HAJAR HOMAYOUNI

hhomayouni@sdsu.edu (970)-844-9638
Assitant Professor
Department of Computer Science, San Diego State University

EDUCATION

Ph.D., Computer Science, Colorado State University, GPA: 3.97 *Feb 2018 - May 2021*
Dissertation: Anomaly Detection and Explanation in Big Data
Committee: Profs. Sudipto Ghosh, Indrakshi Ray (Advisors),
James M. Bieman, Indrajit Ray, and Leo Vijayasarathy

M.S., Computer Science, Colorado State University, GPA: 3.96 *Aug 2015 - Dec 2017*
Thesis: An Approach for Testing the Extract-Transform-Load (ETL) Process in Data Warehouse Systems
Committee: Profs. Sudipto Ghosh, Indrakshi Ray (Advisors), James M. Bieman, and Leo Vijayasarathy

M.S., Computer Science, Alzahra University (Iran), GPA: 3.90 *Aug 2010 - Dec 2013*
Thesis: Automatic Test Case Generation for Web Applications
Committee: Profs. Mohammad Reza Keyvanpour (Advisor) and Eslam Nazemi

B.S., Computer Science, University of Kashan (Iran), GPA: 3.5 *Aug 2004 - Sep 2008*

AREAS OF INTEREST

- Data science, applied Machine Learning, and interpretability of Machine Learning
- Big data quality assurance and anomaly detection and explanation in big data
- Streaming time-series data analysis and testing
- Complex (mixed and non-primitive) data analysis and testing
- Data warehouse and Extract-Transform-Load analysis and testing
- Software testing

AWARDS AND FELLOWSHIPS

- STARS Computing Corps Fellowship to Support Diverse Cohorts of Students, 2023
- SDSU Grants Research and Enterprise Writing (GREW) Fellowship, 2022
- SDSU Inclusive Excellence Faculty Fellowship, 2022
- Faculty Scholarship, ACM Richard Tapia Celebration of Diversity in Computing, 2022
- Faculty Scholarship, Rocky Mountain Advanced Computing Consortium (RMACC), 2022
- Faculty Scholarship, Grace Hopper Celebration of Women in Computing, 2022
- Student Scholarship, ACM Richard Tapia Celebration of Diversity in Computing, 2019 and 2020
- Graduate ACM Student Research Competition in Tapia, 1st place, 2019
- Western Association of Graduate Schools WAGS-Proquest Distinguished Master's Thesis Award, 2019
- Great Minds in Research Award, CSU Graduate Student Showcase, 2019
- P. R. Mukherjee Award in Computer Science, Colorado State University, 2019

- Computer Science Graduate Fellow, Colorado State University, 2019
- Best Graduate Student Talk in Rocky Mountain Celebration of Women in Computing, 2018
- Nominated by CSU for WAGS-ProQuest Distinguished Master's Thesis Award in STEM, 2018
- Robert B. France Fellowship in Computer Science, Colorado State University, 2018

PUBLICATIONS

Journals

1. S. Fernandez, A. Adibfar, H. Homayouni, H. Davani, 2022. "Anomaly Detection in Wastewater Infrastructures using Supervised and Semi-supervised Learning", submitted to the Journal of Advanced Engineering Informatics.
2. H. Homayouni, S. Ghosh, I. Ray, S. Gondalia , M. Kahn 2021. "Anomaly Detection in COVID-19 Time-Series Data", *Springer Nature Computer Science journal special issue on AI for HealthCare*.
3. M.R. Keyvanpour, H. Homayouni, 2014. "Automatic Test Case Generation for Modern Web Applications Using Population-based Automatic Fuzzy Neural Network", *International Journal of Information and Communication Technology Research*, Volume 6, Issue 2.
4. M.R. Keyvanpour, H. Homayouni, and S. Zolfaghari, 2014. Population-based Automatic Fuzzy Neural Network for Online, Knowledge-based Learning", *The Modares Journal of Electrical Engineering*, Volume 14, Issue 3.
5. M.R. Keyvanpour, H. Homayouni, and H. Shirazi, 2013. "A Classification Framework for Automatic Test Case Generation Techniques for web applications" *International Journal of Information Processing and Management*, volume 4, Issue 3.
6. M.R. Keyvanpour, H. Homayouni, and H. Shirazi, 2012. "Automatic Software Test Case Generation: An Analytical Classification Framework." *International Journal of Software Engineering and Its Applications*, volume 6, Issue 4.
7. M.R. Keyvanpour, H. Homayouni, and H. Shirazi, 2011. "Automatic Software Test Case Generation." *Journal of Software Engineering*, Volume 5, Issue 3.

Book Chapter

1. H. Homayouni, S. Ghosh, I. Ray, 2018. "Data Warehouse Testing". *Advances in Computers*, Volume 112.

Conferences

1. S. Kaur, H. Homayouni, S. Justme, 2023. "High-Resolution COVID-19 X-Ray Generator", accepted as a full paper to the ACM Health Informatics and Knowledge Management Conference- HIKM.
2. G. Maurina, H. Homayouni, I. Ray, S. Ghosh, G. P. Duggan, 2022."A Methodology for Energy Usage Prediction in Unforeseen Circumstances", accepted for publication in the IEEE CogMI.
3. J. Cuomo, H. Homayouni, I. Ray, S. Ghosh, 2022. "Detecting Temporal Dependencies in Data", in British International Conference on Databases.
4. H. Homayouni, S. Ghosh, I. Ray, S. Gondalia, J. Duggan, M. Kahn, 2020. "An Autocorrelation-based LSTM-Autoencoder for Anomaly Detection on Time-Series Data", In *IEEE Big Data at the Special Session: Machine Learning on Big Data*, pp. 5068–5077.
5. H. Homayouni, S. Ghosh, I. Ray, M. Kahn 2019. "An Interactive Data Quality Test Approach for Constraint Discovery and Fault Detection", *IEEE Big Data*, Los Angeles, USA, pp. 200–205.

6. H. Homayouni, S. Ghosh, I. Ray, 2019. "ADQuaTe: An Automated Data Quality Test Approach for Constraint Discovery and Fault Detection", In *IEEE* 20^{th} *International Conference on Information Reuse and Integration for Data Science*, Los Angeles, USA, pp. 61–68.

7. H. Homayouni, 2018. "Testing Extract-Transform-Load Process in Data Warehouse Systems". In Doctoral Symposium track of the 29th *IEEE International Symposium on Software Reliability Engineering*, Memphis, USA, pp. 158–161.

8. H. Homayouni, S. Ghosh, I. Ray, 2018. "An Approach for Testing the Extract-Transform-Load Process in Data Warehouse Systems". In 22^{nd} *International Database Engineering & Applications Symposium*, Villa San Giovanni, Italy, pp. 236–245.

Presentations

1. H. Homayouni. "Anomaly Detection and Explanation in Big Data", 2021. C*omputational Science Research Center* (CSRC), San Diego State University.

2. H. Homayouni, S. Ghosh, I. Ray. "Anomaly Detection and Explanation in Big Data", 2020. *Rising Stars*, UC Berkeley.

3. H. Homayouni, S. Ghosh, I. Ray, "IDEAL: Interactive Detection and Explanation of Anomalies using Autocorrelation-based LSTM-Autoencoder for Time-Series Data" 2020, virtual poster presentation at *Rocky Mountain Advanced Computing Consortium*.

4. H. Homayouni, S. Ghosh, I. Ray, "ADQuaTe: An Automated Interactive Data Quality Test Approach", 2019, poster presentations at *Graduate Show Case*, Colorado State University, USA

5. H. Homayouni, S. Ghosh, I. Ray, "ADQuaTe: An Automated Interactive Data Quality Test Approach", 2019, poster presentations at *Grace Hopper Celebration of Women in Computing*, USA

6. H. Homayouni, S. Ghosh, I. Ray, "ADQuaTe: An Automated Interactive Data Quality Test Approach", 2019, poster presentations at *Tapia Celebration of Diversity in Computing*, Sandiego, USA

7. H. Homayouni, S. Ghosh, I. Ray, 2019. "ADQuaTe: An Automated Data Quality Test Approach for Constraint Discovery and Fault Detection", poster presentation at *Rocky Mountain Advanced Computing Consortium*, Boulder, USA.

8. H. Homayouni, S. Ghosh, I. Ray, 2018. "Using Autoencoder to Generate Data Quality Tests", paper presentation at *Rocky Mountain Celebration of Women in Computing* (RMCWIC), Denver, USA, November 2–3.

9. H. Homayouni, S. Ghosh, 2016. "A Study of Evosuite as an Automatic Test Case Generation Approach to Kill First Order Mutants", paper presentation at *Rocky Mountain Celebration of Women in Computing*, Salt Lake City, USA.

GRANTS

Co-PI, **NIH-PA-20-195** (under review), $441,967. "Parsimonious low dimensional biomimetic pump for analysis of hydraulic and rheologic features modifying intraocular and episcleral venous pressure appraised by sparse high throughput flow modeling"

PI, **NSF-CAREER-CISE-SaTC** (under review), $694,422, 2022 Privacy-preserving Multimodal Conditional Generative Adversarial Network for Synthetic Data Generation

Co-PI, **NSF-FM** (unawarded), $389,202, 2022
Cloud-Based Real-Time Defect Monitoring of Laser Powder Bed Fusion Using Multi-Sensor and Physics-Informed Deep Learning

Co-PI, **NIST-MBAMGP-01** (unawarded), $900K, 2022
Developing a data-driven cyber physical manufacturing system for metal additive manufacturing using physics-informed machine learning

Co-Director (awarded) **Division of Research and Innovation (DRI) Equipment Funding** $102,817, 2022
Supporting Deep Learning-based Research Activities with GPU-equipped High Performance Server Computers

PI, **Google Cloud credit grant to support COVID-19 research** (awarded), $2,400, 2020
Assessing the Integrity of COVID-19 Data

RESEARCH PROJECTS

Available at https://github.com/hajarhomayouni

Developing a data-driven cyber physical manufacturing system for metal additive manufacturing using physics-informed machine learning *Present*

- Develop a cloud-based data management system to store and manage sensor data obtained from multiple metal printers.
- Develop ML models to detect anomalies in the printers' data.
  Collaborator: Dr. John Kang, Mechanical Engineering Department at SDSU
  Submited to NIST Additive Manufacturing program

Predict the Future of COVID-19 *Nov 2021–Present*

- Predict new COVID-19 variants
- Predict the super-spreaders of the virus
- Predict the most "at-risk" population
  Collaborators:
  Dr. Henao Tamayo, Department of Microbiology, Immunology, and Pathology, Colorado State University
  Dr. M-Irfan Suleman, School of Medicine, Johns Hopkins

Climate Change Forecasting *Oct 2021–Present*

- Climate forcasting in San Diego based on the Global Circulation Model (GCM) data
- Anomaly detection from videos of underground infrastructures
- Automatic detection of flood situations in San Diego using satellite image data
  Collaborator: Dr. Hassan Davani from Department of Civil and Environmental Engineering at SDSU

Assessing the Integrity of COVID-19 Data *Jun 2020–August 2021*

- Using our proposed data quality test framework to validate COVID-related data
- Validating COVID-19 patient records in the Anschutz Health Data Compass
- Validating COVID-19 records in the Johns Hopkins, New York Times, and COVID Tracking project datasets
  Collaborators: Dr. Michael G. Kahn from the University of Colorado Anschutz and Dr. Saul Lozano from the Centers for Disease Control and Prevention (CDC)
  Awarded the Google Cloud Credit Grant to support COVID-19 research

Published in Springer Nature Computer Science journal special issue on Artificial Intelligence for Health-Care, 2021

Automatic Identification of Grouping Attributes in Sequential Data *Sep 2020–Present*

- Designing a statistical-based model that automatically identifies grouping attributes (i.e., attributes by which grouping results in temporal subsequences of data) in sequential data
- Evaluating the effectiveness of the approach using real-world sequential data

Effects of COVID-19 on Energy Consumption *May 2020–Present*

- Designing different Machine Learning-based models to predict the energy delivered to different premises
- Analyzing the energy consumption before and after the COVID-19 pandemic in the city of Fort Collins
  Collaborator: Jerry Duggan from Energy Institute at Colorado State University

Optimal Selection of Least Significant Bits (LSbs) from Electronic Control Unit Packets *Feb 2020–Present*

- Designing an LSTM-based approach that automatically identifies the LSbs from an input packet sequence without deciphering the data
- Evaluating the approach using datasets from heavy vehicle data generated at the University of Tulsa
  Collaborator: Dr. Jeremy Daily from the Department of Systems Engineering at Colorado State University

Anomaly Detection and Explanation in Big Data *Feb 2018– Sep 2020*

- Proposed a data quality test framework for constraint discovery and anomaly detection and explanation
- Instantiated the framework for non-sequence and sequence data
- Evaluated the instantiations using datasets from the Anschutz Health Data Compass, CSU Plant Diagnostic Clinic database, CSU Energy Institute, and UCI ML repository
  Resulted in two conference publications, and a conference and a journal submission

Testing the Extract-Transform-Load processes in data warehouses *Aug 2016– Dec 2018*

- Proposed an approach that automatically generates balancing tests to check for discrepancies between the data stored in the source databases and the target warehouse
- Evaluated the balancing test approach using data in the Health Data Compass data warehouse
  Resulted in two conference and a book chapter publications
  Won the Western Association of Graduate Schools (WAGS)/ProQuest Distinguished Master's Thesis Award

Automatic Test Case Generation for Web Applications *2010–2013*

- Proposed a novel neural network called Population-based Automatic Fuzzy for automatic generation of test cases for web Applications
  Resulted in five international journals

TEACHING EXPERIENCE

- Instructor, San Diego State University, Principles and Techniques of Data Science, 2021, 2022
- Instructor, San Diego State University, An Introduction to Data Science, 2022
- Adjunct Faculty, Colorado State University, Business Visual Application Development, 2019
- Graduate Teaching Assistant, Colorado State University, Object-Oriented problem solving with Java, 2016
- Graduate Teaching Assistant, Colorado State University, Software Engineering, 2015

- Lecturer, Alzahra University, Database Management System, 2014
- Lecturer, Alzahra University, System Architecture, Electronic circuits, 2013
- Lecturer, Feyz University, Database Management Systems, 2013
- Graduate Teaching Assistant, Alzahra University, Electrical circuits, 2010–2011

SERVICES

- Chair, Undergraduate Consortium at KDD (KDD-UC), 2023
- Point of Contact, Computing Alliance of Hispanic-Serving Institutions (CASHI), 2022–2023
- Reviewer, National Science Foundation, 2023
- Session Chair, IEEE CIC, 2022
- Design and Evaluate SDSU Masters Exam, Fall 2021 and 2022
- Mentor, SDSU Faculty Development Program, Fall 2022
- Advisor, SDSU STEM Pathways, Summer 2022
- Associate Editor, Journal of Machine Learning with Applications (MLWA), 2022
- Reviewer, Journal of Machine Learning with Applications (MLWA), 2022
- Program Committee, IEEE International Conference on Collaboration and Internet Computing (CIC), 2022
- Judge, Applied Computational Sciences and Engineering Student success (ACCESS), 2022
- Reviewer, Scientific Reports, 2022
- Faculty search committee, San Diego State University, 2022
- Reviewer, National Science Foundation, 2022
- Session Chair, IEEE CIC/CogMI/TPS Joint Conferences, 2021
- Reviewer, IEEE Transactions on Power Systems, 2021
- Reviewer, The ACM India Joint International Conference on Data Science & Management of Data (CODS-COMAD), 2022
- Reviewer, 41^{th} IEEE International Conference on Distributed Computing Systems (ICDCS), 2021
- Reviewer, 11^{th} ACM Conference on Data and Application Security and Privacy (CODASPY), 2021
- Session Chair, 2^{nd} Special Session on Machine Learning on Big Data in IEEE Big Data, 2020
- Reviewer, 16^{th} International Conference on Information Systems Security, India, 2020
- Reviewer, IEEE Transactions on Services Computing, 2020
- Reviewer, 19^{th} IEEE International Conference on Software Quality, Reliability, and Security, Bulgaria, 2019
- Reviewer, 3^{rd} Workshop on Attribute Based Access Control, Arizona, USA, 2018
- Reviewer, 23^{rd} ACM Symposium on Access Control Models and Technologies, Indianapolis, USA, 2018

INDUSTRY EXPERIENCE

Health Data Compass, University of Colorado Denver *Summer 2016 and 2017*
Roles: Researcher, developer, and tester

- Improved the design of the existing ETL process in Health Data Compass, developed the new design using BigQuery jobs and tested using my balancing tests against data from two hospitals including UCHealth and Children's Hospital of Colorado
- Established OHDSI data analysis tools for the Health Data Compass on Google BigQuery
- Received certificates for HIPAA training (CITI and Skillsoft)

Noor Company, Tehran, Iran *2013–2015*
Role: Project manager and developer

- Headed a group of five software developers
- Developed and customized Zotero open-source application for organization requirements

Operating System Security Lab (OSSL), Alzahra University, Tehran, Iran *2012–2013*
Role: Tester

- Evaluated Linux Kernel using Autotest tool

IT Research Center, Tehran, Iran *2011–2012*
Role: Researcher and developer

- Studied and developed machine learning algorithms for artificial games

Subsea R&D Center, Isfahan University of Technology, Isfahan, Iran *2008–2009*
Role: Developer

- Developed a web application for aerology dataset management
- Developed a web application for ministerial letters management

ORGANIZATIONS

- Computing Alliance of Hispanic-Serving Institutions (CASHI), Faculty member
- SDSU CSRC Data Science, Faculty member
- Association for Computing Machinery (ACM), Faculty member
- SDSU STEM Pathways, Faculty member
- Data Science Alliance (DSA), Faculty member
- Association for Computing Machinery (ACM), Student member
- Upsilon Pi Epsilon (UPE), President
- Campus Outreach Committee at CSU, Student member
- Iranian Student Organization (ISO) at CSU, Advisor