

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа Инженерная школа природных ресурсов
 Направление подготовки (специальность) 21.04.01 «Нефтегазовое дело»
 Отделение школы (НОЦ) Отделение нефтегазового дела

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Разработка алгоритма автоматизированной интерпретации данных ГИС в целях расчленения и корреляции нижнемеловых отложений Приобского нефтяного месторождения

УДК 528:004:551.763.1(571.122)

Студент

Группа	ФИО	Подпись	Дата
2ТМ91	Абдуллаев Рафаэль Азерович		15.06.2021

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
профессор	Белозеров В.Б.	д.г.-м.н.		15.06.2021

КОНСУЛЬТАНТЫ:

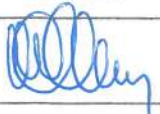
По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Рукавишников В.С.	PhD		15.06.2021

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
профессор	Белозеров В.Б.	д.г.-м.н.		15.06.2021

ДОПУСТИТЬ К ЗАЩИТЕ:

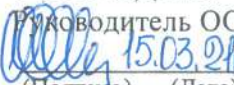
Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
профессор	Чернова О.С.	д.г.-м.н.		15.06.2021

Томск – 2021 г.

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа Инженерная школа природных ресурсов
 Направление подготовки (специальность) 21.04.01 «Нефтегазовое дело»
 Отделение школы (НОЦ) Отделение нефтегазового дела

УТВЕРЖДАЮ:


 Руководитель ООП
 О.С. Чернова
 (Подпись) (Дата)

ЗАДАНИЕ

на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации
(бакалаврской работы, дипломного проекта/работы, магистерской диссертации)

Студенту:

Группа	ФИО
2ТМ91	Абдуллаеву Рафаэлю Азеровичу

Тема работы:

Разработка алгоритма автоматизированной интерпретации данных ГИС в целях расчленения и корреляции нижнемеловых отложений Приобского нефтяного месторождения	
Утверждена приказом директора ИШНПТ	Приказ № 61-6/с от 02.03.2021

Срок сдачи студентом выполненной работы:	15.06.2021
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	Пакет геологической и геофизической информации по ЮЛТ Приобского месторождения, тексты и графические материалы отчетов геолого-технического отдела
Перечень подлежащих исследованию, проектированию и разработке вопросов	Литературный обзор по проблеме исследования, геологическое описание объекта исследования, теоретическое обоснование предлагаемой методики, расчетная часть, социальная ответственность, финансовый менеджмент, ресурсосбережение и ресурсоэффективность

Перечень графического материала	Рисунок 1.1 – Примеры используемых в процессе интерпретации ГИС кросс-плотов Рисунок 1.2 – Сравнение показателей качества работы комплекса простых нейронных сетей и единичной сложной сети Рисунок 1.3 – Принципиальная схема процесса создания алгоритма для автоматизированного расчленения и корреляции горизонтов алгоритмами машинного обучения Рисунок 2.1 – Обзорная карта района работ Рисунок 2.2 – Сводный геологический разрез Приобского месторождения Рисунок 2.3 – Выкопировка из тектонической карты центральной части Западно-Сибирской плиты Рисунок 2.4 – Схема клиноформенного разреза ЮЛТ Приобского месторождения Рисунок 2.5 – Основные геоморфологические элементы ЮЛТ Приобского месторождения Рисунок 3.1 – Условия и последовательность формирования осадочных формаций Рисунок 3.2 – Графическое представление предлагаемого алгоритма Рисунок 3.3 – Диаграмма связи процессов в методологии CRISP-DM Рисунок 3.4 – Итоговая схема автоматизированного алгоритма расчленения и корреляции геологических горизонтов Рисунок 4.1 – Карта кустовой площадки X с нанесенными скважинами Рисунок 4.2 – Матрица скважин и записанных в них методов ГИС Рисунок 4.3 – Представленные литотипы, их обозначения и сокращения Рисунки 4.4-4.5 – Визуальное представление расчленения разреза скважины Рисунок 4.6 – Нормализация данных Рисунок 4.7 – Процесс импортирования и настройки методов классификации Рисунки 4.8-4.11 – Оценка точности произведенного расчленения Рисунки 4.12-4.17 – Процесс оптимизации гиперпараметров моделей Рисунок 4.18 – Матрица ошибок алгоритма случайного леса
---	---

Рисунок 4.19 – Стратиграфическое расчленение скважины 30746 Рисунки 4.21-4.22 – Оценка точности произведенной корреляции	
Консультанты по разделам выпускной квалификационной работы (с указанием разделов)	
Раздел	Консультант
Финансовый менеджмент	Рукавишников В.С., доцент
Социальная ответственность	Белозеров В.Б., профессор
Названия разделов, которые должны быть написаны на русском и иностранном языках:	
Литературный обзор по проблеме автоматизации данных при стратиграфическом расчленении и корреляции меловых отложений / Using of a machine learning methods for facies identifying and correlation based on well logging data	

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	15.03.2021
---	------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
профессор	Белозеров В.Б.	д.г.-м.н.		15.03.2021

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
2TM91	Абдуллаев Рафаэль Азерович		15.03.2021

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И
РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

Группа	ФИО
2ТМ91	Абдуллаеву Рафаэлю Азеровичу

Инженерная школа	Природных ресурсов	Отделение	Нефтегазового дела
Уровень образования	Магистратура	Направление/специальность	21.04.01 Нефтегазовое дело

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Стоимость ресурсов научного исследования: материально-технических, энергетических, финансовых, и человеческих ресурсов проведения исследования по созданию программного продукта автоматизированного расчленения и корреляции продуктивных отложений
--	--

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого и инновационного потенциала НТИ	Оценка стоимостного выражения проведения научного исследования по созданию программного продукта автоматизированного расчленения и корреляции продуктивных отложений
2. Планирование процесса управления НТИ: структура и график проведения, бюджет, риски и организация закупок	Расчет: 1. Затраты на выплату заработной платы сотрудникам 2. Накладные расходы
3. Определение ресурсной, финансовой, экономической эффективности	Расчет затрат на проведение научного исследования по созданию программного продукта автоматизированного расчленения и корреляции продуктивных отложений

Перечень графического материала

Таблица 6.1 – Матрица SWOT-анализа
Таблица 6.2 – Рабочая группа проекта

Дата выдачи задания для раздела по линейному графику	15.03.2021
--	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Рукавишников В.С.	PhD		15.03.2021

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
2ТМ91	Абдуллаев Рафаэль Азерович		15.03.2021

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
2ТМ91	Абдуллаеву Рафаэлю Азеровичу

Инженерная школа	Природных ресурсов	Отделение	Нефтегазового дела
Уровень образования	Магистратура	Направление/специальность	21.04.01 «Нефтегазовое дело»

Исходные данные к разделу «Социальная ответственность»:

1. Описание рабочего места (рабочей зоны) на предмет возникновения: - Вредных проявлений факторов производственной среды (метеосостояние, освещение, шумы, электромагнитные поля) - Опасных проявлений факторов производственной среды (электрической и пожарной природы) - Негативного воздействия на окружающую природную среду (атмосферу, гидросферу и биосферу)	Объект исследования: алгоритм интерпретации данных ГИС на основе методов машинного обучения для автоматизированного расчленения и корреляции отложений нижнего мела. Область применения: разработка нефтяных, газовых и газоконденсатных месторождений. Рабочее место: стол, компьютер.
2. Знакомство и отбор законодательных и нормативных документов по теме:	- Электробезопасность (в т. ч. статическое электричество, средства защиты); - Пожароопасность (причины, профилактические мероприятия, первичные средства пожаротушения)

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Анализ выявленных вредных факторов проектируемой производственной среды в следующей последовательности: - физико-химическая природа вредности, её связь с разрабатываемой темой; - действие фактора на организм человека; - приведение допустимых норм с необходимой размерностью (со ссылкой на соответствующий нормативно-технический документ); - предлагаемые средства защиты (сначала коллективной защиты, затем – индивидуальные защитные средства)	- Действие фактора на организм человека. - Приведение допустимых норм с необходимой размерностью (с ссылкой на соответствующие нормативно-технические документы). - Предлагаемые средства защиты и коллективной защиты, индивидуальные защитные средства
2. Анализ выявленных опасных факторов проектируемой производственной среды в следующей последовательности - механические опасности (источники, средства защиты); - термические опасности (источники, средства защиты); - электробезопасность (в т.ч. статическое электричество, молниезащита – источники, средства защиты); - пожаровзрывобезопасность (причины, профилактические мероприятия, первичные средства пожаротушения)	- Электробезопасность (в т. ч. статическое электричество, средства защиты). - Пожароопасность (причины, профилактические мероприятия, первичные средства пожаротушения)
3. Охрана окружающей среды: - защита селитебной зоны - анализ воздействия объекта на атмосферу (выбросы); - анализ воздействия объекта на гидросферу (сбросы); - анализ воздействия объекта на литосферу (отходы); - разработать решения по обеспечению экологической безопасности со ссылками на НТД по охране окружающей среды.	При строительстве и эксплуатации нефтепроводов будет оказывать негативное воздействие, в основном, на состояние земельных ресурсов и атмосферного воздуха. - Охрана земельных ресурсов и растительного мира - Охрана атмосферного воздуха - Охрана водной среды

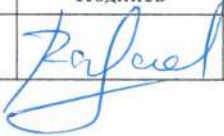
4. Защита в чрезвычайных ситуациях: – перечень возможных ЧС на объекте; – выбор наиболее типичной ЧС; – разработка превентивных мер по предупреждению ЧС; – разработка мер по повышению устойчивости объекта к данной ЧС; – разработка действий в результате возникшей ЧС и мер по ликвидации её последствий.	В районе деятельности возможно возникновение следующих видов чрезвычайных ситуаций: 1. Техногенного характера. 2. Природного характера. 3. Пожарная и взрывная опасность.
5. Правовые и организационные вопросы обеспечения безопасности: – специальные (характерные для проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны	Специальные правовые нормы трудового законодательства
Перечень графического материала:	
При необходимости представить эскизные графические материалы к расчётному заданию (обязательно для специалистов и магистров)	Таблица 5.1 – Нормы освещенности офисных помещений Таблица 5.2 – Оптимальные параметры микроклимата рабочих мест

Дата выдачи задания для раздела по линейному графику	15.03.2021
--	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
профессор	Белозеров В.Б.	д.г-м.н.		15.03.2021

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
2ТМ91	Абдуллаев Рафаэль Азерович		15.03.2021

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное
образовательное учреждение высшего образования
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Школа Инженерная школа природных ресурсов
Направление подготовки (специальность) 21.04.01 «Нефтегазовое дело»
Уровень образования Магистратура
Отделение школы (НОЦ) Отделение нефтегазового дела
Период выполнения (осенний / весенний семестр 2020/2021 учебного года)

Форма представления работы:

Магистерская диссертация
(бакалаврская работа, дипломный проект/работа, магистерская диссертация)

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы

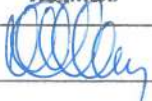
Срок сдачи студентом выполненной работы:	15.06.2021
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
01.04.2021	Литературный обзор	15
13.04.2021	Раздел на английском языке	5
25.04.2021	Теоретическое обоснование предлагаемой методики	35
08.05.2021	Расчетная часть	35
18.05.2021	Финансовый менеджмент	5
24.05.2021	Социальная ответственность	5

Составил преподаватель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
профессор	Белозеров В.Б.	д.г.-м.н.		15.03.21г.

СОГЛАСОВАНО:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
профессор	Чернова О.С.	д.г.-м.н.		15.03.21г.

Результаты освоения образовательной программы
Универсальные компетенции выпускников и индикаторы их достижения

Категория компетенций	Код и наименование компетенции	Индикаторы достижения компетенции
Системное и критическое мышление	УК-1. Способен осуществлять критический анализ проблемных ситуаций на основе системного анализа, вырабатывать стратегию действий	И.УК(У)-1.1. Анализирует проблемную ситуацию как систему, выявляя ее составляющие и связи между ними
		И.УК(У)-1.2. Определяет пробелы в информации, необходимой для решения проблемной ситуации, и проектирует процессы по их устранению
		И.УК(У)-1.3. Разрабатывает стратегию решения проблемной ситуации на основе системного и других современных междисциплинарных подходов; обосновывает выбор темы исследований на основе анализа явлений и процессов в конкретной области научного знания
		И.УК(У)-1.4. Использует логико-методологический инструментарий для критической оценки современных концепций в своей предметной области
Разработка и реализация проектов	УК(У)-2. Способен управлять проектом на всех этапах его жизненного цикла	И.УК(У)-2.1. Определяет проблему и способ ее решения через реализацию проектного управления
		И.УК(У)-2.2. Разрабатывает концепцию проекта в рамках обозначенной проблемы: формулирует цель, задачи, обосновывает актуальность, значимость, ожидаемые результаты и возможные сферы их применения
		И.УК(У)-2.3. Осуществляет мониторинг за ходом реализации проекта, корректирует отклонения, вносит дополнительные изменения в план реализации проекта
Командная работа и лидерство	УК(У)-3. Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели	И.УК(У)-3.1. Планирует и корректирует свою социальную и профессиональную деятельность с учетом интересов, особенностей поведения и мнений людей, с которыми работает и взаимодействует
		И.УК(У)-3.2. Организует дискуссии по заданной теме и обсуждение результатов работы команды
		И.УК(У)-3.3. Планирует командную работу, распределяет поручения и делегирует полномочия членам команды
Коммуникация	УК(У)-4. Способен применять современные коммуникативные технологии, в том числе на иностранном(ых) языке(ах), для академического и профессионального взаимодействия	И.УК(У)-4.1. Решает конкретные задачи профессиональной деятельности на основе академического и профессионального взаимодействия с учетом анализа мнений, предложений, идей отечественных и зарубежных коллег
		И.УК(У)-4.2. Составляет, переводит и редактирует различные академические тексты (рефераты, эссе, обзоры, статьи и т.д.)
		И.УК(У)-4.3. Представляет результаты академической и профессиональной деятельности на различных научных мероприятиях, включая международные
		И.УК(У)-4.4. Планирует и организывает совещания, деловые беседы, дискуссии по заданной теме; аргументированно и конструктивно отстаивает свою точку зрения, позицию, идею в академических и профессиональных дискуссиях на государственном и иностранном языках
Межкультурное взаимодействие	УК(У)-5. Способен анализировать и учитывать разнообразие культур в процессе межкультурного взаимодействия	И.УК(У)-5.1. Осуществляет профессиональную и социальную деятельность с учетом особенностей поведения и мотивации людей различного социального и культурного происхождения, в том числе особенностей деловой и общей культуры представителей других этносов и конфессий

Категория компетенций	Код и наименование компетенции	Индикаторы достижения компетенции
		И.УК(У)-5.2. Выстраивает социальное и профессиональное взаимодействие с учётом особенностей деловой и общей культуры представителей разных этносов и конфессий, других социальных групп
		И.УК(У)-5.3. Обеспечивает создание недискриминационной среды для участников межкультурного взаимодействия при личном общении и при выполнении профессиональных задач
Самоорганизация и саморазвитие (в том числе здоровьесбережение)	УК(У)-6. Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	И.УК(У)-6.1. Анализирует использование рабочего времени в широком спектре деятельности: планирование, распределение, постановка целей, делегирование полномочий, анализ временных затрат, мониторинг, организация, составление списков и расстановка приоритетов
		И.УК(У)-6.2. Сочетает выполнение текущих производственных задач с повышением квалификации; корректирует планы в соответствии с имеющимися ресурсами
		И.УК(У)-6.3. Планирует профессиональную траекторию с учетом особенностей как профессиональной, так и других видов деятельности и требований рынка труда

5.2. Общепрофессиональные компетенции выпускников и индикаторы их достижения

Категория компетенций	Код и наименование компетенции	Индикаторы достижения компетенции
Применение фундаментальных знаний	ОПК-1. Способен решать производственные и (или) исследовательские задачи на основе фундаментальных знаний в нефтегазовой области	И.ОПК(У)-1.1. Демонстрирует навыки физического и программного моделирования отдельных фрагментов процесса выбора оптимального варианта для конкретных условий
		И.ОПК(У)-1.2. Использует фундаментальные знания профессиональной деятельности для решения конкретных задач нефтегазового производства
		И.ОПК(У)-1.3. Анализирует причины снижения качества технологических процессов и предлагает эффективные способы повышения качества производства работ при выполнении различных технологических операций
Техническое проектирование	ОПК(У)-2. Способен осуществлять проектирование объектов нефтегазового производства	И.ОПК(У)-2.1. Использует знание алгоритма организации выполнения работ в процессе проектирования объектов нефтегазовой отрасли
		И.ОПК(У)-2.2. Формулирует цели выполнения работ и предлагает пути их достижения
		И.ОПК(У)-2.3. Выбирает соответствующие программные продукты или их части для решения конкретных профессиональных задач
Работа с информацией	ОПК(У)-3. Способен разрабатывать научно-техническую, проектную и служебную документацию, оформлять научно-технические отчеты, обзоры, публикации, рецензии	И.ОПК(У)-3.1. Анализирует информацию и составляет обзоры, отчеты
		И.ОПК(У)-3.2. Владеет навыками аналитического обзора при подготовке рефератов, публикаций и не менее 50 источников при подготовке магистерской диссертации
		И.ОПК(У)-4.1. Определяет основные направления развития инновационных технологий в нефтегазовой отрасли
		И.ОПК(У)-4.2. Обрабатывает результаты научно-исследовательской, практической технической деятельности, используя имеющееся оборудование, приборы и материалы

Исследование	ОПК(У)-5. Способен оценивать результаты научно-технических разработок, научных исследований и обосновывать собственный выбор, систематизируя и обобщая достижения в нефтегазовой отрасли и смежных областях	И.ОПК(У)-5.1. Определяет на профессиональном уровне особенности работы различных типов оборудования и выявление недостатков в его работе
		И.ОПК(У)-5.3. Интерпретирует результаты лабораторных и технологических исследований применительно к конкретным условиям
Интеграция науки и образования	ОПК(У)-6. Способен участвовать в реализации основных и дополнительных профессиональных образовательных программ, используя специальные научные и профессиональные знания	И.ОПК(У)-6.1. Демонстрирует знания основ педагогики и психологии
		И.ОПК(У)-6.2. Демонстрирует умение общаться с аудиторией, заинтересовать слушателей

Профессиональные компетенции выпускников и индикаторы их достижения

Область и сфера профессиональной деятельности	Задача профессиональной деятельности	Основание - профессиональный стандарт, анализ опыта, форсайт	Код и наименование компетенции	Индикаторы достижения компетенции
Тип задач профессиональной деятельности: Научно-исследовательский				
19. Добыча, переработка, транспортировка нефти и газа 40. Сквозные виды профессиональных стандартов	1. Осуществление научных исследований в области профессиональной деятельности	<i>Профессиональный стандарт «Специалист по добыче нефти, газа и газового конденсата», утвержденный приказом Министерства труда и социальной защиты Российской Федерации от 03 сентября 2018г. № 574н (зарегистрирован Министерством юстиции Российской Федерации 24 сентября 2018 г., регистрационный №52235); ОТФ. Код D. Организация работ по добыче углеводородного сырья</i>	ПК -1. Способен проводить анализ и обобщение научно-технической информации по теме исследования, осуществлять выбор методики и средств решения задачи, проводить патентные исследования в выбранной области нефтегазового инжиниринга	И.ПК-1.1. Анализирует и обобщает научно-техническую информацию по теме исследования, осуществляет выбор методики и средств решения задачи, проводит патентные исследования в выбранной области нефтегазового инжиниринга
			ПК-2. Способен планировать и проводить аналитические, имитационные и экспериментальные исследования, критически оценивать данные и делать выводы	И.ПК-2.1. Планирует и проводит аналитические, имитационные и экспериментальные исследования, критически оценивает данные и делает выводы
	2. Разработка и внедрение новой техники и передовых технологий на объектах нефтегазовой отрасли	19.021 <i>Профессиональный стандарт «Специалист по промысловой геологии», утвержденный приказом Министерства труда и социальной защиты Российской Федерации от 10.03.2015 г. № 151н (зарегистрирован Министерством юстиции Российской Федерации 31.03.2015 г. № 36656) ОТФ. Код В. Организация геолого-промысловых работ</i> <i>Профессиональный стандарт «Специалист-петрофизик», утвержденный приказом</i>	ПК-3. Способен использовать профессиональные программные комплексы в области математического и геолого-геофизического моделирования технологических процессов и объектов	И.ПК-3.1. Использует профессиональные программные комплексы в области математического и геолого-геофизического моделирования технологических процессов и объектов

Область и сфера профессиональной деятельности	Задача профессиональной деятельности	Основание - профессиональный стандарт, анализ опыта, форсайт	Код и наименование компетенции	Индикаторы достижения компетенции
		Министерства труда и социальной защиты Российской Федерации от 29 июня 2017 г. N 534н (зарегистрирован Министерством юстиции Российской Федерации 13 июля 2017 г., регистрационный №47411). <i>ОТФ. Код С.</i> Организация процесса исследований физических свойств кернового материала нефтегазовых месторождений и цифровой обработки полученных петрофизических данных	ПК-10. Способен разрабатывать документацию, планировать и выполнять исследования физических свойств кернового материала осадочных горных пород и цифровую обработку полученных петрофизических данных	И.ПК-10.1. Разрабатывает документацию, планирует и выполняет исследования физических свойств кернового материала осадочных горных пород и цифровую обработку полученных петрофизических данных
	3. Осуществление техниче-ского руководства по выполнению научно-исследовательских и опытно-конструкторских работ, разработки комплексных проектов на всех стадиях и этапах выполнения работ и непосредственное участие в их выполнении	<i>Профессиональный стандарт «Специалист-петрофизик»,</i> утверждённый приказом Министерства труда и социальной защиты Российской Федерации от 29 июня 2017 г. N 534н (зарегистрирован Министерством юстиции Российской Федерации 13 июля 2017 г., регистрационный №47411) <i>ТФ. Код С/01.7.</i> Разработка плановой и проектно-сметной документации на объекты исследований физических свойств кернового материала горных пород и цифровую обработку полученных петрофизических данных <i>Профессиональный стандарт «Специалист по организации и управлению научно-исследовательскими и опытно-конструкторскими работами»,</i> утвержденный приказом Министерства труда и социальной защиты Российской Федерации от 11 февраля 2014 г. N 86н	ПК-10. Способен разрабатывать документацию, планировать и выполнять исследования физических свойств кернового материала осадочных горных пород и цифровую обработку полученных петрофизических данных ПК-11. Способен организовывать и выполнять научно-исследовательские работы в соответствии с тематическим планом организации	И.ПК-10.1. Организует и выполняет плановые задания по исследованию физических свойств кернового материала горных пород и обработку полученных петрофизических данных И.ПК-10.2. Организует и выполняет плановые задания по исследованию физических свойств кернового материала горных пород и обработку полученных петрофизических данных И.ПК-11.1. Организует и выполняет научно-исследовательские работы в соответствии с тематическим планом организации

Область и сфера профессиональной деятельности	Задача профессиональной деятельности	Основание - профессиональный стандарт, анализ опыта, форсайт	Код и наименование компетенции	Индикаторы достижения компетенции
		(зарегистрирован Министерством юстиции Российской Федерации 21 марта 2014 г., регистрационный №31696) <i>ОТФ. Код D.</i> Осуществление руководства разработкой комплексных проектов на всех стадиях и этапах выполнения работ <i>ТФ. Код. D/01/7.</i> Организация выполнения научно-исследовательских работ в соответствии с тематическим планом организации		
Тип задач профессиональной деятельности: Технологический				
19 «Добыча, переработка, транспортировка нефти и газа» 40. Сквозные виды профессиональных стандартов	Осуществление контроля, технического сопровождения и управления технологическими процессами нефтегазового производства	<i>Профессиональный стандарт «Специалист по добыче нефти, газа и газового конденсата», утвержденный приказом Министерства труда и социальной защиты Российской Федерации от 03 сентября 2018г. № 574н (зарегистрирован Министерством юстиции Российской Федерации 24 сентября 2018 г., регистрационный №52235)</i> <i>ОТФ. Код D.</i> Организация работ по добыче углеводородного сырья <i>ОТФ. Код E.</i> Руководство работами по добыче углеводородного сырья	ПК-4. Способен анализировать и обобщать данные о работе технологического оборудования, осуществлять контроль, техническое сопровождение и управление технологическими процессами в нефтегазовом инжиниринге	И.ПК-4.1. Анализирует и обобщает данные о работе технологического оборудования, осуществляет контроль, техническое сопровождение и управление технологическими процессами в нефтегазовом инжиниринге
			ПК-7. Способен контролировать выполнение требований и регламентов для обеспечения добычи нефти, газа и газового конденсата	И.ПК-7.1. Контролирует выполнение требований и регламентов для обеспечения добычи нефти, газа и газового конденсата
	Разработка и внедрение новой техники и передовых технологий на объектах нефтегазовой отрасли		ПК-5. Способен участвовать в управлении технологическими комплексами, принимать решения в условиях неопределенности	И.П-5.1. Участвует в управлении технологическими комплексами, принимает решения в условиях неопределенности
Тип задач профессиональной деятельности: Организационно-управленческий				

Область и сфера профессиональной деятельности	Задача профессиональной деятельности	Основание - профессиональный стандарт, анализ опыта, форсайт	Код и наименование компетенции	Индикаторы достижения компетенции
19 «Добыча, переработка, транспортировка нефти и газа» 40. Сквозные виды профессиональных стандартов	Осуществление маркетинговых исследований, проведение технико-экономического обоснования инновационных решений в профессиональной деятельности, управление коллективом, руководство производственной деятельностью подразделения	<i>Профессиональный стандарт «Специалист по добыче нефти, газа и газового конденсата», утвержденный приказом Министерства труда и социальной защиты Российской Федерации от 03 сентября 2018г. № 574н (зарегистрирован Министерством юстиции Российской Федерации 24 сентября 2018 г., регистрационный №52235) ТФ. Код Е/01.7. Руководство организацией процесса добычи углеводородного сырья</i>	ПК-6. Способен осуществлять руководство по организации производственной деятельности подразделений нефтегазового инжиниринга, применять полученные знания для разработки и реализации проектов различных процессов производственной деятельности, применять методику проектирования	И.ПК-6.1. Осуществляет руководство по организации производственной деятельности подразделений нефтегазового инжиниринга
		<i>Профессиональный стандарт «Специалист по промысловой геологии», утвержденный приказом Министерства труда и социальной защиты Российской Федерации от 10 марта 2015 г. N 151н (зарегистрирован Министерством юстиции Российской Федерации 31 марта 2015 г., регистрационный №36656) ОТФ. Код В. Организация геолого-промысловых работ ТФ. Код В/02.7. Подготовка предложений по дополнительным геолого-промысловым исследованиям для эффективной работы промысла</i>	ПК-8. Способен подготавливать предложения по дополнительным геолого-промысловым исследованиям для эффективной работы промысла	И.ПК-8.1. Подготавливает предложения по дополнительным геолого-промысловым исследованиям для эффективной работы промысла

Область и сфера профессиональной деятельности	Задача профессиональной деятельности	Основание - профессиональный стандарт, анализ опыта, форсайт	Код и наименование компетенции	Индикаторы достижения компетенции
	Разработка оперативных планов и руководство проведением всех видов деятельности, связанной с исследованием, разработкой и реализацией управления технологическими процессами и производствами	<i>Профессиональный стандарт «Специалист по добыче нефти, газа и газового конденсата», утвержденный приказом Министерства труда и социальной защиты Российской Федерации от 03 сентября 2018г. № 574н (зарегистрирован Министерством юстиции Российской Федерации 24 сентября 2018 г., регистрационный №52235) ТФ. Код E/01.7. Руководство организацией процесса добычи углеводородного сырья</i>	ПК-6. Способен осуществлять руководство по организации производственной деятельности подразделений нефтегазового инжиниринга, применять полученные знания для разработки и реализации проектов различных процессов производственной деятельности, применять методику проектирования	И.ПК-6.2. Применяет полученные знания для разработки и реализации проектов различных процессов производственной деятельности, применяет методику проектирования
		<i>Профессиональный стандарт «Специалист по обработке и интерпретации скважинных геофизических данных», утвержденный приказом Министерства труда и социальной защиты Российской Федерации от 28 декабря 2015 г. N 1166н (зарегистрирован Министерством юстиции Российской Федерации 29 июня 2017 г., регистрационный №47457). ОТФ. Код D. Управление процессом обработки и интерпретации полученных скважинных геофизических данных</i>	ПК-9. Способен разрабатывать перспективные планы в области обработки и интерпретации скважинных геофизических данных, руководить производственно-технологическим процессом обработки и интерпретации скважинных геофизических данных	И.ПК-9.1. Разрабатывает перспективные планы в области обработки и интерпретации скважинных и полевых геофизических данных
		<i>ТФ. Код D/01.7. Управление разработкой перспективных планов области обработки и интерпретации скважинных геофизических данных</i>		И.ПК-9.2. Руководит производственно-технологическим процессом обработки и интерпретации скважинных и полевых геофизических данных

Реферат

Выпускная квалификационная работа включает в себя 117 страниц, 34 рисунка, 7 таблиц, 56 источников и 3 приложения.

Ключевые слова: НИЖНИЙ МЕЛ, ПРИОБСКОЕ МЕСТОРОЖДЕНИЕ, МАШИННОЕ ОБУЧЕНИЕ, РАСЧЛЕНЕНИЕ РАЗРЕЗА, КОРРЕЛЯЦИЯ ГОРИЗОНТОВ, ГЕОФИЗИЧЕСКИЕ ИССЛЕДОВАНИЯ.

Объектом исследования являются сложнопостроенные клиноформенные отложения нижнего мела (пласты AC₁₀-AC₁₂) Приобского нефтяного месторождения.

Целью работы является создание алгоритма интерпретации данных геофизических исследований скважин (ГИС) на основе методов машинного обучения для автоматизированного расчленения и корреляции отложений нижнего мела с последующей верификацией результатов на керне.

В результате работы разработано теоретическое обоснование методики автоматической корреляции продуктивных отложений, а также программный комплекс, включающий модули расчленения разреза скважин по данным ГИС, корреляции межскважинного пространства и визуализации данных. Также был проведен сравнительный анализ использованных при построении моделей алгоритмов.

Областью применения является разработка нефтяных, газовых и газоконденсатных месторождений.

Экономическая эффективность (значимость работы) заключается в создании методики, позволяющей значительно сократить временные затраты на сопоставление и прослеживание геологических горизонтов, являющихся целевыми объектами разработки.

Апробация работы. Результаты исследования доложены на XXV Международном научном симпозиуме им. Академика М.А. Усова в апреле 2021 г.

Содержание

Реферат	16
Введение	19
1. Литературный обзор по проблеме автоматизации данных при стратиграфическом расчленении и корреляции меловых отложений	21
2. Геологическое описание объекта исследования	31
2.1. Физико-географическая характеристика	31
2.2. Стратиграфическое описание	33
2.3. Тектоническое строение района работ	36
2.4. Промышленная нефтегазоносность	37
3. Теоретическое обоснование предлагаемой методики	41
3.1. Определение оптимальной последовательности расчленения	41
3.2. Подбор оптимальной последовательности корреляции	46
3.3. Методология решения задач методами машинного обучения	48
4. Расчетная часть	51
4.1. Подготовка исходных данных	51
4.2. Выделение в разрезе глинистых пачек	57
4.3. Выделение аномалий ГИС в пределах глинистых пачек и их систематизация	70
4.4. Определение опорных аномалий и проведение межскважинной корреляции	72
5. Социальная ответственность	77
5.1. Анализ вредных производственных факторов	77
5.2. Анализ опасных производственных факторов	80
5.3. Охрана окружающей среды	82
5.4. Защита в чрезвычайных ситуациях	83
5.5. Обеспечение надлежащей охраны труда	84
6. Финансовый менеджмент, ресурсосбережение и ресурсоэффективность ..	85
6.1. Потенциальные потребители результатов исследований	85

6.2. SWOT-анализ	85
6.3. Цели и результаты проекта	87
6.4. Организационная структура проекта	87
6.5. Экономическая оценка результатов исследования	88
Заключение	93
Список использованной литературы	96
Приложение А	102
Приложение Б	116
Приложение В	117

Введение

Целью работы является создание алгоритма интерпретации данных ГИС на основе методов машинного обучения для автоматизированного расчленения и корреляции отложений нижнего мела с последующей верификацией результатов на керне.

Задачи работы:

- изучить существующие подходы к вопросам ручного и автоматического расчленения и корреляции разрезов скважин;
- подробно ознакомиться с объектом исследования;
- привести теоретическое обоснование предлагаемой методики;
- создать и настроить алгоритмы автоматического расчленения разреза и корреляции межскважинного пространства;
- произвести оценку полученных результатов и верификацию на керновом материале.

Стоит отметить, что подавляющая часть месторождений в настоящее время прошли этап пиковой добычи и вплотную подходят к завершающей стадии разработки либо уже находятся на ней. Ввиду роста потребления углеводородного (УВ) сырья и постоянно ухудшающейся структуры запасов, в нефтегазодобывающей отрасли на первое место выходит увеличение скорости и эффективности производства путем внедрения поэтапной автоматизации и непрерывного наращивания вычислительных мощностей.

Прирост запасов, в первую очередь, происходит за счет ввода в разработку месторождений, содержащих трудноизвлекаемые запасы; альтернативным вариантом является поиск недренированных или неучтенных ранее ресурсов на уже освоенных территориях путем проведения повторных геологоразведочных работ. В обоих случаях ключевыми шагами являются проведение операций по расчленению разреза скважины и последующая корреляция выделенных горизонтов в межскважинном пространстве. Пласты при этом, как правило, имеют сложную структуру и распространение как из-за изменений условий

седиментации в процессе формирования, так и воздействия целого комплекса геологических процессов уже после его завершения. Таким образом, выделение в разрезе скважин отдельных пропластков и их дальнейшая корреляция является сложнейшей задачей, требующей большого объема время- и трудозатрат, а также высокого уровня квалификации.

С другой стороны, разрабатываемые месторождения, особенно вплотную подошедшие к завершающей стадии, характеризуются высокой степенью изученности и информационной обеспеченности. К таким объектам относится и рассматриваемая Южная лицензионная территория (ЮЛТ) Приобского нефтяного месторождения. Появление и дальнейшее распространение устройств с высокой вычислительной мощностью привело к широкомасштабному промышленному внедрению алгоритмов машинного обучения и искусственных нейронных сетей, позволяющих оптимизировать процесс анализа больших объемов данных и повысить качество их обработки.

Объектом исследования являются нижнемеловые отложения (AC_{10} - AC_{12}), являющиеся основными продуктивными объектами месторождения. Горизонты представлены частым и неравномерным переслаиванием нефтенасыщенных песчаников с алевролитами и аргиллитами. В породах часто отмечается включения глауконита и других полупроводящих минералов, а наличие интервалов песчаников, имеющих в своем составе карбонатно-глинистый цемент, вносят дополнительную путаницу в и без того сложную структуру разреза.

Новизна представленной работы заключается в создании комплексного программного продукта на основе алгоритмов машинного обучения. Представленный продукт включает в себя отдельные модули литологического расчленения разрезов скважин, корреляции межскважинного пространства и визуализации полученных результатов. Решение вычислительных задач производилось параллельно несколькими алгоритмами, из которых впоследствии был выбран один, показавший наибольшую точность. В качестве входной информации при построении моделей были использованы данные

геофизических исследований, а также эталонные результаты произведенных интерпретатором расчленений разрезов скважин.

Построение моделей и оценка полученных результатов производилась с помощью математического аппарата высокоуровневого языка программирования «Python».

1. Литературный обзор по проблеме автоматизации данных при стратиграфическом расчленении и корреляции меловых отложений

Литологическая и структурная классификация горных пород, применяемая в российской и общемировой практике, некоторым образом отличается: если в первом случае используется упрощенная типизация на основе морфологии пустотного пространства (присутствие пор, каверн или трещин), то во втором случае породы, как правило, классифицируются на основе их генезиса, химического состава или принадлежности к определенному структурному типу (Петерсилье, 2003; Лусиа, 2010). Так или иначе, вне зависимости от использованной классификации, применение выделенных литотипов будет возможно только при наличии следующих признаков:

- литотипы должны легко идентифицироваться как в керновом материале, так и по данным геофизических исследований соответствующих скважин;
- пористость и проницаемость должны иметь устойчивую корреляцию между собой;
- литотипы должны иметь логичные и геологически объяснимые тренды пространственного распределения (Дюбрьоль, 2009).

Выделение литотипов в основном осуществляется с помощью двух групп методов, имеющих принципиально различную физическую основу: электрическую (как правило, используется метод самопроизвольной поляризации (ПС), реже каротаж сопротивления (КС), боковой (БК) или индукционный каротаж (ИК)) или радиоактивную (гамма-каротаж). Дополнительным инструментом является кавернометрия, которая при сопоставлении ее результатов с номинальным диаметром скважины может послужить отличным индикатором перспективных интервалов (Калантаров, 1971).

Первичное расчленение геологического разреза можно выполнить на основе визуального анализа формы кривых ГИС. Наиболее ярким примером является сравнение поведения кривых нейтронного (НК) и плотностного

каротажа (ГГК-П), отображенных на одном треке в сравнимом масштабе. В таком случае для определения литологии могут быть использованы следующие диагностические признаки:

- Кривая нейтронного каротажа имеет большие значения, чем кривая плотностного. В таком случае рассматриваемый интервал может быть проинтерпретирован как песчаник или газонасыщенная карбонатная порода (чаще всего доломит). Уточнение возможного типа насыщения производится с помощью кривых сопротивления или проводимости.
- Кривые плотностного и нейтронного каротажей примерно совпадают. Такой случай характерен для нефтенасыщенного известкового коллектора. Определение типа насыщаемого флюида можно произвести описанным выше способом.
- Кривая плотностного каротажа имеет большие значения, чем кривая нейтронного. Вероятнее всего, что в данном случае интервал будет представлен либо доломитом, либо глиной в случае повышенных значений естественной радиоактивности. Каверномер в данном случае может помочь в выделении глинистых интервалов.

Для облегчения процесса расчленения разреза и определения минерального состава перспективных горизонтов могут быть привлечены кросс-плоты (Schlumberger, 2009). К наиболее информативным и часто используемым кросс-плотам относят зависимости показаний нейтронного и плотностного каротажей, времени пробега акустической волны и показаний нейтронного метода, а также величины плотности и нейтронной пористости. Примеры подобных кросс-плотов представлены на Рисунке 1.1.

Стоит отметить, что результаты интерпретации вышеперечисленных кросс-плотов напрямую зависят от типа насыщающего породу флюида, поэтому они не всегда являются достоверными. Для достижения максимальной корректности результатов часто необходимо привлекать дополнительные кривые ГИС, например, электрических методов (Wolf, 1982; Busch, 1987; Изотова, 1993; Боганик, 2001).

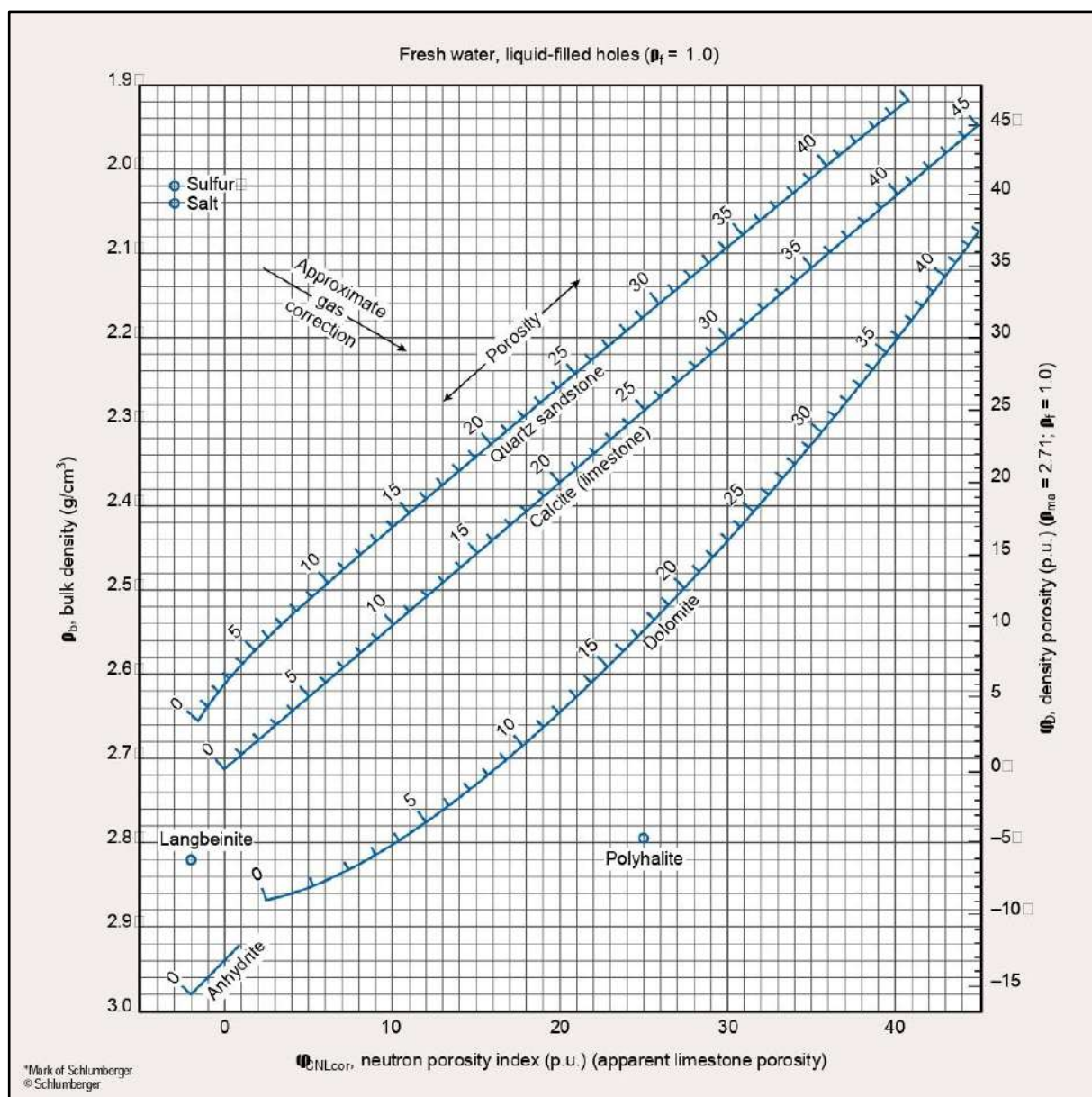


Рисунок 1.1 – Примеры используемых в процессе интерпретации данных ГИС кросс-плотов (Schlumberger, 2009)

Методика интерпретации геофизических исследований немного варьируется в зависимости от геологических условий (Edmundson, 1979; Gardner, 1980; Serra, 1980; Yunsheng, 2017). Так, кривая гамма-каротажа (ГК) может использоваться для выделения глинистых интервалов в случаях недостаточной информативности кривой метода самопроизвольной поляризации, а используя совокупность вышеперечисленных кривых с другими методами каротажа (нейтронными, плотностными, акустическими и т. д.), можно выделить наиболее распространенные литотипы пород (Serra, 1985).

Стоит также отметить, что результаты интерпретации плотностного каротажа позволяют практически напрямую определять литологию вне зависимости от пористости и типа насыщающего флюида (Gardner, 1980; Ellis, 2003). Ввиду того, что породообразующие минералы потенциальных коллекторов не являются многочисленными, может быть определена их плотность в текущий момент времени. Комбинирование результатов интерпретации кривых плотностного каротажа и других методов на основе трехкомпонентной минералогической модели может помочь установить процентное содержание минеральных фракций (Gardner, 1980).

В настоящее время, с появлением высокомоощных рабочих станций компаниям стали доступны специализированные программные продукты, такие как «Techlog» или «Gintel» (от компаний Schlumberger и Roxar соответственно), позволяющие в той или иной степени автоматизировать вышеописанный процесс. Автоматизация же наиболее рутинных действий, таких как расчеты петрофизических параметров, может осуществляться и в программном обеспечении (ПО) более широкого назначения, например, в Microsoft Excel (Ковалевский, 2007).

Следующим шагом после интерпретации данных ГИС является сопоставление полученных в рассматриваемой скважине результатов с соседними в целях корреляции всех, в том числе перспективных, горизонтов. В настоящее время подобная корреляция производится вручную с привлечением сейсморазведки и данных разведочного бурения. Тем не менее, как и в случае интерпретации данных ГИС, некоторые действия могут быть автоматизированы как для увеличения скорости обработки данных, так и для исключения ошибок, вызванных влиянием человека. Дополнительным аргументами в пользу применения программных продуктов является создание параллельно нескольких возможных результатов или вариантов исхода, а также возможность внесения изменений в существующую модель с последующим ее перестроением.

Развитие подобных средств автоматизации является одной из наиболее важных задач, решаемых в настоящее время геологическим сообществом.

Отдельным перспективным направлением в этой области считается интерпретация и корреляция с привлечением алгоритмов машинного обучения, а также нейронных сетей. Прежде всего стоит отметить, что при решении задач, которые требуют использования входных данных, выходящих за граничные интервалы группы признаков обучающей выборки методом экстраполяции, математический аппарат нейронных сетей часто допускает ошибки и в целом является недостаточно точным, а описание полученных с привлечением машинного обучения решений требует сравнительно больших трудозатрат. Таким образом, круг решаемых с помощью методов машинного обучения задач четко определен.

Тем не менее, методы машинного обучения обладают целым рядом достоинств. Так, подобный способ автоматизации позволяет обеспечить высокую гибкость в пределах интервала значений описывающих объект признаков, а также отсеять полученный результат от факторов, который в этот интервал не входили (в случае наличия факторов, которые не влияют на конечный результат, но могут повлиять или исказить процесс ручной интерпретации) (Cannady, 1998). Кроме того, в отличие от часто субъективного визуального анализа входных данных, аппарат методов машинного обучения позволяет улавливать и фиксировать мельчайшие изменения, получая таким образом более однородные результаты. В дополнение к вышесказанному, с помощью автоматических алгоритмов могут быть выделены пропластки, идентификация которых вручную была затруднительна (Validov, 2017).

В настоящее время большая часть исследователей (Cannady, 1998; Chaki, 2015; Validov, 2017), рассматривавших вопросы применения алгоритмов машинного обучения для интерпретации и корреляции горизонтов, использовали в ходе своей работы следующую последовательность действий:

1. Предобработка входных данных – объединение и увязка нескольких наборов данных, их последующая нормировка, исключение взаимнообратных (например, электрических методов сопротивления и проводимости) и

второстепенных (например, номеров скважин) признаков, которые не должны учитываться в модели.

2. Предварительное разделение разреза по известным стратиграфическим отметкам на зоны, разделение набора данных на обучающую, тренировочную и тестовую выборки.
3. Создание нейронной сети или алгоритма машинного обучения для каждой отдельной зоны.
4. Обучение построенного алгоритма и его тренировка на соответствующих выборках.
5. Тестирование полученного алгоритма на данных скважин, где интерпретация известна заранее, и сравнение с полученными в ходе работы алгоритма результатами, а также оценка точности алгоритма и внесение необходимых изменений.

Таким образом, интерпретация геофизических данных и корреляция слоев сначала производится специалистом на некоторой представительной выборке скважин (создается обучающая выборка), затем созданный алгоритм обучается на ней, после чего оценивается на некоторой отложенной группе скважин, для которых задача уже решена (тестовая выборка) для того, чтобы в последствии использоваться на оставшихся скважинах рассматриваемого участка (Ковалевский, 2004; Bishop, 2006; Абдуллаев, 2021).

Особую важность представляет предобработка данных, которая может занимать большую часть времени и от которой зависят оптимальные настройки параметров алгоритма. Так, Chaki отмечает, что в процессе эксперимента по установлению связи между сейсмическими данными и фракционным составом песчаника были опробованы различные методы для нормализации переменных, участвовавших в классификации или являвшихся целевыми. В дальнейшем автором было принято решение использовать для переменных классификатора метод Z-нормализации, а интервал значений признакового описания объекта нормализовать с помощью среднеарифметического и стандартного отклонения по X. После этого преобразованные переменные подверглись дополнительной

нормировке исходя из их максимумов и минимумов, т. е. линейному преобразованию входных данных; соотношение данных до и после преобразования сохраняется. Кроме того, автор подтвердил на практике преимущества применения нескольких менее комплексных нейронных сетей вместо одной более крупной и сложной. Таким образом может быть снижена среднеквадратичная и абсолютная средняя ошибки, а также возрастает коэффициент корреляции и скорость выполнения расчетов. Ниже приведена соответствующая статистика, полученная в ходе проведения экспериментов (Рисунок 1.2).

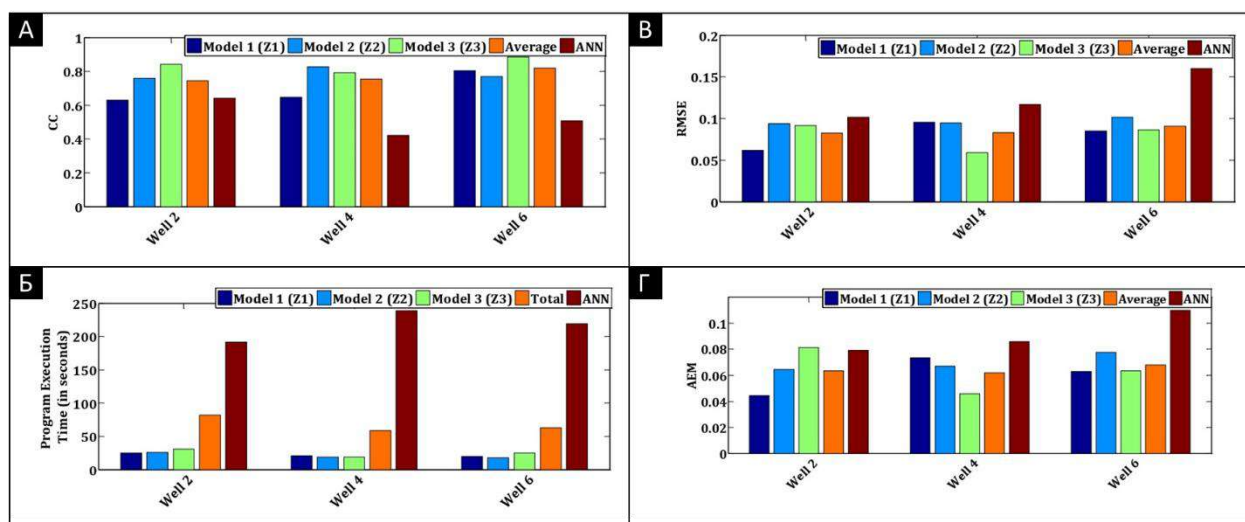


Рисунок 1.2 – Сравнение показателей качества работы комплекса простых нейронных сетей и единичной сложной сети

(А – коэффициенты корреляции; Б – скорость выполнения программ; В – среднеквадратичная ошибка, Г – абсолютная средняя ошибка) (Chaki, 2015)

В большинстве случаев для решения подобных задач, требующих метаматематический аппарат для работы с большим объёмом данных, используется высокоуровневый интерпретируемый язык программирования «Python». Для успешного построения алгоритма машинного обучения должны быть произведены следующие операции:

1. В первую очередь, необходимо создать информативное и в то же время оптимальное с точки зрения размерности описание входных данных на основе признаков (Gurney, 1997; Елисеева, 2002; Дюбрюль, 2009).
2. Выбрать наиболее подходящий для решения задачи алгоритм машинного обучения или структуру нейронной сети, а также подобрать оптимальные параметры для них (функцию активации, количество эпох и т. д.) (Krenker, 2011).
3. Предварительно проанализировать различные виды преобразования геофизических кривых, таких как сглаживание, сегментация и т. п. (Gurney, 1997).
4. Выделить предполагаемые горизонты, для которых будет осуществлен расчет и сравнение коэффициентов корреляции между геофизическими кривыми для того, чтобы выбрать таким образом методы, вносящих максимальный вклад в решение (Hsieh, 2005).
5. Произвести детальную оценку кернового материала (структурные и текстурные особенности, результаты лабораторных испытаний), и в результате сопоставления с данными геофизических исследований определить набор репрезентативных скважин, которые будут использоваться в процессе обучения (Lim, 2003; Luthi, 1997; Li, 2006).
6. Определить оптимальное соотношение для разбиения входных данных на обучающую и тренировочную выборки, а также определить минимальный объем тестовой выборки, применяемой для оценки точности алгоритма (Lim, 1998; Kurgan, 2006; Tianqi, 2016).

На основе общих правил методов машинного обучения могут быть обозначены рекомендации для процесса создания обучающей выборки и непосредственно обучения:

- входные данные по каждой скважине следует представить в виде полного признакового описания свойств горных пород, слагающих разрез, для чего необходимо обозначить некоторый минимальный набор геофизических исследований (Wolf, 1982; Rogers, 1992; Lim, 1999; Yunsheng, 2017);

- продуктивный интервал должен быть представлен максимальным выносом кернового материала (Gardner, 1980);
- в обучающую выборку должны быть включены все характерные для района работ литологические типы пород и области с различным характером насыщения (чисто нефтяная, переходная и другие зоны) (Ковалевский, 2007; Al-Anazi, 2010; Гафуров, 2013).

Вышеперечисленные операции могут быть трансформированы в схему оптимальной последовательности действий по созданию автоматизированного алгоритма литологического расчленения и корреляции скважин алгоритмами машинного обучения, представленную на Рисунке 1.3..

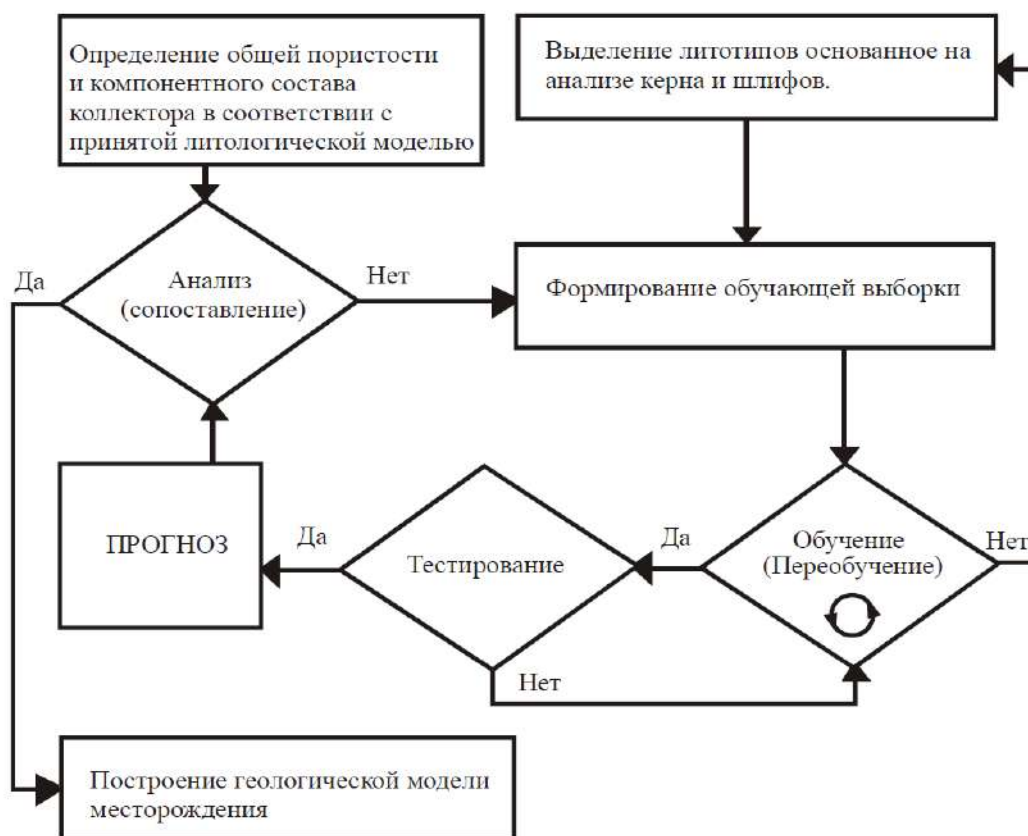


Рисунок 1.3 – Принципиальная схема процесса создания алгоритма для автоматизированного расчленения и корреляции горизонтов алгоритмами машинного обучения (Гафуров, 2013)

2. Геологическое описание объекта исследования

2.1. Физико-географическая характеристика

Южная лицензионная территория (ЮЛТ) была открыта в составе Приобского нефтяного месторождения в 1982 г. и введена в разработку в 1999 г. В административном отношении располагается в Ханты-Мансийском и Кондинском муниципальных районах Ханты-Мансийского автономного округа – Югры (ХМАО), в 26 км к востоку от административного центра округа – города Ханты-Мансийска (Рисунок 2.1).

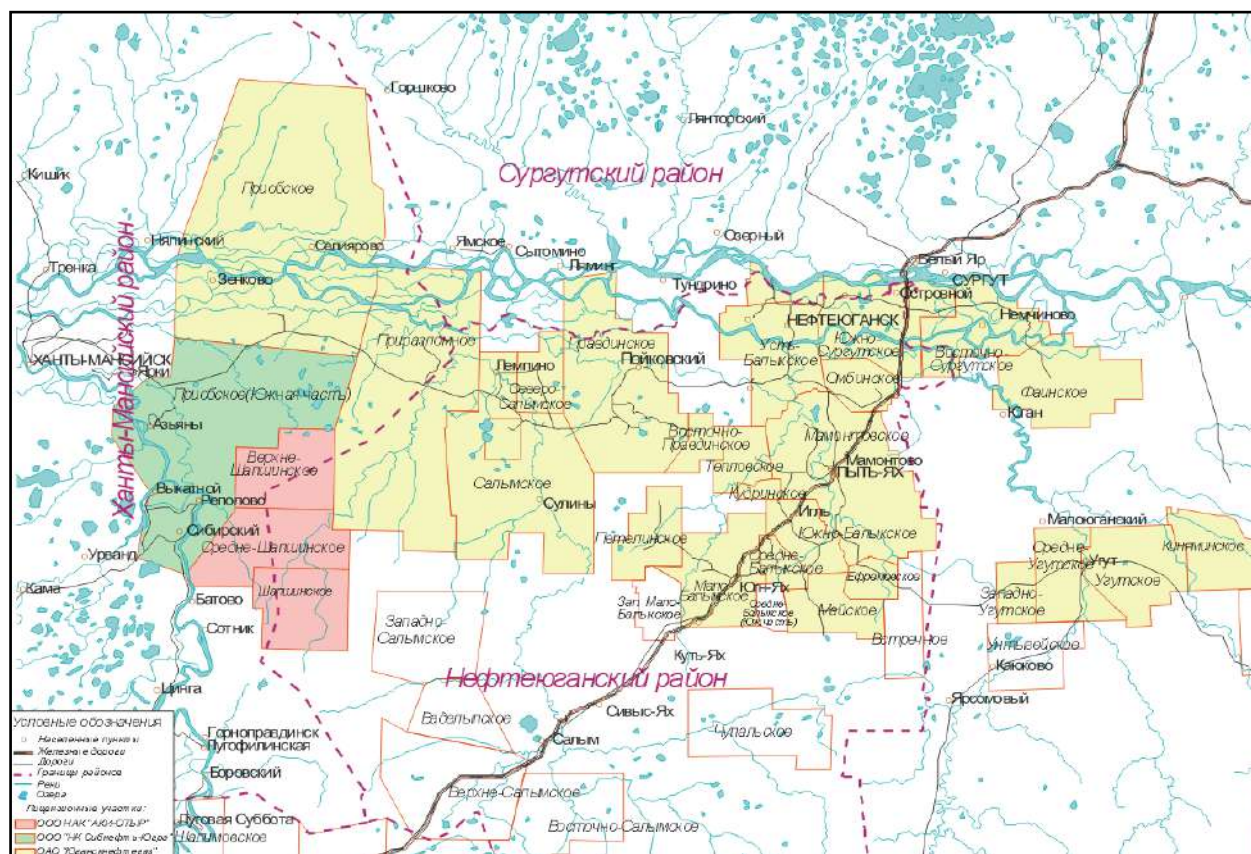


Рисунок 2.1 – Обзорная карта района работ

Территория ЮЛТ Приобского месторождения представляет собой плоскую аллювиальную равнину, включающую слабо выраженные формы речной эрозии и аккумуляции. Стоит также отметить, что поверхность ЮЛТ находится выше Северной лицензионной территории в среднем на 10-15 м. В

непосредственной близости находятся разрабатываемые Шапшинское и Приразломное (на востоке), Чапровское и Кондинское (на юге), Западно-Эргинское (на юго-западе) и Восточно-Камское (на северо-западе) месторождения.

На территории участка располагаются села Базьяны, Тюли, Реполово, поселки Выкатной и Сибирский. Населенные пункты связаны между собой дорогами с асфальтобетонным покрытием, а в северной части участка проходит автодорога федерального назначения Р404 Тюмень – Ханты-Мансийск.

ЮЛТ Приобского месторождения субшироотно пересекается р. Иртыш с развитым судоходством. Ближайший речной порт и международный аэропорт находятся в г. Ханты-Мансийске. Перевозка грузов осуществляется автомобильным и водным транспортом (в летнее время). Ближайшей железнодорожной станцией является Пыть-Ях, расположенный в 260 км к востоку от месторождения.

К юго-востоку от месторождения проходят трассы газопровода Уренгой – Челябинск – Новополюцк и нефтепровода Усть-Балык – Омск.

Климат района резко континентальный, средняя температура самого холодного месяца (января) и самого теплого (июля) составляет -21°C и $+17^{\circ}\text{C}$ соответственно при среднегодовой $-1,5^{\circ}\text{C}$. Осадки в основном выпадают в теплое время года и составляют в среднем до 550 мм в год. Устойчивый снежный покров появляется в октябре и держится до конца апреля, образуя толщу в среднем 2.0 м.

Территория месторождения расположена в зоне разобщенного залегания многолетнемерзлых пород. Приповерхностные мерзлые грунты наблюдаются на водоразделах под торфяниками и достигают 14 м по мощности при постоянной средней температуре 0°C .

На сухих песчаных почвах произрастают хвойные леса, представленные в основном сосной и елью; в состав широко распространенных смешанных лесов входят также осины и березы. Заросли тальника покрывают речные поймы.

2.2. Стратиграфическое описание

Отложения меловой системы на территории объекта исследования распространены повсеместно и представлены нижним и верхним отделами. В геологическом разрезе исследуемого нижнего отдела выделяются отложения ахской, черкашинской, алымской, викуловской и ханты-мансийской свит. Верхний отдел представлен отложениями верхней части ханты-мансийской, а также уватской, кузнецовской, березовской и ганькинской свитами. Фрагмент стратиграфического разреза, соответствующий отложениям нижнемелового возраста, приведен на Рисунке 2.2.

Разрез ахской свиты подразделяется на верхнюю и нижнюю подсвиту, которая, в свою очередь, по литологическому составу подразделяется на подачимовскую и надачимовскую пачки, а также залегающую между ними ачимовскую толщу. Подачимовская пачка сложена темно-серыми до черных аргиллитами и имеет мощность до 40 м. Ачимовская толща характеризуется сложным геологическим строением и представлена частым переслаиванием аргиллитов, алевролитов и песчаников. Песчаные пласты характеризуются сложным строением, литологически не выдержаны по площади и, в целом, имеют чешуйчатое залегание. Выделяемые в разрезе толщи песчаные пласты Ач₁, Ач₂, Ач₃, Ач₄ ограничено распространены на востоке и юго-востоке территории, достигая мощности 150 м. Перекрывающая ачимовскую толщу надачимовская пачка сложена темно-серыми аргиллитами с прослоями песчаников и слоистых алевролитов. К кровле свиты приурочена специфическая пачка темно-серых тонкоотмученных аргиллитов, имеющих региональное развитие на территории Среднего Приобья. Отложения этой пачки, известной как пимская, служат надежным литологическим и электрокаротажным репером. По всему разрезу свиты отмечено наличие аммонитов и фораминифер. Общая мощность свиты достигает 415 м.

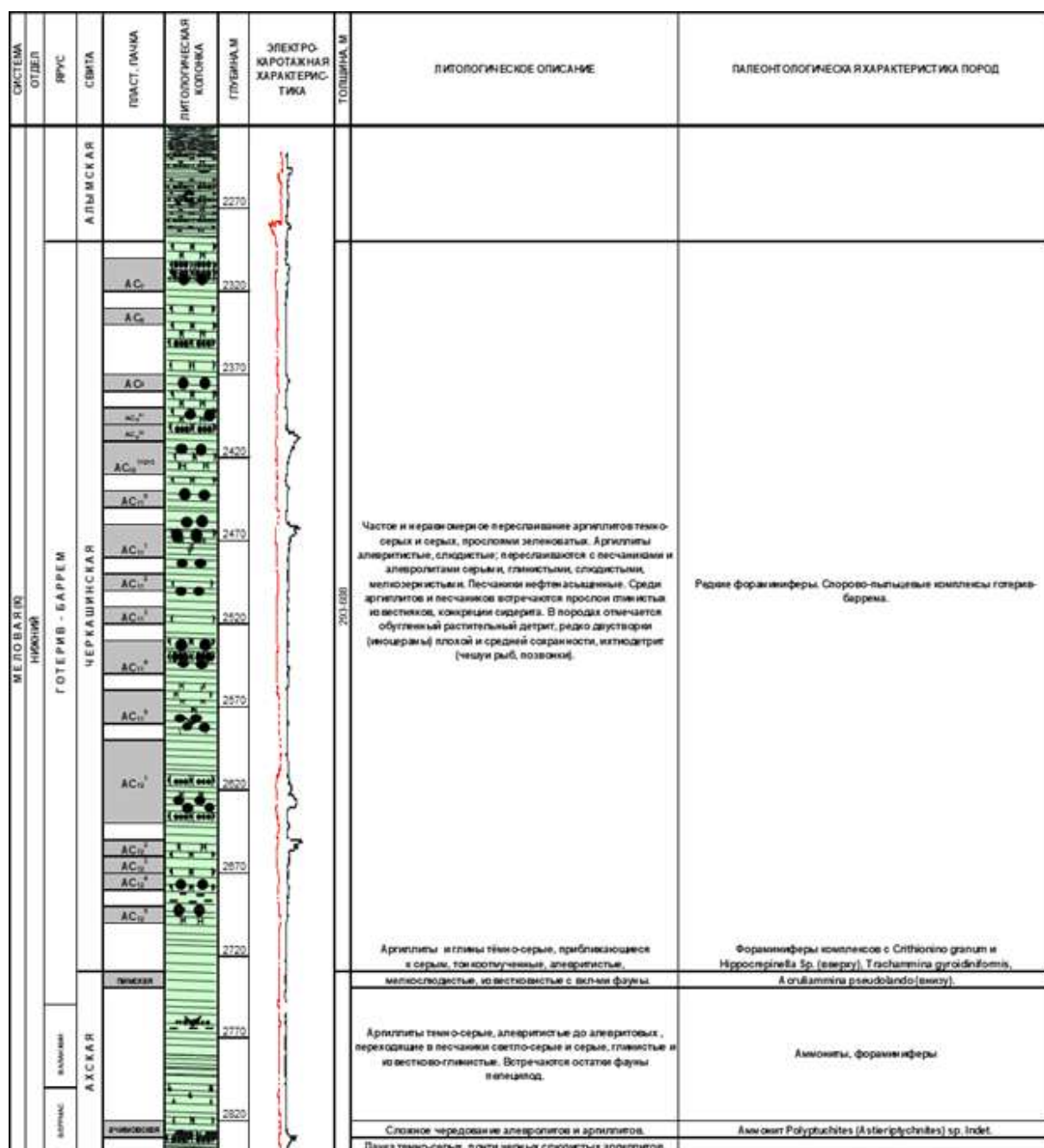


Рисунок 2.2 – Фрагмент стратиграфического разреза Приобского месторождения

Черкашинская свита включает в себя основные продуктивные горизонты объекта исследования. Свита представлена частым переслаиванием серых мелкозернистых песчано-алевритовых пород, местами слюдистых или известковых, и темно-серых, местами зеленоватых и тонкоотмученных, аргиллитов. Отмечаются прослой глинистых известняков, конкреции сидерита. Редко встречаются двустворки (иноцерамы), чешуя и позвонки рыб.

Песчано-алевритовые осадки группируются в пласты группы АС, имеющие существенную неоднородность в направлении простирания, которая выражается в частом переслаивании с глинистыми разностями пород. Песчаные горизонты АС₇-АС₁₂ имеют клиноформный характер, постепенно смещаясь снизу вверх с востока на запад, и являются промышленно нефтеносными. Мощность свиты достигает 610 м.

Алымская свита представлена отложениями темно-серых до черных глин. В верхней части свиты отмечаются прослой битуминозных аргиллитов, а в нижней – алевролитов и светло-серых мелкозернистых песчаников. По всему разрезу распространены прослой глинизированных известняков. Мощность свиты варьируется от 190 до 240 м.

Викуловская свита разделена на нижнюю и верхнюю подсвиты, с преобладанием соответственно глинистых и песчано-алевролитовых пород. Нижняя подсвита сложена аргиллитами и темно-серыми глинами, постепенно замещаясь вверх по разрезу прослоями алевритистых глин и глинистыми алевролитами. Верхняя подсвита, напротив, представлена преимущественно серыми мелкозернистыми алевролитами и песчаниками. Характерной особенностью разреза викуловской свиты являются обильные включения растительного детрита. Мощность свиты на исследуемой территории постепенно уменьшается от 264 до 200 м с запада на северо-восток.

Ханты-Мансийская свита представлена неравномерным переслаиванием песчаных и глинистых пород. Преобладающие в нижней части темно-серые плотные глины включают в себя прослой глинистых известняков и сидеритов. В верхней части напротив, в большей степени развиты алевролиты и светло-серые песчаники, иногда слюдистые с прослоями глин. Для пород свиты характерно обилие углистого детрита. Мощность свиты варьируется от 292 до 306 м.

2.3. Тектоническое строение района работ

В региональном тектоническом плане Приобское месторождение расположено во Фроловской впадине Западно-Сибирской плиты между Сырьегайской террасой на западе и Тундринской седловиной на востоке (Рисунок 2.3). Западно-Сибирская плита представляет собой молодой комплекс в виде зоны прогибания, чья толщина возрастает к центру вплоть до 9 км (Шпильман, 1998). В ее разрезе снизу-вверх могут быть выделены три структурных этажа, от складчатого допалеозойского возраста, к промежуточному параплатформенному и осадочному мезозойско-кайнозойского возраста.

Амплитуда структур антиклинальной формы нижнего горизонта в чехле достигает 150 м и сокращается вверх по разрезу до их полного затухания в отложениях неокомского возраста. Таким образом, в формировании отложений юрского возраста первостепенную роль играла тектоника.

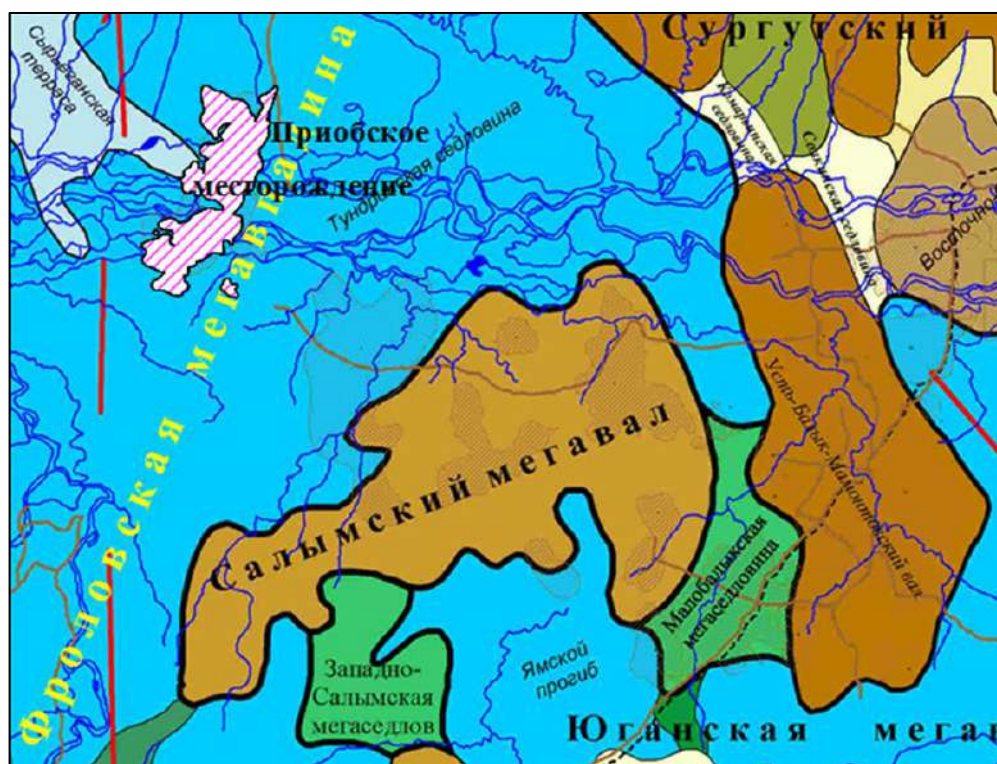


Рисунок 2.3 – Выкопировка из тектонической карты центральной части Западно-Сибирской плиты (под ред. В.И. Шпильмана, 1998)

Современный структурный план доюрского основания изучен по отражающему горизонту «А». На структурной карте по отражающему горизонту «А» находят отображение все структурные элементы. Современный структурный план юрских отложений изучен по горизонтам «Б» и «Т2». Наиболее достоверной и информативной картой является структурная карта по опорному отражающему горизонту «Б».

По отражающему горизонту «Дб», приуроченному к кровле быстринской пачки, прослеживаются Приобское куполовидное поднятие, Западно-Приобское малоамплитудное поднятие, Западно-Сахалинская, Новообская структуры. Все вышеуказанные особенности подтверждаются при изучении структурных карт по различным горизонтам. По кровле березовской, ганькинской и талицкой свит Приобская структурная зона представляет собой пологую моноклиналь, погруженную в восточном направлении. Вверх по разрезу отмечается значительное выполаживание локальных поднятий и депрессий, но их размеры сохраняются или изменяются незначительно.

2.4. Промышленная нефтегазоносность

В разрезе Южной лицензионной территории Приобского месторождения выделяются доюрский, ниже- и среднеюрский, баженовский, ачимовский, неокомский и апт-сеноманский нефтегазоносные комплексы. Основными объектами промышленной эксплуатации месторождения является группа пластов АС₇-АС₁₂ черкашинской свиты, относящейся к неокомскому комплексу. Основными по запасам нефти являются продуктивные горизонты АС₁₀ и АС₁₂, представляющие особый интерес и являющиеся объектами настоящего исследования; на долю пластов АС₇-АС₉ и АС₁₁ приходится около 1 % запасов, поэтому в работе они не рассмотрены. Промышленная нефтеносность также установлена в отложениях тюменской (пласты ЮС₂, ЮС₃) и баженовской свит (пласт ЮС₀).

Продуктивные горизонты нижнемелового возраста характеризуются сложным клиноформным строением и состоят из частого переслаивания маломощных пропластков песчаников и глин. Формирование клиноформных тел происходило в результате сноса кластического материала в западном и северо-западном направлениях. Схема клиноформенного разреза ЮЛТ Приобского месторождения по разрезу широтного профиля приведена на Рисунке 2.4.

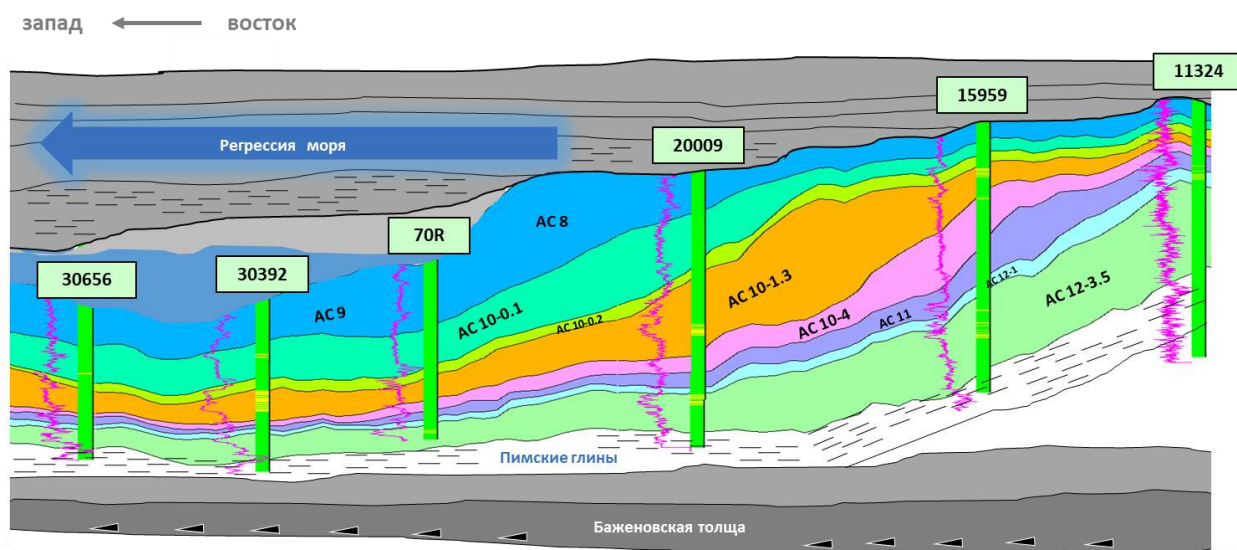


Рисунок 2.4 – Схема клиноформенного разреза ЮЛТ Приобского месторождения по разрезу широтного профиля

Определение морфологических особенностей строения пластов ведет к пониманию их фациального и литологического состава, а также территориальных ограничений. Исходя из того, что рассматриваемый разрез формировался в несколько этапов, территория месторождения может быть условно поделена на зоны, для каждого из которых характерны определенные особенности морфологии песчаных тел. Помимо описанных выше клиноформ, приуроченных к склоновой части, в разрезе могут быть выделены отложения глубоководного шельфа (ундаформная зона) и мелководья (фондоформа), представленные на Рисунке 2.5. Перечисленные типы песчаных тел различаются с точки зрения площадного распространения, степени изменчивости, а также фильтрационно-емкостных свойств и продуктивности в целом. Разделами между

горизонтами служат хорошо трассируемые в геофизических разрезах и данных сейсморазведки реперные глины. Идентификация именно этих глинистых пачек послужит основой для автоматической межскважинной корреляции нижнемеловых отложений.

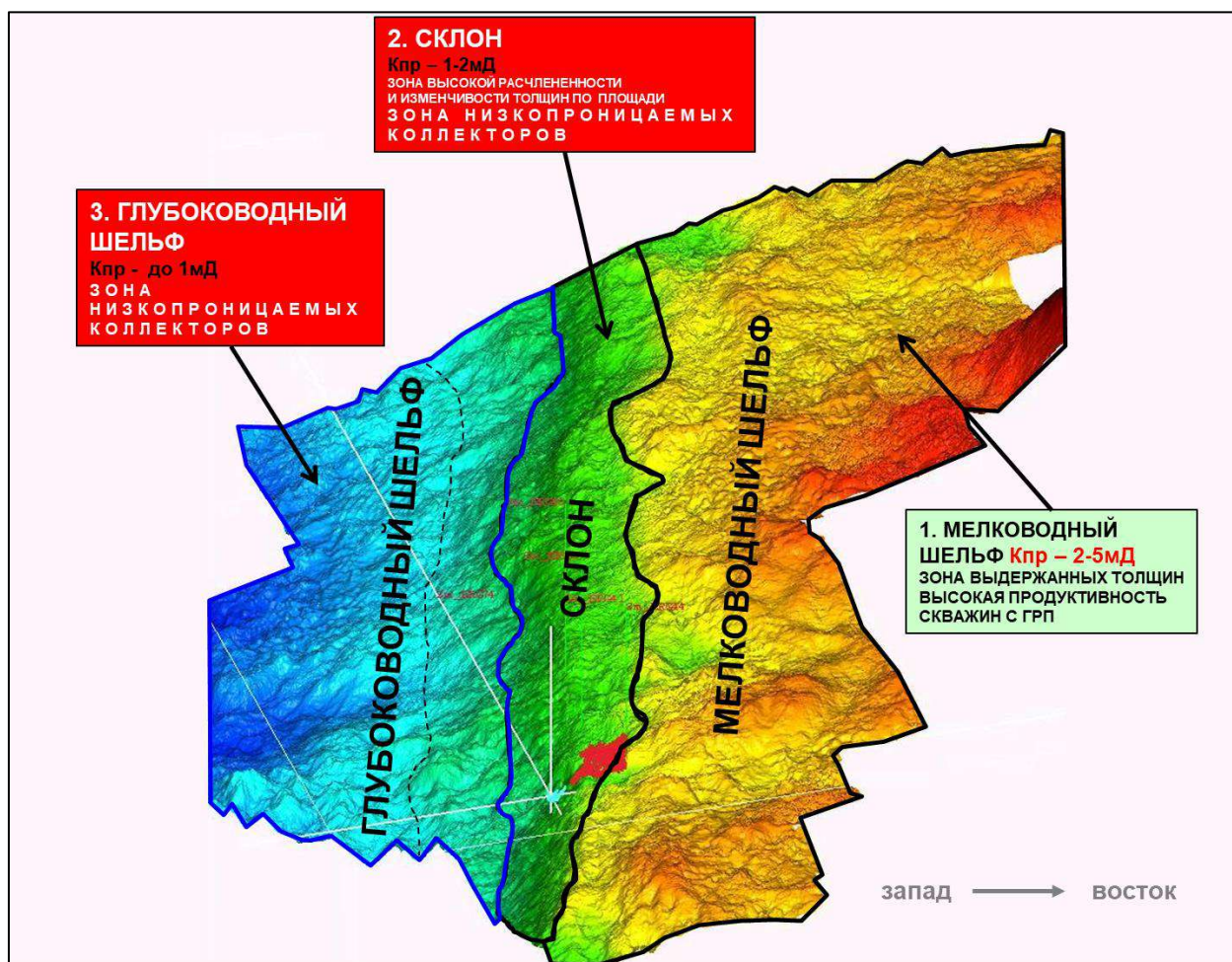


Рисунок 2.5 – Основные геоморфологические элементы ЮЛТ Приобского месторождения

Сводная геолого-физическая характеристика нижнемеловых пластов ЮЛТ Приобского месторождения приведена в Таблице 2.1.

Таблица 2.1 – Сводная характеристика нижнемеловых продуктивных пластов Приобского месторождения

№ п/п	Параметры	Пласты					
		АС ₇	АС ₈	АС ₉	АС ₁₀	АС ₁₁	АС ₁₂
1	Средняя глубина залегания, м	2234	2278	2339	2402	2481	2546
2	Тип залежи	структ.	литологически экранированная				
3	Тип коллектора	терригенный					
4	Средняя общая толщина, м	45	58,6	50,1	135	34,8	81,1
5	Средняя эффективная нефтенасыщенная толщина, м	4,2	1,2	2,6	6,5	2,3	6,5
6	Коэффициент пористости, д. ед.	0,18	0,18	0,17	0,17	0,17	0,17
7	Коэффициент нефтенасыщенности пласта, д. ед.	0,56	0,56	0,57	0,66	0,62	0,61
8	Проницаемость, 10 ⁻³ мкм ²	4,2	4,2	4	5,7	1,6	3,3
9	Коэффициент песчанистости разреза, д. ед.	0,08	0,03	0,06	0,08	0,04	0,12
10	Коэффициент расчлененности разреза, д. ед.	2,6	1,1	2,4	6,4	11,1	8,2
11	Начальная пластовая температура, °С	87	87	90	90	92	92
12	Начальное пластовое давление, МПа	—	—	23	25,7	26,7	27,2
13	Коэффициент продуктивности, т/сут·МПа	2,3	0,7	0,5	3,6	1,9	2,7
14	Коэффициент вытеснения, д. ед.	0,419	0,419	0,419	0,548	0,475	0,475

3. Теоретическое обоснование предлагаемой методики

3.1. Определение оптимальной последовательности расчленения

Строение осадочной формации тесно связано с процессом расчленения и выравнивания областей денудации, а также седиментацией отложений в условиях определённого климата и направленности тектонических движений. Последовательность осадконакопления в пределах профиля седиментации представлена на Рисунке 3.1.

На начальном этапе тектонических преобразований, в зависимости от рассматриваемой геотектонической зоны, формируются области денудации и осадконакопления, в пределах которой выделяются ряд геоморфологических единиц рельефа (Рисунок 3.1 А). Климатическая составляющая контролирует масштабы осадконакопления. В гумидном климате в процесс перераспределения обломочного материала вовлекаются как области сноса, так и области осадконакопления (Рисунок 3.1 Б).

В аридном климате в пределах морского бассейна и приливно-отливной равнины преобладает хемогенная седиментация, в то время как на континенте формируются эоловые осадки и отложения временных потоков речных систем. В ряде случаев аридные области могут дренироваться крупными речными системами, формируя «оазисы» седиментационных обстановок. Перераспределения обломочного материала из области денудации в область осадконакопления прекращается после формирования профиля выравнивания (Рисунок 3.1 Б). При этом значительное снижение поступающего в осадочный морской бассейн грубообломочного материала способствует смещению зоны седиментации глин из удалённой от берега области в прибрежную (Буш, 1977).

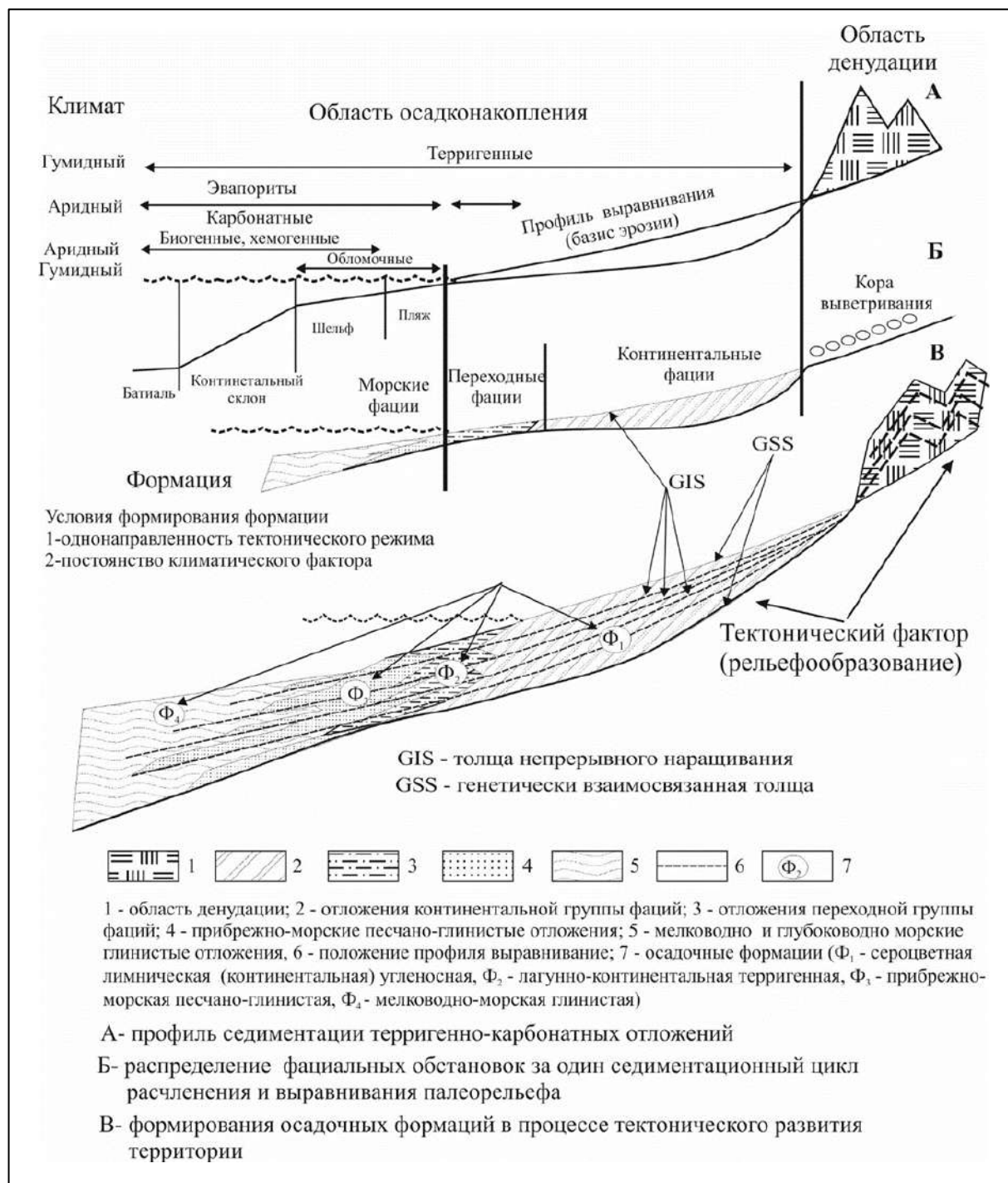


Рисунок 3.1 –Условия и последовательность формирования осадочных формаций (Буш, 1977)

Наиболее активно это перераспределение происходит в условиях жаркого тропического и умеренного гумидного климата с интенсивным проявлением денудационных процессов. В случае отсутствия тектонических подвижек после формирования профиля выравнивания, в области денудации физическое выветривание сменяется химическим с образованием коры выветривания, тип

которой зависит от вещественного состава пород фундамента и профиля выветривания породы.

Это способствует с одной стороны значительному снижению поступления грубообломочного терригенного материала и массовому выносу продуктов выветривания на уровне химических элементов (железо, кремний, и др.) водными потоками в область осадконакопления, а с другой – концентрации продуктов выветривания в преимущественно глинистых отложениях континентальной равнины и морской акватории. Формирование в осадочном разрезе специфических литологических разностей, образованных в условиях минимальных тектонических преобразований и максимального выравнивания профиля седиментации, позволяет выделять в осадочном разрезе закрытых территорий изохронные реперные горизонты (угольные пласты, бентонитовые и монтмориллонитовые глины, аргиллиты, обогащённые железом горизонты и т. д.).

В условиях аридного климата, где основное осадконакопление нефтеперспективных толщ происходит в морском бассейне и связано с карбонатно-терригенными породами, в зависимости от соотношения карбонатной и терригенной составляющей в качестве реперных горизонтов могут быть выделены либо глинистые, либо карбонатные пачки, характеризующие изменение тектоно-климатического режима. Реперные горизонты являются основой для расчленения и территориальной корреляция разрезов формации по данным бурения скважин с целью проведения палеогеографических реконструкций и выделения нефтеперспективных объектов.

При последующем импульсе тектонических однонаправленных преобразований, который преобразует палеорельеф области седиментации, этот процесс повторяется и если климат в области осадконакопления в течение цикла не изменялся, то формируется определённая осадочная формация. В процессе пространственного перемещения фациальных обстановок, связанного с особенностями тектонического развития геоструктурной зоны, в разрезе

образуется последовательность территориально «скользящих» во времени, однотипных по литологическому составу осадочных толщ (формаций), в каждой из которых присутствует своя совокупность реперных горизонтов, свойственных только этой формации.

Пространственную связь формаций можно изучать на основе выделения в разрезе толщ непрерывного наращивания и генетически взаимосвязанных толщ (Буш, 1977). Под первым термином понимается осадочная толща, представляющая собой один седиментационный цикл, в котором литологические компоненты фациальных обстановок в генетическом и пространственном отношении связаны друг с другом (Рисунок 3.1 Б). Генетически взаимосвязанная толща, в свою очередь, представляет сочетание двух или большего количества смежных толщ непрерывного наращивания мощности, сформированных в определённую фазу тектонического цикла осадконакопления. Угловых несогласий с генетически взаимосвязанной толщей быть не может, однако могут иметь место стратиграфические несогласия ограниченного распространения.

Толща непрерывного наращивания может быть сопоставима с литолого-фациальной картой, а генетически взаимосвязанная толща с комплектом литолого-фациальных карт, отражающих процесс пространственного перемещения фациальных обстановок (формаций) во времени.

Таким образом, в основе расчленения терригенных отложений лежит выделение в разрезе литологических элементов, характеризующих профиль выравнивания территории, который соответствует мобилизации продуктов выветривания области денудации (железо, кремнезём, глинистая фракция и т. д.) и переотложению этих элементов в пределы континентальной равнины и моря.

Основные продуктивные горизонты ЮЛТ Приобского месторождения, исходя из приведенного выше описания, принадлежат к отложениям нижнемелового возраста, а по формационным характеристикам относятся к морской терригенно-известковой олигомиктовой субформации (Гольберт, 1968 г.). В подобных отложениях роль маркирующих (реперных) горизонтов, как правило, выполняют пачки низкоомных аргиллитов, хорошо выделяемые по

индукционному каротажу. Высокие значения проводимости этих аргиллитов могут быть связаны с повышенным содержанием в них продуктов выветривания алюмосиликатов (например, монтмориллонита), формирование которых свойственно удалённым от берега зонам морского бассейна, а также более близким к берегу зонам седиментационного профиля в случае повышенных концентраций поступления в осадочный бассейн при выветривании вулканических пород, таких как туфы и пеплы (Арсланов, 1973 г.).

Учитывая, что этапы выравнивания характеризуются накоплением преимущественно глинистой фракции, особый интерес представляет углублённый анализ глинистых пачек и их неоднозначная характеристика по данным ГИС, которые в последующем могут быть использованы как опорные горизонты в процессе корреляции пластов (Изотова, 1993 г.). Дополнительным признаком при выделении опорных горизонтов является включения монтмориллонита или других продуктов выветривания, наличие которых может быть косвенно установлено по повышенным значениям проводимости горизонта и подтверждены в ходе изучения кернового материала. Прослои карбонатизированных песчаников (уплотнённые прослои или «плотняки») в составе песчаных пластов также могут быть использованы в качестве реперных горизонтов при изучении особенностей строения продуктивных коллекторов.

Исходя из вышесказанного, предлагается следующая последовательность действий для осуществления литологического расчленения разреза по реперным горизонтам:

1. Выделение глинистых пачек.
2. Выделение аномалий ГИС в пределах выделенных пачек исходя из ее характера и величины.
3. Систематизация выделенных аномалий по ГИС и связи их с литологическими особенностями по керну.
4. Глубинная привязка аномалий в скважине.

3.2. Подбор оптимальной последовательности корреляции

Основа методов корреляции – выделение в разрезе и прослеживание по скважинам реперных горизонтов. Этими горизонтами могут быть сейсмические отражающие горизонты, геофизические аномалии, выделяемые в разрезах скважин по данным ГИС, породы различного литологического состава (уголь, известняк, аргиллит и т. д.), геохимические и петрографо-минералогические аномалии по керновым данным, а также породные прослои с тождественными фаунистическими остатками.

Методы корреляции предусматривают использование при расчленении разрезов скважин существующий геолого-геофизический комплекс исследования осадочных толщ. Выделение и пространственное прослеживание толщ подразделяется на следующие виды (Арсланов, 1973):

1. Биостратиграфическая – по сходным комплексам органических остатков.
2. Сейсмостратиграфическая – между устойчивыми в сейсмическом разрезе отражающими горизонтами.
3. Стратиграфическая корреляция – по наличию реперных (маркирующих) горизонтов, характерной последовательности напластования в разрезе, цикличности, возрастных определений. Рассматриваемый тип корреляции может проводиться как по литологическому составу пород, так и совместно с геофизической корреляцией, в случае литологической интерпретации по комплексу каротажных диаграмм.
4. Геохимическая корреляция – на основе изучения и миграции химических элементов. Геохимическая корреляция позволяет установить связь тектонического развития между областью седиментации и региональными источниками сноса. Периоды выравнивания областей размыва и осадконакопления характеризуются формированием в зонах денудации коры выветривания, продуктами миграции которой являются химические элементы. Вынос этих элементов в бассейн седиментации и их накопление в

осадочной толщи позволяет проводить корреляцию по характеру их распределения в разрезе.

5. Геофизическая корреляция – по каротажной информации и палеомагнитным данным. В процессе геофизической корреляции проводится сопоставление толщ со схожей каротажной характеристикой, либо участков разреза, отражающих однотипный характер изменения полярности палеомагнитных исследований осадочной толщи.
6. Литологическая параллелизация пластов – по литологическому составу пород (песчаник, глина, карбонаты), их цвету и петрографическим особенностям. В условиях возможной интерпретации литологического состава осадочной толщи по данным каротажа скважин, литологическая параллелизация пластов может выполняться на основе геофизической информации.
7. Микропетрографическая корреляция – сопоставление разрезов на базе изучения микропетрографических особенностей слагающих их пород (гранулометрический, минералогический состав, топоморфические особенности минералов, распределение лёгких и тяжёлых фракций). Наиболее эффективно рассматриваемый вид корреляции может быть проведён на основе терригенно-минералогического анализа. Основное применение этот вид анализа получил при масштабных палеогеографических реконструкциях, где выявляется связь осадконакопления в седиментационном бассейне с областью денудации.

Исходя из того, что основными реперными горизонтами нижнемеловых отложений объекта исследования являются глинистые пропластки, а исходные данные включают в себя в основном результаты геофизических исследований и изучения кернового материала, наиболее подходящим методом корреляции горизонтов является литологическая параллелизация пластов, в первую очередь глинистых интервалов. Соответственно, была сформулирована следующая последовательность действий по межскважинной корреляции:

1. Определение одной или нескольких базовых (опорных) аномалий.
2. Увязка оставшихся выделенных аномалий относительно базовой.

3. Формирование стратиграфического каркаса для проведения корреляции геологических горизонтов на основе литологической параллелизации.

Таким образом, данные по литологическому расчленению разреза на осадочные пачки на базе выделения геофизических аномалий (реперных горизонтов) в глинистых отложениях служат основой для проведения стратиграфической корреляции межскважинного пространства с последующим ее уточнением методом литологической параллелизации пластов. Обоснованная с геологической точки зрения методология расчленения и корреляции горизонтов обобщена на Рисунке 3.2.



Рисунок 3.2 – Графическое представление предлагаемого алгоритма

3.3. Методология решения задач методами машинного обучения

Предложенные выше методики обозначают геологический подход к решению задач расчленения геологического разреза и корреляции выделенных горизонтов. С другой стороны, для того чтобы сделать поиск верного решения

задачи с помощью методов машинного обучения последовательным и прогнозируемым процессом, необходимо придерживаться четкой последовательности действий. Такая последовательность действий описывается в методологиях анализа данных, наиболее распространенной и эффективной из которых считается «CRISP-DM» (<https://ru.wikipedia.org/wiki/CRISP-DM>). Согласно данной методологии, аналитический проект, как правило, состоит из четырех этапов, которые должны выполняться последовательно:

1. Анализ данных – сбор, описание и оценка качества входных данных, их исследование.
2. Подготовка данных – включает в себя отбор необходимых данных, их очистку, интеграцию с остальным объемом информации и приведение их к формату, пригодному для создания модели.
3. Моделирование и оценка результата – выбор алгоритмов, планирование тестирования и обучения моделей, выбор оптимального решения; в случае достижения неудовлетворительного результата необходимо вернуться к одному из предыдущих шагов.
4. Внедрение полученного решения.

В графическом виде описанная методика представлена на Рисунке 3.3.

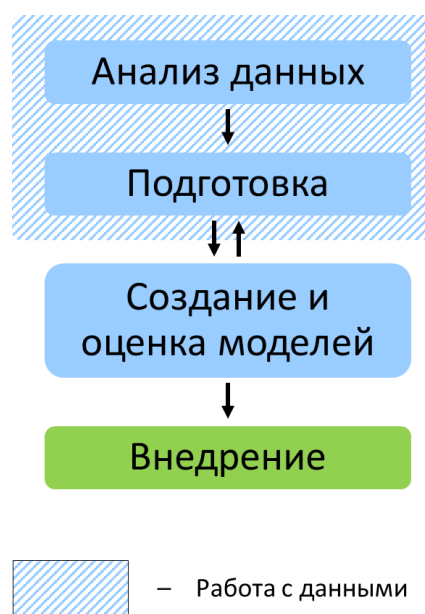


Рисунок 3.3 – Диаграмма связи процессов в методологии CRISP-DM

Описываемая методология обладает рядом явных преимуществ, таких как гибкость в последовательности действий (некоторые шаги можно пропускать или возвращаться к ним вновь после определенного этапа), ее вариативность и простота представления.

Таким образом, перед автором стояла задача объединить два вышеописанных алгоритма (геологическое обоснование расчленения и корреляции и статистическую методологию работы с данными) для создания автоматизированного алгоритма расчленения и корреляции геологических горизонтов. Совмещенные методологии представлены на Рисунке 3.4.

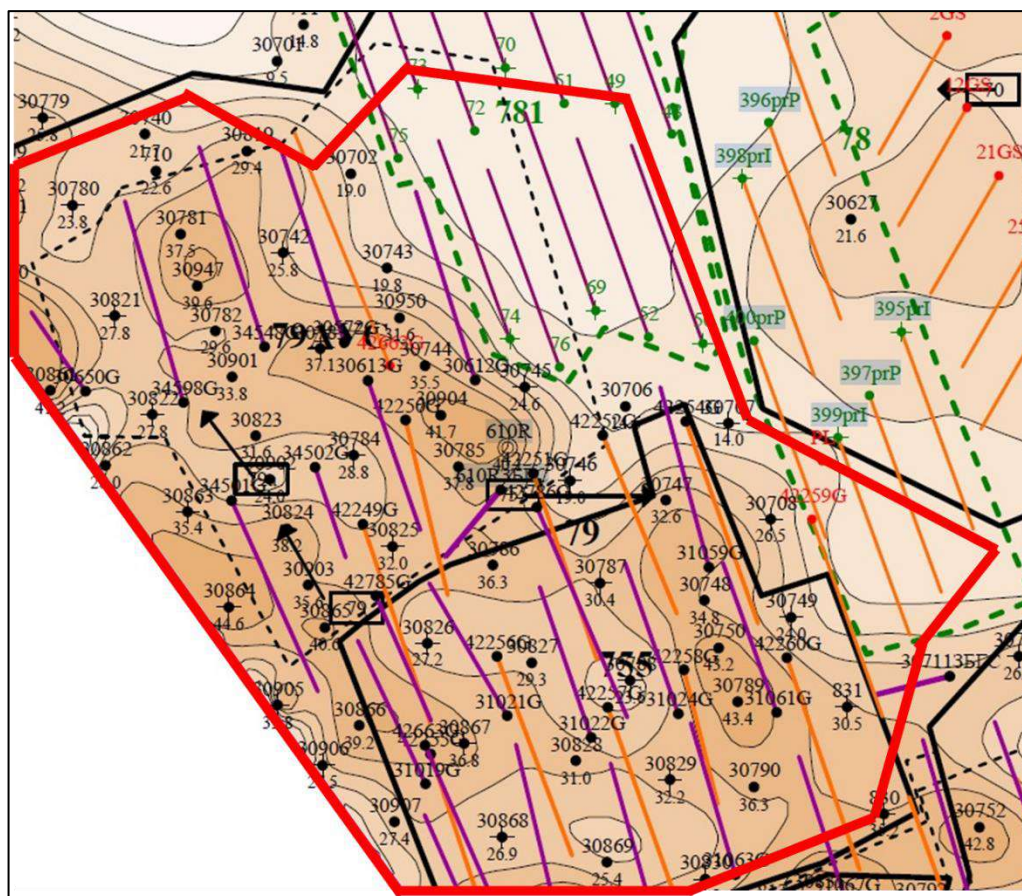


Рисунок 3.4 – Итоговая схема автоматизированного алгоритма расчленения и корреляции геологических горизонтов

4. Расчетная часть

4.1. Подготовка исходных данных

Подготовка и анализ входной информации являются одними из наиболее важных, но в то же время объемных в плане трудозатрат процессами. Первым шагом при построении любых моделей с использованием машинного обучения, как было отмечено выше, является обработка исходного материала. Сюда относится отбор необходимых данных, их очистка и интеграция с остальным объемом информации в единую матрицу объектов-признаков, а также его приведение к формату, приемлемому для использования в программном комплексе. В процессе формирования исходной модели будут использованы данные семи скважин кустовой площадки X, находящейся в юго-восточной части ЮЛТ Приобского месторождения (Рисунок 4.1).



Исходные данные были условно разделены на следующие группы:

1. Общая информация о скважине – порядковый номер, альтитуда и расположение на кустовой площадке.
2. Кривые геофизических исследований группами электрических и радиоактивных методов.
3. Литологическое расчленение разреза, выполненное интерпретатором.
4. Результаты лабораторных исследований, фотографии и описание кернового материала.
5. Карта распространения фаций на участке работ.

Для обеспечения однородности данных в первую очередь необходимо было произвести отбор общих для всех скважин геофизических методов. Итогом данного шага является сформированный набор геофизических методов, использовавшийся для построения моделей. Отобранные методы и их расшифровки приведены на Рисунке 4.2.

Номер	Шифр прибора	Расшифровка метода	Тип
1	BK	Боковой каротаж	Электрический
2	IK	Индукционный каротаж	
3	GZ1-GZ5	Группа последовательных градиент-зондов	
4	OGZ	Обращенный градиент-зонд	
5	PS	Потенциал собственной поляризации	
6	PZ	Потенциал-зонд	
7	RES	Токовая резистивиметрия	
8	RPO	Удельное сопротивление пород	
9	F05-F20	Высокочастотное индукционное каротажное изопараметрическое зондирование (соответственно фазовые и удельные электрические сопротивления)	
10	R05-R20		
11	GK	Естественная гамма-активность	Радиоактивный
12	MZ	Нейтрон-нейтронный каротаж по тепловым нейтронам (м)	
13	NKT	Нейтрон-нейтронный каротаж по тепловым нейтронам (б)	
14	W	Водородосодержание	

Рисунок 4.2 – Матрица скважин и записанных в них методов ГИС

Следующий шаг – добавление в датафрейм информации о литологическом расчленении разреза. Исходная информация была представлена в виде графического отображения геофизических кривых продуктивной части резервуара, по информации которых было произведено расчленение разреза каждой из скважин на литотипы. Пример предоставленного литологического расчленения скважин приведен в Приложении Б.

Ввиду того, что использование визуального представления расчленения для построения модели невозможно, было принято решение разделить разрезы скважин на пропластки толщиной 0.2 м (исходя из шага дискретизации геофизической информации) для того, чтобы в дальнейшем отнести каждый пропласток к одному из 13 представленных литотипов. Названия литотипов, выбранная для их отображения визуальная гамма и сокращения для корректного восприятия средой программирования приведены на Рисунке 4.3. Визуальное представление основных геофизических кривых и расчленение разреза скважины представлено на Рисунке 4.4.

Номер	Литотип	Российское сокращение	Программное сокращение
1	Аргиллит	Арг	Arg
2	Аргиллит кавернозный	АргКав	ArgCav
3	Аргиллит битуминозный	АргБит	ArgBit
4	Аргиллит углистый	АргУгл	ArgC
5	Алеврит	Алевр	Silt
6	Песчаник	Песч	Sand
7	Песчаник на глинистом цементе	ПесчГЦ	SandShaly
8	Песчаник с прослоями глин	ПесчПГ	SandArg
9	Песчаник слабопроницаемый	ПесчСП	SandImp
10	Песчаник уплотненный	ПесчПл	SandComp
11	Плотняки	Уплотн	Tight
12	Уголь	Уголь	Coal
13	Не определено	Неизв	Un

Рисунок 4.3 – Представленные литотипы, их обозначения и сокращения

Скважина 30744

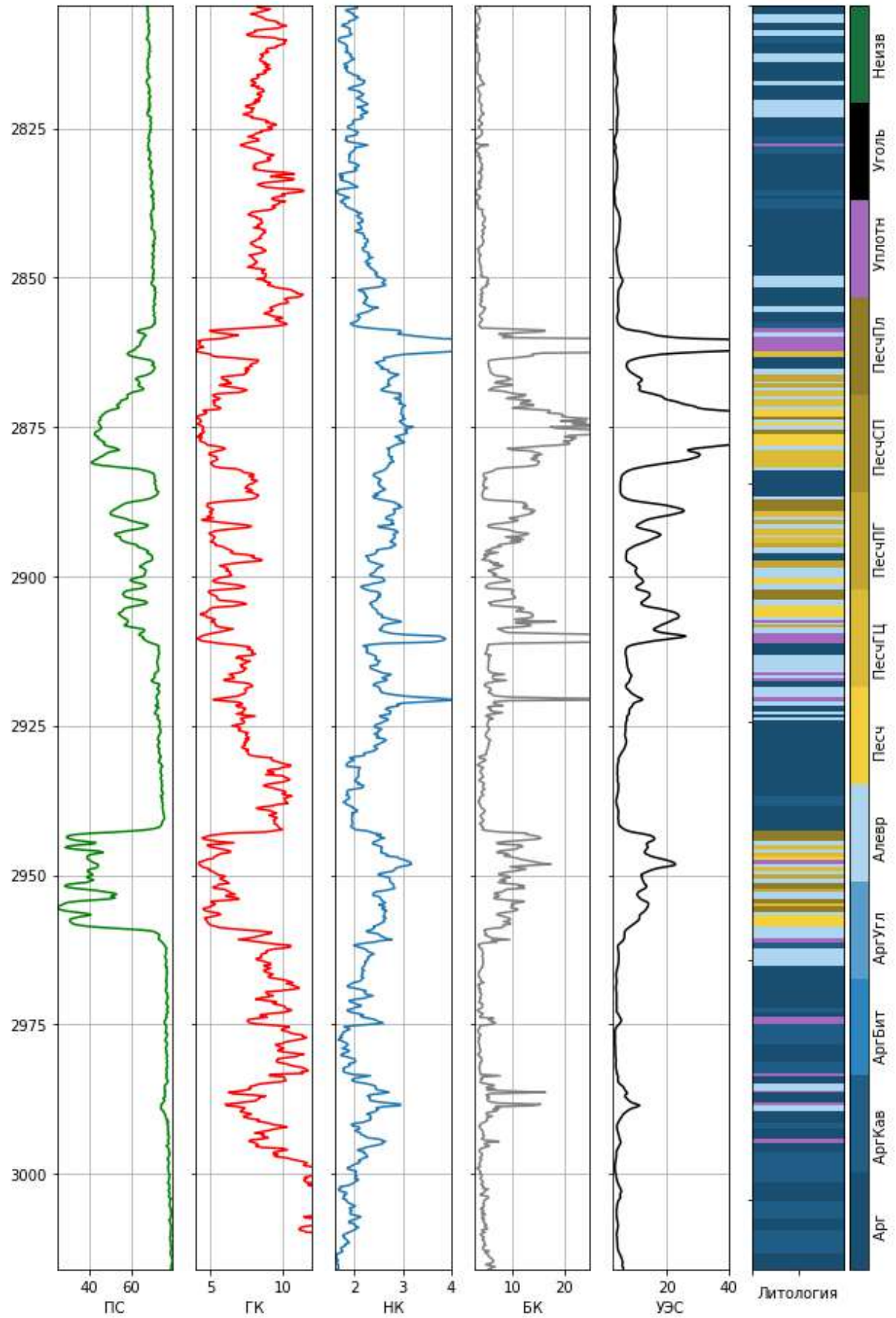


Рисунок 4.4 – Кривые методов ГИС и расчленение разреза скважины 30744

В результате выполнения данного шага числа входных объектов возросло до 7255. В признаковом описании каждого объекта номер скважины был заменен на принадлежность к определенному литологическому классу, а удельные электрические сопротивления ВИКИЗ были исключены в ходе факторного анализа, поэтому общее число признаков осталось равно 23 (Рисунок 4.5).

In [5]:

```
data = data.drop(['Well', 'R05', 'R07', 'R10', 'R14', 'R20'], axis = 1)
data.head()
```

Out [5]:

	Depth	BK	IK	GZ1	GZ2	GZ3	GZ4	GZ5	OGZ	PS	...	F05	F07	F10	F14	F20	Lith
0	3217.2	3.77	229	4.51	3.43	2.90	2.90	2.60	3.90	59.498	...	25.649	24.410	23.88	23.550	23.399	1
1	3217.4	3.61	225	4.95	3.36	2.86	2.77	2.65	3.60	59.400	...	25.649	24.519	23.98	23.560	23.220	1
2	3217.6	3.67	222	4.91	3.48	2.87	2.63	2.49	3.56	59.699	...	25.500	24.220	23.73	23.369	23.040	1
3	3217.8	3.78	220	4.87	3.68	2.80	2.67	2.64	3.73	59.199	...	25.160	23.740	23.19	22.809	22.510	1
4	3218.0	4.04	221	5.14	4.50	3.02	2.88	2.59	3.93	59.400	...	24.520	23.249	22.60	22.200	21.939	11

Рисунок 4.5 – Визуальное представление описания первых пяти объектов

Для выполнения дальнейших операций с данными необходимо было удалить объекты, признаковое описание которых не содержало численного значения. Подобная ситуация встречается в случаях, когда каротаж какого-либо метода был записан не на всю мощность продуктивной части разреза, либо запись началась с абсолютной отметки, лежащей ниже кровли продуктивного интервала. Использование функции «*dropna*» позволило удалить 26 объектов, содержащий нулевые значения признаков, в результате чего общее количество объектов сократилось до 7229.

Исходя из особенностей работы алгоритмов машинного обучения при решении задач «с учителем», входная информация должна быть разделена на обучающую, валидационную и тестовую выборки. Кратко он может быть описан следующим образом: в качестве входной информации алгоритму подается обучающая выборка, которая используется для выявления математических зависимостей между признаками (а также для оптимизации гиперпараметров модели). Полученные зависимости в дальнейшем распространяется на объекты

валидационной выборки для определения значений ключевого признака (задача регрессии), деления объектов на классы (задача классификации) и т. д. Оценка точности построенной модели и ее связей производится на основе объектов тестовой выборки (Абдуллаев, 2021).

Для задачи классификации данные необходимо разделить в таких пропорциях, чтобы в полученных наборах соотношение объектов разных классов было примерно таким же, как и в исходном материале. Кроме того, обучающую выборку ни в коем случае нельзя совмещать с оценочными, т. к. в этом случае модели может грозить переобучение. Строгих требований по размерам каждой из выборок не существует; как правило, размер тестовой выборки может составлять не более 20 % от всего объема исходных данных. Для использования в качестве тестовой была выбрана разведочная скважина 610P исходя из наличия в ней кернового материала, который будет служить дополнительным фактором при оценке точности полученного расчленения. Таким образом, из массива исходных данных были исключены 1224 объекта. Оставшийся объем информации, как правило, делится в пропорции 80 на 20 для обучающей и тестовой выборок соответственно. Для исключения нежелательной предвзятости, когда в выборке могут быть подряд представлены объекты из одной скважины с одинаковой литологией, разбиение данных производилось случайным образом при использовании функции *«train_test_split»* из библиотеки *«Scikit-learn»*. Таким образом, количество объектов в обучающей выборке составило 4804, а в валидационной – 1201.

Необходимо отметить, что каждый из 23 признаков, описывающих объекты, имеет свою размерность и граничные значения. Таким образом, после группировки полученной информации в общий датафрейм, признаковое описание объектов необходимо подвергнуть нормализации. Нормализацией (нормированием) данных в статистике называется приведение признаков значений объекта в модели, измеренных в разных диапазонах (в данном случае – различных геофизических методов), к общей условно шкале. Нормализация значений различных признаков, описывающих объект, предоставляет

возможность сравнивать их между собой, определяя таким образом распределения значений и позволяя выделить среди них аномальные, резко отличающиеся от соседних, что, в свою очередь, является одной из ключевых задач при решении задачи корреляции геологических горизонтов.

Нормализация данных языке программирования «Python» производится автоматически при использовании функции «*StandardScaler*» из библиотеки «*Scikit-learn*», поэтому не требует трудозатрат со стороны исследователя (Рисунок 4.6).

```
B [10]: > from sklearn.preprocessing import StandardScaler  
        scaler = StandardScaler()  
        X_train_scaled = scaler.fit_transform(X_train)  
        X_valid_scaled = scaler.transform(X_valid)
```

Рисунок 4.6 – Нормализация данных с помощью функции «*StandardScaler*»

Таким образом, в результате обработки исходных данных был получен датафрейм, включающий 7229 объектов, каждый из которых описан 23 признаками. Объекты сгруппированы в обучающую, валидационную и тестовую выборки, содержащие соответственно 4804, 1201 и 1224 объект.

4.2. Выделение в разрезе глинистых пачек

Следующим этапом является обучение моделей машинного обучения на подготовленных данных. Расчленение геологического разреза на литотипы является одним из частных вариантов задачи классификации в машинном обучении. Для решения задач подобного типа наиболее часто используются следующие методы:

1. Дерево решений (Decision Tree) – последовательное разделение выборки на две части до момента, когда в каждой части не останется объект одного класса; разделение производится по предикату (условию), который

выбирается из множества элементарных предикатов, на практике в качестве элементарных предикатов чаще всего берут простые пороговые условия на значение какого-либо признака.

2. Случайный лес (Random Forest) – объединение нескольких решающих деревьев в один ансамбль. В процессе построения генерируется M выборок размера N для каждого дерева, затем производится его независимое обучения, затем объектам основной выборке присваиваются классы на каждом из подпространств. В итоге решение о принадлежности объекта одному из классов принимается по большинству результатов.
3. Метод ближайших соседей (k Nearest Neighbors) – после определения расстояния до каждого объекта обучающей выборки и отбора из их числа k объектов с минимальным расстоянием класс объекта будет задан как наиболее часто встречающийся класс среди k соседей.
4. Метод опорных векторов (Support Vector Machine) – предполагает перевод исходных векторов в пространство более высокой размерности для поиска в ней разделяющей гиперплоскости, имеющей максимальный зазор с исходными векторами.

После импортирования соответствующих моделей из библиотеки Scikit-learn им передаются выборки, необходимые для обучения (Рисунок 4.7).

```
B [13]:  from sklearn.tree import DecisionTreeClassifier
         from sklearn.ensemble import RandomForestClassifier
         from sklearn.neighbors import KNeighborsClassifier
         from sklearn import svm

B [14]:  forest = RandomForestClassifier(random_state = 42)
         tree = DecisionTreeClassifier(random_state = 42)
         knn = KNeighborsClassifier(n_neighbors = 10)
         clf = svm.SVC()

         forest.fit(X_train_scaled, y_train)
         tree.fit(X_train_scaled, y_train)
         knn.fit(X_train_scaled, y_train)
         clf.fit(X_train_scaled, y_train)
```

Рисунок 4.7 – Процесс импортирования и настройки методов классификации

После процесса непосредственно обучения для каждого из использованных алгоритмов может быть предварительно оценена точность. Для этого с помощью валидационной выборки для каждого из объектов рассчитывается класс литологии, которая затем сравнивается с фактическими данными, полученными в ходе интерпретации геофизических исследований. Простейшим вариантом оценки точности является использование функции «*accuracy*», показывающей долю объектов с верно определенным классом литологии (Рисунок 4.8).

```
В [12]:  from sklearn.metrics import accuracy_score

         tree_pred = tree.predict(X_valid_scaled)
         forest_pred = forest.predict(X_valid_scaled)
         knn_pred = knn.predict(X_valid_scaled)
         clf_pred = clf.predict(X_valid_scaled)

В [13]:  print('Точность дерева решения: {:.3f}'.format(accuracy_score(y_valid, tree_pred)))
         print('Точность случайного леса: {:.3f}'.format(accuracy_score(y_valid, forest_pred)))
         print('Точность метода ближайших соседей: {:.3f}'.format(accuracy_score(y_valid, knn_pred)))
         print('Точность метода опорных векторов: {:.3f}'.format(accuracy_score(y_valid, clf_pred)))

Точность дерева решения: 0.719
Точность случайного леса: 0.820
Точность метода ближайших соседей: 0.754
Точность метода опорных векторов: 0.729
```

Рисунок 4.8 – Использование функции «*accuracy*» для оценки точности

Таким образом, наибольшую точность при проверке на валидационной выборке показал метод случайного леса, верно определив литологию у более чем 82 % объектов. Использование остальных методов также обеспечивает достаточно высокую точность (более 72 % литологии было установлено верно). Тем не менее, по-настоящему точность можно оценить только при использовании данных скважин, которые не участвовали в процессе построения модели. Для этого построенные моделями зависимости были распространены на данные скважины 610Р. Визуальное представление эталонного (произведенного интерпретатором) и машинного (выполненного с помощью алгоритмов) литологических расчленений разреза скважины, реализованное с помощью встроенных библиотек визуализации, представлено на Рисунке 4.9.

Скважина 610R

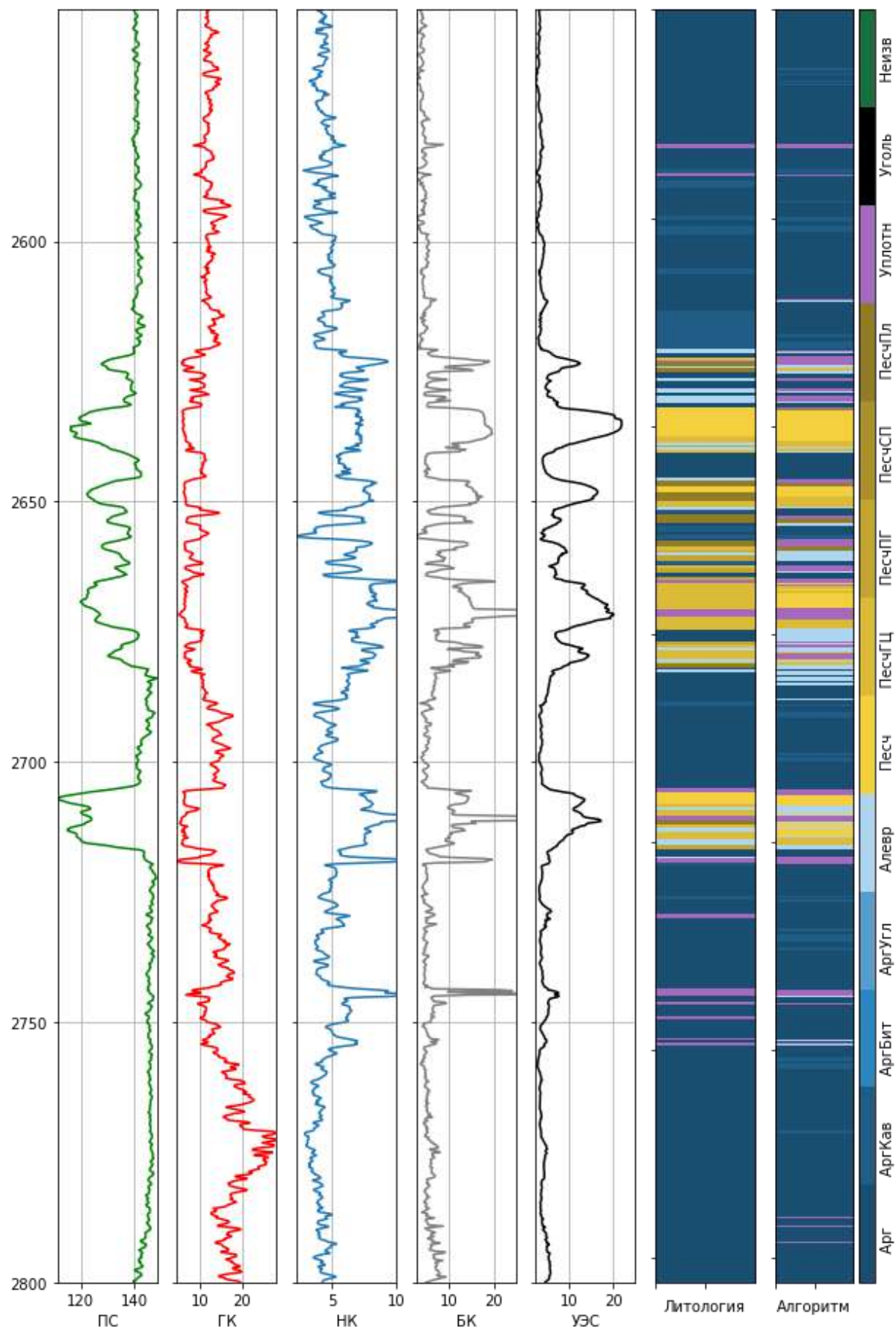


Рисунок 4.9 – Сравнение расчленений интерпретатора и алгоритма

Исходя из визуального сравнения расчленений, алгоритм достаточно точно выделяет аргиллиты всех типов, лишь изредка присваивая отдельным входящим в них пропласткам класс алевролитов. Песчаники в целом также выделяются достаточно уверенно; одновременно с этим, внутри самого класса был отмечен некоторый тренд к замене глинизированных песчаников на обыкновенные. Также стоит отметить значительно увеличившееся число пропластков уплотненных пород. Наблюдения подтверждаются построенными гистограммами распределения литотипов для обоих видов расчленения (Рисунок 4.10).

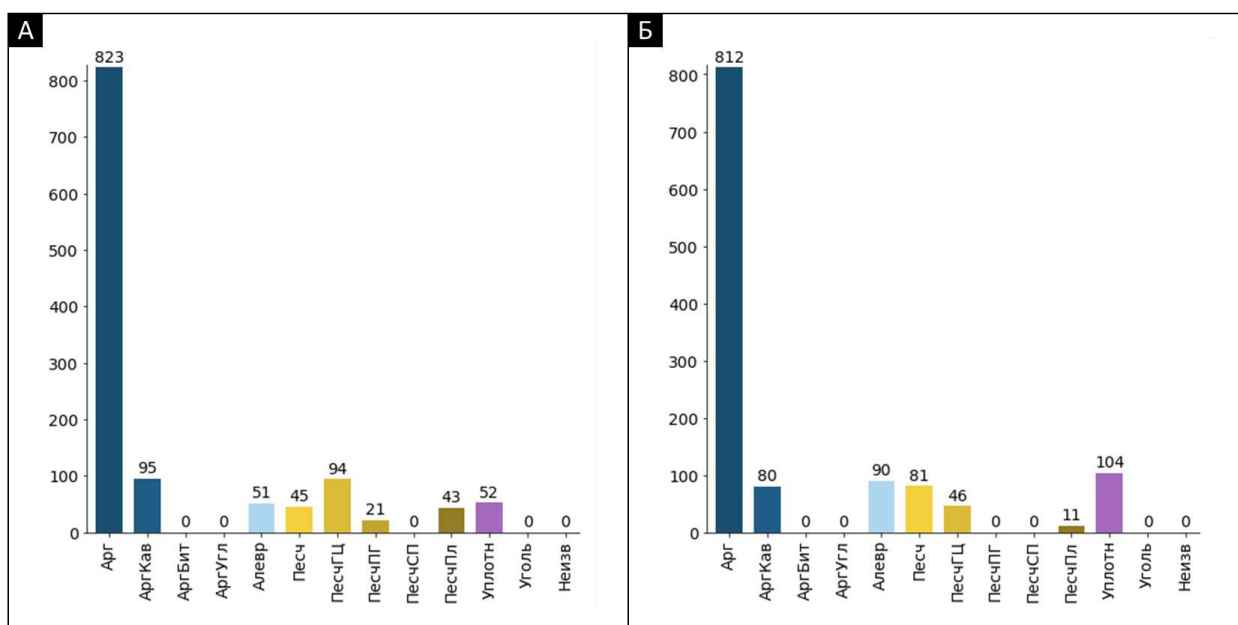


Рисунок 4.10 – Распределение литотипов скважины 610R

(А – эталонное расчленение разреза; Б – машинное расчленение)

Наибольшее расхождение с расчленением интерпретатора алгоритм демонстрирует в интервалах глубин 2630-2635 и 2650-2675 метров. Одновременно с этим, в результатах геофизических исследований скважины 610R некоторые методы демонстрируют завышенные показания зондов, из-за которых возникают неточности в процессе машинного определения литотипов. После ознакомления с фактического каменным материалом (керном), который

представлен 100 % выносом в интересующих интервалах, были сделаны следующие выводы:

1. В интервале глубин 2628-2635 метров песчаник имеет глинисто-карбонатный состав цементного материала, незначительно искажая показания геофизических зондов; предположительно, по этой причине алгоритм присвоил данному интервалу класс «уплотненные прослои», которые также характеризуются высоким содержанием карбонатного материала. Аналогичная ситуация наблюдается в интервалах 2654,9-2657,2 и 2659,9-2661 метров.
2. Интервал глубин 2651,5-2654,9 метров сложен преимущественно алевролитами с прослоем аргиллита в центре. Алгоритм, фактически, выделил пропластки точнее интерпретатора.
3. В интервале 2668,4-2669,5 метров отмечено наличие глауконита. Кроме того факта, что глауконит является полупроводником, достаточно сильно искажая показания геофизических зондов, его наличие свидетельствует о морских условиях седиментации. Включения глауконита также отмечаются в интервале глубин 2670,5-2671,8 метров. Исходя из его отсутствия в вышележащих горизонтах, можно предположить, что ближе к интервалу глубин 2668,4-2669,5 произошла смена условий осадконакопления.
4. Крайне часто в керновом материале продуктивной части отмечаются признаки нефти (запах и выпоты УВ), которая имеет повышенное сопротивление и, соответственно, пониженную проводимость электрического тока. Исходя из того, что большая часть представленных геофизических методов также имеет в своей основе электрическую природу, показания зондов могут быть искажены, в результате чего наблюдается некоторое несоответствие эталонного и машинного расчленения разреза.

Таким образом, некоторое несоответствие эталонного и машинного расчленения разреза может быть объяснено особенностями состава горных пород продуктивной части, которые могут незначительно искажать показания геофизических зондов, приводя к ошибкам в процессе интерпретации

алгоритмами. Более подробно ознакомиться с описанием кернового материала и его особенностями можно в Приложении В.

Незначительное ухудшение точности моделей при их использовании на данных отдельной скважины может свидетельствовать в первую очередь о недостатке исходного материала; одним из возможных последствий этого является отсутствие части представленных в скважине 610Р горизонтов в разрезах скважин, использовавшихся в процессе обучения. Кроме того, записанные в ходе геофизических исследований показания приборов для одной и той же группы пластов могут различаться в разных скважинах, порой достаточно сильно. Причинами подобных различий может служить изменения минерального состава разреза, условий и характера сноса осадков, ухудшения коллекторских свойств или характера насыщения перспективных пластов-коллекторов и т. д.

Первичная численная оценка точности произведенного алгоритмом расчленения была выполнена с помощью функции «*accuracy*» (Рисунок 4.11).

```
B [14]: test_data = pd.read_excel('test_data.xlsx')
test_data = test_data.dropna()
test_data = test_data.drop(test_data[test_data.Depth > 2800].index)

X_from_test = test_data.drop(['Well', 'R05', 'R07', 'R10', 'R14', 'R20', 'Lith'], axis = 1)

scaler = StandardScaler()
X_from_test = scaler.fit_transform(X_from_test)
y_from_test = test_data['Lith']

print('Точность дерева решения: {:.3f}'.format(accuracy_score(y_from_test,
                                                                tree.predict(X_from_test))))
print('Точность случайного леса: {:.3f}'.format(accuracy_score(y_from_test,
                                                                forest.predict(X_from_test))))
print('Точность метода ближайших соседей: {:.3f}'.format(accuracy_score(y_from_test,
                                                                knn.predict(X_from_test))))
print('Точность метода опорных векторов: {:.3f}'.format(accuracy_score(y_from_test,
                                                                clf.predict(X_from_test))))

Точность дерева решения: 0.517
Точность случайного леса: 0.732
Точность метода ближайших соседей: 0.673
Точность метода опорных векторов: 0.700
```

Рисунок 4.11 – Использование функции «*accuracy*» для оценки точности на данных тестовой выборки

Сравнение точностей полученных моделей на данных валидационной и тестовой выборок приведено в Таблице 4.1.

Таблица 4.1 – Сравнение точностей моделей при использовании данных валидационной и тестовой выборок

Алгоритм	Первичная точность	Точность на тестовой выборке	Относительное изменение
Decision tree	0,719	0,517	-28 %
Random forest	0,820	0,732	-11 %
kNN	0,754	0,673	-11 %
SVM	0,729	0,700	-4 %

Алгоритмы в целом демонстрируют незначительное снижение точности по сравнению с использованием данных валидационной выборки. Наибольшую долю верных ответов на отложенных данных, исходя из представленной выше таблицы, показал случайный лес и метод опорных векторов. Исходя из того, что увеличить объем входной информации не представляется возможным, было предложено произвести для вышеперечисленных методов оптимизацию гиперпараметров.

Гиперпараметрами называются параметры алгоритмов, значения которых устанавливаются перед началом запуска процесса обучения. Таким образом, гиперпараметры не входят в подборку исходных данных, но напрямую влияют и управляют процессом обучения. Примерами гиперпараметров являются крутизна активационной функции или число нейронов в нейронной сети, число создаваемых кластеров при использовании карт Кохонена в задачах кластеризации и т. д.

Для оценки изменения точности было решено произвести оптимизации гиперпараметров методов случайного леса и опорных векторов. Так, наибольший вклад в изменение точности алгоритма случайного леса вносят гиперпараметры «*n_estimators*» (число «деревьев» или ответвлений) и «*max_depth*» (максимальная глубина каждого из ответвлений). Процесс оптимизации числа «деревьев» представлен на Рисунке 4.12. Для оценки влияния

количества деревьев на точность были построены соответствующие графики, показанные на Рисунке 4.13.

```
В [17]: ► estimators = np.arange(1, 1000, 20)
scores = []
for n in estimators:
    forest.set_params(n_estimators = n)
    forest.fit(X_train, y_train)
    scores.append(forest.score(X_valid, y_valid))
plt.figure(figsize = (15, 10))
plt.title("Влияние количества деревьев на точность")
plt.xlabel("Число деревьев, шт")
plt.ylabel("Точность, д.е.")
plt.plot(estimators, scores);
```

Рисунок 4.12 – Процесс оптимизации гиперпараметра «*n_estimators*» для интервала значений от 1 до 1000 с шагом 20

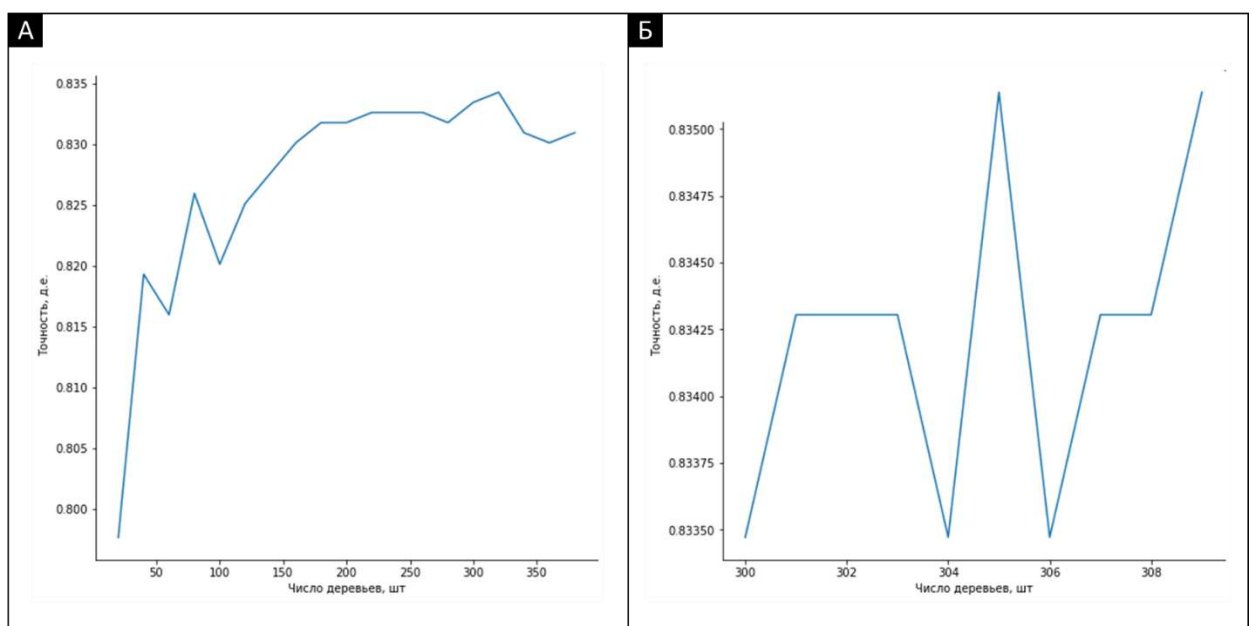


Рисунок 4.13 – Оценка влияния «*n_estimators*» на точность алгоритма (А – в интервале значений от 1 до 400; Б – в интервале значений, показавших наибольшую точность)

Аналогичная операция была проведена для гиперпараметра «*max_depth*». Процесс нахождения оптимального значения представлен на Рисунке 4.14.

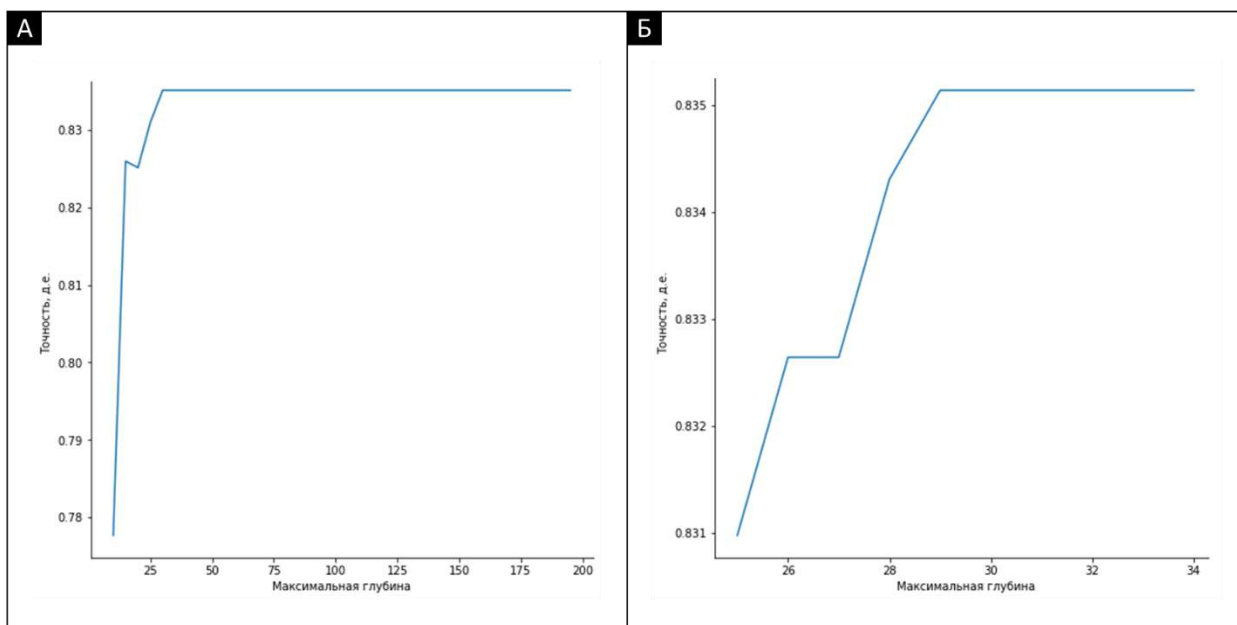


Рисунок 4.14 – Оценка влияния «*max_depth*» на точность алгоритма
(А – в интервале значений от 1 до 200; Б – в интервале выхода на прямую)

Таким образом, оптимальные значения гиперпараметров «*n_estimators*» и «*max_depth*» равны соответственно 305 и 30. Следующий шаг – применение оптимизированных параметров при построении моделей (Рисунок 4.15).

```
В [22]: forest_optimized = RandomForestClassifier(n_estimators = 305,
                                                random_state = 42,
                                                max_depth = 30)
forest_optimized.fit(X_train_scaled, y_train)
print('Точность случайного леса на валидационной выборке: {:.3f}'.
      format(accuracy_score(y_valid,
                             forest_optimized.predict(X_valid_scaled))))
print('Точность случайного леса на тестовой выборке: {:.3f}'.
      format(accuracy_score(y_from_test,
                             forest_optimized.predict(X_from_test))))
```

Точность случайного леса на валидационной выборке: 0.835
Точность случайного леса на тестовой выборке: 0.730

Рисунок 4.15 – Оценка точности алгоритма случайного леса при использовании оптимизированных гиперпараметров

Как видно из представленного выше рисунка, в ходе оптимизации гиперпараметров алгоритма случайного леса точность алгоритма незначительно снизилась, что свидетельствует о переобучении модели.

При оптимизации гиперпараметров метода опорных векторов возникла аналогичная ситуация: точность предсказания значительно возросла при использовании данных валидационной выборки, однако немного снизилась в случае тестовой скважины. Основными гиперпараметрами в данном алгоритме являются «с» (степень гибкости гиперплоскости) и «*gamma*» (расстояние элементов от разделяющей гиперплоскости). Оценка влияния данных параметров на точность приведен на графиках на Рисунке 4.16. Процесс применения оптимизированных параметров при построении моделей и переоценка точности метода опорных векторов представлен на Рисунке 4.17.

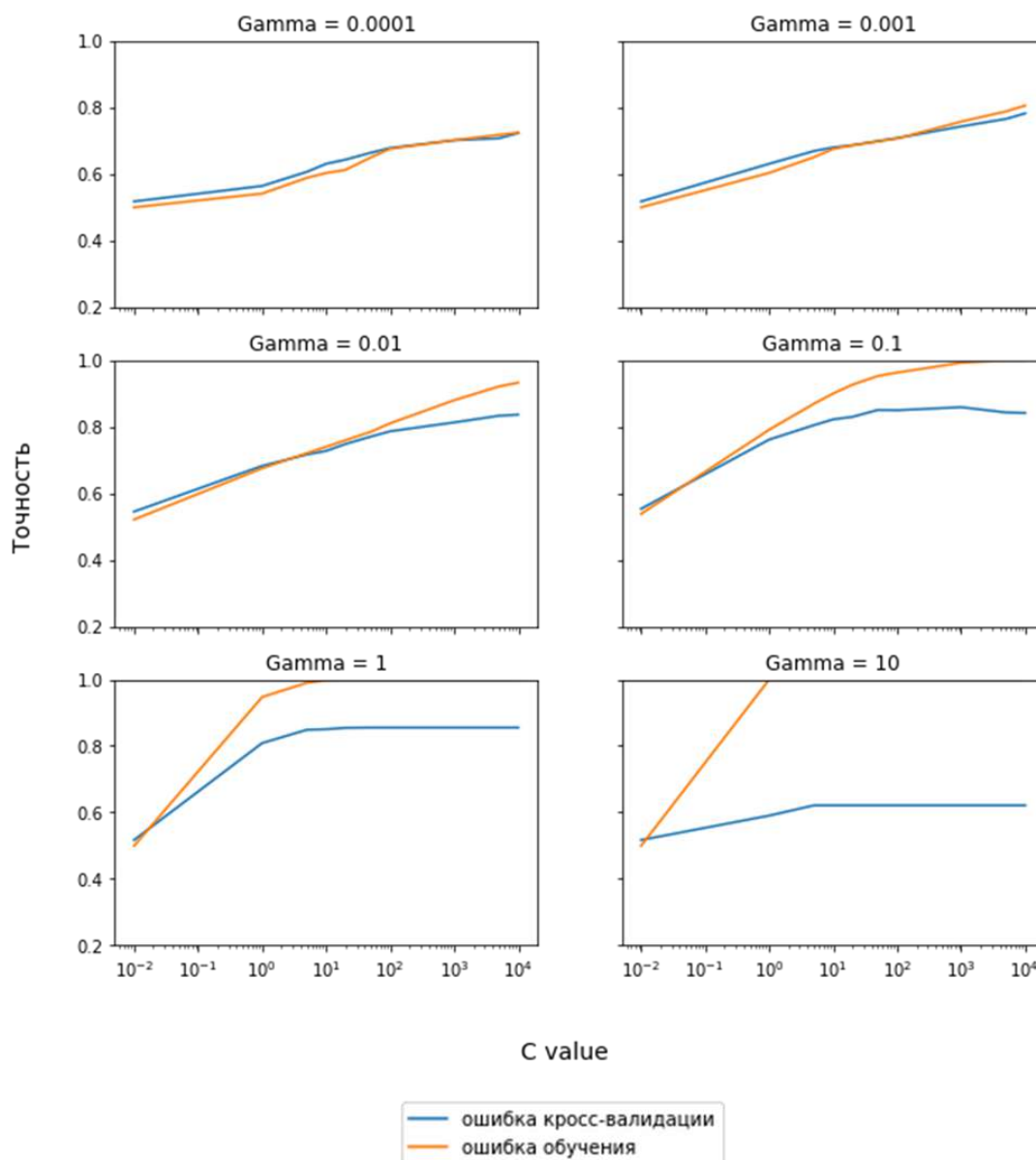


Рисунок 4.16 – Оценка влияния гиперпараметров «с» и «*gamma*» на точность

```

В [26]: In [26]: clf_optimized = svm.SVC(C = 10, gamma = 1)
clf_optimized.fit(X_train_scaled, y_train)
print('Точность метода опорных векторов на валидационной выборке: {:.3f}'.
      format(accuracy_score(y_valid,
                             clf_optimized.predict(X_valid_scaled))))
print('Точность метода опорных векторов на тестовой выборке: {:.3f}'.
      format(accuracy_score(y_from_test,
                             clf_optimized.predict(X_from_test))))

Точность метода опорных векторов на валидационной выборке: 0.850
Точность метода опорных векторов на тестовой выборке: 0.614

```

Рисунок 4.17 – Оценка точности метода опорных векторов при использовании оптимизированных гиперпараметров

Исходя из того, что оптимизация гиперпараметров алгоритмов не обеспечила необходимого прироста в точности, было принято решение произвести углубленный анализ результатов с помощью применения дополнительных методик оценки точности. Одним из таких методов является матрица ошибок, позволяющая оценить ошибки при классификации каждого отдельного литотипа (Рисунок 4.18). Для этой же цели были введены метрики «*precision*» (точность) и «*recall*» (полнота), а также «*F1*» (среднее гармоническое двух предыдущих метрик). Фактически, «*precision*» позволил оценить долю объектов, в которых алгоритм предположительно определил литотип верно и которые действительно оказались верно определенными. Метрика «*recall*», в свою очередь, позволяет оценить долю объектов по каждому правильно оцененному литотипу в общем объеме правильно определенных объектов.

Pred True	Арг	АргКав	АргБит	АргУгл	Алевр	Песч	ПесчГЦ	ПесчПГ	ПесчСП	ПесчПл	Уплотн	Уголь	Неизв	Сумма
Арг	731	51			34						7			823
АргКав	64	29			1						1			95
АргБит														0
АргУгл														0
Алевр	4				23	3	10				11			51
Песч						40				1	4			45
ПесчГЦ					12	33	27			3	19			94
ПесчПГ	1				13		2				5			21
ПесчСП														0
ПесчПл	1				5	5	7			7	18			43
Уплотн	11				2						39			52
Уголь														0
Неизв														0
Precision	0.90	0.36	0.00	0.00	0.26	0.49	0.59	0.00	0.00	0.64	0.38	0.00	0.00	0.75
Recall	0.89	0.31	0.00	0.00	0.45	0.89	0.29	0.00	0.00	0.16	0.75	0.00	0.00	0.73
F1	0.89	0.33	0.00	0.00	0.33	0.63	0.39	0.00	0.00	0.26	0.50	0.00	0.00	0.72

Рисунок 4.18 – Матрица ошибок алгоритма случайного леса

Анализ полученных при использовании метрик результатов позволил сформировать следующие выводы:

1. Чаще и правильнее всего построенная модель выделяет аргиллиты – ключевой литологический класс, который в дальнейшем будет использоваться для проведения межскважинной корреляции. При этом достаточно много ошибок связано с неверным определением литологии у смежного класса – аргиллита кавернозного типа. Одним из вероятных путей решения могло бы стать объединение вышеперечисленных литотипов в единый класс, соответствующий аргиллиту, но в этом случае будут потеряны важные для задач корреляции верно определенные интервалы кавернозного аргиллита, поэтому данные классы были приняты без изменений.
2. Достаточно правильно алгоритм выделяет и песчаные интервалы. Завышенные показания по сопротивлению горных пород, соответствующие возрастающей доле глин в интервале, соответствуют также изменению типа насыщения горной породы. Предположение о насыщении вышеперечисленных интервалов УВ было подтверждено в ходе изучения кернового материала.
3. Алгоритм достаточно часто (примерно в половине случаев) ошибается при выделении алевролитов – промежуточного между аргиллитами и песчаниками литологического класса. В первую очередь это связано с отличными от остальных скважин показаниями геофизических зондов, демонстрирующих в некоторых случаях завышенные или заниженные значения, что может быть связано с особенностями минералогического состава пород или характером насыщения. Кроме того, алевролиты часто выделяются интерпретаторами относительно песчаных и глинистых пластов, т. е. алевролиты не имеют конкретной границы по показаниям геофизических зондов, а определяются как средние по показаниям между песчаными и глинистыми пластами. Также, крайне часто алгоритм не видел разницы между алевролитом и переслаиванием глинистых и песчаных пропластков, которые имеют схожий характер геофизических кривых и соответствующую

амплитуду. Предложенным решением данной проблемы является подтверждение выделенного алгоритмом интервалов аргиллита по керновому материалу с соответствующей корректировкой расчленения.

4. Ошибки при выделении уплотненных прослоев, как было описано выше, связано с наличием в составе песчаных пород глинисто-карбонатного цемента. Характер кривых в таких интервалах схож с таковым у уплотненных прослоев, также имеющих в своем составе карбонатные минералы, приводя к искажениям при записи геофизической информации и ошибкам алгоритма

Таким образом, алгоритм лучше всего справляется с выделением глинистых пачек, определение которых и являлось первоочередной задачей этой части работы. В разрезе выделенных глинистых пачек в дальнейшем будут выделены аномалии, на основании которых будет производиться межскважинная корреляция. Кроме того, алгоритм достаточно точно выделяет и песчаные интервалы, являющиеся основными продуктивными горизонтами исследуемого объекта.

4.3. Выделение аномалий ГИС в пределах глинистых пачек и их систематизация

Выделение аномалий геофизических исследований в расчлененных скважинах производилось на основе визуального анализа характера поведения кривых, изменение которого свидетельствует об изменениях в литологическом составе пород или о смене условий осадконакопления. Кроме того, была проанализирована информация о фактическом каменном материале. Таким образом, на Рисунке 4.19 на примере разреза скважины 30746, находящейся в непосредственной близости от скважины 610Р, отображены основные стратиграфические объекты и выделенные в пределах глинистых пачек аномалии комплекса ГИС.

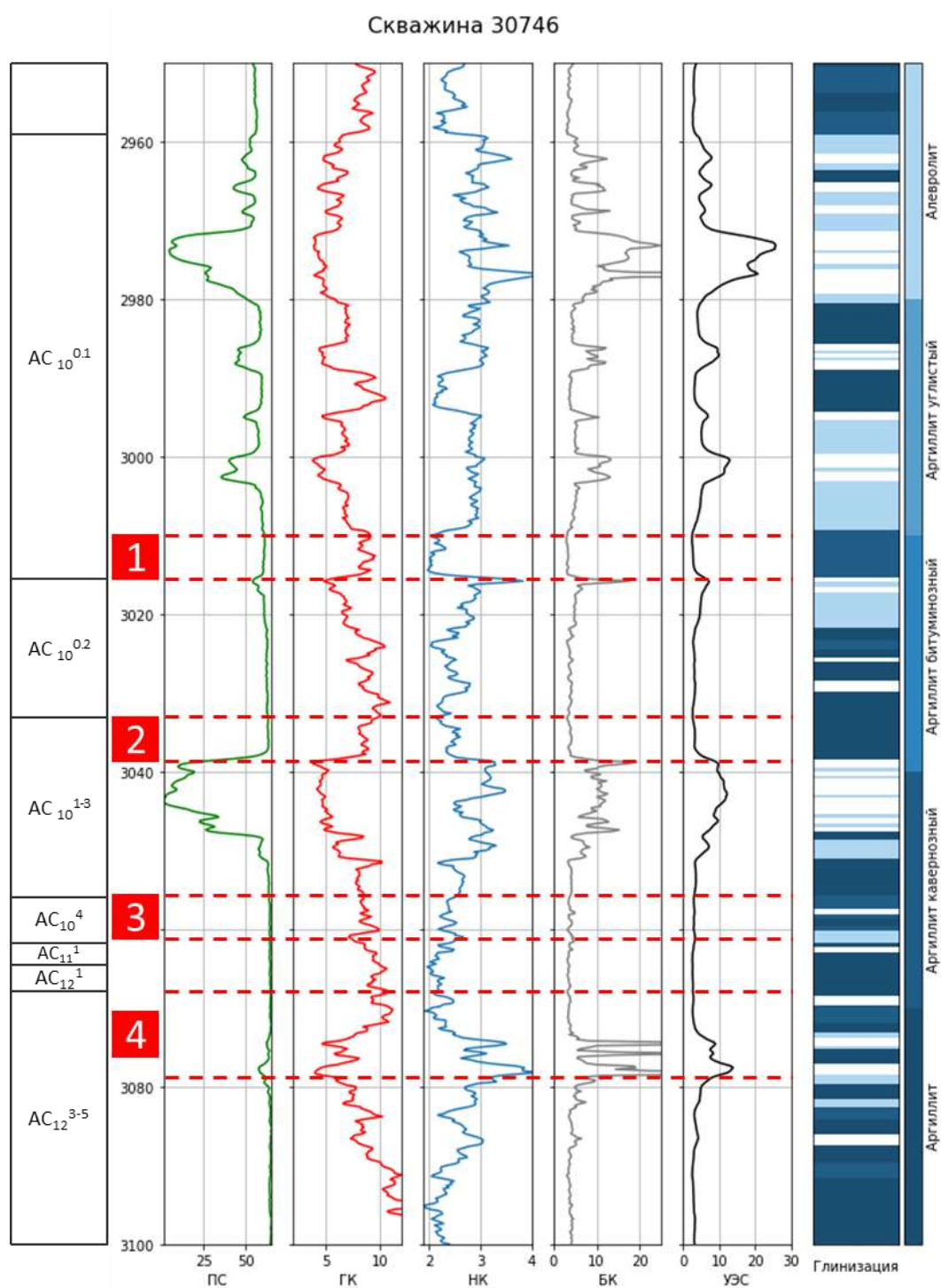


Рисунок 4.19 – Стратиграфическое расчленение скважины 30746
(границы аномалий в пределах глин выделены красной штриховкой)

Таким образом, в ходе анализа особенностей геофизических кривых были установлены четыре типа аномалий, типичных для всех скважин. Все выделенные аномалии соответствуют кровлям продуктивных пластов и представлены в Таблице 4.2.

Таблица 4.2 – Выделенные в пределах глинистых пачек аномалии

Группа	Кровля объекта	Особенности поведения кривых ГИС
1	$AC_{10}^{0.2}$	Высокая дисперсия показаний по ГК и их резкий рост по сравнению с вышележащим интервалом; общий тренд на снижение по НК; плато по ПС, БК и УЭС
2	AC_{10}^{1-3}	Понижение показаний по ГК; резкий рост значений НК; плато по ПС, БК и УЭС
3	AC_{10}^4	Резкий рост дисперсии показаний по ГК и тренд на их увеличение; общий тренд на снижение по НК; плато по ПС, БК и УЭС
4	AC_{12}^{3-5}	Тренд на понижение показаний ГК; повышение значений у НК; плато показаний в верхней части и их скачкообразный рост в нижней по ПС, БК и УЭС

В признаковое описание обучающей выборки была добавлена информация о выделенных аномалиях, соответствующих четырем классам. Таким образом, обновленный датафрейм стал готов ко второму этапу обучения. Построенные на основе классов аномалий зависимости будут также применены к данным разведочной скважины 610Р.

4.4. Определение опорных аномалий и проведение межскважинной корреляции

После окончания процесса непосредственно обучения модели с помощью функции «*accuracy*» была произведена их первичная оценка точности на данных валидационной выборки (Рисунок 4.20).

```

В [13]: ► print('Точность дерева решения: {:.3f}'.
            format(accuracy_score(y_valid, tree_pred)))
          print('Точность случайного леса: {:.3f}'.
                format(accuracy_score(y_valid, forest_pred)))
          print('Точность метода ближайших соседей: {:.3f}'.
                format(accuracy_score(y_valid, knn_pred)))
          print('Точность метода опорных векторов: {:.3f}'.
                format(accuracy_score(y_valid, clf_pred)))

Точность дерева решения: 0.946
Точность случайного леса: 0.971
Точность метода ближайших соседей: 0.867
Точность метода опорных векторов: 0.832

```

Рисунок 4.20 – Использование функции «accuracy» для оценки точности на данных валидационной выборки

Как видно по результатам оценки, алгоритмы демонстрируют впечатляющую точность, угадывая аномалии и, соответственно, границы стратиграфических объектов с точностью от 83 до 97 %. Столь высокие показатели моделей не требуют их предварительной оптимизации, поэтому могут сразу быть распространены на объекты тестовой выборки. Первичная оценка точности приведена на Рисунке 4.21.

```

В [14]: ► test_data = pd.read_excel('test_data_corr.xlsx')
          test_data = test_data.dropna()
          test_data = test_data.drop(test_data[test_data.Depth > 2800].index)

          X_from_test = test_data.drop(['Well', 'R05', 'R07', 'R10', 'R14', 'R20', 'Horizon'], axis = 1)

          scaler = StandardScaler()
          X_from_test = scaler.fit_transform(X_from_test)
          y_from_test = test_data['Horizon']

          print('Точность дерева решения: {:.3f}'.
                format(accuracy_score(y_from_test, tree.predict(X_from_test))))
          print('Точность случайного леса: {:.3f}'.
                format(accuracy_score(y_from_test, forest.predict(X_from_test))))
          print('Точность метода ближайших соседей: {:.3f}'.
                format(accuracy_score(y_from_test, knn.predict(X_from_test))))
          print('Точность метода опорных векторов: {:.3f}'.
                format(accuracy_score(y_from_test, clf.predict(X_from_test))))

Точность дерева решения: 0.430
Точность случайного леса: 0.643
Точность метода ближайших соседей: 0.643
Точность метода опорных векторов: 0.708

```

Рисунок 4.21 – Использование функции «accuracy» для оценки точности на данных тестовой выборки

Стоит отметить, что в случае задачи корреляции числовые значения не позволяют корректно оценить достоверность полученного расчленения. Исходя из этого, результаты было предложено оценить визуально (Рисунок 4.22).

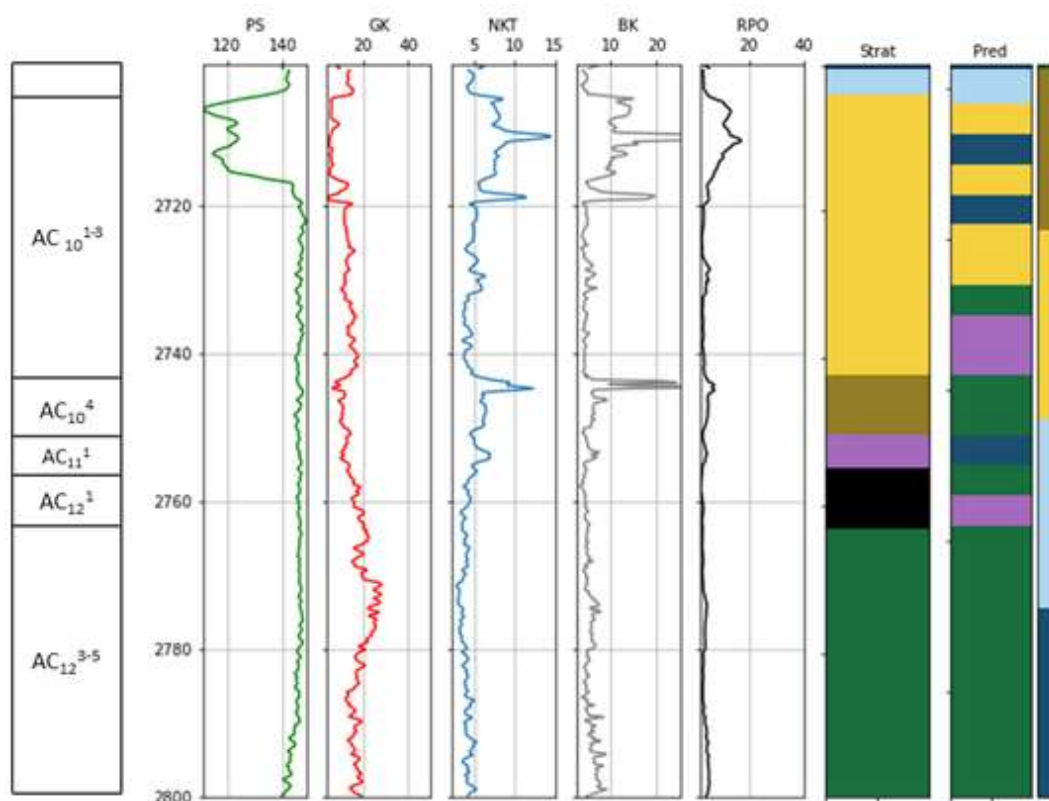


Рисунок 4.22 – Сравнение корреляций интерпретатора и алгоритма

Исходя из оценки результатов функцией «*accuracy*», наибольшую точность при выделении опорных аномалий в пределах глинистых интервалов показал метод опорных векторов. Чуть меньшую значения функции (разница в пределах 10 %) показали методы ближайших соседей и случайного леса. Таким образом, все вышеперечисленные алгоритмы демонстрируют снижение точности не менее чем на 25%. Визуальный анализ, в свою очередь, показал, что алгоритм уверенно справился с выделением половины реперных горизонтов, располагающихся соответственно в кровле объектов $AC_{10}^{0.2}$ и AC_{12}^{3-5} . Ошибки возникли при выделении реперных горизонтов, соответствующих кровлям пластов AC_{10}^{1-3} и AC_{10}^4 . Вероятной причиной подобной проблемы является достаточно сильное изменение мощности данного горизонта в направлении тестовой скважины, из-за которого в алгоритме начинает возникать путаница.

Стоит отметить, что из-за особенностей геологического строения месторождения, приуроченного к лавинным обстановкам седиментации (клиноформам), горизонты могут быть не выдержаны по мощности, выклиниваясь

в определенных направлениях; кроме того, исходя из особенностей формирования палеорельефа, может изменяться также и глубина залегания горизонтов. С другой стороны, не менее важным фактором является объем и представительность входных данных: для получения корректных результатов работы алгоритма требуется достаточно крупная и представительная выборка обучающего материала. Соответственно, целевым объектом применения данной методики являются месторождения с большим объемом накопленных эксплуатационных данных, как правило, находящиеся на поздних стадиях разработки. В таком случае полученная методика может быть приведена к следующему виду:

1. Создание стратиграфического каркаса – выделение глинистых пачек в разрезе скважин, а также аномалий в их составе, которые в дальнейшем будут использованы как реперные горизонты. Подобный каркас послужит основанием для разукрупнения эксплуатационных объектов в процессе их перекорреляции.
2. Классификация скважин в зависимости от количества представленных в них реперных горизонтов.
3. Проведение зонирования территории месторождения по полученной классификации. Зональная группировка в данном случае будет осуществляться по количеству выделенных реперных горизонтов. Опорными в каждой из зон будут являться скважины, разрез которых представлен наибольшим количеством реперных горизонтов.
4. Проведение стратиграфической корреляции, основанной на прослеживании выдержанных (базовых) реперных горизонтов в разрезе в пределах рассматриваемой площади
5. Выполнить расчленение разреза на основе литологической параллелизации пластов, предусматривающее корреляцию второстепенных реперных горизонтов между базовыми в пределах проведённой территориальной зональности, предусмотренной пунктом 2. В соответствии с полученным «каркасом» распределения реперных горизонтов в осадочном разрезе

проводится корреляция пластов-коллекторов и выделяются самостоятельные резервуары.

6. По наличию в коллекторе непроницаемых разделов и их пространственной увязке осуществляется внутрирезервуарная корреляция, которая отражает внутреннюю фильтрационную неоднородность продуктивного пласта, влияющую на разработку залежей углеводородов.

5. Социальная ответственность

5.1. Анализ вредных производственных факторов

Определение потенциальных опасных и вредных производственных факторов выполнялось в соответствии с ГОСТ 12.0.003-74 «Система стандартов безопасности труда (ССБТ). Опасные и вредные производственные факторы. Классификация».

На производительности труда негативно сказывается влияние следующих факторов:

- отклонение температуры и влажности воздуха от нормы;
- недостаточная освещенность рабочего места;
- повышенный уровень электромагнитных излучений.

Работник также подвержен опасности из-за возможности поражения электрическим током и возникновения пожаров в результате короткого замыкания. Негативное воздействие на окружающую среду в процессе работы практически отсутствует. Производственные аварии и пожары являются наиболее вероятными источниками чрезвычайных ситуаций техногенного характера.

Вредные факторы при работе в помещениях были приведены согласно ГОСТ 12.0.003-74, после чего был произведён анализ мероприятий, необходимых для их устранения.

Для защиты от шума на рабочем месте необходимо в первую очередь соблюдать требования и нормы шума в местах работы людей в соответствии с ГОСТ 12.1.003-83 «Шум. Общие требования безопасности», а также ГОСТ 12.1.029-80 «Средства и методы защиты от шума. Классификация» помещений».

Согласно ГОСТ 12.1.003-83, уровень шума, устанавливаемого при рабочих частотах от 63 до 8000 Гц, не должен превышать 80 дБ, а участки с повышенным уровнем шума должны обозначаться соответствующими визуальными знаками или табличками.

Согласно ГОСТ 12.1.003-83 и ГОСТ 12.1 029-80, защита от шума предполагает создание шумобезопасной техники, применение средств индивидуальной и коллективной защиты и применение строительно-акустических методов. Коллективная защита характеризуется снижением силы шума от источника, проведением строительно-акустических мероприятий, в том числе применением звукоизоляционных материалов. Индивидуальная защита, в свою очередь, предполагает применение наушников, противοшумных шлемов и вкладышей.

Недостаточная освещенность рабочей места является одним из наиболее опасных негативных факторов. Нормы освещения рабочих мест и помещений подбираются в соответствии со СНиП 23-05-95 «Строительные нормы и правила. Нормы проектирования. Естественное и искусственное освещение». Отклонение от естественного освещения как в большую, так и в меньшую сторону ведет к раннему переутомлению, а в долгосрочной перспективе – к ухудшению зрения. Нормы освещенности помещений приведены в Таблице 5.1.

Таблица 5.1 – Нормы освещенности офисных помещений

Вид помещения	Норма освещенности, лк
Офисы, в которых осуществляется проведение чертежных работ	500
Офисы большой площади (свободная планировка)	400
Офисы общего назначения	200-300
Конференц-залы	200
Эскалаторы / лестницы	50-100
Холлы, коридоры	50-75
Архивные помещения	75
Кладовые помещения	50

Оптимальные характеристики микроклимата в помещении приведены в Таблице 5.2 согласно СанПиН 2.2.4.548-96 «Гигиенические требования к микроклимату производственных помещений».

Таблица 5.2 – Оптимальные параметры микроклимата рабочих мест.

Время года	Категория работ по уровню энергозатрат, Вт	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость воздуха, м/м
Холодное	1А (до 139)	22-24	21-25	40-60	0,1
	1Б (140-174)	21-23	20-24		
Теплое	1А (до 139)	23-25	22-26		
	1Б (140-174)	22-24	21-25		

Работа за компьютером в течение всего рабочего дня имеет ряд негативных последствия, таких как резь в глазах, регулярные головные боли, а также боли в мышцах рук, шеи и спины; при регулярном выполнении профессиональных задач с использованием персонального компьютера у сотрудников могут возникнуть мигрени, частичная потеря зрения, сколиоз. Таким образом, необходимо грамотно подходить к организации труда, чтобы избежать ухудшения здоровья сотрудников. При оборудовании рабочих мест необходимо следовать требованиям СанПиН 2.2.2/2.4.1340-03:

- естественный свет должен попадать на рабочее место сбоку, лучше слева;
- жалюзи, занавески и прочие регулируемые устройства должны быть на окнах в помещении с визуальными дисплейными терминалами и другими электронно-вычислительными устройствами;
- если на рабочем месте установлен видеомонитор, то расстояние до другого стола с видеомонитором должно быть не менее двух метров; кроме того,

необходимо соблюдать расстояние не менее 1,2 метра между боковыми поверхностями видеомониторов;

- рабочая поверхность должна иметь возможность регулироваться по высоте в диапазоне от 680 мм до 800 мм; в ином случае на уровне 725 мм должна быть установлена высота рабочей поверхности стола;
- оператор должен соблюдать рекомендуемое минимальное расстояние до монитора (60 до 70 см), который должен находиться на 20 градусов ниже уровня глаз.
- рекомендуется организовывать перерывы длиной от 10 до 15 минут через каждые 45-60 рабочих минут в случаях, если работа предусматривает непрерывное взаимодействие с дисплеями без возможности периодически менять виды трудовой деятельности.

5.2. Анализ опасных производственных факторов

Помещение осуществления работ, опираясь на Правила устройства электроустановок (издание 7), относится к категории без повышенной опасности в случае наличия следующих факторов:

- установленная влажность воздуха менее 75%;
- установленная температура менее 350 °С;
- не представляется возможным контакт человека с имеющими соединения с землей металлоконструкциями;
- не представляется возможным контакт с металлическими корпусами электрооборудования;
- отсутствуют токопроводящие пыль и полы.

В зависимости от вероятной степени поражения электрическим током, помещениям могут быть присвоены категории «без повышенной опасности», «с повышенной опасностью» и «особо опасные».

В процессе работы над проектом источником опасного фактора также может выступать персональный компьютер, в первую очередь из-за возможности поражения электрическим током. Согласно ГОСТ 12.1.038-82 «ССБТ. Электробезопасность. Предельно допустимые уровни напряжений прикосновения и токов», возникновение вероятности поражения сотрудников электрическим током происходит при неисправностях электрического оборудования, а также при несоблюдении правил безопасности, при этом степень воздействия увеличивается с ростом мощности тока. Электрическое сопротивление тела человека и приложенное к нему напряжение (так называемое напряжение прикосновения) также влияют на исход поражения, так как они определяют значение тока, протекающего через тело человека. К основным способам и средствам электрозащиты относятся:

- защитное заземление – предназначено для защиты людей от поражения током при прикосновении к этим нетоковедущим частям и выполняется в виде намеренного соединения металлических нетоковедущих частей, которые могут оказаться под напряжением, с землей или ее эквивалентом;
- защитное зануление – преднамеренное электрическое соединение с помощью нулевого защитного проводника с заземлённой наглухо нейтралью электрической сети нетоковедущих частей электроустановок, которые в аварийных ситуациях могут оказаться под напряжением;
- электрическое разделение сетей;
- уравнивание потенциалов – применяют в помещениях, имеющих заземлённые или занулённые электроустановки для повышения уровня безопасности;
- использование малых напряжений;
- защитное отключение;
- предупредительная сигнализация;
- средства индивидуальной электрозащиты.

Рабочие места должны быть оборудованы отдельными щитами с общим рубильником электропитания, который должен находиться в легкодоступном

месте, иметь закрытый зануленный металлический корпус и четкую надпись с указанием величины номинального напряжения. Кроме того, при использовании электроприборов рекомендуется соблюдение следующих правил безопасности:

- электропроводка и все электроприемники должны иметь исправную токоизоляцию;
- нельзя подвешивать провода на гвоздях, металлических и деревянных предметах, перекручивать или завязывать их в узел;
- все токоведущие элементы, электроприборы, розетки, должны быть удалены от труб отопления, водопровода и других металлических коммуникаций;
- протирать осветительную арматуру от пыли можно только сухой тряпкой;
- при возгорании электроприборов или электрических проводов нельзя их гасить водой. Кроме того, в первую очередь необходимо обеспечить их отключение от источника электрического тока и только затем приступить к тушению пожара;
- при включении любого электрооборудования в сеть шнур сначала подключается к прибору, а затем уже в сеть; отключение электроприбора нужно производить в обратном порядке;
- нельзя прикасаться мокрыми или влажными руками к электроприборам, находящимся под напряжением.

5.3. Охрана окружающей среды

Политика государства в области охраны окружающей среды может быть тезисно обозначена следующим образом:

- рациональное использование сырьевой базы;
- переход к экологически чистым и энергосберегающим технологиям;
- минимизация отрицательного воздействия производственной деятельности на естественную природную среду;
- восстановление уже нарушенной природной среды.

Параллельно происходит реализация долгосрочных программ, направленных на восстановление экологической обстановки в районах, в которых осуществляют свою деятельность добывающие предприятия. В свою очередь, в нефтегазодобывающей отрасли происходит внедрение новейших технологий, направленных на рекультивацию территорий, подвергшихся загрязнению нефтепродуктами или другими токсичными веществами. На данный момент постоянно ведутся реконструкции и строительство новых природоохранных объектов.

Стоит отметить, что на территории Российской Федерации постоянно возрастает уровень загрязнения окружающей среды, приводя к возникновению реальной угрозы жизни и здоровью населения. Среди добывающих отраслей в Российской Федерации нефте- и газодобыча занимает одно из лидирующих мест по уровню отрицательного воздействия на окружающую среду. В процессе своей деятельности нефтегазодобывающее предприятие последовательно загрязняет атмосферу, гидросферу (в том числе подземные воды) и нарушает целостность биосферы.

5.4. Защита в чрезвычайных ситуациях

В закрытых помещениях особую опасность представляют пожары в компьютерных кабинетах, в частности из-за того, что они влекут за собой значительные материальные затраты. Источниками огня в подобных случаях, как правило, являются электрические схемы и устройства кондиционирования воздуха.

К мерам предотвращения значительных потерь в случае чрезвычайной ситуации в первую очередь относятся:

1. Разработка планов эвакуации случае чрезвычайных ситуаций и ознакомление с ним сотрудников;
2. Информирование сотрудников о правилах пожарной безопасности;

3. Разработка инструкций о мерах поведения при пожаре.

Ежегодно организации необходимо проводить профилактические мероприятия, направленные на проверку огнетушителей, пожарных гидрантов, шлангов и других средств пожаротушения. Кроме того, все без исключения сотрудники обязаны ежегодно проходить инструктаж по технике безопасности.

5.5. Обеспечение надлежащей охраны труда

Охрану труда можно определить, как систему обеспечения безопасности жизни, здоровья и работоспособности сотрудников в процессе трудовой деятельности. Охрана труда включает в себя комплекс правовых, социально-экономических, лечебно-профилактических, реабилитационных и иных мероприятий.

Основной функцией охраны труда является исследование условий труда, и организация мер по снижению рисков, а также факторов, которые отрицательно влияют на организм работников в процессе труда. Применение техники безопасности – это основной метод охраны труда. В соответствии с законодательством обязанности по обеспечению безопасных условий и охране труда возлагаются на работодателя. Каждый работник обязан:

- Соблюдать требования охраны труда;
- Проходить обучение безопасным методам и приёмам выполнения работ, инструктаж и проверку знаний по охране труда;
- Правильно применять средства индивидуальной и коллективной защиты;
- Немедленно извещать своего непосредственного руководителя о любой ситуации, угрожающей жизни и здоровью людей, о каждом несчастном случае, происшедшем на производстве, или об ухудшении состояния своего здоровья, в том числе о проявлении признаков острого профессионального заболевания или отравления;
- Проходить обязательные предварительные и периодические медосмотры.

6. Финансовый менеджмент, ресурсосбережение и ресурсоэффективность

6.1. Потенциальные потребители результатов исследований

Основной целью оценки любого проекта является высокий уровень экономической ценности произведенного исследования, так как именно эта оценка в дальнейшем определит вероятность получения финансирования со стороны спонсоров, а также увеличит шансы получения гранта для проведения дальнейших научных исследований. Для демонстрации финансовой ценности необходимо подготовить грамотное и профессиональное маркетинговое и финансовое предложение, которое должно состоять из:

- оценки коммерческого потенциала исследования, а также оценки перспективы проведения дальнейших научных исследований;
- перечисления существующих альтернативных вариантов проведения научных исследований, которые будут отвечать новейшим требованиям в области ресурсоснабжения и эффективности;
- указаний финансовой, социальной и экономической эффективности исследования.

6.2. SWOT-анализ

SWOT-анализ — это метод, позволяющий исследовать внешнюю и внутреннюю среду проекта. SWOT-анализ по данной работе представлен в Таблице 6.1.

Таблица 6.1 – Матрица SWOT-анализа

	Сильные стороны: С₁: автоматизация процесса расчленения и корреляции горизонтов. С₂ – визуализация полученных результатов С₃: возможность быстро оценить корректность работы алгоритмов.	Слабые стороны: Сл₁: часто случайное распределение значений. Сл₂: большой объем входных данных для построения корректной модели.
Возможности: В₁: спрос на продукт	Передача интеллектуальных прав на использование программного продукта нефтегазодобывающим компаниям	Предоставление бесплатной лицензии нефтегазодобывающим компаниям при соблюдении условия предоставления с их стороны обратной связи по результатам применения
Угрозы: У₁: отсутствие потребности в продукте	Написание кода в петрофизических программах с реализацией автоматического расчленения и корреляции горизонтов	Продолжение разработки с возможностью применения других алгоритмов

6.3. Цели и результаты проекта

Устав проекта – это документ, закрепляющий основную информацию о проекте, такую как цели проекта, его начальные финансовые ресурсы, планируемые результаты и прочее. Таким образом, цель данного проекта – разработка алгоритма для его дальнейшего использования с целью автоматического расчленения и корреляции нижнемеловых отложений.

Ожидаемым результатом является реализованный на базе высокоуровневого языка программирования «Python» программный продукт, а также оценка полученных результатов. Критерий приема результата выполнения проекта заключается в адекватности полученных результатов и возможности их дальнейшего применения.

6.4. Организационная структура проекта

Проект должен содержать в себе данные об основных участниках разработки, а также их ролях, функциях и времязатратах. Подробная информация представлена в Таблице 6.2.

Таблица 6.2 – Рабочая группа проекта

ФИО, место работы и должность	Роль	Функции	Времязатраты, час.
Белозеров Владимир Борисович, д.г.-м.н., профессор ОНД	Руководитель	Распределение ресурсов, координация деятельности участников проекта	520
Абдуллаев Рафаэль Азерович, магистрант ИШПР НИ ТПУ	Исполнитель	Выполнение отдельных работ по проекту	940

6.5. Экономическая оценка результатов исследования

Предварительная оценка эффективности от внедрения разработанного программного комплекса была определена расчетным путем. Так, прибыль от внедрения алгоритма может быть рассчитана по Формуле 6.1.

$$\Pi = \Xi - \mu, \quad (6.1)$$

где Π – возможная прибыль, руб;

Ξ – стоимость применения алгоритма, руб/год;

μ – затраты на использование алгоритма, руб.

Для оценки стоимости применения алгоритма (притока средств) использовалась Формула 6.2:

$$\Xi = (Z_{\text{руч}} - Z_{\text{алг}}) + \Xi_{\text{алг}}, \quad (6.2)$$

где $Z_{\text{руч}}$ – затраты на проведение ручного расчленения и корреляции, руб;

$Z_{\text{алг}}$ – затраты на автоматизацию процессов, руб;

$\Xi_{\text{алг}}$ – экономический эффект от внедрения программного комплекса (сокращение временных и ресурсных затрат).

Алгоритм будет поставляться предприятию в виде цифрового программного обеспечения для персонального компьютера. Средняя заработная плата специалиста-петрофизика в России на 2021 год составляет около 88400 рублей в месяц. Согласно трудовому законодательству, стандартный рабочий день в организация составляет 8 часов в день или 40 часов в неделю при 5-дневной рабочей неделе. Таким образом, средняя почасовая оплата специалиста-петрофизика может быть найдена по Формуле 6.3:

$$P_{\text{ч}} = \frac{P_{\text{зп}}}{N_{\text{рч}}} = \frac{P_{\text{зп}} \cdot N_{\text{дн}} \cdot N_{\text{м}}}{N_{\text{дг}} \cdot N_{\text{рд}} \cdot T_{\text{сут}}}, \quad (6.3)$$

где $P_{\text{зп}}$ – заработная плата сотрудника, 88400 руб/месяц;

$N_{\text{рч}}$ – количество рабочих часов в одном месяце;

$N_{\text{дн}}$ – количество дней в неделе;

N_m – количество месяцев в году;

$N_{дг}$ – количество дней в году;

$N_{дг}$ – количество рабочих дней в неделе;

$T_{сут}$ – рабочее время в сутки.

Соответственно, средняя почасовая оплата специалиста-петрофизика оценена в 508,6 рублей (Формула 6.4):

$$П_ч = \frac{88400 \cdot 7 \cdot 12}{365 \cdot 5 \cdot 8} = \frac{88400}{173,81} = 508,6 \text{ руб/час} \quad (6.4)$$

Следующим шагом была произведена оценка затрат компании на производство расчленения и корреляции по геофизическим данным. Принято допущение, что специалист-петрофизик затрачивает на эти операции 40 % рабочего времени. Таким образом, годовые затраты компании на проведение расчленения и корреляции по геофизическим данным могут быть вычислены по Формуле 6.5:

$$З_{руч} = П_{зп} \cdot N_m \cdot K_d \cdot 40 \% , \quad (6.5)$$

где $З_{руч}$ – затраты компании на проведение расчленения и корреляции вручную;

K_d – коэффициент, учитывающий человеческий фактор и время на логические операции (принимается равным 1,5).

Соответственно, годовые затраты компании на проведение ручного расчленения и корреляции по геофизическим данным принимаются равными 636480 рублей (Формула 6.6):

$$З_{руч} = 88400 \cdot 12 \cdot 1,5 \cdot 0,4 = 636480 \text{ руб} \quad (6.6)$$

На следующем этапе необходимо рассчитать затраты компании на проведение расчленения и корреляции по геофизическим данным в автоматическом режиме с использованием разработанного программного продукта. Так как алгоритм не имеет человеческого фактора, его скорость работы выше скорости человека, поэтому коэффициент K_d не учитывается.

В первую очередь необходимо оценить затраты на электроэнергию для работы вычислительного комплекса. Расчеты будут проводиться при эквивалентном ручному методу времени, т. е. время работы составит 8 часов в сутки. Персональный компьютер средней мощности потребляет 135 Ватт в час. Соответственно, годовые затраты на электроэнергию можно оценить по Формуле 6.7:

$$Z_{эл} = W_{эвм} \cdot N_{рч} \cdot \nabla \cdot K_{со} , \quad (6.7)$$

где $Z_{эл}$ – затраты в год на электроэнергию для одного компьютера, руб;

$W_{эвм}$ – средняя электрическая мощность одного компьютера, 135 Вт/ч;

∇ – стоимость электроэнергии, 2,95 руб/кВт;

$K_{со}$ – коэффициент сложных операций. Его введение необходимо, потому что при обработке информации компьютер тратит в 2-4 раза больше электроэнергии (принят равным 3);

Таким образом, годовые затраты на электроэнергию для работы вычислительного комплекса составляют 2492 рубля (Формула 6.8):

$$Z_{эл} = 135 \cdot 173,81 \cdot 12 \cdot 0,00295 \cdot 3 = 2492 \text{ руб} \quad (6.8)$$

Необходимо учитывать, что алгоритм производит увязку гораздо быстрее человека. Стандартная расчетная операция для одной скважин, в зависимости от размера массива данных, может занимать до 2 часа в день. Для примера было взято максимальное значение времени. Также в расчет был введен коэффициент K_p , учитывающий случайность распределения значений, из-за которого программу необходимо запускать 2-3 раза для нахождения максимальной корреляции.

Предположим, что компания имеет 40 скважин, данные которых надо увязать. Тогда затраты на работу алгоритма, с учетом его скорости, будут составлять (Формула 6.9):

$$Z_{алг}^{v\uparrow} = t_a^{v\uparrow} \cdot \nabla^{\uparrow} \cdot K_p = 0.17 \cdot 9,75 \cdot 3 \cdot 40 = 198,9 \text{ руб/мес} , \quad (6.9)$$

где $t_a^{v\uparrow}$ – время работы алгоритма на одну скважину;

$\nabla^{\uparrow}$ – затраты на 1 кВт с учетом затрат мощности компьютера;

K_p – коэффициент, учитывающий перезапуск программы.

Затраты за год будут составлять (Формула 6.10):

$$198,9 \cdot 12 = 2387 \text{ руб/год.} \quad (6.10)$$

Следующим шагом был произведен расчет годового притока денежных средств во время использования алгоритма с учетом дополнительного экономического эффекта ($\Xi_{\text{доп}}$). Эффект от внедрения алгоритма рассчитан по Формуле 6.10:

$$\Xi = (Z_{\text{руч}} - Z_{\text{алг}}) - \Xi_{\text{доп}} = (636480 - 2387) - 0 = 634093 \text{ руб.} \quad (6.10)$$

Кроме того, по Формуле 6.11 были оценены затраты на техническое обслуживание и ремонт компьютера:

$$Z_{\text{тех}} = C \cdot \frac{\theta}{100} \cdot \frac{t_{\text{тех}}}{t_{\text{тех,год}}}, \quad (6.11)$$

где $Z_{\text{тех}}$ – затраты на техническое обслуживание компьютера, руб;

C – балансовая стоимость рабочей техники, руб;

θ – норма отчислений на ремонт, %;

$t_{\text{тех}}$ – время работы компьютера, ч;

$t_{\text{тех,год}}$ – годовой фонд времени работы компьютера, ч;

Для расчета времени работы техники с учетом ремонта вводится коэффициент затрат времени на ремонт и профилактику ПК (Формула 6.12):

$$t_{\text{тех}} = t_a^{v\uparrow} \cdot 22 \cdot K_{\text{с.о.}} \cdot \alpha_{\text{п}} = 0,17 \cdot 3 \cdot 22 \cdot 1,15 = 12,9 \text{ ч.} \quad (6.12)$$

Балансовая стоимость компьютера составляет 100 тысяч рублей. Норма отчислений на ремонт (θ) равна 7 %. Годовой фонд времени компьютера при учете коэффициента сложности операции будет составлять 154,8 часа в год на операцию увязки. С помощью Формулы 6.13 производится расчет затрат на техническое обслуживание и текущий ремонт компьютера:

$$Z_{\text{тех}} = 100000 \cdot 0,07 \cdot \frac{12,9}{154,8} = 583 \text{ руб/год.} \quad (6.13)$$

Расчет затрат на амортизацию вычислительной техники (Z_{AM}) производится по Формуле 6.14:

$$Z_{AM} = C \cdot \frac{\beta_{отч}}{100} \cdot \frac{t_{тех}}{t_{тех\text{год}}}, \quad (6.14)$$

где C – балансовая стоимость рабочей техники, руб;

$\beta_{отч}$ – норма отчислений на амортизацию техники, %;

$t_{тех}$ – время работы техники (компьютера), ч;

$t_{тех\text{год}}$ – годовой фонд времени работы компьютера, ч;

Норма отчислений на амортизацию принимается равной 20 %.

Таким образом, затраты на амортизацию вычислительной техники составили 1667 рублей в год (Формула 6.15):

$$Z_{AM} = 100000 \cdot 0,2 \cdot \frac{12,9}{154,8} = 1667 \text{ руб/год}. \quad (6.15)$$

Далее необходимо рассчитать эксплуатационные затраты компании при внедрении алгоритма в работу (Формула 6.16):

$$\mu = Z_{э} + Z_{тех} + Z_{AM} = 2387 + 583 + 1667 = 2438 \text{ руб/год} \quad (6.16)$$

Таким образом, компания после внедрения программного комплекса автоматического расчленения и корреляции получит прибыль в размере 631655 рублей (Формула 6.17):

$$\Pi = 634093 - 2438 = 631655 \text{ руб}. \quad (6.17)$$

Заключение

В процессе создания автоматизированного алгоритма интерпретации данных ГИС для автоматизированного расчленения и корреляции на основе методов машинного обучения были решены следующие задачи:

- изучены существующие подходы к вопросам ручного и автоматического расчленения и корреляции разрезов скважин;
- приведено теоретическое обоснование предлагаемой методики;
- созданы и оптимизированы алгоритмы автоматического расчленения разреза и корреляции межскважинного пространства;
- произведена оценка полученных результатов с последующей верификацией на керновом материале.

Прежде всего стоит отметить, что чаще и правильнее всего построенные модели выделяет литотипы аргиллитов, которые в дальнейшем использовались для проведения межскважинной корреляции; достаточно точно определяются и песчаные интервалы. Хуже всего алгоритм справляется с выделением алевролитов (промежуточного между аргиллитами и песчаниками литологического класса) и уплотненных прослоев.

Ошибки в процессе расчленения были связаны с изменением типа насыщения интервалов, наличием карбонатного цемента в составе горных пород, включениями глауконита и других минералов-полупроводников, а также относительный характер выделения алевролитов. Вышеперечисленные факторы (кроме последнего) приводят к искажениям в процессе записи геофизической информации и, соответственно, ошибкам алгоритма. Для снижения их влияния была проведена оптимизация гиперпараметров, которая позволила незначительно повысить точность определения литотипов горных пород. Подводя итог, алгоритм случайного леса, показавшего наибольшую точность, уверенно определяет литотип образца в 73 % случаев при использовании отложенных данных, показывая незначительное снижение в сравнении с первичной оценкой на валидационной выборке. Выделенные литотипы, а также

характерные ошибки были верифицированы на фактическом (керновом) материале.

В случае задачи корреляции была отмечена следующая закономерность: в случае использования данных верификационной выборки скважин все использованные типы алгоритмов демонстрируют впечатляющую точность, угадывая аномалии и, соответственно, границы стратиграфических объектов с точностью от 83 до 97 %. Наибольшую точность при выделении опорных аномалий в пределах глинистых интервалов показал метод опорных векторов, немного меньшие значения отмечены у методов ближайших соседей и случайного леса. Столь высокие показатели моделей не требовали их предварительной оптимизации, были незамедлительно распространены на объекты тестовой выборки. В последнем случае вышеперечисленные алгоритмы продемонстрировали снижение точности не менее чем на 25%. Проведенный визуальный анализ позволил установить, что алгоритм уверенно справился с выделением половины реперных горизонтов, располагающихся соответственно в кровле объектов $AC_{10}^{0.2}$ и AC_{12}^{3-5} .

Вероятной причиной ошибок является изменение как мощности, так и глубины залегания горизонтов в соответствии с особенностями формирования палеорельефа, приводя к выклиниваниям в определенных направлениях. Кроме того, необходимо отметить крайне низкий объем и слабую представительность входных данных, которых оказалось недостаточно для получения корректных результатов работы алгоритма. Учитывая, что перекорреляция, проводимая с разукрупнения эксплуатационных объектов для пересчета содержащихся в них запасов и оптимизации стратегии их выработки целью требует значительных трудо- и времязатрат, целевым объектом применения описанной методики являются месторождения с большим объемом накопленных эксплуатационных данных, как правило, находящиеся на поздних стадиях разработки, таких, как Приобское нефтяное месторождение. В таком случае полученная методика может быть приведена к следующему виду:

1. Создание стратиграфического каркаса с помощью выделения аномалий в составе глинистых пачек, представленных в разрезе скважин. Подобные горизонты их аномалии в дальнейшем будут использованы как реперные горизонты в процессе создания каркаса, применяемого для разукрупнения эксплуатационных объектов в процессе их перекорреляции.
2. Классификация скважин в зависимости от количества представленных в них реперных горизонтов, выделенных в составе глинистых пачек.
3. Проведение зонирования территории месторождения в зависимости от представленного в скважинах количества выделенных реперных горизонтов. Опорными в процессе определения границ каждой из зон будут являться скважины, разрез которых представлен наибольшим количеством реперных горизонтов.
4. Проведение стратиграфической корреляции, основанной на прослеживании базовых реперных горизонтов в разрезе в пределах рассматриваемой площади.
5. Расчленение разреза на основе литологической параллелизации пластов – в этом случае осуществляется корреляция второстепенных реперных горизонтов между базовыми в пределах проведённой территориальной зональности, предусмотренной пунктом 2. Таким образом в осадочном разрезе проводится корреляция пластов-коллекторов и выделяются самостоятельные резервуары.
6. Внутрирезервуарная корреляция, осуществляемая по наличию в коллекторе непроницаемых разделов и их пространственной увязке. Подобная корреляция отражает внутреннюю фильтрационную неоднородность продуктивного пласта, влияющую на разработку залежей углеводородов.

Дальнейшее повышение точности созданных алгоритмов также может быть произведено как за счет увеличения количества объектов, используемых для получения математических зависимостей, так и с помощью расширения признакового описания представленных объектов и использования новейших методик их интерпретации.

Список использованной литературы

1. Абдуллаев Р.А. «Разработка автоматизированного алгоритма для определения фильтрационно-емкостных свойств в скважинах без отбора керна» / Труды XXV Международного научного симпозиума студентов и молодых ученых имени академика М.А. Усова «Проблемы геологии и освоения недр». – Томск, 2021.
2. Арсланов Х.А. «Геологический словарь»/ Х.А. Арсланов, М.Н. Голубчин, А.Д. Искандеров // М.: Недра, 1973. Т1.- 486 с.
3. Боганик В.Н. «Способы повышения достоверности обработки данных ГИС» / В.Н. Боганик, А.И. Медведев, С.Н. Григорьев // НТВ "Каротажник". Тверь: Изд. АИС. 2001. Вып. 86. с. 99-110.
4. Буш А.Д. «Стратиграфические ловушки в песчаниках». – М.: Мир, 1977. – 211 с.
5. Верхотурова О.М. «Идентификация границ слабоструктурированных трехмерных объектов на основе методов дискретной оптимизации: дис. канд. техн. наук». – Уфа, 2009. – 170 с.
6. Гафуров Д.С. «Методика комплексной интерпретации геолого-геофизических данных обучаемыми нейронными сетями на примере Талаканского нефтегазоконденсатного месторождения: автореферат дис. канд. геол.-минерал. наук». – Томск: ТПУ, 2013. – 22 с.
7. Гольберт А.В. «Палеоландшафты Западной Сибири в юре, мелу и палеогене» / А.В. Гольберт, Л.Г. Маркова, И.Д. Полякова, В.Н. Сакс, Ю.В. Тесленко // М.: Наука, 1968 г. – 152 с.
8. ГОСТ 12.0.003-74 «Система стандартов безопасности труда (ССБТ). Опасные и вредные производственные факторы. Классификация». М.: Издательство стандартов, 1974 г.– 25 с.
9. ГОСТ 12.1.003-83 (СТ СЭВ 1930-79). «Шум. Общие требования безопасности» – М.: Госкомсанэпиднадзор, 2003 – 28 с.

- 10.ГОСТ 12.1.029-80. «Средства и методы защиты от шума. Классификация» – М.: Госкомсанэпиднадзор, 2003 – 28 с.
- 11.ГОСТ 12.1.038-82 «ССБТ. Электробезопасность. Предельно допустимые уровни напряжений прикосновения и токов» – М.: Издательство стандартов, 1982 – 35 с.
- 12.Гутман И.С. «Программный комплекс ACDV для изучения осадконакопления в залежах углеводородов сложного геологического строения» / И.С. Гутман, И.Ю. Балабан, О.В. Постнова, С.А. Руднев, М.И. Саакян // Геофизика, № 4, 2010, с. 17-25.
- 13.Дюбрюль О. "Геостатистика в нефтяной геологии.". Издательство: Институт компьютерных исследований, 2009 г.
- 14.Елисеева И.И. «Общая теория статистики: Учебник» / И.И. Елисеева, М.М. Юзбашев // Москва: Финансы и Статистика, 2002. – 480 с.
- 15.Изотова Т.С. «Седиментологический анализ данных промысловой геофизики» / Т.С. Изотова, С.Б. Денисов, Б.Ю. Вендельштейн // М.: Недра, 1993. – 175 с.
- 16.Калантаров А.П. «Спутник нефтегазопромыслового геолога» / А. П. Калантаров, М. Г. Ованесов, Г. П. Ованесов. // М.: Недра, 1971. – 333 с.
- 17.Ковалевский Е.В. «Программные средства поддержки детальной корреляции скважин по данным ГИС» / Е.В. Ковалевский, Г.Н. Гогоненков // Геофизика. – 2004. – Вып.1. – с. 21–26.
- 18.Ковалевский Е.В. «Уточнение геологических моделей посредством использования автоматической корреляции скважин» / Е.В. Ковалевский, Г.Н. Гогоненков, М.В. Перепечкин // Недропользование – XXI век. – 2007 (№ 4), с. 28–31.
- 19.Лусиа Д. «Построение геолого-гидродинамической модели карбонатного коллектора: интегрированный подход». / Под редакцией Ботвиновской О.А., Меркушкиной Ю.В., Охотиной С.В., Тверитнева А.А., Яценко В.М. // – Издательство «ИКИ», 2010 г.

20. Петерсилье В.И. «Методические рекомендаций по подсчету запасов нефти и газа объемным методом» / В.И. Петерсилье, В.И. Порожун, Г.Г. Яценко. // Москва-Тверь: ВНИГНИ, НПЦ «Тверьгеофизика», 2003 г.
21. СанПиН 2.2.1/2.1.1.1200-03. «Санитарно-защитные зоны и санитарная классификация предприятий, сооружений и иных объектов» М.: Госкомсанэпиднадзор, 2005 – 11 с.
22. СанПиН 2.2.4.548-96. «Гигиенические требования к микроклимату производственных помещений» – М.: Госкомсанэпиднадзор, 2003 – 11 с.
23. СНиП 23-05-95. «Строительные нормы и правила. Нормы проектирования. Естественное и искусственное освещение» – М.: Госкомсанэпиднадзор, 2003 – 28 с.
24. Трудовой кодекс Российской Федерации от 30.12.2001 № 197-ФЗ (ред. от 01.04.2019) // Собрание законодательства РФ. - 07.01.2002. - № 1 (ч. 1). – ст. 212, 214.
25. Федеральный закон от 21.12.1994 № 69-ФЗ «О пожарной безопасности» // СЗ РФ. 26.12.1994.
26. Al-Anazi A. «On the capability of support vector machines to classify lithology from well logs» / A. Al-Anazi, I.D. Gates // Nat. Resour. Res. – № 19, pp. 125-139, June (2), 2010.
27. Amirgaliyev Y. «Automated method of 3D plane models creation» / Y. Amirgaliyev, A. Mukhamedgaliyev, R. Yunussov, T. Merembayev // Kyrgyz Republic National Academy of Science, 2009, vol. 3, p. 17-22.
28. Bishop C.M. «Pattern recognition and machine learning (information science and statistics)». – Springer-Verlag, 2006.
29. Busch J. «Determination of lithology from well logs by statistical analysis» / J. Busch, W. Fortney, L. Berry // SPE formation evaluation, pp. 412-418, 1987.
30. Cannady J. «National information systems security conference», 1998.
31. Chaki S. «Reservoir characterization: a machine learning approach» – Karagpur.: Indian institute of technology, 2015. – 74 p.

32. Dubois M.K. «Comparison of four approaches to a rock facies classification problem» / M.K. Dubois, G.C. Bohling, S. Chakrabarti // Comput. Geosci., 2007, № 33, p. 599-617.
33. Edmundson H. «Radioactive Logging Parameters for Common Minerals» /H. Edmundson, L.L. Raymer // SPWLA 20th Annual Logging Symposium. – Tulsa, Oklahoma, 1979.
34. Ellis D. «Formation Porosity Estimation from Density Logs». Society of Petrophysicists and Well-Log Analysts, 2003.
35. Gardner J. «Litho-Density Log Interpretation» / J.S. Gardner, J.L. Dumanoir // SPWLA 21st Annual Logging Symposium, Lafayette, 1980.
36. Gardner J.S. «Litho-Density Log Interpretation» / J.S. Gardner, J.L. Dumanoir // Society of Petrophysicists and Well-Log Analysts. 1980.
37. Gurney K. «An Introduction to Neural Networks». – London: University College London Press, 1997. – 248 p.
38. Hsieh B.Z. «Lithology identification of aquifers from geophysical well logs and fuzzy logic analysis» / B.Z. Hsieh, C. Lewis, Z.S. Lin, // Comput. Geosci. – № 31, pp. 263-275, April (3), 2005.
39. <https://dprom.online/oilngas/aspekty-primeneniya-ingibiruyushhego-biopolimernogo-burovogo-rastvora-pri-burenii-bokovyh-stvolov-na-yuzhno-priobskom-mestorozhdenii/>
40. <https://ru.wikipedia.org/wiki/CRISP-DM>
41. <https://www.bibliofond.ru/view.aspx?id=655679#text>
42. Krenker A. «Introduction to the Artificial Neural Networks»/ A. Krenker, J. Bešter, A. Kos // Artificial Neural Networks – Methodological Advances and Biomedical Applications. INTECH Open Access Publisher, 2011. p. 3-18.
43. Kurgan L. «A survey of Knowledge Discovery and Data Mining process models» / L. Kurgan, P. Musilek // The Knowledge Engineering Review. Cambridge University Press, New York. – 2006, vol. 21, p. 1–24,

44. Li Y. «Facies identification from well logs: a comparison of discriminant analysis and native Bayes classifier» / Y. Li, R. Anderson-Sprecher // J. Pet. Sci. Eng, 2006, p. 149-157.
45. Lim J. «Artificial intelligence approach for well-to-well log correlation» / J.S. Lim, J.M. Kang, K.I. Jungwhan // Journal of petroleum technology. – 1998, №. 11. – p. 30–32.
46. Lim J. «Interwell log correlation using artificial intelligence approach and multivariate statistical analysis» / J.S. Lim, J.M. Kang, J. Kim // Artículo SPE. – 1999, №. 9.
47. Lim J. «Multivariate statistical techniques including PCA and rule-based systems for well log correlation» / Developments in Petroleum Science, 2003, vol. 51 – p. 673–688.
48. Luthi S. «Well-log correlation using a back-propagation neural network» / S. Luthi, I.D. Bryant // Mathematical geology. – 1997, №. 3. – p. 413–425.
49. Rogers S.J. «Determination of lithology from well logs using a neural network» / S.J. Rogers, J.H. Fang, C.L. Karr, D. Stanley // AAPG bulletin, 1992, p. 731-739.
50. Schlumberger «Log Interpretation Charts» / Texas: Schlumberger, 2009.
51. Serra O. «Sedimentological Analysis of Shale-Sand Series from Well Logs» / O. Serra, L. Sulpice // SPWLA 16th Annual Logging Symposium, New Orleans, Louisiana, 1985.
52. Serra O. «Theory, Interpretation, And Practical Applications of Natural Gamma Ray Spectroscopy» / O. Serra, J. Baldwin, J. Quirein // SPWLA 21st Annual Logging Symposium, Lafayette, Louisiana, 1980.
53. Tianqi Ch. «XGBoost: A Scalable Tree Boosting System» / Ch. Tianqi, C. Guestrin // pp.785-794, KDD, 2016.
54. Validov M.F. «Development of the Approach for Automatic Well Logging Interpretation for Big Number of Wells with the Use of Machine Learning» / M.F. Validov, A.R. Ismagilov, D.S. Voloskov, M.S. Magdeev, A.A. Nazarov // EAGE Geomodel, 2017.

55. Wolf M. «Faciologautomatic electrofacies determination» / M. Wolf, J. Pelissier-Combescure // Presented at the SPWLA 23rd Annual Logging Symposium, Society of Petrophysicists and Well-Log Analysts, 1982.
56. Yunsheng W.L. «Integrated Workflow on Lithofacies Modeling» / W.L. Yunsheng, R.S. Bansal // SPE Middle East Oil & Gas Show and Conference, Manama, Kingdom of Bahrain, 6-9 March, 2017.

Приложение А

(обязательное)

Using of a machine learning methods for facies identifying and correlation based on well logging data

Студент

Группа	ФИО	Подпись	Дата
2ТМ91	Абдуллаев Рафаэль Азерович		28.05.2021

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор	Белозеров В.Б.	д.г.-м.н.		28.05.2021

Консультант-лингвист отделения иностранных языков ШБИП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Уткина А.Н.	к.филос.н.		28.05.2021

There has been much excitement recently about big data and the dire need for data scientists who possess the ability to extract meaning from it. Geoscientists, meanwhile, have been doing science with voluminous data for years, without needing to brag about how big it is. But now that large, complex data sets are widely available, there has been a proliferation of tools and techniques for analyzing them. Many free and open-source packages now exist that provide powerful additions to the geoscientist's toolbox, much of which used to be only available in proprietary (and expensive) software platforms.

One of the best examples is machine learning toolbox, a set of data-analysis methods that includes classification, clustering, and regression. These algorithms can be used to discover features and trends within the data without being explicitly programmed, in essence learning from the data itself. So, this term will demonstrate how to use a classification algorithm to identify lithofacies based on well-log measurements.

Determining the optimal sequence of actions for constructing a machine learning algorithm

Knowledge of stratigraphy and lithology boundaries is important for hydrogeology, oil and gas and in-situ recovery areas. Stratigraphy boundaries are provided two main information for detail study: water level horizon, which is important for monitoring of environment situation in underground and defining production reservoir that give bottom and top level for detail geology study. Lithology boundaries are provided main information about geology and detail information about type of geology – permeable or impermeable rocks, it is major data which give to engineer to make decision – to mine ore or not mine. Stratigraphy and lithology are obtained by drilling and geophysics logging. Interpretation of data is based on resistivity and spontaneous measurement.

Well logging interpretation is based on mathematical and physical modelling of the processes under study (solution of direct problems of geophysics), statistical methods (correlation and discrimination), solving systems of non-linear equations petrophysics (inverse problem of geophysics) and some other linear statistical methods.

Depending on the logging signals from the calculated physical and geological parameters, such as permeability, shaliness and radioactivity of host rocks, water content, salinity of aquifers, etc. Often geophysical raw acquired data have complex non the geophysics measurement error signal, specificity of the geological environment. All these analytical conditions and parameters lead to the assumption that the solutions for the problems of geophysical techniques will be more easily obtained by using algorithms such as: random forest, logistic regression, gradient boosting and k-nearest neighbor. These methods have the property of adaptability, generalizations, knowledge extraction and modelling of complex non-linear dependencies in the data.

First of all, underground stratification of geology is studied through geophysics logging. Significant changing of geology stratification defines boundaries of stratigraphy, interpretation geophysics logging provides rock types, for example clay, sand, organic matter, etc. – lithological facies. The targets are view generalization of boundaries, compare accuracy of algorithms for stratigraphy and lithological facies data. To solve it, for example, the next data may be used:

- Resistivity logging (R) – measures electric resistivity of rocks;
- Spontaneous potential logging (SP) –measures small electric potentials of rocks;
- Depth (D) – values of distance between geodesic elevation and down points
- Elevation (EL) – values of distance between well head (0 value) and down points

Base on this matrix of data (R, SP, D, EL), the goal is to find best machine learning algorithms which will associate the appropriate stratigraphy class according to input parameters values with best accuracy. And check if the selected algorithms can show the same accuracy for lithology data.

There are several models and algorithms were applied to improve remote sensing data analysis procedures, such as Regression, Random Forest, KNN and XGBoost. During the experiments, the accuracy of selected algorithms and its time performance were obtained. To measure the quality of algorithm, some metrics were selected. To validate the results of training, a three-fold cross-validation strategy was selected. The tuning of algorithm meta-parameters is important process, so the best parameters for each of the model were configured. To train the models, 4 features (R, SP, EL, D) were

used. Another train model is done with adding of two extra 6 features:(R` and R`` – the first and second derivatives of RM curve respectively). However, addition of 2 extra features have not significantly improved the accuracy.

All algorithms showed high accuracy on test stratigraphy data and less for lithology data. This approach helps to create a self-automatic definition of stratigraphy levels and reduces the time for data interpretation (Merembayev, 2018).

Usage of Python classification algorithms for automatic identifying and correlation of lithofacies

The example data set consists of seven features (five wireline log measurements and two indicator variables) and a facies label at half-foot depth intervals. In machine learning terminology, the set of measurements at each depth interval comprises a feature vector, each of which is associated with a class (the facies type). The Pandas library from Python will be used to load the data into a dataframe, which provides a convenient data structure to work with well-log data.

Data.describe function may be used to provide a quick overview of the statistical distribution of the training data (Figure 1).

	Facies	Depth	GR	ILD_log10	DeltaPHI	PHIND	PE	NM_M	RELPOS
count	3232	3232	3232	3232	3232	3232	3232	3232	3232
mean	4.42	2875.82	66.14	0.64	3.55	13.48	3.73	1.50	0.52
std	2.50	131.00	30.85	0.24	5.23	7.70	0.89	0.50	0.29
min	1	2573.50	13.25	-0.03	-21.83	0.55	.020	1	0.01
25%	2	2791.00	46.92	0.49	1.16	8.35	3.10	1	0.27
50%	4	2932.50	65.72	0.62	3.5	12.15	3.55	2	0.53
75%	6	2980.00	79.63	0.81	6.43	16.45	4.30	2	0.77
max	9	3122.50	361.15	1.48	18.60	84.40	8.09	2	1.00

Figure 1 – Statistical distribution of the training data set

As it can be seen from the count row in Figure 1, there is a total of 3232 feature vectors in the data set. The feature vectors consist of the following variables:

1. Gamma ray (GR).
2. Resistivity (ILD_log10).
3. Photoelectric effect (PE).

4. Neutron-density porosity difference (DeltaPHI).
5. Average neutron-density porosity (PHIND).
6. Nonmarine/marine indicator (NM_M).
7. Relative position (RELPOS).

There are nine facies classes (numbered from 1 to 9) identified in the data set. Figure 2 contains the descriptions associated with these classes. Note that not all of these facies are completely discrete; some gradually blend in to one another. Misclassification of these neighboring facies can be expected to occur. The Adjacent Facies column in Figure 2 lists these related classes.

Facies	Description	Label	Adjacent facies
1	Nonmarine sandstone	SS	2
2	Nonmarine coarse siltstone	CSiS	1,3
3	Nonmarine fine siltstone	FSiS	2
4	Marine siltstone and shale	SiSh	5
5	Mudstone	MS	4,6
6	Wackestone	WS	5,7,8
7	Dolomite	D	6,8
8	Packstone-grainstone	PS	6,7,9
9	Phylloid-algal bafflestone	BS	7,8

Figure 2 – Facies labels with their descriptions

To evaluate the accuracy of the classifier, one well from the training set will be removed so the predicted and actual facies labels can be compared. Cross-plots are a familiar tool to visualize how two properties vary with rock type. This data set contains five log measurements, and the Seaborn library can be employed to create a matrix of cross-plots to visualize the variation between the log measurements in the data set.

Each pane in Figure 3 shows the relationship between two of the variables on the x and y axis, with a stacked bar plot showing the distribution of each point along the diagonal. Each point is colored according to its facies. It is not clear from these crossplots what relationships exist between the measurements and facies labels. This is where machine learning will prove useful.

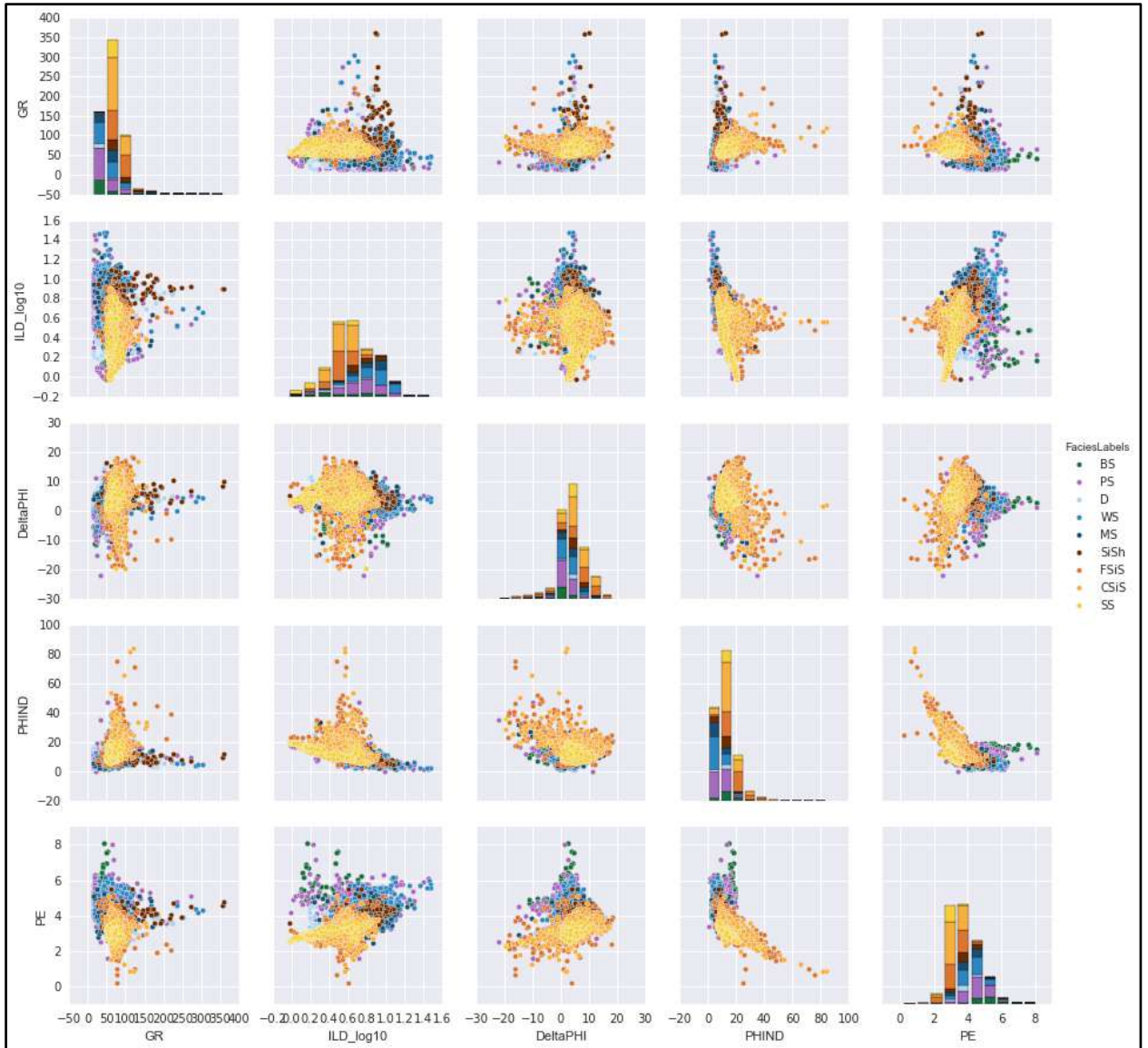


Figure 3 – Cross-plot matrix generated with the Seaborn library

Many machine-learning algorithms assume the feature data are normally distributed (i.e., Gaussian with zero mean and unit variance). Figure 1 shows that this is not the case with the training data. The training data will be conditioned or standardized, so that it has this property. The same factors used to standardize the training set must be applied to any subsequent data set that will be classified. Scikit-learn includes a handy `StandardScaler` class that can be applied to the training set and later used to standardize any input data.

A standard practice when training supervised-learning algorithms is to separate some data from the training set to evaluate the accuracy of the classifier. The data from

single well has already been removed for this purpose. It is also useful to have a cross-validation data set that can be used to tune the parameters of the model. Scikit-learn includes a handy function to randomly split the training data into subsets, so 5 % of the data will be used for the cross-validation set.

After that, the conditioned data set may be used to train a support vector machine to classify facies. The concept of a support vector machine is straight forward. If the data were linearly separable, it would be easy to draw boundaries between the input data points that identified distinct classes. Figure 3 suggests that the data is not linearly separable. Using a technique known as the “kernel trick,” the data is projected into a higher dimensional space where it is separable. Boundaries, or margins, can be drawn between the data in this space. These boundaries are generated during the training step.

The SVM implementation in scikit-learn takes a number of important parameters. These can be used to control the learning rate and the specifics of the kernel functions. The choice of parameters can affect the accuracy of the classifier, and finding optimal parameter choices is an important step known as model selection. A succession of models is created with different parameter values, and the combination with the lowest cross-validation error is used for the classifier.

After the model has been trained on the data, it can be used for the facies predicting of any well with the same set of features. To evaluate the accuracy of the classifier the well that was kept for a blind test will be used for comparing the predicted facies with the actual ones. The facies labels and features of this dataset should be extracted. Now the trained classifier may be used to predict facies labels for this well, and store the results in the Prediction column of the test well dataframe.

Because the true facies labels of the vectors in the test dataset are known, the results may be used to evaluate the accuracy of the classifier on this well (Figure 4).

Class	precision	recall	f1-score	support
SS	0.30	0.08	0.12	89
CSiS	0.36	0.72	0.48	89
FSiS	0.62	0.54	0.58	117
SiSh	0.25	0.29	0.27	7
MS	0.17	0.11	0.13	19
WS	0.66	0.54	0.59	71
D	0.71	0.29	0.42	17
PS	0.41	0.60	0.49	40
BS	0.00	0.00	0.00	0
avg/total	0.47	0.46	0.43	449

Figure 4 – Cross-plot matrix generated with the Seaborn library

Precision and recall are metrics that tell us how the classifier is performing for individual facies. Precision is the probability that, given a classification result for a sample, the sample actually belongs to that class. Recall is the probability that a sample will be correctly classified for a given class. For example, if the classifier predicts that an interval is fine siltstone (FSiS), there is a 62 % probability that the interval is actually sandstone (precision). If an interval is actually fine siltstone, there is a 54 % probability that it will be correctly classified (recall). The F1 score combines both accuracy and precision to give a single measure of relevancy of the classifier results. The used classifier achieved an overall F1 score of 0.43 on the test well, so there is room for improvement (Hall, 2016).

Figure 5 demonstrates the classifier results in log-plot form. The five logs used as features are plotted along with the actual and predicted lithofacies class log.

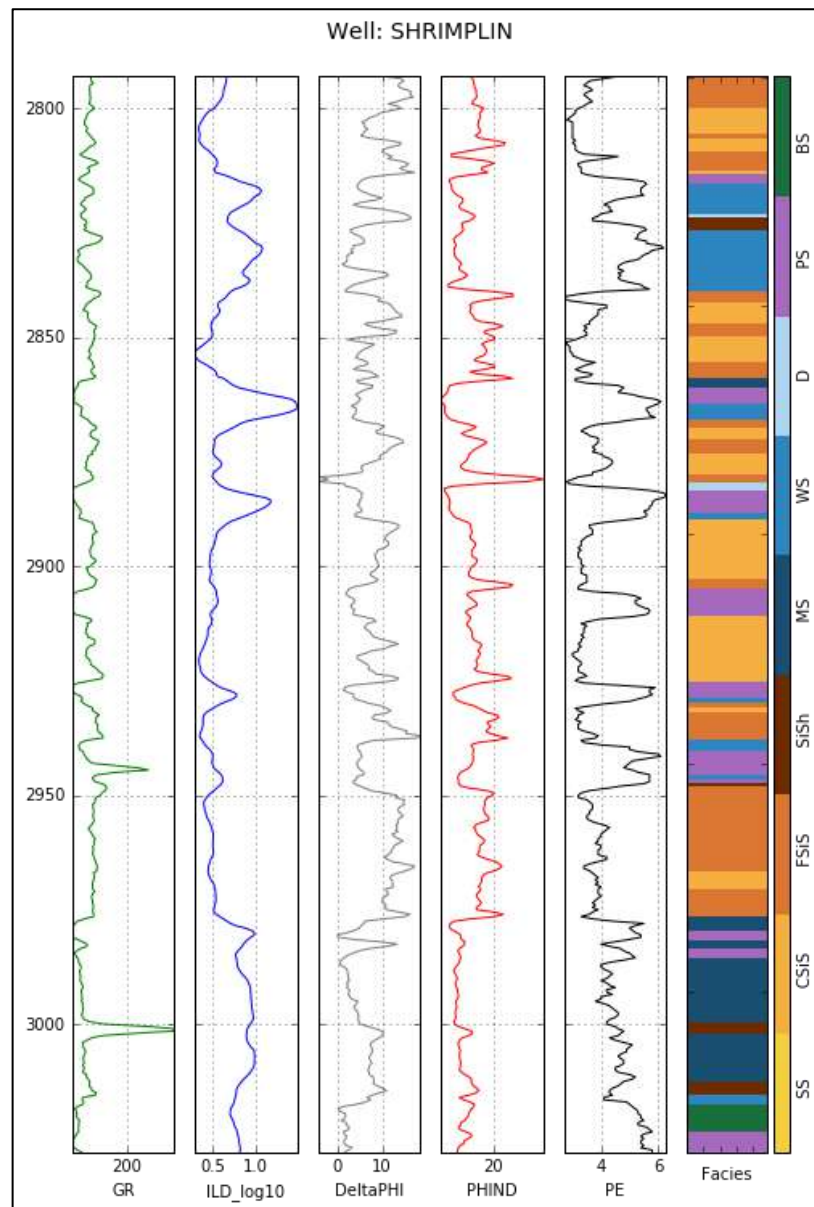


Figure 5 – Well logs and facies classification results from a single well

To increase the accuracy of the classifier, may be used additional data such as seismic information. The seismic horizons can be treated as the stratigraphic boundaries which can represent the depositional environments and geological features. Generally, the horizons are tracked by picking the similar configuration of samples such as peaks, troughs, or zero-crossing points through the consistently seismic traces. Manually tracking is the most familiar but the least efficient interpretation techniques.

The process of seismic horizon interpretation can be viewed as the segmentation of seismic traces into different piece wise segments. In this paper, CNN was employed to segment the seismic traces into different zones automatically to increase accuracy of

model prediction. One percent of the seismic traces was randomly selected within the seismic survey as the training data. The interpreted horizons divide the seismic traces into different segment. The same labels were given to the segment of seismic traces which are bounded the same horizons. The rest seismic traces function as the testing data set used for validating the effectiveness of our algorithm. The curve patterns of the seismic amplitude can be treated as the image features.

The proposed network consists of two main parts: encoder and decoder networks. The role of encoder network is learning the features imbedded in each segment of seismic traces of the training set. The decoder network is automatically segment the seismic traces using the learned features in the decoder process. The decoder process consists four layers and each layer contains convolution layer, batch-normalization layer, activation layer and max-pooling layer. The decoder process also consists four layers which corresponded to the encoder network. Each layer in the decoder contains up sampling layer, convolution layer, batch-normalization layer and activation layer. The pixel-wise classification layer is the last process in our workflow. The output of the workflow is segmented seismic traces and the boundary of each segments are the automatically tracked horizons.

The effectiveness of the workflow demonstrated on the the F3 block seismic data acquired over the offshore North Sea, Netherland. The seismic survey consists of 651 inline and 951 crosslines. The time sample increment is 4 ms. The project already has four interpreted horizons (Figure 6).

To train a deep convolutional encoder-decoder neural network, the dataset should be organized. First of all, the seismic traces should be labeled with manually picking four horizons to build a horizon model (Figure 6). Then they labeled for training according to interpreted horizons. There are four horizons, so the seismic traces are segmented into four segments. Moreover, there are 5 labels corresponding to the segmented seismic traces. Figures 7a and 7b are the seismic inline sections before and after labeling according to the interpreted horizons.

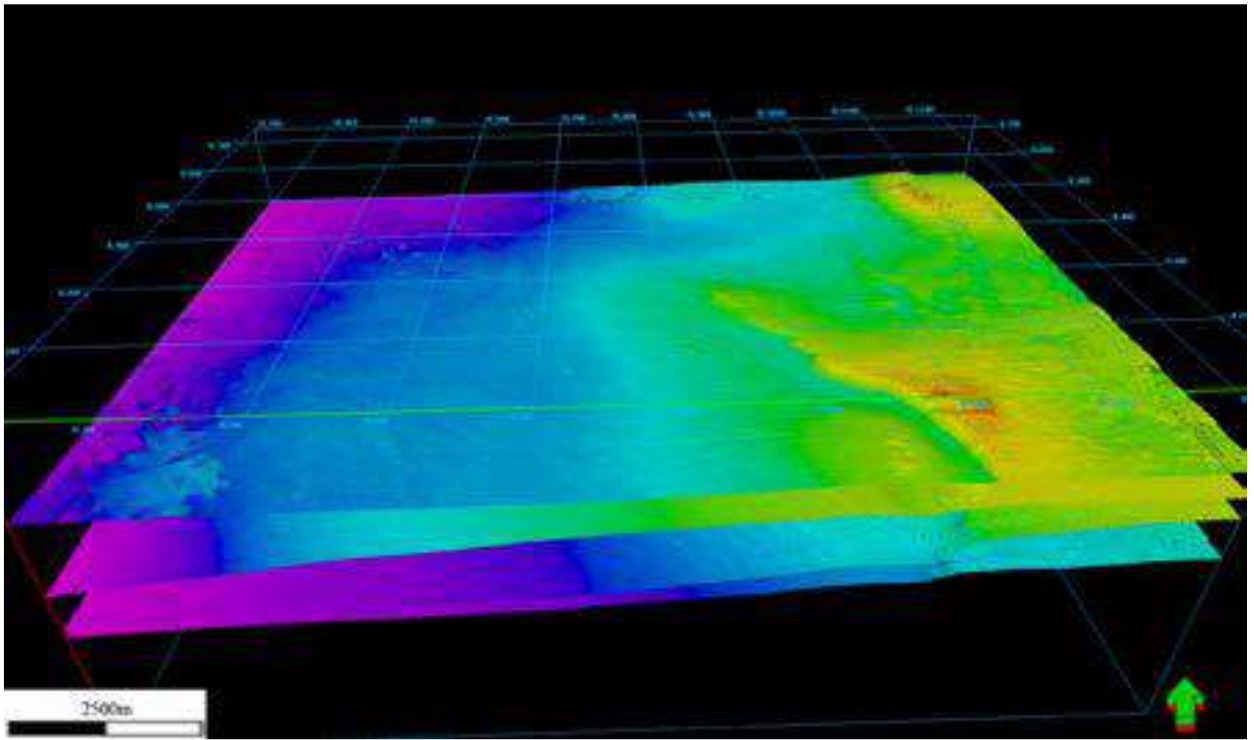


Figure 6 – Manually picked horizon volume (Wu, 2018)

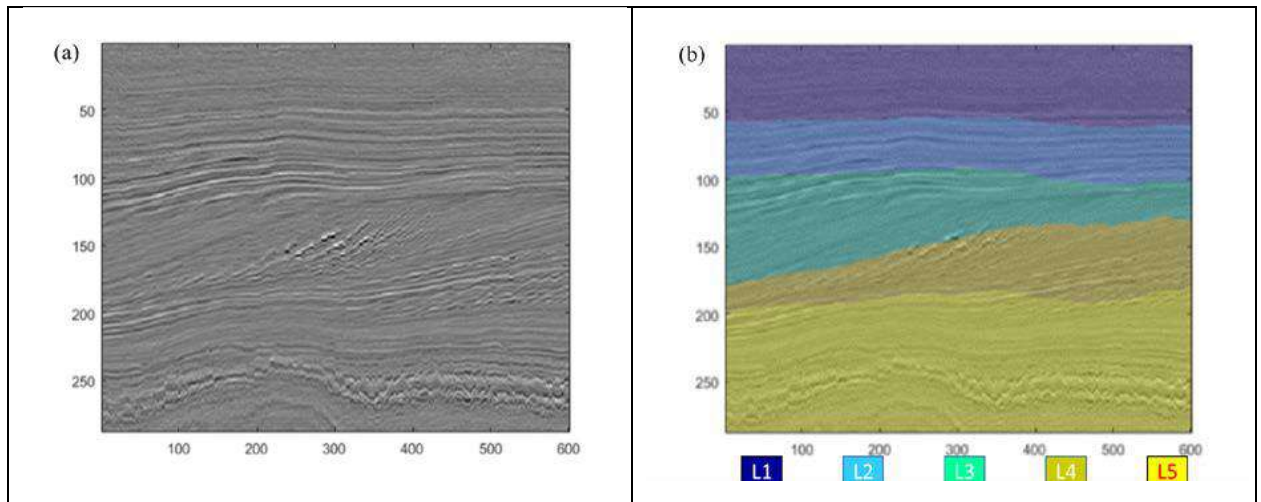


Figure 7 – A section of seismic traces and labeled traces

(a – the seismic traces, b – the seismic traces and their corresponding labels) (Wu, 2018)

After that, a deep convolutional encoder-decoder neural network was constructed. There are four layers in the encoder network and size of convolution kernels are 32, 16, 8, and 3, respectively. In the decoder network, four layers that correspond to the encoder network were also built. The size of the convolution kernels

in the decoder network are 3, 8, 16, and 32, respectively. The training seismic traces and corresponding segment labels are the inputs for the encoder-decoder neural network training. The training process converges after 10 epochs. The validation accuracy in the training process is above 99 %.

To evaluate the result, the predicted horizons were compared with the ground truth horizons. First of all, trained neural network was used to segment the test seismic traces. After that, the horizons were extracted according to the boundary between two segments. Figure 8 show the extracted four horizons in 3D view. Figure 9 shows predicted and “ground truth” horizons overlaid on a representative inline section. The red and blue curves are the predicted and “ground truth” horizons, respectively. It should be noted that there is perfect match between red and blue horizons. The average absolute error between the ground truth horizons and predicted horizons is 1.13 ms. Figure 10 compares interpolated and predicted horizons. The blue, red, and green curves in Figure 9 are the “ground truth”, predicted, and interpolated horizons, respectively. It should be noted that the prediction of horizons successfully follows the seismic reflections across the faults. However, the interpolated horizon fails to follow the local reflectors.

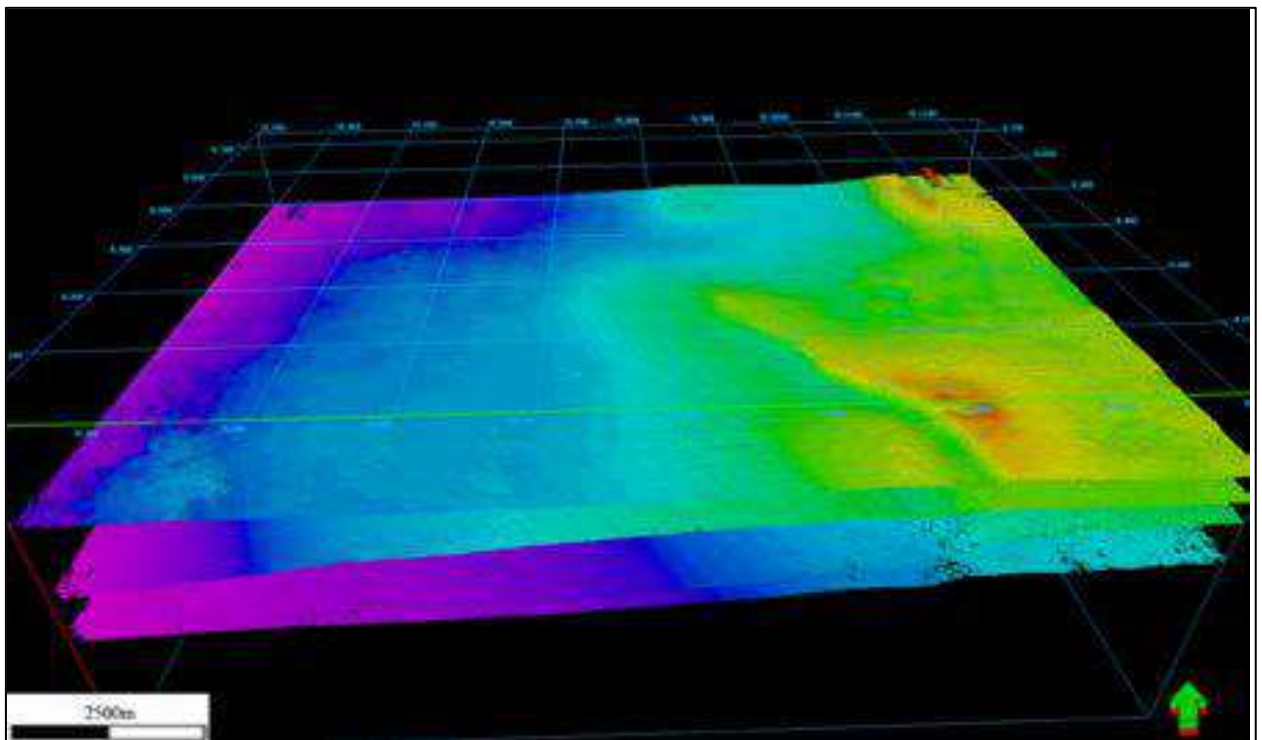


Figure 8 – Predicted horizon volume (Wu, 2018)

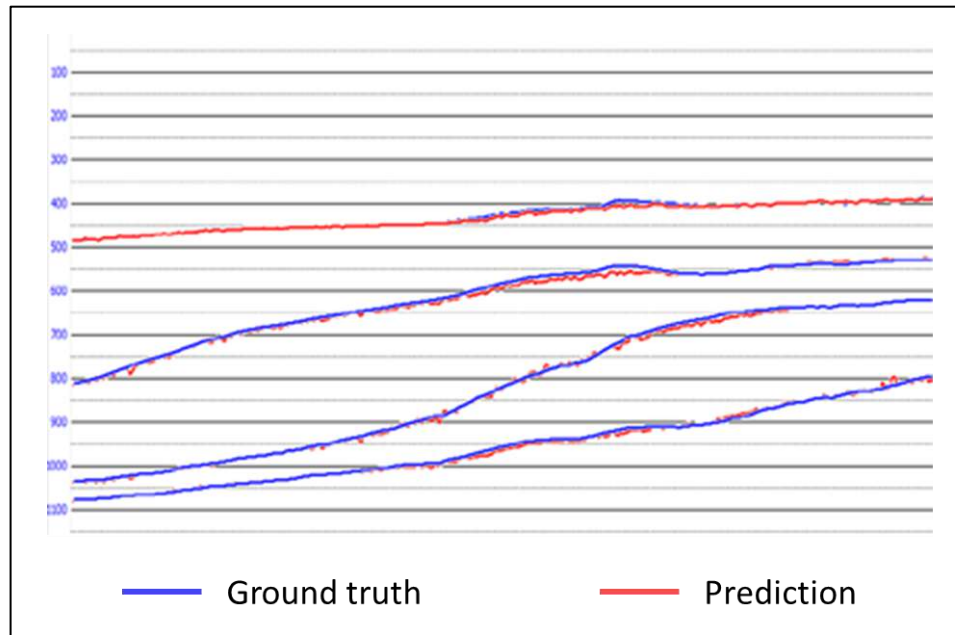


Figure 9 – Comparison between the ground truth horizons and predicted horizons (Wu, 2018)

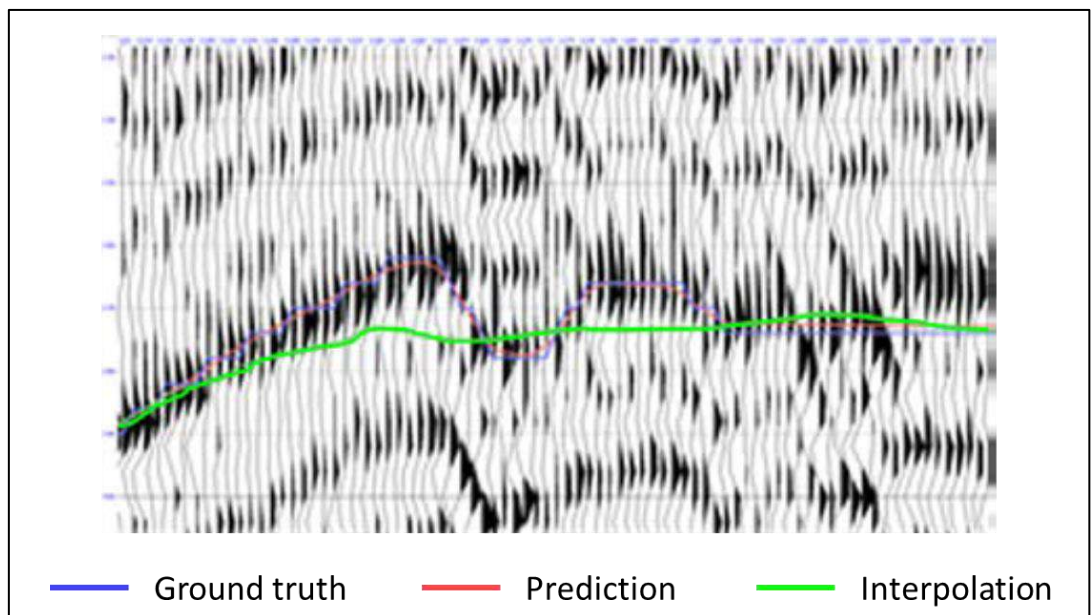


Figure 10 – Comparison between the interpolated and predicted horizons (Wu, 2018)

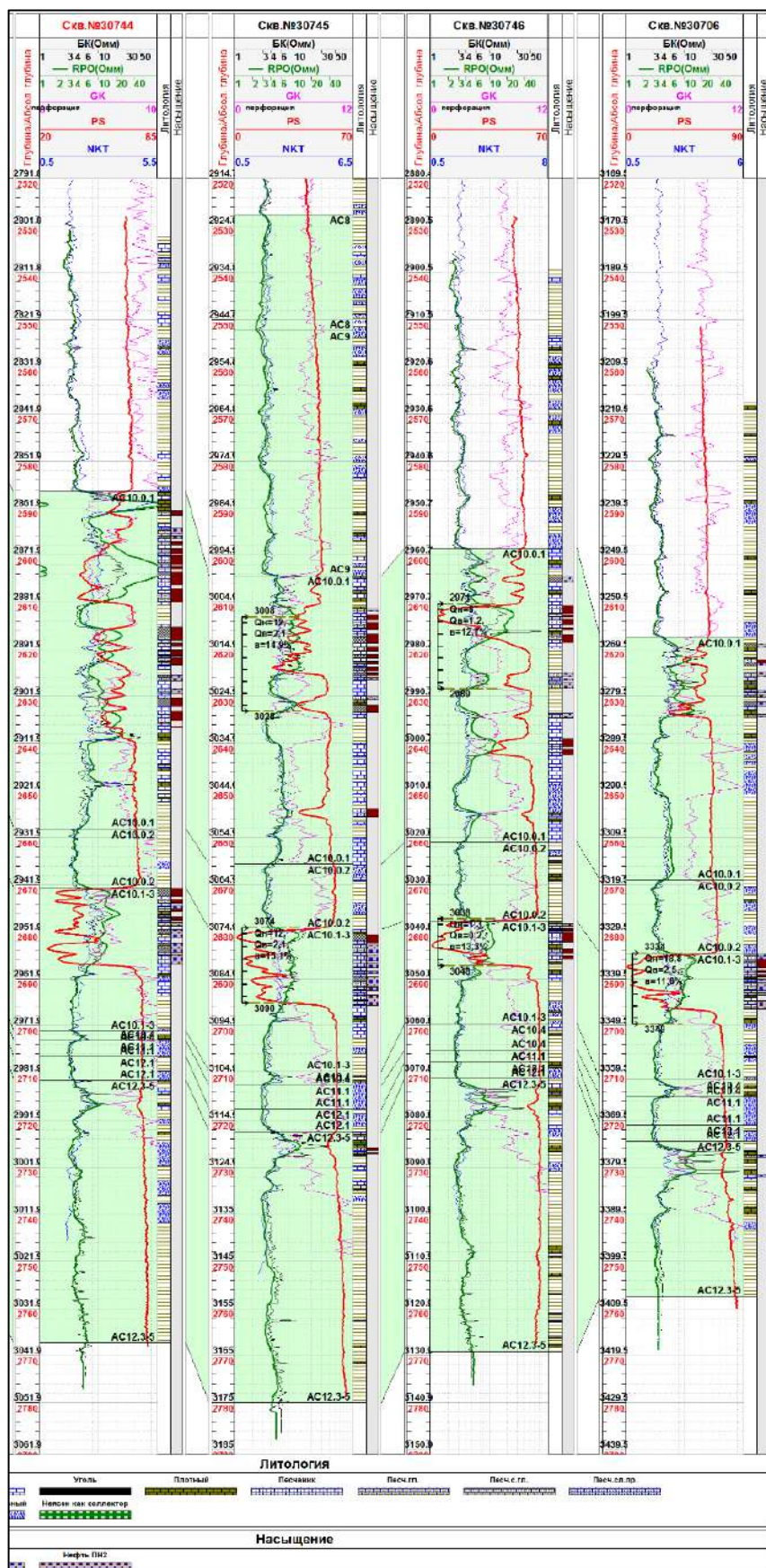
The comparison between the predicted horizons and interpolated horizons illustrates that the proposed method has a better performance during the correlation process at the complicated geological structure zone (Wu, 2018).

References

1. Hall B. Facies classification using machine learning / B. Hall // The Leading Edge, № 35 (2016), p. 906–909.
2. Merembayev T. «Machine Learning Algorithms for Classification Geology Data from Well Logging» / T. Merembayev, R. Yunussov, Y. Amirgaliyev // 14th International Conference on Electronics Computer and Computation (ICECCO). – Almaty, 2018.
3. Wu H. «A deep convolutional encoder-decoder neural network in assisting seismic horizon tracking» / H. Wu, B. Zhang // Deep learning in assisting seismic horizon tracking, 2018.

Приложение Б

Пример предоставленного литологического расчленения скважин



Приложение В

Описание кернового материала скважины 610Р и его особенности

Интервал, м	Описание кернового материала	Особенности
2628- 2635	Песчаник бурый, мелкозернистый, глауконит-кварцевого состава, на карбонатно-глинистом цементе, с запахом УВ и выпотами нефти; в конце интервала наблюдается смятие и тонкая слоистость алевролита	Карбонатно-глинистый цемент, признаки УВ
2635- 2635,5	Тонкое переслаивание аргиллита, алевролита и песчаника; аргиллит светло-серый, средней крепости; алевролит серый, глинистый, в центре слоя встречен прослой песчаника бурого, мелкозернистого, нефтенасыщенного мощностью 15 см.	Признаки УВ в прослое песчаника
2643,5- 2647,7	Песчаник бурый, мелкозернистый, кварцевый, на глинистом цементе порово-базального типа, средней крепости, нефтенасыщенный, с включениями сидерита	Включения сидерита
2647,7- 2648,2	Алевролит серый с бурым оттенком, крупнозернистый, глинистый, средней крепости, с признаками УВ.	Признаки УВ
2648,2- 2649,5	Аргиллит серый, плотный, слоистый, средней крепости, с тонкими прослоями алевролита серого, глинистого	—
2649,5- 2651,5	Песчаник бурый, мелкозернистый, кварцевый, на глинистом цементе, средней крепости, нефтенасыщенный	Признаки УВ
2651,5- 2654,9	Алевролит серый, плотный, массивный, крепкий, в центре интервала встречен прослой аргиллита темно-серого, слоистого, трещиноватость под углом 70-30° к оси керна наблюдаются зеркала скольжения;	Зеркала скольжения
2654,9- 2657,2	Песчаник бурый, мелкозернистый, слюдисто-кварцевый на карбонатно-глинистом цементе, нефтенасыщенный	Карбонатно-глинистый цемент, признаки УВ
2657,2- 2659,9	Алевролит серый, массивный, плотный, с прослоями песчаника бурого, мелкозернистого, на глинистом цементе, с выпотами нефти, наблюдается трещиноватость под углом 60-70° к оси керна, по трещинам распространен углистый детрит;	Трещиноватость, углистый детрит
2659,9- 2661	Песчаник бурый, мелкозернистый, слюдисто-кварцевый, на карбонатно-глинистом цементе, базального типа средней крепости, нефтенасыщенный, с тонкими прослоями алевролита и аргиллита	Карбонатно-глинистый цемент, признаки УВ
2661- 2661,8	Аргиллит темно-серый, слоистый, трещиноватый, средней крепости, зеркала скольжения под углом 40-50°	Зеркала скольжения
2661,8- 2668,4	Песчаник бурый, мелкозернистый, массивный, слюдисто-кварцевый, на глинистом цементе, базально-порового типа средней крепости, нефтенасыщенный	Признаки УВ
2668,4- 2669,5	Песчаник светло-серый, мелкозернистый, глауконит-кварцевого состава, на глинисто-карбонатном цементе базального типа, сильной крепости, глинизированный, с фрагментами аргиллита	Карбонатно-глинистый цемент, глауконит
2669,5- 2670,5	Песчаник бурый, средне-мелкозернистый, слюдисто-кварцевый, на глинистом цементе, базально-порового типа средней крепости, нефтенасыщенный, с выпотами нефти;	Признаки УВ
2670,5- 2671,8	Песчаник бурый, средне-мелкозернистый, слюдисто-глауконит-кварцевый, на глинистом цементе, базально-порового типа, средней пористости, средней крепости, нефтенасыщенный	Глауконит, признаки УВ
2671,8- 2674,6	Алевролит серый, мелкозернистый, слюдистый, плотный, средней крепости, с тонкими прослоями аргиллита	—