

**Київський національний університет імені Тараса Шевченка**  
**Фізичний факультет**  
**Кафедра загальної фізики**

**Звіт**  
**про науково-виробничу практику із фізики наноматеріалів**

**Тема роботи:** «Параметризація рухливості носіїв заряду у кремнії з використанням машинного навчання»

**Студент 2 курсу магістратури**  
освітньої програми «Фізика наносистем»  
кафедри загальної фізики  
Куш Іван

**Керівник практики:**  
завідувач кафедри загальної фізики, д. ф.-м. н.  
Оліх Олег

**Науковий керівник:**  
завідувач кафедри загальної фізики, д. ф.-м. н.  
Оліх Олег

**Київ – 2024**

## **1 План**

науково-виробничої практики з фізики наноматеріалів Куш І.О. (ОКР “магістр” , II курс)

#	Вид роботи	Кабінет	Час виконання
1	Знайомство з методами розрахунку рухливості	432 Кабінет	25.01.2024–17.02.2024
2	Розрахунок рухливості носіїв заряду за допомогою методу Klassen	432 Кабінет	17.02.2024–14.03.2024
3	Параметризація рухливості носіїв заряду за допомогою Випадкового лісу	432 Кабінет	14.03.2024–01.04.2024
4	Параметризація рухливості носіїв заряду за допомогою Символічної регресії	432 Кабінет	01.04.2024–20.04.2024
5	Написання звіту	432 Кабінет	20.04.2024–20.05.2024

Табл. 1: Розклад виконання робіт

## 2 Методики розрахунку рухливості

### 2.1 Наближення Klassen

Метод наближення Klassen є одним із підходів для оцінки рухливості носіїв заряду в напівпровідниках, враховуючи різні механізми розсіювання. Він передбачає паралельне складання впливу розсіювання на ґратці та розсіювання на домішках. Рухливість  $\mu$  визначається через зворотну величину, що є сумою зворотних рухливостей, обумовлених кожним із цих механізмів. Це дозволяє точніше моделювати поведінку електронів та дірок при різних умовах температури і концентрації домішок. Метод використовує численні коефіцієнти, які визначають температурну залежність та взаємодію з концентрацією носіїв заряду. У напівпровіднику визначається через паралельне складання впливу різних механізмів розсіювання. Це виражається через зворотну рухливість:

$$\mu^{-1} = \mu_L^{-1} + \mu_{DA}^{-1}$$

де:  $\mu_L$  — рухливість, обмежена розсіюванням на ґратці (фононах),  $\mu_{DA}$  — рухливість, обмежена розсіюванням на домішках (іонізовані домішки).

Рухливість, обмежена розсіюванням на ґратці:

$$\mu_L = \mu_{\max} \cdot \left( \frac{300}{T} \right)^{2.25}$$

Рухливість, обмежена розсіюванням на домішках:

$$\mu_{DA} = \frac{\mu_{\max}^2}{\mu_{\max} - \mu_{\min}} \cdot \frac{N_{sc}}{N_{eff}} \cdot \left( \frac{N_{eff}}{s_c} \right)^{\alpha} \cdot \left( \frac{T}{300} \right)^{3\alpha-1.5} + \frac{\mu_{\max} \cdot \mu_{\min}}{\mu_{\max} - \mu_{\min}} \cdot \frac{n+p}{N_{eff}} \cdot \left( \frac{300}{T} \right)^{0.5}$$

Для електронів ( $n$ ):

$$\mu_{\max} = 0.1414, \quad \mu_{\min} = 0.00685, \quad N_{eff} = 9.2 \times 10^{22}$$

Для дірок ( $p$ ):

$$\mu_{\max} = 0.04705, \quad \mu_{\min} = 0.000449, \quad N_{eff} = 2.28 \times 10^{23}$$

Для електронів ( $n$ ):

$$N_{sc} = N_d + N_a + p$$

Для дірок ( $p$ ):

$$N_{sc} = N_d + N_a + n$$

Для електронів ( $n$ ):

$$N_{\text{eff}} = N_d + N_a \cdot G_n(P_n, T) + \frac{P}{F_n(P_n, T)}$$

Для дірок ( $p$ ):

$$N_{\text{eff}} = N_a + N_d \cdot G_p(P_p, T) + \frac{n}{F_p(P_p, T)}$$

Концентрація донорів ( $N_d$ ) та акцепторів ( $N_a$ ) — це параметри, що визначають кількість іонізованих атомів у напівпровіднику, які можуть передавати або отримувати електрони:

- $N_d$  — концентрація донорів, тобто атомів або молекул, які здатні віддавати електрони в зону провідності, створюючи електронні носії заряду.
- $N_a$  — концентрація акцепторів, тобто атомів або молекул, які здатні приймати електрони з валентної зони, створюючи дірки (недостаючи електрони) в кристалі.

Ці концентрації суттєво впливають на властивості напівпровідника, оскільки вони визначають кількість носіїв заряду в системі, а також її електричні та оптичні характеристики.

$$G_n = 1 - \left( \frac{S_1}{S_2 + \left( \frac{m_0}{m_e} \cdot \frac{T}{300} \right)^{S_4} \cdot P_n} \right)^{S_3} + \frac{S_5}{\left( \frac{m_e}{m_0} \cdot \frac{300}{T} \right)^{S_7}} \cdot P_p^{S_6}$$

Коефіцієнт	Значення
$S_1$	0.89233
$S_2$	0.41372
$S_3$	0.19778
$S_4$	0.28227
$S_5$	0.005978
$S_6$	1.80618
$S_7$	0.72169

Табл. 2: Значення коефіцієнтів для коефіцієнта іонізації

$$F_n = \frac{r_1 \cdot P_n^{r_6} + r_2 + r_3 \cdot \frac{m_e}{m_0}}{P_n^{r_6} + r_4 + r_5 \cdot \frac{m_e}{m_0}}$$

Коефіцієнт	Значення
$r_1$	0.7643
$r_2$	2.2999
$r_3$	6.5502
$r_4$	2.3670
$r_5$	-0.01552
$r_6$	0.6478

Табл. 3: Значення коефіцієнтів для функції  $F_n$

$$P = \frac{1}{P_b + P_c}$$

Для електронів ( $n$ ):

$$P_b = \frac{3.83}{\left( \frac{1.36 \cdot 10^{26}}{(N_{\text{dop}} + P) \cdot \left( \frac{T}{300} \right)^2} \right) \cdot \frac{m_e}{m_0}}$$

$$P_c = \frac{2.46}{\left( 3.97 \cdot 10^{19} \cdot \left( \frac{T}{300Z} \right)^3 \cdot \frac{1}{N_{\text{dop}}} \right)^{2/3}}$$

Для дірок ( $p$ ):

$$P_b = \frac{3.83}{\left( \frac{1.36 \cdot 10^{26}}{(N_{\text{dop}} + n) \cdot \left( \frac{T}{300} \right)^2} \right) \cdot \frac{m_e}{m_0}}$$

Для електронів ( $Z_n$ ):

$$Z_n = 1 + \frac{1}{0.21 + \left( \frac{4 \cdot 10^{26}}{N_{\text{dop}}} \right)^2}$$

Для дірок ( $Z_p$ ):

$$Z_p = 1 + \frac{1}{0.5 + \left( \frac{7.2 \cdot 10^{26}}{N_{\text{dop}}} \right)^2}$$

[?] [?]

## 2.2 Наближення Aroga

Метод Aroga є альтернативним підходом для оцінки рухливості носіїв заряду в напівпровідниках, який також враховує вплив різних механізмів розсіювання, таких як розсіювання на ґратці та на домішках. Однак, на відміну від методу Klassen, метод Aroga більше фокусується на температурних залежностях, пропонуючи інші математичні вирази для рухливості. В Aroga, основна увага приділяється розрахунку ефективної рухливості, враховуючи фактори, що змінюються в залежності від температури та концентрації носіїв заряду, зокрема акцепторів і донорів. Однією з головних відмінностей між методами є те, що метод Klassen передбачає більш комплексну модель, яка поєднує різні механізми розсіювання в одну формулу з використанням зворотних рухливостей. У той час як метод Aroga може використовувати більш окремі підходи для кожного механізму розсіювання, зокрема, для температурної залежності рухливості електронів та дірок. Крім того, метод Aroga часто включає інші коефіцієнти, які специфічні для конкретних умов і матеріалів, що може давати точніші результати в певних випадках.

$$\mu = \mu_{\min} + \frac{\mu_{\max}}{1 + \left( \frac{N}{N_0} \right)^\alpha}$$

Ця формула виражає рухливість носіїв заряду в залежності від концентрації  $N$ . Коефіцієнт  $\mu_{\min}$  є мінімальною рухливістю, тоді як  $\mu_{\max}$  — максимальною рухливістю. Параметр  $N_0$  визначає критичну концентрацію, після досягнення якої відбувається насичення рухливості.

$$N = N_{a^-} + N_{d^+}$$

$N$  виражає загальну концентрацію носіїв заряду, як суму концентрацій акцепторів  $N_a$  та донорів  $N_d$ .

$$\mu_{\min} = \mu_{\min}^{\text{eff}} \cdot \left( \frac{T}{T_{\text{eff}}} \right)^{\beta_1}$$

$$\mu_0 = \mu_0^{\text{eff}} \cdot \left( \frac{T}{T_{\text{eff}}} \right)^{\beta_2}$$

$$N_0 = N_0^{\text{eff}} \cdot \left( \frac{T}{T_{\text{eff}}} \right)^{\beta_3}$$

$$\alpha = \alpha_0^{\text{eff}} \cdot \left( \frac{T}{T_{\text{eff}}} \right)^{\beta_4}$$

**Значення параметрів:**

$$T_{\text{eff}} = 300 \text{ K}, \quad \alpha_0 = 0.88, \quad \beta_1 = -0.57$$

Для  $n$  (електронів):

$$\beta_2 = -2.33$$

Для  $p$  (дірки):

$$\beta_2 = -2.23$$

$$\beta_3 = 2.4, \quad \beta_4 = -0.146$$

Параметр	n	p
$\mu_{\min} (\text{eff})$	88	54,3
$\mu_0 (\text{eff})$	1252	407
$N_0 (\text{eff})$	$1,26 \times 10^{17}$	$2,35 \times 10^{17}$

Табл. 4: Параметри для основних носіїв

Для розрахунку рухливості не основних зарядів змінюються параметри

Параметр	$\mu_e (\text{Teff} = 292\text{K})$	$\mu_n (\text{Teff} = 295\text{K})$
$\mu_{\min}$	130	232
$\mu_0$	370	1180
$N_0$	$8 \times 10^{17}$	$8 \times 10^{16}$
$\alpha$	1,25	0,9

Табл. 5: Значення параметрів неосновних носіїв заряду для  $\mu_e$  та  $\mu_n$

## 2.3 Наближення Fletcher

Ще один із відомих методів пошуку рухливості основних носіїв заряду є методі Fletcher . Основна перевага в цьому методі полягає в доволі простому обчисленні .

$$\frac{1}{\mu_{n,\text{fl}}} = \frac{1}{\mu_{\text{in},n}} + \frac{1}{\mu_{\text{cc}}}$$
$$\frac{1}{\mu_{p,\text{fl}}} = \frac{1}{\mu_{\text{in},p}} + \frac{1}{\mu_{\text{cc}}}$$
$$\mu_{\text{cc}} = \frac{\left(\frac{T}{T_{\text{eff}}}\right)^{3/2} F_1}{\sqrt{np} \cdot \ln \left(1 + \left(\frac{T}{T_{\text{eff}}}\right)^2 \cdot (np)^{-1/3} F_2\right)}$$

Де:

- $T$  — температура решітки,
- $\mu_{\text{in},n}$  і  $\mu_{\text{in},p}$  — вхідна рухливість електронів і дірок,
- $n$  — концентрація електронів,
- $p$  — концентрація дірок,
- $F_1$  (одиниця SI:  $\text{с}^2\text{А}/(\text{м}^3\text{кг})$ ) і  $F_2$  (одиниця SI:  $1/\text{м}^2$ ) — властивості матеріалу,
- $T_{\text{eff}}$  — характеристична температура матеріалу.

## 3 Параметризація рухливості за допомогою Random Forest

Метод Random Forest — це ансамблевий алгоритм машинного навчання, який будує кілька рішень дерев на випадкових підмножинах даних і комбінує їх для покращення точності та стабільності моделі. Кожне дерево в лісі приймає рішення незалежно, і кінцевий результат визначається голосуванням або середнім значенням передбачень усіх дерев. Цей метод допомагає знижувати ризик перенавчання, підвищує точність прогнозів і може працювати з великою кількістю вхідних змінних.

В коді використовувався метод пошуку найкращих гіперпараметрів Optuna

Optuna — це потужна бібліотека для автоматичного налаштування гіперпараметрів, яка використовує інтелектуальні методи пошуку для оптимізації моделей машинного навчання. ptuna використовує методи послідовного вибору гіперпараметрів, базуючись на попередніх результатах. Це дозволяє скоротити кількість ітерацій у порівнянні з випадковим пошуком.

Код та результати обчислень знаходяться за посиланням. Модель передбачення демонструє загалом хороший рівень точності. На основі графіка, що порівнює передбачені значення з реальними, видно, що більшість точок розташовані близько до лінії ідеального співпадіння, що свідчить про здатність моделі ефективно описувати основні закономірності даних. Однак, на графіку також помітні окремі значні відхилення (виброси).

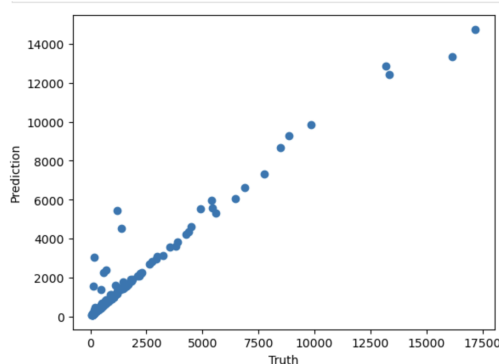


Рис. 1: лінія ідеальної відповідності Випадкового лісу

## 4 Параметризація рухливості за допомогою Symbolic Regression

PySR (Python Symbolic Regression) є потужним інструментом для виконання символічної регресії, яка полягає у пошуку математичних виразів, що найкраще описують вхідні дані. PySR використовує еволюційні алгоритми для генерування та оптимізації формул, що дозволяє знаходити компактні та інтерпретовані математичні моделі[3]. Основний компонент пакету - PySRRegressor, який надає інтерфейс для тренування та оцінки моделей символічної регресії. PySRRegressor автоматично шукає найкращу формулу шляхом ітеративного еволюційного процесу, що включає генерацію можливих виразів, їх оцінку та вибір найкращих серед них на основі заданих критеріїв, таких як точність та складність моделі. Протягом семестру я працював над кількома наборами даних, застосовуючи PySRRegressor для розробки точних та інтерпретованих моделей. Цей процес включав попередню обробку даних, налаштування параметрів PySRRegressor, а також аналіз отриманих результатів для забезпечення найкращої можливої апроксимації даних. Завдяки цьому я здобув глибокі знання в області символічної регресії та її практичного застосування для аналізу даних.

Навчання проходило в пару етапів. Перший етап полягає в предобробці даних. Розбиття на тренувальний набір, для основного тренування моделі, та тестування, на якому відбувається оцінка вже навченої моделі. Потім відбувається створення самої моделі та саме навчання за нею. Весь код та результати з результатами знаходяться за посиланням. Особливої уваги хочеться звернути на створенні оператори та залучення їх, а саме "Zved" та "PowInv". Пакет Pysr дозволяє створювати свої оператори для навчання моделі, для більш швидкого та точного підбору апроксимуючих формул. Ці формули, "Zved" та "PowInv" зумовлені теоретичними даними, що були представлені в звіті. І представляють собою:

$$\text{Zved}(x, y) = \frac{x \cdot y}{x + y} \quad (1)$$

$$\text{PowInv}(x) = \frac{1.0}{1.0 + x} \quad (2)$$

Було виконано декілька спроб навчання з різними параметрами та часом навчання. В ході роботи було виявлено, що для нашого набору даних для основний приріст точності відбувається 30 годинним навчанням, після проходження порогу в 30 годин, приріст точності дуже поступається приросту складності та інтерпретації запропонованих формул. Основним показником було вирішено брати MAE (mean absolute error), через

великий діапазон самих значень , саме цей критерій допомагав найбільш точно та наглядно інтерпретувати результати моделі. Найкращий результат якого змогла досягти модель полягає в  $MAE = 56,2749$  . Що є доволі хорошим та значущим результатом .

Сама запропонована формула має такий вигляд :

$$e^{\left(\frac{44.2}{\log(T-1.82)}\right) \sin\left(\log\left(\left(\frac{T-63.1}{0.328Nd+T^{6.15}}\right)^{-0.0606}\right)\right)}$$

З результату видно , що модель користується запропонованими нашими функціями , що дає потенційно гарні результати .

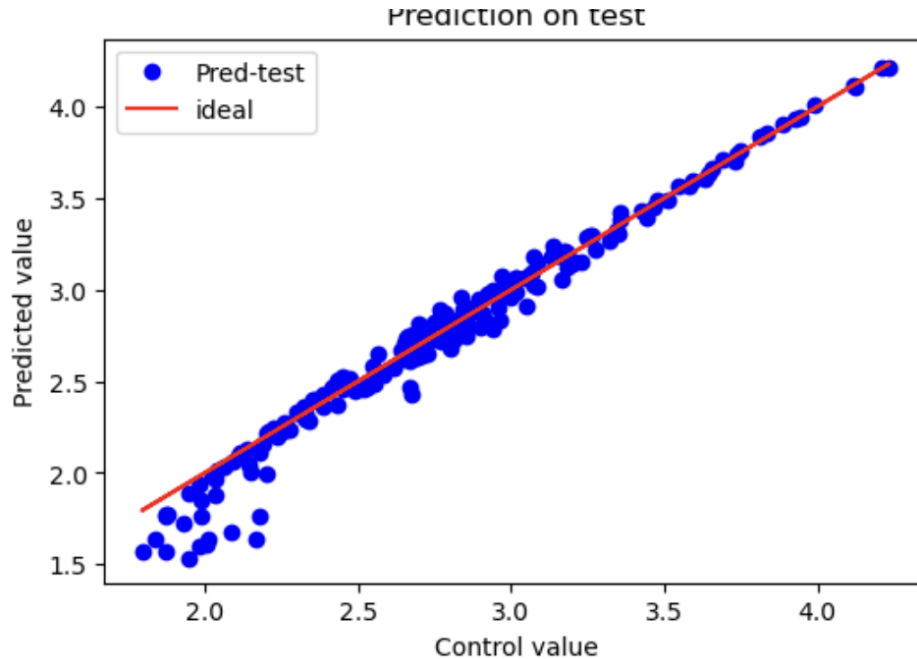


Рис. 2: графік розсіювання передбачуваних значень , отриманих на моделі символічної регресії

## 5 Висновок

Протягом цієї практики я вивчав теоретичні основи методів розрахунку рухливості основних та неосновних носіїв заряду в кремнії, зокрема методи Классена, Аврори та Флетчера. Це забезпечило розуміння фундаментальних фізичних принципів та математичних підходів, які лежать в основі цих методів. Окрім теоретичного опрацювання, значна увага була приділена практичній частині, а саме застосуванню методів машинного навчання для параметризації даних. Використання алгоритмів Random Forest та символічної регресії дозволило порівняти ефективність різних підходів для моделювання та аналізу залежностей, пов'язаних із рухливістю носіїв заряду. Такі методи дали змогу отримати точні результати та виявити ключові закономірності в даних.

## Список літератури

### Література

- [1] D.B.M. Klaassen. A Unified Mobility Model for Device Simulation. Philips Research Laboratories, 5600 JA Eindhoven, The Netherlands.



- [2] D.B.M. Klaassen. A Unified Mobility Model for Device Simulation–II. Temperature Dependence of Carrier Mobility and Lifetime. Philips Research Laboratories, P.O. Box 80000, 5600 JA Eindhoven, The Netherlands.
- [3] Sofos F., Dritselis C., Misdanitis S., Karakasidis T., Valougeorgis D. Computation of flow rates in rarefied gas flow through circular tubes via machine learning techniques. 2023.
- [4] Kronberger G., Burlacu B., Kommenda M., Winkler S. M., Affenzeller M. Symbolic Regression. Chapman & Hall, 2024. 312 c.
- [5] PySR: High-Performance Symbolic Regression in Python and Julia. Available at: <https://astroautomata.com/PySR/> (accessed June 3, 2024).