

Prithvi Rajan R

M.Sc. Bioinformatics, Freie Universität Berlin | Berlin, Germany

prithvi.rajan300@gmail.com | +49 176 7909 5187 | prithvirajan.com | github.com/prithvirajanR

SUMMARY

Bioinformatics M.Sc. student at Freie Universität Berlin with a healthcare data-systems background and hands-on project work in protein language models, molecular docking, single-cell benchmark auditing, and reproducible genomics workflows. Interested in RA, working-student, or internship roles where biological data analysis, Python workflow engineering, and careful validation matter.

PROFESSIONAL EXPERIENCE

Research Assistant – Freie Universität Berlin October 2025 – Present

- Conducting a supervised computational biology research project on protein sequence evolution using protein language models, embedding-based scoring, and reproducible Python workflows.

Data Analyst – Kaiser Permanente Jan 2024 – Oct 2024

- Cleaned, queried, and prepared recurring analytics outputs from hospital admissions and claims datasets.
- Built SQL reporting pipelines and Power BI dashboards for healthcare utilization and claims-pattern review.
- Documented metric definitions and validation checks so recurring reports were easier to reproduce and audit.

Financial Analyst – Havenspire Jul 2021 – Jun 2024

- Supported part-time investment-analysis work through market research, financial data review, and reporting.
- Built a business-facing analysis habit alongside technical database and healthcare-data roles.

Oracle Database Administrator – Accenture Dec 2021 – Jan 2024

- Supported production Oracle 11g/19c database systems in high-availability healthcare environments.
- Worked across migration support, SQL tuning, Data Guard, access controls, patching, and recovery checks.
- Assisted with backup/recovery validation, environment checks, and release support for sensitive database systems.

IBM DB2 Database Administrator – Accenture Feb 2021 – Sept 2021

- Maintained DB2 database environments with focus on performance, schema reliability, and secure operations.
- Supported database capacity planning, OS-level maintenance coordination, patching, and monitoring workflows.
- Reviewed query behavior, indexing patterns, and operational alerts to support stable production performance.

TECHNICAL SKILLS

Programming	Python, R, SQL, PL/SQL, Bash, JavaScript
Bioinformatics	ESM models, variant-effect scoring, epistasis, Vina, GNINA, fpocket, AnnData, Snakemake
Data/ML	pandas, NumPy, scikit-learn, PyTorch, Jupyter, Power BI, statistical analysis
Systems	Linux, Git, conda, Docker, Oracle 11g/19c, IBM DB2, ETL, SQL optimization

EDUCATION

Freie Universität Berlin, *Master of Science in Bioinformatics* Present

- Current coursework: Statistical Modeling, Machine Learning, Algorithm Development

Karunya University, *Bachelor of Technology in Bioinformatics*

- Relevant technical skills: Python, IBM DB2, Oracle RDBMS, drug discovery studio, PyRx, R programming, virtual molecular screening, pharmacophore modeling, toxicity analysis, molecular dynamics, cheminformatics, medical coding, 3D QSAR, health informatics

BIOINFORMATICS WORK

Protein Variant-Effect Analysis with ESM Models [GitHub](#)

Benchmarked ESM-1v and ESM-2 on 14,483 APP deep-mutational scanning variants from ProteinGym, comparing model scores against experimental DMS measurements and reviewing score behavior across model variants.

Molecular Docking Workflow Engineering [GitHub](#)

Built StrataDock, a Python/Streamlit workflow for receptor cleanup, ligand preparation, pocket selection, Vina/GNINA docking, validation cases, 3D review, and exportable run reports with traceable inputs and outputs.

Single-Cell Perturbation Benchmark Auditing [GitHub](#)

Developed PerturbGuard, a Python package for AnnData quality checks, split leakage detection, controls, confounding, weak target effects, unsupported benchmark claims, CLI reports, tests, and PyPI releases.

Reproducible Omics and Genome Workflows [BRCA](#) | [WGS](#) | [viral](#) | [RNA-seq](#)

Worked on BRCA multi-omics subtype analysis and Snakemake workflows for bacterial WGS, adenovirus assembly, and RNA-seq processing, emphasizing declared inputs, traceable outputs, and reviewable reports.