

Д.Ю. Зубенко, В.В. Ліньков

Харківський національний університет міського господарства імені О.М. Бекетова, Україна

ПОВТОРНЕ ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ З ДЕКІЛЬКОМА ЕКЗЕМПЛЯРАМИ В ЗАДАЧАХ ЕКСПЛУАТАЦІЇ І РЕМОНТУ ЕЛЕКТРОТРАНСПОРТУ

Останнім часом нейронні мережі і багаторазові навчання (MIL) є привабливими темами в дослідницьких областях, пов'язаних зі штучним інтелектом. Глибокі нейронні мережі домоглися великих успіхів в контрольованих проблеми навчання, а MIL як типовий слабо контрольованій метод навчання є ефективний для багатьох додатків в області комп'ютерного зору, біометрії, і т. д.

Ключові слова: експлуатація електротранспорту, багаторівневе навчання, нейронні мережі, глибоке навчання, наскрізне навчання

Постановка проблеми. Аналіз останніх досліджень і публікацій

Спочатку було запропоновано кілька примірників (MIL) для прогнозування активності відмов [1]. Тепер він широко застосовується до багатьох доменів і є важливою проблемою в комп'ютерному навчанні. Багато мультимедійних даних мають структуру множиливого примірника (MI); наприклад, текстова стаття містить кілька абзаців, зображення можна розділити на кілька локальних областей, а дані експресії генів містять кілька генів. MIL корисний для обробки і розуміння даних MI.

Вивчення декількох екземплярів - це вид слабого контрольованого навчання (WSL). Кожен зразок виконаний у вигляді мічених даних, що складаються з широкої різноманітності примірників, пов'язаних з функціями введення. Мета MIL в подвійній задачі - підготувати класифікатор для прогнозування ярликів тестових пакетів, заснований на припущенні, що є позитивний пакет та містить хоча б один позитивний екземпляр, тоді як данні негативні, якщо він складається тільки з негативних екземплярів. Таким чином, суть MIL полягає в тому, щоб мати справу з двозначністю ярликів предметів, особливо в позитивних даних, які мають безліч випадків з різними позиціями.

Для вирішення проблеми MIL було запропоновано багато алгоритмів. Згідно з опитуванням Amores [2], алгоритми MIL знаходяться в трьох варіантах: парадигма простору-зразка, парадигма простору даних і парадигма вкладеного простору. Парадигма простору примірників вивчає класифікатор примірників і виконує класифікацію пакетів шляхом агрегування відповідей класифікатора рівня екземпляра.

Парадигма простору Bag-space використовує відносини з даними і обробляє сумку в цілому; зокрема, він обчислює відстань / подобу даних до кляйстеру; то найближчий сусід або байесовський класифікатор виконує класифікацію задач, засновану на відстанях / подібності. Парадигма вкладеного простору вбудовує пакет в просторове простір на основі словника, щоб отримати компактне представлення для даними, наприклад, векторне подання; то класичні класифікатори можуть застосовуватися для вирішення проблеми класифікації пакетів. Глибокі нейронні мережі були застосовані для вирішення багатьох проблем машинного навчання. Для контрольованого навчання існує кілька видів нейронних мереж.

Мета статті

Вирішення задачі повторного використання нейронних мереж з декількома екземплярами в задачах експлуатації і ремонту електротранспорту.

Виклад основного матеріалу

Розглянемо визначення MIL. З огляду на набір пакетів $X = \{ X_1, X_2, \dots, X_N \}$ і особливості примірника i th пакет $X_i = \{ x_{i1}, x_{i2}, \dots, x_{im_i} \}$, $x_{ij} \in R^{d \times 1}$, де N чи m_i позначають кількість даних і кількість примірників в кляйстері X_i відповідно. припускати $Y_i \in \{0, 1\}$ чи $u_{ij} \in \{0, 1\}$ ярлик пакета X_i і примірник x_{ij} окремо, де 1 означає позитивний, а 0 означає негативний. У MIL під час навчання даються тільки деталі з даними, і є два обмеження MIL: • Якщо пакет X_i є негативний, то всі випадки в X_i буде негативними,

За умови, якщо $Y_i = 0$, значить $u_{ij} = 0$;

• якщо пакет X_i є позитивний, то хоча б один екземпляр в X_i буде позитивним, тобто, якщо $Y_i = 1$, then $\sum_{j=1}^{m_i} y_{ij} \geq 1$.

Найскладніша проблема в MIL полягає в тому, що мітка примірника не вказана. У MINN існує дві стратегії: перша полягає в тому, щоб вивести мітку примірника в мережі, тобто поставити ймовірності примірника позитивного в якості прихованого рівня в мережі; другий - вивчити уявлення кляйстера в мережі і безпосередньо виконувати класифікацію даних без обчислення ймовірності примірника. Перша стратегія була вивчена в [8,9,15]. Друга стратегія знову пропонується в цій статті. У наступних підрозділах ми дамо опису MINN.

Розглянемо налаштування одного пакета X_i з декількома екземплярами x_{ij} який проходить через MINN. MINN складається з L шарів, кожна з яких містить нелінійне перетворення H^l , де L індексує шар. $H^l(\cdot)$ може бути складовою частиною таких операцій, як внутрішній продукт (або повне з'єднання), випрямлені лінійні одиниці (ReLU) [2-9] або пропонований пул MIL. Позначимо висновок 1-го шару примірника x_{ij} як x_{ij}^1 .

Спочатку ми розглядаємо традиційні нейронні мережі з декількома екземплярами [8,9,14], які називаються mi-Net. Як показано, кожен екземпляр в кляйстері спочатку подається на кілька повністю пов'язаних (FC) шарів з функцією активації (в цій статті ми використовуємо чотири шари FC з активацією ReLU [9]). Ми отримуємо функцію примірника, позначену як x_{ij}^{L-2} в $(L-2)$ -й рівня і ймовірності примірника, позначеної як p_{ij}^{L-1} . p_{ij}^{L-1} є скаляром в діапазоні від $[0, 1]$ і виводиться з x_{ij}^{L-2} індивідуально. В останньому шарі є пул MIL Pooling Layer, який приймає ймовірності примірника в якості входних і виведених сумарною ймовірності, позначених як $P^L(X_i)$.

Ці перші шари $L-2$ можуть вивчити деякі більш семантичні особливості примірника в порівнянні з оригінальним x_{ij} (Вищий рівень, що відповідає більш високим семантичним ознаками). Після вивчення цих особливостей примірника використовується рівень FC, який має тільки один нейрон з сигмовидною активацією, для прогнозування позитивності примірників.

На відміну від традиційних нейронних мереж, для mi-Net у нас є тільки данні відмов для навчання, але ярлики примірників недоступні. Щоб вирішити цю проблему, ми розглядаємо мітки екземпляра як приховані змінні і виводимо їх під час мережевого навчання. Ми створюємо шар для підсумовування балів примірників в кляйстері. Тут MILL Pooling Layer використовується для

об'єднання цих примірників в кінцеву позитивність пакета.

Метод об'єднання MIL задовольняє обмеженням MIL: якщо кляйстер позитивний, повинен бути хоча б один екземпляр з великою позитивністю. В іншому випадку всі екземпляри в кляйстері повинні мати низьку позитивність. Коли об'єднаний шар інтегрований в нейронну мережу, функція об'єднання повинна бути диференційованою. Типовий пул MIL представлений таким чином, mi-Net можна

Можна сформулювати так:

$$\begin{cases} x_{ij}^l = H^l(x_{ij}^{l-1}), \\ p_i^l = M^L(p_{ij}^{l-1}_{j=1\dots m_i}). \end{cases} \quad (1)$$

У mi-Net, формулювання останніх двох шарів: $P_i^l = M^l(p_{ij}^{l-1}_{j=1\dots m_i})$. P_i^l це ймовірність пакета даних та M^l є MIL оператором. Таким чином, нейрони від другого до останнього шару (тобто the $(L-1)$ го рівня) представляють ймовірності примірників.

Ми пропонуємо серію нових нейронних мереж з декількома екземплярами, які не покладаються на висновок ймовірності примірника. Мережі безпосередньо вивчають уявлення кляйстера і виробляють кращу точність класифікації даних. Ці методи відносяться до категорії вбудованих просторових алгоритмів MIL, визначених в огляді [2]. Дотримуючись стилю іменування в [10], ми називаємо цю мережу як MI-Net.

Показана проста MI-Net з трьома повністю пов'язаними шарами і одним пулом MIL. Зміна структури мережі призводить до того, що мережа фокусується на поданні даних навчання, а не на прогнозі ймовірності примірника. Незалежно від того, скільки входних примірників є, MILL Pooling Layer об'єднує їх в один вектор функцій як уявлення пакета. Нарешті, шар FC з тільки однієї нейронної і сигмоїдною активацією приймає подання даними в якості входного сигналу і прогнозує ймовірність даними. Цей простий MI-Net сформульований як:

$$\begin{cases} x_{ij}^l = H^l(x_{ij}^{l-1}), \\ x_i^l = M^L(x_{ij}^{l-1}_{j=1\dots m_i}). \end{cases} \quad (2)$$

Різниця між MI-Net або mi-Net. По-перше, ми можемо порівняти 1 і 2, щоб знайти різницю між mi-Net і MI-Net. У mi-Net є кілька вузлів, що представляють екземпляри. В MI-Net немає лічильників та примірників; замість цього він містить вектор функції з даними. З точки зору вивчення особливостей, mi-Net фокусується на поданні навчального примірника; в той час як MI-Net вивчає подання примірників і уявлення сум. У нас є явна мотивація проектування MI-Net. Оскільки

mi-Net пророкує рахунок примірника на основі окремого примірника, і оцінка сум залежить від кількості примірників, класифікація пакетів не буде виконана, якщо класифікатори примірників помиляються. Наша мотивація MI-Net полягає в тому, щоб отримати більш багате уявлення.

Агрегуючи всі можливості примірника, а потім давати більш надійне передбачення, засноване на уявленні підсумовування. Коли навчання з використанням декількох екземплярів - це слабо контрольована проблема навчання; або явно, або неявно викликаючи ймовірність примірника, він завжди ризикує зробити помилку. Однак в MI-Net основна увага приділяється завданню пакетів; таким чином, слабо контрольована проблема MIL стає повністю контрольованою меншою ціною робіт по ремонту. Саме з цієї причини MI-Net має тенденцію давати кращу точність класифікації даних, як показано в експериментах. Однак в MI-Net є обмеження; він не може дати можливість примірника. У додатках, що вимагають ймовірності примірника,

Натхненні глибоко контрольованими мережами (DSN) [13], ми додаємо глибокі спостереження в MI-Net. Тобто кожен середній рівень FC, який може вивчати функції примірника, супроводжується шаром пулу MIL і рівнем FC для прогнозування сумарного балу. Під час навчання додається на кожен рівень. Крім того, під час тестування ми обчислюємо середній бал для кожного рівня. MI-Net з глибоким спостереженням формулюється як:

$$\begin{cases} x_{ij}^l = H^l(x_{ij}^{l-1}), \\ x_i^l = M^L(x_{ij|j=1\dots m_i}^{l-1}), k \in \{1,2,3\}, \end{cases} \quad (3)$$

де індекс k в $X^{l,k}$, k і означає, що ми вивчаємо безліч функцій сум з усіх рівнів можливостей примірника за допомогою пулу MIL. MI-Net з глибоким спостереженням може використовувати кілька ієрархій, щоб отримати кращу точність класифікації даних. Його можна інтерпретувати з двох точок зору: У процесі навчання функція примірника в нижніх шарах може отримувати більш ретельний контроль; і (2) при тестуванні ми можемо усереднити ймовірність множинних сум, щоб отримати більш міцні дані. У цій статті ми встановлюємо ваги різних рівнів однаково.

Останнім часом глибоке залишкове навчання було запропоновано в [14] і продемонструвало вражаюче поліпшення розпізнавання зображень за рахунок використання дуже глибоких нейронних мереж. Ми вивчаємо залишкові з'єднання в MI-Net, як показано MI-Net з залишковими сполуками формулюється як:

$$\begin{cases} x_{ij}^l = H^l(x_{ij}^{l-1}), \\ x_i^l = M^L(x_{ij|j=1\dots m_i}^{l-1}) \\ x_i^l = M^L(x_{ij|j=1\dots m_i}^{l-1}) + X^{l-1}, l > 1, \end{cases} \quad (4)$$

На відміну від вихідного залишкового навчання в [14], в якому вивчаються залишки репрезентацій з використанням згортки, нормалізації партії і ReLU, ми вивчаємо залишки уявлення сум через повністю зв'язані шари, об'єднання ReLU і MIL. В кінці мережі уявлення фінальної клястера пов'язано з етикеткою даними через шар FC з однієї нейронної і сигмоїдною активацією.

Висновки

У цьому дослідженні ми переглядаємо проблему наскрізного навчання MINN і пропонуємо серію нових MINN з найсучаснішою продуктивністю. На відміну від існуючих MINN, наш метод фокусується на навчанні уявлення сумарного рівня замість оцінки мітки рівня на рівні примірника. Експерименти показують, що наші мережі на рівні сум показують чудові результати по кільком контрольних показників MIL в порівнянні з мережами на рівні примірників. Більш того, ми інтегруємо в наші мережі найпопулярніші глибокі навчальні трюки (глибокий контроль і залишкові з'єднання), які можуть підвищити продуктивність. Більш того, наш метод займає близько 0,0003 с для тестування (вперед) і 0,0008 с для навчання в клястері, що дуже ефективно.

References

1. Dietterich, T.G., Lathrop, R.H., Lozano-Pérez, T. (1997) Solving the multiple instance problem with axis-parallel rectangles, *Artif. Intell.*, 89 (1), 31–71.
2. Amores, J. (2013) Multiple instance classification: review, taxonomy and comparative study, *Artif. Intell.*, 201, 81–105.
3. Hinton, G., Osindero, S., Teh, Y.W. (2006) A fast learning algorithm for deep belief nets, *Neural Comput.*, 18 (7), 1527–1554.
4. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998) Gradient-based learning applied to document recognition, *Proc. IEEE*, 86 (11), 2278–2324.
5. Krizhevsky, A., Sutskever, I., Hinton, G.E. (2012) Imagenet classification with deep convolutional neural networks, in: NIPS, 1097–1105.
6. Williams, R.J., Zipser, D. (1989) A learning algorithm for continually running fully recurrent neural networks, *Neural Comput.*, 1 (2), 270–280.
7. Hochreiter, S., Schmidhuber, J. (1997) Long short-term memory, *Neural Comput.*, 9 (8), 1735–1780.
8. Ramon, J., De Raedt, L. (2000) Multi instance neural networks, in: Proceedings of the ICML-2000 Workshop on Attribute-Value and Relational Learning, 53–60.
9. Zhou, Z.-H., Zhang, M.-L. (2002) Neural networks for multi-instance learning, in: Proceedings of the International

Conference on Intelligent Information Technology, Beijing, China, 455–459.

10. Andrews , S. , Tsochantaridis , I. , Hofmann , T. (2002) Support vector machines for multiple-instance learning, in: NIPS, 561–568 .

11. Srivastava , N. , Hinton , G. , Krizhevsky , A. , Sutskever , I., Salakhutdinov , R. (2014) Dropout: a simple way to prevent neural networks from overfitting, *JMLR*, 15 (1), 1929–1958.

12. Nair , V. , Hinton , G. (2010) Rectified linear units improve restricted Boltzmann machines, in: ICML, 807–814.

13. Lee , C.Y. , Xie , S. , Gallagher , P. , Zhang , Z. , Tu , Z. (2015) Deeply-supervised nets, in: AIS-TATS, 562–570.

14. He , K. , Zhang , X. , Ren , S. , Sun , J. (2015) Deep residual learning for image recognition, arXiv:1512.03385.

15. Zhang , M.-L. , Zhou , Z.-H. (2004) Improve multi-instance neural networks through feature selection, *Neural Process. Lett.*, 19 (1), 1–10.

Рецензент: д-р техн. наук, проф. М.Ф. Смирний, Харківський національний університет міського господарства імені О.М. Бекетова, Україна.

Автор: ЗУБЕНКО Денис Юрійович

доцент, к.т.н.

Харківський національний університет міського господарства імені О.М. Бекетова

E-mail – Denis04@ukr.net

ID ORCID: <http://orcid.org/0000-0002-6736-7849>

Автори: ЛІНЬКОВ Віктор Васильович

доцент, к.т.н.

Харківський національний університет міського господарства імені О.М. Бекетова

E-mail – Denis04@ukr.net

ID ORCID: <http://orcid.org/0000-0003-0246-0513>

RE-USE OF NEURAL NETWORKS WITH A DIFFERENT PLANTS IN THE PROBLEMS OF OPERATION AND REPAIR OF ELECTRIC TRANSPORT

D. Zubenko, V. Linkov

O.M. Beketov National University of Urban Economy in Kharkiv, Ukraine

Recently, neural networks and multiple learning (MIL) are attractive topics in research areas related to artificial intelligence. Deep neural networks have achieved great success in controlled learning problems, and MIL as a typical poorly controlled learning method is effective for many applications in computer vision, biometrics, natural language processing, etc. In this article, we review several neural networks with multiple instances (MINN), which neural networks seek to solve MIL problems. MINNs perform MILs in the end, which take bags with different numbers of instances as input and directly output the tags of the bags. All parameters in MINN can be optimized by back propagation. In addition to revising old MINNs, we offer a new type of MINN for exploring bag representations, which differs from existing MINNs that focus on the evaluation of an instance label. In addition, recent tricks developed in deep learning have been studied in MINN; we find that deep supervision is effective for a better understanding of bag views. In experiments, the proposed MINNs achieve the most advanced or competitive performance on several MIL tests. Moreover, for testing and learning it is very fast, for example, it takes only 0.003 s to predict the bag and a few seconds to learn on the MIL datasets on a moderate processor.

Initially, several instances (MILs) were proposed to predict bounce activity [1]. Now it is widely applied to many domains and is an important problem in computer training. Many multimedia data have a multiplier (MI) structure; For example, a text article contains several paragraphs, the image can be divided into several local areas, and gene expression data contains several genes. MIL is useful for processing and understanding MI data.

Studying multiple instances is a type of weak controlled learning (WSL). Each sample is executed in the form of labeled data, which consist of a wide variety of instances associated with the functions of input. The purpose of MIL in the dual task is to prepare a classifier for prediction of test packet labels, based on the assumption that there is a positive packet and contains at least one positive instance, while the data is negative if it consists only of negative instances.

Keywords: electric transport operation, multi-level learning, neural networks, deep learning, end-to-end learning.