

ВИКОРИСТАННЯ МЕТОДУ СВІТЛЯЧКІВ ДЛЯ СИНТЕЗУ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ

Синтез штучної нейронної мережі (ШНМ) – це процес проектування, структурної оптимізації та параметричного налаштування (навчання ШНМ) для подальшого використання з метою виконання конкретного завдання. Тобто, це включає визначення архітектури ШНМ, вибір відповідних функцій активації та оптимізацію параметрів, щоб мінімізувати функцію помилок або втрат у даному наборі даних.

Структурний синтез фокусується на визначенні архітектури ШНМ (структури або топології). Структурна оптимізація спрямована на пошук архітектури, яка забезпечує баланс між складністю моделі та здатністю до узагальнення, гарантуючи, що ШНМ зможе ефективно вивчати основні закономірності даних без перенавчання.

Параметричний синтез або ж процес навчання передбачає налаштування параметрів нейронної мережі, зокрема ваг і зсувів, пов'язаних з кожним з'єднанням між нейронами. Цей процес синтезу спрямований на мінімізацію функції помилок або втрат у навчальному наборі даних шляхом налаштування параметрів для покращення продуктивності моделі. Методи параметричної оптимізації зазвичай використовують ітераційні методи, такі як варіанти градієнтного спуску (наприклад, стохастичний градієнтний спуск, зворотне поширення помилки) або методи метаевристичної оптимізації (наприклад, генетичні алгоритми, оптимізація рою частинок) для пошуку оптимальних значень параметрів, що мінімізують цільову функцію.

Методи метаевристичної оптимізації, що засновані на природні процесах, такі як генетичні алгоритми (ГА), методи ройового інтелекту (РІ), зокрема і метод світлячків (МС), широко використовуються для синтезу ШНМ завдяки ряду переваг, які вони надають:

- глобальна оптимізація: ці методи дозволяють ефективно досліджувати великий простір пошуку і часто дозволяють знаходити глобальні оптимальні або близькі до них рішення. Це має вирішальне значення для синтезу ШНМ, адже важливо синтезувати оптимальну структуру, для якої в подальшому налаштувати параметри;

- високі адаптивність та гнучкість незалежно від прикладної задачі. Такі методи можуть працювати з різними типами змінних, обмежень та завдань, що робить їх придатними для оптимізації різних архітектур та параметрів ШНМ;

–стійкість до зашумлених або неповних вхідних даних або ж недосконалих цільових функцій і можуть працювати з нелінійними, неопуклими та переривчастими ландшафтами оптимізації;

–адаптивність – такі методи часто включають механізми адаптації та самоналаштування, що дозволяє їм динамічно реагувати на зміни в ландшафті оптимізації або обмеження, пов'язані з проблемою.

МС – це метод з групи PI, який імітує блимання світлячків. Він широко використовується для вирішення проблем оптимізації, а також може бути адаптований для синтезу ШНМ. Ось загальний підхід до використання методу МС для нейросинтезу, що пропонується авторами:

–завдання метапараметрів популяції нейромереж (кількість шарів, кількість нейронів в шарі, функції активації);

–кодування: визначення схеми подання для кодування простору рішень;

–визначення цільової функції, яка кількісно оцінює продуктивність ШНМ;

–ініціалізація популяції ШНМ, що імітує агентів (окремих особин – світлячків);

–оцінка придатності кожного агента в популяції, використовуючи цільову функцію;

–рух: агенти коригують своє положення відповідно до привабливості, яка визначається їх яскравістю та відстанню до інших агентів. У контексті нейросинтезу налаштовується архітектура та параметри кожної ШНМ на основі її продуктивності та продуктивності сусідніх рішень;

–оновлення стану агентів, залежно від їх руху та відповідно налаштування параметру їх яскравості;

–перевірка критерію зупинки (наприклад, збіжність цільової функції, досягнення заданого порогу продуктивності, тощо);

–відбір найбільш ефективної ШНМ з кінцевої сукупності в якості рішення задачі синтезу;

–тестування продуктивність вибраної ШНМ на окремому наборі даних (частині набору) для перевірки, щоб переконатися в її здатності до узагальнення.

Дотримуючись цих кроків, МС може ефективно синтезувати ШНМ з оптимізованою архітектурою та параметрами для заданого завдання. Однак варто зазначити, що продуктивність синтезованої нейронної мережі значною мірою залежить від якості цільової функції, схеми подання та початкових параметрів методу. Тому для успішного застосування необхідні ретельне проєктування та налаштування.