имеются примеры достаточно глубокого физического анализа параметров резервуара. Также не стоит забывать о более простых путях с использованием необычных методов, где даже они требуют обоснования применения каждой системы сетей и метода обработки данных (дата сет).

В случае работы с данным проектом выявлены основные факторы, влияющие на результаты, также как сложное воздействие факторов друг на друга.

Данная работа проведена в целях поиска и применимости новейших подходов в анализе данных, где следующими этапами требуется продолжить исследования в нахождении путей решения задач с помощью интеграции новых методов, работающих с различных данных как пример: пороговые и дискретные значения. Требуется также изыскать новые подходы и методы работы с неоднородностью и подземного резервуара т.к. в Нефтегазовой промышленности - имеющиеся скважинные и геолого-геофизические данные покрывают очень малую часть подземного мира.

Практическая значимость полученных результатов

Наличие накопленных огромных массивов данных в нефтяной промышленности дает возможность использовать их для принятия более эффективных, рациональных решений и прогнозирования будущих результатов. На ряду с тем, что на данный момент увеличение производительности компьютеров актуализировал вопрос обработки больших данных и применение нейронных сетей в научно-исследовательских работах получило широкое распространение практически во всех сферах деятельности человека.

Можно сказать, что в Нефтегазовой отрасли на текущий момент проведено недостаточное количество исследований с использованием машинного обучения и нейронных сетей что открывает простор к исследованиям и тестированию различных методов.

По результатам изучения опыта обработки аналогичных данных в других компаниях, таких, как Газпром, Schlumberger, IBM, Total, Шелл и опыта других коллективов со всего мира в смежных направлениях, выявлена острая необходимость более точного прогнозирования при добыче УВ сырья.

Необходимо отметить большой интерес крупных НГ компаний к данному направлению отмечается попытками ускорения развития направления, где зачастую они являются спонсорами и инициаторами при проведении различных конкурсов по всему миру для анализа данных нефтегазового сектора с применением новейших подходов обработки.

В период исследования по теме данной магистерской диссертации:

- опубликована статья в Научно-практическом журнале «Вестник нефтегазовой отрасли Казахстана» №2 (3) 2020, на тему «ПРИМЕНЕНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ В ЛИТОЛОГИЧЕСКОМ ОПИСАНИИ КЕРНА».

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Zeng, B., Yang, H., Xu, R., Zhou, W., & Han, X. (2019). Lcf: A local context focus mechanism for aspect-based sentiment classification. *Applied Sciences*, 9(16), 3389.
2. Tseng, B. H., Kreyszig, F., Budzianowski, P., Casanueva, I., Wu, Y. C., Ultes, S., & Gasic, M. (2018). Variational cross-domain natural language generation for spoken dialogue systems. arXiv preprint arXiv:1812.08879.
3. Yan, X., Rastogi, A., Villegas, R., Sunkavalli, K., Shechtman, E., Hadap, S., ... & Lee, H. (2018). Mt-vae: Learning motion transformations to generate multimodal human dynamics. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 265-281).
4. Makarenkov, V., Rokach, L., & Shapira, B. (2019). Choosing the right word: Using bidirectional LSTM tagger for writing support systems. *Engineering Applications of Artificial Intelligence*, 84, 1-10.
5. Cui, Z., Ke, R., Pu, Z., & Wang, Y. (2018). Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction. arXiv preprint arXiv:1801.02143.
6. Xiong, H., Pandey, G., Steinbach, M., & Kumar, V. (2006). Enhancing data analysis with noise removal. *IEEE Transactions on Knowledge and Data Engineering*, 18(3), 304-319.
7. Gwak, J., Choy, C. B., Chandraker, M., Garg, A., & Savarese, S. (2017, October). Weakly supervised 3d reconstruction with adversarial constraint. In *2017 International Conference on 3D Vision (3DV)* (pp. 263-272). IEEE.
8. Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017). Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 652-660).
9. Yang, H., Zhu, G., Wang, Y., Su, J., & Zhang, B. (2014). The geological characteristics of reservoirs and major controlling factors of hydrocarbon accumulation in the Ordovician of Tazhong area, Tarim Basin. *Energy exploration & exploitation*, 32(2), 345-368.
10. Cao, F., Luo, H. & Lake L. W. (2014). Development of a Fully Coupled Two-phase Flow Based Capacitance Resistance Model (CRM). Paper presented at SPE Improved Oil Recovery Symposium, Oklahoma.
11. Kim, J. S. (2011). Development of Linear Capacitance-Resistance Models for Characterizing Waterflooded Reservoirs. PhD dissertation, The University of Texas Austin, Texas.
12. Kim, J. S., Lake, L. W. & Edgar, T. F. (2012). Integrated Capacitance-Resistance Model for Characterizing Waterflooded Reservoirs. Paper presented at the 2012 IFAC Workshop on Automatic Control in Offshore Oil and Gas Production, Trondheim
13. https://github.com/jhihan/GARCH_Neural_Network/blob/master/GARCH_NN.ipynb

14. <https://github.com/yohanesnuwara/reservoir-engineering>
15. https://github.com/rmojgani/PhysicsAwareAE/blob/master/Experiments/LSTM/Wave/main_eulerian_wave_small.ipynb
16. Odi, U., & Nguyen, T. (2018, August). Geological facies prediction using computed tomography in a machine learning and deep learning environment. In SPE/AAPG/SEG Unconventional Resources Technology Conference. Unconventional Resources Technology Conference.
17. Oil Reservoir Simulations, https://dataplatform.cloud.ibm.com/analytics/notebooks/v2/b5e9870d-09e7-46e7-81f6-8fda491528ee/view?access_token=c41287b79fafacff3644c93bdac4006a79175161cbcfb5630c304781f641cfd4&cm_sp=ibmdev-_-developer-blogs-_-cloudreg