

# Omid Jafari



Science Hall 159, NMSU, Las Cruces, NM 88003



<http://omidjafari.com>



<https://linkedin.com/in/omidjafari>



[ojafari@nmsu.edu](mailto:ojafari@nmsu.edu)

## Education

New Mexico State University, USA  
Ph.D., Computer Science, GPA: 3.9  
| Jan. 2018 - Jan. 2022

Azad University of Mashhad, Iran  
M.Sc., Software Engineering, GPA: 4.0  
| Sept. 2014 - Dec. 2017 | Thesis:  
*Optimized MapReduce-based  
K-Means++ algorithm in Hadoop-GPU  
framework*

## Skills

Languages: Python, C++, Bash, SQL,  
Matlab, TCL, VB6

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

Web: HTML, CSS, Flask, Django

Platforms: Apache Hadoop, MPI,  
Weka, RapidMiner, Genie

## Awards

Scholarships: Distinguished Student  
Leadership Scholarship | 2019-2021

Best Student Paper: KMIS2020

Travel: ASNMSU to attend ICMR2019  
and SIGI2019, GSC to attend  
ICMR2019 and SIGI2019

## Services

External Reviewer: KDD'21,  
DASFAA'21, ACSW'21, TKDE'20,  
KMIS'20, SFDI'20, PVLDB'19,  
SIGMOD'19, CIKM'19, ICMR'19

Reviewer: International Journal of  
Intelligent Systems'21

Invited Panelist: SISAP'20

College Senator at Associated  
Students of New Mexico State  
University | Fall 2019 - Present

Vice-president of Iranian Students  
Organization | Aug. 2019 - Aug. 2020

## Interests

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

## Work and Research Experience

Graduate Research Assistant                      New Mexico State University                      1/18 - Now  
Developed an optimized cost-based index structure for retrieving fair approxi-  
mate nearest neighbors (using C++).

Evaluated LSH-based algorithms for similarity searches in High-dimensional  
spaces (using C++).

Developed an efficient machine learning based technique to improve any LSH-  
based data-independent algorithm (using Python and C++).

Developed an LSH-based index structure for query processing with a content-  
based image retrieval application (using C++).

Developed bitmap-based indexing for similar image retrieval (using C++).

Benchmarked LSH algorithms for real-time processing (using C++ and Python).

Developed cache-conscious indexing for LSH algorithms (using C++).

Summer Intern                      Vigilant Technologies, Tempe, AZ                      6/21-9/21  
Developed a chatbot using Microsoft Bot Framework SDK.  
Implemented a MongoDB database.

Summer Intern                      Vigilant Technologies, Tempe, AZ                      6/20-9/20  
Developed an NLP data science engine in the domain of satellite launches.  
Utilized RabbitMQ message bus in a cluster of worker nodes.

Graduate Research Assistant                      Azad University of Mashhad                      12/14-10/15  
Simulated an M/M/1 queue in Matlab and did a case study on a bank call center.

## Peer-Reviewed Publications

1. "Optimizing Fair Approximate Nearest Neighbor Searches using Threaded B+-Trees." in *SISAP2021*.
2. "A Survey on Data Annotation and Labeling in Machine Learning." in *Arxiv2021*.
3. "A Survey of Performance Optimization in Neural Network-Based Video Analytics Systems." in *Arxiv2021*.
4. "A Survey on Locality Sensitive Hashing Algorithms and their Applications." in *Arxiv2021*.
5. Tutorial: "Exploring State-of-the-Art Nearest Neighbor (NN) Search Techniques." in *CODS-COMAD2021*.
6. "Experimental Analysis of Locality Sensitive Hashing Techniques for High-Dimensional Approximate Nearest Neighbor Searches." in *ADC2021*.
7. "Improving Locality Sensitive Hashing by Efficiently Finding Projected Nearest Neighbors." in *SISAP2020*.
8. "mmLSH: A Practical and Efficient Technique for Processing Approximate Nearest Neighbor Query Workloads in High-Dimensional Spaces." in *SISAP2020*.
9. "SatelliteNER: An Effective Named Entity Recognition Model for the Satellite Domain." in *KMIS2020*.
10. "Efficient Bitmap-based Indexing and Retrieval of Similarity Search Image Queries." in *SSIAI2020*.
11. "qwLSH: Cache-conscious Indexing for Processing Similarity Search Query Workloads in High-Dimensional Spaces." in *ICMR2019*.
12. "Drawbacks and Proposed Solutions for Real-time Processing on Existing State-of-the-art Locality Sensitive Hashing Techniques." in *SIGI-2019*.
13. "M/M/1 queue analysis and case study of a call center." in *ICKIS2015*.