

Д.Ю. Зубенко, О.В. Донець, В.В. Лінков

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

ПОДВІЙНЕ МАТРИЧНЕ РЕГУЛЮВАННЯ ТА ФАКТОРИЗАЦІЯ З ГЛИБОКИМИ НЕЙРОННИМИ МЕРЕЖАМИ ДЛЯ СИСТЕМ ВИЗНАЧЕННЯ ТЕПЛООВОГО СТАНУ ЕЛЕКТРОДВИГУНА

У багатофакторних системах з використанням текстової та графічної інформації в матричній факторизації для полегшення проблеми роздільної обробки даних. Останнім часом в деяких роботах дослідження нейронних мереж, щоб глибше зрозуміти зміст текстових та графічних елементів і досягти ефективності шляхом створення більш точних моделей розпізнавання елементів. Однак, залишається відкрите питання про те, як ефективно використовувати графічні дані з тепловізорів при матричній факторизації. У цій роботі ми запропонували подвійно-регуляризовану матричну факторизацію з глибокими нейронними мережами (DRMF) для вирішення цієї проблеми.

Ключові слова: електротранспорт, рухомий склад, електродвигуни, тепловий стан електродвигуна, факторизація матриць, глибокі нейронні мережі.

Постановка проблеми. Аналіз останніх досліджень і публікацій

Системи, які спрямовані на пом'якшення негативного впливу інформаційної перевантаженості шляхом фільтрації та надання користувачам найбільш привабливих та релевантних елементів (таких як відео з дисплея тепловізора при визначенні теплового стану електродвигуна), і таким чином вирішили проблему для великих даних [1]. Запропоновано різні методи побудови систем рекомендацій за останні десять років з різних систем [2], наприклад, спільна фільтрація [3] та методами мережі [4], [7]; [9]; [11], б). Серед них, матрична факторизація на основі співпраці фільтрування (CF) є домінуючим методом завдяки його успішному застосуванню в системах [5-6]. У традиційних методах СФ використовується лише матриця зворотного зв'язку, яка містить явні (наприклад, оцінки) або неявні зворотні зв'язки для навчання та прогнозування ресурсу електродвигуна. Як правило, матриця зворотного зв'язку є рідкою, а це означає, що більшість користувачів стикаються з кількома елементами. Виходячи з цього був представлений в Proc. BigData Congress [10]. Проте це подання було суттєво розширене, використовуючи нову модель глибокої нейронної мережі та додаючи нові експериментальні вкладення у порівнянні з публікацією конференції.

Багато дослідників запропонували використовувати змістовну інформацію для пом'якшення проблеми роздільної здатності даних у [12-13]. Зокрема, репрезентативні робочі праці полягають у вилученні семантичної інформації

текстового змісту за допомогою тематичної моделі та моделі глибокої нейронної мережі [14-15].

Мета статті

Вирішення задачі подвійного матричного регулювання та факторизація з глибокими нейронними мережами для систем визначення теплового стану електродвигуна

Виклад основного матеріалу

У матричній факторизації, користувачі та елементи обидва представлені як латентні вектори в загальному латентному k -вимірному просторі, \mathbb{R}^k , де користувач i представлений як латентний вектор $u_i \in \mathbb{R}^k$ і предмет j представлений як латентний вектор $v_j \in \mathbb{R}^k$. Прогноз, чи є користувач i буде подобатися предмету j дається внутрішнім твором між їх латентними уявленнями, $r_{ij} = u_i^T v_j$. Для того, щоб використовувати факторизацію matrix для спільної фільтрації, приховані представлення користувачів та предметів потрібно вивчити, надаючи оцінену матрицю. Загальний підхід полягає в тому, щоб звести до мінімуму нормалізовану втрату квадратів помилок щодо коефіцієнтів користувача $= (u_i)_{i=1}^I$ and item factors $V = (v_j)_{j=1}^J$.

$$\min_{u,v} \sum_{i,j} (r_{i,j} - u_i^T v_j)^2 + \lambda_u \|u_i\|^2 + \lambda_v \|v_j\|^2 \quad (1)$$

де λ_u і λ_v є параметрами регуляризації. $r_{ij} > 0$ якщо користувач i номінальна позиція j , і $r_{ij} = 0$ інакше. Матрична факторизація може бути узагальнена як імовірнісна модель шляхом розміщення нульового середнього сферичного гаусового попереднього на

латентні фактори користувачів і предметів [1-5], який може бути далі описаний як наступний генеративний процес, Для кожного користувача i , намалювати векторний латентний користувач $u_i \sim N(0, \lambda_u^{-1} E_k)$;

Для кожної позиції j , зобразити елемент латентного вектор $v_j \sim N(0, \lambda_v^{-1} E_k)$;

Для кожної пари користувача-елемента (i, j) , набрати данні $r_{ij} \sim N(u_i^T v_j, c_{ij}^{-1})$;

де c_{ij} служить параметром надійності для r_{ij} (Hu et al., 2008). Якщо c_{ij} є забагато, r_{ij} вірне рівняння. Загалом, $c_{ij} = a$ if $r_{ij} > 0$ і $c_{ij} = b$ if $r_{ij} = 0$, a та b є налаштування параметрів satisfying $ab \geq 0$. Таким чином, імовірнісна матрична факторизація (PMF) може мати справу з незаперечними рейтингами. PMF можна легко розширити, щоб включити упередження для різних користувачів, елементів і контекстів, щоб отримати більш надійні моделі латентних факторів [6-12].

Традиційні періодичні нейронні мережі, що виникають із проблеми довготривалої залежності в моделюванні послідовності. Gated RNN (GRNN) вводять рекурентні одиниці (наприклад, LSTM і GRU), щоб полегшити кожну одиницю, щоб запам'ятати наявність певної функції в потоці введення для довгої серії кроків, а також автоматично закриті контекстні шляхи, які обходять декілька тимчасових кроків до уникнути зникаючих градієнтів [13-15]. Тут ми приймаємо замкнуту рекурентну одиницю (GRU).

Формально, враховуючи послідовність S , оновлення повторюваного прихованого стану в GRU-RNN узагальнюється, як показано нижче [3-9]. Для кожного j -го підрозділу GRU активації h^j як шаг t є:

$$h_t^j = (1 - z_t^j) h_{t-1}^j + z_t^j \tilde{h}_t^j, \quad (2)$$

де z_t^j є воротами оновлення і обчислено за:

$$z_t^j = \delta(W_z s_t + M_z h_{t-1}^j) \quad (3)$$

Де функція δ може брати *sigmoid* або *tanh*. Активация параметра h_t^j обчислюється шляхом:

$$h_t^j = \tanh(W_s s_t + M(d_t \circ h_{t-1}^j)) \quad (4)$$

де d_t є собою набір перезавантажувальних воріт і означає елементний помножувач. Ворота скидання d_t обчислюється як:

$$d_t^j = \delta(W_d s_t + M_d h_{t-1}^j) \quad (5)$$

З огляду на набір навчальних програм, ми хочемо визначити оцінку Maximum a posteriori (MAP) U, V, W^* , тому ми можемо використовувати U та V для прогнозування відсутніх записів у R та використовувати прогнози для надання рекомендацій.

Для вивчення параметрів DRMF ми розробляємо алгоритм EM-стилю, подібний до [15]. Максимізація заднього рівня еквівалентна максимізації повної логарифмічної вірогідності формули 5, тому ми можемо вивести цільову функцію наступним чином,

$$L(U, V, W_1, W_2) = - \sum_i \sum_j \frac{c_{ij}}{2} (r_{ij} - u_i^T r_{ij} v_j)^2 - \frac{\lambda_u}{2} \sum_i (u_i - \phi_i)^T (u_i - \phi_i) - \frac{\lambda_v}{2} \sum_j (v_j - \theta_j)^T (v_j - \theta_j) - \frac{\lambda_u}{2} \left(\sum_k W_1^k + \sum_k W_2^k \right).$$

Спочатку виправляємо значення θ та ϕ (а саме, виправити параметри W_1, W_2 в перетворенні $dnn(W_1, X), dnn(W_2, Y)$) вирішити та V . Для u_i, v_j , максимізація виконується аналогічно матричному факторизації (Hu et al., 2008). Взяття градієнта L з відношення до u_i, v_j і встановлення його до нуля допомагає знайти u_i, v_j . Вирішення відповідних рівнянь приведе до оновлення правил наступним чином,

$$u_i \leftarrow (VC_i V^T + \lambda_u E_k)^{-1} (VC_i R_i + \lambda_u \phi_i) \\ v_j \leftarrow (UC_j U^T + \lambda_v E_k)^{-1} (UC_j R_j + \lambda_v \theta_j) \quad (6)$$

де C_i є діагональна матриця з c_{ij} як діагональний елемент з $R_i = (r_{ij})_{j=1}^J$ для використання i . Для предметів j , C_j and R_j аналогічно визначено. c_{ij} є показником надійності роботи r_{ij} . Ми використовуємо таку ж стратегію, як зазначено в [15] to set c_{ij} : $c_{ij} = a$ if $r_{ij} > 0$ та $c_{ij} = b$ if $r_{ij} = 0$. Комбінація a та b є різницею для завдань прогнозування надійності та рекомендацій провідних.

Для малих наборів даних фіксація значень векторних слів у модулях DNN також забезпечує кращу точність прогнозів w.r.t DRMF.

Висновки

Останні дослідження виявили, що точність рекомендацій зменшується в довгостроковій перспективі, якщо всі користувачі повністю покладаються на рекомендації щодо вибору елементів. Оскільки цей процес, як правило, робить деякі елементи системи більш надійними.

Ми використовуємо лише глибинну нейронну мережу для прогнозування ресурсу рухомого складу, яка дозволяє підвищити надійність експлуатації транспортної техніки.

References

1. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
2. Bobadilla, J., Ortega, F., Hernando, A., & Gutierrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, 46, 109–132.
3. Cao, D., He, X., Nie, L., Wei, X., Hu, X., Wu, S., & Chua, T. (2017a). Cross-platform app recommendation by jointly modeling ratings and texts. *ACM Trans. Inf. Syst.*, 35, 37:1–37:27.
4. Gao, J., & Zhou, T. (2017). Evaluating user reputation in online rating systems via an iterative group-based ranking

- method. *Physica A Statistical Mechanics Its Applications*, 473, 546–560.
5. He, R., & McAuley, J. (2016). Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *Proceedings of the 25th International Conference on World Wide Web (WWW'16)* (pp. 507–517). ACM.
6. Johnson, R., & Zhang, T. (2014). Effective use of word order for text categorization with convolutional neural networks. *arXiv preprint arXiv:1412.1058*, .
7. Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, 42, 30–37.
8. Lu, L., & Zhou, T. (2011). Statistical mechanics and its applications - link prediction in complex networks: A survey. *Physica A-statistical Mechanics Its Applications*, 390, 1150–1170.
9. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Proceedings of the 27th Annual Conference on Neural Information Processing Systems (NIPS'13)* (pp. 3111–3119).
10. Salakhutdinov, R., & Mnih, A. (2007). Probabilistic matrix factorization. In *Proceedings of the Twenty-First Annual Conference on Neural Information Processing Systems (NIPS'07)* (pp. 1257–1264).
11. Tang, D., Qin, B., & Liu, T. (2015b). Document modeling with gated re-current neural network for sentiment classification. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP'15)* (pp. 1422–1432).
12. Wang, C., & Blei, D. M. (2011). Collaborative topic modeling for recommending scientific articles. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'11)* (pp. 448–456).
13. Zhang, F., Yuan, N. J., Lian, D., Xie, X., & Ma, W.-Y. (2016). Collaborative knowledge base embedding for recommender systems. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'16)* (pp. 353–362). ACM.
14. Zhao, D.-D., Zeng, A., Shang, M.-S., & Gao, J. (2013). Long-term effects of recommendation on the evolution of online systems. *Chinese Physics Letters*, 30, 118901.
15. Zhou, T., Kuscsik, Z., Liu, J.-G., Medo, M., Wakeling, J. R., & Zhang, Y.-C. (2010). Solving the apparent diversity-accuracy dilemma of recommender systems. *The Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 107, 4511–4515.

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: ДОНЕЦЬ Олександр Вадимович
доцент, к.т.н.
Харківський національний університет міського господарства імені О.М. Бекетова
E-mail – Denis04@ukr.net
ID ORCID: <http://orcid.org/0000-0002-6265-8508>

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/0000-0003-0246-05>

DOUBLE MATRIX REGULATION AND FACTORIZATION WITH DEEP NEURAL NETWORKS FOR ELECTRODIVIGUE HEATING SYSTEMS

D. Zubenko, O. Donets, V. Linkov

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

In multifactorial systems using textual and graphical information in matrix factorization to facilitate the problem of separate data processing. Recently, in some studies, the study of neural networks to understand the content of text and graphic elements more deeply and to achieve efficacy by creating more accurate patterns of recognition of elements. However, the open question remains about how to effectively use graphic data from the thermal imager in matrix factorization. In this paper, we proposed a double-regularized matrix factorization with deep neural networks (DRMF) to solve this problem. DRMF applies a multilayered neural network model by stacking a convolutional neural network and a secured repetitive neural network to create independent distributed views of user content and objects. Then representations serve to regularize the generation of hidden models for both users and for elements of matrix factorization. So the proposed new model of the neural network works better than a model with a single convergent neural network. Systems that are aimed at mitigating the negative effects of information overload by filtration and providing users with the most attractive and relevant elements (such as video from the display of the thermal imager in determining the thermal state of the electric motor), thus solving the problem for large data. Different methods of building recommendations for the last ten years from different systems, for example, joint filtration and network methods. Among them, matrix factorization on the basis of co-filtration (CF) is the dominant method due to its successful application in systems. In traditional SF methods, only a feedback matrix that contains explicit (eg, estimates) or implicit feedback for learning and prediction of an electric motor resource is used. Typically, the feedback matrix is liquid, which means that most users are faced with several elements. Proceeding from this, was presented in Proc. BigDataCongress. However, this presentation was substantially expanded, using the new model of the deep neural network and adding new experimental contributions compared to the publication of the conference. Many researchers have suggested using content to mitigate the problem of data resolution in. In particular, representative work jobs consist of extracting semantic information from text content using a thematic model and deep neural network model.

Keywords: electric transport, rolling stock, electric motors, thermal state of the electric motor, factorization of matrices, deep neural networks.