

Cheminformatics Pipeline Automation & Reproducibility — Hands-on

Build **production-grade cheminformatics pipelines** for QSAR/ADMET, docking, and virtual screening that are **automated, reproducible, and auditable**. You will orchestrate workflows using **Snakemake/Nextflow**, containerize tools with **Docker/Singularity**, track data & models via **Git/DVC**, and publish **FAIR**, re-runnable reports with provenance and CI checks.

Cheminformatics Pipeline Automation & Reproducibility

[Help Desk](#) · [WhatsApp](#)

Session Index

[Session 1 — Workflow Orchestration](#) [Session 2 — Containers & Environments](#) [Session 3 —
Versioning, CI & FAIR Provenance](#) [Session 4 — Mini Capstone: One-Click Run](#)

Session 1

Fee: Rs 19800 [Apply Now](#)

Workflow Orchestration: Snakemake / Nextflow

DAG thinking & rule design

[inputs/outputs & wildcards](#) [resources & checkpoints](#)
[caching & resumability](#)

Cheminformatics tasks as stages

[standardize](#) → [featurize](#) → [model](#) [dock](#) → [rescore](#) →
[rank](#) [report & archive](#)

Parallelism & executors

local/HPC/cloud runners **job arrays & retries** **cost-aware scheduling**

Session 2

Fee: Rs 22800 Apply Now

Containers, Environments & Parameter Sweeps

Portable environments

Docker/Singularity images **conda/mamba + lockfiles**
GPU runtime notes

Parameterization at scale

grid/random/bayes sweeps **config files & templating**
artifact naming & hashing

Operational concerns

logging & metrics **quotas & storage hygiene** **alerts & failure hooks**

Session 3

Fee: Rs 25800 Apply Now

Data/Model Versioning, CI & FAIR Provenance

Git/DVC best practices

data/model remotes **params.yaml & metrics.json**
reproducible seeds

CI pipelines & tests

lint/unit/smoke tests **container build checks** **artifact promotion**

FAIR & provenance

metadata/data cards **lineage graphs** **immutable archives**

Session 4

Fee: Rs 29800 Apply Now

Mini Capstone: One-Click QSAR/VS Run

Assemble an end-to-end pipeline (std → featurize → model/dock → report)

Theory + Practical

Re-runs & parameterized batches

configs & sweeps **CI-triggered builds** **artifact registry**

Deliverables

pipeline repo (Git/DVC) **container image** **run report (HTML/PDF)**