Ideation Phase
Brainstorm & Idea Prioritization Template

Date	31 January 2025
Team ID	SWTID-2026-5418
Project Name	Insurance Fraud Detection Using Maching Learning
Maximum Marks	4 Marks

Project Brainstorm: Insurance Fraud Detection System

1. Problem Statement

Insurance companies face a significant challenge in the accurate and timely detection of fraudulent claims. The high volume of claims coupled with increasingly complex fraud patterns leads to substantial financial losses and delays in processing genuine claims.

2. Idea Generation

This section outlines potential solutions categorized by area of focus:

Category	Ideas
Detection Techniques (Models)	Machine Learning (Random Forest), Deep Learning (Neural Networks), Rule-based systems, Hybrid systems (ML + rules), Anomaly detection, Logistic Regression baseline.
Data Improvement & Preprocessing	Feature Selection (identifying top predictors), Feature Engineering (date extraction, encoding), Handling Class Imbalance (SMOTE), Outlier Detection, Data Scaling & preprocessing pipeline.
System Features & Output	Web-based Dashboard, Real-time Prediction API, Confidence Score Display, Feature Importance Visualization, Explanation System (justification for fraud flag).
Advanced Enhancements	Fraud Risk Scoring System (0–100 scale), Auto-alert for high-risk claims, Model Retraining Pipeline, Performance Monitoring Dashboard, Bias Detection & Fairness Check.

3. Idea Grouping

Group	Associated Ideas
Core ML Solution	Random Forest, Feature Selection, Standard Scaling
Accuracy Improvement	SMOTE, Outlier Removal, Feature Engineering
Deployment	Flask Web App, API, Dashboard
Explainability	Confidence Score, Feature Importance
Advanced Features	Risk Score, Monitoring, Bias Detection

4. Idea Prioritization Matrix (Importance vs. Feasibility)

Priority Quadrant	Ideas
High Importance + High Feasibility (IMPLEMENT)	Random Forest Model, Feature Selection (Top 8 features), Standard Scaling, Web Application, Confidence Score.
High Importance + Medium Feasibility	SMOTE for imbalance, Explainable AI (SHAP values), Fraud risk scoring meter.
Medium Importance + High Feasibility	Model comparison (KNN, SVM, Logistic Regression), Improved UI/UX dashboard.
Future Enhancements (High Effort)	Deep Learning model, Real-time enterprise integration, Continuous model retraining.

5. Final Selected Concept

Develop a Machine Learning-based Insurance Fraud Detection System utilizing the **Random Forest** algorithm, coupled with **feature optimization**. The system will be deployed via a **web dashboard** to provide fraud predictions, a **confidence score**, and **explanatory insights**.