

СУЧАСНІ МЕТОДИ ТА МОДЕЛІ МАШИННОГО І ГЛИБОКОГО НАВЧАННЯ З ПАРАМЕТРАМИ

Шовкун П. О.

Науковий керівник – доц. Адамов О. С.

Харківський національний університет радіоелектроніки (61166, Харків,
пр. Науки, 14, каф. АПОТ, тел. (057) 702-13-26)

e-mail: pavlo.shovkuna@nure.ua

This work explores contemporary methods and models in machine and deep learning with parameters. The study encompasses a comprehensive review of parameterized models, their applications, and the impact of varying parameters on model performance. Additionally, the paper investigates the latest trends and innovations in the rapidly evolving landscape of machine and deep learning, providing insights into the optimization and enhancement of model capabilities through parameter tuning.

Вступ. В сучасному інформаційному суспільстві, де дані стають все більш об'ємними та складними, розвиток методів та моделей машинного та глибокого навчання набуває особливого значення. Метою дослідження є останні досягнення в галузі машинного та глибокого навчання, зосереджуючись на параметрах, які визначають ефективність цих методів.

Зміст дослідження: Ансамблеві методи є ефективним підходом до машинного навчання, де кілька моделей об'єднуються для отримання більш точних та стабільних прогнозів. Такий підхід розглядається у контексті розв'язання різноманітних завдань. Види ансамблевих методів: В ансамблевих методах виділяються різні стратегії, такі як випадковий ліс, градієнтний бустінг та багатокласові класифікатори. Вони взаємодіють з даними та вдосконалюються в процесі навчання. Ансамблеві методи широко використовуються в різних галузях, включаючи фінанси (прогнозування ринків), медицину (діагностика захворювань) та промисловість (якість контролю продукції). В контексті рекомендаційних систем, персоналізовані моделі грають ключову роль у виведенні індивідуалізованих пропозицій, незалежно від того, чи йдеться про товари, контент або послуги. Цей підхід сприяє збільшенню задоволення користувачів та підвищенню їхньої взаємодії з платформою.

Основи навчання з підкріпленням: Навчання з підкріпленням ґрунтується на взаємодії агента з оточенням та отриманні винагороди за правильні дії. Застосування у обробці природних мов: Навчання з підкріпленням виявляється корисним для розв'язання завдань обробки природних мов, таких як вдосконалення систем автоматичного перекладу, створення більш ефективних голосових помічників та покращення синтезу

тексту в різноманітних контекстах. Архітектури нейронних мереж визначають спосіб, яким модель взаємодіє з даними. Згорткові нейронні мережі (CNN) ефективні для обробки зображень, в той час як рекурентні нейронні мережі (RNN) використовуються для послідовностей, таких як текст або мовлення. Вплив архітектур на точність: CNN, завдяки своїм конволюційним шарам, добре впорається із просторовою ієрархією зображень. RNN дозволяє враховувати контекст у послідовностях. Трансформери дозволяють взаємодіяти з контекстом на всій довжині послідовності.

Параметри моделей, які визначають ефективність: Глибина нейронної мережі (Depth): Глибші мережі може вивчати складніші патерни у вхідних даних. Однак занадто велика глибина може призвести до перенавчання. Кількість нейронів у шарах (Width): Збільшення кількості нейронів може збільшити потужність моделі, але також може призвести до перенавчання, особливо при обмежених ресурсах. Швидкість навчання (Learning Rate): Визначає, наскільки значущим чином змінюються ваги під час навчання. Занадто висока або низька швидкість навчання може впливати на збіжність моделі.

Функція активації: Впливає на вираження та нелінійність моделі. Популярні функції включають ReLU, Sigmoid та Tanh, кожна з яких підходить для різних сценаріїв. Регуляризація: Допомогає уникнути перенавчання шляхом додаткових обмежень на ваги моделі. L1 та L2 регуляризація є загальними методами. Батч-розмір (Batch Size): Кількість прикладів даних, які використовуються для одного оновлення ваг моделі. Впливає на швидкість та стабільність навчання.

Функція втрат: Визначає, як модель вимірює точність своїх прогнозів порівняно з фактичними даними. Залежно від задачі може використовуватися різна функція втрат. Проблеми у навчанні глибоких моделей: Обмежені ресурси можуть включати нестачу даних для навчання, низьку обчислювальну потужність та обмежену кількість пам'яті. Ці обмеження можуть призвести до перенавчання, погіршення якості моделі та подовження часу навчання. Рішення та стратегії: Для подолання цих проблем можна використовувати попереднє навчання, яке дозволяє моделі використовувати заздалегідь навчені ваги. Квантизація ваг та компресія моделей можуть допомогти зменшити об'єм пам'яті та прискорити інференс.

Основи оптимізації глибоких нейронних мереж: Стандартні методи оптимізації, такі як стохастичний градієнтний спуск, використовуються для навчання глибоких моделей. Вони оновлюють ваги моделі, враховуючи градієнти від втрат. Глибокі нейронні мережі є основою глибокого навчання та використовуються для різноманітних завдань, включаючи класифікацію, регресію, та визначення зображень та мовлення. Принцип їхнього навчання включає проходження вхідних даних через

кілька шарів нейронів з вагами, які оновлюються під час зворотного поширення помилки (backpropagation) за допомогою градієнтного спуску. SVM використовується для завдань класифікації та регресії. Принцип навчання полягає в пошуку гіперплощини, яка найкраще розділяє дані. SVM спирається на опорні вектори, які представляють найбільш важливі екземпляри в навчальному наборі. Застосовується для класифікації та виявлення аномалій. Випадковий ліс є ансамблевим методом, який використовує кілька дерев рішень для прийняття рішення. Кожне дерево навчається на випадковому підмножині даних та голосує за результат. Принцип навчання включає створення та об'єднання багатьох дерев для зменшення перенавчання та покращення стабільності моделі. Використовується для класифікації та виявлення паттернів. Логістична регресія використовується для бінарної класифікації. Під час навчання ваги оновлюються за допомогою градієнтного спуску з метою максимізації ймовірності правильного класифікаційного результату. Застосовується для виявлення вірусів та загроз.

Градієнтний бустінг (Gradient Boosting) – ансамбль, де нові моделі навчаються для виправлення помилок попередніх. Використовується для покращення точності та зменшення перенавчання. Метод головних компонент (Principal Component Analysis) використовується для зменшення розмірності даних та виділення основних компонент. Може служити для виявлення аномалій. K-Means кластеризація використовується для групування подібних об'єктів. Використовується для виявлення аномалій та невідомих паттернів. Автокодування – нейронні мережі для створення компактних представлень даних. Використовуються для виявлення вірусів та аномальних змін у вхідних даних. Алгоритми виявлення аномалій (Isolation Forest, One-Class SVM) – спеціалізовані алгоритми для виявлення незвичайних паттернів, що можуть вказувати на наявність вірусів або загроз.

Висновки. Розвиток методів та моделей машинного та глибокого навчання є актуальним завданням у сучасному інформаційному середовищі. Висвітлені у тезі аспекти дозволяють краще зрозуміти сучасні тенденції та виклики у цій області, параметри ефективності, та також визначити можливі шляхи подальших досліджень, та їх використання для виявлення зловмисних програм.

Список використаних джерел:

1. Machine Learning: A brief introduction to Random Forest. URL:<https://www.einsteinmed.edu/uploadedfiles/centers/ictr/new/intro-to-random-forest.pdf>.
2. Ensemble deep learning: A review: M.A. Ganaie, Minghui Hu, A.K. Malik, M.Tanveer, P.N.Suganthan. URL: <https://arxiv.org/abs/2104.02395>.
3. Understanding LSTM Networks. URL: <https://colah.github.io/posts/2015-08-Understanding-LSTMs>.