Mohsen Botlani

Resume

📞 *(813) 817-4190*
✉ *m.botlani@gmail.com, mohsenb@mail.usf.edu*
🖳 *//mohbot.github.io/researchProfile/*

Summary

I am a data scientist and have a PhD in Cell and Molecular Biology, particularly interested in application of **machine learning and data mining** in **bioinformatics, computational biology and health informatics**. I analyze gigabytes of data and write algorithms to mine in these huge data for meaningful information and tie the results back to the biology.

Education

8/12 - 8/17   **Ph.D. in Cell and Molecular Biology (Computational Biology)**, Department of Cell Biology, Microbiology and Molecular Biology, University of South Florida (USF), Tampa, Florida. Advisor: Dr. Sameer Varma

9/03 - 9/08   **B.Sc. in Materials Science and Engineering**, Department of Material Science and Engineering, Isfahan University of Technology, Isfahan, Iran.

Research Interests

Computational Biology, Bioinformatics, System Biology, *omics Data Analysis, Next Generation Sequencing (NGS) Data Analysis, Pharmacogenomics, Computational Oncology, Health Informatics, Machine Learning (Deep Learning), Data Mining, Big Data, Statistical Modeling, Network Analysis, Bayesian Inference, Statistical Mechanics

Work and Research Experiences

4/18 - present   **Bioinformatics Staff Scientist** at M2Gen-Moffitt Cancer Center, *Tampa, Florida*

M2Gen® is a health informatics solutions company focused on accelerating the discovery, development and delivery of precision medicine. Our Company was founded in 2006 at the Moffitt Cancer Center to operationalize the Total Cancer Care® Protocol, a unique approach to studying patients throughout their lifetime. In this effort, we partner with the nation's leading cancer centers via the Oncology Research Information Exchange Network (ORIEN).

Development of scalable pipelines for RNAseq and Whole Exome sequencing (WES) NGS data analysis both upstream and downstream

- Processing raw WES via the Sentieon and RNAseq STAR, STAR-Fusion and RSEM on AWS machines supported by DNAnexus

- Automated QC pipeline development via integration with DNAnexus' API in the cloud to check the output files of both sequencing and alignment for WES and RNAseq data

- Development of QC Metrics and Classifiers for QC of raw and processed NGS data by statistical and machine learning methods (Gender Prediction, Tumor-Normal concordance, RNA-WES concordance, Tissue of orogin prediction)

- Pipeline development for Copy Number Variant detection (CNVKit) for different cancer types

- Internal Tandem Duplication (IDT) detection (Pindel) on WES data for a specific cohort of patients

- Microsatellite instability (MSI) analysis (MSIsensor) for a specific cohort of patients

- Multi-omics analysis for downstream analysis as well as QC control considering effect of latent variables such as cancer types, preservation methods, batch effects

- Pan-cancer classification and exploration for gene signatures, using gene expression data as well a gene mutation data

- Testing different hypothesis for pharmaceutical companies as well as cancer researchers of ORIEN members

1/18 - 3/18   **Data Scientist** at WellCare, *Tampa, Florida*

- Customer retention predictive model development by using machine learning methods (Gradient boosting, Logistic regression)

- Development of all-cause mortality rate predictive models of members based on their age, gender and common laboratory tests by machine learning methods (Gradient boosting, Random forest, Logistic regression)

7/17 - 1/18 **Research Assistant** at Morsani College of Medicine, Dep. Molecular Medicine, USF (Dr. George Blanck), *Tampa, Florida*

- Pipeline development for identification of systemic mutations (in whole exome sequence (TCGA)) associated with cancer.

- Expression analysis of secretome of cancer cells by using deep learning to characterize chronic acidosis of tumor microenvironment.

- Expression analysis of proteome of well differentiated and dedifferentiated liposarcoma.

2/15 - 3/15 **Graduate Research Assistant** at Microbiology Lab at USF, *Tampa, Florida*

- Motif co-occurrence analysis on proteome of S. aureus, B. subtilis and E. coli.

8/12 - 8/17 **Graduate Research Assistant** at Computational Biophysics Lab at USF, *Tampa, Florida*

- Method development to model dynamic allostery in proteins enabled by machine learning (SVM) and graph theory on molecular dynamics (MD) simulation models.

- Modeling of allosteric signaling of Nipah virus binding, fusion by applying the developed method on MD models.

- Development of methods to quantify ensemble changes by a direct comparison of MD models enabled by SVM and deployment of method.

- Investigating the effect of different MD force fields on simulating the D-band Length of type I collagen.

8/11 - 8/12 **Research Assistant** at Bioinformatics Research Center at Isfahan University of Medical Sciences *Isfahan, Iran*

- Molecular docking of drug like ligands into G-protein coupled receptor (GPCR) specifically human adenosine receptors.

8/10 - 8/11 **Research Assistant** at Isfahan University of Technology, *Isfahan, Iran*

- Modeling of freezing behavior and structures of cobalt nanoclusters by MD simulation.

2/10 - 8/10 **Data Analyst** at Dadeh Pardaz Esfahan, *Isfahan, Iran*

- A comparative performance analysis of the genetic algorithm and backward propagation in neural network predictive models.

8/8 - 8/10 **Data Analyst** at Mobarakeh Steel Company, *Isfahan, Iran*

- Developing models to predict mechanical properties of steel products by Bayesian Neural Networks.

Technical Skills

Prog. Lang. | Python, R, Matlab, Julia, C++, C#, Bash scripting, SQL, Cloud-app development

Software | Python(Numpy, Pandas, scikit-learn, H2O, PySpark, Keras, TensorFlow, Scipy, StatsModels, Jupyter, seaborn), MATLAB(neural networks, genetic algorithm and optimization, bioinformatics, parallel computing, symbolic math, curve Fitting). LaTeX

Bioinformatics | Biopython, bowtie2, Pysam, samtools, BCFtools, bedtools, maftools bcbio-nextgen(bwa, freebayes, GATK), Sentieon, CNVKit, DNAnexus, GDC, UCSC genome browser, IGV, cbioportal

Computational Biology | Gromacs, LAMMPS, Pdynamo, AutoDock, Schrödinger, CACHe, Moldock, Modeller and Materials Explorer

Teaching Experiences

Biology of Human, Biology 1 lab, General Genetics

Honors and Awards

Media Coverage of My Last Publication, eurekalert.org, sciencedaily.com

Machine learning techniques may reveal cause-effect relationships in protein dynamics data

3rd Best Poster Presentation in *Intelligent Systems For Molecular Biology (ISMB)*, 2016 Orlando

Summer Scholarship, 2013, USF

Best researcher in R&D Metal Forming Branch of *Mobarakeh Steel Company*, 2010, Iran.

Publications

Book Chapter, more than 8400 Downloads

- **M. Botlani**, M. R. Toroghinejad, "Application of Bayesian Neural Networks to Predict Strength and Grain Size of Hot Strip Low Carbon Steels", *Artificial Neural Networks - Industrial and Control Engineering Applications, Intech, Austria* , 2010.

Journal Papers

- **M. Botlani**, G. Blanck, "Systematic analysis of germline mutations of tumor suppressor genes impacting earlier age of onset in 11 cancer types. *In preparation.*

- **M. Botlani**, A. Siddiqui, S. Varma, "Machine Learning Enabled Approach to Incorporate Multi-State Information in Molecular Modeling of Dynamic Allostery: A Case Study of the PDZ2 Domain", *Journal of Chemical Physics*, 2018.

- P. Dutta, A. Siddiqui, **M. Botlani** and S. Varma, Inter-monomer Rearrangements in Allosteric Stimulation of Nipah Virus Host Binding Protein. *Biophys. J.*, 2016.

- S. Varma, **M. Botlani**, J. Hammond, H.L. Scott, J.P.R.O Orgel and J.D. Schieber, "Effect of Intrinsic and Extrinsic Factors on the Simulated D band Length of Type I Collagen", *Proteins*, 2015.

- C.N. Krute, R. K. Carroll, F.E. Rivera, A. Weiss, R.M. Young, A. Shilling, **M. Botlani**, S. Varma, Bill J. Baker and Lindsey N. Shaw, "The Disruption of Prenylation Leads to Pleiotropic Rearrangements in Cellular Behavior in Staphylococcus aureus", *Molecular Microbiology*, 2014.

- P. Dutta, **M. Botlani**, S. Varma, "Water Dynamics at Protein-Protein Interfaces: A Molecular Dynamics Study of Virus-Host Receptor Complexes", *J. of Phys. Chem. B*, 2014.

- S. Varma, **M. Botlani** and R.E. Leighty, "Discerning intersecting fusion-activation pathways in the Nipah virus using machine learning", *Proteins*, 2013.

- **M. Botlani**, M. R. Toroghinejad, "Application of Bayesian MLP Networks and RJMCMC to Predict the Grain Size of Hot Strip Low Carbon Steels", *J. Ser. Chem. Soc.*, 2010.

- **M. Botlani**, M. R. Toroghinejad, A. R. Key Yeganeh, "Modeling the Yield Strength of Hot Strip Low Carbon Steels by Artificial Neural Network", *J. of Mat. and Des.*, 2009.

- **M. Botlani**, M. R. Toroghinejad, A. R. Key Yeganeh, "Modeling the Yield Strength of Hot Strip Low Carbon Steels by Artificial Neural Network", *ISIJ*, 2009.

Conference presentations

- **M. Botlani**, S. Varma, "Machine learning enabled approach to determine correlations in ensemble shifts: a case study of the PDZ2 domain" *Biophys. Soc. Meet., New Orleans*, 2017.

- **M. Botlani**, A. Siddiqui, S. Varma., "Qantifying Conformational Ensemble Changes in Proteins Using Inverse Machine Learning", *Intelligent Systems For Molecular Biology(ISMB), Orlando*, 2016.

- **M. Botlani**, S. Varma, R.E. Leighty, "Discerning intersecting fusion-activation pathways in the Nipah virus using machine learning", *Biophys. Soc. Meet., San Francisco*, 2014.

- S. Varma, **M. Botlani**, R.E. Leighty, P. Dutta, "Using machine learning to understand the allosteric regulation of biomolecules", *Gordon research conference, Maryland*, 2014.