Md Abid **Hasan**

PH.D. CANDIDATE · GRADUATE STUDENT RESEARCHER

☐ (+1) 951-907-2393 | ✉ mhasa006@ucr.edu | ⌂ mdahasan | in mdabidhasan

Research Interests

Computational biology, Machine learning and statistical model in bioinformatics, Cancer genomics, Epigenetics, Big data analysis and Deep learning in computational biology

Education

University of California Riverside *California, USA*

PH.D. CANDIDATE, COMPUTER SCIENCE AND ENGINEERING *Sept 2014 - Sept 2019 (Expected)*

- Thesis: Accurate Prediction Models for Biomolecular Data

Islamic University of Technology *Gazipur, Bangladesh*

M.S., COMPUTER SCIENCE AND ENGINEERING (GPA: 4.0/4.0) *Jan 2011 - Dec 2012*

- Thesis: Linear Regression Based Feature Selection for Microarray Data Classification

Islamic University of Technology *Gazipur, Bangladesh*

B.S., COMPUTER SCIENCE AND INFORMATION TECHNOLOGY (GPA: 3.98/4.0) *Jan 2007 - Nov 2010*

- First class first with honors (IUT Gold Medalist)

Publication

- **MA Hasan**, S Lonardi "A Deep Neural Network for Predicting Essential Genes in Microbes" [Submitted]
- **MA Hasan**, S Lonardi "mClass: Cancer Type Classification with Somatic Point Mutation Data", RECOMB International conference on Comparative Genomics, 131-145, 2018, Quebec, Canada.
- A. Polishko, **M. A. Hasan**, W. Pan, Evelien M. B. Karine L. R. Stefano L. "ThIEF: Finding Genome-wide Trajectories of Epigenetic Marks", The 17th Workshop of Algorithm in Bioinformatics (WABI), 2017, Boston MA.
- **Hasan M. A.**, Hasan M. K. and Mottalib M. A. "Linear Regression based Feature Selection for Microarray Data Classification", Int. J. Data Mining and Bioinformatics, 2015, Vol 11, Issue 2, pp. 167-179.
- K. M. Mutakabbir, S. S. Mahin, **M. A. Hasan**, "Mining Frequent Pattern Within a Genetic Sequence Using Unique Pattern Indexing and Mapping Techniques", IEEE International Conference on Informatics, Electronics & Vision (ICIEV), pp. 1–5, 2014, Dhaka, Bangladesh.
- **Hasan A.**, Adnan M. A., "High Dimensional Microarray Data Classification Using Correlation Based Feature Selection", International Conference on Biomedical Engineering (ICoBE), pp. 319 – 321, 2012, Penang, Malaysia.

Skills

Programming Language

- Skilled in Java, Python, C++
- Statistical tools: R, MATLAB
- Good understanding of HTML, PHP, CSS, SQL, Linux Shell
- Basic knowledge of PHP web framework, JavaScript, git

Bioinformatics Tools

- Experience with Samtools, Bowtie2, BWA, Genome Browser, BLAST, IGV
- Experience with large datasets, i.e. Microarray data, Sequencing data, PacBio data

Documentation Tools

- MS Office, LaTeX

Work Experience

Roche Sequencing Solutions *California, USA*

BIOINFORMATICS INTERN *Jun 2018 - Sept 2018*

- I was a summer intern at Roche working with the Bioinformatics team on variant calling. I have worked extensively on a tool written in Java and extended the functionality of the tool.
- I have also used Python scripts for data analysis during the internship.

University of California Riverside
California, USA

GRADUATE STUDENT RESEARCHER
Sept 2014 - Present

- Worked on ThIEF, A tool for finding genome wide trajectories of epigenetic marker
  - We addressed the problem of comparing multiple genome-wide maps representing nucleosome positions or specific histone marks. The tool generates an alignment of nucleosomes/histone marks across time points, allowing small movements and gaps in some of the maps. The output of ThIEF can also be used to produce supervised classifier that can accurately predict the position of stable and unstable nucleosomes from primary DNA sequence.
- mClass: Cancer type classification using somatic point mutation data
  - We used normalized mutual information based feature selection to classify multiple types of cancer using only somatic mutation data with logistic regression classifier.
- DeeplyEssential: Application of deep neural network in prediction of bacterial essential genes
  - A deep neural network architecture to predict gene essentiality in 30 bacterial species using exclusively sequence based features.
- HiC interaction prediction from Epigenetics data (Ongoing project)
  - A convolutional neural network based architecture to predict HiC interactions from epigenetic signals across multiple cell lines.

Islamic University of Technology
Gazipur, Bangladesh

LECTURER
Dec 2010 - Aug 2014

- Worked as a lecturer for multiple courses i.e. Bioinformatics, Computer Programming, Data Structure, System Analysis and Design, Information System, Statistical and Mathematical Analysis, Artificial Intelligence and Expert System
- Co-supervised undergraduate research groups

Software

DeeplyEssential
Coded in Python

- A feed forward neural network for predicting essential genes in 30 bacterial species.
Github Repository

mClass
Coded in Python

- Multiple cancer type classification using only Single Nucleotide Polymorphism (SNP) data
Github Repository

NUCLIUS
Coded in C++

- Supervised classification tool to predict stable/unstable nucleosomes position from primary DNA sequences
Github Repository

Cancer Classification from Microarray Data
Coded in Matlab

- Linear regression based feature selection for microarray data classification
Github Repository

References

Stefano Lonardi Ph.D.
University of California Riverside

PROFESSOR & ASSOCIATE CHAIR, COMPUTER SCIENCE AND ENGINEERING
Riverside, CA

- Tel: (951) 827 2203 • Email: stelo@cs.ucr.edu • LinkedIn Profile

Amrita Pati Ph.D.
Roche Sequencing Solutions

RESEARCH LEADER & BIOINFORMATICS GROUP LEADER
Pleasanton, CA

- Tel: (925) 353 7057 • Email: amrita.pati@roche.com • LinkedIn Profile

Hamid Mirebrahim Ph.D.
Roche Sequencing Solutions

PRINCIPAL SCIENTIST II
Pleasanton, CA

- Tel: (951) 880 8532 • Email: hamid.mirebrahim@roche.com • LinkedIn Profile