Zhaorui Ding

+86 15311720025 dingzr2001@gmail.com github.com/dingzr2001

Education

Huazhong University of Science and Technology(HUST) **9/2020 – 6/2024**

Bachelor of Engineering in Computer Science ***Overall GPA****: 3.81/4.00,* ***Major GPA****: 3.88/4.00* *Wuhan, China*

Relevant Coursework

- Data Structures
- Operating Systems
- Algorithms Analysis(97)
- Database Systems(93)
- C++ Programming(91)
- Machine Learning(97)
- Computer Network(94)
- Computer Architecture(91)

Research Experience

Database Partitioning and Indexing Joint Optimization System **5/2023 – 10/2023**

Research Assistant, Advisor: Prof. ***Hua Wang****, Prof.* ***Ke Zhou*** *Wuhan National Lab for Optoelectronics*

- Conceived and modeled a novel ML-based OLAP system partitioning and indexing algorithm to search for the global optimal partitioning and customized fine-grained indexing scheme jointly under a memory budget.
- Implemented the optimization agent and search space decoder with **Python** to iteratively search for the partitioning and indexing scheme, build the partitions and indexs accordingly, run workloads, and calculate the objective function value.
- Conducted experiments on **Linux** server to test its performance, reproduced codes of state-of-the-art partitioning and indexing works(Slalom, Casper,...) as benchmarks. The scheme we built outperforms theirs by 2x in query latency.

Multi-modal Contrastive Learning Model for De Novo Peptide Sequencing **3/2023 – 9/2023**

Research Intern, mentor: Dr. ***Zhiqiang Gao****, Prof.* ***Siqi Sun*** *Shanghai AI Lab(Pujiang Lab)*

- Summarized research papers about state-of-the-art de novo deep learning models and multi-modal models(OpenAI CLIP, Facebook DPR), conducted experiments to train these models and test their performance.
- Collaborated with team members to construct a multi-modal deep learning model for de novo peptide sequencing using **PyTorch** framework and extract the similarity between spectra and their annotations by contrastive learning.
- Conducted experiments on GPU cluster to pre-train the model and test the performance.

Publication

SOPIOS: Synergistic Optimizer for Partitioning and Indexing in OLAP Systems

Fengrui Liu, ***Zhaorui Ding****, Rukai Wei, Hua Wang, Ke Zhou* *VLDB 24'(in preparation)*

Projects & Internships

High Dimensional ANN Graph Construction | *ACM SIGMOD Programming Contest* **3/2023 - 4/2023**

- Implemented Locality Sensitive Hashing and KD-Tree algorithms with **C++** for Approximate Nearest Neighbor Graph building, compared their performance through experiments, and enhanced their efficiency with multi-threading.
- Employed NN-Descent algorithm to optimize the initial ANN Graph by iteratively exploring neighbors of the current neighbors set of a specific point.
- Conducted experiments on **Linux** server to test its performance on 10M points, reaching 96% recall and ranking 7th.

Transfer Learning for Minority Disk Failure Predictioon | *Tencent Cloud Storage Research Center* **9/2022 - 3/2023**

- Summarized research papers about AI for storage and learned about the application of deep learning, reinforcement learning and transfer learning in storage failure prediction and web service anomaly root cause diagnosis.
- Employed transfer learning model for minority disk failure detection that pre-trains on a large majority disk datasets and leverages TrAdaBoost algorithm to transfer it to small minority disk datasets.

Web Application for Sharing Ordered Learning Documents | *Student Team Project* **7/2022 - 9/2022**

- Led the design of the prototype of an open-source online learning platform that allows for sharing documents and managing them in a tree structure; used Gantt chart and **Git** to manage project progress and team collaboration.
- Built the front-end of the application with **JavaScript** and **Vue** framework to provide GUI, handle user requests, send HTTP requests to back-end and receive and parse back-end responses.
- Tested the robustness and efficiency of the application, deployed it on public network.

Technical Skills

Languages: C, C++, Java, JavaScript, Python, HTML/CSS, SQL
Developer Tools Visual Studio, VS Code, Eclipse, IntelliJ IDEA, PyCharm, Matlab
Technologies/Frameworks: Linux, Git, Latex, Pytorch, Tensorflow, Gym, Vue, SpringBoot