**Capstone Project Proposal: Fraud Detection System for JPMorgan Chase**

**1. Project Title:**  
Fraud Detection System for JPMorgan Chase

**2. Problem Statement:**  
JPMorgan Chase, a global financial leader, processes millions of transactions each day. With the rise of sophisticated cyber threats, traditional fraud detection methods struggle to identify new and evolving fraud patterns in real-time. This project will develop a real-time fraud detection system leveraging machine learning models to enhance security and reduce financial losses.

**3. Objectives:**

* Ingest and preprocess transaction data from varied sources.
* Develop and evaluate supervised and unsupervised fraud detection models.
* Create a real-time anomaly detection and alerting pipeline.
* Build interactive dashboards for fraud monitoring.
* Optimize models to maximize recall while minimizing false positives.

**4. Scope:**

* Data ingestion from CSV, SQL, API, and NoSQL.
* ML techniques: Logistic Regression, Random Forest, Isolation Forest, Autoencoders.
* Real-time scoring simulation with streaming tools.
* Dashboard creation using Power BI or Tableau.
* Deployment on local or cloud-based infrastructure.

**5. Methodology:**

1. **Data Ingestion & Cleaning** – Handle missing data, normalize formats, and remove inconsistencies.
2. **Exploratory Data Analysis** – Identify fraud-prone transaction types, locations, devices.
3. **Feature Engineering** – Transaction velocity, IP-location mismatches, device change flags.
4. **Model Development** – Train supervised models on labeled fraud data; apply unsupervised methods for anomaly detection.
5. **Real-Time Pipeline** – Process incoming transactions, compute fraud risk scores, trigger alerts.
6. **Visualization** – Develop dashboards for transaction trends, fraud alerts, and KPIs.
7. **Evaluation & Optimization** – Assess using precision, recall, F1-score, ROC-AUC; fine-tune hyperparameters.

**6. Deliverables:**

* Processed dataset ready for analysis.
* EDA report with key insights.
* Fraud detection models with documented metrics.
* Real-time monitoring pipeline.
* Interactive dashboard.
* Final project report and presentation.

**7. Expected Outcomes:**

* Robust real-time fraud detection with high detection accuracy.
* Actionable intelligence for security teams.
* Scalable and adaptable architecture for evolving fraud tactics.

**8. Timeline:**

* Weeks 1–2: Data ingestion and preprocessing.
* Weeks 3–4: EDA and feature engineering.
* Weeks 5–6: Model building and evaluation.
* Weeks 7–8: Real-time pipeline and dashboard.
* Week 9: Model optimization.
* Week 10: Report writing and presentation.

####