# Big Data Technologies in Precision Medicine Training Course

Professional Development Training Course Outline

| | | | |
|---|---|---|---|
| **Category** | Research and Data Analysis | **Duration** | 10 Days |
| **Base Fee** | $3,000 USD | **Delivery Mode** | Online / On-site Hybrid |

## Course Introduction

Big Data Technologies in Precision Medicine Training Course is an advanced, interdisciplinary program designed to equip healthcare professionals, bioinformaticians, data scientists, and IT experts with the skills necessary to harness big data analytics, artificial intelligence (AI), machine learning (ML), and cloud computing in transforming the landscape of precision medicine. As healthcare becomes increasingly data-driven, this course offers practical and technical know-how in processing, analyzing, and interpreting vast and complex datasets—ranging from genomic to clinical data—to deliver personalized, predictive, and preventative care.

This hands-on course integrates real-world medical case studies, high-performance computing tools, and cutting-edge technologies such as Apache Hadoop, Apache Spark, Python for bioinformatics, data visualization, and health informatics. Learners will explore the intersection of biomedical data science, genomic sequencing, and EHR integration, building robust solutions for complex clinical problems. With growing investments in digital health, this course is vital for those aiming to stay competitive and impactful in modern healthcare ecosystems.

## Programme Curriculum

**Big Data Technologies in Precision Medicine Training Course**

**Introduction**

Big Data Technologies in Precision Medicine Training Course is an advanced, interdisciplinary program designed to equip healthcare professionals, bioinformaticians, data scientists, and IT experts with the skills necessary to harness big data analytics, artificial intelligence (AI), machine learning (ML), and cloud computing in transforming the landscape of precision medicine. As healthcare becomes increasingly data-driven, this course offers practical and technical know-how in processing, analyzing, and interpreting vast and complex datasets—ranging from genomic to clinical data—to deliver personalized, predictive, and preventative care.

This hands-on course integrates real-world medical case studies, high-performance computing tools, and cutting-edge technologies such as Apache Hadoop, Apache Spark, Python for bioinformatics, data visualization, and health informatics. Learners will explore the intersection of biomedical data science, genomic sequencing, and EHR integration, building robust solutions for complex clinical problems. With growing investments in digital health, this course is vital for those aiming to stay competitive and impactful in modern healthcare ecosystems.

**Course Objectives**

1. Understand the foundations of big data analytics in healthcare.
2. Explore data governance, ethics, and regulatory frameworks in precision medicine.
3. Utilize Apache Hadoop and Spark for biomedical data processing.
4. Apply machine learning algorithms to genomic and clinical datasets.
5. Integrate AI tools in personalized treatment plans.
6. Perform real-time analytics with streaming data in healthcare.
7. Master data wrangling and feature engineering for EHR systems.
8. Develop predictive models for patient risk stratification.
9. Analyze large-scale omics datasets using Python and R.
10. Visualize multidimensional biomedical data using advanced tools.
11. Implement cloud-based solutions for scalable data storage.
12. Collaborate in interdisciplinary precision medicine projects.
13. Design end-to-end precision medicine workflows with big data tools.

**Target Audiences**

1. Healthcare Data Scientists
2. Bioinformatics Researchers
3. Clinical IT Professionals
4. Precision Medicine Program Leaders
5. Medical Researchers and Geneticists
6. Health Informatics Analysts
7. Graduate Students in Biomedical Fields
8. AI and Machine Learning Engineers in Health Tech

**Course Duration:** 10 days

**Course Modules**

**Module 1: Introduction to Big Data in Precision Medicine**

- Definition and significance of big data in healthcare
- Overview of precision medicine landscape
- Sources of biomedical big data
- Key challenges in big data integration
- Opportunities in data-driven care models
- **Case Study**: Integrating genomic and EHR data for personalized diabetes management

**Module 2: Data Acquisition and Management in Healthcare**

- Data formats in genomics, proteomics, and EHR
- Interoperability standards (FHIR, HL7)
- Data cleaning and transformation
- Metadata and annotation standards
- Tools for biomedical data acquisition
- **Case Study**: Handling high-throughput genomic data from multiple platforms

**Module 3: Hadoop for Biomedical Data Processing**

- Overview of Hadoop ecosystem
- HDFS and MapReduce fundamentals
- Data ingestion with Sqoop and Flume
- Using Hive and Pig for structured queries
- Performance tuning in Hadoop clusters
- **Case Study**: Processing cancer genomics data using Hadoop

**Module 4: Apache Spark for Clinical Data Analytics**

- Spark architecture and components
- PySpark for large-scale data processing
- Spark SQL and DataFrames
- MLlib for machine learning pipelines
- Real-time streaming with Spark Streaming
- **Case Study**: Predicting ICU patient deterioration using Spark MLlib

**Module 5: Cloud Computing in Precision Medicine**

- Benefits of cloud adoption in healthcare
- AWS, Google Cloud, and Azure for genomics
- Cloud-based genomics pipelines
- Security and compliance in cloud computing
- Scalable storage and compute environments
- **Case Study**: Cloud-based genome assembly using AWS EC2 and S3

**Module 6: Machine Learning for Precision Medicine**

- Supervised vs unsupervised learning
- Feature selection in biomedical datasets
- Classification and clustering techniques
- Model validation and evaluation metrics
- Use of ML tools: Scikit-learn, TensorFlow
- **Case Study**: ML-based breast cancer subtype prediction

**Module 7: Artificial Intelligence and Decision Support**

- Deep learning architectures in medicine
- Natural language processing (NLP) in EHRs
- AI-driven diagnostics and treatment suggestions
- Reinforcement learning for health policies
- Explainable AI and ethical considerations
- **Case Study**: AI-based clinical decision support for sepsis prediction

**Module 8: Omics Data Integration and Analysis**

- Genomics, transcriptomics, proteomics overview
- Integrative omics platforms and tools
- Network analysis and pathway mapping
- Dimensionality reduction techniques
- Statistical methods for multi-omics data
- **Case Study**: Integrating RNA-seq and proteomic data in cancer prognosis

**Module 9: Electronic Health Records (EHR) Analytics**

- EHR data structure and types

- Temporal data analysis and visualization
- Data linkage and longitudinal studies
- Predictive modeling using EHR
- NLP for unstructured EHR data
- **Case Study**: Predicting hospital readmission from EHR analytics

**Module 10: Data Visualization for Biomedical Insights**

- Visual analytics tools: Tableau, Python, R
- Genomic and patient data dashboards
- Interactive visualization for clinical use
- Heatmaps, scatterplots, networks
- Best practices in health data storytelling
- **Case Study**: Visualizing rare disease trends across population cohorts

**Module 11: Ethical, Legal, and Regulatory Considerations**

- HIPAA, GDPR, and patient data rights
- Informed consent and data ownership
- Bias in algorithms and data sets
- Ethical AI and fairness in precision medicine
- Data sharing and reproducibility
- **Case Study**: Ethical challenges in sharing pediatric genomic data

**Module 12: Predictive Modeling in Precision Medicine**

- Risk prediction and scoring systems
- Time-to-event (survival) analysis
- Disease progression modeling
- ML for drug response prediction
- Causal inference methods
- **Case Study**: Predicting cardiovascular event risks using multimodal data

**Module 13: Mobile and Wearable Data in Precision Health**

- Types of mobile and wearable devices
- Continuous health monitoring
- IoT data integration with EHR
- Behavioral and lifestyle data analytics
- Mobile health interventions
- **Case Study**: Using Fitbit data to monitor post-surgery recovery

**Module 14: Building Scalable Pipelines and Workflows**

- Workflow management tools (Nextflow, Snakemake)
- Docker and containerization
- CI/CD for health analytics
- Parallel processing and job scheduling
- Workflow reproducibility
- **Case Study**: Scalable COVID-19 genomic surveillance pipeline

**Module 15: Capstone Project and Practical Implementation**

- Problem identification and scoping
- Dataset acquisition and cleaning
- Tool and model selection

- Implementation and validation
- Reporting and interpretation
- **Case Study**: Building a full-stack predictive model for oncology therapy response

**Training Methodology**

- Interactive lectures with real-time demonstrations
- Hands-on lab sessions with guided walkthroughs
- Capstone project for applied learning
- Peer discussion forums and expert Q&A
- Downloadable tools, datasets, and code templates

**Register as a group from 3 participants for a Discount**

**Send us an email: info@fineskilltrainingcenter.org or call** *+254769199797*

***Certification***

*Upon successful completion of this training, participants will be issued with a globally- recognized certificate.*

***Tailor-Made Course***

*We also offer tailor-made courses based on your needs.*

**Key Notes**

**a.** The participant must be conversant with English.

**b.** Upon completion of training the participant will be issued with an Authorized Training Certificate

**c.** Course duration is flexible and the contents can be modified to fit any number of days.

**d.** The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

**e.** One-year post-training support Consultation and Coaching provided after the course.

**f.** Payment should be done at least a week before commence of the training, to Fineskill Training Center account, as indicated in the invoice so as to enable us prepare better for you.

## Upcoming Scheduled Sessions

| Start Date | End Date | Location | Price |
|---|---|---|---|
| **23 Nov 2026** | 04 Dec 2026 | Online | $2,000 |
| **23 Nov 2026** | 04 Dec 2026 | Physical | $3,000 |
| **23 Nov 2026** | 04 Dec 2026 | Physical | $6,000 |
| **23 Nov 2026** | 04 Dec 2026 | Physical | $6,000 |

| Start Date | End Date | Location | Price |
|---|---|---|---|
| **23 Nov 2026** | 04 Dec 2026 | Physical | $12,000 |
| **23 Nov 2026** | 04 Dec 2026 | Physical | $9,000 |
| **23 Nov 2026** | 04 Dec 2026 | Physical | $12,000 |
| **23 Nov 2026** | 04 Dec 2026 | Physical | $12,000 |

Ready to enroll or request a customized program? Book online or contact us:

**Register Online Now**

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com