Ronak Pradeep

Email 1 : ronakice@hotmail.com
Email 2 : rpradeep@uwaterloo.ca
Website : ronakice.github.io
Mobile : 519-781-6942

Education

- **University of Waterloo** Waterloo, ON
  PhD - Computer Science
  - Neural Information Retrieval and Large Language Models with Professor Jimmy Lin
  - Fellowship - Apple PhD Fellow
  - Coursework: Differential Privacy, Optimization for Data Science, High Stakes Information Retrieval, Affective Computing, Deep Learning for Biotechnology
  - Teaching Assistant: Algorithms, Data Structures, Introduction to Computer Science II

- **University of Waterloo** Waterloo, ON
  BMath - Double Major in Computer Science and Combinatorics and Optimization
  - Graduate Level Coursework: Deep Reinforcement Learning, Randomized Algorithms, Formal Languages and Parsing (100%), Computational Vision, Statistical Learning, Dependent Types and Software Verification (100%)
  - Part of the Term Dean's Honours List; Graduate Level Coursework Average 96.5%

Research Interests

Evaluation, Benchmarking, Retrieval-Augmented Large Language Models, Reranking,Factuality, Document Ranking, Fact Verification, Biomedical NLP

Experience

- **Yupp AI** Waterloo, ON
  Research Scientist *Jan 2025 - Present*
  - Working on agents, models, labelers, evaluations, and leaderboards.

- **Apple - Siri Information Intelligence** Seattle, WA
  Research Intern *May 2023 - Aug 2023*
  - Worked with Yunyao Li on the Siri Team on Knowledge Graphs and Large Language Models

- **Google AI** Remote
  Student Researcher *Feb 2022 - Feb 2023*
  - Worked with Donald Metzler on scaling Differentiable Search Indexes to large corpora and models

- **University of Waterloo** Waterloo, ON
  Undergraduate Researcher *Apr 2017 - Aug 2020*
  - Worked with Professor Jimmy Lin and Rodrigo Nogueira on Paragraph Retrieval and Ranking
  - Worked with Professor Pascal Poupart on Reading Comprehension tasks
  - Worked with Professor Jeff Orchard on a Deep Biologically Plausible Vision Model

- **Montreal Institute for Learning Algorithms (MILA)** Montreal, QC
  Visiting Researcher *May 2019 - Dec 2019*
  - Worked with Professor Chris Pal and Dr. Jie Fu on Open Domain Question Answering and Graph Representation Learning

- **Wish** San Francisco, CA
  AI Research Intern *Jan 2018 - Apr 2018*
  - Worked on Neural Title Generation for e-Commerce Products and Attribute Categorization
  - Curated the iMaterialist Challenge for the FGVC Workshop at CVPR 2018

- **Royal Bank of Canada** Toronto, ON
  Research Developer *Aug 2016 - Dec 2016*
  - Worked on Document Ranking and Question Answering using Dual Embedding Space & Seq2Seq models

- **University of Waterloo** Waterloo, ON
  Undergraduate Teaching Assistant for Math 136 - Linear Algebra *Jan 2017 - Apr 2017*

Notable Professional Activities

- **TREC 2025–2026 Retrieval-Augmented Generation Track**
  Co-organizer
  - The focus of this track is on building a realistic test collection to facilitate scalable, diverse, and reproducible Retrieval-Augmented Generation (RAG) pipelines.
  - Organized with Shivani Upadhyay (lead organizer), Nandan Thakur, Daniel Campos, Ian Soboroff, Nick Craswell, and Jimmy Lin.

- **TREC 2024 Retrieval-Augmented Generation Track**
  Lead Organizer
  - The focus of this track is on building a realistic test collection to facilitate scalable, diverse, and reproducible Retrieval-Augmented Generation (RAG) pipelines.
  - Done in collaboration with Nandan Thakur, Shivani Upadhyay, Daniel Campos, Ian Soboroff, Nick Craswell, and Jimmy Lin.

Publications

- **NeuCLIRBench: A Modern Evaluation Collection for Monolingual, Cross-Language, and Multilingual Information Retrieval** (Under Review)
  Dawn Lawrie, James Mayfield, Eugene Yang, Andrew Yates, Sean MacAvaney, ***Ronak Pradeep****, Scott Miller, Paul McNamee, Luca Soldani*

- **Humanity's Last Exam** (arXiv 2025)
  Long Phan, Alice Gatti, Ziwen Han, ..., ***Ronak Pradeep****, ... [and other co-authors]*

- **RankLLM: A Python Package for Reranking with LLMs** (SIGIR 2025)
  Sahel Sharifymoghaddam, ***Ronak Pradeep****, Andre Slavescu, Ryan Nguyen, Andrew Xu, Zijian Chen, Yilin Zhang, Yidi Chen, Jasper Xian, Jimmy Lin*

- **Gosling Grows Up: Retrieval with Learned Dense and Sparse Representations Using Anserini** (SIGIR 2025)
  Jimmy Lin, Arthur Haonan Chen, Carlos Lassance, Xueguang Ma, ***Ronak Pradeep****, Tommaso Teofili, Jasper Xian, Jheng-Hong Yang, Brayden Zhong, Vincent Zhong*

- **Support Evaluation for the TREC 2024 RAG Track: Comparing Human versus LLM Judges** (SIGIR 2025)
  Nandan Thakur, ***Ronak Pradeep****, Shivani Upadhyay, Daniel Campos, Nick Craswell, Jimmy Lin*

- **Chatbot Arena Meets Nuggets: Towards Explanations and Diagnostics in the Evaluation of LLM Responses** (Under Review)
  Shivani Upadhyay, ***Ronak Pradeep****, Nandan Thakur, Daniel Campos, Nick Craswell, Ian Soboroff, Hoa Trang Dang, Jimmy Lin*

- **The Great Nugget Recall: Automating Fact Extraction and RAG Evaluation with Large Language Models** (SIGIR 2025)
  ***Ronak Pradeep**, Nandan Thakur, Shivani Upadhyay, Daniel Campos, Nick Craswell, Jimmy Lin*
- **A FIRST Reproduction and Improvements to Single-Token Decoding for Fast Listwise Reranking** (SIGIR 2025)
  *Zijian Chen, **Ronak Pradeep**, Jimmy Lin*
- **Initial Nugget Evaluation Results for the TREC 2024 RAG Track with the AutoNuggetizer Framework** (arXiv 2024)
  ***Ronak Pradeep**, Nandan Thakur, Shivani Upadhyay, Daniel Campos, Nick Craswell, Jimmy Lin*
- **A Large-Scale Study of Relevance Assessments with Large Language Models: An Initial Look** (arXiv 2024)
  *Shivani Upadhyay, **Ronak Pradeep**, Nandan Thakur, Daniel Campos, Nick Craswell, Ian Soboroff, Hoa Trang Dang, Jimmy Lin*
- **Ragnarok: A Reusable RAG Framework and Baselines for TREC 2024 Retrieval-Augmented Generation Track** (ECIR 2025)
  ***Ronak Pradeep**, Nandan Thakur, Sahel Sharifymoghaddam, Eric Zhang, Ryan Nguyen, Daniel Campos, Nick Craswell, Jimmy Lin*
- **Prompts as Auto-Optimized Training Hyperparameters: Training Best-in-Class IR Models from Scratch with 10 Gold Labels** (Under Review for a Suitable Conference)
  *Jasper Xian, Saron Samuel, Faraz Khoubsirat, **Ronak Pradeep**, Md Arafat Sultan, Radu Florian, Salim Roukos, Avirup Sil, Christopher Potts, Omar Khattab*
- **UMBRELA: UMbrela is the (Open-Source Reproduction of the) Bing RELevance Assessor** (Under Review for a Suitable Conference)
  *Shivani Upadhyay, **Ronak Pradeep**, Nandan Thakur, Nick Craswell, Jimmy Lin*
- **ConvKGYarn: Spinning Configurable and Scalable Conversational Knowledge Graph QA datasets with Large Language Models** (ACL 2024 KaLLM Workshop, EMNLP 2024 Industry Track)
  ***Ronak Pradeep**, Daniel Lee, Ali Mousavi, Jeffrey Pound, Yisi Sang, Jimmy Lin, Ihab Ilyas, Saloni Potdar, Mostafa Arefiyan, Yunyao Li*
- **Entity Disambiguation via Fusion Entity Decoding** (NAACL 2024)
  *Junxiong Wang, Ali Mousavi, Omar Attia, **Ronak Pradeep**, Saloni Potdar, Alexander Rush, Umar Farooq Minhas, Yunyao Li*
- **Zero-Shot Cross-Lingual Reranking with Large Language Models for Low-Resource Languages** (ACL 2024)
  *Mofetoluwa Adeyemi, Akintunde Oladipo, **Ronak Pradeep**, Jimmy Lin*
- **LiT and Lean: Distilling Listwise Rerankers Into Encoder-Decoder Models** (ECIR 2025)
  *Manveer Singh Tamber, **Ronak Pradeep**, Jimmy Lin*
- **RankZephyr: Effective and Robust Zero-Shot Listwise Reranking is a Breeze!** (Under Review for a Suitable Conference)
  ***Ronak Pradeep**, Sahel Sharifymoghaddam, Jimmy Lin*
- **RankVicuna: Zero-Shot Listwise Document Reranking with Open-Source Large Language Models** (Under Review for a Suitable Conference)
  Ronak Pradeep, Sahel Sharifymoghaddam*, Jimmy Lin*
- **Vector Search with OpenAI Embeddings: Lucene Is All You Need** (WSDM 2024 Demo)
  *Jimmy Lin, **Ronak Pradeep**, Tommaso Teofili, Jasper Xian*

- **How Does Generative Retrieval Scale to Millions of Passages?** (EMNLP 2023, SIGIR 2023 GenIR Workshop)
  Ronak Pradeep*, Kai Hui*, Jai Gupta, Adam D Lelkes, Honglei Zhuang, Jimmy Lin, Donald Metzler, Vinh Q Tran**
- **naverloo @ TREC Deep Learning and NeuCLIR 2023: As Easy as Zero, One, Two, Three — Cascading Dual Encoders, Mono, Duo, and Listo for Ad-Hoc Retrieval** (TREC 2023)
  Carlos Lassance, ***Ronak Pradeep****, Jimmy Lin*
- **Towards Automated End-to-End Health Misinformation Free Search with a Large Language Model** (ECIR 2024)
  Ronak Pradeep*, Jimmy Lin*
- **Zero-Shot Listwise Document Reranking with a Large Language Model** (arXiv)
  Xueguang Ma, Xinyu Zhang, ***Ronak Pradeep****, Jimmy Lin*
- **ReadProbe: A Demo of Retrieval-Enhanced Large Language Models to Support Lateral Reading** (arXiv)
  Dake Zhang, ***Ronak Pradeep***
- **Pre-Processing Matters! Improved Wikipedia Corpora for Open-Domain Question Answering** (ECIR 2023 Reproducibility)
  Manveer Singh Tamber, ***Ronak Pradeep****, Jimmy Lin*
- **PyGaggle: A Gaggle of Resources for Open-Domain Question Answering** (ECIR 2023 Reproducibility)
  Ronak Pradeep*, Haonan Chen, Lingwei Gu, Manveer Singh Tamber, Jimmy Lin*
- **Neural Query Synthesis and Domain-Specific Ranking Templates for Multi-Stage Clinical Trial Matching** (SIGIR 2022)
  Ronak Pradeep*, Yilin Li, Yuetong Wang, Jimmy Lin*
- **Document Expansion Baselines and Learned Sparse Lexical Representations for MS MARCO V1 and V2** (SIGIR 2022 Resource)
  Xueguang Ma, **Ronak Pradeep***, Rodrigo Nogueira, Jimmy Lin*
- **Squeezing Water from a Stone: A Bag of Tricks for Further Improving Cross-encoder Effectiveness for Reranking** (ECIR 2022 Reproducibility)
  Ronak Pradeep*, Yuqi Liu, Xinyu Zhang, Yilin Li, Andrew Yates, Jimmy Lin*
- **Another Look at DPR: Reproduction of Training and Replication of Retrieval** (ECIR 2022 Reproducibility)
  Xueguang Ma, Kai Sun, ***Ronak Pradeep****, Minghan Li, Jimmy Lin*
- **New Nails for Old Hammers: Anserini and Pyserini at TREC 2021** (TREC 2021 Proceedings)
  Jimmy Lin, Haonen Chen, Chengcheng Hu, Sheng-Chieh Lin, Yilin Li, Xueguang Ma, ***Ronak Pradeep****, Jheng-Hong Yang, Chuan-Ju Wang, Andrew Yates, Xinyu Zhang*
- **Vera: Prediction Techniques for Reducing Harmful Misinformation in Consumer Health Search** (SIGIR 2021)
  Ronak Pradeep*, Xueguang Ma, Rodrigo Nogueira, Jimmy Lin*
- **Chatty Goose: A Python Framework for Conversational Search** (SIGIR 2021 Demo)
  Edwin Zhang, Sheng-Chieh Lin, Jheng-Hong Yang, ***Ronak Pradeep****, Rodrigo Nogueira, and Jimmy Lin*
- **Pyserini: An Easy-to-Use Python Toolkit to Support Replicable IR Research with Sparse and Dense Representations** (SIGIR 2021 Resource)
  Jimmy Lin, Xueguang Ma, Sheng-Chieh Lin, Jheng-Hong Yang, ***Ronak Pradeep****, and Rodrigo Nogueira*

- **H_2oloo at TAC 2020: Epidemic Question Answering** (TAC 2020 Proceedings)
  Justin Borromeo, **Ronak Pradeep***, Jimmy Lin*
- **Exploring Listwise Evidence Reasoning with T5 for Fact Verification** (ACL 2021)
  Kelvin Jiang, **Ronak Pradeep***, Jimmy Lin*
- **H_2oloo at TREC 2020: When all you got is a Hammer... Deep Learning, Health Misinformation, and Precision Medicine** (TREC 2020 Proceedings)
  ***Ronak Pradeep**, Xueguang Ma, Xinyu Zhang, Hang Cui, Ruizhou Xu, Rodrigo Nogueira, Jimmy Lin*
- **Scientific Claim Verification with VerT5erini** (LOUHI 2021: The 12th International Workshop on Health Text Mining and Information Analysis colocated with EACL 2021)
  ***Ronak Pradeep**, Xueguang Ma, Rodrigo Nogueira, Jimmy Lin*
- **A Replication Study of Dense Passage Retriever** (Will be submitted to a suitable venue)
  *Xueguang Ma, **Ronak Pradeep**, Kai Sun, Jimmy Lin*
- **Covidex: Neural Ranking Models and Keyword Search Infrastructure for the COVID-19 Open Research Dataset** (Scholarly Document Processing @ EMNLP 2020)
  *Edwin Zhang, Nikhil Gupta, Raphael Tang, Xiao Han, **Ronak Pradeep**, Kuang Lu, Yue Zhang, Rodrigo Nogueira, Kyunghyun Cho, Hui Fang, Jimmy Lin*
- **The Expando-Mono-Duo Design Pattern for Text Ranking with Pretrained Sequence-to-Sequence Models** (arXiv)
  ***Ronak Pradeep**, Rodrigo Nogueira, Jimmy Lin*
- **Document Ranking with a Pretrained Sequence-to-Sequence Model** (EMNLP 2020 Findings)
  *Rodrigo Nogueira, Zhiying Jiang, **Ronak Pradeep**, Jimmy Lin*
- **Modular Diversity-Seeking Query Reformulation for Open-Domain Question Answering**
  Ronak Pradeep, Jie Fu*, Xingdi Yuan, Zhouhan Lin, Yi Tay, Chris Pal*
- **Foveated Down-Sampling Techniques** (CVIS 2020)
  *Parsa Torabian, **Ronak Pradeep**, Jeff Orchard, Bryan Tripp*

Accomplishments

- **TREC NeuCLIR 2023**: A track that studies cross-lingual information retrieval — Top Submission.
- **SCIVER: Verifying Scientific Claims with Evidence (Scholarly Document Processing @ NAACL 2021)**: Top Submission based on primary metric.
- **Fact Extraction and VERification (FEVER) - 1st (As of Jan 14th 2021)**: State-of-the-Art model in a widely popular Fact Verification dataset
- **TREC Clinical Trials 2021/2022**: A task that matches patient descriptions to clinical trials — Top Submission.
- **TREC Health Misinformation 2020/2021/2022**: A task that studies search technologies that promote credible and correct information over incorrect information — Top Submission in the AdHoc Retrieval task.
- **TREC Deep Learning 2020/2023**: A track that studies information retrieval in a large training data regime - Top Submission in the Document Ranking task
- **TREC-COVID 2020**: A multi-round COVID-19 Literature Ranking Task - Best Round 4, 5 Automatic Run, Best Round 3 Feedback run
- **Fact Extraction and VERification (FEVER) - 1st (As of Jan 14th)**: State-of-the-Art model in a widely popular Fact Verification dataset
- **MS MARCO Document Ranking - 1st (As of Sep 8th 2020)**: State-of-the-Art model in a widely popular Neural Document Ranking dataset
- **MS MARCO Passage Ranking - 1st (As of May 20th 2020)**: State-of-the-Art model in a widely popular Neural Passage Ranking dataset

- **DiMarco Undergraduate Scholarship in Computational Rhetoric**: Annually awarded to a single student based on academic achievement combined with a well-demonstrated interest in the area of Computational Rhetoric
- **Terminal AI - Winner**: Developed an heuristic-based AI game bot that placed 1st among teams of top Waterloo students. Globally ranked 2nd among 15k players (at the time of submission)
- **Citadel Datathon - NYC**: Placed 2nd among teams from top universities in North America
- **HackPrinceton - Top 10**: Implemented a tool for the Sentiment Analysis of Twitter and Guardian News using Vader Lexicon and Encoder-Decoder LSTMs and visualized the trends
- **University of Waterloo President's Scholarship of Distinction and Research Award**: Awarded based on high academic average and research terms