biotecnika
Your Bio Resource

INDUSTRIAL HANDS-ON TRAINING PROGRAM

AI/ML for CRISPR
Genome Engineering

A 45-Day Hands-On Live Training + PI-Mentored Research Projects

START DATE	DURATION	TIMINGS
3 June 2026	**45 Days + Project Phase**	**7:00 PM – 8:00 PM IST**

Biotecnika Info Labs Pvt. Ltd.
www.biotecnika.org
Bangalore | India

CRISPR Changed Biology. AI Is Changing CRISPR.

The next generation of genome engineers will not just design guide RNAs — they will train models that design them better, predict off-targets faster, and turn weeks of wet-lab trial-and-error into hours of in silico precision.

This 45-day intensive program is built for life sciences students and early-career professionals who want to enter the highest-paying frontier of modern biotech: AI-driven genome editing.

No prior coding experience needed. By the end, you will have built your own end-to-end CRISPR design pipeline — using Python, deep learning, and the same tools that companies like **Profluent, Metagenomi, and insitro** are using right now.

Course Summary — At a Glance

Program Name	AI/ML for CRISPR Genome Engineering
Start Date	3rd June 2026
Duration	45 working days of live training + project work
Class Timings	7:00 PM – 8:00 PM IST
Mode	Live online sessions (with lifetime recorded access)
Daily Format	20–25 min Theory + 35–40 min Hands-on Practical
Project Work	3-month / 6-month / 12-month tracks under PI mentorship
Faculty	5 industry experts and senior researchers in CRISPR & AI/ML
Certification	Certificate of Completion from Biotecnika
Eligibility	Life sciences, biotech, pharma, bioinformatics students & professionals
Prerequisites	Basic biology; no prior coding required

Why Learn AI in CRISPR?

CRISPR was supposed to make gene editing simple. In practice, it created a new problem: choosing the right guide RNA out of thousands of possibilities, predicting whether it will cut the wrong place, and designing therapies safe enough for human trials.

Traditional rule-based design tools (like Rule Set 3) get you only so far. They miss context, ignore chromatin state, and cannot scale to genome-wide screens. This is exactly where AI/ML steps in.

Today, AI is reshaping every layer of CRISPR research:

- **DeepCRISPR** predicts on-target activity with accuracy traditional methods cannot match.
- **DNABERT and transformer models** read DNA the way GPT reads language.
- **Profluent's OpenCRISPR-1** is among the first AI-generated, openly licensed gene editors.
- **LLM agents like CRISPR-GPT** are automating gRNA selection, off-target analysis, and experiment planning across 22+ tasks.

Companies are increasingly hiring professionals who can combine biology with AI skills. They are hiring people who can sit between biology and AI. People who understand both the Cas9 mechanism and the loss function of the model predicting where it cuts.

This program creates exactly that person.

Why Enrol For This Program

▸ You learn both sides

CRISPR biology and machine learning, taught in a way that connects them at every step. Most courses teach one or the other.

▸ You build, not just watch

Every day has a 35–40 minute hands-on practical. By the end, you will have written code for one-hot encoding, GC content scoring, RF/SVM classifiers, CNNs, LSTMs, DNABERT fine-tuning, and a full automated gRNA design pipeline.

▸ You get a real research project under PI mentorship

Not a self-paced capstone, but a structured project with a named faculty mentor and a 3, 6, or 12-month timeline.

▸ You work with the real tools used in industry

Biopython, scikit-learn, Keras, ViennaRNA, CRISPick, MAGeCK, DeepCRISPR, DNABERT, LangChain, and CRISPR-GPT-style LLM agents.

▸ You learn the business and ethics too

FDA guidelines, OpenCRISPR-1 patents, regulatory landscape, and responsible AI in gene editing. This is what employers actually ask about in interviews.

- **You finish with a portfolio, not just a certificate**

  Including your own gRNA design pipeline, model benchmarking reports, off-target predictions, a written policy brief on responsible AI, and a PI-mentored research project.

- **It is built for Indian life sciences students**

  Fresher-friendly, conversational, and priced for accessibility, but with the depth that gets you taken seriously by global biotech companies.

Who Is This Program For?

This program is designed for:

- M.Sc. / B.Tech / M.Tech students in Biotechnology, Bioinformatics, Life Sciences, Genetics, Molecular Biology, Microbiology, Biochemistry, or Pharmacy
- Ph.D. scholars working in molecular biology, gene editing, or computational biology who want to add AI/ML to their toolkit
- Early-career researchers and lab professionals working with CRISPR who feel limited by manual gRNA design tools
- Final-year B.Sc. students with a strong interest in genome engineering and a willingness to learn programming
- Working professionals from biotech, pharma, or clinical research who want to transition into AI-driven biology roles
- Faculty members upgrading their teaching and research portfolios

Prerequisites

- Basic understanding of molecular biology and DNA (high school level is enough — we revise the rest in the initial sessions)
- A laptop with internet access
- Willingness to learn Python — we teach from scratch starting Day 4
- No prior coding experience required
- No prior machine learning experience required

What You Will Learn — Key Outcomes

By the end of the program, you will be able to:

- Explain the molecular biology of CRISPR-Cas9, base editing, and prime editing in depth
- Write Python code to parse genomic data, encode DNA sequences, and compute biological features
- Build and train classical ML models (logistic regression, random forest, SVM) for gRNA efficiency prediction
- Build deep learning models (CNN, LSTM, transformer-based) for on-target and off-target prediction
- Fine-tune DNABERT and use transfer learning on biological sequence data
- Design base and prime editing strategies using AI tools
- Run pooled CRISPR screen analysis with MAGeCK
- Build LLM agents (LangChain) that automate gRNA design pipelines

- Analyse single-cell CRISPR screen data (Perturb-seq, CITE-seq)
- Evaluate models using AUC-ROC, MCC, F1, and SHAP for interpretability
- Write policy briefs on regulatory and ethical aspects of AI in gene editing
- Independently execute a research project under PI mentorship

Program Structure — 7 Weeks, 45 Days

The 45-day live training is organised into 7 thematic weeks that take you from CRISPR fundamentals to advanced transformer-based models and LLM agents. Each session is 1 hour long: 20–25 minutes of focused theory followed by 35-40 minutes of hands-on practical work.

Week	Days	Theme
1	**1 – 7**	Introduction to CRISPR Biology & The Need for AI/ML
2	**8 – 14**	Python & Data Science Bootcamp for Biotech
3	**15 – 21**	Sequence Representation & Core ML for On-Target Prediction
4	**22 – 28**	Deep Learning for gRNA Design
5	**29 – 35**	Off-Target Prediction, Transformers & Advanced Gene Editing
6	**36 – 42**	Industry Integration, Automation & Ethics
7	**43 – 45**	Advanced Topics & Global Trends

Detailed Day-Wise Curriculum

Below is the complete day-by-day breakdown of the 45-day live training. Each day combines structured theory with a hands-on coding or analysis exercise.

Week 1: CRISPR Biology, Applications & Why AI/ML Matters

Day	Topic	Theory (20–25 min)	Practical (35–40 min)
1	**CRISPR-Cas9 Biology**	Bacterial immune system, Cas9 domains, PAM (NGG), DNA repair (NHEJ vs HDR)	Annotate all NGG PAM sites in a 500bp region (manual + simple Python later)
2	**CRISPR Applications**	Gene knockout, knock-in, activation (CRISPRa), interference (CRISPRi), base/prime editing	Literature search: find 3 real-world CRISPR therapies (e.g., sickle cell)
3	**Why AI/ML for CRISPR?**	Limitations of rule-based design, scale of gRNA space, off-target prediction challenge	Compare Rule Set 3 vs. experimental data — see the performance gap
4	**Python Basics for Biologists**	Variables, data types, loops, conditionals	Write reverse complement of a DNA string
5	**Data Structures & File Parsing**	Lists, dicts, sets; reading FASTA/GFF/CSV	Parse FASTA; extract sequence lengths
6	**NumPy for Numerical Data**	Arrays, vectorization, random sampling	Simulate random DNA; compute GC content vectorized
7	**Mini-Project: GC Content Dashboard**	Consolidate Week 1 — but now with Python	Script that computes length, GC%, dinucleotide frequencies for 5 genes

Week 2: Python & Data Science Bootcamp for Biotech

Day	Topic	Theory (20–25 min)	Practical (35–40 min)
8	**Pandas for Genomic Tables**	DataFrame, filtering, groupby, merging	Load CRISPR dataset; filter top 5% efficient gRNAs
9	**Data Visualization**	Matplotlib/seaborn: histograms, boxplots, heatmaps	Plot on-target activity vs. GC content (with regression)
10	**Introduction to Biopython**	Seq, SeqIO, Entrez, pairwise alignment	Fetch human EMX1 gene from NCBI; write FASTA
11	**One-Hot Encoding for DNA**	Convert ATCG to vectors; positional importance	Write one-hot encoding function; visualize as heatmap
12	**K-mer & Dinucleotide Features**	K-mers as features; dinucleotide odds ratio	Compute dinucleotide frequencies for 100 gRNAs
13	**Secondary Structure Features**	RNAfold basics; stem-loops, ΔG	Use ViennaRNA to predict gRNA secondary structure
14	**Data Sources & QC**	CRISPick, CRISPRbrain, CRISPy; batch effects	Query CRISPick API; compare efficiency distributions

Week 3: Core Machine Learning for On-Target Prediction

Day	Topic	Theory (20–25 min)	Practical (35–40 min)
15	**Feature Engineering for gRNAs**	Positional nucleotides, structure counts, GC%	Extract 50+ features from 1000 gRNAs
16	**Train/Test Split & Logistic Regression**	Overfitting, cross-validation, AUC-ROC	Train logistic regression; plot ROC curve
17	**Random Forest**	Ensemble learning; feature importance	Train RF; compare to logistic regression
18	**Support Vector Machines**	Kernel tricks; RBF kernel for sequences	SVM with sequence-derived features
19	**Hyperparameter Tuning**	Grid search, cross-val, MCC, F1 score	GridSearchCV to optimize random forest
20	**Feature Importance & SHAP**	Global vs local interpretability	Identify top 5 predictive features using SHAP
21	**Model Benchmark**	Compare LogReg, RF, SVM	Table of AUC-ROC, MCC, F1 on held-out test set

Week 4: Deep Learning for gRNA Design

Day	Topic	Theory (20–25 min)	Practical (35–40 min)
22	**Neural Networks Basics**	Perceptron, activation functions, backprop	Build dense neural network (Keras) for regression
23	**CNNs for Sequences**	1D convolutions as motif detectors	CNN with Embedding + Conv1D + pooling
24	**DeepCRISPR Architecture**	Multi-channel inputs; on-target + off-target	Load pre-trained DeepCRISPR; score gRNAs
25	**RNNs & LSTMs**	Recurrent connections; long-range dependencies	LSTM model for on-target prediction
26	**CNN vs LSTM Comparison**	Strengths, weaknesses, hybrid models	Compare CNN and LSTM on same dataset
27	**Transfer Learning for CRISPR**	Pre-training on large DNA corpus; fine-tuning	Fine-tune a pre-trained DNA CNN on gRNA data
28	**DL Benchmark**	Compare DL vs traditional ML	Bar chart of R^2 and Pearson correlation

Week 5: Off-Target, Transformers & Advanced Gene Editing

Day	Topic	Theory (20–25 min)	Practical (35–40 min)
29	**Off-Target Effects & Data Sources**	GUIDE-seq, Digenome-seq; mismatches, bulges	Use CRISTA web server to visualize off-target sites
30	**ML for Off-Target Prediction**	Mismatch features; CRISPRoff, CnnCrispr	Run CnnCrispr (Colab) to predict off-target scores
31	**Transformers for DNA**	Self-attention; position encodings; DNABERT	Load DNABERT; extract embeddings for a gRNA
32	**Fine-Tuning DNABERT**	Fine-tuning strategy; adding epigenetic features	Fine-tune DNABERT on a small off-target dataset
33	**Base Editing & AI**	ABE, CBE; bystander edits; CRISPRon-BE	Use BE-Designer to design base editor for a point mutation
34	**Prime Editing & GNNs**	pegRNA design; Graph-CRISPR for structure	Run PrimeDesign; simulate Graph-CRISPR predictions
35	**Week 5 Review**	Compare off-target models; base vs prime editing	Create decision flowchart for model selection

Week 6: Industry Integration, Automation & Ethics

Day	Topic	Theory (20–25 min)	Practical (35–40 min)
36	**High-Throughput CRISPR Screens**	Pooled screens; MAGeCK algorithm	Run MAGeCK on sample screen data; identify essential genes
37	**Delivery Systems & AI**	AAV capsid engineering; LNP design; ML for tropism	Predict AAV capsid tropism (simulated model)
38	**End-to-End Automated Pipeline**	APIs (NCBI, CRISPy); version control (Git)	Script: gene → fetch sequence → design gRNAs → score → ranked CSV
39	**LLM Agents for CRISPR (CRISPR-GPT)**	Planner, Executor, Tool provider; 22+ tasks	Prompt engineering for gRNA selection workflow (LLM as agent)
40	**Regulatory & Ethical Landscape**	FDA guidelines; patents (OpenCRISPR-1); ELSI	Write 1-page policy brief on responsible AI in gene editing
41	**Multi-Omics Integration**	Epigenomics, transcriptomics + gRNA sequence	Train random forest with chromatin accessibility features
42	**Industry Case Studies**	Metagenomi, insitro, Profluent (OpenCRISPR-1)	Analyze published pipeline; design a hypothetical experiment

Week 7: Advanced Topics & Global Trends

Day	Topic	Theory (20–25 min)	Practical (35–40 min)
43	**Protein Engineering for Cas Enzymes + Single-Cell CRISPR Screens**	Generative AI for Cas variants (OpenCRISPR-1); PAMmla; Perturb-seq & CITE-seq	Use PAMmla to predict PAM specificity; run scanpy + MAGeCK-Flute on public Perturb-seq data
44	**In Vivo Delivery Prediction + Custom LLM Agents**	ML for AAV/LNP tropism; LangChain tool calling & memory	Predict tissue tropism from capsid sequence (simulated); build LangChain agent that calls CRISPick + NCBI APIs
45	**Clinical Trial Design Simulators + Future Directions**	Patient selection, off-target risk per genotype; FDA model-informed drug development; epigenetic editing, CRISPRa/i, ethical AI	ML risk stratification on simulated patient genotypes (100 patients); final Q&A

How the Program Works

This is a live, instructor-led program — not a pre-recorded video library. You learn in real time, ask questions in real time, and build alongside the faculty.

PHASE 1 Live Training (45 Days)

Daily 1-hour live sessions from 7:00 PM to 8:00 PM IST, starting 3rd June 2026. Every session is split into structured theory followed by a hands-on practical, so you are not just watching — you are coding, designing, and analysing from Day 1.

All live sessions are recorded and available for lifetime access, so you never lose a class even if you miss it live.

PHASE 2 Project Work Under PI Mentorship

After completing the 45-day training, you move into the project phase — and this is where the program becomes genuinely career-defining.

You choose a research project from a curated list of real, faculty-designed projects across three duration tracks. Each project is assigned to a Principal Investigator (PI) — a faculty mentor who guides you through problem definition, methodology, execution, and final deliverables. You are not figuring things out alone; you are working under structured research supervision.

PHASE 3 Certification & Portfolio

On successful completion of training and your chosen project, you receive your Biotecnika certification along with a documented project portfolio that you can showcase in interviews, Ph.D. applications, and on LinkedIn.

Project Tracks — Choose Your Depth

Every learner selects one project track based on their goals, timeline, and the depth of work they want to commit to. Each track is mentored by a named Principal Investigator from our faculty team.

TRACK 1

3-Month Projects

Focused, single-objective projects ideal for building a strong portfolio piece. Perfect for learners targeting industry roles, internship applications, or a quick high-quality deliverable.

Example projects: *guide RNA efficiency prediction models, AI-based variant identification, CNN-based on-target predictors, AI-driven sgRNA design for cancer driver gene knockout.*

TRACK 2

6-Month Projects

Deeper, multi-stage projects suitable for dissertation-level work or serious portfolio development. Ideal for Ph.D. aspirants and learners wanting research-grade outputs.

Example projects: *off-target site prediction using ML, deep learning-based gRNA design, transformer-based off-target prediction for sickle cell therapy, multi-omics CRISPR efficiency modelling, CRISPR-GPT agents for designing gene therapies.*

TRACK 3

12-Month Projects

Research-grade, publication-ready projects for those targeting Ph.D. applications, research publications, or specialised industry roles. The most rigorous and rewarding track.

Example projects: *multi-task deep learning models for CRISPR design optimisation, AI-driven discovery of novel PAM-specific Cas variants, multi-omics AI platforms for personalised cancer therapy, generative AI for novel CRISPR enzymes, CRISPR-ML benchmarking platforms.*

Project Categories Span Across:

- gRNA design and on-target efficiency prediction
- Off-target prediction and explainable AI
- Deep learning architectures for CRISPR (CNN, LSTM, transformers)
- Base editing and prime editing outcome prediction
- Multi-omics integration (epigenomics, transcriptomics, chromatin accessibility)
- Generative AI for novel Cas enzymes and PAM specificity
- LLM agents and CRISPR-GPT-style automation

- Disease-specific CRISPR therapy design (sickle cell, cystic fibrosis, Duchenne muscular dystrophy, cancer)
- Benchmarking and comparative ML studies

Why the Project Phase Matters

Most online courses end with a certificate. This one ends with proof of work.

Recruiters and Ph.D. admission committees increasingly want to see what you have actually built — not just what you have attended. The project phase is designed exactly for this:

- You work on real research questions, not toy datasets
- You are mentored by a named PI, which adds credibility
- Your final output is publication-ready or interview-ready
- You build a portfolio that differentiates you from candidates with generic ML certifications

This is the difference between "I took a course on CRISPR and AI" and ***"I built a transformer-based off-target predictor under PI mentorship over six months."*** *Employers and admission committees notice.*

Complete List of Available Research Projects

Below is the full catalog of 28 research projects available for the post-training project phase. Each project has been designed by a Principal Investigator from our faculty team and is mapped to a clear duration track. You will work directly under your assigned PI through every stage — problem definition, methodology, execution, and final deliverables.

Track distribution: *9 projects in the 3-Month track • 14 projects in the 6-Month track • 7 projects in the 12-Month track*

#	Project Title	Duration
1	AI-Based Variance Identification and Modelling of Cas Protein	3 Months
2	Analyze the Results of a High-Throughput Screen to Find Genes Linked to a Specific Trait	3 Months
3	Model the Distribution of DNA Repair Outcomes (Indels) After a CRISPR Cut	6 Months
4	Use a Pre-Existing AI Agent to Plan and Design a Simple CRISPR Experiment	6 Months
5	Compare the Performance of Different Machine-Learning Models for a CRISPR Prediction Task	12 Months
6	Guide (gRNA) Efficiency Prediction Model	3 Months
7	Off-Target Site Prediction Using ML	6 Months
8	CRISPR-Cas9 Target Site Classification Using Deep Learning	6 Months
9	Predict Functional Impact of CRISPR-Induced Mutations	6 Months
10	Multi-Task Deep Learning Model for CRISPR Design Optimization	12 Months
11	AI-Based sgRNA Design for Cancer Driver Gene Knockout	3 Months
12	Transformer-Based Off-Target Prediction for CRISPR Therapy in Sickle Cell Disease	6 Months
13	Prediction for Correction of Point Mutations in Cystic Fibrosis	6 Months
14	CRISPR-GPT Agent for Designing Gene Therapies for Duchenne Muscular Dystrophy	6 Months
15	Multi-Omics AI Platform for Personalized CRISPR Therapy in Cancer	12 Months
16	Task-Oriented AI Agent for Automating Data Cleaning in Bioinformatics	3 Months
17	CRISPR-Cas System Classification Using Sequence-Based ML Models	3 Months
18	Ensemble Learning Model for Off-Target Prediction with Explainability	6 Months
19	Autonomous AI Agent for Hypothesis Generation from Research Papers	6 Months
20	Off-Target Effect Prediction in CRISPR Gene Editing with Deep Learning	6/12 Months
21	CNN-Based On-Target gRNA Efficiency Predictor for Human Genomes	3 Months
22	Machine Learning-Based Prediction of CRISPR-Cas9 On-Target gRNA Efficiency Using Sequence-Derived Features and Explainable AI Approaches	3 Months

#	Project Title	Duration
23	Deep Learning-Based Prediction of CRISPR-Cas9 gRNA Efficiency Using CNN and LSTM Models with Sequence and Structural Features	6 Months
24	Deep Learning Model for gRNA Prediction	6 Months
25	AI-Driven Discovery and Characterization of Novel PAM-Specific Cas Variants	12 Months
26	Base-Editing Outcome Predictor: Implement the CRISPRon-BE architecture to predict the efficiency of single-base changes and identify potential bystander edits in a target sequence	3 Months
27	Multi-Omics CRISPR Efficiency Model: Integrate chromatin accessibility data and epigenetics with gRNA sequences to train a Random Forest model that predicts cell-type-specific editing activity	6 Months
28	Custom Cas9 Variant Predictor: Utilize the PAMmla framework to predict the Protospacer Adjacent Motif (PAM) preferences for novel or engineered Cas9 variants	6 Months
29	CRISPR-ML Benchmarking Platform: A Systematic Comparison of CNN, RNN, and Classical ML Performance Across Multiple Datasets	12 Months
30	Generative AI for Novel Enzymes: Designing Custom CRISPR-Cas9 Variants via Scalable ML Engineering	12 Months

Meet Your Faculty

This program is delivered by a team of five senior researchers and Biotecnika CRO scientists who work daily at the intersection of AI/ML and computational biology. Each Principal Investigator brings published research, hands-on industry experience, and deep teaching expertise to your training and project mentorship.

Dr. Elamathi Natarajan

Bioinformatics & AI Specialist | *Bioinformatics CRO Scientist, Biotecnika*

Dr. Elamathi Natarajan is a seasoned Bioinformatics and AI expert with 13+ years of experience at the intersection of life sciences, computational biology, and data science. She currently leads projects in NGS data analysis, transcriptomics, multi-omics, and AI-driven drug discovery at Biotecnika, Bangalore.

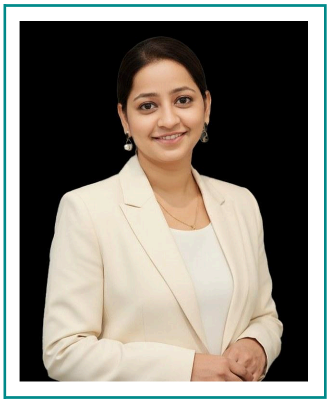
Her expertise lies in transforming complex biological data into actionable insights using machine learning, bioinformatics pipelines, and advanced data analytics. She has strong hands-on experience in RNA-seq analysis, differential gene expression, variant calling, biomarker discovery, disease prediction, precision medicine applications, and machine learning model development for biological data.

Previously, she has served in academic leadership roles including Assistant Professor and Head (In-Charge), and has been associated with premier research organizations such as ICMR-NIMR, NIPGR, and DRDO. She holds a Ph.D. in Bioinformatics and was awarded the Excellence in Research Award (2024).

Core Expertise:

AI & Machine Learning in Life Sciences • RNA-seq, NGS & Multi-omics Analysis • Python, R & Bioinformatics Pipelines • Drug Discovery (Virtual Screening & Docking) • Biomarker Discovery & Disease Modeling • Transcriptomics & Precision Medicine

Google Scholar: https://scholar.google.com/citations?user=SntWOEcAAAAJ&hl=en



Dr. Nilofer K. Shaikh

AI/ML & Bioinformatics Specialist | Biotecnika CRO Scientist

Dr. Nilofer K. Shaikh, Ph.D. (Bioinformatics, MIT ADT University, Pune), is an accomplished computational biologist specialising in big data analysis of cancer omics using advanced computational and AI-driven approaches. Her expertise spans genomics, proteomics, transcriptomics, metabolomics, and pharmacogenomics, with a focus on biomarker discovery, precision cancer medicine, and data-driven drug development.

She has extensive experience in multi-omics integration (WGS, RNA-seq, scRNA-seq, proteogenomics) and applies machine learning, deep learning, and NLP for cancer biomarker identification and drug sensitivity prediction. She is proficient in Python, R, Biopython, MySQL, Weka, Orange, Tableau, Seurat, Scanpy, Bioconductor, and AWS cloud computing.

Her research is reflected in publications in Current Topics in Medicinal Chemistry, Journal of Biomolecular Structure & Dynamics, and International Journal of Swarm Intelligence Research. She holds an Indian Patent (Design No. 384325-001) and has presented at the NVIDIA GPU Technology Conference (Silicon Valley, 2019), among other international forums.

Core Expertise:

AI/ML & Deep Learning for Cancer Omics • Multi-omics Integration • Precision Oncology • Biomarker Discovery • Computational Drug Discovery • Cloud Computing for Genomics

Google Scholar: https://scholar.google.co.in/citations?hl=en&user=iH5LZyUAAAAJ

Mr. Prodyot Banerjee

CADD, Bioinformatics & Genomics Specialist | Biotecnika CRO Scientist

Mr. Prodyot Banerjee holds a Master of Technology in Bioinformatics from Delhi Technological University, with national qualifications in GATE and CSIR-NET. He is a seasoned practitioner of Computer-Aided Drug Design (CADD), employing in-silico methodologies — from structure and ligand-based virtual screening to pharmacophore modelling, molecular docking, and Molecular Dynamics simulations — to navigate vast chemical spaces and identify high-potency lead compounds.

His research delves deeper into the blueprint of life through advanced genomics data analysis. He applies sophisticated algorithms for phylogenetic reconstruction and genome-wide association studies (GWAS) to decipher the molecular etiology of complex diseases.

His work has been published in SCI-indexed peer-reviewed journals with a cumulative impact factor of 43.87. He has presented at IIT Kharagpur and the National Institute of Immunology, New Delhi, and has contributed to projects at CSIR-IGIB, IIT Madras, AIIMS Delhi, and CSIR-CLRI Chennai. He has mentored over 200 students from undergraduates to postdoctoral scholars.

Core Expertise:

Computer-Aided Drug Design (CADD) • Molecular Docking & Dynamics Simulations • Genomics & GWAS • Phylogenetics • Virtual Screening • Pharmacophore Modeling

Google Scholar: https://scholar.google.com/citations?hl=en&user=KG_2tiIAAAAJ
LinkedIn: https://www.linkedin.com/in/prodyot-banerjee-86728a185

Dr. Shubhi Singh

Bioinformatics & Computational Drug Design Specialist | Biotecnika CRO Scientist

Dr. Shubhi Singh holds a Ph.D. in Biotechnology from SRM Institute of Science and Technology and brings a strong interdisciplinary background in both computational and experimental research. Her expertise spans molecular docking, molecular dynamics simulations, structural bioinformatics, immunoinformatics, and whole-genome analysis, along with AI and machine learning applications in drug discovery.

She leverages ML models for virtual screening, drug repurposing, and prediction of protein-ligand interactions, enabling faster identification of therapeutic candidates against antibiotic-resistant pathogens. She is proficient in Python (NumPy, Pandas, Matplotlib) and works with cheminformatics tools like RDKit, Open Babel, and DeepChem.

Her research has been published in journals including Molecular Diversity, Journal of Biomolecular Structure and Dynamics, and Current Microbiology. She is a recipient of the Tata Trust Scholarship for Women, has earned multiple oral presentation awards at national and international conferences, and serves as a peer reviewer for journals in bioinformatics and AI-driven drug discovery.

Core Expertise:

Molecular Docking & Dynamics Simulations • Structural Bioinformatics • Immunoinformatics • AI/ML for Drug Discovery • Cheminformatics (RDKit, DeepChem) • Whole-Genome Analysis

Google Scholar: https://scholar.google.com/citations?hl=en&user=UDPFzpsAAAAJ

Dr. Bhupender Singh

AI/ML, Bioinformatics & Computational Drug Design Specialist | CRO Scientist & AI-ML Trainer, Biotecnika

Dr. Bhupender Singh holds a Ph.D. in Bioinformatics from Lovely Professional University, with specialisation in Machine Learning and in-silico Drug Discovery. He currently serves as a CRO Scientist & AI-ML Trainer at Biotecnika, where he trains professionals in AI/ML for Computational Biology, NGS, and multi-omics data analysis.

With nearly a decade of research experience, he has previously worked at the Bioinformatics Centre, CSIR-Institute of Microbial Technology (IMTECH), Chandigarh, contributing to high-impact national projects. His expertise spans AI/ML model development for biological datasets, cheminformatics, in-silico drug discovery, protein sequence and structural analysis, database development, and Python/R-based bioinformatics workflows.

Dr. Bhupender has authored 20+ peer-reviewed publications, received 700+ citations, and holds an h-index of 10. He has contributed to book chapters in machine learning, drug discovery, metagenomics, and omics sciences, and serves as a peer reviewer for journals in Bioinformatics and Computational Biology.

Core Expertise:

AI & Machine Learning in Life Sciences • Computational Biology & Bioinformatics • NGS & Multi-omics Analysis • Python & R for Biological Data Science • Cheminformatics & In-silico Drug Discovery • Protein Sequence & Structural Analysis

Google Scholar: https://scholar.google.co.in/citations?hl=en&user=705KZLIAAAAJ

Tools & Technologies You Will Master

This program is built on the actual tools used in industry and academic research. By the end, these will be standard parts of your daily workflow.

Category	Tools
Programming	Python, NumPy, Pandas, Matplotlib, Seaborn
Biology Libraries	Biopython, ViennaRNA, Scanpy
Machine Learning	scikit-learn, SHAP, XGBoost
Deep Learning	Keras, TensorFlow, PyTorch
Transformer Models	DNABERT
CRISPR-Specific Tools/Concepts/Datasets	CRISPick, CRISPRbrain, DeepCRISPR, CnnCrispr, CRISTA BE-Designer, PrimeDesign, MAGeCK, MAGeCK-Flute
LLM Agents	LangChain, CRISPR-GPT-style workflows
Data Sources	NCBI Entrez, CRISPRbrain, CRISPick API
Version Control	Git (basics)

Hands-On Projects You Will Build During Training

These projects are completed during the 45-day live training, in addition to your final PI-mentored research project:

- A GC content and dinucleotide frequency dashboard (Week 1)
- A gRNA feature extractor processing 1000+ guides (Week 3)
- Benchmark report comparing logistic regression, random forest, and SVM (Week 3)
- A CNN-based on-target activity predictor (Week 4)
- A fine-tuned DNABERT model for off-target prediction (Week 5)
- An automated end-to-end pipeline: gene name → fetch sequence → design gRNAs → score → ranked CSV (Week 6)
- A LangChain LLM agent that calls CRISPick and NCBI APIs (Week 7)
- A 1-page policy brief on responsible AI in gene editing (Week 6)
- ML risk stratification model on simulated patient genotypes (Week 7)

Career Outcomes — Where This Takes You

Roles This Program Prepares You For:

- Computational Biologist (CRISPR / Gene Editing focus)
- AI/ML Scientist in Biotech
- gRNA Design Specialist
- Bioinformatics Scientist (Genome Engineering)
- Research Associate in CRISPR therapeutics
- Cell Engineering Computational Scientist
- Genomic Data Scientist
- Computational Geneticist
- Faculty / Research Scholar in computational genome engineering

Industries Hiring for These Skills:

Gene therapy companies, CRISPR diagnostics, agricultural biotech, pharmaceutical R&D, contract research organisations (CROs), cell and gene therapy startups, academic research labs, and AI-biotech crossover companies.

Certification

On successful completion of all 45 days of training and your assigned PI-mentored project, participants receive a Certificate in AI/ML for CRISPR Genome Engineering from Biotecnika Info Labs Pvt. Ltd.

The certificate is verifiable and suitable for adding to LinkedIn, your CV, and Ph.D. or job applications. It serves as documented evidence of both your training completion and your PI-mentored research work.

Frequently Asked Questions

Q. I am from a pure biology background. Can I really learn AI/ML in 45 days?

Yes. The first two weeks are designed exactly for you — we start from "what is a variable" and build up to neural networks. The pacing is fresher-friendly.

Q. I already know Python. Will this still be useful?

Yes. Weeks 3 onwards assume Python knowledge and go deep into ML, DL, transformers, and LLM agents — content that is rarely available in a single Indian program.

Q. How does the project phase work?

After the 45-day training, you select a project from our curated list across 3-month, 6-month, or 12-month tracks. Each project is assigned to a Principal Investigator who mentors you through the entire research process.

Q. Can I choose my project duration?

Yes. You select the track that fits your timeline and goals — a 3-month project for portfolio-building, 6-month for dissertation-level depth, or 12-month for publication-ready research.

Q. Do I need a powerful computer?

No. Most practicals run in Google Colab (free GPU access). A standard laptop with internet is sufficient.

Q. How is this different from a generic AI/ML course?

Generic ML courses teach you to predict house prices and classify images. This program teaches you to predict gRNA efficiency and detect off-target cuts — using real biological data, real tools, and real research papers, with a real research project at the end.

Q. What if I miss a live class?

All sessions are recorded and available for lifetime access.

Q. Will the recordings be available afterwards?

Yes, lifetime access to all recordings and course material.

Why Biotecnika

Biotecnika has trained over thousands of life sciences students and professionals across India. Our certifications are designed by working scientists and educators who understand both Indian academic realities and global industry expectations.

This program is part of our growing AI-in-Life-Sciences vertical, alongside our PG Diploma in AI/ML & Data Science, AI for Drug Designing workshops, and bioinformatics certifications.

READY TO BEGIN?

Genome Editing Is No Longer Just A Wet-Lab Discipline.

The future belongs to scientists who can write the algorithms that design the experiments.

Starting 3rd June 2026

Limited seats. Closed-door cohort.

ENROL NOW

Get In Touch

Biotecnika Info Labs Pvt. Ltd.
www.biotecnika.org
Bangalore, India