

1.2. Моделі для опису рухливості носіїв заряду у кремнії

Задача оцінки рухливості електронів μ_n та дірок μ_p у напівпровіднику за певних умов є достатньо поширеною у різноманітних фізичних дослідженнях. Один з варіантів її вирішення полягає у використанні загального підходу, згідно з яким

$$\mu = \frac{e\tau_p}{m_\sigma}, \quad (1.1)$$

де e — елементарний заряд, τ_p — середній час вільного пробігу носія заряду, m_σ — ефективна маса електропровідності.

Час вільного пробігу обмежується розсіянням носіїв заряду, яке може бути викликане декількома причинами, пов'язаними з порушеннями періодичності потенціалу. Зокрема виділяють розсіяння на коливаннях ґратки (акустичних та оптичних фононах), заряджених та нейтральних домішках, дислокаціях, границях зерен та інших неоднорідностях структури, поверхнях та межах розділу, інших носіях. Кожен із цих механізмів має свою залежність від температури, рівня легування та розміру напівпровідникової структури і може бути визначальним для величини рухливості за певних умов. Проте найчастіше необхідно враховувати декілька можливих шляхів розсіяння носіїв заряду. В такому випадку для оцінки рухливості використовується правило Маттісена:

$$\mu^{-1} = \sum_i \mu_i^{-1}, \quad (1.2)$$

де сумування відбувається за механізмами розсіяння, μ_i — рухливість носіїв за наявності лише i -го механізму розсіяння. Для оцінки μ_i можна використовувати вирази, аналогічні формулі (1.1), розрахувавши відповідний час вільного пробігу.

Для переважної більшості механізмів розсіяння вирази для оцінки рухливості відомі. Так, при розсіянні на іонізованих домішках нерідко використовується вираз Brooks & Herring [?]:

$$\mu_I = \frac{3,68 \cdot 10^{20} \text{ см}^{-3}}{N_I Z^2} \left(\frac{\varepsilon}{16} \right)^2 \left(\frac{T}{100 \text{ К}} \right)^{1,5} \frac{(m_0/m)^{0,5}}{\left[\lg(1 + \beta_{\text{ВН}}^2) - \frac{0,434\beta_{\text{ВН}}^2}{1+\beta_{\text{ВН}}^2} \right]}, \quad (1.3)$$

де N_I — концентрація домішок із зарядом Ze , ε — діелектрична проникність напівпровідника, T — температура, m — ефективна маса носія, m_0 — маса вільного електрону, а величина $\beta_{\text{ВН}}$ має вигляд

$$\beta_{\text{ВН}} = \left(\frac{\varepsilon}{16} \right)^{0,5} \frac{T}{100 \text{ К}} \left(\frac{m}{m_0} \right)^{0,5} \left(\frac{2,08 \cdot 10^{18} \text{ см}^{-3}}{n_c} \right)^{0,5}, \quad (1.4)$$

n_c — концентрація носіїв заряду. Іншим наближенням для такого випадку є формула Conwell & Weisskopf [?]:

$$\mu_I = \frac{3,68 \cdot 10^{20} \text{ см}^{-3}}{N_I Z^2} \left(\frac{\varepsilon}{16} \right)^2 \left(\frac{T}{100 \text{ K}} \right)^{1,5} \frac{(m_0/m)^{0,5}}{\lg(1 + \beta_{CW}^2)}, \quad (1.5)$$

$$\beta_{CW} = \frac{1}{Z} \frac{\varepsilon}{16} \frac{T}{100 \text{ K}} \left(\frac{2,35 \cdot 10^{19} \text{ см}^{-3}}{N_I} \right)^{1/3}.$$

Для іншого механізму, розсіяння носій-носії, застосовується підхід, розвинутий Fletcher [?]:

$$\mu_{cc} = \frac{\left(\frac{T}{T_{ref}^{Fl}} \right)^{3/2} F_1}{(np)^{1/2} \ln \left[1 + \left(\frac{T}{T_{ref}^{Fl}} \right)^2 (np)^{-1/3} F_2 \right]}, \quad (1.6)$$

де n та p — концентрації електронів та дірок, відповідно; T_{ref}^{Fl} , F_1 та F_2 — певні константи, які залежать від матеріалу. Як показано в роботах [?, ?], для кремнію доцільно застосовувати $T_{ref}^{Fl} = 300 \text{ K}$, $F_1 = 1,04 \cdot 10^{21} \text{ см}^{-1} \text{ В}^{-1} \text{ с}^{-1}$, $F_2 = 7,45 \cdot 10^{12} \text{ см}^2$ і тоді вираз (1.6) перетворюється на наступний

$$\mu_{cc} = \frac{2 \cdot 10^{17} T^{3/2}}{\sqrt{n p} \ln \left[1 + 8,28 \cdot 10^8 T^2 (np)^{-1/3} \right]}. \quad (1.7)$$

Проте в більшості випадків для більш-менш точного опису рухливості реального матеріалу необхідно враховувати значну кількість механізмів. Як наслідок, більше поширення отримав підхід оцінки величини μ з використанням апроксимаційної функції, яка часто базується на результатах експериментальних вимірювань. Як правило, для кожного матеріалу вигляд функції або наявні в ній коефіцієнти відрізняються.

Розглянемо декілька подібних підходів до опису рухливості носіїв заряду у монокристалічному кремнії, обмежуючись випадками об'ємного напівпровідника (без врахування впливу поверхні) та слабких полів. Одним з перших подібних наближень був вираз, запропонований Scaughey & Thomas [?]:

$$\mu = \mu_{\min} + \frac{\mu_{\max} - \mu_{\min}}{1 + (N/N_{ref})^\alpha}, \quad (1.8)$$

де N — концентрація легантів, а значення констант для випадків, коли розглядаються рухливості електронів та дірок, наведені у Табл. 1. Вираз насамперед призначений для оцінки залежності рухливості основних носіїв від концентрації легуючої домішки поблизу 300 K. З іншого

Табл. 1. Коефіцієнти для розрахунку рухливості відповідно до моделі Caughey–Thomas (1.8)

Тип носіїв	Параметр			
	$\mu_{\max}, \text{см}^2/(\text{В} \cdot \text{с})$	$\mu_{\min}, \text{см}^2/(\text{В} \cdot \text{с})$	α	$N_{ref}, \text{см}^{-3}$
Електрони	1330	65	0,72	$8,5 \cdot 10^{16}$
Дірки	495	47,7	0,76	$6,3 \cdot 10^{16}$

боку, формула (1.8) не дозволяє оцінити рухливості неосновних носіїв та не враховує температурну залежність μ .

Модифікований варіант такого підходу був запропонований Masetti [?]:

$$\begin{aligned}\mu_n &= \mu_0 + \frac{\mu_{\max} - \mu_0}{1 + \left(\frac{n}{C_r}\right)^a} - \frac{\mu_1}{1 + \left(\frac{C_s}{n}\right)^b}, \\ \mu_p &= \mu_0 \exp\left(-\frac{p_c}{p}\right) + \frac{\mu_{\max}}{1 + \left(\frac{p}{C_r}\right)^a} - \frac{\mu_1}{1 + \left(\frac{C_s}{p}\right)^b}.\end{aligned}\quad (1.9)$$

Як видно з виразу (1.9), в такому випадку передбачено застосування різних виразів для опису рухливості електронів та дірок. Більше того, коефіцієнти мають залежати від того, які атоми були використані для легування — у Табл. 2 приведені їхні значення для бору та фосфору, які є найбільш типовими домішками для отримання кремнію з дірковою та електронною провідностями, відповідно. Теорія Masetti краще описує експериментальні дані, ніж підхід Caughey–Thomas, проте так само не дозволяє отримати інформацію для неосновних носіїв та температур, відрізняючись від кімнатної.

Спроба взяти до уваги температурні залежності була зроблена Н. Агога зі співавторами [?]. При цьому розглядалися розсіяння на фононах та іонізованих домішках і було запропоновано використання узагальненого виразу Caughey–Thomas

$$\mu = \mu_{\min} + \frac{\mu_0}{1 + \left(\frac{N}{N_0}\right)^\alpha}, \quad (1.10)$$

де $N = N_a^- + N_d^+$ — загальна концентрація іонізованих легантів (N_a^- та N_d^+ — кількості іонізованих акцепторів та донорів у одиниці об'єму, відповідно), а решта коефіцієнтів у виразі (1.10) залежать від темпе-

Табл. 2. Коефіцієнти для розрахунку рухливості відповідно до моделі Masetti (1.9)

Параметр	Легант	
	Фосфор	Бор
$\mu_0, \text{см}^2/(\text{В} \cdot \text{с})$	68,5	44,9
$\mu_{\max}, \text{см}^2/(\text{В} \cdot \text{с})$	1414	470,5
$\mu_0, \text{см}^2/(\text{В} \cdot \text{с})$	56,1	29,0
$C_r, \text{см}^{-3}$	$9,2 \cdot 10^{16}$	$2,23 \cdot 10^{17}$
$C_s, \text{см}^{-3}$	$3,41 \cdot 10^{20}$	$6,10 \cdot 10^{20}$
a	0,711	0,719
b	1,98	2,00
$p_c, \text{см}^{-3}$	—	$9,23 \cdot 10^{16}$

ратури схожим чином:

$$\begin{aligned}
 \mu_{\min} &= \mu_{\min}^{ref} \left(\frac{T}{T_{ref}} \right)^{\beta_1}, \\
 \mu_0 &= \mu_0^{ref} \left(\frac{T}{T_{ref}} \right)^{\beta_2}, \\
 N_0 &= N_0^{ref} \left(\frac{T}{T_{ref}} \right)^{\beta_3}, \\
 \alpha &= \alpha^{ref} \left(\frac{T}{T_{ref}} \right)^{\beta_4}.
 \end{aligned} \tag{1.11}$$

Тобто, відбулося повернення до ситуації, коли вираз для опису рухливості не залежить від типу носіїв, проте різними є константи, що мають бути використані під час розрахунків. Зауважимо, що у оригінальній роботі [?] вирази отримані (та визначені значення констант) на основі аналізу експериментальних даних щодо рухливості основних носіїв при різних значеннях температури та ступеню легування. Проте нерідко вирази (1.10)-(1.11) поширюють і для неосновних носіїв, використовуючи значення констант, отримані в роботах [?] (для неосновних електронів) та [?] (для неосновних дірок). Проте у вказаних роботах приведені дані лише для однієї температури і тому поширення виразів на широкий температурний діапазон дещо самовпевнено. Значення констант для всіх випадків наведено у Табл. 3.

Проте найбільш точним з точки зору збіжності з експериментальними даними вважається підхід до оцінки рухливості, розвинутий Klaassen

Табл. 3. Коефіцієнти для розрахунку рухливості відповідно до моделі Arora (1.10)-(1.11)

Параметр	Тип носіїв			
	електрони		дірки	
	основні [?]	неосновні [?]	основні [?]	неосновні [?]
T_{ref} , К	300	295	300	292
μ_{min}^{ref} , см ² /(В·с)	88	232	54.3	130
μ_0^{ref} , см ² /(В·с)	1252	1180	407	370
N_0^{ref} , см ⁻³	$1,26 \cdot 10^{17}$	$8 \cdot 10^{16}$	$2,35 \cdot 10^{17}$	$8 \cdot 10^{17}$
α^{ref}	0,88	0,9	0,88	1,25
β_1	-0,57			
β_2	-2,33		-2,23	
β_3	2,4			
β_4	-0,146			

[?] (на який в літературі також посилаються як на модель Philips). В цьому випадку вважається, що визначальними для рухливості є процеси розсіяння на коливаннях ґратки, іонізованих домішках та інших носіях, а температурну та концентраційні залежності μ як для основних, так і неосновних носіїв заряду можна описати наступними виразами:

$$\mu = \frac{\mu_L \mu_{DA}}{\mu_L + \mu_{DA}}, \quad (1.12)$$

де

$$\mu_L = \mu_{max} \left(\frac{300}{T} \right)^{2,25} \quad (1.13)$$

$$\begin{aligned} \mu_{DA} = & \frac{\mu_{max}^2}{\mu_{max} - \mu_{min}} \cdot \frac{N_{sc}}{N_{eff}} \cdot \left(\frac{N_{ref}}{N_{sc}} \right)^\alpha \cdot \left(\frac{T}{300} \right)^{3\alpha - 1,5} \\ & + \frac{\mu_{max} \mu_{min}}{\mu_{max} - \mu_{min}} \cdot \frac{n + p}{N_{eff}} \cdot \left(\frac{300}{T} \right)^{0,5} \end{aligned} \quad (1.14)$$

а константи, що входять до рівнянь (1.13)-(1.14) наведені у Табл. 4. Величини N_{sc} та N_{eff} є функціями, вигляд яких залежить від типу носіїв:

$$\text{електрони: } N_{sc} = N_d^+ + N_a^- + p, \quad (1.15a)$$

$$\text{дірки: } N_{sc} = N_a^- + N_d^+ + n, \quad (1.156)$$

Табл. 4. Коефіцієнти для розрахунку рухливості відповідно до моделі Klaassen, формули (1.13)-(1.14)

Тип носіїв	Параметр			
	$\mu_{\max}, \text{см}^2/(\text{В} \cdot \text{с})$	$\mu_{\min}, \text{см}^2/(\text{В} \cdot \text{с})$	α	$N_{ref}, \text{см}^{-3}$
Електрони	1414	68,5	0,711	$9,2 \cdot 10^{16}$
Дірки	495	44,9	0,719	$2,23 \cdot 10^{17}$

$$\text{електрони: } N_{\text{eff}} = N_d^+ + N_a^- \cdot G_n(P_n, T) + \frac{p}{F_n(P_n, T)}, \quad (1.16a)$$

$$\text{дірки: } N_{\text{eff}} = N_a^- + N_d^+ \cdot G_p(P_p, T) + \frac{n}{F_p(P_p, T)}. \quad (1.16b)$$

В свою чергу, для функцій $G_{n(p)}$, $F_{n(p)}$ та $P_{n(p)}$ справедливі наступні вирази:

$$G_n = 1 - \frac{S_1}{\left[S_2 + \left(\frac{m_0}{m_e} \frac{T}{300} \right)^{S_4} P_n \right]^{S_3}} + \frac{S_5}{\left[\left(\frac{m_e}{m_0} \frac{300}{T} \right)^{S_7} P_n \right]^{S_6}}, \quad (1.17a)$$

$$G_p = 1 - \frac{S_1}{\left[S_2 + \left(\frac{m_0}{m_h} \frac{T}{300} \right)^{S_4} P_p \right]^{S_3}} + \frac{S_5}{\left[\left(\frac{m_h}{m_0} \frac{300}{T} \right)^{S_7} P_p \right]^{S_6}}, \quad (1.17b)$$

$$F_n = \frac{r_1 P_n^{r_6} + r_2 + r_3 \frac{m_e}{m_h}}{P_n^{r_6} + r_4 + r_5 \frac{m_e}{m_h}}; \quad F_p = \frac{r_1 P_p^{r_6} + r_2 + r_3 \frac{m_h}{m_e}}{P_p^{r_6} + r_4 + r_5 \frac{m_h}{m_e}}, \quad (1.18)$$

$$P_n = (P_{bn} + P_{cn})^{-1}; \quad P_p = (P_{bp} + P_{cp})^{-1}, \quad (1.19)$$

де m_e та m_h — ефективні маси електрона та дірки, відповідно, а константи, що входять до виразів (1.17) та (1.18), наведені в Табл. 5. У випадку, коли мова йде про неосновні носії, то $P_{cn} = 0$ ($P_{cp} = 0$). При розрахунках для основних носіїв

Табл. 5. Коефіцієнти для розрахунку рухливості відповідно до моделі Klaassen, формули (1.17) та (1.18)

Параметр	S_1	S_2	S_3	S_4	S_5	S_6	S_7
Значення	0,89233	0,41372	0,19778	0,28227	0,005978	1,80618	0,72169
Параметр	r_1	r_2	r_3	r_4	r_5	r_6	
Значення	0,7643	2,2999	6,5502	2,3670	-0,01552	0,6478	

$$P_{cn} = \frac{2,46}{3,97 \cdot 10^{19} \left[\left(\frac{T}{300 Z_n} \right)^3 \frac{1}{N_d^+} \right]^{2/3}}; \quad P_{cp} = \frac{2,46}{3,97 \cdot 10^{19} \left[\left(\frac{T}{300 Z_p} \right)^3 \frac{1}{N_a^-} \right]^{2/3}}, \quad (1.20a)$$

$$Z_n = 1 + \frac{1}{0,21 + \left(\frac{4 \cdot 10^{20}}{N_d^+} \right)^2}; \quad Z_p = 1 + \frac{1}{0,5 + \left(\frac{7,2 \cdot 10^{20}}{N_a^-} \right)^2}, \quad (1.20б)$$

$$P_{bn} = \frac{3.83}{\frac{1,36 \cdot 10^{26}}{N_d^+ + p} \left(\frac{T}{300} \right)^2 \frac{m_e}{m_0}}; \quad P_{bp} = \frac{3.83}{\frac{1,36 \cdot 10^{26}}{N_a^- + n} \left(\frac{T}{300} \right)^2 \frac{m_h}{m_0}}. \quad (1.20в)$$

Як можна бачити, підхід Klaassen (1.12)-(1.20) передбачає застосування достатньо комплексних виразів та більше двадцяти констант. Це, звичайно, не створює непереборних складностей, але призводить до того, що на практиці для оцінки рухливості носіїв заряду в кремнії нерідко використовується більш простий підхід Agora (1.10)-(1.11), незважаючи на його меншу точність.

Зазначимо, що питання опису рухливості продовжує привертати увагу і періодично з'являються нові моделі. Наприклад, пропонуються вирази, які дозволяють врахувати додаткові механізми розсіювання за рахунок модифікації моделі Masetti [?], або використовують апроксимаційні поліноми [?].

2. Методика обчислень

2.1. Створення тренувального та тестового наборів

Метою цієї роботи було створення моделей для оцінки рухливості носіїв заряду у кремнії при певних температурі T та рівні легування N_D з використанням алгоритмів машинного навчання (МН). Іншими словами, з точки зору МН, ми вирішували задачу у якій набір ознак (дескрипторів) складався з двох величин (T, N_D) , а цільовим результатом була величина μ . Зрозуміло, що для досягнення поставленої мети насамперед необхідні дані і для їх отримання ми використовували модель Klaassen. Тобто, передбачення моделі Klaassen вважалися істинними і ми намагалися створити регресійні моделі, передбачення яких максимально узгоджуються з підходом, запропонованим у роботі [?].

У роботі ми розглядали наступні діапазони температур та рівнів легування:

$$T \in [200, 500] \text{ K}; \quad N_D \in [10^{13}, 10^{19}] \text{ cm}^{-3}. \quad (2.1)$$

Під час створення тренувального набору даних випадковим чином вибиралися 500 пар значень (T, N_D) з діапазону (2.1) і обчислювалися значення μ , використовуючи вирази (1.12)-(1.20) та дані таблиць 4 та 5. Тестовий набір, який використовувався для оцінки якості передбачень моделей, складався з 2550 зразків, що відповідали всім комбінаціям 51 значень температури (від 200 K до 500 K з кроком 6 K) та 50 значень рівня легування, рівномірно розподіленим в логарифмічному масштабі від 10^{13} cm^{-3} до 10^{19} cm^{-3} . У роботі розроблялися моделі, здатні передбачати рухливість як основних так і неосновних електронів та дірок у кремнії і для кожного з цих чотирьох випадків створювалися окремі тренувальний та тестові набори. Надалі рухливість електронів у n-Si позначатиметься $\mu_{n,n}$, у p-Si — $\mu_{n,p}$. Аналогічно, $\mu_{p,p}$ та $\mu_{p,n}$ позначатимуть рухливості дірок, коли вони є основними та неосновними носіями заряду, відповідно.

Під час розрахунків величини рухливості вважалося, що: 1) у кристалі наявні легуючі домішки лише одного типу (випадок компенсованого напівпровідника не розглядався); 2) концентрація нерівноважних (інжектованих) носіїв набагато менша концентрації електронів у n-Si та дірок у p-Si.

У випадку електронного напівпровідника для обчислення концентрацій носіїв заряду використовувався вираз

$$n = \frac{N_d^+}{2} + \sqrt{\left(\frac{N_d^+}{2}\right)^2 + n_i^2}, \quad (2.2)$$

$$p = \frac{n_i^2}{n}, \quad (2.3)$$

де n_i — концентрація електронів у власному напівпровіднику, для обчислення температурної залежності якої, використовувався підхід з роботи [?]. В свою чергу, кількість іонізованих донорів у одиниці об'єму обчислювалася з урахуванням статистики Фермі-Дірака:

$$N_d^+ = N_d \left(1 - \frac{1}{1 + 0,5 \exp \left[\frac{E_F - E_d}{kT} \right]} \right), \quad (2.4)$$

де E_F — положення рівня Фермі, E_d — енергія розташування донорного рівня. Оцінка розташування рівня Фермі здійснювалася з використанням умови електронейтральності

$$n = p + N_d^+, \quad (2.5)$$

при цьому вважалося, що

$$n = N_C F_{1/2} \left(-\frac{E_F}{kT} \right), \quad (2.6)$$

$$p = N_V F_{1/2} \left(-\frac{E_G - E_F}{kT} \right), \quad (2.7)$$

де N_C та N_V — ефективні густини станів біля дна зони провідності та вершини валентної зони, обчислювалися відповідно до [?]; E_G — ширина забороненої зони, яка розраховувалася згідно з наближенням Passler [?]; $F_{1/2}$ — інтеграл Фермі-Дірака ступеня 1/2, для обчислення якого використовувався підхід [?].

У випадку р-Si розрахунки проводилися аналогічно.

І для електронного, і для діркового напівпровідника вважалося, що енергія активації легуючої домішки складає 45 меВ. Ця величина відповідає енергіям іонізації атомів бору та фосфору, які є основними легуючими домішками для кремнію.

Результати розрахунків істинних значень рухливості представлені на Рис. 2.1.

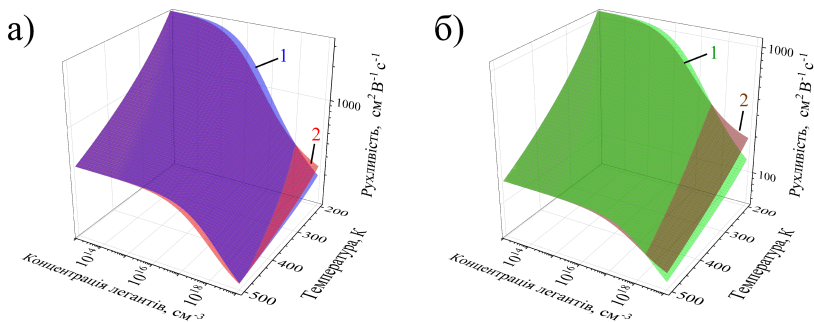


Рис. 2.1. Залежності рухливості електронів (а) та дірок (б) у випадку, коли вони є основними (поверхні 1) та неосновними (поверхні 2) носіями заряду від температури та концентрації легуючої домішки. Розрахунки виконані відповідно до теорії Klaassen.

2.2. Налаштування Symbolic Regression

Символьна регресія (Symbolic Regression, SR) — це достатньо новий алгоритм МН, який апроксимує залежності між вхідними та вихідними параметрами та дозволяє отримати аналітичний зв'язок між ними у вигляді математичної формули [?]. При побудові функціональної залежності та пошуку числових параметрів використовуються еволюційні алгоритми. SR є одним із методів, які об'єднують дані, моделі машинного навчання та наукові теорії, дозволяючи отримати аналітичні вирази, які можуть бути інтерпретовані, дозволяють отримати передбачення дешевши з точки зору обчислень порівняно з класичними моделями МН та легше інтегруються у вирішення різноманітних фізичних проблем.

При реалізації цього алгоритму використовувався open source пакет PySR, орієнтований на використання мови Python. Для побудови рівнянь використовувалися стандартні математичні функції (+, −, ×, ÷, exp, ln, x^y , tanh, sin) та дійсні константи. Крім того використовувалися спеціально сконструйовані оператори $f_1(x) = \frac{1}{1+x}$ та $f_2(x, y) = \frac{x \cdot y}{x+y}$. Ці оператори відповідають функціям, які часто зустрічаються в існуючих моделях для опису рухливості носіїв заряду — див. Розділ 1.2.

SR може призвести до складних виразів, які важко інтерпретувати через те, що вони, що містить небажані ознаки, такі як, наприклад, вкладені операції. Щоб запобігти цьому застосовувалися обмеження, які забороняли дві послідовні операції sin (тобто sin(sin(•))), f_1 , f_2 та

\tanh , а також вкладеність \tanh в f_1 та f_2 , та \ln і x^y в \exp . Також було заборонено комплексність (складність) показника експоненти вище 3 та показника функції x^y вище 5.

Попередня обробка даних полягала у їхній нормалізації, тобто на вхід SR моделі надходили величини T_n N_{Dn} , пов'язані з реальними температурою та концентрацією легуючої домішки наступним чином:

$$T_n = \frac{T}{300 \text{ K}}, \quad (2.8a)$$

$$N_{Dn} = \frac{N_D}{10^{17} \text{ см}^{-3}}. \quad (2.8b)$$

2.3. Класичні регресійні моделі

Окрім символічної регресії, в роботі для оцінки рухливості також були реалізовані декілька моделей, які базуються на використанні більш класичних алгоритмів МН. А саме, були використані:

1) *Random Forest* (RF) — алгоритм, який досягає покращення точності прогнозування завдяки навчанню декількох дерев рішень на різних підмножинах тренувального набору даних; RF агрегує прогнози всіх дерев, використовуючи голосування під час класифікації або усереднення у випадку вирішення регресійних завдань (як в нашому випадку), що дозволяє зменшити перенавчання та підвищуючи стійкість до шуму [?];

Gradient Boosting (GB) — об'єднує декілька слабких моделей, як правило, дерев рішень, з метою покращення ефективності прогнозування; кожна нова модель виправляє помилки попередньої, тим самим підвищуючи загальну точність; остаточний прогноз отримують шляхом агрегування прогнозів, зазвичай за допомогою зваженої суми [?];

Support Vector Regression (SVR) — передбачає знаходження функції, яка наближує залежність між вхідними змінними та цільовими значеннями з максимально допустимою похибкою, водночас забезпечуючи якнайменшу складність моделі; може використовувати різні функції ядра (наприклад, радіально-базисні чи поліноміальні), щоб працювати з нелінійними залежностями [?];

Deep Neural Network (DNN) — складається з декількох шарів взаємопов'язаних нейронів, які обробляють вхідні дані шляхом послідовних нелінійних перетворень [?].

Підкреслимо, що хоча подібні моделі також дозволяють передбачати певні величини (наприклад, рухливість носіїв, як це відбувалося

Табл. 6. Пошуковий простір гіперпараметрів для RF моделей

Гіперпараметр	Можливі значення (діапазон пошуку)
# estimators	100 ÷ 700
max depth	10 ÷ 70
min samples leaf	1, 2, 3, 4, 5
min samples split	2, 3, 4, 5
bootstrap	True, False
max features	0,5, 0,8, 1

в нашому випадку), проте вони є фактично чорною скринькою, що суттєво утруднює обґрунтування отриманих результатів і поступово перестає задовольняти дослідників.

Вказані моделі були імплементовані на мові Python, використовуючи пакети Keras (для DNN) та Scikit-learn (для RF, GB та and SVR).

Розглядалися два варіанти кожного типу моделей, які відрізнялися попередньою обробкою даних. В одному випадку як вхідні ознаки використовувалися нормовані значення температури та концентрації легуючої домішки (визначені відповідно до (2.8)), а прогнози були орієнтовані на отримання величини μ . Відповідні моделі надалі позначатимуться шляхом додавання до аббревіатури використаного алгоритму суфіксу “_N” (RF_N, DNN_N тощо) Проте ознаки достатньо сильно відрізняються за величиною між собою, а N_D ще й змінюється у межах шести порядків, що може бути перешкодою для точних прогнозів. Тому для другого типу моделей проводилася попередня обробка вхідних ознак, яка передбачала нормалізацію (Z-стандартизацію, вже вибачте за усталений термін) як значень температури T , так і логарифмів ступеню легування $\lg N_D$ (здійснювалися лінійні перетворення, в результаті яких тренувальні набори кожної з ознак характеризувалися нульовим середнім і одиничним стандартним відхиленням). Моделі були орієнтовані на передбачення нормалізованих значень $\lg \mu$. Процес нормалізації був реалізований за допомогою функції StandartScaler з пакету Scikit-learn і позначення відповідних моделей міститиме суфікс “_S”.

Відомо, що налаштування гіперпараметрів є надзвичайно важливим для оптимізації продуктивності моделі [?]. Перелік гіперпараметрів, які налаштовувалися для кожної з моделей, а також діапазони пошуку вказані в Таблицях 6-9.

Зазначимо, що у випадку DNN вхідний шар складався з двох вузлів, вихідний містив один вузол, а структура схованих повнозв'язних

Табл. 7. Пошуковий простір гіперпараметрів для GB моделей

Гіперпараметр	Можливі значення (діапазон пошуку)
# estimators	200 ÷ 900
max depth	2 ÷ 40
min samples leaf	1, 2, 3, 4, 5, 6, 7
min samples split	2, 3, 4, 5, 6, 7
learning rate	$10^{-3} \div 1$
max features	0,5, 0,6, 0,7, 0,8, 0,9, 1

Табл. 8. Пошуковий простір гіперпараметрів для SVR моделей

Гіперпараметр	Можливі значення (діапазон пошуку)
kernel	linear, poly, rbf, sigmoid
degree*	2, 3, 4, 5, 6
C_0	0 – 8
Tolerance	$10^{-5} \div 10^{-2}$
C	$10^{-2} \div 15$
ε	$10^{-4} \div 15$

* лише для kernel=poly

Табл. 9. Пошуковий простір гіперпараметрів для DNN моделей

Гіперпараметр	Можливі значення (діапазон пошуку)
# nodes for first hidden layer	5 ÷ 150
batch size	1, 4, 16
activation function	ReLu, sigmoid, tanh, SELU, ELU
optimizer	SGD, RMSprop, Adam, Adadelata, Adagrad, Adamax, Nadam, Ftrl
learning rate	$10^{-4} \div 10^{-2}$
weight initializer	Xavier Normal, Xavier Uniform, He Normal, He Uniform, Random Normal, Random Uniform

шарів наведена на Рис. 2.2. Як видно з рисунку, використовувалося шість схованих шарів, кількість вузлів у яких поступово зменшувалося: від шару до шару на 10% від кількості вузлів у першому схованому N_{nll}), яка була одним із гіперпараметрів, що налаштовувалися.

Налаштування моделей здійснювалося за допомогою пакету Optuna [?] з використанням TPE самплера та Hyperband прюнера для ефе-

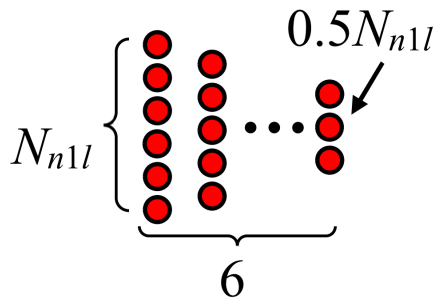


Рис. 2.2. Конфігурація схованих шарів для використаних DNN моделей.

ктивного вибору гіперпараметрів. Під час налаштування використувалась 5-кратна крос-валідація, яка передбачала по чергове використання кожних 20% тренувального набору як валідаційних даних для оцінки ефективності моделей, навчених з використанням решти 80%.

2.4. Метрики оцінювання

3. Отримані результати

3.1. Символьна регресія

$$\mu_{n,n} = \frac{N_{Dn} + 1413,2}{T_n^{2,2508} + \frac{a_{n,n} \cdot b_{n,n}}{a_{n,n} + b_{n,n}}}, \quad (3.1a)$$

$$a_{n,n} = \left(\frac{N_{Dn}}{0,873T_n + 0,0727} \right)^{0,7212}, \quad (3.1б)$$

$$b_{n,n} = 0,215N_{Dn}^{0,58} + 10,5T_n - 1,342, \quad (3.1в)$$

$$\mu_{n,p} = 26,3 \cdot \left(\frac{N_{Dn}}{N_{Dn} + 192} \right)^{0,62} + \frac{1412,3}{T_n^{2,25} + \frac{a_{n,p} \cdot b_{n,p}}{a_{n,p} + b_{n,p}}}, \quad (3.2a)$$

$$a_{n,p} = 1,92 \cdot \left(\frac{N_{Dn}}{T_n + 0,071} \right)^{0,717}, \quad (3.2б)$$

$$b_{n,p} = 5,7N_{Dn}^{0,103} \cdot T_n, \quad (3.2в)$$

$$\mu_{p,p} = N_{Dn}^{0,448} + \frac{470}{T_n^{2,2505} + \frac{a_{p,p} \cdot b_{p,p}}{a_{p,p} + b_{p,p}}}, \quad (3.3a)$$

$$a_{p,p} = 2,652N_{Dn}^{0,1669} \cdot T_n, \quad (3.3б)$$

$$b_{p,p} = 0,578N_{Dn}^{0,7587} \cdot T_n^{-0,755}, \quad (3.3в)$$

$$\mu_{p,n} = \frac{0,0226N_{Dn} \cdot \ln T_n}{\ln T_n + 0,63} + \frac{113,56}{\left[0,15617T_n^{2,94} + \left(\frac{a_{p,n} \cdot b_{p,n}}{1,151a_{p,n} + 4,21b_{p,n}} \right)^{1,177} \right]^{0,76547}}, \quad (3.4a)$$

$$a_{p,n} = N_{Dn}^{0,6085}, \quad (3.4б)$$

$$b_{p,n} = T_n^{1,143}. \quad (3.4в)$$

Табл. 10. Використані значення гіперпараметрів для RF моделей

Модель	RF_N				RF_S			
Цільова величина	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$
Гіперпараметр	Значення							
# estimators	468	500	658	510	278	564	304	204
max depth	19	58	37	19	30	15	55	55
min samples leaf	1	1	1	1	1	1	1	1
min samples split	2	2	2	2	2	2	2	2
bootstrap	True	True	True	True	True	True	True	True
max features	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0

3.2. Класичні регресійні моделі

в Таблицях 10-13.

Табл. 11. Використані значення гіперпараметрів для GB моделей

Модель	GB_N				GB_S			
Цільова величина	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$
Гіперпараметр	Значення							
# estimators	700	800	800	570	840	770	810	850
max depth	4	4	4	4	4	4	2	4
min samples leaf	6	2	5	4	5	3	5	4
min samples split	2	2	3	3	5	5	5	6
learning rate, 10^{-2}	6,44	9,78	7,69	4,68	9,49	4,63	9,67	5,29
max features	0,8	1,0	1,0	0,6	1,0	0,7	1,0	0,8

3.3. Порівняння точності різних алгоритмів машинного навчання

Текст перед таблицею...

Табл. 12. Використані значення гіперпараметрів для SVR моделей

Модель	SVR_N				SVP_S			
Цільова величина	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$
Гіперпараметр	Значення							
kernel	poly	poly	poly	poly	rbf	rbf	rbf	rbf
degree	6	6	6	6	—	—	—	—
C_0	7,52	7,54	7,08	7,02	4,92	2,73	3,88	6,78
Tolerance, 10^{-3}	5,64	2,22	0,60	0,015	0,75	0,32	0,083	0,068
C	9,79	7,04	3,00	4,19	15	15	15	15
ε , 10^{-3}	33,2	11,6	10,4	0,91	0,55	0,50	2,44	0,84

Табл. 13. Використані значення гіперпараметрів для DNN моделей

Модель	DNN_N				DNN_S			
Цільова величина	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$
Гіперпараметр	Значення							
# nodes for first hidden layer	25	100	25	75	25	75	100	25
batch size	16							
activation function	SELU		ReLU	SELU	sigmoid	ReLU	sigmoid	
optimizer	Nadam	Adamax				Nadam	Adamax	
learning rate, 10 ⁻³	0,58	1,25	8,05	5,17	13,0	4,89	12,8	2,84
weight initializer	Xavier Normal	He Normal		Xavier Normal	He Normal			Xavier Normal
# epoch	1000							

Табл. 14. Метрики відхилень розрахованих значень рухливості у кремнії від теорії Klaassen. Жирним виділено найкращі результати для кожної метрики при передбаченнях рухливості електронів та дірок, коли вони є основними та неосновними носіями заряду

Модель	MAPE, %				APE _{MAX} , %				APE _{MED} , %				MAE, cm ² /(B · c)			
	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$	$\mu_{n,n}$	$\mu_{n,p}$	$\mu_{p,p}$	$\mu_{p,n}$
Arora [?]	5,35	20,3	5,62	37,8	38,2	70,7	17,6	111	3,67	16,3	5,01	32,6	45,7	103	16,0	82,6
RF_N	1,29	1,69	1,52	1,47	9,33	14,5	4,60	10,9	0,805	0,994	0,264	0,913	11,0	13,5	4,50	4,76
RF_S	1,42	1,77	1,46	1,41	12,3	14,4	14,5	12,6	0,852	1,04	0,744	0,872	11,6	14,1	4,40	4,76
SVR_N	46,5	51,5	32,6	33,4	1630	1750	2070	1510	14,9	17,4	13,1	12,9	220	244	68,8	73,6
SVR_S	0,108	0,087	0,098	0,082	1,71	2,16	2,26	3,56	0,088	0,064	0,069	0,050	1,20	1,03	0,394	0,410
GB_N	2,20	1,46	1,60	1,55	58,2	10,8	17,3	11,0	1,46	0,939	0,905	1,15	15,43	10,85	4,75	5,24
GB_S	1,32	1,60	1,82	1,23	1,42	11,7	15,0	8,48	0,841	1,104	1,19	0,874	11,3	13,4	5,92	4,19
DNN_N	0,925	0,942	2,01	0,529	9,39	20,0	8,07	6,02	0,621	0,702	1,893	0,383	5,23	5,37	4,22	1,39
DNN_S	0,374	0,623	0,655	0,671	2,23	3,75	5,96	2,97	0,323	0,392	0,556	0,631	3,58	3,49	2,88	1,97
SR	0,094	0,137	0,140	0,050	1,19	0,992	1,48	0,596	0,044	0,077	0,079	0,026	0,475	0,708	0,314	0,115

Текст після таблиці...

Висновки