Basic Security and Compliance Framework for a Big Data System

1. Use Case

This project presents a simple framework for securely handling sensitive data in a Big Data system (such as Spark or Hadoop), with basic compliance with data privacy laws like GDPR (Europe) and HIPAA (USA).

Use Case Example:

A hospital stores thousands of patient records using a big data system. These records must be protected from unauthorized access, and data privacy regulations must be followed.

2. Architecture Diagram (Text Description)

[Data Sources] -> [Data Ingestion Layer] (TLS Encryption) -> [Big Data Storage] (AES Encryption at rest) -> [Data Processing (Spark/Hadoop)] -> [Access Layer] (Role-Based Access Control) -> [Analytics/Reporting]

3. Security & Compliance Practices

- Encryption in Transit: Use TLS to protect data while it is being transferred.
- Encryption at Rest: Store data securely using AES encryption.
- Role-Based Access Control: Only authorized users (e.g., doctors, admins) can access sensitive records.
- Anonymization of Data: Remove or mask names, phone numbers, etc. when used for analytics.
- Logging & Monitoring: Log who accesses what and detect unusual activity.
- Compliance Audits: Run checks regularly to ensure GDPR and HIPAA rules are followed.

4. Conclusion

This simple framework shows how basic encryption, access control, and data masking can help secure a Big Data system and ensure privacy law compliance. While more complex systems exist, this forms a strong foundation for secure data handling.