Harmeet Kaur

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

SKILLS

Programming Languages: Python(Pandas, Matplotlib, NumPy), SQL, HTML
Methods: Random Forest, Feature selection, Multi-layer perceptron, High Performance Computing (HPC), ETL, Sentiment Analysis (Lexicon-based NLP)
Tools: Scikit-learn, Tensorflow, Keras, Jupyter notebook, Tableau, Git, Flask, Streamlit
Bioinformatics: Molecular dynamics simulation, antibody sequence analysis, predictive structure modeling
Database: PostGreSQL, mongoDB (introductory)

EXPERIENCE

Health Data Fellow, 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**

- Analysed antibody-antigen complexes to understand structural evolution in antibodies for different antigens
- Collated and mined X-ray crystallographic data of antibody-antigen complexes from PDB, conducted 3D modeling and molecular dynamics simulation to generate a landscape of structural conformations, implemented k-means clustering in AMBER14 to identify dominant conformation involved in antigen interaction
- Collaborated with a peer to study bioactive behaviour of MP-4 protein using docking and simulation studies
- Mentored an Indian National Science Academy trainee for a 3-month project
- 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

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 to predict the potential risk of heart disease based on associated risk factors
- Analysed big data, feature selection, formulated support vector classifier and mulit-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 of four peers, extracted time-series data from CDC API, cleaned and loaded to cloud mongoDB Several API routers were created using flask for interactive visualizations
- Deployed online, enables monitoring trends across the country and decision making

EDUCATION

PhD (Bioinformatics) | UNESCO Regional Centre for Biotechnology **Jan 2019**
MSc (Biotechnology) | Gauhati University, India. **Jul 2007**