PULKIT ANUPAM SRIVASTAVA

"I want to positively mutate our Lifestyles"

Pulkit Anupam | Pulkit Anupam | PulkiD

pulkit.srivastava@qriousatom.com

+91-962XXXXXXX, 862XXXXXXX

Gorakhpur, Uttar Pradesh, India

👉 Freelancer | Bioinformatician cum Data Scientist | Developing 'T' shaped skills | Interdisciplinary Research

TECHNICAL STRENGTH & RESEARCH INTEREST

Exploratory Data Analysis	●●●●●	Network Analysis	●●●●
Deep Learning	●●●	Convolutional Neural Network	●●
Panomics Data Analysis	●●●●	Natural Language Processing	●●●
Database (MongoDB, ES)	●●●	Programming (Python, R)	●●●●●

WORK EXPERIENCE

Freelancer | Oct, 2021 - Present

Objective: Predict phage insertion sites in bacteria
Contribution: Developed an end-to-end novel approach to identify integrase proteins & phage insertion sites in bacterial genome
Techstack: Python|BLAST|Web scraping through BeautifulSoup|Pandas

Objective: Knowledgebase: Bioethanol
Contribution: Developing one-stop solution for Bioethanol based studies & prediction of novel molecular signatures involved in Bioethanol production
Techstack: Python|NLP|ElasticSearch|NGS Analysis|MongoDB|Web scraping|Bio-BERT|Network Analysis

Objective: Investment Survey Analysis
Contribution: Analysed risk aversion of participants with respect to psychological factors
Techstack: Python|Pandas|Seaborn|Scipy

Consultant - Data Science | Aug, 2021 - Oct, 2021 | ViralOps, India

Objective: Personalised F&B recommender system
Contribution: Developed a system to recommend combination of F&B products based on customer reviews
Techstack: Python|MySQL|MongoDB|Scikit-learn|Seaborn|SVM|Random Forest|Market-Basket Analysis

Associate Scientific Manager - Data Science | Jan, 2021 - Aug, 2021 | Innoplexus Consulting Services, India

Objective: Biomarker Continuous Analytics Dashboard Development
Contribution: Led & expanded our dashboard to run on-fly analysis to predict potential biomarkers through user provided experimental data
Techstack: Python|ElasticSearch|NGS Analysis|MongoDB|Flask|Backend API Development|Omics Data Analysis

Objective: Network based molecular signature identification
Contribution: Developed a method to identify molecular signatures on top of biological networks and experimental data
Techstack: Python|NGS Analysis|Transcriptomics|Network Analysis|Graph Neural Network|Neo4j|MongoDB

Senior Research Analyst - Data Science | Apr, 2020 - Dec, 2020 | Innoplexus Consulting Services, India

Objective: Biomarker Continuous Analytics Dashboard Development
Contribution: Led and developed analytics dashboard that served as one stop solution for present clinical Biomarkers and prediction of novel potential Biomarkers
Techstack: Python|ElasticSearch|NGS Analysis|MongoDB|Flask|Backend API Development|Omics Data Analysis|Scikit-learn|Random Forest

Objective: Collation of insights from multi-layer data
Contribution: Developed a method to associate experimental insights and reported results in research papers to identify and prioritize molecular signatures for a given disease or treatment therapy
Techstack: Python|Pandas|Scikit-learn|Spacy|NLP|Ontology|Random Forest|TF-IDF|ElasticSearch

Knowledge Associate - Data Science | Aug, 2019 - Mar, 2020 | Innoplexus Consulting Services, India

Objective: Automate analysis of experimental data gathered from NCBI GEO and EBI ArrayExpress
Contribution: Developed end-to-end pipeline to analyze RNA-seq and Microarray based Transcriptomics data
Techstack: Python|Pandas|Seaborn|Scipy|Numpy|Biopython

Objective: Client Delivery Projects
Contribution: To provide assistance in analysis of experimental data gathered as a part of client delivery projects
Techstack: Python|Exploratory Data Analysis|Seaborn|Matplotlib|Pygal|Scipy

Continued.

Research Articles

Srivastava P. A., Hegg E. L., Fox B. G., Yennamalli R. M. PreDSLpmo: A neural network-based prediction tool for functional annotation of lytic polysaccharide monooxygenases. Journal of Biotechnology - 2020.

Chauhan D., **Srivastava P. A.**, Ritzl B., Yennamalli R. M., Cava F., Priyadarshini R. Amino acid Induced Pleomorphism in *Deinococcus indicus* Strain DR1. Frontiers in Microbiology - 2019.

Chauhan D., **Srivastava P. A.**, Agnihotri V., Yennamalli R. M., Priyadarshini R. Structure and Function Prediction of Arsenate Reductase from *Deinococcus indicus* DR1. Journal of Molecular Modeling - 2019.

Bansal A., **Srivastava P. A**., and Singh T. R. An integrative approach to develop computational pipeline for drug-target interaction network analysis. Scientific Reports - 2018.

Chauhan D., **Srivastava P. A.**, Yennamalli R. M., Priyadarshini R. Draft Genome Sequence of *Deinococcus indicus* DR1, a Novel Strain Isolated from a Freshwater Wetland. Genome Announcements - 2017.

Book Chapters

Srivastava PA and Yennamalli RM. OMIC Technologies in Bioethanol Production: An Indian Context. OMICS-based Approaches for Plant Biotechnology - 2019.

Bansal A. and **Srivastava P. A**. Transcriptomics to Metabolomics: A Network Perspective for Big Data. Applying Big Data Analytics in Bioinformatics and Medicine - 2018.

Review Articles

Srivastava PA, Kalra S, Yennamalli RM. Structural Bioinformatics and Big Data Analytics: A mini-review. International Journal of Computational Biology - 2017.

Poster Presentation | Conferences

Poster titled "Evaluation of gene co-expression in metabolic pathways, and its application in prediction of biological process/pathway of unannotated genes" at **International Conference on Bioinformatics (InCoB) 2018**, New Delhi.

Poster titled "Systems Biology approach to identify genes involved in ligno-cellulosic deconstruction" at **ISCB's Intelligent Systems for Molecular Biology (ISMB) 2018**, Chicago, USA. (Presented by co-author)

Education

Bachelor's of Technology | Jul 2015 - May, 2019 | Jaypee University of Information Technology, India
Area of study: Bioinformatics
CGPA: 9.1

Internships

Intern - Data Science | May, 2019 - Jul, 2019 | Innoplexus Consulting Services, India

Objective: Establish Proof-of-Concepts for publicly available datasets
Contribution: Developed an automated pipeline to extract expression samples from GEO and annotate its meta information using named entity recognition.
Techstack: Python|Spacy|Ontology|MongoDB|ElasticSearch|Web scraping through BeautifulSoup|Pandas|Seaborn

Summer Intern | May, 2018 - Jul, 2018| International Crops Research Institute for the Semi-Arid Tropics, India

Objective: Development of Know Your Gene Dashboard
Contribution: Developed an automated pipeline to functionally annotate genes with unknown function through co-expression with other molecular signatures deciphered from experimental studies.
Techstack: Python|Pandas|Seaborn|Scipy|Numpy|Gene Set Enrichment Analysis|Exploratory Data Analysis

Learnings

Domain knowledge, though can be learned, plays a crucial role in deriving insights from the data.

In-depth research and planning is good, but "Bias for execution" is the game changer.

Data collection and it's preprocessing is the most important step in analysis.

Getting the first version deployed is more important than making it perfect on staging.

"Contemplating, questioning, associating, networking & experimenting" skill-set is more impactful than any technical skill.