

МІНІМАКСНА АПРОКСИМАЦІЯ ЯК МЕТОД ПОДОЛАННЯ ФЕНОМЕНУ ПОДВІЙНОГО СПУСКУ

Вінницький національний технічний університет

Анотація

У цій роботі розглядаються методи мінімаксної апроксимації як інструменту для подолання феномену подвійного спуску в машинному навчанні. Подвійний спуск є важливою проблемою, що виникає, коли моделі починають демонструвати покращення результатів після перенавчання. Це явище спостерігається при навчанні складних моделей з великою кількістю параметрів, коли зростає ризик перенавчання. Мінімаксна апроксимація дозволяє зменшити максимальні втрати, що виникають в процесі навчання, і стабілізує результати моделі навіть в умовах, коли дані можуть бути шумними або непередбачуваними. У роботі описуються основні стратегії мінімізації максимальних втрат, оптимізації складності моделі, контролю перенавчання та впровадження регуляризації. Особливу увагу приділено визначенню оптимальної точки другого спуску, що дозволяє підвищити узагальнюючі властивості моделей та знизити їх схильність до перенавчання. Дослідження показують, що використання мінімаксної апроксимації може суттєво підвищити ефективність навчання моделей, особливо в умовах, коли дані мають високий рівень шуму або коли моделі мають велику кількість параметрів. Отримані результати можуть бути корисними для науковців і практиків у сфері машинного навчання, оскільки вони демонструють нові підходи до покращення загальної продуктивності моделей і підвищення їх надійності.

Ключові слова: мінімаксна апроксимація, подвійний спуск, машинне навчання, перенавчання, оптимізація моделі, регуляризація.

Вступ

У сучасному машинному навчанні, яке активно застосовується у багатьох галузях, таких як фінанси, охорона здоров'я, комп'ютерне бачення та обробка природної мови, виникає безліч викликів, пов'язаних зі складністю моделей та обсягами даних. Однією з найважливіших проблем є феномен подвійного спуску, що проявляється, коли моделі, які спочатку демонструють покращення результатів, починають зазнавати зниження продуктивності внаслідок перенавчання. Це явище є особливо актуальним при навчанні складних моделей, що мають велику кількість параметрів, оскільки в таких випадках ризик перенавчання суттєво зростає.

Мінімаксна апроксимація, як метод, намагається подолати цю проблему, зосереджуючись на мінімізації максимальних втрат, що дозволяє стабілізувати результати моделі навіть в умовах високого рівня шуму в даних. У цій роботі ми пропонуємо докладний аналіз застосування мінімаксної апроксимації в контексті подвійного спуску, а також описуємо основні стратегії, які допоможуть у покращенні загальної продуктивності моделей машинного навчання.

Метою даної роботи є дослідження методів мінімаксної апроксимації, їх впливу на подолання подвійного спуску, а також пропозиція нових підходів для покращення ефективності навчання моделей у різних умовах.

Актуальність

Актуальність цього дослідження зумовлена необхідністю розробки нових методів, що дозволяють ефективно справлятися з проблемою подвійного спуску в машинному навчанні. У сучасних реаліях, коли обсяги даних зростають, а вимоги до моделей стають все більш складними, важливо мати інструменти, які можуть зменшити ризики перенавчання і покращити узагальнюючі властивості моделей.

Сучасні підходи, які використовують складні методи, такі як BPMN для моделювання бізнес-процесів, можуть виявитися недоступними для багатьох практиків. Це створює бар'єри для їх впровадження у реальні проекти, оскільки високий рівень технічних знань не завжди є у бізнес-аналітиків та фахівців. Водночас, зростає потреба у методах, які забезпечують точність і деталізацію моделювання, але при цьому є простими і зрозумілими.

Запропонований метод мінімаксної апроксимації може значно знизити поріг входження для фахівців, що працюють у галузі машинного навчання. Використання цього підходу дозволяє ефективно покращувати продуктивність моделей, особливо в умовах, коли дані мають високий рівень шуму або при роботі з великими наборами даних. Це особливо важливо в проектах, де співпрацюють міждисциплінарні команди, адже ефективна комунікація між технічними фахівцями та експертами предметної області є критично важливою. Запропонований метод, заснований на мінімаксній апроксимації, вирішує ці проблеми, забезпечуючи баланс між простотою і точністю моделювання.

Результати дослідження.

Мінімізація максимальних втрат: мінімаксна апроксимація здатна суттєво зменшити максимальні помилки, що виникають у процесі навчання моделей. Цей підхід дозволяє фокусуватися на гірших

сценаріях, що може привести до більш стабільних і надійних результатів. Важливо зазначити, що при традиційних методах навчання моделі часто орієнтуються на середні показники, не враховуючи крайні випадки, які можуть суттєво вплинути на загальну продуктивність. Застосування мінімаксної апроксимації дозволяє враховувати ці крайні випадки, знижуючи ризик отримання високих помилок у реальних умовах.

Оптимізація складності моделі: Завдяки мінімаксній апроксимації вдалося знайти оптимальний рівень складності моделей. Дослідження показують, що поступове збільшення кількості параметрів у моделі, з одночасним контролем максимальних втрат, дозволяє уникнути перенавчання та підвищити здатність моделі узагальнювати нові дані. Це важливо в умовах реальних завдань, де доступні обмежені обсяги навчальних даних. Таким чином, завдяки мінімаксному підходу, можливо досягти балансу між складністю моделі та її узагальнюючими властивостями.

Контроль перенавчання: застосування мінімаксу дозволило зменшити ризики перенавчання в моделях з великою кількістю параметрів. Перенавчання часто відбувається, коли модель занадто добре підходить до навчальних даних, але погано узагальнює на нових даних. Мінімаксна апроксимація забезпечує механізми, які контролюють цю поведінку, забезпечуючи стабільність моделі навіть у складних ситуаціях. Це підтверджується стабільними результатами на тестових даних, що свідчить про підвищення узагальнюючих можливостей моделі.

Регуляризація і мінімакс: інтеграція регуляризаційних технік (L1, L2) з мінімаксним підходом допомогла знизити втрати на тестових даних та забезпечити кращу робастність моделі. Регуляризація дозволяє зменшити вплив надмірної складності, додаючи штрафи за складність моделі до функції втрат. Поєднання цих двох підходів дозволяє досягти оптимальних результатів, забезпечуючи при цьому стійкість моделі до варіацій у даних.

Оптимальна точка другого спуску: мінімаксна апроксимація допомогла точно визначити момент другого спуску, коли модель після перенавчання знову почала покращувати результати. Це дозволило підвищити ефективність моделей, особливо при роботі з великими нейронними мережами.

У порівнянні з традиційними підходами, такими як рання зупинка (early stopping) і dropout, мінімаксна апроксимація показала кращі результати на тестових даних. Це стало можливим завдяки тому, що мінімаксна апроксимація фокусується на мінімізації максимальних втрат

Висновки

Мінімаксна апроксимація виступає потужним інструментом для подолання феномену подвійного спуску в машинному навчанні. Цей метод дозволяє ефективно контролювати складність моделі та зменшувати максимальні втрати, що виникають під час навчання, особливо в найгірших сценаріях. Вона допомагає знизити ризики перенавчання, оптимізувати кількість параметрів і забезпечити стабільність результатів на тестових даних. Поєднання мінімаксної апроксимації з методами регуляризації, такими як L1 або L2, дозволяє досягти кращих результатів за умов високої складності моделей. Проте, мінімаксна апроксимація не є універсальним рішенням і повинна використовуватися разом з іншими підходами, такими як крос-валідація та адаптивне навчання, для досягнення максимальних результатів у різних задачах машинного навчання.

Список літератури

- [1] Belkin, M., Hsu, D., Ma, S., & Mandal, S. (2019). Reconciling modern machine learning practice and the classical bias-variance trade-off. *Proceedings of the National Academy of Sciences*,
- [2] Дубовой В. М., Кветний Р. Н., Михальов О. І., Усов А. В., Моделювання та оптимізація систем : підручник. Вінниця : ПП «ТД«Едельвейс», 2017. 804 с.
- [3] Lafon, M., & Thomas, A. *Understanding the Double Descent Phenomenon in Deep Learning*. *arXiv preprint*, arXiv:2403.10459, 2024 – URL: <https://arxiv.org/abs/2403.10459>

References

- [1] Belkin, M., Hsu, D., Ma, S., & Mandal, S. (2019). Reconciling modern machine learning practice and the classical bias-variance trade-off. *Proceedings of the National Academy of Sciences*.
- [2] Dubovoy, V. M., Kvyetnyy, R. N., Mykhailov, O. I., & Usov, A. V. (2017). *Modeling and Optimization of Systems: A Textbook*. Vinnytsia: ТД «Edelweiss», 804 p.
- [3] Lafon, M., & Thomas, A. (2024). *Understanding the Double Descent Phenomenon in Deep Learning*. arXiv preprint, arXiv:2403.10459. Retrieved from <https://arxiv.org/abs/2403.10459>

Кривошея Михайло Ігорович — студент кафедри автоматизації та інтелектуальних інформаційних технологій, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця. e-mail: mishakryvoshea@gmail.com

Науковий керівник: Квєтний Роман Наумович – доктор технічних наук, професор, професор кафедри автоматизації та інтелектуальних інформаційних технологій, Вінницький національний технічний університет, м.Вінниця, e-mail: rkvetny@vntu.edu.ua

M. I. Kryvosheia

MINIMAX APPROXIMATION AS A METHOD FOR OVERCOMING THE DOUBLE DESCENT PHENOMENON

Vinnitsia National Technical University, Vinnitsia

***Abstract.** This paper explores minimax approximation methods as a tool for overcoming the double descent phenomenon in machine learning. Double descent is an important issue that arises when models begin to demonstrate improved results after overfitting. This phenomenon is observed when training complex models with a large number of parameters, which increases the risk of overfitting. Minimax approximation allows for the reduction of maximum losses that occur during training and stabilizes model results even in situations where data may be noisy or unpredictable. The paper describes the main strategies for minimizing maximum losses, optimizing model complexity, controlling overfitting, and implementing regularization. Special attention is given to identifying the optimal point of the second descent, which enhances the generalization capabilities of models and reduces their susceptibility to overfitting. The research shows that using minimax approximation can significantly improve the effectiveness of model training, especially in conditions where data has a high level of noise or when models have a large number of parameters. The obtained results may be beneficial for researchers and practitioners in the field of machine learning, as they demonstrate new approaches to improving overall model performance and enhancing their reliability.*

Keywords: minimax approximation, double descent, machine learning, overfitting, model optimization, regularization.

Kryvosheia Mykhailo I. — Department of intelligent information technologies and automation, Vinnitsia National Technical University, Vinnitsia, email : mishakryvoshea@gmail.com

Supervisor: Kvyetnyy Roman N.– Doctor of Technical Sciences, Professor, Professor of the Department of Automation and Intelligent Information Technologies, Vinnitsia National Technical University, Vinnitsia, e-mail: rkvetny@vntu.edu.ua