Harmeet Kaur

Franklin Park, NJ | 469-514-4840 | meet.academia@gmail.com | harmeet2504.github.io/
linkedin.com/in/harmeet-kaur-2504 | github.com/Harmeet2504 | tryanalyticsblog.com/timeseries-rnn/

SKILLS

Programming Languages:	Python (Pandas, Matplotlib, NumPy), SQL, HTML, Plotly
Methods:	Predictive modeling, Random Forest, Feature selection, Multi-layer perceptron (Deep learning), High Performance Computing (HPC), ETL, NLP
Tools:	Scikit-learn, Tensorflow, Keras, Jupyter notebook, Tableau, Git, Flask, Streamlit
Bioinformatics:	Molecular dynamics simulation, antibody sequence analysis, homology modeling, docking
Database:	PostGreSQL, mongoDB (basic), PySpark (basic), AWS (basic)

EXPERIENCE

Fellow, **Health Data Science,** Insight, Boston **May 2020 - Present**

- Built a web application for risk stratification of gestational diabetes in expectant mothers
- Formulated random forest algorithm to train models using electronic health records of 133 patients, implemented RFECV to identify fewer accessible features, tuned hyperparameters facilitating model optimization by 25%
- Identified high risk categories to introduce early lifestyle interventions, saves $5800 per pregnancy
- Deployed on Heroku: https://sweet-expectations.herokuapp.com/

Graduate Research Fellow, Regional Centre for Biotechnology, India **Feb 2012 - Jan 2019**

- Analyzed antibody-antigen (ab-ag) complexes to understand structural evolution in ab for different ag
- Collated X-ray crystallographic data of ab-ag complexes from PDB, conducted 3D modeling, sequence analysis, molecular dynamics simulation, implemented k-means clustering to analyze landscape of structural conformations
- Collaborated with a peer to study bioactive behaviour of MP-4 protein using docking and simulation studies
- Published 1 book chapter, 1 review article and 2 research articles in peer-reviewed journals, awarded best scientific poster presenter at the Program Advisory Committee meeting

Junior Research Fellow, CSIR-Central Drug Research Institute, India **Dec 2011 - Feb 2012**

- Standardized molecular biology techniques including plasmid extraction, gel electrophoresis, induction
- Facilitated lab establishment

Research Trainee, Bioinformatics Infrastructure Facility (BIF), Gauhati University **May 2010 - Oct 2010**

- Performed sequence analysis and phylogenetic studies of GCH1 gene product in some vertebrate species
- Prepared a project draft, developed pipeline and implemented the pipeline and presented to relevant stakeholder
- Communicated findings in technical report

PROJECTS

LubDub (https://lubdub-heartsense.herokuapp.com/) | Rutgers University, New Jersey **Feb 2020**

- Built a flask app from data deployed on AWS S3 PostgreSQL server to assess the risk of heart disease
- Analyzed big data using PySpark, feature selection, formulated multi-layer perceptron model
- Leveraged tableau for analytics and visualizations

Paradise (http://livebetter.herokuapp.com/) | Rutgers University, New Jersey **Dec 2019**

- Built an interactive dashboard to analyze 10 leading death causes in the United States from 2010-2017
- Led a team, extracted time-series data from CDC API, deployed to cloud mongoDB, created API routers
- Enables monitoring trends across the country, through JavaScript visualizations, facilitates decision making

EDUCATION

PhD (Bioinformatics) | UNESCO Regional Centre for Biotechnology **Jan 2019**