

УДК 628.9.06

О. С. Бельський, студент гр. ПН-41мп, к.т.н., Ківа І.Л.
КПІ ім. Ігоря Сікорського

РЕКУРЕНТНІ НЕЙРОННІ МЕРЕЖІ ЯК МЕТОД ПРОГНОЗУВАННЯ ФІЗИЧНИХ ПРОЦЕСІВ

Анотація. В даному документі представлені результати аналітичного дослідження РНМ, дослідження нового класу – резервуарних РНМ, що дало змогу виділити основні динамічні режими роботи РНМ та визначити найбільш перспективні напрямки в розвитку методів навчання РНМ.

Ключові слова: рекурентні нейронні мережі, прогнозування, методи навчання, нейронні мережі, метод динамічного резервуару

ВСТУП

Рекурентні нейронні мережі (РНМ) - нейронні мережі (НМ) із зворотними зв'язками, завдяки яким вони мають власну динаміку, що відрізняє їх від статичних НМ, наприклад, нейронних мереж прямого поширення. Стан РНМ в поточний момент часу визначається не тільки вхідними сигналами, але і попереднім станом. За рахунок цього РНМ володіють нескінченною пам'яттю, що відрізняє їх від НМ із затримками, які запам'ятовують інформацію про входах тільки за обмежений проміжок часу. З огляду на те що РНМ є динамічною системою, доцільно провести аналіз можливості залучення результатів дослідження динамічних систем в суміжних областях науки, таких, як фізика, теорія нелінійних динамічних систем, теорія хаосу і інших, для синтезу РНМ і вирішення питань їх навчання. Необхідно відмітити, що при цьому терміни «нейрон» і «НМ» можуть взагалі не використовуватися - замість цього використовуються такі поняття, як пов'язана решітка, структура, середовище, кінцевий автомат з пам'яттю, клітинний автомат. Результатом такого аналізу стане виявлення перспективних напрямків розвитку РНМ в частині їх синтезу і навчання, що актуально з огляду на великих потенційних можливостей РНМ, зазначених у багатьох наукових роботах. [1]

СТРУКТУРА РНМ

Структура РНМ є її основною характеристикою, від якої безпосередньо залежать всі її властивості. В залежності від типів зв'язків виділяють повнозв'язні та локальні РНМ. У повнозв'язних РНМ кожен нейрон пов'язаний з кожним. Це призводить до підпорядкування всіх нейронів мережі якомусь одному глобальному режиму. У РНМ з локальними зв'язками кожен нейрон пов'язаний лише з нейронами в деякій околиці навколо нього. Такі зв'язки формуються детермінованим або випадковим чином. Локальна зв'язність дає нейронам більше незалежності, в різних областях РНМ можуть переважають різні динамічні режими. Існують приклади РНМ зі змінною структурою: в ході роботи нейронної мережі у неї можуть з'являтися (зникати) нейрони і зв'язку. При невеликій кількості елементів структура РНМ являє собою грати в одно-, дво- і максимум тривимірному дискретно просторі. Стан РНМ характеризується сукупністю значень виходів її нейронів. Для математичного опису таких РНМ використовуються системи керування. Якщо кількість елементів досить велика, то від дискретного простору станів переходять до

безперервного, і структура РНМ являє собою безперервне середовище з розподіленими параметрами. Динаміка таких нескінченновимірних в просторі систем описується спеціальними хвильовими рівняннями - рівняннями в локальних похідних або інтегрально диференціальними рівняннями. Для аналізу структури РНМ може застосовуватися підхід, заснований на теорії графів. Він полягає в обчисленні різних показників, таких, як матриця досяжності, шляхи, цикли, маршрути, кластерний індекс та ін. За допомогою цих показників можна приблизно оцінити взаємодію елементів в РНМ. [2][3]

Однією з головних властивостей РНМ як динамічної системи є стійкість. РНМ може бути стійка (нестійка) в малому і в великому. Стійкість у великому гарантує стійкість у всьому просторі станів, а стійкість в малому - лише в певних точках. Теореми Ляпунова дають аналітичні критерії стійкості нелінійних систем: для стійкості у великому необхідно підібрати спеціальну функцію, що задовольняє умовам другий теореми Ляпунова; для стійкості в малому необхідно провести лінеаризацію системи в точці рівноваги і перевірити в ній стійкість.

Залежно від стійкості розрізняють три основних типи динаміки РНМ: стійка динаміка, при якій за кінцеве час РНМ сходиться до стійкого стану рівноваги (в даному режимі при зміні вхідного сигналу можливі переходи системи з одного стану в інший); коливальна динаміка: стан РНМ описує замкнуту циклічну траєкторію - граничний цикл; з однієї боку можна вважати, що граничний цикл кодує певну інформацію, а з іншого - що відбуваються коливання, при цьому інформація передається в фазових співвідношеннях; нестійка динаміка: траєкторії РНМ при як завгодно близьких початкових станах розходяться з часом; якщо зростання статків системи ніяк не обмежений, то воно йде в нескінченність, в іншому випадку встановлюється хаотична динаміка, при якій стан системи спонтанно змінюється в часі всередині певній галузі фазового простору. [2]

Різні динамічні режими РНМ можуть розглядатися як окремі етапи вирішення конкретного завдання. Наприклад, хаотичний режим відповідає початковому пошуку рішень - вибору між різними гіпотезами, коливальний режим відображає процеси перемикавання уваги з одного рішення на інше, а стійкий режим відповідає знайденому рішенням.

Динаміка РНМ залежить як від її структури, так і від властивостей складових її елементів.

Теорія біфуркацій, катастроф та інші розділи теорії нелінійної динаміки дозволяють аналітично досліджувати динаміку досить простих систем (2-3 порядку). Це рівень нейрона (синапсу). На рівні РНМ ці теорії не працюють. Тому в загальному випадку для РНМ довільної структури єдиний можливий спосіб аналізу динаміки - чисельне моделювання та обчислення різних показників динаміки. [2]

АНАЛІЗ ДИНАМІКИ РНМ

Динаміку РНМ можна представити у вигляді суми двох складових - усталеною і перехідною. Перехідна складова згодом згасає, після чого

залишається тільки постійна складова. Для розпізнавання образів використовуються перехідні і сталі складові (режими): в перехідному режимі після подачі вхідного образу його динаміка накладається на динаміку РНМ - по «знімку» результуючої динаміки через деякий час можна розпізнати вхідний образ; в даному випадку використовується властивість короткочасної пам'яті - обліку передісторії зміни вхідного сигналу (стимулу); в сталому режимі після подачі вхідного образу через якийсь час система виходить на аттрактор (фіксований стан, граничний цикл або хаотичний аттрактор) - з цього аттрактору можна розпізнати вхідний образ.

РНМ з стійкими режимами є стійкими в великому. Після зміни їх стану або після подачі вхідного сигналу через деякий час вони сходяться до стійким точкам рівноваги - фіксованим аттрактором.

РНМ з коливальними режимами або осциляторний НМ (ОНМ) називаються НМ, стан яких змінюється в часі по періодичному або близькій до періодичного закону. Іноді їх ще називають НМ з аттракторами типу «граничний цикл». Тип зв'язків залежить від моделі осцилятора. Зв'язки можуть мати властивості саморегуляції (внутрішня динаміка) і пластичності (Самонавчання). [3]

НАВЧАННЯ РНМ

Для навчання РНМ допустимий вибір між наступними групами:

- використання відомих взаємозв'язків між параметрами значення РНМ та її динамікою;
- використання алгоритмів навчання з викладачем на основі методів оптимізації за типом алгоритму зворотного розповсюдження; приклад - алгоритми зворотного розповсюдження в часі (Зворотне поширення через час, ВРТТ), періодичне навчання в режимі реального часу (повторне навчання в режимі реального часу, RTRL), повторне зворотне розмноження (Періодичне зворотне поширення, RBP) [4], алгоритми з використанням фільтрації Каллмана; через те, параметри на кожному кроці встановлюються однаково, функціонал оптимізації формується як сума помилок у деяких інтервалах часу, виникають проблеми з часом роботи, зі сходженням методів та в обчислювальних затратах;
- з використанням алгоритму навчання з викладачем, коли розглядає РНМ як частково повторюваний - сигнали від зворотних зв'язків розглядаються як окремі вхідні сигнали - контекстні нейрони, як результат, навчальне завдання спрощується і зведений до звичайних алгоритмів навчання НМ подача вперед;
- використання алгоритмів без навчання (синаптична пластичність імпульсних НС) [5]
- метод резервуарного навчання, при якому ваги встановлюються випадковими значеннями, при цьому функції навчання проходять через спеціальний пристрій для класифікації динаміки РНМ.
- Модернізація в останньому варіанті навчання полягає в тому, щоб використовувати РНМ, як динамічний резервуар з потужними

обчислювальними можливостями. При дії неперервного сигналу на такий резервуар він потрапляє до певного динамічного процесу або стану, також резервуар формується таким чином щоб для схожих сигналів чи процесів цей стан також був схожий. Виходи резервуару зазвичай під'єднують до пристроїв виведення, які в залежності від стану резервуару класифікують, прогнозують чи кластеризують. [6]

МОДЕРНІЗАЦІЯ СИСТЕМ ВИМІРЮВАННЯ

Вимірювання по суті це один неперервний сигнал на виході. При умові що ведеться вимірювання на основі інерціального процесу який може прямувати до неприпустимих точок екстремуму (наприклад вимірювання температури при якому вона досягає критичного максимуму) можна спрогнозувати появу цих точок за допомогою РНМ з досить високою точністю. Це дає змогу побачити можливу проблему та запобігти їй запустивши процес протидії. При підвищенні температури – запустити процес її зменшення використовуючи вихід РНМ для підтримання сталих значень.

ВИСНОВОК

Використання РНМ дає широкий спектр можливостей в прогнозування фізичних явищ використовуючи рінманітні методи регуляції. РНМ являються досить гнучкими в навчанні що робить їх універсальним інструментом у вирішенні задачі контролю фізичних процесів, так як на виході ми отримуємо динамічний стан який можна класифікувати, кластеризувати чи спрогнозувати.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Dasgupta B., Siegelmann H., Sontag E.D. On the Complexity of Training Neural Networks with Continuous Activation Functions / IEEE Transactions on Neural Networks. – 1995. – Vol. 6. – № 6. – P. 1490–1504.
- [2] Dunin-Barkowski W.L., Osovets N.B. HebbHopfield neural networks based on one-dimensional sets of neuron states / Neural Processing Letters. – 1995. – Vol. 2. – № 5. – P. 28–31.
- [3] Feng J., Brown D. Fixed-point attractor analysis for a class of neurodynamics / Neural Computation. – 1998. – Vol. 10. – P. 189–213.
- [4] Maass W., Natschläger T., Markram H. Real-time computing without stable states: a new framework for neural computations based on perturbations / Neural Computation. – 2002. – Vol. 11. – P. 2531–2560.
- [5] Steil J.J. Backpropagation-decorrelation: recurrent learning with O(N) complexity / Proc. IJCNN. – 2004. – Vol. 1. – P. 843–848.
- [6] Jaeger H. The «echo state» approach to analyzing and training recurrent neural networks: GMD report 148.– St. Augustin: German National

Наук. керівник – к.т.н., Ківа І. Л.