NILOOFAR A.ABIANE
Senior Data Scientist, Ph.D. in Computer Science

✉ niloofar.abiane@gmail.com ✉ n.a.abiane@gmail.com 📟 (929)310-4410

Webpage: https://web.njit.edu/na396/ LinkedIn: https://www.linkedin.com/in/na396/

GitHub: https://github.com/na396

Last update: October 13, 2023

EDUCATION

New Jersey Institute of Technology (NJIT) — Newark, NJ
Ph.D. in Computer Science — *September 2017 - April 2023*
- Emphasis on Data Science, Statistical Modeling and Analysis, Biostatistics, Machine Learning, Clustering, Network Embedding, Algorithm Development in Computational Systems Biology.

Massachusetts Institute of Technology (MIT) — Cambridge, MA
MicroMaster in Statistics and Data Science — *May 2020 - October 2021*
- Emphasis on Artificial Intelligence, Probability, Statistics, Data Analysis, and Machine Learning.

University of Tehran (UT) — Tehran, Iran
M.Sc. in Computer Science, Minor Algorithms and Computation — *September 2011 - January 2014*
- Strong emphasis on Data Science, Data Analysis, Machine Learning, Classification, Biostatistics, Statistical Modeling and Analysi in Computational Systems Biology.

Relevant Courses: Data Structures & Algorithm Design, Advanced Algorithms, Graph & Network Algorithms, Artificial Intelligence, Data Mining, Machine Learning, Fundamental of Probability & Statistics, Linear Algebra, Data Management System Design.

SKILLS

- **Recent Experience in**: R, Python, MATLAB, SQL, LaTeX
- **Others**: Perl, JAVA, C++
- **Technology**: Linux, git, Microsoft SQL server, SPSS, MiniTab
- **Data Science Library**: Pandas, NumPy, Pytorch, SciPy, Scikit-learn, dplyr, ggplot2, tidyr, mlr3
- **Statistical Skills**: Hypothesis Testing & Confidence Interval, Maximum Likelihood Estimation (MLE), Various Test Statistics, Bayesian Statistics, Linear Regression, Generalized Linear Model (GLM), Experiment Design, Omitted Variable Bias, Control Variable, Treatment Effect.
- **Machine Learning Skills**: Linear Regression, Naive Bayes, K Nearest Neighbor (or K Nearest Mean), Logistic Regression, Feed-forward and Backpropagation Neural Networks (NN) or Multi-Layer Perceptron (MLP), Support vector Machine (SVM), Supervised Self Organized Map (SOM), Bagged Decision Stump Tree, K-means Clustering, Unsupervised Self Organized Map (SOM), Principal Component Analysis (PCA), Anomaly Detection, and Recommender System for Movies, kernelization, regularization.
- **Deep Learning Skills**: Autoencoders, Regularization, Feed-Forward Neural Networks, Multi-layered Perceptron (MLP), Convolution Neural Networks (CNN), Recurrent Neural Networks (RNN), Gated Recurrent Units (GRUs), Long Short Term Memory (LSTM) Cell, Word Vectorization, Encoder-Decoder Models, Transformers, Attention Mechanism.
- **Data Base Skills**: Conceptual Modeling, The Relational Model, The Standard Database Language SQL
- **Computation Biology Skills** : networking, DNA-microarray, RNA-seq, Gene Expression, Next Generation Sequencing (NGS), single-cell RNA-seq, Gene Co-Expression Networks
- **Diverse background in Computer Science, Probability & Statistics, Math, and Biology allows me to communicate to a wide scientific and general audience and begin contributing to any group immediately**.
- **I have worked in many places. I am self-driven and can readily learn and adapt to new disciplines, areas, or environments and start pushing for real results quickly.**

WORK EXPERIENCE

- **JP Morgan & Chase Co. - <https://www.jpmorganchase.com/>** NYC, NY
  Senior Data Scientist *March 2023 - present*
    - Capture latent structure, novelties, and derive transformations & enhancement using AI, machine learning, and deep learning techniques.
    - Give technical talks to the client about the products & solution we proposed at JPM Chase Co.

- **Seagate Technology - <https://www.seagate.com>** Remote, CA
  Data Scientist - Intern *May 2022 - December 2022*
    - **Project**: Classify and estimate hard drive failure riskiness using supervised learning, and identify essential hard drive failure patterns using unsupervised learning.
    - **Steps**: Feature Engineering, Feature Selection & Dimensionality Reduction, Unsupervised Learning, Pattern identification, Supervised Learning, and model prediction.
    - **Impact**: Minimize parentage of returning hard drives and the associated cost.

- **New Jersey Institute of Technology - <https://www.njit.edu>** Newark, NJ
  Adjunct Professor & Research Assistant, in Deep Machine Learning & Statistics *September 2017 - May 2023*
    - Proposed a novel preprocessing step on raw data of gene expression using statistical tools.
    - Improving the network construction using linear algebraic tools, biostatistics, and machine learning.
    - **Converting the unsupervised problem into semi-supervised using self-training approach**
    - Proposing a novel framework for network construction in gene co-expression networks.
    - Implementing a package for gene co-expression network construction in R.
    - paper link and package repository

- **CVS Health - <https://www.cvs.com>** Remote, RI
  Data Science Intern *Summer 2021*
    - **Project**: Unsupervised model for customer segmentation and identified high-risk customers.
    - **Model Selection**: Selected hierarchical clustering, - complete linkage model. Performed several clustering models and analyzed the results. Competently, developed a model that evaluates and distinguishes the customers based on their prescription behaviors using R & Pyhton. Successfully, identified a cluster of high-risk customers.
    - **Prepossessing Steps**: Data Cleaning& Feature Engineering, Feature Selection & Dimensionality Reduction.
    - **Impact**: Minimize the potential risk of regulatory action against CVS by identifying high-risk customers.

- **Ershad University of Damavand - <http://e-damavandihe.ac.ir>** Tehran, Iran
  University of Applied Science and Technology - <https://www.uast.ac.ir/en> Tehran, Iran
  University Lecturer *Fall 2013, Fall 2015 - Summer 2016*
    - Led classes with 40+ participants for Discrete Mathematics, Data Base Management System Design, Operating Systems, and Boolean Circuits courses.
    - Supervised three undergraduate thesis students in data structures and algorithm design. Successfully, carried efficient management, organizational, teamwork, and supervisory skills.

- **Parseh Institute - <https://parseh.ac.ir>** Tehran, Iran
  Head of Design Algorithm Team *February 2008 - September 2012*
    - **Publishing books:** Led the team for Data Structures and Algorithm Design team resulted in co-authored three books and found significant gains in different metrics across analyzing the algorithm efficiency and problem-solving in Divide & Conquer, Dynamic Programming, Greedy, and Graph algorithms approaches using different data structures.
    - **Teaching and Design Mock Testing**: Taught classes with 50+ participants for data structures & algorithm design and designed weekly mock questions similar to what was expected to see on Nationwide Graduate University Entrance Exam (NGUEE) for data structures & algorithm design course for 4000+ participants. Scrutinized Computer Science & Computer Engineering-hardware mock testing for errors and oddities

prior to administering the exam. Strengthen problem-solving and analytical thinking by creating a road map to hack the questions.

– **Leadership:** Led and directed the Data Structures and Algorithm Design team as the head of the team. New-hire mentor in Parseh. Successfully managed projects from inception to on-time completion; balancing tradeoffs and coordinating teamwork. Carried efficient management, organizational, teamwork, and supervisory skills resulting in publishing books and more.

PUBLICATION

- Ehsan Beikihassan, Amy K. Hoover, Ioannis Koutis, Ali Parviz, **Niloofar Aghaieabiane** *Resource-constrained knowledge diffusion processes inspired by human peer learning*, **26th European Conference on Artificial Intelligence**, ECAI (2023), Published.

- **Niloofar Aghaieabiane**, Ioannis Koutis, *SGC: A semi-supervised pipeline for gene clustering using the self-training approach in gene co-expression network*, arxiv, (2022), Preprint.

- **Niloofar Aghaieabiane**, Ioannis Koutis, *A novel calibration step in gene co-expression network,* Frontier in Bioinformatics 1 (2021), 66. Published.

- **Niloofar Aghaieabiane**, Henk Koppelaar, P. Nasehpour, *A novel algorithm to determine the leaf (leaves) of a binary tree from its preorder and postorder traversals,* JAC 49 (2) (2017), 1-11. Published.

- **Niloofar Aghaieabiane**, Henk Koppelaar, P. Nasehpour, *An improved algorithm to reconstruct a binary tree from its inorder and postorder traversals*, JAC 49 (1) (2017), 93-113. Published

- **Niloofar Aghaieabiane**, Henk Koppelaar, P. Nasehpour, PREPOS*: A novel algorithm to reconstruct a binary tree(s) from its preorder and postorder traversals and determining type of each node*. Published.

- **Book:** "*Data Structures and Algorithm Design*"**,** Book in Persian   Parseh Press, Tehran, Iran
Supervised by Dr. Mohsen Tourani   *May 2011*

- **Book:** "*Key Solutions Introduction to Algorithms*"**,** Book in Persian   Parseh Press, Tehran, Iran
Supervised by Dr. Mohsen Tourani   *Jun 2010*

SELECTED RESEARCH EXPERIENCE

- **Princeton University**   Princeton, NJ
Research in Systems Biology using Biostatistics under Prof. Wingreen   *March 2015 - Fall 2018*
  – Investigated the pairwise association between physical gene properties and metabolite flux.
  – Created a dataset of structural properties of the genes; gene length, gene frequencies, adaptation index, GC-content and GC-bias.

- **Laboratory of Systems Biology and Bioinformatics (LBB), University of Tehran**   Tehran, Iran
M.Sc. Graduate Thesis Under Dr. Ali Masoudi-Nejad   *Spring 2012 - October 2014*
  – Used machine learning algorithm to classify transposons sequences, identify transcription binding sites, and identify promoter polymerase III using likelihood function.
  – Researched in gene regulatory networks to find motifs and reduce the network to its kernel using ordinary differential equations and statistical models respectively.

PRESENTATION

- *"Classifying Transposons Using Machine Learning Algorithms"*, M.Sc. Thesis Defense, University of Tehran (Talk) - Tehran, Iran, *January 2014*.

AWARDS & HONORS

- CVS Award for Internship Completion, RI, Summer 2021.
- Graduate Stipend Award, Graduate Tuition Award, New Jersey Institute of Technology, NJ, Fall 2017 - Spring 2022.
- Elected for Computer Science Department Representative in Graduate Student Association (GSA), New Jersey Institute of Technology, NJ, Fall 2018 - present.

- Offered to study Ph.D. at Institute of Research in Fundamental Sciences (IPM) as an Exceptional Talent Without the Nationwide Graduate University Entrance Exam, Tehran, Iran, Spring 2013.
- Offered to study Ph.D. at University of Tehran (UT) as an Exceptional Talent Without the Nationwide Graduate University Entrance Exam, Tehran, Iran, Spring 2013.
- Ranked 3th among 50+ students in Computer Engineering program, University of Tehran (UT), Tehran, Iran, Spring 2010.
- Graduate Tuition Award (IRR20,000), University of Tehran, Tehran (UT), Iran, Fall 2011 - Fall 2013.
- Ranked 34th among almost 100,000 in Nationwide Graduate University Entrance Exam, Iran, Spring 2010.

PARTICIPATION IN WORKSHOPS AND CONFERENCES

- **Fintech Event at Microsoft, Workshop**   New York, NY
  Microsoft   *October 2019*
- **CSLAW: ACM Inaugural Symposium on Computer Science and Law Conference**   New York, NY
  Association for Computing Machinery (ACM)   *October 2019*
- **Oracle Modern Marketing Express**   New York, NY
  Oracle Team   *October 2019*
- **AWS + Databricks HLS Dev Day Workshop**   Morisstown, NJ
  The Databricks Team   *September 2019*

COURSE CERTIFICATION

- **Neural Networks and Deep Learning**   May 2021
  DeepLearning.AI
- **Machine Learning with Python - From Linear Models to Deep Learning**   December 2020
  Massachusetts Institute of Technology (MIT)
- **Probability - The Science of Uncertainty and Data**   December 2020
  Massachusetts Institute of Technology (MIT)
- **Fundamentals of Statistics**   September 2020
  Massachusetts Institute of Technology (MIT)
- **Data Analysis in Social Science**   September 2020
  Massachusetts Institute of Technology (MIT)
- **Fundamentals of Scalable Data Science**   February 2020
  IBM
- **Machine Learning**   August 2019
  Stanford University