

Omid Jafari

CONTACT INFORMATION Science Hall 159, NMSU, Las Cruces, NM 88003 ojafari@nmsu.edu
<https://www.linkedin.com/in/omidjafari> <http://omidjafari.com>

RESEARCH INTERESTS Data Science, Big Data, Query Optimization, Indexing, Distributed Computing

EDUCATION **New Mexico State University**, Las Cruces, NM, USA

Ph.D., Computer Science, GPA: 3.9 [Jan. 2018 - Present]
– Advisor: Parth Nagarkar

Azad University of Mashhad, Mashhad, Iran

M.Sc., Computer Software Engineering, GPA: 4.0 [Sept. 2014 - Dec. 2017]
– Ranked 2nd in my class
– Thesis: *Optimized MapReduce-based K-Means++ algorithm in Hadoop-GPU framework*
– Advisor: Hossein Deldari

Quchan University of Technology, Quchan, Iran

B.Sc., Chemical Engineering, GPA: 2.72 [Jan. 2008 - Jan. 2014]
– 2 semesters at “**Ferdowsi University of Mashhad**, Iran” as a transfer student
– 4 semesters at “**University of Tehran**, Iran” as a transfer student
– Thesis: *Modelling of vacuum membrane distillation*
– Advisor: Majid Mahdavian

WORK EXPERIENCE **Summer Intern** June 2021 - Aug. 2021
at Vigilant Technologies, Tempe, AZ
– Developed a chatbot using Microsoft Bot Framework SDK and Microsoft Language Understanding (LUIS) service.
– Implemented a MongoDB database to be used by the data science tasks.

Summer Intern June 2020 - Aug. 2020
at Vigilant Technologies, Tempe, AZ
– Developed a data science engine capable of performing several NLP tasks in the domain of satellite launches.
– Utilized RabbitMQ and developed a message bus architecture to be used in a cluster of worker nodes.

Founder and Developer [2012 - 2016]
Founded and developed the **HyperIRC** project, which provided customized bouncers and bots to OnlineGamesNet IRC users, and did Linux server administration and network programming.

Translator [2012 - 2014]
Worked as the official English-Persian translator at GameForge company.

Game Operator, Mobile Games Tester, and Support Manager [2012 - 2014]
Worked as a volunteer in several game communities of GameForge company.

RESEARCH EXPERIENCE **Graduate Research Assistant** [Jan. 2018 - Present]
at New Mexico State University
– **Fair-LSH**
Developed an optimized cost-based index structure for retrieving fair approximate nearest neighbors based on Locality Sensitive Hashing. Languages used: C++

- **LSH Benchmark**
Did an in-depth evaluation of LSH-based algorithms for similarity searches in High-dimensional spaces. Languages used: C++, Python, Bash
- **roLSH**
Utilized a Neural Network to optimize the performance of LSH-based techniques. Result accepted to be published in SISAP 2020. Languages used: C++, Python
- **mmLSH**
Developed an LSH-based index structure for query processing with a content-based image retrieval application. Result accepted to be published in SISAP 2020. Languages used: C++
- **bitmapLSH**
Developed a bitmap-based indexing structure for retrieval of similarity search image queries. Result published in SSIAI2020. Languages used: C++
- **rtLSH**
Compared different LSH-based techniques for real-time processing. Result published in SIGI-2019. Languages used: C++, Python
- **qwLSH**
Developed a cache-conscious indexing for processing similarity search query workloads in high-dimensional spaces. Result published in ICMR2019. Languages used: C++

Graduate Research Assistant [Dec. 2014 - Oct.2015]

at Azad University of Mashhad

- **M/M/1 Queue Analysis**
Simulated an M/M/1 queue and did a case study on a bank call center. Result published in ICKIS2015. Languages used: Matlab

TECHNICAL SKILLS

Programming Languages

- Python (proficient), C++ (proficient), Bash (proficient), SQL (proficient), Matlab (experience), TCL (experience), VB6 (experience)

Python Data Science and Machine Learning Libraries

- Pandas, Numpy, Scipy, GraphLab, Scikit-Learn, Keras, NLTK, Requests, BeautifulSoup, Scrapy, Matplotlib, SpaCy, Gensim

Professional Softwares/Platforms

- Apache Hadoop, MPI, Weka, RapidMiner, Genie

Web Technologies

- HTML, CSS, Bootstrap, Flask, Django

General Softwares/Platforms

- Git, LaTeX, EndNote, Microsoft Office, Photoshop

Operating Systems

- Linux (Ubuntu, Centos, Debian), Microsoft Windows

PUBLICATIONS

Peer-Reviewed Conference Papers

1. **Omid Jafari**, Preeti Maurya, Khandker Mushfiqul Islam, and Parth Nagarkar. "Optimizing Fair Approximate Nearest Neighbor Searches using Threaded B+-Trees." in *14th International Conference on Similarity Search and Applications (SISAP 2021)*, 2021.
2. Shikun Zhang, **Omid Jafari**, and Parth Nagarkar. "A Survey on Data Annotation and Labeling in Machine Learning." in *Arxiv*, 2021.

3. Nada Ibrahim, Preeti Maurya, **Omid Jafari**, and Parth Nagarkar. “A Survey of Performance Optimization in Neural Network-Based Video Analytics Systems.” in *Arxiv*, 2021.
4. **Omid Jafari**, Preeti Maurya, Parth Nagarkar, Khandker Mushfiqul Islam, and Chidambaram Crushev . “A Survey on Locality Sensitive Hashing Algorithms and their Applications.” in *Arxiv*, 2021.
5. **Omid Jafari** and Parth Nagarkar. “Experimental Analysis of Locality Sensitive Hashing Techniques for High-Dimensional Approximate Nearest Neighbor Searches.” in *Australasian Database Conference (ADC 2021)*, 2021.
6. Parth Nagarkar, Arnab Bhattacharya, and **Omid Jafari**. “Exploring State-of-the-Art Nearest Neighbor (NN) Search Techniques.” in *ACM India Joint International Conference on Data Science & Management of Data (CODS-COMAD 2021)*, 2021.
7. **Omid Jafari**, Parth Nagarkar, Bhagwan Thatte, and Carl Ingram. “SatelliteNER: An Effective Named Entity Recognition Model for the Satellite Domain.” in *12th International Conference on Knowledge Management and Information Systems (KMIS 2020)*, 2020.
8. **Omid Jafari**, Parth Nagarkar, and Jonathan Montaño. “Improving Locality Sensitive Hashing by Efficiently Finding Projected Nearest Neighbors.” in *13th International Conference on Similarity Search and Applications (SISAP 2020)*, 2020.
9. **Omid Jafari**, Parth Nagarkar, and Jonathan Montaño. “mmLSH: A Practical and Efficient Technique for Processing Approximate Nearest Neighbor Query Workloads in High-Dimensional Spaces.” in *13th International Conference on Similarity Search and Applications (SISAP 2020)*, 2020.
10. **Omid Jafari**, Parth Nagarkar, and Jonathan Montaño. “Efficient Bitmap-based Indexing and Retrieval of Similarity Search Image Queries.” in *2020 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI 2020)*, 2020.
11. **Omid Jafari**, John Ossorgin, and Parth Nagarkar. “qwLSH: Cache-conscious Indexing for Processing Similarity Search Query Workloads in High-Dimensional Spaces.” in *Proceedings of the 2019 on International Conference on Multimedia Retrieval (ICMR2019)*, 2019.
12. **Omid Jafari**, Khandker M. Islam, and Parth Nagarkar. “Drawbacks and Proposed Solutions for Real-time Processing on Existing State-of-the-art Locality Sensitive Hashing Techniques.” in *The 5th International Conference on Signal and Image Processing (SIGI-2019)*, 2019.
13. **Omid Jafari**, Donya Ghasvari, and Majid Vafaei Jahan. “M/M/1 queue analysis and case study of a call center.” in *The 2nd International Conference on Knowledge, information and software engineering (ICKIS2015)*, 2015.

PAPERS IN
PREPARATION

–

TEACHING
EXPERIENCE

Teaching Assistant

at New Mexico State University

CS 482/502 - Database Management Systems I [Fall 2018 - Fall 2019 - Summer 2020]
 CS 582 - Database Management Systems II [Spring 2019]
 CS 479/579 - Special Topics: Distributed DB [Fall 2019]

Instructor

Taught Matlab Simulink in a workshop held at Azad University of Mashhad. [Fall 2016]

Video Tutorial Instructor [Spring 2016]
Made tutorials about “Computer Modeling”, available on <http://vafaeijahan.com>

RELEVANT
COURSEWORK

CS 519 - APPLIED ML I [A+]

– **Projects**

Single-layer Linear Neural Networks, Comparing classifiers in scikit-learn library, Comparing regression methods, Comparing clustering methods, Dimensionality reduction techniques, Convolutional Neural Networks, Ensemble approaches.

CS 508 - DATA MINING [A]

CS 582 - DATABASE MANAGEMENT SYSTEMS II [A-]

– **Projects**

Implemented a columnar approach in the Java Minibase DBMS.

DISTRIBUTED SYSTEMS [A+]

CS 512 - COMPUTER SYSTEMS MODELING AND SIMULATION [A+]

SERVICE

External Reviewer for Conference Publications

– 27th SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2021

– 26th International Conference on Database Systems for Advanced Applications (DASFAA) 2021

– Australasian Computer Science Week (ACSW) 2021

– IEEE Transactions on Knowledge and Data Engineering (TKDE) 2020

– Fourth Workshop on Software Foundations for Data Interoperability (SFDI) 2020

– 12th International Conference on Knowledge Management and Information Systems (KMIS) 2020

– Proceedings of Very Large Databases (PVLDB) 2019

– ACM Special Interest Group on Management of Data (SIGMOD) Demonstration Track 2019

– Conference on Information and Knowledge Management (CIKM) 2019

– ACM International Conference on Multimedia Retrieval (ICMR) 2019

Reviewer for Journal Publications

– International Journal of Intelligent Systems 2021

Invited Panelist 2020

at the 13th International Conference on Similarity Search and Applications (SISAP)

Graduate College Senator

[Fall 2019 - Present]

at Associated Students of New Mexico State University

Vice president of Iranian Students Organization

[Aug. 2019 - Aug. 2020]

at New Mexico State University

AWARDS

Scholarships

– Distinguished Student Leadership [Fall 2019 - Fall 2021]

Best Paper

– Best Student Paper award in the 12th International Conference on Knowledge Management and Information Systems (KMIS) 2020

Travel Awards

– The Associated Students of New Mexico State University, Travel Grant (to attend ICMR 2019) 2019

- Graduate Student Council of New Mexico State University, Travel Grant (to attend ICMR 2019) 2019
- The Associated Students of New Mexico State University, Travel Grant (to attend SIGI 2019) 2019
- Graduate Student Council of New Mexico State University, Travel Grant (to attend SIGI 2019) 2019