244 New Computer Science Building
Stony Brook University
Stony Brook, NY 11794
☎ (631) – 974 – 8156
✉ falmodaresit@cs.stonybrook.edu
https://www.linkedin.com/in/fataltes
https://github.com/fataltes

Fatemeh Almodaresi

Research Interests

Machine Learning, Pattern Recognition, and Data Analysis
Computational Biology

Education

2015-present **Ph.D**, *Computer Science Department, Stony Brook University (SBU)*, NY, USA.
Advisor – Prof. Rob Patro

2009-2011 **MS**, *School of Computer Engineering, Iran University of Science and Technology (IUST)*, Tehran, Iran.
Advisor – Prof. Jahed Motlagh

2004-2009 **BS**, *School of Electrical & Computer Engineering (ECE), University of Tehran*, Tehran, Iran.

Research Projects

Apr. 2016 - present **A Tool to Optimize Colored de Bruijn Graph Storage**, Computational Biology Lab., SBU,
Under the supervision of Prof. Rob Patro.
This tool provides a new data structure to store and query colored de Bruijn graph that in case of large data sets improves storage by more than twenty times compared to state-of-the-art tools without hurting performance of queries.

Nov. 2016 - Mar. 2016 **RapClust**, *"Rapid Clustering" project*, Computational Biology Lab., SBU,
Under the supervision of Prof. Rob Patro
`https://github.com/fataltes/RapClust_OrphanReads/tree/orphan`.
RapClust is a fast tool for clustering contigs of a de novo transcriptome assembly. In RapClust2 we focus on improving the accuracy of clustering by making use of orphaned reads.

Jan 2017 **DDFactorization**, *"Improving Transcript Abundance Estimation Using Eq. Class Factorization (A Data Driven Approach)" project*, Computational Biology Lab., SBU,
Under the supervision of Prof. Rob Patro
`https://github.com/COMBINE-lab/salmon/tree/factorizations`.
Exploring the effect that factorizations of the likelihood function can have on accurately estimating transcript abundance, and designing new efficient and accurate factorizations.

Aug 2016-Jan 2017 **MLDD**, *"Multi-Level Distribution Detection" project*, Data Science Lab., SBU.,
Under the Supervision of Professor Andrew Schwartz.
Using statistical tests and classification models such as NaiveBayes we show how distribution of NLP features in social media changes in different levels of analysis (county, user, and message). This can highly affect prior assumptions for further text analysis.

2013-2014 **AutismFD**, *"A game to improve face emotion detection in children with Autism" project*,
Under the supervision of Professor Pouretemad.
Beside collaboration with psychology students to design the method, I also implemented the idea as a tool in C# language. This package was used in a treatment center to collect validation data.

Publications

[1] M. Zakeri, A. Srivastava, F. Almodaresi, and R. Patro. Data driven likelihood factorizations improve lightweight transcript abundance estimation. April 2017. Accepted in ISMB 2017.

[2] F. Almodaresi, L. Ungar, V. Kulkarni, M. Zakeri, S. Giorgi, and H.A. Schwartz. On the distribution of lexical features in social media. March 2017. Accepted in ACL 2017.

[3] F. Almodaresi, P. Pandey, and R. Patro. Rainbowfish, a succinct colored de bruijn graph. June 2017. Accepted in WABI 2017.

[4] F. Almodaressi, N. Mozayani, M.R. Jahed, and M. Ahmadi. The relation between friendship and academic

performance in university students: Role of personality. *Journal of Technology of Education*, 8(2):81–92, 2013. In Persian.

Work Experiences

Jun-Aug 2016 **Member of the NLP Team**, *Third Frederick Jelinek Memorial Summer Workshop (JSALT)*, Baltimore.
JSALT is a well-known summer workshop in Language and Speech organized by JHU each year. .
During the project, we worked on analyzing and forecasting social media user's psychological state based on their language in their posts using statistical methods such as significance tests and time series models such as ARMA and ARIMA.

Jan-Aug 2015 **Senior Designer and Developer**, *Nexeven AB*.
Nexeven AB is a Swedish company and a niche player in the online video broadcasting field.

2011-2015 **Team Supervisor, Senior Designer and Developer**, *Tosan Intelligent Data Miners Co. (TIDM)*, Data Mining Development Team.
TIDM is the first solution provider for fraud detection and anti-money laundry in banking section in Iran, a Subsidiary of Tosan Company.

- **Customer Relationship Management System** [2014]
  In this project we use statistical and data mining methods to calculate customer's RFM, CLV, and churn probability.
- **Data mining Module, Operational Intelligence System** [2014]
  This module, developed in PLSQL, uses Statistical and Mining Methods such as regression models, error functions, k-means, and SVM to detect fraudulent transactions online in the stream of transactions.
- **Customer Name Similarity Detection Module** [2013]
  As a part of Anti-money Laundry System, this module uses natural language algorithms to detect accounts with similar names. The whole system is developed in PLSQL and now operational in many private banks in Iran including Eghtesad-Novin and Ansar Bank.
- **Unsupervised Fraud Detection System, Version 1 & 2** [2011-2014]
  Version 1 which is fully designed and developed by myself is now operational in Saman Bank, Ansar Bank, and Mehr-e-Eghtesad Bank in Iran. Version 2 is now installed in Eghtesad-Novin Bank.

2009-2011 **Java and UI Developer** , *Tosan Co.*
Tosan Company is a pioneer company for total banking solutions with more than fifteen Iranian financial institutes in its customer list. As a member of a team of nearly 20 people, I participated in developing the UI of Internet banking system.

Honors & Awards

2016 **CS Department Best TA Award**.

2015 **Special CS Department Chair Fellowship**.

Teaching Experiences

Spring 2017 **Teaching Assistant**, *Machine Learning*, Stony Brook University.

2013 **Teacher**, *C++ Programming Language*, Farzanegan High School [NODET].

2013 **Teacher**, *Developing simple motion detection algorithms in MATLAB*, Farzanegan High School [NODET].

Fall 2008 **Teaching Assistant**, *Artificial intelligence*, University of Tehran.

Skills

Programming Languages **Python (expert), Java Core(expert), C++ (familiar), R (familiar), MATLAB (familiar), NetLogo (familiar), C# (familiar)**.

Libraries and Frameworks **Standard Python Libraries (numpy, pandas, scipy.stats, sklearn), Spring Framework, Hibernate, Play Framework**.

Databases **Oracle (expert), MySQL (expert), MongoDB (familiar)**.

Other Tools **IntelliJ IDEA, Pycharm, Jupyter Notebook, git, Atlassian Jira, Atlassian Confluence, Thought-Works Go, Anaconda Platform**.