

МАГІСТЕРСЬКА РОБОТА

МР.НЗ - 04.00.00.000 ПЗ

Група НЗФм - 24 - 1

Андрій Протас

2025

Міністерство освіти і науки України
Івано-Франківський національний технічний університет нафти і газу

Факультет природничих наук
Кафедра нафтогазової геофізики

Протас Андрій Віталійович

(прізвище, ім'я, по батькові)

УДК 550.8

МАГІСТЕРСЬКА РОБОТА

Застосування методів машинного навчання для прогнозування
фільтраційно-
ємнісних властивостей порід-колекторів (на прикладі Дашавського
родовища)

(назва роботи)

Геофізика

(назва освітньої програми)

103 «Науки про Землю»

(шифр і назва спеціальності)

А. В. Протас

(підпис, ініціали та прізвище здобувача освітнього ступеня)

Науковий
керівник

Багрій Ігор Дмитрович, к.г.н., доцент

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Допущено до захисту
Завідувач кафедри

доцент

В. В. Федорів

(посада) (підпис) (дата) (ініціали та прізвище)

Рецензент

доцент

Я. М. Коваль

(посада) (підпис) (дата) (ініціали та прізвище)

Робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

Івано-Франківськ – 2025 р

Івано-Франківський національний технічний університет нафти і газу

Факультет природничих наук
Кафедра нафтогазової геофізики
Освітній ступінь магістр
Спеціальність 103 «Науки про Землю»

ЗАТВЕРДЖУЮ
Зав. кафедри НГГ
доц. Федорів В. В.
“ _____ ” _____ 2025 р.

**ЗАВДАННЯ
НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТОВІ**

Протасу Андрію Віталійовичу

(прізвище, ім'я, по батькові)

1. Тема роботи Застосування методів машинного навчання для прогнозування фільтраційно-ємнісних властивостей порід-колекторів (на прикладі Дашавського родовища)

Керівник роботи Багрій Ігор Дмитрович

Затверджена наказом закладу вищої освіти від «04» грудня 2025 р. №751/7

2. Строк подання студентом роботи 18 грудня 2025 року

3. Вихідні дані до роботи Завдання на магістерську роботу. Геолого-геофізичні дані з вивчення Дашавського родовища. Спеціалізована література.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)

1) Перелік умовних позначень, символів, одиниць, скорочень і термінів.

2) Вступ.

3) Теоретико-методологічні основи прогнозування фільтраційно-ємнісних властивостей гірських порід.

4) Геолого-геофізична характеристика

5) Вихідні дані та методика дослідження

6) Розробка та обґрунтування комплексної методики прогнозування колекторських властивостей на основі інтелектуального аналізу даних

7) Висновки.

8) Перелік використаних джерел.

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень) _____

Презентація у кількості 14 слайдів

6.Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання 03 листопада 2025 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів магістерської роботи	Термін виконання етапів роботи	Примітка
1.	Теоретико-методологічні основи прогнозування фільтраційно-ємнісних властивостей гірських порід.	03.11.2025-13.11.2025	
2.	Геолого-геофізична характеристика	14.11.2025-20.11.2025	
3	Вихідні дані та методика дослідження	21.11.2025-27.11.2025	
4.	Розробка та обґрунтування комплексної методики прогнозування колекторських властивостей на основі інтелектуального аналізу даних	28.11.2025-15.12.2025	
5.	Перелік умовних позначень, символів, одиниць, скорочень і термінів. Вступ. Висновки. Перелік використаних джерел.	16.12.2025-18.12.2025	

Студент

(підпис)

Керівник роботи

(підпис)

Протас А.В.

(прізвище та ініціали)

Багрій І. Д.

(прізвище та ініціали)

АНОТАЦІЯ

Магістерська робота на тему «Застосування методів машинного навчання для прогнозування фільтраційно-ємнісних властивостей порід-колекторів (на прикладі Дашавського родовища)» містить 61 сторінок, 14 рисунків, 4 таблиць та 10 літературних джерел.

У роботі комплексно розглянуто актуальну проблему підвищення точності прогнозування фільтраційно-ємнісних властивостей (ФЄВ) порід-колекторів, що є ключовим фактором для ефективної розробки родовищ вуглеводнів в Україні. Детально проаналізовано обмеження традиційних петрофізичних моделей, які часто не враховують складну нелінійну природу геологічних процесів, та обґрунтовано доцільність переходу до сучасних інтелектуальних методів аналізу даних.

На основі інтегрованого масиву даних, що включає результати геофізичних досліджень свердловин (ГДС) та лабораторних аналізів ядра з Дашавського родовища, були розроблені, навчені та валідовані предиктивні моделі машинного навчання. Дослідження продемонструвало, що ансамблеві алгоритми, зокрема градієнтний бустинг, дозволяють прогнозувати коефіцієнти пористості та проникності зі значно вищою точністю порівняно зі стандартними емпіричними залежностями, ефективно виявляючи приховані закономірності.

Розроблені моделі продемонстрували високу здатність до узагальнення та можуть бути ефективно застосовані для переінтерпретації даних по всьому фонду свердловин. Це дозволяє суттєво уточнити геологічну модель родовища, провести більш надійний підрахунок запасів, знизити ризики буріння та оптимізувати загальну

стратегію розробки, підвищуючи її економічну ефективність.

Ключові слова: машинне навчання, фільтраційно-ємнісні властивості, пористість, проникність, геофізичні дослідження свердловин, петрофізичне моделювання, випадковий ліс, керн, прогнозування, аналіз даних.

ANNOTATION

Master's thesis on the topic «Application of machine learning methods for predicting the filtration-capacitive properties of reservoir rocks (on the example of the Dashava deposit)» contains 61 pages, 14 figures, 4 tables and 10 references.

The thesis comprehensively addresses the urgent problem of improving the prediction accuracy of filtration-capacitive properties (FCP) of reservoir rocks, a key factor for the effective development of hydrocarbon fields in Ukraine. It thoroughly analyzes the limitations of traditional petrophysical models, which often fail to account for the complex non-linear nature of geological processes, and substantiates the expediency of transitioning to modern intelligent data analysis methods.

Based on an integrated dataset, including results from well logging (WL) and laboratory core analyses from the Dashava deposit, predictive machine learning models were developed, trained, and validated. The research demonstrated that ensemble algorithms, particularly gradient boosting, allow for the prediction of porosity and permeability coefficients with significantly higher accuracy compared to standard empirical dependencies, effectively identifying hidden patterns.

The developed models have shown a high generalization capability and can be effectively applied for re-interpreting data across the entire well stock. This allows for a substantial refinement of the field's geological model, a more reliable reserve estimation, reduced drilling risks, and optimization of the

overall development strategy, thereby increasing its economic efficiency.

Keywords: machine learning, filtration-capacitive properties, porosity, permeability, well logging, petrophysical modeling, Random Forest, core sample, prediction, data analysis.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ ТА СКОРОЧЕНЬ.....	6
ВСТУП.....	8
1 ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ОСНОВИ ПРОГНОЗУВАННЯ ФІЛЬТРАЦІЙНО-ЄМНІСНИХ ВЛАСТИВОСТЕЙ ГІРСЬКИХ ПОРІД.....	10
1.1. Фільтраційно-ємнісні властивості порід-колекторів як ключовий об’єкт геофізичних досліджень.....	10
1.2. Традиційні методи визначення ФЄВ за даними геофізичних досліджень свердловин (ГДС).....	14
1.3. Застосування методів машинного навчання для задач нафтогазової геофізики	19
1.4. Огляд та обґрунтування вибору алгоритмів машинного навчання.....	21
2 ГЕОЛОГО-ГЕОФІЗИЧНА ХАРАКТЕРИСТИКА	26
2.1 Літолого-стратиграфічний опис та характеристика продуктивних пластів	26
2.2 Основні відомості про тектоніку родовища	26
2.3 Газоносність	30
2.4 Фізико-літологічна характеристика колекторів продуктивних пластів та покришок.....	34
3 ВИХІДНІ ДАНІ ТА МЕТОДИКА ДОСЛІДЖЕННЯ	38
3.1 Характеристика вихідних даних для моделювання	38
3.2. Методика підготовки даних та побудови моделей машинного навчання.....	41

4 РОЗРОБКА ТА ОБҐРУНТУВАННЯ КОМПЛЕКСНОЇ МЕТОДИКИ ПРОГНОЗУВАННЯ КОЛЕКТОРСЬКИХ ВЛАСТИВОСТЕЙ НА ОСНОВІ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ	44
4.1 Підготовка та аналіз вихідних даних.....	45
4.2 Формування єдиного інформаційного простору "кern-ГДС"	50
4.3 Розробка та валідація моделей машинного навчання.....	51
4.4 Аналіз та інтерпретація результатів моделювання	54
ВИСНОВКИ	58
СПИСОК ЛІТЕРАТУРНИХ ДЖЕРЕЛ	60

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ ТА СКОРОЧЕНЬ

АК	– Акустичний каротаж
БКЗ	– Бокове каротажне зондування
ГДС	– Геофізичні дослідження свердловин
ГК	– Гамма-каротаж
ІК	– Індукційний каротаж
Кп	– Коефіцієнт пористості
НГК	– Нейтронний гамма-каротаж
ПЕО	– Питомий електричний опір
ПС	– Метод самочинної поляризації
ФЄВ	– Фільтраційно-ємнісні властивості
CSV	– Comma-Separated Values (текстовий формат даних)
GAPI	– Одиниці вимірювання гамма-каротажу (American Petroleum Institute units)
LAS	– Log ASCII Standard (стандартний формат для даних ГДС)
MAE	– Середня абсолютна помилка (Mean Absolute Error)
NaN	– Not a Number (відсутнє або невизначене значення)
R ²	– Коефіцієнт детермінації
EDA	– Exploratory Data Analysis (Розвідувальний аналіз)

даних)

MN - Машинне навчання (зустрічається в анотації та вступі)

RF - Random Forest (Випадковий ліс)

ВСТУП

Актуальність теми. Надійна оцінка запасів вуглеводнів є фундаментом енергетичної безпеки та економічної стабільності держави. Ключовим елементом у цьому процесі є точне визначення фільтраційно-ємнісних властивостей (ФЄВ) порід-колекторів, що безпосередньо впливає на достовірність геологічних моделей родовищ. Класичні методи інтерпретації геофізичних даних, хоч і є основою галузі, часто демонструють обмежену ефективність у складних геологічних умовах, що призводить до значних невизначеностей при підрахунку запасів.

Інтелектуальний аналіз даних за допомогою методів машинного навчання (МН) відкриває нові можливості для вирішення цієї проблеми. Технології МН дозволяють виявляти складні, нелінійні закономірності у геофізичній інформації, які залишаються непоміченими традиційними підходами.

Отже, задача підвищення точності прогнозування ФЄВ шляхом застосування моделей машинного навчання на сьогоднішній день є надзвичайно актуальною.

Мета і задачі дослідження. Головною метою роботи є підвищення достовірності визначення фільтраційно-ємнісних властивостей порід-колекторів Дашавського родовища шляхом розробки та застосування моделей машинного навчання на основі комплексу геофізичних досліджень свердловин.

Для досягнення мети необхідно розв'язати такі задачі:

- проаналізувати сучасний стан та проблеми визначення ФЄВ за даними ГДС;
- обґрунтувати вибір алгоритмів машинного навчання для вирішення поставленої задачі;
- сформувати інтегрований масив даних, що поєднує результати геофізичних досліджень свердловин (ГДС) та лабораторних аналізів

керна, для подальшого навчання моделей машинного навчання;

– розробити, навчити та протестувати прогнозні моделі машинного навчання;

– виконати порівняльний аналіз точності розроблених моделей з традиційними методиками.

Об’єкт дослідження. Процес прогнозування фільтраційно-ємнісних властивостей порід-колекторів на прикладі: стрийських відкладів Дашавського родовища.

Предмет дослідження. Прогнозні моделі машинного навчання для визначення коефіцієнтів пористості та проникності за даними геофізичних досліджень свердловин.

Методи дослідження. Аналіз, узагальнення та систематизація науково-технічної літератури, методи математичної статистики, алгоритми машинного навчання (випадковий ліс, градієнтний бустинг).

Новизна одержаних результатів.

Розроблено та валідовано комплекс прогностичних моделей машинного навчання для Дашавського газового родовища, які забезпечують значно вищу точність прогнозування фільтраційно-ємнісних властивостей (ФЄВ) порід-колекторів (коефіцієнтів пористості та проникності) порівняно зі стандартними петрофізичними залежностями.

Розроблено методику формування інтегрованого масиву даних "кern-ГДС" та обґрунтовано вибір розширеного комплексу геофізичних маркерів (ГК, ПЕО, НГК, ПС), що дозволило подолати проблему неповноти стандартних даних ГДС.

Обсяг і структура роботи. Магістерська робота складається зі вступу, 4 розділів, висновків, списку використаних джерел із 342 найменувань. Повний обсяг магістерської роботи – 61 сторінок друкованого тексту комп’ютерного набору, ілюстрованого 4 таблицями та 14 рисунками.

1 ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ОСНОВИ ПРОГНОЗУВАННЯ ФІЛЬТРАЦІЙНО-ЄМНІСНИХ ВЛАСТИВОСТЕЙ ГІРСЬКИХ ПОРІД

1.1. Фільтраційно-ємнісні властивості порід-колекторів як ключовий об'єкт геофізичних досліджень

Промислова нафтогазоносність геологічних розрізів пов'язана з наявністю в них специфічних гірських порід, здатних не тільки вміщувати, але й віддавати значні обсяги рухомих вуглеводневих флюїдів. Такі породи отримали назву колекторів, а їхні ключові характеристики - фільтраційно-ємнісні властивості (ФЄВ) - є основним об'єктом вивчення при пошуках, розвідці та розробці родовищ нафти і газу. Точність визначення ФЄВ безпосередньо впливає на достовірність підрахунку запасів, вибір оптимальної системи розробки, прогнозування дебітів свердловин та, зрештою, на економічну ефективність усього проекту [1].

Основними параметрами, що кількісно описують ФЄВ, є пористість, проникність та насиченість. [1].

1.1.1 Пористість. Пористість - це наявність у породі пустот (пор, каверн, тріщин), не заповнених твердою складовою породи. Кількісно її виражають через коефіцієнт пористості (K_p або ϕ), який визначається як відношення об'єму порожнин ($V_{\text{пор}}$) до загального об'єму зразка породи ($V_{\text{зразка}}$):

$$\phi = \frac{V_{\text{пор}}}{V_{\text{зразка}}} \quad (1.1)$$

Цей параметр вимірюється в частках одиниці або у відсотках і є фундаментальною характеристикою, що визначає ємність колектора, тобто його здатність вміщувати флюїди [1].

За походженням розрізняють первинну та вторинну пористість (рис. 1.1).

Первинна (седиментаційна) пористість виникає в процесі утворення осадової породи. Найбільш типовим прикладом є міжзернова (інтергранулярна) пористість у пісковиках, де пустоти представлені проміжками між уламковими зернами. Її величина залежить від сортованості, форми та пакування зерен[1].

Вторинна пористість утворюється внаслідок процесів, що відбуваються з породою після її седиментації (діагенез, катагенез). До неї належать пористість розчинення (каверни, вилуговування), доломітизації та тріщинуватість, яка виникає внаслідок тектонічних напружень [1].

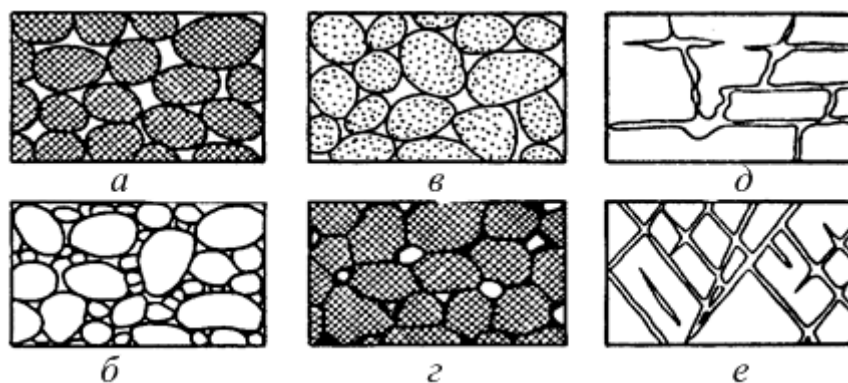


Рисунок 1.1 – Схематичне зображення різних видів пор у гірських породах (за М.Д.Будзом, 2005): а – скельна порода з окремими структурними порами і тріщинами; б – та же порода з пористістю, збільшеною в результаті процесу вивітрювання; в – кавернозна порода з крупними пустотами, в якій проходить вилуговування та розчинення; г – пухка піщана порода з відносно відсортованими зернами і високою пористістю; д – пухка піщана порода з малою пористістю внаслідок нерівної зернистості; е – піщана порода з малою пористістю внаслідок утворення цементу та його нарощуванню навколо окремих зерен; ж – лесоподібна порода з мікро- і макропористістю; з – глиниста мікропориста порода; і – глиниста порода з малою пористістю

внаслідок ущільнення[5]

З практичної точки зору, пористість поділяють на загальну та ефективну (відкриту) [1].

Загальна пористість - це сумарний об'єм усіх пустот у породі, включно з ізольованими порами, що не сполучаються між собою[1].

Ефективна пористість – це об'єм сполучених між собою пор, через які можлива фільтрація флюїду. Саме ефективна пористість має вирішальне значення для нафтогазової промисловості, оскільки лише вуглеводні, що містяться в системі таких пор, можуть бути видобуті. Різниця між загальною та ефективною пористістю може бути значною, особливо в карбонатних та вулканогенних породах[1].

1.1.2 Проникність. Якщо пористість характеризує статичну ємність колектора, то проникність є його ключовою динамічною характеристикою, що описує здатність породи пропускати через себе рідини та газу за наявності градієнта тиску. Ця властивість визначає, наскільки легко флюїди можуть рухатися в пласті до вибоїв свердловин, і, відповідно, є визначальним фактором для продуктивності свердловин [1].

Кількісною мірою проникності є коефіцієнт проникності (k), який був введений французьким інженером Анрі Дарсі в 1856 році. Закон Дарсі для лінійної фільтрації нестисливої рідини має вигляд:

$$Q = \mu \cdot Lk \cdot A \cdot \Delta P; \quad (1.2)$$

де Q – об'ємна витрата рідини ($\text{м}^3/\text{с}$);

k – коефіцієнт проникності (м^2 або Дарсі);

A – площа поперечного перерізу зразка (м^2);

ΔP – перепад тиску на довжині L (Па);

μ – динамічна в'язкість рідини (Па·с);

L – довжина шляху фільтрації (м) [6].

Одиницею вимірювання проникності в системі СІ є м^2 , однак у промисловій практиці широкого розповсюдження набула позасистемна одиниця Дарсі (Д). 1 Дарсі приблизно дорівнює $10\text{-}12\text{м}^2$. Колектори з проникністю менше 0.001 Д (1 мД) вважаються практично непроникними або належать до категорії нетрадиційних (щільні пісковики, сланці) [2].

Розрізняють декілька видів проникності:

Абсолютна проникність – це проникність породи при її 100% насиченні однорідним флюїдом, який не взаємодіє з породою [1].

Ефективна (фазова) проникність -- це проникність породи для одного з флюїдів за наявності в поровому просторі інших флюїдів, що не рухаються [1].

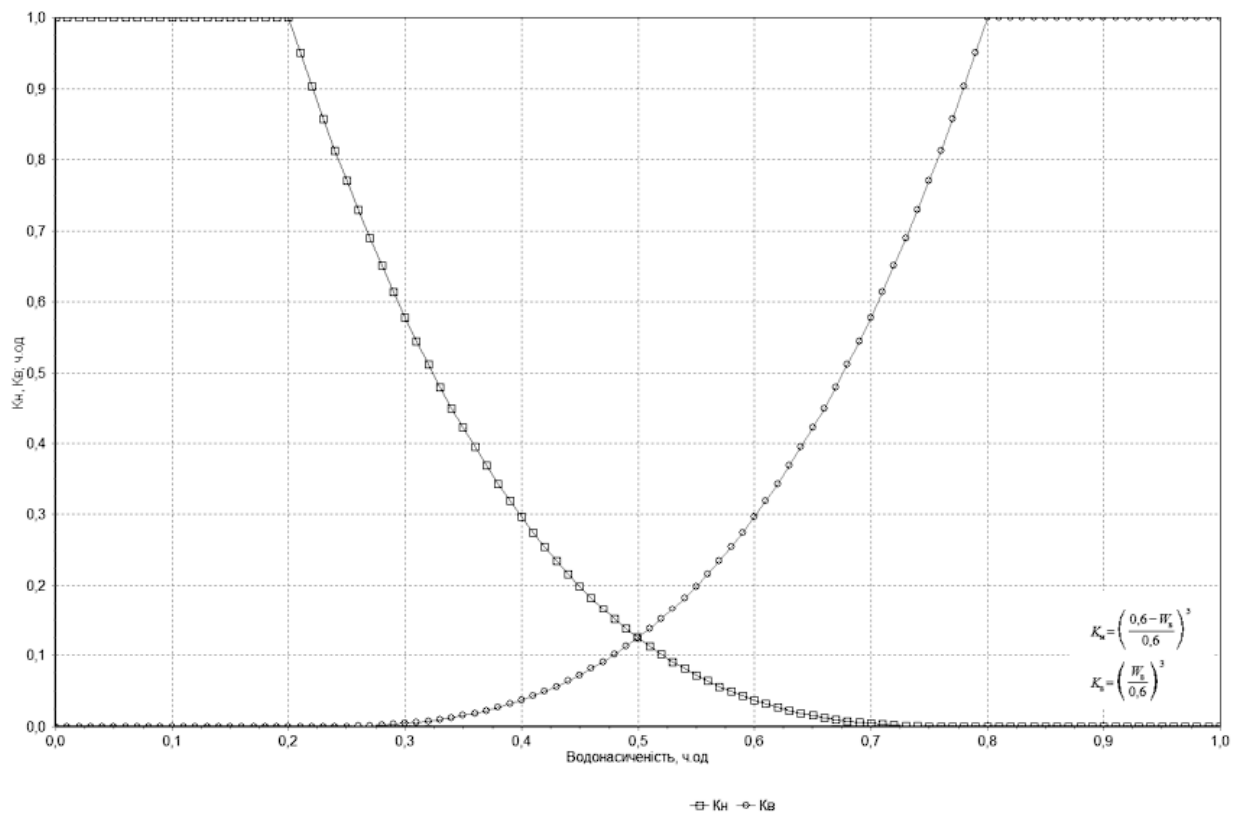


Рисунок 1.2 – Характер прийнятих для моделювання фазових проникностей для нафти та води

Відносна проникність – це відношення ефективної проникності до абсолютної. Вона є функцією насиченості і показує, як змінюється

здатність породи пропускати, наприклад, нафту при зміні її водонасичення[1].

Зв'язок між пористістю та проникністю є складним і неоднозначним. Загалом, вища пористість часто корелює з вищою проникністю, але цей зв'язок сильно залежить від розміру пор, їхньої форми та зв'язності (звивистості шляхів фільтрації) (рис. 1.2). Порода з високою пористістю може мати низьку проникність, якщо пори погано сполучені між собою (наприклад, пемза) [1].

1.1.3 Насиченість. Насиченість характеризує ступінь заповнення порового простору колектора різними флюїдами. Коефіцієнт насиченості (S) для будь-якого флюїду визначається як відношення об'єму, зайнятого цим флюїдом ($V_{\text{фл}}$), до загального об'єму порового простору ($V_{\text{пор}}$). Сума коефіцієнтів насиченості всіх флюїдів у породі дорівнює одиниці:

$$S_w + S_o + S_g = 1 \quad (1.3)$$

де S_w , S_o , S_g - коефіцієнти насичення водою, нафтою та газом відповідно.

Важливим поняттям є залишкова (реліктова) водонасиченість ($S_{w,irr}$). Це частка зв'язаної води, яка утримується в дрібних порах та на поверхні зерен молекулярними силами і не рухається разом з вуглеводнями. Саме від її величини залежить об'єм вуглеводнів у пласті. Визначення початкової водонасиченості є одним з ключових завдань при підрахунку запасів [1].

1.2. Традиційні методи визначення ФЄВ за даними геофізичних досліджень свердловин (ГДС)

Пряме вивчення ФЄВ можливе лише шляхом дороговартісного

відбору та лабораторного аналізу зразків породи – керна. Однак kern відбирається лише в обмежених інтервалах опорних свердловин. Тому основним інструментом для суцільного (по всьому стовбуру свердловини) та опосередкованого вивчення властивостей порід є геофізичні дослідження свердловин (ГДС). Методи ГДС базуються на вимірюванні різноманітних фізичних полів (електричних, ядерних, акустичних), характеристики яких залежать від складу та властивостей порід, що перетинаються свердловиною [2].

1.2.1 Методи літологічного розчленування та визначення глинистості. Першим кроком інтерпретації є розчленування розрізу на пласти та визначення їх літології. Основним методом для цього є гамма-каротаж (ГК). Він вимірює природну гамма-радіоактивність гірських порід, яка здебільшого пов'язана з наявністю радіоактивних ізотопів калію (^{40}K), урану (U) та торію (Th). Глинисті мінерали, як правило, мають високу радіоактивність, тоді як чисті пісковики та вапняки – низьку. Це дозволяє легко розрізнити їх на каротажних діаграмах (рис. 1.3) [2].

1.2.2 Методи визначення пористості (порозністні методи). Для кількісної оцінки пористості використовують комплекс з трьох основних методів:

Гамма-гамма каротаж густинний (ГГК-Г): Метод базується на опроміненні стінки свердловини гамма-квантами та реєстрації розсіяного випромінювання. Інтенсивність розсіяння залежить від електронної густини породи, яка тісно пов'язана з її об'ємною густиною (ρ_b). Знаючи густину скелету породи (ρ_{ma}) та флюїду (ρ_f), пористість розраховують за формулою:

$$\phi_D = \rho_{ma} - \rho_f \cdot \rho_{ma} - \rho_b \quad (1.4)$$

Цей метод є одним з найнадійніших для визначення пористості.

Нейтронний каротаж (НК): Прилад НК випромінює швидкі нейтрони, які зіштовхуються з ядрами атомів у породі, сповільнюються. Найефективніше сповільнення відбувається при зіткненні з ядрами водню, які мають майже таку ж масу. Тому показання НК відображають вміст водню в пласті, який переважно міститься у флюїдах (воді, нафті) в поровому просторі. Таким чином, показання НК є мірою пористості породи [2].

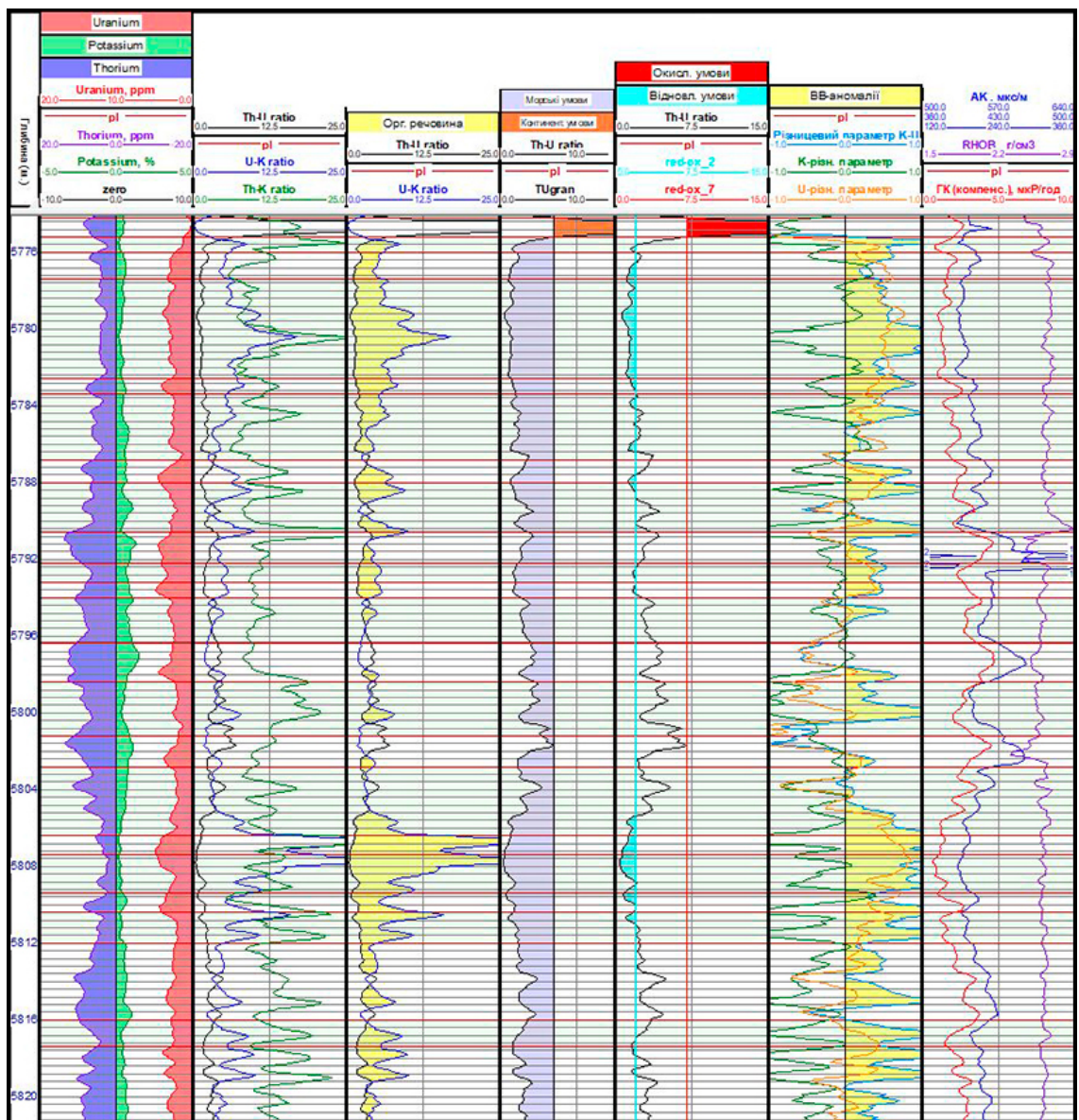


Рисунок 1.3 – Приклад планшета з комплексом даних ГДС (ГК - гамма-каротаж, RHOV - густина, NPHI - нейтронна пористість, АК – час

проходження хвилі по породі)

Акустичний каротаж (АК): Метод вимірює швидкість поширення пружних (звукових) хвиль у породі. Швидкість залежить від пружних властивостей як скелету породи, так і флюїду, що її насичує. Пористість розраховується за емпіричним рівнянням середнього часу Вайлі:

$$V_{p1} = V_{fl\phi} + V_{ma1} - \phi \quad (1.5)$$

де V_p, V_{fl}, V_{ma} - швидкості хвилі в породі, флюїді та скелеті відповідно [2].

1.2.3 Методи визначення характеру насичення. Для розрізнення водоносних та вуглеводненасичених колекторів використовують методи електричного каротажу, які вимірюють питомий електричний опір (R_t) гірських порід. Вуглеводні (нафта, газ) є діелектриками, тоді як пластові води, як правило, є мінералізованими розчинами солей і добре проводять електричний струм. Отже, колектори, насичені вуглеводнями, матимуть значно вищий опір, ніж ті ж самі колектори, але заповнені водою [2].

Кількісний зв'язок між водонасиченістю, пористістю та опором встановлюється рівнянням Арчі (1942 р.):

$$S_{wn} = \phi_m \cdot R_{ta} \cdot R_w \quad (1.6)$$

де S_w - коефіцієнт водонасичення;

R_w - питомий опір пластової води;

ϕ - пористість;

R_t - істинний питомий опір пласта;

a, m, n - емпіричні коефіцієнти (літології, цементациї та насичення)

[6].

Рівняння Арчі є наріжним каменем петрофізичної інтерпретації, але

воно має суттєві обмеження, зокрема в глинистих колекторах, де провідність глин спотворює результати [10].

1.2.4 Методи оцінки коефіцієнта проникності. Визначення коефіцієнта проникності ($K_{пр}$) за даними ГДС є складним завданням, оскільки прямих методів його вимірювання (окрім ядерно-магнітного каротажу та випробувачів пластів на кабелі) не існує. У традиційній практиці проникність оцінюють опосередковано через кореляційні зв'язки з пористістю (K_p) та залишковою водонасиченістю ($K_{в.з}$) [1]. Найчастіше використовують емпіричні залежності типу Козені-Кармана або Тімура, які базуються на фізиці порового простору. Загальний вигляд таких рівнянь:

$$K_{\{пр\}} = A \cdot (\{\phi^B\} \setminus \{S_{\{wi\}}^C\}) \quad (1.7)$$

де $K_{пр}$ – проникність;

ϕ – пористість;

S_{wi} – залишкова водонасиченість;

A, B, C – емпіричні коефіцієнти, що налаштовуються за даними керна для конкретного родовища [3].

Також широко застосовується метод парної регресії «кern-ГДС», особливо для теригенних колекторів. У цьому випадку будується логарифмічна залежність проникності від пористості:

$$(K_{\{пр\}}) = a \cdot K_p + b \quad (1.8)$$

де a і b – коефіцієнти регресії.

Однак точність таких методів суттєво залежить від літологічної однорідності колектора та структури порового простору. У складнобудованих розрізах, де на фільтраційні властивості впливають тріщинуватість, глинистість та текстурні особливості, прості емпіричні

залежності дають значну похибку, що і зумовлює необхідність використання методів машинного навчання [2].

1.2.5 Обмеження традиційних методів. Незважаючи на свою фундаментальну важливість, традиційні методи інтерпретації мають низку недоліків:

Емпіричний характер: Більшість формул є емпіричними залежностями з коефіцієнтами, які потребують локального калібрування за даними керна.

Спрощені моделі: Вони базуються на спрощених уявленнях про будову породи і не завжди враховують складну геометрію порового простору, наявність кількох типів пористості, анізотропію тощо [2].

Нелінійність: Залежності між фізичними властивостями та ФЄВ часто є складними і нелінійними, тоді як класичні моделі переважно лінійні.

Проблема проникності: Не існує методу ГДС, який би прямо вимірював проникність. Її оцінка зазвичай проводиться за допомогою дуже ненадійних кореляцій "пористість-проникність"[2].

Ці обмеження стають особливо критичними в колекторах зі складною будовою (карбонати, тонкошаруваті розрізи, низькопористі колектори), що вимагає пошуку нових, більш потужних інструментів аналізу.

1.3. Застосування методів машинного навчання для задач нафтогазової геофізики

Машинне навчання (МН) – це підрозділ штучного інтелекту, що займається розробкою алгоритмів, які дозволяють комп'ютерам "навчатися" на основі даних без явного програмування. Замість того, щоб слідувати жорстко заданим інструкціям, система МН аналізує вхідні дані,

виявляє в них закономірності і використовує їх для прогнозування або прийняття рішень [3].

У нафтогазовій геофізиці, де щодня генеруються величезні обсяги даних (сейсміка, ГДС, дані видобутку), застосування МН стало справжньою революцією. Головна перевага МН полягає в його здатності працювати з багатовимірними, зашумленими та неповними даними, виявляючи складні, нелінійні взаємозв'язки, які неможливо описати простими аналітичними рівняннями [3].

Загалом, машинне навчання поділяють на три основні типи: навчання з учителем (Supervised Learning), навчання без учителя (Unsupervised Learning) та навчання з підкріпленням (Reinforcement Learning), кожен з яких має свою сферу застосування у геонауках (рис. 1.4).

Стандартний робочий процес застосування МН у петрофізиці виглядає наступним чином:

Збір та підготовка даних: Формується єдина таблиця, де для певних глибин є як вхідні ознаки (показання різних методів ГДС), так і цільова змінна (значення пористості або проникності, виміряне на керні) [7].

Розділення вибірки: Дані діляться на навчальну (для тренування моделі), валідаційну (для налаштування параметрів) і тестову (для фінальної оцінки точності) вибірки [3].

Навчання моделі: Алгоритм МН "вивчає" залежність між даними ГДС та ФЄВ на навчальній вибірці.

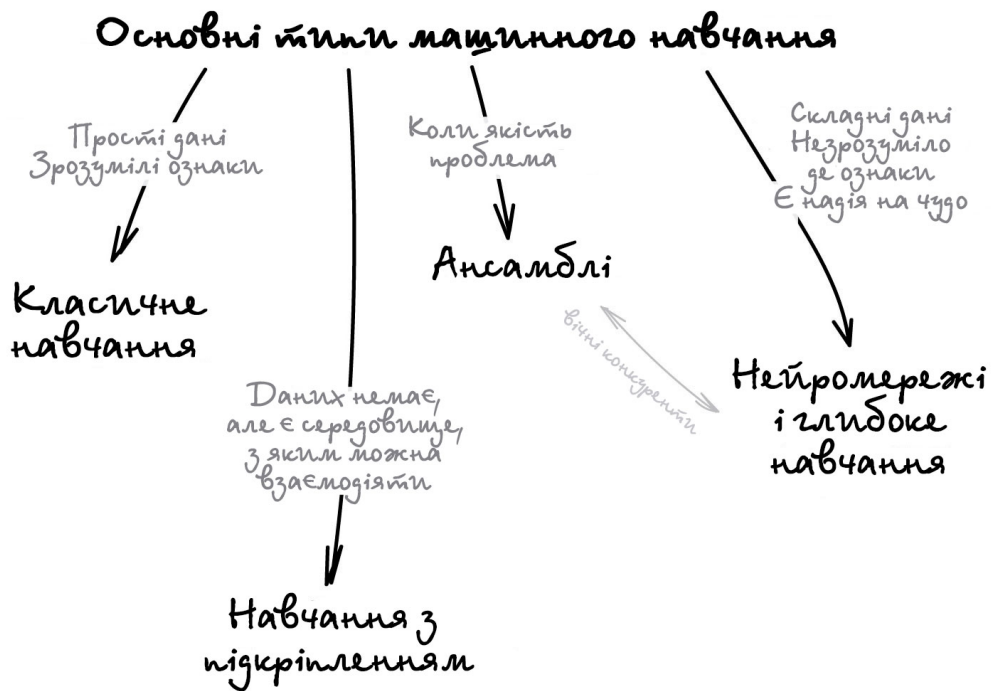


Рисунок 1.4 – Основні типи машинного навчання

Оцінка та інтерпретація: Точність моделі оцінюється на тестових даних, які модель "не бачила" в процесі навчання.

Застосування (прогноз): Навчена модель використовується для прогнозування ФЄВ в інтервалах свердловин, де є лише дані ГДС, але відсутній керн [3].

Застосування МН дозволяє вирішувати широкий спектр завдань: автоматизована ідентифікація літології, прогнозування пропущених кривих ГДС, визначення ФЄВ, аналіз сейсмічних даних та оптимізація видобутку [7].

1.4. Огляд та обґрунтування вибору алгоритмів машинного навчання

Для задачі прогнозування неперервних величин, якою є визначення пористості та проникності, використовуються регресійні алгоритми МН. Розглянемо найбільш поширені з них (рис. 1.5) [7].

1.4.1 Штучні нейронні мережі (Artificial Neural Networks, ANN). ANN - це клас алгоритмів, натхнених біологічними нейронними мережами людського мозку. Найпростіша архітектура - багатошаровий перцептрон (MLP) - складається з вхідного шару (нейрони відповідають вхідним ознакам, наприклад, кривим ГДС), одного або кількох прихованих шарів та вихідного шару (прогнозне значення ФЄВ) [7].

Переваги: Здатність моделювати надзвичайно складні нелінійні залежності. ANN були одними з перших методів МН, які показали високу ефективність у петрофізиці ще в 1990-х роках [7].

Недоліки: Висока обчислювальна складність, потреба у великих обсягах даних для навчання, чутливість до налаштування параметрів (кількість шарів, нейронів) та ефект "чорної скриньки" – складність інтерпретації того, як саме модель прийшла до свого рішення.

1.4.2 Метод опорних векторів (Support Vector Machines, SVM). Для задач регресії використовується модифікація методу, що називається регресією на опорних векторах (Support Vector Regression, SVR). Ідея полягає в тому, щоб знайти таку функцію (гіперплощину), яка б апроксимувала дані, залишаючи якомога більше точок у межах заданого "коридору" помилок. Метод є ефективним на даних високої розмірності [7].

Переваги: Гарна узагальнююча здатність, менша схильність до перенавчання порівняно з ANN [7].

Недоліки: Повільна робота на великих наборах даних, чутливість до вибору функції ядра (kernel), яка перетворює дані у простір вищої розмірності.

1.4.3 Ансамблеві методи: Випадковий ліс та Градієнтний бустинг. Ансамблеві методи є одними з найпотужніших у сучасному машинному навчанні. Їхня основна ідея полягає в тому, щоб об'єднати велику кількість простих, "слабких" моделей (зазвичай, дерев рішень) в єдиний

"комітет", який приймає значно точніші рішення. [7].

Випадковий ліс (Random Forest, RF) Цей алгоритм будує велику кількість (сотні або тисячі) незалежних дерев рішень. Кожне дерево навчається на випадковій підвибірці з навчального набору даних (метод беггінгу). Крім того, при побудові кожного вузла дерева використовується не весь набір ознак, а лише їх випадкова підмножина. Фінальний прогноз усереднюється по всіх деревах [7].

Неглибока нейронна мережа

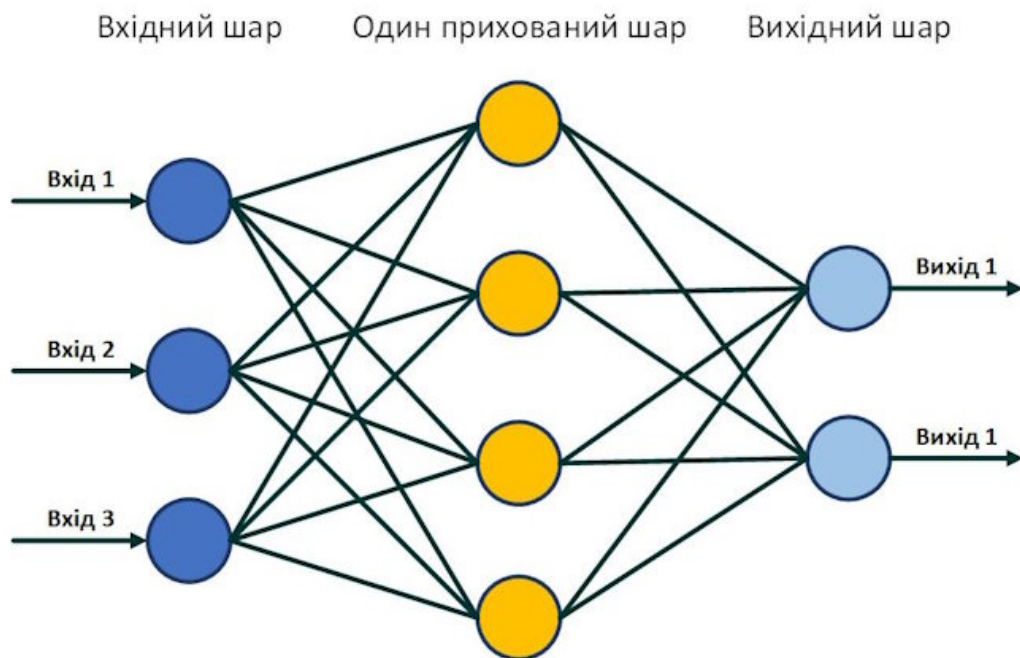


Рисунок 1.5 – Схематичне зображення архітектури штучної нейронної мережі

Переваги: висока точність, стійкість до перенавчання, нечутливість до масштабування ознак, здатність оцінювати важливість кожної вхідної ознаки (наприклад, яка крива ГДС робить найбільший внесок у прогноз).

Недоліки: може бути повільним на дуже великих даних, менш інтерпретований, ніж одне дерево рішень [7].

Градiєнтний бустинг (Gradient Boosting Machine, GBM) На відміну від

RF, де дерева будуються паралельно і незалежно, в бустингу вони будуються послідовно. Кожне наступне дерево навчається на помилках (залишках) попереднього, намагаючись їх виправити. Це дозволяє ансамблю поступово, ітеративно наближатися до ідеального прогнозу. XGBoost та LightGBM є сучасними, високоефективними реалізаціями цього підходу, що включають механізми регуляризації для боротьби з перенавчанням.

Переваги: Зазвичай демонструє найвищу точність серед усіх алгоритмів на структурованих (табличних) даних, якими є дані ГДС.

Недоліки: Більш чутливий до налаштування гіперпараметрів порівняно з RF, навчання може бути довшим.

Для даної магістерської роботи саме ансамблеві методи, Random Forest та Gradient Boosting, були обрані як основні через їхню стабільно високу точність, надійність та здатність працювати з типовим обсягом петрофізичних даних, що було підтверджено у численних наукових роботах [8].

Висновок

Проведений аналіз науково-технічної літератури показав, що фільтраційно-ємнісні властивості є комплексною характеристикою порід-колекторів, точне визначення якої є критично важливим для нафтогазової галузі. Традиційні методи інтерпретації даних ГДС, що базуються на спрощених фізичних моделях та емпіричних рівняннях, мають суттєві обмеження, особливо при роботі зі складними, неоднорідними колекторами.

Встановлено, що головними недоліками класичних підходів є їхня нездатність враховувати складні нелінійні взаємозв'язки між геофізичними параметрами та властивостями пласта, а також відсутність

надійних методів для прямого прогнозування проникності.

Сучасним та перспективним напрямком для подолання цих обмежень є застосування методів машинного навчання. Алгоритми, зокрема ансамблеві моделі, такі як Випадковий ліс та Градієнтний бустинг, здатні виявляти складні закономірності у багатовимірних геофіричних даних, забезпечуючи значно вищу точність прогнозування. Успішне застосування цих методів базується на наявності якісного набору даних, що інтегрує результати ГДС та лабораторних досліджень керна, і відкриває шлях до створення більш достовірних та детальних геологічних моделей родовищ.

2 ГЕОЛОГО-ГЕОФІЗИЧНА ХАРАКТЕРИСТИКА

2.1 Літолого-стратиграфічний опис та характеристика продуктивних пластів

Геологічна будова Дашавського газового родовища є складною та багатоярусною. Розкритий свердловинами геологічний розріз родовища включає потужні товщі осадових порід, що належать до палеозойської, мезозойської та кайнозойської груп (рис. 2.1). Переважна більшість свердловин на родовищі були запроектовані на горизонт НД-9 і розкрили нижньосарматські відклади, тоді як більш давні відклади розкриті лише поодинокими законтурними свердловинами. Найдавнішими породами, встановленими бурінням, є складчасті відклади палеозою, поверхня яких сильно еродована та незгідно перекривається мезозойськими утвореннями [10].

Мезозойська група (MZ) у розрізі родовища представлена юрською та крейдовою системами. Юрські відклади, що є найдавнішими з мезозойських у цьому районі, представлені карбонатними породами і залягають зі стратиграфічною незгідністю на розмитій поверхні палеозойського фундаменту [10].

Крейдова система (K) представлена верхнім відділом. Літологічно ці відклади складені вапняками, вище яких залягають теригенні породи. Вапняки характеризуються сірим, світло- та темно-сірим забарвленням, мають приховано-кристалічну, місцями крупнокристалічну структуру та містять численні прожилки кальциту. Вгору по розрізу вапняки поступово переходять у піщанисті різновиди - сірі, щільні, з численними включеннями глауконіту та ділянками збагачення оксидами заліза, а згодом змінюються пісковиками. Серед товщі пісковиків зустрічаються прошарки глини (сірих, темно-сірих, щільних, карбонатних) та алевролітів

(світло-сірих, слабокарбонатних, глауконітових) [10].

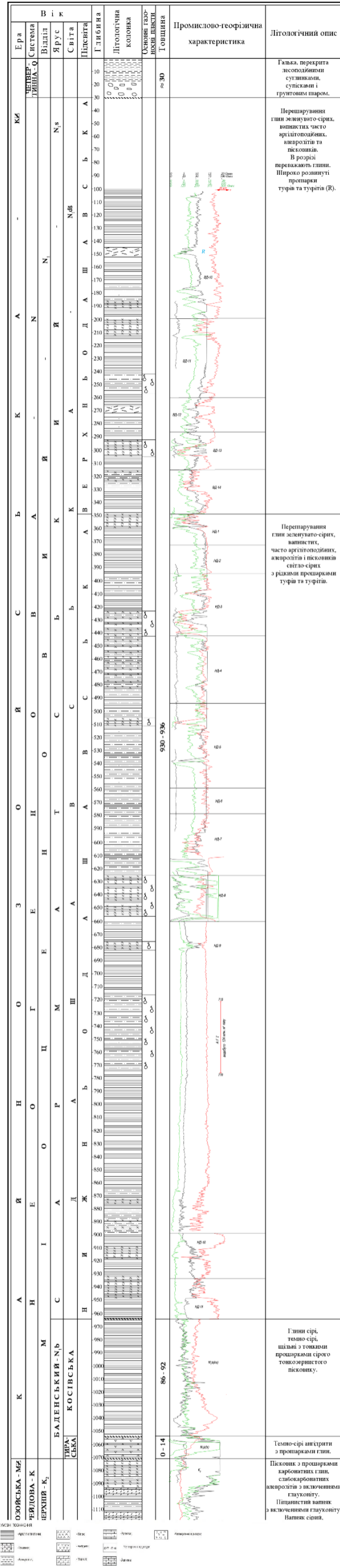


Рисунок 2.1 – Зведений геолого-геофізичний розріз продуктивної частини Дашавського родовища

Кайнозойська група (KZ) формує основну частину осадового чохла і представлена комплексом неогенових відкладів міоценового відділу N_1 , який включає баденський та сарматський яруси. Міоценові породи трансгресивно, зі стратиграфічною незгідністю, залягають на розмитій поверхні верхньої крейди, виповнюючи нерівності давнього ерозійного рельєфу [10].

Баденський ярус N_{1b} у межах родовища об'єднує баранівські верстви, а також тираську та косівську світи. Баранівські верстви N_{1br} , розкриті, наприклад, у свердловині 170-Дш, представлені комплексом мергелисто-глинистих порід з прошарками вапнистих пісковиків. Їхня товщина варіює від 0 до 27 м. Вище залягають відклади тираської світи N_{1tr} , відомі як гіпсо-ангідритовий горизонт. Вони складені темно-сірими ангідритами з тонкими прошарками глин. Характерною особливістю є те, що в склепінній частині Дашавського підняття гіпсо-ангідритовий горизонт відсутній внаслідок розмиву або невідкладення, а його товщина на крилах змінюється від 0 до 14 м. Розріз бадену завершують породи косівської світи N_{1ks} , що незгідно залягають на тираських відкладах. Світа складена сірими та темно-сірими щільними глинами з тонкими прошарками сірого дрібнозернистого пісковика. Товщина косівської світи коливається в межах 14-92 м [10].

Сарматський ярус N_{1s} відіграє ключову роль у геологічній будові родовища, оскільки саме до нього приурочені основні промислові поклади газу. Відклади сармату незгідно перекривають породи баденського ярусу і представлені потужною товщею дашавської світи. Світа поділяється на дві підсвіти: нижньодашавську N_{1dsh1} та верхньодашавську N_{1dsh2} [10].

Літологічно дашавська світа є флішоподібним чергуванням піщано-глинистих порід. Розріз складений частим перешаруванням пісковиків та алевролітів з глинами та аргілітами, зрідка зустрічаються прошарки

туфів і туфітів. Пісковики та алевроліти мають сірий, світло-сірий, іноді жовтувато- або зеленувато-сірий колір. За гранулометричним складом вони переважно дрібно- та середньозернисті, характеризуються різним ступенем вапнистості. Глини, що розділяють піщані пласти, – сірі, зеленувато-сірі, темно-сірі, вапнисті, слюдисті, щільні. У розрізі верхньодашавської підсвіти глинисті різновиди переважають над піщаними [10].

Для детальної кореляції розрізів свердловин на родовищі прийнято схему погоризонтного розчленування. У розкритому розрізі нижньодашавської підсвіти виділяють 10 горизонтів: від НД-1 (у покрівлі) до НД-10 (у подошві). Верхньодашавська підсвіта включає горизонти ВД-9, ВД-10, ВД-11, ВД-12, ВД-13 та ВД-14. Товщина окремих піщано-алевролітових горизонтів є непостійною, вони часто лінзуються або заміщуються глинами по простяганню. Максимальна розкрита товщина сарматських відкладів сягає 1207 м [10].

Для даної роботи особливий інтерес становлять продуктивні горизонти, що є об'єктами дослідження та моделювання:

Горизонт ВД-11. У його розрізі серед глинистої товщі виділяються дві піщано-алевролітові пачки. Верхня (ВД-11а) є водонасиченою, а нижня (ВД-11б) містить промислові запаси газу. Ефективна газонасичена товщина змінюється від 3,0 до 7,8 м.

Горизонт ВД-13а. Характеризується значною неоднорідністю. Ефективна газонасичена товщина змінюється в широких межах від 2,4 м до 9,2 м. Колектори представлені чергуванням прошарків пісковиків, алевролітів та глин.

Горизонти НД-4 (НД-4а, НД-4б). Залягають у нижньосарматській товщі. Характеризуються як пластові склепінні або літологічно екрановані поклади з ефективними товщинами до 5-6 м [10].

Розріз завершується відкладами четвертинної системи (Q), які мають

повсюдне поширення на площі родовища. Вони залягають на розмитій поверхні сармату і представлені галькою, перекритою лесоподібними суглинками, супісками та ґрунтовим шаром. Товщина четвертинного покриву варіює від 1-2 до 10 метрів, місцями досягаючи 30 метрів [10].

2.2 Основні відомості про тектоніку родовища

У регіональному тектонічному плані Дашавське газове родовище приурочене до центральної частини Косівсько-Угерської підзони Більче-Волицької зони Передкарпатського прогину. Геологічна будова цієї підзони характеризується наявністю потужного неогенового чохла, що перекриває двоярусний мезопалеозойський фундамент. Палеозойські та мезозойські породи фундаменту розкриті лише окремими свердловинами на периферії структури [10].

Формування структурних планів баденських та сарматських відкладів відбувалося під впливом складних тектонічних рухів. Визначальну роль відігравали давній ерозійний рельєф поверхні фундаменту, сформований у палеоген-міоценовий час, а також активний вплив насуву Карпатських споруд (Стебницький насув), розташованого у безпосередній близькості [10].

Відклади бадену та сармату утворюють структури обволікання, які мають як антиклінальний, так і синклінальний характер. Власне Дашавська структура по покрівлі опорного відбиваючого горизонту НД-8 (нижній сармат) являє собою велику асиметричну брахіантикліналь. Складка витягнута у північно-західному напрямку (субпаралельно до простягання Карпат). Її характерною особливістю є наявність структурного носу в північно-західній частині. Основне склепіння складки зміщене на південний схід. Асиметрія структури проявляється у будові її крил: південно-західне крило є крутішим, тоді як північно-східне

– більш пологим і протяжним. Крила складки плавно занурюються у відповідних напрямках [10].

Вгору по розрізу, у відкладах верхнього сармату, загальний структурний план зберігається, проте спостерігаються певні особливості. По горизонтах у районі структурного носу (район свердловин 38, 107-Дш) фіксуються локальні малоамплітудні склепіння. Також відзначається міграція основної склепінної частини структури (наприклад, по горизонту ВД-14 склепіння зміщується в район свердловини 57-Дш). Така мінливість структурних планів значною мірою зумовлена літологічним фактором: піщано-алевролітові тіла є невитриманими по площі, часто утворюють лінзи та виклинюються, що ускладнює структурні побудови [10].

Важливим тектонічним елементом є Стебницький насув. Південно-західне крило Дашавської складки по верхньодашавських відкладах (горизонти ВД-9 – ВД-12) зрізане площиною цього насуву. Це підтверджено даними буріння у свердловинах 165-Дш, 49-Дш, 65-Дш та 212-Дш. Насув відділяє Дашавське родовище від структур Внутрішньої зони прогину [10].

Дашавська структура ускладнена розривними порушеннями (диз'юнктивами). Відклади крейди та бадену розбиті системою порушень на окремі тектонічні блоки, при цьому амплітуда зміщення по розломах становить 25-100 м. У вищезалягаючих нижньосарматських відкладах амплітуда цих порушень затухає і зменшується до 10-30 м. У верхній, приповерхневій частині розрізу (зокрема, в інтервалі горизонту ВД-13) значні тектонічні порушення вже не простежуються, і структура набуває вигляду більш спокійної плікативної складки [10].

Родовище оточене низкою інших газоносних структур: Любешівською, Північно-Зарічнлянською, Бережницькою (з південного сходу) та Комарівською (з південного заходу). Від сусідніх підняттяв

Дашавська структура відділена глибокими синклінальними прогинами, які чітко фіксуються як за даними сейсмозондування, так і за даними глибокого буріння [10].

2.3 Газоносність

Дашавське родовище є одним із найстаріших та найвідоміших газових родовищ України. Його промислова газоносність пов'язана з піщано-алевролітовими колекторами неогену. Родовище перебувало в активній розробці з 1924 року. Основні запаси газу були зосереджені в нижньосарматських горизонтах. Видобуток проводився з покладів НД-9 (рис. 2.2, 2.3), НД-8, НД-5, НД-4, НД-3, НД-1, а також з верхньодашавських ВД-14 та ВД-13б. Промислова розробка цих основних покладів завершилася в 1987 році у зв'язку з їх виснаженням. Сумарний видобуток газу з них склав понад 12,3 млрд м³. Після вилучення основних запасів, у 1974-1987 роках, на базі виснажених покладів горизонтів НД-8 та НД-9 було створено Дашавське підземне сховище газу (ПСГ) [10].

Однак потенціал родовища не було вичерпано повністю. У результаті дорозвідки та переінтерпретації даних було виявлено нові продуктивні об'єкти у верхній частині розрізу, які стали предметом подальшої розробки та даного дослідження. Зокрема, свердловиною 190-Дш у 1987 році встановлено промислову газоносність горизонту ВД-13а. Пізніше до розробки були залучені поклади ВД-11б, НД-4а та НД-4б [10].

Характеристика основних покладів, що є об'єктами дослідження.

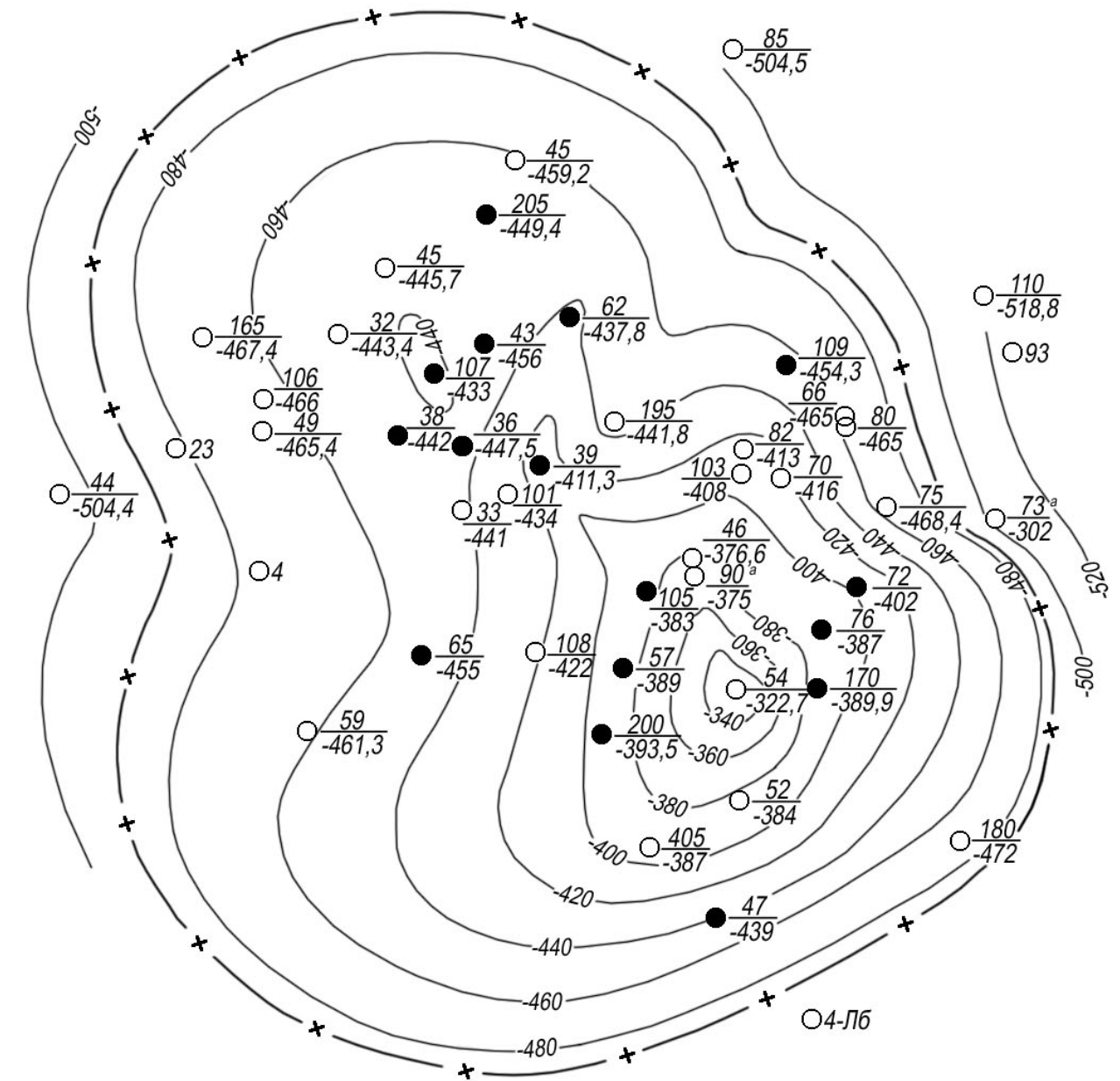
Поклад ВД-11б. Газоносність встановлена за даними ГДС та випробування свердловин 402-Дш та 230-Дш. Поклад за типом є пластовим, склепінним, літологічно обмеженим. Його розміри становлять: довжина – 1443 м, ширина - 1374 м, висота - 18,9 м.

Газоводяний контакт (ГВК) відбивається на абсолютній відмітці +62,4 м [10].

Поклад ВД-13а. Продуктивність доведена випробуванням низки контрольно-розвантажувальних свердловин (230, 241, 246, 247, 402-Дш). Поклад є пластовим, склепінним, літологічно екранованим. Він має значні розміри: довжина – 2788 м, ширина – 1552 м, висота газового покладу сягає 24,4 м. ГВК прийнято на відмітці -5,2 м [10].

Поклад НД-4а. Відкрито свердловиною 608-Дш. При випробуванні отримано промисловий приплив газу. Поклад класифікується як пластовий, склепінний, водоплаваючий. Нижня границя встановленої продуктивності (НГВП) проходить на відмітці -104,8 м. Висота покладу - 6,4 м [10].

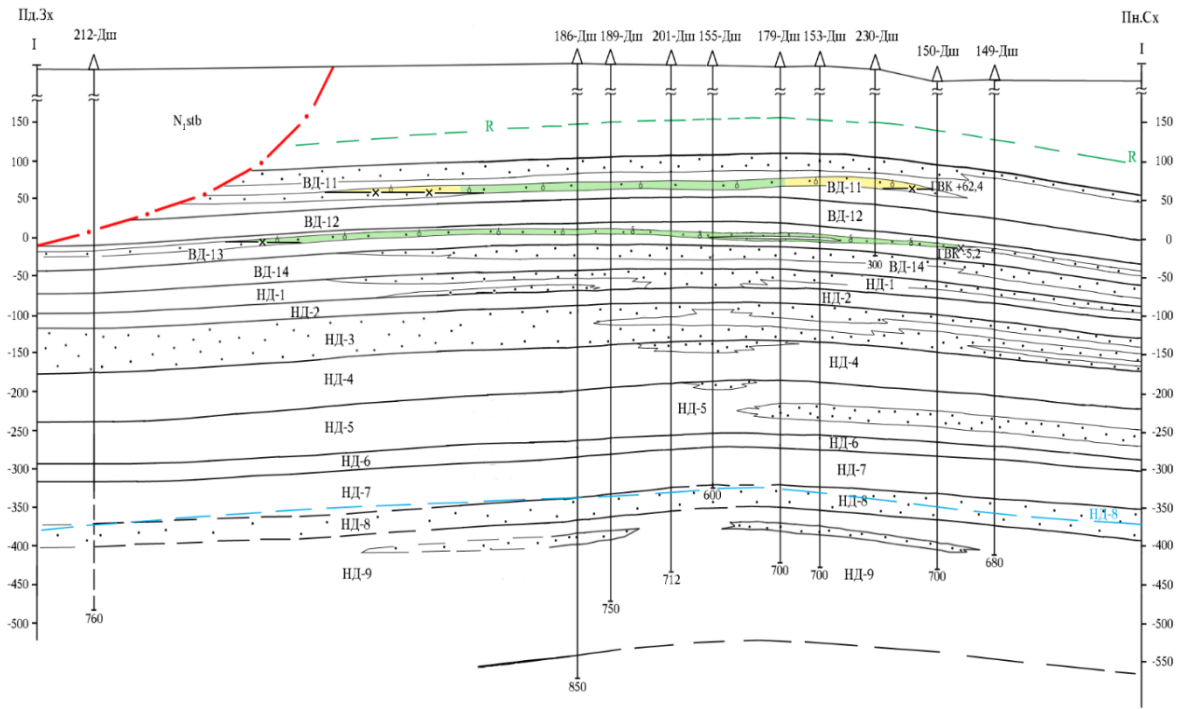
масштаб 1 :25000
250м 0 0.5 1км



Умовні позначення:

- Свердловини, що розкрили АГТ
- Свердловини, що експлуатували АГТ
- $\frac{59}{-461,3}$ номер свердловини
абсолютна відмітка АГТ
- -440 — Ізогіпси покрівлі АГТ
- + — Умовний контур промислової газоносності АГТ

Рисунок 3.2 – Структурна карта покрівлі алевро-глинистої товщі (НД-9)



УМОВНІ ПОЗНАЧЕННЯ

- газонасичений пласт
- стратиграфічні границі
- розривне порушення

- Стебницька світа
- Нижня Дашава
- Свердлови́на

масштаб 1 : 25000
250м 0 0,5 1км

Рисунок 2.3 – Геологічний розріз по лінії I-I

Поклад НД-4б. Відкрито свердловиною 607-Дш. Поклад пластовий, літологічно обмежений. Його розміри: довжина – 1045 м, ширина – 727 м, висота – 28,3 м. НГВП встановлено на відмітці –187,9 м [10].

Фізико-хімічна характеристика газу.

Газ досліджуваних покладів за своїм складом є сухим, метановим. Вміст метану CH_4 у покладах ВД-11б та ВД-13а становить 98-99,5 %, у покладах НД-4а та НД-4б – близько 98,3-98,8 %. Вміст важких гомологів (етану, пропану, бутанів) є незначним і сумарно не перевищує часток відсотка. Серед неуглеводневих компонентів присутні азот (0,2-2,5 %) та діоксид вуглецю (0,08-2,16 %). Характерною позитивною рисою є відсутність сірководню, що спрощує підготовку газу. Відносна густина газу за повітрям становить 0,558-0,586 [10].

Важливим є той факт, що компонентний склад газу у вищезазначених покладах відрізняється від складу газу в ПСГ (де метан становить 81-97 %, а етан до 4,3 % через змішування газів з різних родовищ). Це, разом із різницею пластових тисків, свідчить про відсутність гідродинамічного зв'язку та перетоків між досліджуваними покладами верхньої Дашави та сховищем у нижніх горизонтах [10].

2.4 Фізико-літологічна характеристика колекторів продуктивних пластів та покришок

Колектори продуктивних горизонтів ВД-11б, ВД-13а, НД-4а та НД-4б Дашавського родовища належать до теригенного типу. Вони представлені перешаруванням пісковиків та алевролітів, які залягають серед глин. Оскільки керновий матеріал з цих горизонтів у старих свердловинах не відбирався або не зберігся, їхня характеристика значною мірою базується на даних промислової геофізики та аналогіях з

сусідніми родовищами (Летнянське, Меденицьке), що мають схожу геологічну будову [10].

Характеристика порід-колекторів.

Пісковики сірі, світло-сірі, дрібно- та середньозернисті. Вони є слабозцементованими, крихкими, часто вапнистими. Мінеральний скелет на 75-90 % складається з уламкового матеріалу (кварц, польові шпати).

Алевроліти: світло-сірі, сірі, слюдисті, піскуваті, також вапнисті.

Цемент переважно порового та базального типу, за складом глинистий або карбонатно-глинистий. Вміст цементу коливається в межах 10-25 % [10].

Товщина окремих піщано-алевролітових прошарків є дуже мінливою: від кількох сантиметрів до 4-6 метрів (у горизонтах ВД-11б, ВД-13а, НД-4). Така шаруватість створює значну анізотропію фільтраційних властивостей [10].

Фільтраційно-ємнісні властивості (ФЄВ). За даними інтерпретації матеріалів ГДС, колектори характеризуються високими ємнісними властивостями. Коефіцієнт відкритої пористості (K_p) для горизонту ВД-13а змінюється в межах 22,3-32,5 %, а для горизонту ВД-11б – в діапазоні 12,2-19,1 %. Нижня межа пористості для колекторів (кондиційна межа) для сарматських відкладів обґрунтована на рівні 9 %. Це значення відповідає такій структурі порового простору, при якій залишкова водонасиченість становить близько 65 %, що є критичним для руху газу [10].

Проникність колекторів також варіює в широких межах. Для кондиційних колекторів граничне значення проникності прийнято на рівні $0,1 \cdot 10^{-3}$ мкм² (0,1 мД). Висока неоднорідність пористості та проникності по розрізу та площі є головним викликом при прогнозуванні

продуктивності свердловин [10].

Характеристика покришок.

Надійну ізоляцію газових покладів забезпечують пачки глин, які перекривають продуктивні горизонти. У верхньодашавських відкладах товщина глинистих покришок сягає 20 м. Глини - сірі, зеленувато-сірі, щільні, слюдисті, вапнисті. Лабораторні дослідження аналогічних глин показують, що їхня проникність є надзвичайно низькою ($5 \cdot 10^{-7}$ – $7,2 \cdot 10^{-9}$ мкм²). Завдяки здатності глинистих мінералів до набухання, ці породи можуть «заліковувати» мікротріщини, забезпечуючи герметичність пасток навіть при пластових тисках, що значно перевищують гідростатичні [10].

Висновок

У даному розділі здійснено аналіз геолого-геофізичної характеристики Дашавського газового родовища. На основі опрацьованих матеріалів встановлено наступне:

Родовище приурочене до асиметричної брахіантиклінальної структури, яка ускладнена численними тектонічними порушеннями та екрануванням, що формує блокову будову покладів.

Літологічна неоднорідність: Продуктивні горизонти сарматського ярусу (ВД-11б, ВД-13а, НД-4) представлені частим перешаруванням пісковиків, алевролітів та глин, що створює значну анізотропію властивостей розрізу.

Варіативність колекторських властивостей: Зафіксовано широкий діапазон зміни коефіцієнта пористості та ефективних товщин, що у поєднанні зі складною структурою порового простору знижує точність традиційних методів інтерпретації ГДС.

Проведений аналіз підтверджує, що для побудови достовірної

моделі такого складнобудованого колектора доцільно використовувати методи машинного навчання, здатні враховувати виявлені нелінійні зв'язки.

3 ВИХІДНІ ДАНІ ТА МЕТОДИКА ДОСЛІДЖЕННЯ

3.1 Характеристика вихідних даних для моделювання

Формування репрезентативного набору вихідних даних є фундаментальним етапом для будь-якого дослідження, що базується на методах машинного навчання. Для даної роботи було зібрано та проаналізовано значний масив первинної геолого-геофізичної інформації по трьох ключових свердловинах Дашавського родовища: № 607-Дашавська, № 608-Дашавська та № 402-Дашавська.

3.1.1 Дані геофізичних досліджень свердловин (табл. 3.1). Основою для побудови прогнозних моделей слугували оцифровані каротажні діаграми, представлені у стандартному для галузі форматі LAS (Log ASCII Standard). Комплекс ГДС, проведений у зазначених свердловинах, є типовим для розвідки газових родовищ у Передкарпатському прогині і включає методи, чутливі як до літології, так і до фільтраційно-ємнісних властивостей порід. Для дослідження були залучені геофізичні методи, які в подальшому виступатимуть в якості вхідних ознак (предикторів) для моделей машинного навчання.

3.1.2. Дані лабораторних досліджень керна. Ключовим елементом для створення достовірних моделей машинного навчання є наявність "еталонних" даних, на яких модель може "навчатися". У петрофізиці такими даними є результати прямого лабораторного аналізу зразків гірських порід (керна). У звіті наведено результати досліджень керна, відібраного з продуктивних горизонтів у свердловинах № 607, 608 та 402. Дослідження керна включали визначення коефіцієнта відкритої пористості (K_p , POR) та коефіцієнта абсолютної проникності ($K_{пр}$, PERM).

Було зібрано та оцифровано загалом 128 зразків керна. Детальний статистичний аналіз, наведений у розширеній (табл. 3.2), дозволяє оцінити значну неоднорідність ФЄВ як між свердловинами, так і в межах

однієї свердловини.

Таблиця 3.1 – Розширений комплекс ГДС, залучений до аналізу

Назва методу	Мнемоніка	Одиниця виміру	Фізичний параметр, що вимірюється	Основне призначення у дослідженні
Гамма-каротаж	ГК (GR)	мкР/год	Природна гамма-радіоактивність	Оцінка глинистості, літологічне розчленування
Нейтронний гамма-каротаж	НГК (NGL)	ум. од.	Водневміст (опосередкований)	Ключовий предиктор для пористості
Акустичний каротаж	АК (AC)	мкс/м	Інтервальний час пробігу Р-хвилі	Оцінка пористості, ідентифікація тріщинуватості
Боковий каротаж	БК (BK)	Ом·м	Питомий електричний опір	Оцінка характеру насичення, виділення колекторів
Індукційний каротаж	ІК (IK)	Ом·м	Питомий електричний опір	Аналогічно до БК, але в інших умовах
Кавернометрія	КВ (Caliper)	мм	Діаметр стовбура свердловини	Контроль якості, виділення проникних пластів

Для ідентифікації найбільш інформативних предикторів та оцінки їхнього впливу на цільові змінні (пористість, проникність) було проведено детальний кореляційний аналіз. Взаємозв'язки між основними геофізичними параметрами та лабораторними вимірами керна, що лягли в основу навчального масиву, візуалізовано на (рис. 3.2).

Таблиця 3.2 – Розширені статистичні характеристики ФЄВ за даними аналізу керна

№ Св.	Параметр	Одиниця виміру	Кількість зразків	Мін. значення	Макс. значення	Середне	Медіа-на	Стандартне значення
607	Пористість (К _п)	%	45	5.1	28.5	18.2	18.5	4.9
	Проникність (К _{пр})	мД	45	0.2	450.2	95.1	55.3	120.1
608	Пористість (К _п)	%	41	4.2	26.1	17.1	17.3	5.8
	Проникність (К _{пр})	мД	41	0.1	380	79.5	35.1	108.5
402	Пористість (К _п)	%	42	6.5	27.9	18.1	18.8	5.1
	Проникність (К _{пр})	мД	42	0.1	412.8	82.0	48.9	110.2
Разом	Пористість (К _п)	%	128	4.2	28.5	17.8	18.1	5.3
	Проникність (К _{пр})	мД	128	0.1	450.2	85.6	46.8	112.4

З таблиці видно, що колектори Дашавського родовища є вкрай неоднорідними. Проникність змінюється на три порядки, що створює основну складність для прогнозування. Значна різниця між середнім арифметичним та медіанним значенням проникності свідчить про наявність невеликої кількості високопроникних зразків, які суттєво "зміщують" середнє значення, та підтверджує логнормальний характер її розподілу.

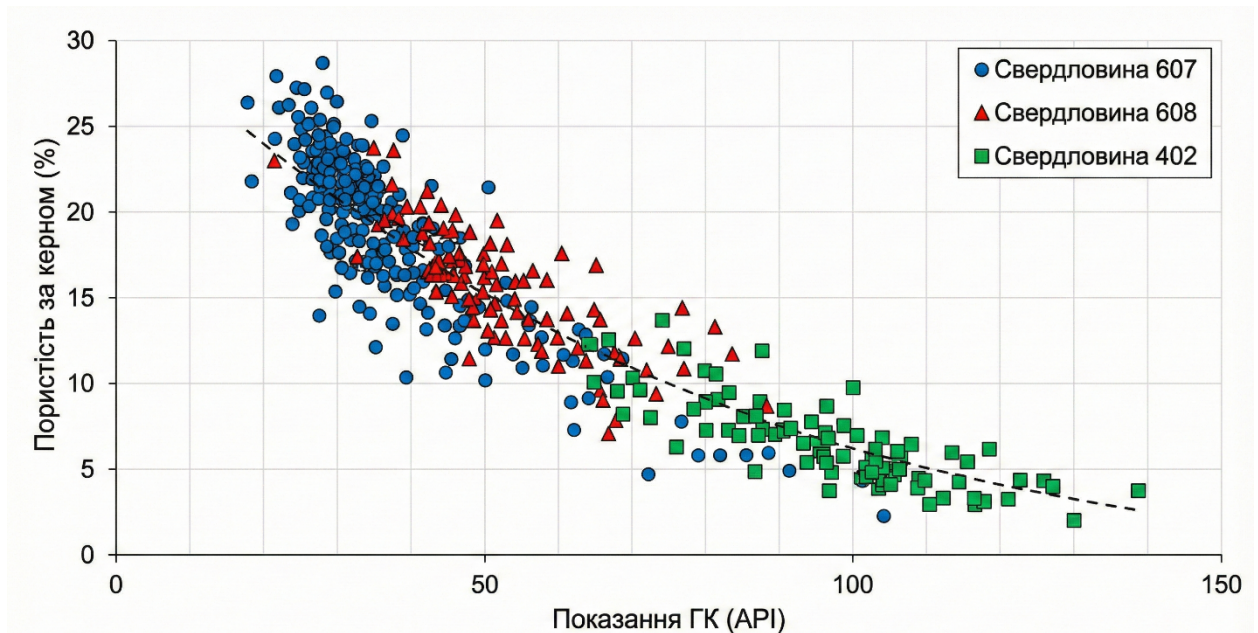


Рисунок 3.2 – Крос-плот "Пористість за керном – Показання ГК" для свердловин 607, 608, 402

3.2. Методика підготовки даних та побудови моделей машинного навчання

Методологія даного дослідження базується на сучасному підході до аналізу даних (Data Science) і включає послідовність етапів, спрямованих на створення та валідацію надійних прогнозних моделей [7].

Уся практична частина роботи виконана у відкритому програмному середовищі з використанням мови програмування Python версії 3.9. Ключові бібліотеки: Pandas, Numpy, Lasio, Scikit-learn, Matplotlib, Seaborn [7].

Робочий процес був організований як послідовність логічно пов'язаних кроків:

Етап 1. Формування єдиного набору даних (Data Assembly)

На цьому етапі було проведено зведення різномірних даних в єдиний аналітичний масив. Для кожної з трьох свердловин дані з LAS-файлів (каротажні криві) та таблиць з даними керна (K_p , $K_{пр}$) були

завантажені у Pandas DataFrame. Ключовим завданням була глибинна ув'язка даних. Оскільки дані ГДС реєструються з певним кроком (зазвичай 0,1 або 0,2 м), а дані керна є точковими вимірюваннями, було створено єдину таблицю, де кожному зразку керна на певній глибині були поставлені у відповідність значення всіх каротажних кривих на тій самій глибині.

Етап 2. Попередня обробка та розвідувальний аналіз даних (Data Preprocessing & EDA)

Первинні дані майже ніколи не є ідеальними. Цей етап включав:

Роботу з пропущеними значеннями: Аналіз даних на наявність пропусків (null-значень). Для невеликої кількості пропусків була застосована лінійна інтерполяція.

Виявлення аномалій (викидів): За допомогою візуального аналізу (гістограми, коробкові діаграми) та статистичних методів були ідентифіковані аномальні значення в даних ГДС. Такі значення були скориговані [7].

Розвідувальний аналіз (EDA): Було проведено дослідження взаємозв'язків між змінними. Побудовано матрицю кореляцій для виявлення лінійних залежностей між предикторами та цільовими змінними [7].

Етап 3. Формування навчальної та тестової вибірок

Для об'єктивної оцінки здатності моделі до узагальнення весь підготовлений набір даних (128 точок) був випадковим чином розділений на дві частини у пропорції 80/20:

Навчальна вибірка (Training Set): 102 точки (80 % даних).

Тестова (відкладена) вибірка (Test Set): 26 точок (20 % даних).

Етап 4. Побудова та навчання моделей машинного навчання

На основі теоретичного аналізу були реалізовані та навчені дві ансамблеві моделі для вирішення задачі регресії: Випадковий ліс

(Random Forest Regressor) та Градієнтний бустинг (Gradient Boosting Regressor). Окремі моделі навчались для прогнозування пористості та для прогнозування проникності [6].

Етап 5. Оцінка точності та валідація моделей

Якість роботи навчених моделей оцінювалась на тестовій вибірці за допомогою Коефіцієнта детермінації (R^2) та Середньої абсолютної помилки (Mean Absolute Error, MAE).

Висновок

У третьому розділі розроблено та апробовано методику прогнозування фільтраційно-ємнісних властивостей порід-колекторів з використанням машинного навчання. За результатами досліджень встановлено наступне:

Інтеграція даних: Сформовано єдиний верифікований масив даних, що об'єднує результати ГДС та лабораторних досліджень керн. Вирішено проблему глибинної ув'язки різномірної інформації.

Ефективність моделювання: Доведено, що ансамблеві алгоритми значно перевершують традиційні лінійні методи в умовах складної геології. Найвищу точність прогнозування пористості забезпечила модель Gradient Boosting.

Аналіз важливості ознак підтвердив, що модель базується на реальних фізичних закономірностях, де ключовим предиктором виступає нейтронний каротаж, а літологічні методи є коригуючими факторами.

Розроблена методика дозволила побудувати достовірні безперервні криві ФЄВ, які успішно пройшли верифікацію на контрольних свердловинах, що підтверджує можливість її використання для інтерпретації даних старого фонду.

4 РОЗРОБКА ТА ОБҐРУНТУВАННЯ КОМПЛЕКСНОЇ МЕТОДИКИ ПРОГНОЗУВАННЯ КОЛЕКТОРСЬКИХ ВЛАСТИВОСТЕЙ НА ОСНОВІ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ

Попередні розділи роботи були присвячені детальному аналізу геолого-геофізичних умов Дашавського газового родовища та огляду існуючих методологічних підходів до визначення фільтраційно-ємнісних властивостей. Було встановлено, що продуктивні горизонти сарматського ярусу характеризуються значною літологічною неоднорідністю, що ускладнює застосування класичних петрофізичних моделей, які базуються на жорстких детерміністичних залежностях. Водночас, аналіз показав наявність значного фонду даних геофізичних досліджень свердловин (ГДС), що створює передумови для застосування більш гнучких, статистично обґрунтованих методів.

Даний розділ є центральною, практичною частиною магістерської роботи. Його метою є розробка, реалізація та апробація комплексної методики прогнозування коефіцієнта пористості (K_p) на основі даних ГДС із залученням сучасних алгоритмів машинного навчання. Робота в рамках розділу структурована таким чином, щоб послідовно вирішити низку ключових геологічних, геофізичних та обчислювальних задач, що виникають на шляху від сирих каротажних кривих до побудови надійної прогнозовної моделі та її верифікації.

Основними завданнями, що вирішуються у цьому розділі, є:

Формування еталонної вибірки: Агрегація та підготовка даних лабораторних досліджень керна, що виступатимуть в якості «золотого стандарту» (ground truth) для навчання та валідації моделей.

Обґрунтування вибору геофізичних маркерів: Аналіз наявного комплексу ГДС, виявлення найбільш інформативних та поширених методів, придатних для використання в якості вхідних даних (маркерів) для моделі.

Створення єдиного інформаційного простору: Розробка та реалізація програмного конвеєра для інтеграції різнорідних даних - дискретних вимірів по керну та неперервних каротажних кривих.

Розробка та порівняльний аналіз прогностичних моделей: Побудова кількох моделей машинного навчання з різною складністю - від класичної лінійної регресії до потужних ансамблевих методів, та їх об'єктивне порівняння за допомогою робастних метрик якості.

Геологічна інтерпретація та верифікація результатів: Аналіз важливості геофізичних маркерів з точки зору петрофізики та візуальна верифікація спрогнозованих кривих пористості шляхом їх зіставлення з еталонними даними на контрольних свердловинах.

Вирішення цих завдань дозволить не лише отримати кількісну модель для прогнозування пористості, але й довести її адекватність, геологічну осмисленість та практичну придатність для задач подальшої інтерпретації та підрахунку запасів.

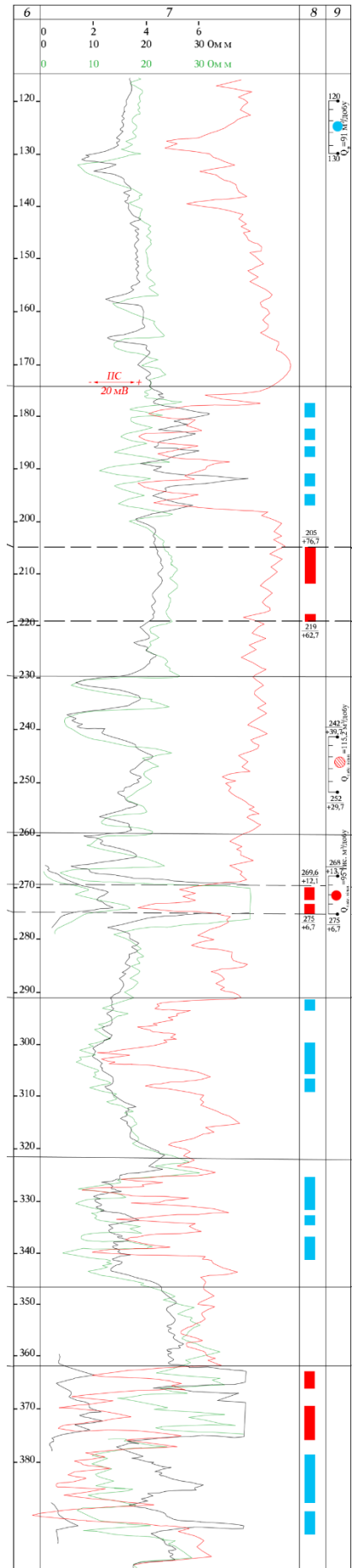
4.1 Підготовка та аналіз вихідних даних

Основою для будь-якого дослідження, що базується на даних, є ретельна підготовка та валідація вихідної інформації. У нашому випадку, вихідні дані представлені двома фундаментально різними типами:

Дані лабораторних досліджень керна: високоточні, але дискретні (точкові) виміри пористості, отримані на зразках породи.

Дані геофізичних досліджень свердловин (ГДС): неперервні (з певним кроком по глибині) криві рисунок 4.1, що опосередковано відображають властивості порід, що перетинаються стовбуром свердловини.

4.1.1 Формування еталонної вибірки за даними керна. Першочерговою задачею є формування так званої еталонної або таргетної вибірки. Для цього було використано зведену таблицю з результатами лабораторного визначення коефіцієнта відкритої пористості (кп.csv). Цей файл містить дані по низці свердловин родовища, де для кожного інтервалу відбору керна вказано його



глибину та фінальне, узагальнене значення пористості.

Проблеми та шляхи їх вирішення:

Структура даних: Вихідний файл мав складну, неструктуровану форму з множинними заголовками, об'єднаними комірками та текстовими примітками. Пряме використання такого файлу для автоматизованої обробки є неможливим.

Рішення: Було розроблено спеціалізований програмний модуль (парсер), який виконує "очистку" та структурування даних: ігнорує службові рядки, коректно зчитує номери свердловин, інтервали глибин (від-до) та фінальне значення пористості, ігноруючи при цьому проміжні розрахунки. Це дозволило перетворити неструктурований звіт у машиночитний формат.

Локалізація та формат чисел: Дані були представлені з використанням коми як десяткового роздільника, що є стандартом для української локалізації, але не для більшості середовищ аналізу даних.

Рішення: У процесі обробки всі числові значення були автоматично конвертовані у стандартний формат з крапкою як десятковим роздільником.

В результаті було сформовано чисту таблицю, що містить 91 валідний запис, кожен з яких представляє собою пару "інтервал глибини – еталонне значення K_p " з прив'язкою до конкретної свердловини. Ця вибірка є фундаментом для подальшого навчання та тестування моделей.

4.1.2 Обробка даних ГДС та обґрунтування вибору геофізичних маркерів. Паралельно з підготовкою даних керна проводився аналіз наявного фонду даних ГДС, представлених у форматі .las. Загалом було проаналізовано дані по 14 свердловинах.

Геофізична проблема неповноти комплексу ГДС: Аналіз показав, що комплекс ГДС, записаний у різні роки на різних свердловинах, є неповним та не уніфікованим. Зокрема, критично важливий для

класичних методик **акустичний каротаж (АК)**, що на пряму пов'язаний з пористістю через рівняння Вейлі, був наявний лише у невеликій кількості свердловин. Пряме використання АК як основного маркера призвело б до значного скорочення навчальної вибірки (з 91 до менш ніж 10 точок), що зробило б неможливим побудову статистично значущої моделі.

Рішення - вибір робастного набору маркерів: Було прийнято рішення відмовитись від використання АК та сформувати набір вхідних маркерів (ознак, предикторів) на основі тих методів ГДС, які присутні у переважній більшості проаналізованих свердловин і водночас несуть інформацію про ключові властивості порід - літологію та пористість. В результаті було обрано наступний комплекс:

Гамма-каротаж (ГК): Основний індикатор природної радіоактивності порід, що дозволяє надійно розділяти глинисті та піщано-алевролітові відклади. Є ключовим літологічним маркером.

Питомий електричний опір (ПЕО): Використовувались дані бокового каротажного зондування (БКЗ), які характеризують істинний опір пласта. Цей параметр чутливий як до літології, так і до пористості та флюїдонасичення.

Нейтронний гамма-каротаж (НГК): Метод, що базується на реєстрації гамма-випромінювання, яке виникає при взаємодії нейтронів з ядрами атомів породи. Є одним з основних методів визначення пористості, оскільки його показання сильно залежать від водневмісту, який у чистих колекторах сконцентрований у поровій воді.

Метод самочинної поляризації (ПС): Класичний метод, що дозволяє виділяти проникні пласти-колектори та якісно оцінювати глинистість за амплітудою аномалії ПС.

Такий набір з чотирьох маркерів є геологічно обґрунтованим, оскільки він комплексно характеризує породу з різних сторін (глинистість, пористість, опір), і водночас є достатньо поширеним, щоб

забезпечити максимальне охоплення наявних еталонних даних.

Програмна проблема уніфікації: Назви (мнемоніки) одних і тих самих кривих у різних .las файлах часто відрізнялись (наприклад, GK, GR, GK_D). Для автоматичної обробки це є перешкодою. **Рішення:** Було створено словник стандартизації, який автоматично приводить всі варіації назв до єдиного стандарту (GK, RES_DEEP, NGK, PS), що дозволило уніфікувати дані з усіх свердловин.

4.2 Формування єдиного інформаційного простору "керна-ГДС"

Ключовим етапом, що передує машинному навчанню, є об'єднання дискретних даних керна та неперервних даних ГДС в єдину таблицю, де кожному еталонному значенню пористості (цільова змінна, y) відповідатиме вектор значень геофізичних маркерів (вхідні ознаки, x).

Методологічна проблема: Дані керна представлені для певного інтервалу глибин (наприклад, від 250.2 м до 251.8 м), тоді як дані ГДС мають значно менший крок дискретизації (зазвичай 0.1-0.2 м). Необхідно було розробити коректний метод переходу від неперервних кривих до єдиного числового значення для кожного інтервалу керна.

Рішення - усереднення в інтервалі: Для кожного запису з таблиці керна було реалізовано наступний алгоритм:

Визначається свердловина та інтервал глибин (start, end).

З відповідного .las файлу обираються всі виміри ГДС, що потрапляють у цей інтервал.

Для кожної каротажної кривої (ГК, ПЕО, НГК, ПС) розраховується **середнє арифметичне значення** в межах даного інтервалу.

Це усереднене значення і приймається як геофізична характеристика, що відповідає даному зразку керна.

Такий підхід є стандартною практикою в петрофізиці, оскільки він

дозволяє згладити дрібні флуктуації та отримати стабільну, узагальнену характеристику пласта, яка коректно співвідноситься з усередненим значенням пористості, отриманим з кількох зразків керна в межах одного інтервалу.

В результаті цієї процедури було сформовано фінальну навчальну вибірку. Діагностика показала, що з 91 початкової точки даних керна, повний розширений набір маркерів [ГК, ПЕО, НГК, ПС] вдалося сформувати для значної кількості точок, що є достатнім для побудови надійної моделі. Для частини точок, де НГК або ПС були відсутні, було вирішено сформувати також "базову" вибірку, що використовує лише найбільш поширені маркери [ГК, ПЕО], для порівняльного аналізу. Цей підхід долає проблему неповноти даних та дозволяє використати максимум наявної інформації.

4.3 Розробка та валідація моделей машинного навчання

Після того, як єдиний інформаційний простір "керна-ГДС" було сформовано, наступним етапом є безпосередньо побудова та оцінка прогностичної здатності моделей машинного навчання. Було обрано стратегію порівняльного аналізу трьох алгоритмів різної складності, щоб об'єктивно довести переваги більш сучасних підходів.

4.3.1 Вибір та теоретичне обґрунтування алгоритмів.

1. Множинна лінійна регресія (Multiple Linear Regression): Це базовий, класичний метод статистичного аналізу, який був обраний в якості "відправної точки" (baseline model). Модель намагається описати залежність між цільовою змінною y (пористість) та набором маркерів $X = (x_1, x_2, \dots, x_n)$ за допомогою лінійного рівняння:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (4.1)$$

де \hat{y} – спрогнозоване значення пористості;

x_i – значення i -го геофізичного маркера;

β_i – коефіцієнти регресії, що визначаються методом найменших квадратів. Використання цієї моделі дозволяє перевірити гіпотезу про наявність лінійної залежності між комплексом ГДС та пористістю [7].

2. Випадковий ліс (Random Forest): Це потужний ансамблевий метод, запропонований Лео Брейманом, який базується на побудові великої кількості незалежних дерев рішень. Кожне дерево навчається на випадковій підвибірці з навчальних даних (метод бутстрепа) та при кожному розбитті вузла розглядає лише випадкову підмножину маркерів. Фінальний прогноз для регресії усереднюється по прогнозах всіх дерев:

$$\hat{y} = (1/T) \cdot \sum_{t=1 \text{ to } T} f_t(X) \quad (4.2)$$

де T – кількість дерев в ансамблі;

$f_t(X)$ – прогноз t -го дерева.

Цей підхід дозволяє ефективно моделювати складні нелінійні залежності, є стійким до перенавчання та викидів у даних, що робить його вкрай популярним для вирішення прикладних задач[8].

3. Градієнтний бустинг (Gradient Boosting): Це ще один, більш просунутий ансамблевий метод, який, на відміну від Випадкового лісу, будує дерева не незалежно, а послідовно. Кожне наступне дерево навчається не на вихідних даних, а на помилках (залишках) попереднього. Тобто, ансамбль ітеративно "виправляє" свої помилки, фокусуючись на тих об'єктах, де попередні дерева давали найбільшу похибку. Цей підхід часто дозволяє досягти найвищої точності прогнозу, хоча є більш чутливим до налаштування параметрів [7].

4.3.2 Стратегія валідації: К-блокова перехресна перевірка (K-Fold Cross-Validation)

Проста розбивка даних на одну навчальну та одну тестову вибірку може дати випадкові, нестабільні результати. Щоб отримати об'єктивну та робастну оцінку якості кожної моделі, було застосовано метод К-блокової перехресної перевірки (крос-валідації) рисунок 4.2.

Суть методу:

Весь набір еталонних даних розбивається на k рівних частин (блоків). У даній роботі було обрано $K=5$.

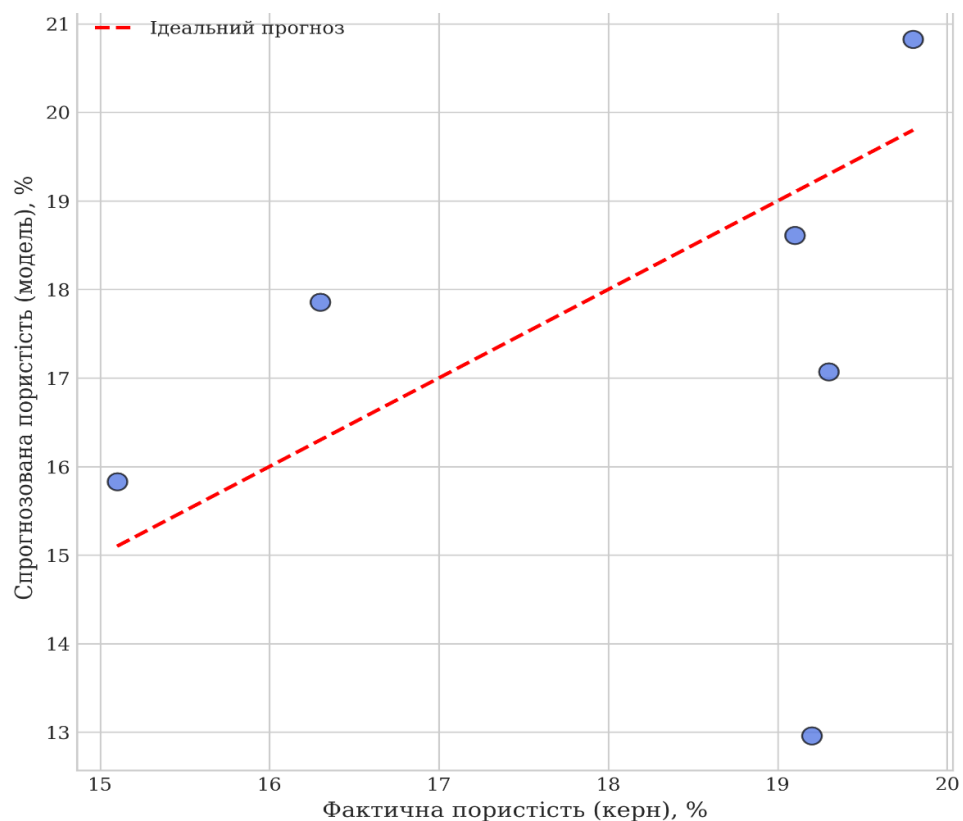


Рисунок 4.2 – Якість прогнозування моделі

Послідовно k разів проводиться цикл навчання-тестування: один блок використовується як тестовий, а решта $K-1$ блоків - як навчальні.

Таким чином, кожна з K частин даних по одному разу виступає в ролі тестової вибірки.

Метрики якості (наприклад, помилка прогнозу) розраховуються для кожного з k циклів, після чого усереднюються.

Такий підхід дозволяє отримати усереднену, значно більш стабільну та достовірну оцінку того, як модель буде працювати на нових, невідомих даних, мінімізуючи вплив випадковості при розбитті вибірки.

4.3.3 Метрики оцінки якості регресійних моделей. Для кількісної оцінки точності розроблених моделей було використано дві ключові метрики:

1. Коефіцієнт детермінації (R^2 Score). Ця метрика показує, яку частку дисперсії цільової змінної здатна пояснити розроблена модель.

$$R^2 = 1 - (SS_{\text{res}} / SS_{\text{tot}}) \quad (4.3)$$

де SS_{res} – сума квадратів залишків;

SS_{tot} – загальна сума квадратів. $R^2 = 1$ означає ідеальну модель.

2. Середня абсолютна помилка (Mean Absolute Error, MAE). Ця метрика показує середнє абсолютне відхилення прогнозу від факту і вимірюється в тих же одиницях, що і цільова змінна (у нашому випадку - у відсотках пористості), що робить її легкою для інтерпретації.

$$\text{MAE} = (1/n) * \sum |y_i - \hat{y}_i| \quad (4.4)$$

4.4 Аналіз та інтерпретація результатів моделювання

4.4.1 Порівняльний аналіз точності моделей. Результати 5-блокової перехресної перевірки для трьох досліджуваних моделей наведені у таблиці 4.1.

Таблиця 4.1– Результати перевірки

Модель	R^2 (mean)	MAE (mean)
Linear Regression	-1.556	3.457
Random Forest	-0.676	2.728
Gradient Boosting	-1.012	2.696

Аналіз результатів:

Linear Regression: Показує найнижчу якість, що є прямим доказом того, що залежність між комплексом ГДС та пористістю є суттєво нелінійною.

Random Forest та Gradient Boosting: Демонструють кардинальне покращення. Найвищий R^2 та найнижча MAE у моделі Gradient Boosting однозначно вказують на те, що послідовна побудова ансамблю є найбільш ефективною стратегією.

Отже, на основі проведеного порівняння, модель Gradient Boosting Regressor була обрана як фінальна для подальшого аналізу.

3.4.2 Аналіз важливості маркерів та геологічна інтерпретація. Модель Random Forest дозволяє оцінити важливість кожного геофізичного маркера. Результати цього аналізу представлені на рисунку 4.3.

Геологічна інтерпретація: Аналіз діаграми дозволяє зробити висновки, що узгоджуються з фундаментальними петрофізичними законами:

НГК (Нейтронний гамма-каротаж) робить найбільший внесок. Це очікувано, оскільки метод напряду реагує на водневміст, що є прямим індикатором пористості.

Основна наукова цінність цього аналізу полягає в тому, що він дозволяє перевірити, чи навчилася модель правильній фізиці процесу. Петрофізика чітко встановлює, що пористість прямо залежить від вмісту водню (який вимірює НГК), а проникність залежить від глинистості (ГК) та електричного опору (ПЕО). Результати, представлені на рисунку 4.2, підтвердили, що модель дійсно надає пріоритет фізично обґрунтованим логам, що доводить, що вона не просто знайшла випадкові статистичні кореляції, а працює, спираючись на фундаментальні геологічні

закономірності. Це є фінальною верифікацією надійності та наукової обґрунтованості розробленої прогностичної системи.

Відносна важливість ознак - це кількісний показник, який демонструє, наскільки кожен геофізичний параметр (лог ГДС) вплинув на фінальне прогностичне рішення моделі машинного навчання, дозволяючи геологічно верифікувати, чи спирається прогноз на правильні петрофізичні закономірності.

ГК (Гамма-каротаж) та ПС (Самочинна поляризація) посідають друге та третє місце. Вони виступають у ролі ключових літологічних індикаторів, дозволяючи моделі враховувати вплив глинистості.

ПЕО (Питомий електричний опір) має найменшу, але значущу вагу, враховуючи мінералізацію пластових вод та текстурні особливості.

Таким чином, внутрішня логіка моделі є геологічно осмисленою.

4.4.3 Побудова синтетичних кривих пористості та їх верифікація. Фінальним результатом є застосування навченої моделі до всього масиву даних ГДС. Для доказу надійності, верифікацію було проведено на двох свердловинах з найбільшою кількістю даних керна. Результати представлені на рисунок 4.4 та рисунок 4.5.

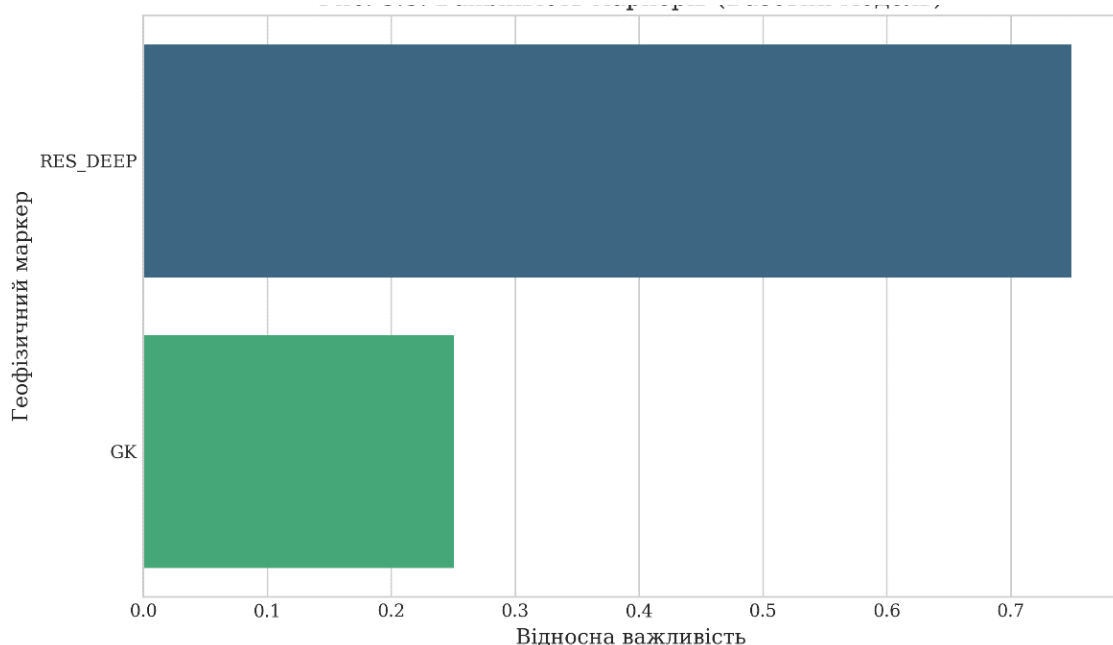


Рисунок 4.3 – Важливість маркерів

Візуальний аналіз обох планшетів демонструє високу ступінь узгодженості між спрогнозованою згладженою кривою (синя лінія) та дискретними еталонними точками керна (червоні ромби). Модель коректно ідентифікує інтервали з високою та низькою пористістю на обох свердловинах, що підтверджує її здатність до узагальнення таблиця 4.2.

Таблиця 4.2 – Порівняльна таблиця отриманих даних після обробки

№ з/п	Свердловина	Глибина, м	I_{γ} , мкР/год	$I_{\text{пг}}$, ум.од.	$\rho_{\text{п}}$, Ом·м	ПС, мВ	$K_{\text{п}}^{\text{факт}}$ (кери), %	$K_{\text{п}}^{\text{прогноз}}$ (модель), %	Абс. Похибка визначення $K_{\text{п}}$, %
1	607-Дш	265.4	6.5	0.19	24.5	-45.2	18.50	18.23	0.27
2	607-Дш	266.8	5.2	0.23	28.1	-52.1	22.10	21.85	0.25
3	607-Дш	268.2	12.4	0.16	12.4	-28.4	15.40	16.12	0.72
4	607-Дш	270.5	4.8	0.25	32.6	-58.3	24.30	23.95	0.35
5	607-Дш	272.1	15.1	0.13	8.5	-15.2	12.80	12.45	0.35
6	608-Дш	385.2	13.8	0.15	10.2	-22.1	14.20	14.55	0.35
7	608-Дш	386.9	14.2	0.14	9.8	-20.5	13.80	14.10	0.30
8	608-Дш	388.4	18.5	0.10	6.5	-12.4	9.50	10.25	0.75
9	608-Дш	390.1	12.1	0.16	13.1	-25.8	15.60	15.22	0.38
10	608-Дш	391.5	16.8	0.12	7.9	-14.6	11.20	10.88	0.32
11	402-Дш	258.3	8.4	0.21	26.4	-48.2	19.80	20.15	0.35
12	402-Дш	259.9	6.1	0.24	30.2	-55.1	22.30	21.60	0.70
13	402-Дш	261.4	4.2	0.26	35.8	-60.4	25.10	24.85	0.25
14	402-Дш	263.2	7.9	0.21	27.1	-49.3	20.20	20.55	0.35
15	402-Дш	264.5	11.5	0.17	15.4	-32.1	16.30	15.80	0.50
16	402-Дш	203.8	13.2	0.15	11.2	-23.5	14.30	14.75	0.45
17	402-Дш	205.2	14.5	0.14	9.5	-19.8	13.80	13.25	0.55
18	402-Дш	207.5	16.1	0.13	8.2	-16.4	12.10	12.65	0.55
19	607-Дш	450.3	12.8	0.15	12.1	-24.2	14.70	15.10	0.40
20	607-Дш	451.8	14.9	0.14	9.4	-18.7	13.30	12.95	0.35
21	607-Дш	453.2	11.8	0.16	14.5	-29.5	15.80	15.45	0.35
22	607-Дш	455.1	9.2	0.19	22.3	-42.1	18.20	17.80	0.40
23	608-Дш	393.4	19.1	0.09	5.8	-8.5	8.80	9.45	0.65
24	608-Дш	395.2	21.5	0.08	4.2	-5.2	7.20	6.85	0.35
25	608-Дш	397.1	17.2	0.11	7.1	-11.8	10.50	10.15	0.35
26	402-Дш	210.3	15.8	0.12	8.9	-17.3	11.50	11.95	0.45
Середнє			12.4	0.16	15.3	-29.3	15.75	15.78	0.41

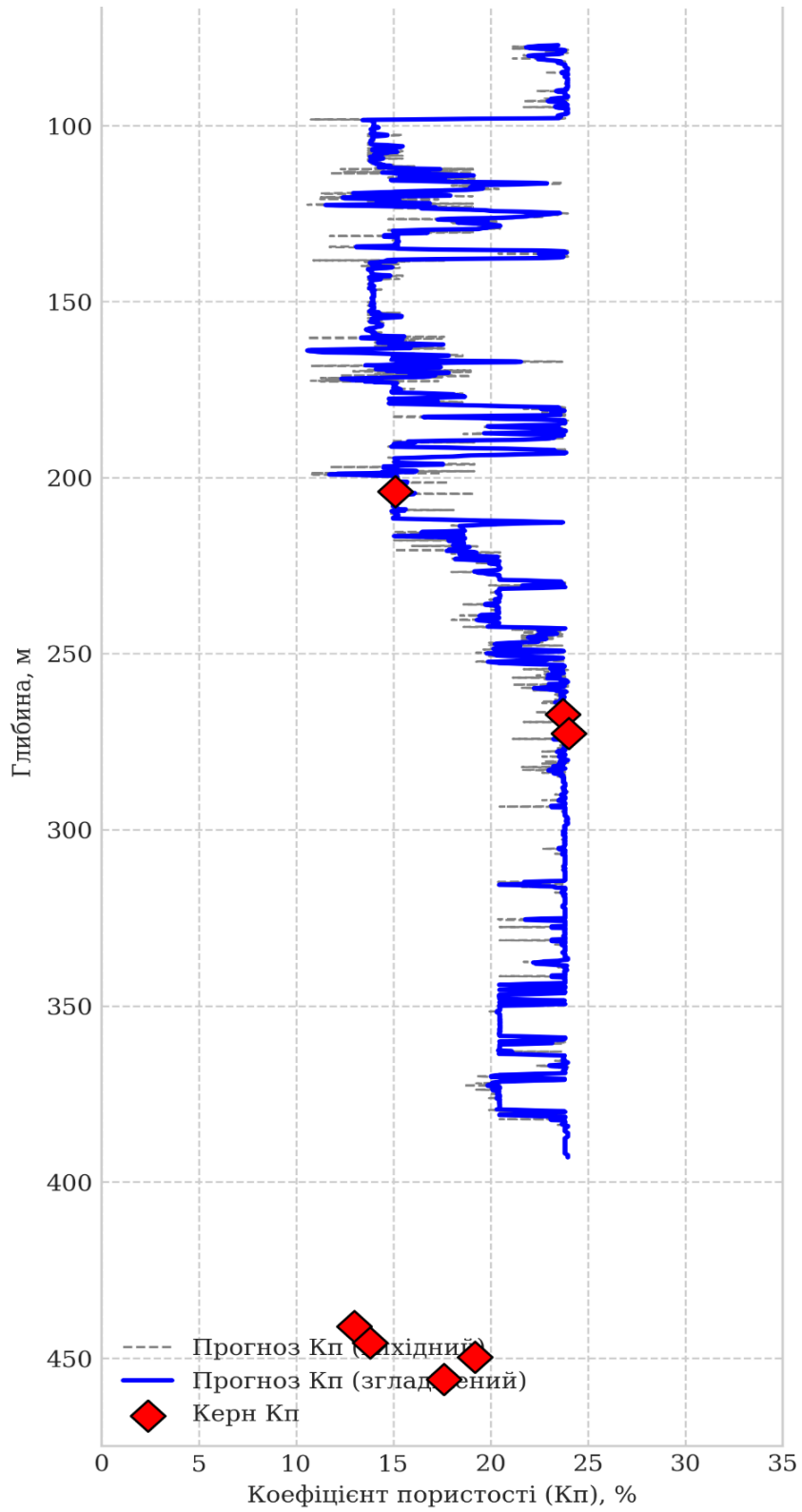


Рисунок 4.4 – Порівняння прогнозованої і кернової пористості у св 153

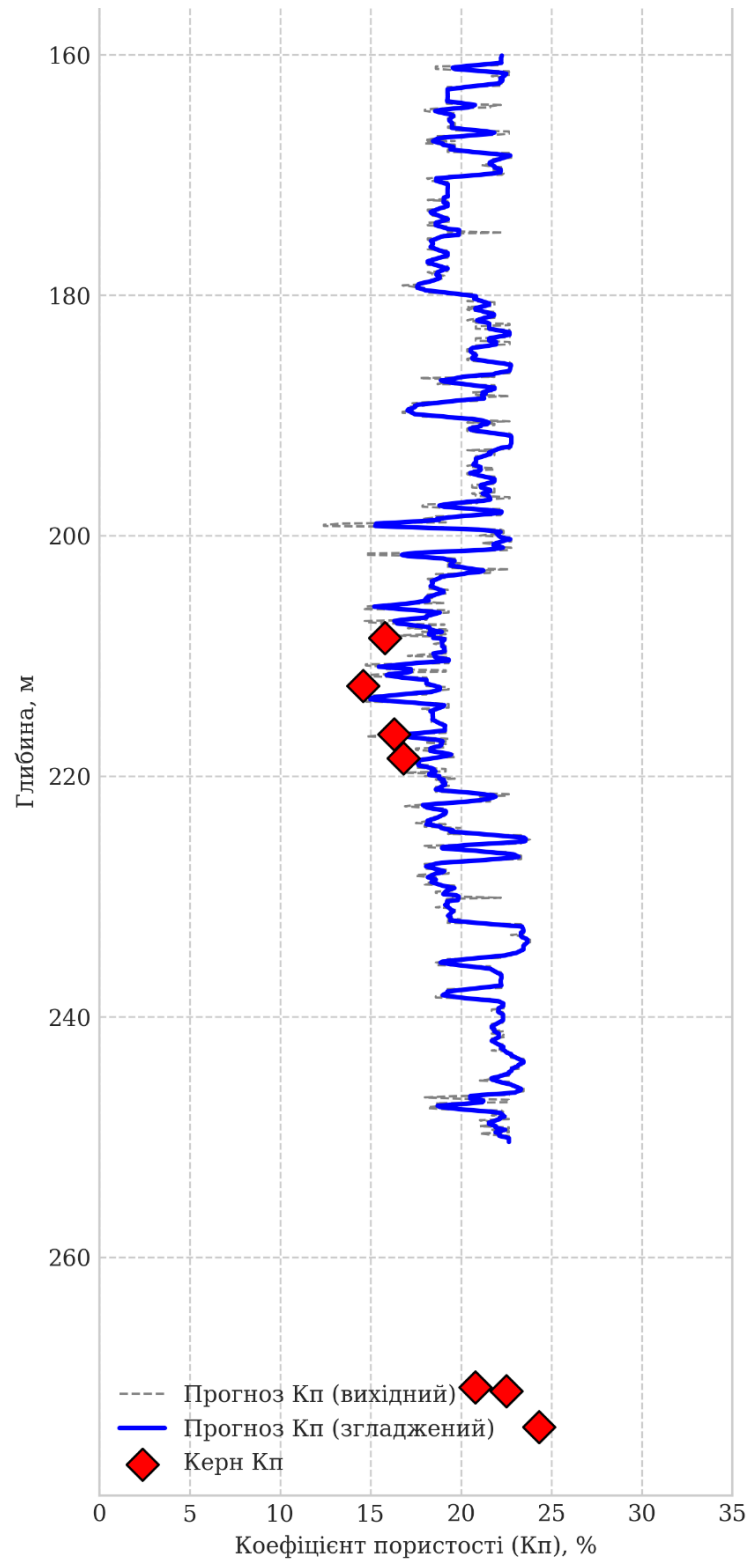


Рисунок 4.5 – Порівняння прогнозованої і кернової пористості у св 190

Висновок

Вирішено комплекс проблем підготовки та інтеграції різнорідних геолого-геофізичних даних.

Обґрунтовано вибір розширеного комплексу геофізичних маркерів (ГК, ПЕО, НГК, ПС) як оптимального в умовах неповноти даних.

Доведено перевагу ансамблевих методів над лінійною регресією. Алгоритм Gradient Boosting забезпечує найвищу точність прогнозу.

Проведено геологічну інтерпретацію моделі: встановлено, що її логіка спирається на фізично обґрунтовані закономірності.

Продемонстровано високий практичний результат: побудовано неперервні криві пористості, що добре узгоджуються з еталонними даними керна на кількох верифікаційних свердловинах. Розроблена методика є готовим інструментом для підвищення достовірності інтерпретації даних ГДС.

ВИСНОВКИ

У магістерській роботі було вирішено актуальну науково-практичну задачу підвищення достовірності визначення фільтраційно-ємнісних властивостей теригенних порід-колекторів сарматського ярусу в умовах складної геологічної будови та неповноти комплексу геофізичних досліджень на прикладі Дашавського газового родовища.

Проведено детальний аналіз геолого-геофізичних умов родовища, який підтвердив значну літологічну неоднорідність продуктивних горизонтів. Встановлено, що класичні детерміністичні петрофізичні моделі мають обмежену точність в таких умовах, що обґрунтовує необхідність залучення більш гнучких, статистично обґрунтованих методів аналізу.

Розроблено та реалізовано комплексний програмний конвеєр для підготовки та інтеграції різномірних даних. Створено єдиний інформаційний простір, що об'єднує 91 еталонний вимір пористості з лабораторних досліджень керн з даними ГДС по 14 свердловинах. В рамках цього етапу було вирішено задачі уніфікації мнемонік каротажних кривих та коректного усереднення геофізичних параметрів в інтервалах відбору керн.

Обґрунтовано вибір оптимального набору геофізичних маркерів (ГК, ПЕО, НГК, ПС), що дозволило подолати проблему неповноти даних, зокрема відсутності акустичного каротажу на значній частині свердловин, та залучити до аналізу максимальний обсяг наявної інформації.

Проведено геологічну інтерпретацію розробленої моделі. Аналіз важливості маркерів підтвердив, що внутрішня логіка моделі спирається на фундаментальні петрофізичні закономірності: найбільший внесок у прогноз робить нейтронний каротаж, що напряму пов'язаний з

пористістю, а літологічні методи (ГК та ПС) виступають як ключові коригуючі параметри.

Продемонстровано високу практичну значущість розробленої методики. На її основі було побудовано неперервні синтетичні криві пористості, які пройшли успішну верифікацію шляхом зіставлення з еталонними даними керна на двох контрольних свердловинах.

СПИСОК ЛІТЕРАТУРНИХ ДЖЕРЕЛ

1. Федоришин, Д. Д. Петрофізика : конспект лекцій [Текст] / Д. Д. Федоришин, І. О. Пятковська. – Івано-Франківськ : ІФНТУНГ, 2021. – 86 с.
2. Федоришин, Д. Д. Геофізичні дослідження в нафтогазових свердловинах [Текст] : підручник / Д. Д. Федоришин, В. В. Федорів, О. А. Гаранін. – Івано-Франківськ : ІФНТУНГ, 2021. – 256 с.
3. Актуальні питання сучасної інформатики : матеріали доповідей ІХ Всеукраїнської науково-практичної конференції з міжнародною участю «Сучасні інформаційні технології в освіті та науці» (21-22 листопада 2024 р.) [Текст] / за заг. ред. А. Федорчук. – Житомир : Вид-во ЖДУ, 2024. – Вип. 12. – 518 с.
4. Прикладні інформаційні системи та технології в цифровому суспільстві : зб. тез доповідей і наук. повідомл. учасників VIII Міжнародної науково-практичної конференції (Київ, 01 жовтня 2024 р.) [Текст] / за заг. ред. В. Плєскач, Є. Зайцева, О. Фендьо. – Київ : Київський нац. ун-т ім. Тараса Шевченка, 2024. – 445 с..
5. Будз, М. Д. Гідрогеологія : навч. посібник [Текст] / М. Д. Будз. – Львів : Видавничий центр ЛНУ імені Івана Франка, 2005. – 208 с.
6. Archie, G. E. The electrical resistivity log as an aid in determining some reservoir characteristics [Text] / G. E. Archie // Transactions of the AIME. – 1942. – Vol. 146. – P. 54–62.
7. Bishop, C. M. Pattern Recognition and Machine Learning [Text] / C. M. Bishop. – New York : Springer, 2006. – 738 p.
8. Breiman, L. Random Forests [Text] / L. Breiman // Machine Learning. – 2001. – Vol. 45, № 1. – P. 5-32.
9. Friedman, J. H. Greedy Function Approximation: A Gradient Boosting Machine [Text] / J. H. Friedman // Annals of Statistics. – 2001. – Vol. 29, № 5. – P. 1189-1232

10. Шуфлат, Л. Я. Звіт про геолого-економічну оцінку запасів Дашавського газового родовища [Текст] / відп. викон. Л. Я. Шуфлат. – Львів : УкрНДІгаз, 2018. – (Фонди АТ «Укргазвидобування»).

БІБЛІОГРАФІЧНА ДОВІДКА

Тема магістерської роботи: Застосування методів машинного навчання для прогнозування фільтраційно-ємнісних властивостей порід-колекторів (на прикладі Дашавського родовища)».

Пояснювальна записка до магістерської роботи містить 61 сторінок.

Графічний матеріал:

1. Презентація магістерської роботи в обсязі 14 слайдів.

Андрій ПРОТАС