biotecnika
Your Bio Resource

DATA SCIENCE FOR BIOLOGISTS

Hands-On Internship

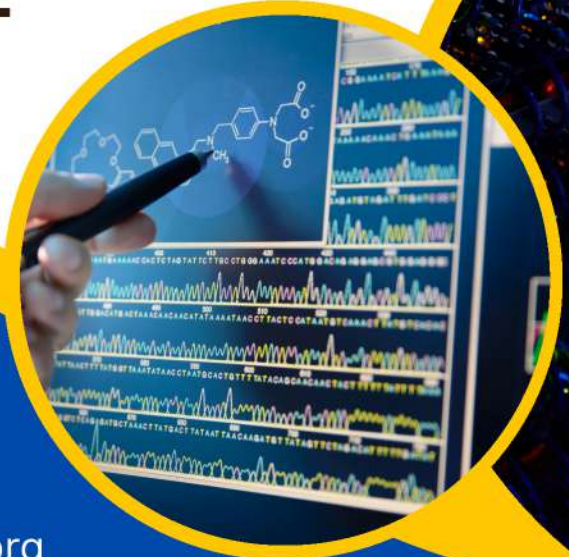
PROJECT | PAPER PUBLICATION | WORK EXPERIENCE

Extensive 120 Days Hands-on Training With 3,6 & 12 Months Project Work

Exclusively Crafted For B.Sc., M.Sc., B.Tech M.Tech. & PhD Biotech, Life Science, Chemistry + B.Pharm, M.Pharm Students

GET WORK EXPERIENCE

PAPER PUBLICATION ASSISTANCE



stores.biotecnika.org
Mail: info@biotecnika.org

biotecnika
Your Bio Resource

Embark on a transformative journey with the Data Science Hands-On Internship for Biologists: From Basics to Advanced with Projects & Publication! Tailored for biologists, this program offers expert mentorship, hands-on training, and practical experience with essential data science tools. From foundational skills to real-world projects and publications, you'll gain everything you need to thrive in a data-driven biology career. Don't miss this chance to innovate, discover, and grow—your future in data science starts here!

Extensive 120 Days Hands-on Training With 3,6 & 12 Months Project Work

WITH ONLINE + OFFLINE PROJECT

Session No.	Unit No. and Topic
Unit 1: Introduction to Biological Data Analysis	
Day 1	Getting Started with Python and R Python vs. R for Biological Data AnalysisInstalling Python, R, and Required Libraries
Day 2	Data Visualization Introduction to Data VisualizationCreating Basic Plots with Matplotlib (Python) and ggplot2 (R)
Day 3	Data Import and Manipulation Working with Biological Data FormatsBasic Data Manipulation with Python and R
Day 4	Exploratory Data Analysis (EDA) Understanding Your Biological DataEDA Techniques in Python and R
Day 5	Introduction to Bio Python What is Bio Python?Basic Bio Python Functions for Sequence Analysis
Unit 2: Molecular Data Analysis	
Day 6	Sequence Feature Analysis : Identifying and Annotating Sequence Features Sequence Motif Search with Biopython
Day 7	Protein Structure Analysis : Introduction to Protein Structure AnalysisUsing Biopython for Protein Structure Data
Day 8	Phylogenetic Tree Analysis: Phylogenetic tree construction, visualization and interpretation.

Day 9	Sequence Alignment Introduction to Sequence AlignmentPairwise Sequence Alignment with Bio Python
Day 10	Multiple Sequence Alignment : Multiple Sequence Alignment with Biopython Sequence Alignment Tools and Techniques

Unit 3: Clinical Data Analysis

Day 11	Clinical Data Preprocessing Cleaning and Organizing Clinical Data Handling Missing Data
Day 12	Survival Analysis Introduction to Survival Analysis Kaplan-Meier Estimator and Cox Proportional Hazards Model in R
Day 13	File Parsing and Data Retrieval Reading and Writing FASTA Files Parsing GenBank Files
Day 14	CADD data analysis using Python and BioPython Working with molecular structures and visualization for drug designing Analyzing drug bioactivity data to screen potential drug candidate
Day 15	Project and Presentation Applying Biological Data Analysis in a Research Project

Day 16 **Open Discussion and Project Discussion for coding**

Unit 4- Introduction to Biostatistics

Day 17	Overview of Biostatistics: Definition, importance in public health and medicine, roles in research.Basic Concepts: Variables (categorical, continuous), data types, populations vs. samples.Descriptive vs. Inferential St
Day 18	Data Types and Measures of Central Tendency Data Types Review: Nominal, ordinal, interval, and ratio data. Central Tendency: Mean, median, mode—when to use each. Variation and Dispersion: Range, variance, standard deviation, interquartile range.

Day 19	Hypothesis Testing Basics Null and Alternative Hypothesis: Definitions and examples. Type I and Type II Errors: Explanation of errors in hypothesis testing. Significance Levels (α): Understanding p-values.
Day 20	Confidence Intervals Understanding Confidence Intervals: Meaning, interpretation, and significance. Confidence Interval for Mean: How to compute and interpret. Relationship to Hypothesis Testing: Using CIs to assess statistical significance.

Bioinformatics

UNIT-5: Core Bioinformatics & Sequence Analysis Foundations

Day 21	Introduction to Bioinformatics Resources Sub-topics- What is bioinformatics? Major branches and applications II Navigating NCBI portal – PubMed, Nucleotide, Protein databases
Day 22	Sequence Databases Deep Dive Sub-topics- UniProtKB: understanding protein annotations (Swiss-Prot vs. TrEMBL) II Sequence file formats – FASTA, GenBank, EMBL formats
Day 23	Pairwise Alignment Concepts, Sub-topics- Scoring matrices (BLOSUM, PAM) and gap penalties II BLAST algorithms – understanding E-values and bit scores
Day 24	Multiple Sequence Alignment, Sub-topics- MSA algorithms overview (Clustal, MUSCLE, MAFFT) II Alignment visualization and manual editing

UNIT-6: Structural Bioinformatics & Protein Analysis

Day 25	3D Structure Databases, Sub-topics- Protein Data Bank (PDB) – structure determination methods II Analyzing PDB file headers and understanding experimental data
Day 26	Specialized Molecular Databases, Sub-topics- Domain databases: Pfam, SMART, InterPro II Structure classification databases – CATH, SCOP2

Day 27	Visualization Fundamentals, Sub-topics- Molecular visualization principles: ribbons, surfaces, cartoons II PyMOL basics – loading structures, coloring schemes, saving images
Day 28	Structure-Function Relationships, Sub-topics- Active sites, binding pockets, and functional motifs II Ligand-protein interactions in PDB structures
Day 29	Chemical Databases, Sub-Topics- Exploration of chemical entities and extracting usefull information out of it
Day 30	Protein Structure Modelling, Sub-Topics- Types of protein modelling: Homology, Ab-Initio and threading II Mutant protein modelling

UNIT-7: Evolution, Function & Interaction Biology

Day 31	Phylogenetics Fundamentals, Sub-topics- Tree terminology: nodes, branches, clades, outgroups II Distance matrix methods (Neighbor-Joining)
Day 32	Protein Family Classification, Sub-topics- Hierarchical classification systems (CATH, SCOP) II Superfamily and family relationships
Day 33	Functional Annotation, Sub-topics- Gene Ontology (GO): structure and relationships II GO term annotation and propagation
Day 34	Protein-Protein Interactions, Sub-topics- Experimental methods for detecting interactions II Interaction databases – STRING, BioGRID, IntAct

Unit-8- Genomics, Sequencing & Systems-Level Biology

Day 35	Sequencing advancements and workflow overview, Sub-Topics- Sanger sequencing, NGS, 3rd gen sequencing techniques II Advantages and disadvantages of each technique II Basic NGS data analysis workflow
Day 36	Fundamentals of Drug Discovery, Sub-topic- Pre-clinical and Clinical Trials for drug discoveryII Genomics segregation: Structural, functional and comparative

Day 37	Metabolic Pathways, Sub-topics- Pathway databases: KEGG, Reactome, MetaCyc II Navigating pathway maps and understanding reactions
Day 38	Fundamentals of Genomics Sub-Topics- Human Genome project II Bacterial Genomics II Epigenetics

Unit-9: Translational Bioinformatics & Disease Applications

Day 39	Druglikeness and ADMET Analysis, Sub-Topics- Lipinski's, Pfizer, Ghose, Veber rule etc. II Pharmacokinetics and Pharmacodynamics concept- ADME/T
Day 40	Molecular Docking, Sub-Topics- Fundamental concepts of docking and energy function II Types of Docking II Site-specific and blind docking
Day 41	Molecular Dynamics Simulations, Sub-Topics- Fundamental concepts of MDS II Difference between MD and MDS II Defining the steps and stages II Command or RMSD/RMSF/Rg etc. analysis explanation
Day 42	Functional Enrichment Analysis, Sub-topics- Over-representation analysis concepts II Enrichment tools – DAVID, g:Profiler, Enrichr
Day 43	Variant Databases & Analysis, Sub-topics- Genetic variation resources: dbSNP, ClinVar, gnomAD II Variant interpretation and pathogenicity prediction
Day 44	Disease Bioinformatics, Sub-topics- Disease databases: OMIM, DisGeNET, GWAS Catalog II Gene-disease associations and network medicine
Day 45	Integrative Data Portals, Sub-topics- Multi-omics data integration platforms II cBioPortal for cancer genomics
Day 46	Human Microbiome databases, Sub-Topics- Microbiome overview, Good bacteria concept, gut-brain axis *Tools: MicrobiomeDb, IndMDB etc, Hands-on: Overview of the tools

Unit-10: AI, CADD & Future Trends

Day 47	AI in biology overview Sub-Topics- Fundamental concepts of AI, ML, DL, NN, NLP, LLM, GenAI etc II Application of AI in biological data analysis overview only
Day 48	Computer Aided Drug Designing (CADD)- Part I Sub-Topics- SBDD with standard workflow explanation
Day 49	Computer Aided Drug Designing (CADD)- Part II Sub-Topics- LBDD with QSAR modelling explanation
Day 50	Emerging Resources & Future Trends Sub-topics- Single-cell databases, spatial transcriptomics resources II Metabolomics and proteomics databases II Future prospects of Bioinformatics

Day 51 Open Discussion & Project topic discussion for Bioinformatics

Unit 11- Scientific Writing, Publication & Research Fundingatics

Day 52	Publishing Research – From Manuscript to Acceptance : Types of research publications, Journal categories & indexing, Publisher overview (Elsevier, Springer, Wiley, MDPI, Bentham)Journal selection tools, Manuscript submission workflow, Peer review, revision & resubmission, Editorial Manager demo
Day 53	AI-Assisted Research Writing & Strategy : AI/ML basics for researchers, ChatGPT for literature review, Manuscript structuring & language editing, AI for cover letters & reviewer responses, Ethical use of AI in research
Day 54	Proposal Writing, Grants & Funding Ecosystem : Proposal structure & key components, Generating preliminary data, Funding agencies & collaborators, Grant submission portals, Common proposal mistakes

Unit 12- Biostatistics Part 2

Day 55	Sampling Methods Sampling Techniques: Simple random sampling, stratified, cluster, and systematic sampling. Sampling Bias and Its Impact: Discuss selection bias, non-response bias. Sample Size Determination: Power analysis and sample size calculation.
Day 56	Study Design: Observational Studies, Types of Observational Studies: Cross-sectional, case-control, cohort studies, Biases in Observational Studies: Confounding, selection bias, Advantages and Limitations: Pros and cons of each study design.
Day 57	Study Design: Experimental Studies Randomized Controlled Trials (RCTs): Key features and methodology. Blinding and Randomization: Importance in reducing bias. Ethical Considerations: Ethics in clinical trials and human subject research.

UNIT-13: Introduction: Foundation Concepts of Coding & AIML

Day 58	Overview of Artificial Intelligence (AI) and Machine Learning (ML) Introduction to Biology & Bioinformatics applications of AI/ML
Day 59	Introduction to Integrated development environment Software for R and python
Day 60	Introduction to python, Installation, writing first code in python
Day 61	Introduction to R, Installation, writing first code in R
Day 62	Data representation and visualization in python and R
Day 63	Importance and significance of AI/ML in biological research. Types of ML algorithms and their relevance in Biology & Bioinformatics

Unit 14: Machine Learning Algorithms

Day 64	Overview of supervised learning algorithms Applications of supervised learning in Biology & Bioinformatics (e.g., sequence classification, protein structure prediction)
Day 65	Classification and regression algorithms in details
Day 66	Dimensionality reduction techniques (e.g., Principal Component Analysis) and their applications.
Day 67	Unsupervised learning applications in Biology & Bioinformatics
Day 68	Feature Engineering for Biological Data Encoding categorical biological features (e.g., DNA bases, amino acids), One-hot encoding, embedding, k-mer representation, why feature engineering is critical for biological predictions.

Unit 15: Deep Learning Algorithms

Day 69	Introduction to Deep Learning
Day 70	Introduction to Neural Networks
Day 71	Artificial Neural Network for biological data analysis
Day 72	Convolutional Neural Networks (CNNs) for biological analysis
Day 73	Recurrent Neural Networks (RNNs) for sequence data analysis

Unit 16: Data Preprocessing in handling large Data Set

Day 74	Data cleaning techniques for biological data Handling missing data in bioinformatics datasets
Day 75	Data normalization transformation methods Feature selection dimensionality reduction techniques in bioinformatics
Day 76	Evaluation Metrics for AI/ML Models in Biology : Accuracy, Precision, Recall, F1-score, ROC–AUC, Why certain metrics matter more in bioinformatics (e.g., sensitivity in disease classification), Cross-validation and overfitting in biological datasets

Unit 17: Applications of AI ML in Life Science

Day 77	Applications of AI/ML in Genomics
Day 78	Applications of AI/ML in Proteomics. Protein structure prediction techniques: AI-based modeling, AlphaFold etc
Day 79	AI/ML tools for CADD and its applications
Day 80	Application of AI/ML in Drug Discovery
Day 81	Applications of AI/ML in Multi-omics
Day 82	Practical Sessions using GitHub repository Hands-on exercises using relevant tools and datasets

Day 83 Open discussion & Project Topic Discussion for AIML

Unit 18 - Descriptive statistics

Day 84	Comparison of Means : Independent t-test: Assumptions, application, and interpretation.Paired t-test: When to use it, examples, and calculations.One-Way ANOVA: Introduction to analysis of variance.
Day 85	Non-Parametric Tests : When to Use Non-Parametric Tests: For non-normal data or ordinal variables.Mann-Whitney U Test, Wilcoxon Signed-Rank Test, Kruskal-Wallis Test.Chi-Square Test: For categorical data.
Day 86	Regression Analysis : Linear Regression: Simple linear regression, model interpretation, assumptions.Multiple Regression: Multiple predictors, multicollinearity, model selection.Model Diagnostics: Checking residuals, goodness-of-fit.
Day 87	Correlation and Association : Pearson Correlation: Measures of linear correlation.Spearman's Rank Correlation: Non-parametric measure of association.Chi-Square for Association: In categorical data analysis.

NGS and Multi-Omics Data Analysis

Unit 19 : Introduction & Fundamental Concepts

Day 88	Orientation and Introduction to NGS: Principle & Importance, Historical context and evolution from Sanger sequencingComparison with traditional sequencing methods
Day 89	Overview of the NGS workflow: library preparation, sequencing, and data analysis. Detailed explanation of library preparation steps (DNA fragmentation, adapter ligation, amplification), Quality control measures at each step of the workflow
Day 90	Introduction to major NGS platforms (e.g., Illumina, Ion Torrent, Pacific Biosciences)Differences in sequencing chemistry, read lengths (Differences between short-read vs. long-read sequencing), throughput, and applicationsConsiderations for platform selection based on research needs
Day 91	Library preparation techniques for different types of samples (DNA, RNA, ChIP, etc.)Protocols for preparing libraries for whole-genome sequencing, RNA-seq, ChIP-seq, etc.Common challenges and troubleshooting tips
Day 92	Manipulation of Linux & Python/R in NGS data analysis Basic Linux commands line: File and folder navigation (cd, mkdir, rm, ls, mv); Basic file handling (cat, more, wc); File manipulation (cut, awk, sed, grep, sort). Python/R basics (variables, loops, installing libraries like Bioconductor).

Unit 20 : Preprocessing, Alignment, Data Formats, & Variant Calling

Day 93	User Interface : Location, Directory, path in computational terms; Basic configuration check of Computers; Switching between different UIs
Day 94	Insight into navigating and utilizing the SRA database for NGS data
Day 95	Adapter contamination, trimming algorithms (e.g., Trimmomatic, Cutadapt). Practical: Adapter contamination, trimming algorithms (e.g., Trimmomatic, Cutadapt).
Day 96	Burrows-Wheeler Transform (BWA), alignment algorithms (BWA-MEM, Bowtie2) **Practical:** Align reads to a reference genome (e.g., hg19, hg38) using BWA

Day 97	SAM/BAM format, flags, and sorting/indexing **Practical:** Use SAMtools to sort, index, and filter BAM files
Day 98	SNVs vs. indels, germline vs. somatic variants, VCF format, annotations (e.g., QUAL, DP), filtering variants
Day 99	**Practical:** Filter VCF files with bcftools and annotate with VEP(variant effect predictor), Varsome

Unit 21: Gene Expression Analysis

Day 100	Transcript quantification, differential expression analysis
Day 101	**Practical:** Demo on Data preprocessing
Day 102	**Practical:** Align RNA-seq reads with STAR aligner
Day 103	Gene-count matrices, normalization (TPM, RPKM)
Day 104	**Practical:** Generate a count matrix using featureCounts
Day 105	Differential Expression : DESeq2/edgeR workflow, fold change, p-value adjustment
Day 106	**Practical:** Run DESeq2 on a count matrix to identify DEGs
Day 107	**Practical:** Demo on DEG Visulaization and Reporting

Unit 22 : Applications of NGS	
Day 108	Cancer Genomics, Inherited Genetic Disorders , Microbiology and infectious diseasesRole of NGS in cancer and biomedical researchMicrobial genomics - Pathogen Identification & Antimicrobial Resistance
Day 109	NGS for Target Identification, Validation & Clinical TrialsExamples of some computational algorithms, Future Directions of NGS in Drug Discovery & Challenges, PharmacogenomicsOverview of NGS applications in clinical settings (diagnostics, personalized medicine)
Day 110	**Practical:** Design sgRNAs using online tools (e.g., CHOPCHOP, CRISPOR); visualize off-targets.
Day 111	Genomic Breeding and Crop Improvement, Disease Resistance and Pest Management, Soil Microbiome and Agricultural Sustainability
Unit 23: Cloud based NGS analysis	
Day 112	Introduction to Galaxy platform and cloud-based NGS analysis; **Practical:** Live RNA seq walkthrough
Day 113	Basics of single-cell RNA sequencing; Introduction to Python packages: scanpy, anndata, seaborn;Data preprocessing: filtering, normalization, and scaling;Dimensionality reduction (PCA, UMAP) **Practical:** Demo on running a scRNA-seq workflow on Google Colab using a sample dataset
Day 114	Overview of tools: QIIME2, Kraken2, MetaPhlAn;Sequence preprocessing and quality control;Taxonomic classification and diversity analysis; **Practical:** Demo on Practical demo using Galaxy/QIIME2 or Google Colab with sample microbiome data
Day 115	**Practical:** Demo on Epigenomics analysis by Google colab

Unit 24 : Exploring the Future - Prospects, Challenges, Ethics, Reglutions and Career Guidance

Day 116	Emerging trends and technologies in NGS , Challenges and opportunities in advancing NGS technology, Ethical implications of NGS technology, Regulatory frameworks governing NGS research and clinical applications, Best practices for responsible conduct of NGS research
Day 117	NGS opportunities in India and abroad, Institutions and companies for NGS expertise , Career guidance and counselling

Day 118 - Open Discussion Session

Day 119 - Doubt Solving Session

Day 120 - Project Discussion Session

About the Instructor

Ms. Nilofer K Shaikh , PhD

Biotecnika CRO scientist/ ALML & Bioinformatics Specialist

With a strong background in big data analysis using computational approaches in cancer omics data, Ms. Nilofer K Shaikh brings a wealth of experience from MIT ADT University. Her expertise spans cancer research, drug design, molecular dynamics simulation, data mining, and various omics technologies. Proficient in Python, R, and computational methodologies, she has a deep understanding of genomics, metabolomics, proteomics, transcriptomics, pharmacogenomics, and AI for cancer treatment. Her skillset also includes machine learning, MySQL database management, and natural language processing (NLP

About the Instructor

1.
2.

Mr. Prodyot Banerjee

Biotecnika CRO scientist/ CADD, Bioinformatics & Genomics Specialist

Prodyot Banerjee is a seasoned professional in Computer-Aided Drug Designing, Bioinformatics Analysis, and Genomics, boasting rich experience from institutions like CSIR-IGIB, CSIR-CLRI, IIT Madras, and Delhi Technological University.

With an M.Tech in Bioinformatics from Delhi Technological University, Prodyot has excelled in research and development roles, presenting his work at prestigious venues like IIT Kharagpur. His research is published in esteemed journals such as IEEE and Frontiers in Pharmacology, with more underway. Prodyot's GATE 2019 qualification from IIT Madras underscores his dedication to both academic excellence and professional growth. With a proven track record and relentless pursuit of knowledge, he is a valuable asset in bioinformatics, genomics, and computer-aided drug design endeavors.

About the Instructor

Dr. Elamathi Natrajan

Biotecnika CRO scientist/ ALML & Bioinformatics Specialist

Elamathi Natarajan is a dedicated bioinformatician with a robust background in computational biology, data analysis, and genomics. Holding a Doctorate in Bioinformatics from Dr. A.P.J Abdul Kalam Technical University and an MBA in Information Systems Management, she has made significant contributions to the field through both research and teaching.

She has served as an Assistant Professor and Head of Department (HOD) In-Charge at Kalinga University, Raipur, where she excelled in lecturing, research, and departmental management. At Biotecnika Info Labs Pvt Ltd, Bangalore, she played a key role in academic support, enhancing student success through coaching and program development.

Elamathi's expertise includes developing bioinformatics pipelines, conducting quality assessments, and applying machine learning algorithms to genomics data. Recognized for her work, including a Senior Research Fellowship from the Indian Council of Medical Research (ICMR), she continues to drive innovation in bioinformatics and is seeking a new challenge to further advance scientific discoveries.

About the Instructor

Ms. Shubhi Singh

Biotecnika CRO scientist/ Bioinformatics & Computational Drug Design Specialist

Ms. Shubhi Singh is a dedicated researcher and educator having completed her Ph.D. in Biotechnology from SRM Institute of Science and Technology, bringing a strong interdisciplinary background in both computational and experimental research.

She is proficient in molecular docking, ADMET prediction, protein–ligand interaction analysis, and bioinformatics workflows for target identification and drug design. Her expertise includes integrating in silico modeling with wet-lab techniques such as qPCR-based gene expression profiling, microbial culture under stress conditions, and biofilm inhibition assays.

Ms. Singh is also experienced in academic teaching and laboratory supervision, fostering scientific skills and critical thinking among students. She is dedicated to advancing research in antimicrobial resistance and therapeutic development through innovative, data-driven approaches.

About the Instructor

Dr. Neeraj Kumar

Biotecnika CRO scientist/ Computational Biologist & Bioinformatics Specialist

Dr. Neeraj Kumar is a computational biologist and bioinformatician with expertise in AI-driven drug discovery, cheminformatics, and structural bioinformatics. He holds a Ph.D. in Bioinformatics from CSIR-IHBT and AcSIR, India, specializing in machine learning (ML) and deep learning (DL) for virtual screening, drug repositioning, and lead optimization.

His postdoctoral research at Pennsylvania State University focused on developing computational algorithms for immunoglobulin analysis. He has extensive experience in ML/DL frameworks (TensorFlow, PyTorch), cheminformatics tools (RDKit, OpenBabel), and molecular modeling platforms (AlphaFold, Rosetta, GROMACS). He has contributed primarily to AI-guided virtual screening pipeline development, HIV drug discovery, and structural bioinformatics projects.

His research work has been published in esteemed journals including Journal of Cheminformatics, Medicinal Research Reviews, and Computers in Biology and Medicine, with more publications in progress. He has qualified UGC-NET, GPAT, NIPER-JEE, and GATE.

About the Instructor

1.

Dr. Bhupender

Biotecnika CRO scientist/AI/ML, Bioinformatics & Computational Drug Design Specialist

Dr Bhupender Singh is currently working as a Scientist/Trainer at Biotecnika info Labs. He holds a PhD degree specialising in Bioinformatics from Lovely Professional University. During his PhD, he grew as a dynamic researcher with experience as Teaching Assistant and has published his thesis work on Machine Learning and in-silico Drug Discovery in peer-reviewed journals.
With extensive experience at CSIR-Institute of Microbial Technology (IMTECH) as a Project Scientist and Senior Project Associate, he has proven his skills in ML/AI, Cheminformatics, in-silico Drug Discovery, Protein-sequence and Structural Analysis, Database development, supervision and organisation of the Biological-sampling events, gamification of scientific concepts for awareness of children and others. Apart from this, he has also taken scientific administrative roles at Lovely Professional University contributing to PhD thesis submission processes and Plagiarism verification.

Dr Bhupender has published in several peer-reviewed journals and has more than 700 citations. He has active collaborations with reputed National and International Scientists, Academicians, and Researchers. Further, he is also serving as a Reviewer in peer-reviewed scientific journals of Bioinformatics and Computational Biology for more than four years.
Dr Bhupender has joined Biotecnika with an objective to share his learning and experience among growing professionals.

Leading Biopharma Companies Hiring Biologists for Data Science Abroad

Illumina:

Illumina leverages AI and genomic data analytics to improve precision medicine, cancer genomics, and agricultural biotechnology.

Google (DeepMind, Verily)

Google has strong AI/ML research arms (DeepMind) and a dedicated healthcare subsidiary (Verily) actively involved in bioinformatics, drug discovery, and personalized medicine.

Microsoft (Microsoft Research)

Microsoft Research conducts cutting-edge research in AI, including areas relevant to biomedicine. They also offer cloud services and tools for life sciences.

Johnson & Johnson

J&J integrates AI and machine learning across pharmaceutical R&D, medical devices, and healthcare services. They use predictive analytics to streamline clinical trials, optimize patient recruitment, and personalize treatments.

Pfizer

An American multinational pharmaceutical and biotechnology company. They hire bioinformaticians for roles in drug discovery, clinical research, and vaccines development.

23andMe/AncestryDNA

These companies specialize in consumer genomics and utilize data science for genetic research and personalized health insights.

GSK

GSK invests heavily in AI for vaccine development and drug discovery. Their AI-powered analytics platforms enhance clinical trial design, monitor real-world data, and optimize supply chains for pharmaceutical manufacturing.

AstraZeneca

AstraZeneca employs AI and data science for precision medicine, improving drug discovery efficiency and clinical trial accuracy. They integrate machine learning tools to analyze complex disease pathways and predict patient responses to treatments.

Flatiron Health

Flatiron integrates real-world data analytics into oncology research, bridging the gap between clinical trials and routine patient care. They structure data from Electronic Health Records (EHRs) to accelerate cancer treatment insights and improve healthcare delivery outcomes.

Recursion Pharmaceuticals

Recursion combines machine learning, bioinformatics, and automation to analyze massive biological datasets. Their AI models identify new drug targets and repurpose existing drugs for rare diseases, significantly shortening the discovery timeline.

Interdisciplinary Nature

Data Science offers biologists unique chances to integrate their expertise with computational tools. Combining biology, statistics, and machine learning allows them to tackle challenges like genomic analysis, ecosystem modeling, and drug interaction predictions. This synergy fosters innovative solutions in healthcare, agriculture, and environmental science, positioning biologists as essential contributors to data-driven research.

Research Opportunities

Data Science provides biologists with research opportunities for complex questions with computational precision. From bioinformatics to biodiversity modeling, options such as drug discovery through machine learning and omics integration for personalized medicine enhance understanding and drive innovations in healthcare, agriculture, and sustainability.

Data Science is transforming biology, enabling biologists to analyze complex datasets and drive innovation. From genomic data in personalized medicine to enhancing crop yields, this integration has created exciting career paths. Biologists possess essential domain expertise, allowing them to utilize data science effectively. In healthcare, environmental science, and drug discovery, the ability to interpret biological data is crucial for research and technological advancement.

CAREER SCOPE IN DATA SCIENCE FOR BIOLOGISTS

Toll Free - 1800-1200-1818

www.Biotecnika.org

Pharmaceutical and Biotechnology Industries

Pharmaceutical and Biotechnology Industries are transforming drug discovery and precision medicine through Data Science. Skilled biologists analyze vast datasets to identify drug targets and improve development. Machine learning models extract insights from molecular interactions, speeding up innovation. This dynamic field offers biologists various career opportunities, establishing them as key contributors to future healthcare.

Healthcare and Clinical Applications

Data Science transforms healthcare by analyzing large patient datasets to uncover disease trends and treatment outcomes. Skilled biologists can develop AI diagnostic tools, predict patient responses, and identify biomarkers for precision medicine. This approach improves understanding of complex diseases and promotes personalized treatment strategies, enhancing patient care and healthcare efficiency.

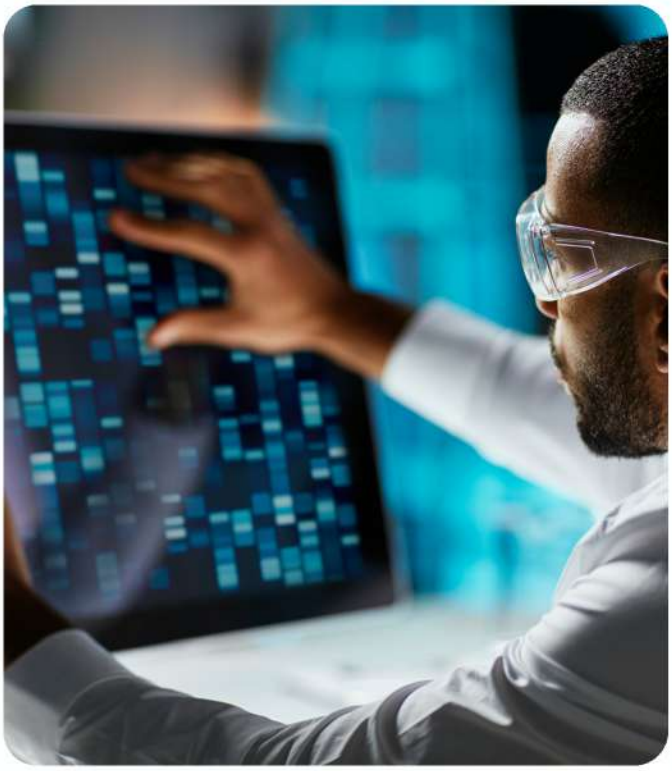
Academia and Education

Biologists skilled in Data Science are in high demand in academia, leading interdisciplinary research that combines computational techniques with biology. They enhance genomics, systems biology, and ecological modeling, while also educating future scientists in necessary computational and analytical skills. This expertise fosters impactful research and innovation in the life sciences.

Data Science and Big Data Analytic

Data Science and Big Data Analytics enable researchers to extract insights from large datasets, addressing challenges in biology, genomics, and healthcare. Utilizing machine learning and statistical analysis, data scientists improve agriculture, analyze patient trends, and predict environmental changes, creating diverse career opportunities in academia, industry, and entrepreneurship within the life sciences.

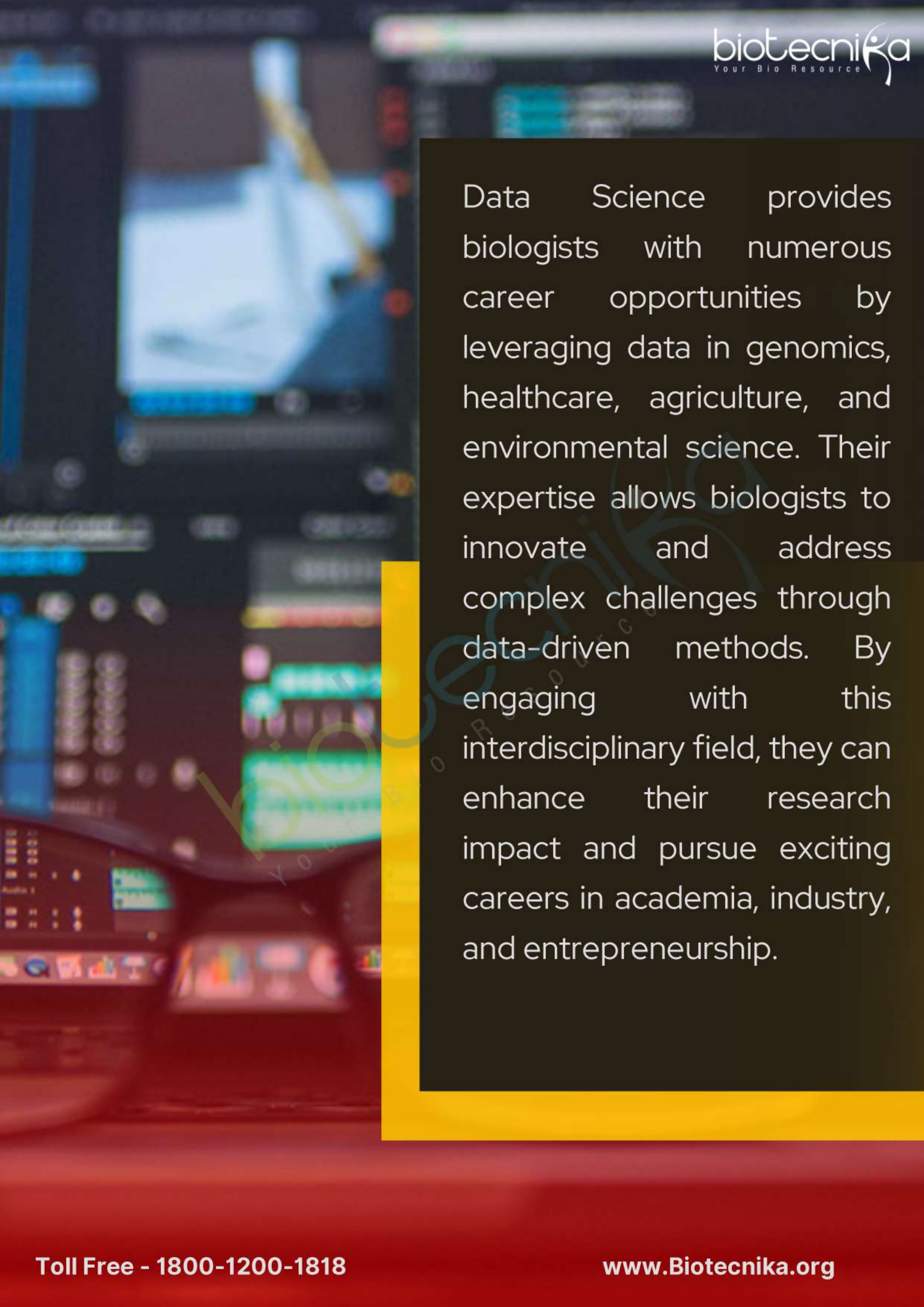
Emerging Technologies



Emerging technologies in data science are transforming biology, creating exciting career opportunities. Biologists are using AI, machine learning, and bioinformatics to revolutionize fields like genomics, drug discovery, agriculture, and healthcare. These innovations enable more precise predictions, enhance research, and drive breakthroughs, opening new paths for biologists to tackle complex challenges and drive innovation.

Global Impact

Data Science is revolutionizing biology globally, enabling breakthroughs in personalized medicine, drug discovery, agriculture, and environmental conservation. By combining computational techniques with biological research, biologists are driving innovation and contributing to global efforts in improving health, sustainability, and well-being.



Data Science provides biologists with numerous career opportunities by leveraging data in genomics, healthcare, agriculture, and environmental science. Their expertise allows biologists to innovate and address complex challenges through data-driven methods. By engaging with this interdisciplinary field, they can enhance their research impact and pursue exciting careers in academia, industry, and entrepreneurship.

How Will This Internship Help You? And What will you learn?

biotecnika
Your Bio Resource

Interns will learn how to use Data Science tools and software to interpret complex biological data, and develop the skills necessary to design and implement their own computational pipelines.

Dive into the refreshing side of science with 30 days, 3 months & 6 months of Data science training! Get a complete 360-degree approach to Data Science with LIVE practicals, assignments & projects

By the end of the program, interns will have gained hands-on experience with cutting-edge techniques and will be well-prepared to pursue a career in Data Science or related fields.

Troubleshooting sessions with seasoned professionals will ensure you have all the support you need to explore this exciting field.

Whether you're just starting out or already have some experience, our Data Science Hands-On Internship is designed to enhance your skills. You'll have the opportunity to learn from industry experts, work with cutting-edge tools and software, and engage in real-world projects. Join us for an exciting journey of learning, professional growth, and hands-on experience!

biotecnika

Your Bio Resource