笔杆检测报告单（全文对照）

全文标明引文　　全文对照　　　　打印　　检测说明

标题：Project Proposal-202018010116.docx
作者：Jack
报告编号：BG202311031402439439
提交时间：2023-11-03 14:03:05

去除引用文献复制比：1.9%

去除本人已发表文献复制比：1.9%

单篇最大文字复制比：0.1%

重复字数:**213**　　总字数：**11,254**　　　　检测范围

复制比部分	章
3.298%（213）	Project Proposal-202018010116.docx_第1部分（总6,458字）
0%（0）	Project Proposal-202018010116.docx_第2部分（总4,796字）

（ 无问题部分　复制比部分　引用部分）

Project Proposal-202018010116.docx_第1部分

相似文献列表

文献名	复制比	是否引证
1.The rate of transient beta frequency events predicts behavior across tasks and species Hyeyoung Shin;Robert Law;Shawn Tsutsui;Christopher I Moore;Stephanie R Jones - 《eLife 》- 2017	1.7%(108字)	否
2.The Allele Frequency Spectrum in Genome-Wide Human Variation Data Reveals Signals of Differential Demographic History in Three Large World Populations Marth; G. T. - 《Genetics 》- 2004	1.7%(108字)	否
3.Isolation Modeling and Analysis Based on Mobility Jiang, Jian-Min;Zhu, Huibiao;Li, Qin;Zhao, Yongxin;Hong, Zhong;Zhang, Shi;Gong, Ping - 《ACM Transactions on Software Engineering and Methodology 》- 2019	1.5%(100字)	否
4.Content-aware rate allocation for efficient video streaming via dynamic network utility maximization Mohammad H. Hajiesmailia; b; hajiesmaili@ipm.ir; Ahmad Khonsaria; b; ak@ipm.ir; Ali Sehatia; b; a.sehati@ipm.ir; Mohammad Sadegh Talebib; mstalebi@ipm.ir - 《Journal of Network and Computer Applications 》- 2012	1.4%(91字)	否
5.Content-aware rate allocation for efficient video streaming via dynamic network utility maximization Mohammad H. Hajiesmaili; Ahmad Khonsari; Ali Sehati; Mohammad Sadegh Talebi - 《Journal of Network and Computer Applications 》- 2012	1.4%(90字)	否

查看更多相似文献

全文对照

原文内容	相似内容来源
1. making it easier for product users to find products, services, or informa tion of interest,	*1.*search engine results pages), making it easier for consumers to find you r products, services or information.¦The content you publish - be it on yo ur website, ——百科 -《Houston SEO Copywriting Services》-1500-2020 （是否引证：否）
2. Markov chains are a mathematical model in which the occurrence of an event only depends on the state of the previous event, and is independen t of the earlier state. But it also has some shortcomings,	*1.* This process is Markovian, in that the length of time until the next coal escent event depends¦only on the current state and is independent of the previous states. Due to molecular mutation processes, ——Genetics Marth； G. T.-《The Allele Frequency Spectrum in Genome-Wide Human Variation Data Reveals Signals of Differential Demographic History in Three Large World Populations》-2004 （是否引证：否） *2.*31].¦Let e1, e2 be two events. The notation e1→ e2 is called a dependen cy, which denotes that the¦occurrence of the event e2 depends on the pr evious occurrence of the event e1[47]. Note that a¦dependency actually c orresponds to a transformation dependency in a dependency structure (s ee¦Section 4).¦We give a running example here. The example is a simple y et typical mobile system in which¦a passenger John needs to take a bus i n a road ——ACM Transactions on Software Engineering and Methodology Jiang, Jian-Min； Zhu, Huibiao； Li, Qin； Zhao, Yongxin； Hong, Zhong； Zhang, Shi； Gong,

Ping -《Isolation Modeling and Analysis Based on Mobility》-2019 （是否引证：否）

3. and assuming that any one dosing event depends only on the occurrence of the previous dosing event, given the individual' s probability density function of dosing frequencies,

——Annual Review of Pharmacology and Toxicology Kastrissios and, H.; Blaschke, T. F.-《MEDICATION COMPLIANCE AS A FEATURE IN DRUG DEVELOPMENT》-1997 （是否引证：否）

Project Proposal-202018010116.docx_第2部分

相似文献列表

文献名	复制比	是否引证
1.A Whole Sentence Maximum Entropy Language Model Roni Rosenfeld - 《 》 -	1.8%(87字)	否
2.进语言的知识集成当模特儿：一条随机的森林途径。 Su, Yi. - 《Ph.D.硕士论文 》 -	1.8%(86字)	否
3.Whole-sentence exponential language models: a vehicle for linguistic-statistical integration Rosenfeld, Ronald; Chen, Stanley F.; Zhu, Xiaojin - 《Computer Speech & Language 》 - 2001	1.8%(85字)	否
4.Improving Reliability Performance of Diffusion-based Molecular Communication With Adaptive Threshold Variation Algorithm Peng He;Yuming Mao;Qiang Liu;Kun Yang - 《Int. J. Communication Systems 》 -	1.8%(85字)	否
5.A tutorial on Bayesian models of perception Vincent; Benjamin T. - 《Journal of Mathematical Psychology 》 - 2015	1.8%(85字)	否

查看更多相似文献

全文对照

说明：1.指标是由系统根据《学术论文不端行为的界定标准》自动生成的
2.本报告单仅对您所选择比对资源范围内检测结果负责

笔杆检测报告单（全文标明引文）

全文标明引文　　全文对照　　　　打印　　检测说明

标题：Project Proposal-202018010116.docx
作者：Jack
报告编号：BG202311031402439439
提交时间：2023-11-03 14:03:05

去除引用文献复制比：1.9%

去除本人已发表文献复制比：1.9%

单篇最大文字复制比：0.1%

重复字数:**213**　　总字数：**11,254**　　检测范围

Project Proposal-202018010116.docx_第1部分

原文内容

UNDERGRADUATE PROJECT PROPOSAL

Project Title: Federated Convolutional Generative Network for Next Item Recommendation

Surname: Long

First Name: Yicheng(Jack)

Student Number:202018010116

Supervisor Name: Joojo Walker

Module Code: CHC 6096

Module Name: Project

Date Submitted:2023.11.3

Table of Contents

1Introduction

1.1Background

In recent years, in the field of recommendation systems, the Next Item Recommendation System has gained increasing attention due to its ability to infer dynamic user preferences through sequential user interactions, as well as its real-time and personalized advantages. Especially in applications that focus on real-time or continuous user experiences, the Next Item Recommendation System has become very popular. For example, users of Last.fm or Weishi typically enjoy a series of songs/videos over a period of time [1]. Convolutional neural network (CNN), as a deep learning neural network architecture, has excellent feature extraction ability and sequence data processing performance. Therefore, the application of CNN in recommendation systems is very common. In addition, federated learning, as a decentralized machine learning method, allows models to be trained on local devices in recommendation systems, thereby protecting user data privacy and reducing data transportation costs. It is now widely used in recommendation systems.

1.2Aim

This project is the next recommendation system based on federated learning and convolutional generative networks. This system solves the user data privacy and data transmission cost issues of traditional recommendation systems by introducing federated learning, and processes data by introducing convolutional generation networks. Finally, a recommendation system that combines the above advantages can surpass existing recommendation systems to achieve better recommendation results.

1.3Objectives

• Collection of relevant literature

• Evaluation and Comparison of Models

Evaluate and compare the advantages and disadvantages of different federated learning recommendation systems and the next recommendation mo del based on CNN.

• Improved model

Propose an improvement to the existing next recommendation model based on CNN, introducing federated learning to the recommendation model.

• Selection and Preprocessing of Datasets

Select a suitable dataset from public resources and preprocess the data to meet the input requirements of the model.

• Implementation of the model

Implement a federated learning framework through programming, based on the next recommendation model of CNN, and then combine them.

• Experimental testing and optimization

Use the dataset as input to the recommendation system, conduct repeated testing, and adjust parameters and models according to different situation s.

• Analysis and Summary

Analyze and summarize the experimental results, evaluate the performance of the recommendation system, including advantages and disadvantages, and complete the paperwork.

• Presentation Prepare

Create PowerPoint presentations, demonstrate videos, and prepare speeches.

1.4Project Overview

1.4.1Scope

The purpose of the study is to improve the next recommendation system by introducing a federated convolutional generation network model to over come the shortcomings of traditional CNN recommendation systems. This model aims to better capture the features and sequential patterns in user proje ct interaction sequences, thereby providing more personalized and accurate recommendations. Regarding importance, it has technological innovation an d addresses the shortcomings of traditional methods. It also has significant importance for practical applications, improving user experience and having a positive impact in commercial applications.

1.4.2Audience

Users and producers of products that require real-time recommendations will be one of the main beneficiaries, such as applications in music streamin g, video on demand, e-commerce, online games, and other fields. For their product users, the Next Item Recommendation System will provide more perso nalized and real-time recommendations that users are interested in, making it easier for product users to find products, services, or information of interes t, This saves time and effort, and product developers gain more users as a result. In addition, some companies or advertisers are the main beneficiaries. By increasing transaction volume, businesses can benefit, and advertisers can place more targeted advertisements to improve advertising efficiency. Technical researchers will also benefit from new technologies and methods.

2Background Review

The earliest work and ideas for sequence recommendation mainly relied on Markov chains [2] and feature based matrix decomposition [3] methods. Markov chains are a mathematical model in which the occurrence of an event only depends on the state of the previous event, and is independent of the earlier state. But it also has some shortcomings, especially when dealing with complex sequence data, its ability to model complex nonlinear relationships and patterns in sequence data is limited and lacks long-term memory. Afterwards, deep learning models gradually began to demonstrate advanced reco mmendation accuracy. In 2016, Hidasi et al.[4] proposed a DL based SBR system, commonly known as GRU4Rec. This is the first model to use RNN, which i ntroduces session parallel small batch, output sampling based on small batch, and sorting loss function, resulting in significant results due to popular base lines. In 2018, Tang and Wang [5] proposed a new sequence recommendation called Caser. They abandoned the RNN structure and proposed a convolutio nal sequence embedding model, demonstrating that this CNN based recommendation can achieve similar or superior performance in the popular RNN m odel's top-N sequence recommendation. Not long after the same year, Yuan et al.[1] proposed a simple, efficient, and efficient convolutional generation m odel for session based top-N project recommendations. This model is suitable for short-term and long-term project dependencies and simplifies deeper n etwork optimization. Ultimately, the model's recommendation accuracy and effectiveness are significantly better than existing technologies at the time. In 2021, Song et al.[6] designed an effective SBRS called Intersessional Collaborative Recommendation Network (Insert) to recommend the next project in sh ort sessions, and designed a Session Retrieval Network (SSRN) to identify sessions similar to the current short session from the historical sessions of the c urrent user and other users, resulting in better recommendation performance than the most advanced series of recommendations at the time. In the same year, Chai et al.[7] proposed a secure matrix decomposition framework under federated learning settings, called FedMF, which to some extent prevents us ers' raw data leakage and ensures user data privacy, but has not been applied in recommendation systems. In 2023, Li et al.[8] proposed Federated Recom mendation with Additive Personalization (FedRAP), which introduced federated learning. This recommendation system effectively avoids user information leakage, reduces communication costs, and solves the problem of poor personalization in other federated recommendations. Afterwards, Kumar et al.[9] p roposed a Horizontal Vertical Convolutional Neural Network (HV-CNN) embedded with Word2Vec technology, which outperformed state-of-the-art meth ods on 30 publicly available music datasets.

Project Proposal-202018010116.docx_第2部分

原文内容

3Methodology

3.1Approach

3.1.1 Federated Learning Recommendation System

Federated learning is a privacy preserving distributed learning scheme proposed by Google, which allows machine learning models to use intermediat e model parameters for training, avoiding direct use of user real data and thus protecting user privacy. Due to the fact that traditional recommendation sy stems typically store user data centrally on a server and use it directly for training and testing, introducing federated learning into recommendation syste ms can effectively protect user privacy data. In this article, we refer to it as the Federated Recommendation System (FedRS). FedRS uses a federated learni ng architecture based on local storage of participant data, as shown in Figure 1. Its communication architecture can be divided into client server architectu re and peer architecture. The client server architecture relies on a central server to perform initialization and model aggregation tasks. The server is responsible for distributing the global recommendation model to selected clients, and then the clients use the received model and local data for local training in each round. Finally, the client sends the updated intermediate parameters (such as model parameters and gradients) back to the server for global model aggregation. The peer-to-peer architecture does not have a central server participating in the communication process. In each round of communication, each participant broadcasts the updated intermediate parameters to some random online neighbors, and then aggregates the received parameters into their own global model. This architecture can avoid single point of failure issues and privacy issues related to central servers.

Figure 1: Communication architecture of FedRS

3.1.2 A Convolutional Generation Model for Item Recommendation Based on Conversation

This model directly processes the sequence of items that have been previously interacted with. Its goal is to estimate the distribution of the original item interaction sequence in order to reliably calculate the probability of the item and generate future items that users may enjoy interacting with. The model introduces a probability distribution p (x), which represents the joint distribution of item sequence x={x0,..., xt}. To model p (x), the chain rule can be used to decompose it into a product of a series of conditional distributions, taking into account the probability distribution of each item, as shown in Figure 2:

Figure 2: Conditional probability of an item

Next, K proposes to use an extended 1D convolutional neural network (CNN) to model the conditional distribution of user item interaction, as shown in Figure 3(b). Compared with the standard CNN (Figure 3(a)), the extended convolutional neural network uses the convolutional operation of "l-dialated convolution" to obtain a larger receptive field without introducing more parameters. This enables the network to better capture long-range dependencies in input data.

Figure 3: Conditional distribution prediction

3.2Technology

Hardware includes: computer (GPU: NVIDIA RTX2070S, RAM:500G), software includes: deep learning framework: TensorFlow, programming language: Python 3.9, editor: PyCharm3.3, data preprocessing tool: MySQL.

3.3Version management plan

Use the Git repository and Feishu to manage project code or multiple versions of models that have been developed. The code will continue to be updated on Github's personal homepage, and backup management will also be carried out every time it is submitted through Flybook to facilitate subsequent changes.

4Project Management

4.1Activities

Objectives Tasks

Collection of relevant literature Read at least 20 articles in relevant fields, select and take notes

Project Proposal Complete the proposal with clear structure and logic, including a reference list

Understanding models and mathematical methods Analyze and compare the differences between models and interpret relevant methods

Improved model According to the shortcomings of the existing model, search for possible improved technologies and methods through the literature

Experimental data processing According to relevant research fields, the data set was searched and preprocessed

Experiment and Test The model was implemented by code and applied to the dataset, trial and error

Summary Analyze and summarize the experimental results to reach a conclusion, and complete the remaining writing

Paper Modify Revise the format and improve the article

Presentation Prepare Prepare PPT and review research work

4.2Schedule

Table 1: Gantt chart of thesis plan

4.3Data management plan

This article uses Zotero to manage relevant literature, including storing the literature in the Zotero manager and annotating and taking notes on each literature, making it easy to quickly search for the desired literature. For code files and dataset files, the code files will be uploaded to a personal account through Github for backup, and each version and date will be classified. Finally, regarding report file management, literature will be stored in cloud documents through the use of WPS office.

Literature management Zotero

Code management Github

Report management WPS office

Table 2: Management

4.4Project Deliverables

Project Resources Deadline

weekly reports Every week from the fourth week of the first semester

Project proposal,ethic form,plagiarism report November 3,2023

Progress Report December 22,2023

final report April 12,2024

presentation poster, video, and demo May 29th to 31st,2024

Tech Show June 6th -7th,2024

Table 3: Resources and Submission Date

5References

[1]F. Yuan, A. Karatzoglou, I. Arapakis, J. M. Jose, and X. He, ‘A Simple Convolutional Generative Network for Next Item Recommendation’. arXiv, Nov. 28, 2018. Accessed: Oct. 27, 2023. [Online]. Available: http://arxiv.org/abs/1808.05163

[2]C. Cheng, H. Yang, M. R. Lyu, and I. King, ‘Where You Like to Go Next: Successive Point-of-Interest Recommendation’.

[3]X. He and T.-S. Chua, ‘Neural Factorization Machines for Sparse Predictive Analytics’. arXiv, Aug. 16, 2017. Accessed: Oct. 27, 2023. [Online]. Available: http://arxiv.org/abs/1708.05027

[4]B. Hidasi, A. Karatzoglou, L. Baltrunas, and D. Tikk, ‘Session-based Recommendations with Recurrent Neural Networks’. arXiv, Mar. 29, 2016. Accessed: Oct. 27, 2023. [Online]. Available: http://arxiv.org/abs/1511.06939

[5]J. Tang and K. Wang, ‘Personalized Top-N Sequential Recommendation via Convolutional Sequence Embedding’. arXiv, Sep. 19, 2018. Accessed: Oct. 27, 2023. [Online]. Available: http://arxiv.org/abs/1809.07426

[6]W. Song, S. Wang, Y. Wang, and S. Wang, ‘Next-item Recommendations in Short Sessions’. arXiv, Jul. 20, 2021. doi: 10.48550/arXiv.2107.07453.

[7]D. Chai, L. Wang, K. Chen, and Q. Yang, ‘Secure Federated Matrix Factorization’, IEEE Intell. Syst., vol. 36, no. 5, pp. 11–20, Sep. 2021, doi: 10.1109/MIS.2020.3014880.

[8]Z. Li, G. Long, and T. Zhou, ‘Federated Recommendation with Additive Personalization’. arXiv, May 17, 2023. doi: 10.48550/arXiv.2301.09109.

[9]C. Kumar and M. Kumar, ‘Next-item recommendation within a short session using the combined features of horizontal and vertical convolutional neural network’, Multimed. Tools Appl., 2023, doi: 10.1007/s11042-023-17201-z.

说明：1.指标是由系统根据《学术论文不端行为的界定标准》自动生成的
2.本报告单仅对您所选择比对资源范围内检测结果负责