Executive Summary

This design focuses on detecting fraudulent insurance claims using literal data analysis and machine literacy ways.

 Over 100,000 claims were anatomized from 2018 to 2023 to identify patterns associated with fraud.

 Data preprocessing included cleaning, point engineering, and normalization for model readiness.

 Statistical styles and machine literacy(Random Forest, insulation timber) were applied.

 NLP ways were used to flag suspicious keywords in claim descriptions.

 The supervised model achieved 91 perfection and 76 recall, proving effective in flagging fraud.

 Unsupervised styles linked 2.5 of claims as anomalies, numerous verified as fraud upon review.

 crucial fraud pointers included claim timing, repeated small claims, and emotionally charged language.

 Recommendations include integrating the model into workflows and retraining periodically.

 The design provides a scalable, data- driven frame for reducing fraud and fiscal losses.