Sajad Ebrahimi

SADJADEB@GMAIL.COM | [LinkedIn] | [GitHub] | [Website] | [Scholar] | ⌂ Toronto, ON

Summary

- Graduate researcher in **NLP and Information Retrieval** with 6+ peer-reviewed publications (SIGIR, ECIR, CIKM); focus on Generative approaches for ranking, Diffusion Models, and Automated Scholarly Peer Review.
- Industry experience as a Data Scientist and Data Analyst, with extensive experience in **LLM training and deployment** across various scales (from 1B to 34B parameters), with advanced proficiency in PyTorch, Transformers, Accelerate, Ollama, and other related tools.

Research Interests

- Natural Language Processing
- Information Retrieval
- Diffusion Models
- Generative Ranking
- Query Performance Prediction
- Reviewer Assignment Problem

Education

M.A.Sc. in Computer Engineering | University of Guelph · Canada | Fall 2023 -

- **CGPA**: 89/100
- **Thesis**: DiffusionRank: A Denoising Diffusion Framework for Learning to Rank

B.Sc. in Computer Engineering | Ferdowsi University of Mashhad · Iran | Fall 2018 - Winter 2023

- **CGPA**: 81/100
- **Selected courses**: Fundamentals of Information Retrieval(A+), Data Mining(A+), Database(A+)

Research Experience

Graduate Research Assistant | University of Guelph · Fall 2023 -
Supervisors: Fattane Zarrinkalam and Ebrahim Bagheri
Affiliated with Vector Institute for Artificial Intelligence as Faculty Affiliate Researcher

- Proposed a novel Query Performance Prediction (QPP) framework using Nearest Neighbor Queries, achieving up to 4% improvement in Kendall's τ on benchmark datasets.
- Prototyped a Generative Learning-to-Rank model inspired by Denoising Diffusion Probabilistic Models (DDPMs).

Research Intern | Toronto Metropolitan University · Fall 2022 - Summer 2023
Supervisor: Ebrahim Bagheri

- Improving dense retriever utilizing Entity Linking - Enhancing the performance of QPP models by enriching query representations, incorporating information from similar queries into the input.

Research Assistant | Ferdowsi University of Mashhad · Summer 2021 - Summer 2022
Supervisor: Behshid Behkamal

- Designing an effective data structure - Developing scrapers to find universities' course catalogs data - Text summarizing and Keyword extracting by state-of-the-art models

Work Experience

Data Scientist | Reviewerly · Toronto, Canada | Nov 2023 -

- Built and now maintain a Milvus-backed index of 260M+ papers and 96M+ authors, delivering sub-second similarity search that powers Reviewerly's matching engine, adopted by two peer-review venues and one grant agency in its first release cycle.
- Parallelized a keyphrase extraction model across 6 GPUs to process 200M records, accelerating the extraction process by 10×.

Data Analyst | Kava Investment Analysis   Tehran, Iran | Sept 2021 - Aug 2022

- Built a scalable web scraper that collected data from 4+ sources across the Tehran Stock Exchange, reducing manual data collection time by 70%.
- Designed and implemented a PostgreSQL data warehouse: defined tables, indexes, and ETL pipelines for equities, bonds, and derivatives, which reduced query latency by 80%.

Software Engineer Intern | Zima & Iliasystem Co.   Iran | Aug 2019 - Feb 2021

- Acquired essential knowledge in backend development, REST API implementation, and database design by contributing to 3 internal tools and APIs used by a team of 10+ developers.
- Learned the fundamentals of web development and built 2 internal-facing pages using HTML, CSS, and JavaScript.

Skills

Languages: Python, C, C++

Python Frameworks and Packages:

- Accelerate, DeepSpeed, PyTorch, TensorFlow, Transformers, Pandas, NumPy, SKlearn, NLTK, spaCy, Beir
- Django, Django Rest Framework, Flask, Fastapi
- Selenium, Pytest, Doctest, BeatifulSoup4

Databases: PostgreSQL, MongoDB, SQLite, MySQL, Milvus

Misc: Git, Docker, Elasticsearch, Nginx, GraphQL, HTML/CSS, LaTeX, Unix-based OS

Soft Skills: Communication, Leadership, Teamwork, Flexibility

Publications

exHarmony: Authorship and Citations for Benchmarking the Reviewer Assignment Problem   ECIR'25
S.Ebrahimi[†], S.Salamat[†], N.Arabzadeh, M.Bashari, E.Bagheri

Reinforcement Learning for Effective Few-Shot Ranking   SIGIR'25
Sh.Soleimany, **S.Ebrahimi**, Sh.SeyedSalehi, F.Zarrinkalam, E.Bagheri

Gender Disentangled Representation Learning in Neural Rankers   MLJ'24
Sh.SeyedSalehi, S.Salamat, N.Arabzadeh, **S.Ebrahimi**, M.Zihayat, E.Bagheri

Estimating Query Performance Through Rich Contextualized Query Representations   ECIR'24
S.Ebrahimi, M.Khodabakhsh, N.Arabzadeh, E.Bagheri

Reviewerly: Modeling the Reviewer Assignment Task as an Information Retrieval Problem   CIKM'24
N.Arabzadeh, **S.Ebrahimi**, S.Salamat, M.Bashari, E.Bagheri

Estimating Query Performance Using Neural Query Space Proximity   *Under Review
A.Bigdeli, **S.Ebrahimi**, N.Arabzadeh, S.Salamat, Sh.SeyedSalehi, F.Zarrinkalam, E.Bagheri

Selected Projects

RottenReviews   2025
Benchmarking Review Quality with Human and LLM-Based Judgments

QR Code Ticket System   2022
An app that generates a QR code for users and provides a QR reader for the receptionist to verify each ticket.

Academic Services

- **Organizer**: ProActLLM: Proactive Conversational Information Seeking with Large Language Models @ CIKM'25
- **Program Committee**: TOIS, WSDM'26, SIGIR'25, ICTIR'25, ReasoningandPlanningforLLMs@ICLR'25, COLING'25