УДК 004.8

М. І. Кривошея; Р. Н. Квєтний, д-р техн. наук, проф.

МІНІМАКСНА АПРОКСИМАЦІЯ В БАЙЄСІВСЬКИХ НЕЙРОННИХ МЕРЕЖАХ

У статті досліджується застосування мінімаксної стратегії до байєсівських нейронних мереж (BNN) як ефективного підходу до підвищення стійкості моделей машинного навчання в умовах невизначеності, шуму та адверсарних впливів. Актуальність дослідження зумовлена обмеженнями класичних нейронних мереж і навіть стандартних BNN, які можуть демонструвати нестабільність навчання, переоцінювати впевненість у передбаченнях та втрачати узагальнювальну здатність поза межами тренувальних даних. Запропонований підхід поєднує байєсівське моделювання невизначеності з мінімаксною оптимізацією, що дозволяє враховувати найгірші сценарії збурення даних у процесі навчання. У роботі наведено теоретичне обґрунтування методу, де задача навчання формалізується, як двоосібна гра між моделлю та адверсарним середовищем, яке генерує несприятливі збурення вхідних даних. Такий підхід дозволяє мінімізувати максимальні втрати та підвищити робастність моделі. Реалізація мінімаксної BNN здійснюється через інтеграцію адверсарного навчання у варіаційний байєсівський підхід, що дає змогу одночасно враховувати як епістемічну, так і алеаторну невизначеність. Експериментальна частина дослідження базується на задачі апроксимації функцій різного характеру, зокрема періодичних та експоненційних залежностей, за наявності гаусівського шуму. Проведено порівняльний аналіз класичної BNN та мінімаксної BNN за такими критеріями, як стабільність навчання, динаміка функції втрат, точність апроксимації та дисперсія передбачень. Результати показують, що мінімаксна модель демонструє більш стабільну збіжність, зменшення коливань втрат, кращу поведінку на краях області визначення та підвищену узагальнювальну здатність. Продемостровано, що використання мінімаксної стратегії дозволяє ефективно пригнічувати вплив шуму, зменшувати перенавчання та забезпечувати більш надійні передбачення в умовах обмежених або зашумлених даних. Запропонований підхід може бути застосований у задачах, де критичною є стійкість моделі до невизначеності, зокрема в аналізі ризиків, обробці сигналів та інтелектуальних системах прийняття рішень.

Ключові слова: *байєсівська нейронна мережа, мінімаксна оптимізація, апроксимація, перенавчання, невизначеність, регуляризація.*

Вступ

У сучасному штучному інтелекті ключовим викликом залишається стабільність і надійність передбачень нейронних мереж у непередбачуваному середовищі. Традиційні нейронні мережі мають фіксовані параметри та не дають змоги оцінити невизначеність у своїх передбаченнях. Байєсівські нейронні мережі (BNN) частково вирішують цю проблему, вводячи стохастичність у ваги та моделюючи розподіли [1].

Однак класичні BNN також мають низку обмежень: вони можуть переоцінювати впевненість, особливо за межами тренувальних даних, або давати нестабільні результати під дією адверсарних шумів. У таких умовах доцільно застосувати мінімаксну оптимізацію – метод, що дозволяє зробити модель стійкішою до найгірших сценаріїв [2].

Актуальність дослідження

Аналіз літературних джерел свідчить, що байєсівські нейронні мережі (BNN) є одним з базових інструментів моделювання епістемічної та алеаторної невизначеності в задачах машинного навчання. Значна кількість робіт присвячена теоретичним аспектам байєсівського навчання, а також практичним застосуванням BNN у задачах класифікації, регресії та оцінювання ризиків [3, 4].

Окремий напрям досліджень стосується стратегій прийняття рішень в умовах невизначеності, зокрема мінімаксних, байєсівських та ризик-орієнтованих підходів [5, 6]. Мінімаксні методи активно застосовуються для підвищення робастності моделей до шуму,

збурень даних та змагальних впливів [7, 8]. Разом із тим, більшість наявних досліджень зосереджені на задачах класифікації або послідовного прийняття рішень, тоді як питання застосування мінімаксної апроксимації у задачах функціональної регресії з використанням BNN залишаються недостатньо вивченими.

Таким чином, попри наявність значної кількості робіт з байєсівських нейронних мереж і стратегій прийняття рішень, поєднання мінімаксної апроксимації з байєсівським підходом у задачах обробки даних та аналізу похибок потребує подальших досліджень, що і зумовлює актуальність цієї роботи.

Теоретичне підґрунтя

Байєсівська нейронна мережа – це нейромережа, в якій ваги моделюються як випадкові змінні [9], [10]. Замість фіксованих параметрів ми маємо розподіли, що дозволяє оцінювати апостеріорну ймовірність ваг за даними:

$$\mathrm{P(w|D)} \propto \mathrm{P(D|w)} \cdot \mathrm{P(w)} \quad (1)$$

де:

- $\mathrm{P(w)}$ – апріорний розподіл ваг,
- $\mathrm{P(D|w)}$ – правдоподібність даних,
- D – навчальна вибірка.

Мінімаксна стратегія – це стратегія, яка мінімізує максимальні втратигравця у найгіршому випадку.

$$\min_\theta \max_\xi L(\theta, \xi) \quad (2)$$

У конткксті BNN:

- θ – розподіл ваг,
- ξ – адверсарний шум

У запропонованому підході класичну функцію втрат замінює адверсарна функція, де враховується вплив несприятливих сценаріїв (наприклад, шуму або адверсарних прикладів). Це формалізується як двоосібна гра:

$$\min_{q(w)} \max_{noise} E_{w\sim q(w)}[L(w, D)] \quad (3)$$

Мета та структура експерименту

Метою серії експериментів є порівняння поведінки класичної байєсівської нейронної мережі (BNN) і мінімаксноїбайєсівської мережі (Minimax BNN) в умовах шуму [4].

Для оцінювання властивостей мінімаксної апроксимації було обрано набір тестових функцій різного характеру.

1. $y = \sin x$використовується як приклад гладкої періодичної залежності з регулярною зміною кривизни.
2. $y = \cos x$дозволяє проаналізувати вплив фазового зсуву на поведінку байєсівської моделі.
3. $y = \exp(-x^2)$характеризується швидким зростанням та підвищеною чутливістю до крайових значень області визначення, що є типовим випадком для аналізу невизначеності та екстраполяції.

Такий вибір функцій забезпечує комплексну перевірку властивостей мінімаксної апроксимації в умовах різної гладкості, масштабу значень та поведінки похибки.

Кожна функція апроксимується в межах $x\in[-3, 3]$ на 100 точках з додаванням гаусівського шуму до цільових значень.

Опис алгоритму

1. Варіаційна BNN

Модель має варіаційні лінійні шари, де ваги не фіксовані, а описані як розподіли з параметрами μ, ρ. Кожного разу при передбаченні модель бере випадкову вибірку ваг, що дозволяє моделювати епістемічну невизначеність [5].

2. Мінімаксна BNN

Модель навчається на адверсарно збурених вхідних даних. При кожному кроці вона генерує невелике збурення δ, яке збільшує втрати. Модель оптимізується проти цього "противника", що робить її стійкішою до зашумлених або викривлених даних [6].

$$\delta = \alpha \cdot sign(\nabla_x L) \text{ та } x' = x + \delta \quad (4)$$

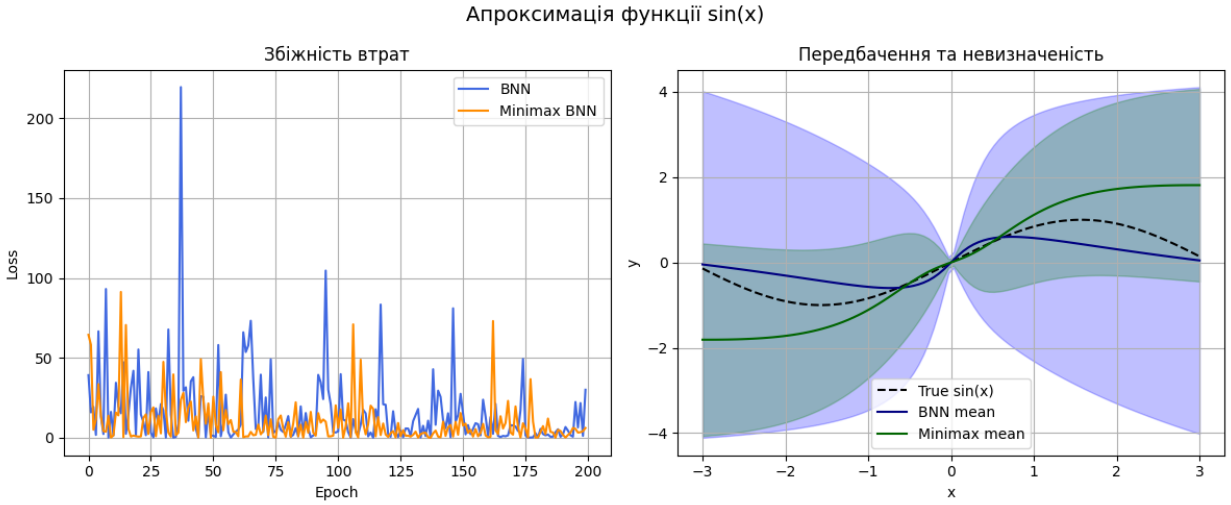

Рис. 1. Апроксимація функції sin(x)

Ліва частина – графік збіжності втрат: синя лінія – класична BNN, помаранчева – Minimax BNN.

Класична BNN демонструє високу волатильність втрат, особливо на ранніх етапах, із окремими піками понад 200. Це свідчить про нестійкість навчання.

Мінімаксна BNN має стабільнішу динаміку зниження втрат, без різких коливань. Це результат регуляризації через адверсарне навчання.

Права частина – передбачення:

– класична BNN має високу невизначеність на краях (x≈±3), що візуалізується широкою напівпрозорою зоною;

– Minimax BNN точніше наближає форму синусоїди на краях і має нижчу дисперсію.

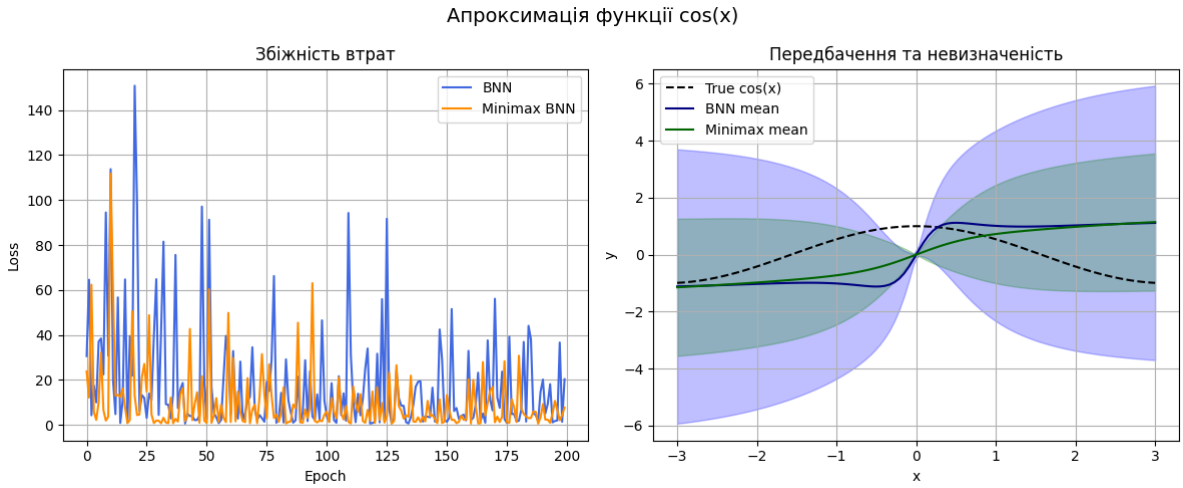

Рис. 2. Алгоритм функції cos(x)

У випадку апроксимації cos(x), спостерігається схожа тенденція:

– класична BNN демонструє значно вищу нестабільність втрат з піками до 150;

– мінімаксна модель стабільніша, з поступовим спадом втрат і меншою дисперсією.

У передбаченнях видно, що:

– класична BNN неадекватно оцінює значення в зоні x>2, де спостерігається значне відхилення від істинної кривої;

– Minimax BNN зберігає правильну фазу та амплітуду навіть за малих значень функції.

– Мінімаксна модель ефективніше справляється з формами функцій, де є тонкі симетрії і зміни кривизни, краще контролює межі області визначення і має нижчу невизначеність по всій довжині.

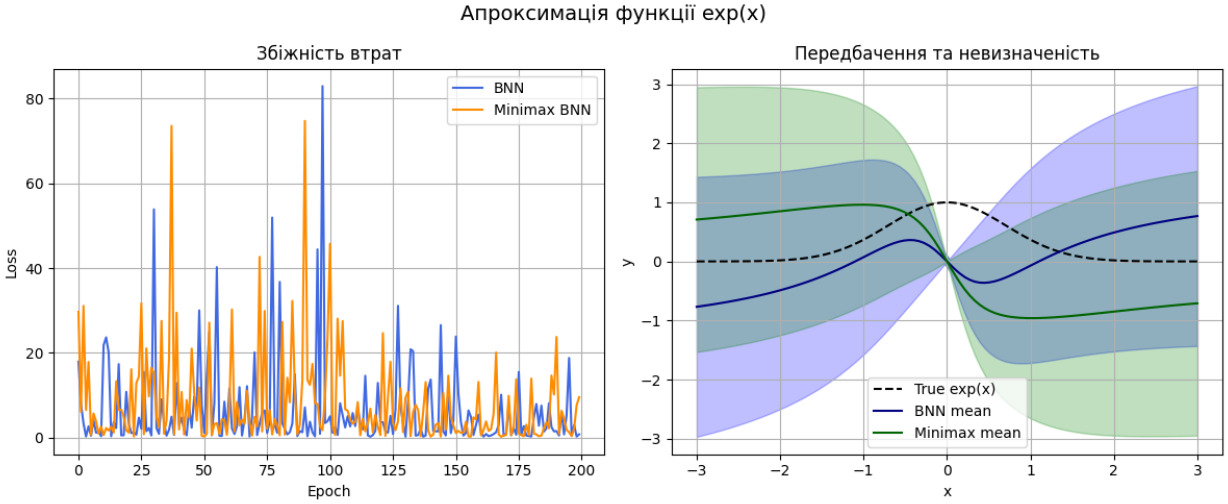

Рис. 3. Апроксимація функції exp(x)

Цей експеримент оцінює здатність моделей відтворити гладку, немонотонну, симетричну функцію з різким спадом.

Класична BNN у лівій частині графіка демонструє коливання втрат з піками до 80;

Мінімаксна BNN – знову стабільніша, хоча тут коливання ближчі, ніж у sin/cos.

Передбачення:

Класична BNN завищує значення на краях, через що отримує розмиту форму експоненти;

мінімаксна модель зберігає падіння до нуля за x→±3, що відповідає фізичній природі функції.

Висновки

Minimax BNN краще зберігає спадну форму та пригнічує шумові викривлення на границях, чим перевершує класичну модель у задачах із "схлопуванням" значень.

Мінімаксна байєсівська нейромережа:

- тренується стабільніше;
- має нижчу втрату на кожному кроці;
- демонструє меншу дисперсію передбачень;
- краще узагальнює в умовах невизначеності або викривлень;
- стійкіша до шуму та до нестачі даних на краях.

Мінімаксна стратегія є ефективним доповненням до BNN, особливо в застосуваннях, де модель має працювати з ризикованими або непередбачуваними даними.

СПИСОК ЛІТЕРАТУРИ

1. Мокін Б. І., Мокін О. Б., Мокін В. Б. Машинне навчання та інтелектуальний аналіз даних. Вінниця: ВНТУ, 2024. 264 с. https://pdf.lib.vntu.edu.ua/books/2024/Mokin_2024_263.pdf.

2. Attacking Bayes: On the Adversarial Robustness of Bayesian Neural Networks / Y. Feng et al. 2024. *arXiv:2404.19640*. https://arxiv.org/abs/2404.19640

3. Adversarially Robust Fault Zone Predictionin Smart Grids with Bayesian Neural Networks / E. Efatinasab et al. *Research Gate*. 2024. P. (99): 1-1. https://www.researchgate.net/publication/383598032_Adversarially_Robust_Fault_Zone_Prediction_in_Smart_Grids_with_Bayesian_Neural_Networks.

4. On the Robustness of Bayesian Neural Networks to Adversarial Attacks / L. Bortolussi et al. *arXiv:2207.06154*. 2022. https://arxiv.org/abs/2207.06154.

5. Bayesian Inference with Certifiable Adversarial Robustness / M. Wicker et al. *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics. arXiv:2102.05289*. 2021. https://arxiv.org/abs/2102.05289.

6. Hong J., Kuruoglu E. E. Minimax Bayesian Neural Networks. *MDPI Entropy*. 2025. №27(4). P. 340. https://doi.org/10.3390/e27040340.

7. Making Substitute Models More Bayesian Can Enhance Transferability of Adversarial Examples / Q. Li et al. *arXiv:2302.05086v3*. 2023. https://arxiv.org/pdf/2302.05086

8. Minimax-Bayes Reinforcement Learning / T. K. Buening et al. *Proceedings of the 40th International Conference on Machine Learning. arXiv:2302.10831*. 2023. https://arxiv.org/abs/2302.10831.

9. Cheng S., Gokhale T., Yang Y. Adversarial Bayesian Augmentation for Single-Source Domain Generalization. *ICCV 2023*. 2023. P. 11400–11410. https://openaccess.thecvf.com/content/ICCV2023/papers/Cheng_Adversarial_Bayesian_Augmentation_for_Single-Source_Domain_Generalization_ICCV_2023_paper.pdf.

10. Yu T., Zhang Z., Wang C. Flat Seeking Bayesian Neural Networks. *arXivpreprint arXiv:2302.02713*. 2023. https://arxiv.org/abs/2302.02713.

Стаття надійшла до редакції 20.02.2026.
Стаття пройшла рецензування 25.03.2026.
Стаття опублікована 31.03.2026.

Кривошея Михайло Ігорович – аспірант кафедри Автоматизації та інтелектуальних інформаційних технологій, факультет інтелектуальних інформаційних технологій та автоматизації, ORCID: 0009-0000-4365-5937 e-mail: mishakryvoshea@gmail.com.

Квєтний Роман Наумович – д-р техн. наук, професор кафедри Автоматизації та інтелектуальних інформаційних технологій, факультет інтелектуальних інформаційних технологій та автоматизації, ORCID: 0000-0002-9192-9258, e-mail: rkvetny@vntu.edu.ua.

Вінницький національний технічний університет.