BHANU SHARMA

+1 617-560-0716 | sharma.bhan@northeastern.edu | Linkedin | Boston, MA

EDUCATION

MS in Bioinformatics (GPA: 4.0) — Sept 2024 - Expected
Northeastern University — Boston, MA, USA
Relevant coursework: Python, Statistics, Algorithms, Computational Biology

Bachelor of Technology in ECE (GPA: 7.7/10) — Aug 2016 - May 2020
Punjab Engineering College (PEC) — Chandigarh, India
Relevant coursework: DBMS, Data Structures Algorithms, Operating Systems

TECHNICAL SKILLS

Programming Languages: Python, R, SQL, UNIX/Shell, Scala, Java, SAS, HTML, CSS, Node.js
Libraries: Numpy, Pyspark, Pandas, Matplotlib, Seaborn, ggplot2, scikit-learn, Pytorch, Tensorflow
Tools: Power BI, Tableau, Azure, Databricks, Snowflake, Apache Spark, Docker, Singularity, AWS, MS Excel, Github
Skills: Data Wrangling, Data Visualisation, Data modelling, Business Intelligence, Big Data, Big Query, ML, ETL pipelines, CI/CD, Agile, Cloud Technologies

EXPERIENCE

UnitedHealth Group (Optum), Gurgaon, India: Data Engineer 2 — Sept 2023 - Aug 2024
- Designed numerous interactive dashboards using Power BI and MySQL delivering essential business insights and KPI results for 50+ clients spread across 4 departments
- Developed a comprehensive notification & alerting system in Python that enabled users to send and schedule SQL result sets as emails via Outlook, saving an estimated 6 hours weekly for end-users

UnitedHealth Group (Optum), Gurgaon, India: Data Engineer 1 — Aug 2020 - Aug 2023
- Migrated legacy ETL pipeline to Spark/Scala and Snowflake framework on Microsoft Azure, enabling a transition to a cost-efficient cloud-based platform and reducing database costs by up to 90%
- Redesigned on-prem ETL architecture from SAS to Spark/Scala, PySpark, and MongoDb(NoSQL) architecture hosted on Azure Databricks and Azure Data Factory allowing clients to adopt an efficient workflow handling an estimated 3TB+ of daily data
- Engineered a novel Data Quality checking system using Great Expectations library in Python with an accuracy of 99%

Ituple Technologies, Gurgaon, India: Software Engineering Intern — Jan 2019 - May 2019
- Implemented Identity Access and Management (IAM) System with CRUD operations using Amazon Web Services(AWS) and Node.js to streamline secure access for 100+ users

PROJECTS

Cervical Cancer Risk Prediction using Machine Learning, Northeastern University, Boston, MA - Github
- Built, trained and evaluated decision trees based XG-Boost classifier ML model using Python's Scikit-learn library to predict whether a person has a risk of having cervical cancer with an accuracy of 97%

DNA Methylation Analysis Pipeline, Northeastern University, Boston, MA - Github
- Reproduced an end-to-end DNA methylation analysis pipeline using Singularity/Docker from the following research paper: https://doi.org/10.1186/s12859-022-04925-2 hosted on HPC allowing users to generate DNA Methylation Profiles for any species and sample for advanced genomic analysis

Eye Controlled Wheelchair, Punjab Engineering College, Chandigarh, India
- Created a prototype wheelchair controlled by eye movements for paraplegic patients using Python and C source code employed by Arduino UNO microcontroller

LEADERSHIP

Unitedhealth Group (Optum), Gurgaon, India: Data Engineer 1 — May 2021 - May 2022
- Led a team of 5 to develop ETL pipeline, resulting in recognition in the Optum Newsletter for improving data processing efficiency and bringing down database cost