Curriculum Vitae - Shraddha Dumawat

+1-2164181762 shraddha.dumawat@gmail.com linkedin.com/in/shraddha-dumawat github.com/Shraddha-awat

Education

Case Western Reserve University **August 2024 – Current**
PhD in Biomedical and Health Informatics

Birla Institute of Technology and Science **September 2018 – December 2021**
Bachelor of Pharmacy *GPA : 8.12/10*

Experience

GSK (GlaxoSmithKline) **Feb 2023 – Present**
Digital Data Scientist II *Bangalore, India*

- Developed algorithms and engineered features for building of **Digital Biomarkers** for diagnosing Parkinson's' patients through a clinically sourced pilot study.
- Utilized the data from **passive monitoring such as sleep and active assessments such as voice and hand rotation** by intensive utilization of **digital signal processing** to derive essential features for baseline and longitudinal analysis of the condition.
- Conducted complex statistical analyses, covering **univariate, bivariate, and multivariable approaches**. Utilized various statistical tests, including but not limited to **p-values** and the **Mann-Whitney U test**, to quantify the strength of observed associations in potential digital biomarkers between Parkinson's patients and healthy controls.
- Deployed and crafted classification algorithms such as **Self Organized Maps, Auto-encoders, CNN** to curate and quality control the real-world data, resulting in a **60 percent increase** in the importance of accurate features and **reducing the turnaround time by 1800 times**.
- Presented in various internal workshops and competitions to gauge understanding of application Machine learning in real world data.

Elucidata **Jan 2022 – Feb 2023**
Data Analyst *Bangalore , India*

- Executed **data analysis** and **data audits** related to bioinformatics using **python** to connect with 25+ potential prospects from major and new drug discovery companies.
- Worked on multiple pipelines of different clinical, pre-clinical and gene expression data types including but not limited to **RNA sequencing, flow cytometry, single cell sequencing and drug response** using **python and R** to demonstrate **downstream analysis and data preprocessing** such as PCA, t-SNE etc for prospect's ML pipelines and marketing collaterals.
- Contributed to the **model building** for identifying and curating healthcare data from public sources using various ML models.
- Explored **Auto-ML** integration on the platform for proof of concept of **clean data** leading to better outcomes. Curated and crafted a running algorithm to predict lung cancer stages through RNA-sequencing data from the Auto-ML algorithm, achieving 93 percent AUC.

Kyoto University **April 2021 – Jan 2022**
Undergraduate Research Intern *Remote*

- Worked under **Dr Ganesh Namasivayam** on Leber Hereditary optic nephropathy to model drug candidates. Explored various techniques for wet lab analysis of real-time patient data collaborating with an eye hospital in Bangalore.
- Explored the potential of targeting **G-quadruplex** structures in mitochondria for therapeutic interventions.
- Used **Python** for Data Mining to build a database on host-microbiome interaction. Used **SQL snd MS Excel** to curate 90000+ rows of data and for further analysis.

Epigenetic Research Lab, BITS Pilani **Jan 2020 – Dec 2021**
Dissertation and Project Research Scholar *Hyderabad, India*

- Worked under **Dr. Balaram Ghosh** , Evaluated **Pharmacophore and QSAR** of histone deacetylase inhibitors through utilising Schrodinger.
- Constructed and refined databases and compared ML models for the analysis of the activity of histone deacetylase inhibitor activity for a **regression-based model** for prediction of activity (IC50 value) against various HDAC isoforms.
- Worked on Fragment based structural analysis of HDAC3 inhibitor, leading to a publication.

Publications

Banerjee S, **Dumawat S**, Jha T, Lanka G, Adhikari N, Ghosh B. Fragment-based structural exploration and chemico-biological interaction study of HDAC3 inhibitors through non-linear pattern recognition, chemical space, and binding mode of interaction analysis. J Biomol Struct Dyn. 2023 Aug 23:1-23. doi: 10.1080/07391102.2023.2248509. Epub ahead of print. PMID: 37608752.

Projects

Fragment based structural analysis of HDAC3 inhibitor | *Python,QSAR,Machine Learning* **Aug - Dec 2022**

* Implemented a **fragment-based non-linear machine learning (ML)** method along with chemical space exploration followed by a structure-based binding mode of interaction analysis on some HDAC3 inhibitors to obtain the key structural features modulating HDAC3 inhibition.

Molecular Modelling and docking for LHON | *Cheminformatics, Python* **Aug - Dec 2022**

* Modeled towards the treatment for a rare genetic disorder, Leber Hereditary Optic Neuropathy (LHON) by designing molecules through **Molecular Modelling** by conjugation of desired DNA intercalators to cPIP and Docking studies on a mutated set of DNAs.

Modelling of therapeutic candidates for Anti-SARS CoV 2 | *QSAR, Python, Drug design* **Jan-May 2021**

* Comprehensive Literature review on various drug receptors of SARS COV-2 and existing drugs for the same.
* Analyzing the receptors, proteins and ligands in order to look into the possible therapeutic agents against the virus utilizing the softwares schrodinger, Argus Labs and YASARA.
* Drug Designing by targeting Proteins that are reported for SARS type of virus using drug repurposing strategy.

Drug Bioactivity Prediction | *Machine Learning, Python, Data Processing* **Aug-Dec 2022**

* **Installed the Decision Tree Regression** Machine Learning model for the **prediction of bio-activity of drugs** in terms of IC50 value against HDAC1. Operated CHEMBL for data sourcing and Remodeled the data using RDKIT

Data analysis of pregnant women during COVID-19 pandemic | *Advanced Excel, Python* **May - July 2020**

* Studied the effects on Mental Health by review of the existing the literature on mental health and neurological effects on mothers and pregnant women were performed by survying 100 volunteers and analysing the data and providing solutions for betterment of their mental health. The data was mined as processed from **Lenest Hospitals, Mumbai** under the supervision of Dr. Mukesh Gupta (OB-GYN).

Technical Skills

Languages: Python, R, C, SQL

Technology: Digital Signal Processing , Deep learning, Machine Learning, Molecular Docking, QSAR, Advanced Excel, MATLAB , AWS, Schrodinger

Select Libraries : TensorFlow, Pandas, OpenCV, Scikit-learn , Librosa, SciPy, NumPy

Relevant Coursework: Pharmacology, Medicinal Chemistry, C, Linear Algebra, Genomics

Bioinformatics Data sets : RNAseq, Single cell RNA-seq, flow cytometry, ImmPort, Drug Response

Wet-lab Skills: Western Blotting, Tissue and Cell culture, PCR, RNA-extraction, Multi-step molecule synthesis

Interests: : Swimming, Poetry, Oration, Creative Writing, Video Editing

Soft Skills: Event Management, Public Speaking, Leadership, Time Management

Leadership / Extracurricular

Student Welfare Division **Spring 2020 – Fall 2022**
Coordinator *BITS Pilani*

* Coordinator of the Student Welfare Division, leading a team of 8 students to monitor and take care of the finances, hygiene and health of 6000 + students.
* Employed programming languages such as Python and SQL to manage databases of university students, create platform for mess registration and bio metric entry at campus gates.

National Service Scheme **August 2018 - December 2019**
Member *BITS Pilani*

* Coordinated donation drives for school books, clothing, and stationery in underprivileged schools.
* Provided tutoring sessions to enhance computer skills, covering painting software as well as Microsoft Word and Excel for students unfamiliar with these applications.

Hindi Tarang (Club) **August 2019 - August 2020**
Treasurer *BITS Pilani*

* Handled Finances of the club and organised competitions, trips and workshops for the cultivation of Hindi Language.

Achievements

- Awarded the prestigious **Transformational Medicines and Vaccines award** at GSK, awarded to top employees.
- Winner of Digispark 2023 at GSK for automated quality control of real world voice data using **ML for Digital Biomarkers of Parkinsons' disease.**
- Amongst the 0.5 percent of students to complete graduation one semester early.

Certifications and Workshops

Linear Algebra for Machine Learning and Data Science : Leveraged online learning to bridge gaps and build stronger foundation of Linear Algebra in Machine learning.

Statistics : Revisited statistical concepts, including probability distributions and regression analysis, for their practical application in real-world clinical data.

Computer-Aided Drug Design/Discovery and Development-011: Mastered tools such as Schrodinger for Molecular Docking studies essential in drug development.

Professional Data Science: Upskilled in the concepts of Data Science and Machine learning along with AI modules. (Organisation : 360DigiTMG, India)

Medical Innovation and Design Thinking: Attended 2 day long conference and workshop on Medical Innovation in a collaborative effort between BITS Pilani and Cardiff University.