様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　２ 年　 ６ 月 １５ 日現在

機関番号：１２１０２
研究種目：若手研究(B)
研究期間：2017～2019
課題番号：１７Ｋ１２７８４

研究課題名（和文）A Novel Interactive Information Retrieval System Using Deep Neural Network

研究課題名（英文）A Novel Interactive Information Retrieval System Using Deep Neural Network

研究代表者
于　海涛（Yu, Haitao）

筑波大学・図書館情報メディア系・助教

研究者番号：３０７５１０５２

交付決定額（研究期間全体）：（直接経費）　　3,000,000円

研究成果の概要（和文）：ディープニューラルネットワークに基づいて2つのモデルを開発した。1つ目は、最適輸送理論に基づいて新しいランキング学習モデルを提案し、WSDMというトップ国際会議に発表した。2つ目は、ユーザーの検索行動を分析するための新しいクリックモデルを提案し、CHIIRという国際会議に発表した。さらに、私はPT-Rankingというオープンソースプロジェクトを公開した。従来のランク付け学習パッケージへの補完性が高く、様々な分野の学習ランク付けモデルを評価・開発するための便利なツールを提供することで、様々なバックグラウンドを持つ研究者を支援できるようになることを期待する。

研究成果の学術的意義や社会的意義
The proposed learning-to-rank models shed new light on how to solving the ranking problem. The released open-source project makes it reasonable to envision that PT-Ranking will lower the technical barrier and provide a convenient open-source platform for examining ranking models in different fields.

研究成果の概要（英文）：This year I developed two models based on deep neural networks. The first one is a novel learning-to-rank model based on the theory of optimal transport, which is published at the 12th international conference on web search and data mining. The second one is a new click model for decoding users' search behaviour, which is published at the 2019 conference on human information interaction & retrieval. Based on a series of experiments using benchmark datasets, the experiments have demonstrated their effectiveness for information retrieval. Moreover, I released the open-source project titled as PT-Ranking. PT-Ranking is highly complementary to the previous packages for learning-to-rank. I envision that PT-Ranking will lower the technical barrier and provide a convenient open-source platform for evaluating and developing learning-to-rank models in different fields, and thus facilitate researchers from various backgrounds.

研究分野：情報検索

キーワード：learning-to-rank　click modelling　optimal transport

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

Information Retrieval (IR) has been playing an important role in our daily information access. Significant research efforts have been made on different aspects of IR. However, accurately and efficiently providing desired information to users is far from being resolved. Facing the richest types of information needs and the biggest amount of data than ever, most of the time, the returned search results remain rather poor. Recently, the deep learning techniques have attracted great attention as a breakthrough and yielded the state-of-the-art performance in many fields. However, its application in improving interactive information retrieval has not been explored yet.

２．研究の目的

Inspired by the newly developed embedding techniques, queries, documents and users are represented as low-dimensional real-valued vectors in this research. Besides the flexibility of encoding them in the same space, a far more compelling advantage is that all the representations are automatically learnt, and many patterns and linguistic regularities can be encoded. The deep neural network models, e.g., Recurrent Neural Networks (RNN), have shown promising performance in modeling complex low-dimensional sequences. In this proposal a novel deep neural network model is devised to discover intricate dependencies among factors during the interactive search process and provide adaptive search results. Overall, this project aims to develop a novel interactive information retrieval system, which is capable of “automatic representation learning and dynamic ranking”.

The success of this project will lead to the following impact on the IR field: (1) A number of effective learning-to-rank models will be developed. The test collections and models generated by user studies will be publicly released. Other researchers can replicate or extend the proposed methods. (2) The theoretical contribution will be the expansion of the territory of the techniques, such as Representation Learning and Deep Learning, for information retrieval. Therefore, other fields like digital library and e-commerce also can benefit from this research.

３．研究の方法

In order to achieve the research objectives, I focus on two research challenges. The first challenge is how to effectively and efficiently learn the vector representations for queries, documents and users? Moreover, the learned vector representations can be used for understanding users’ search behaviors and ranking search results. The second challenge is that how to adaptively rank results using deep neural network corresponding to users’ dynamic search behaviors.

Regarding the first challenge, the proposed method is to learn the vector representations for queries, documents and users based on the search log. Then click models can be further developed to understand users’ search behaviors.

For the second challenge, the key problem is a deep neural network model for interactive ranking. Since the search process is essentially dynamic, an interactive information retrieval system should provide adaptive results. In real search scenarios, a search task usually involves multiple queries, which span one or more sessions. The user needs under the interactive behaviors are essentially associated with one another. This means that the underlying patterns among users’ interactive behaviors must be well captured. Moreover, previously browsed documents indicate the potential interests of users, while currently returned documents must return marginally useful information rather than redundant information. In other words, users’ preferences are essentially dynamic rather than static. The designed deep neural network model must be able to effectively deal with these factors. To address the aforementioned problems, the proposed method is to develop a listwise learning-to-rank method based on deep neural networks. By using deep neural networks as the basis to construct a scoring function, it is easy to incorporating different factors through the technique of embedding.

４．研究成果

The research achievements consist of three parts.

The first part is a novel learning-to-rank model based on deep neural networks. The work has been published at the 12th international conference on web search and data mining (WSDM), which is a top international conference in the field of information retrieval. In particular, I proposed a new listwise loss function for computing the ranking loss based on the theory of optimal transport.

The second part is a novel click model based on deep neural networks. The work has been published at the 2019 conference on human information interaction & retrieval. The key idea is to deploy different weight matrices across different rank positions. Furthermore, I introduced a new method for automatically learning the vector representations for both queries and documents within the same low-dimensional space.

The third part is that I proposed *PT-Ranking*[1], an open-source project based on PyTorch for developing and evaluating learning-to-rank methods using deep neural networks as the basis to construct a scoring function. On one hand, PT-Ranking includes many representative learning-to-rank methods. Besides the traditional optimization framework via empirical risk minimization, adversarial optimization framework is also integrated. Furthermore, PT-Ranking's modular design provides a set of building blocks that users can leverage to develop new ranking models. On the other hand, PT-Ranking supports to compare different learning-to-rank methods based on the widely used datasets (e.g., MSLR-WEB30K, Yahoo!LETOR and Istella LETOR) in terms of different metrics, such as precision, MAP, nDCG, nERR. By randomly masking the ground-truth labels with a specified ratio, PT-Ranking allows to examine to what extent the ratio of unlabeled query-document pairs affects the performance of different learning-to-rank methods. PT-Ranking is highly complementary to the previous open-source projects for learning-to-rank. It is reasonable to envision that PT-Ranking will lower the technical barrier and provide a convenient open-source platform for evaluating and developing learning-to-rank models in different fields, and thus facilitate researchers from various backgrounds.

[1] https://pt-ranking.github.io/

５．主な発表論文等

〔雑誌論文〕　計0件

〔学会発表〕　計2件（うち招待講演　0件／うち国際学会　2件）

１．発表者名 Haitao Yu, Adam Jatowt, Roi Blanco, Joemon Jose, Zhou Ke
２．発表標題 A rank-biased neural network model for click modeling
３．学会等名 The 2019 Conference on Human Information Interaction & Retrieval (CHIIR)（国際学会）
４．発表年 2019年

１．発表者名 Haitao Yu, Adam Jatowt, Hideo Joho, Joemon Jose, Xiao Yang, Long Chen
２．発表標題 WassRank: listwise document ranking using optimal transport theory
３．学会等名 The 12th International Conference on Web Search and Data Mining (WSDM)（国際学会）
４．発表年 2019年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

Project's website: Neural_IIR https://github.com/ii-research/Neural_IIR Learning to Rank in PyTorch https://pt-ranking.github.io

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考