



## LEARNATHON 4.0

# GROUP-14 (TEAM NAME - CONCEPT SQUARD)

Title Name

AI-Powered Fraud Detection in Auto Insurance: Predictive Modelling for Smarter  
Claims Management

### PRESENTED BY

Mr. Ayushman Panigrahy

Mr. Debasish Mishra

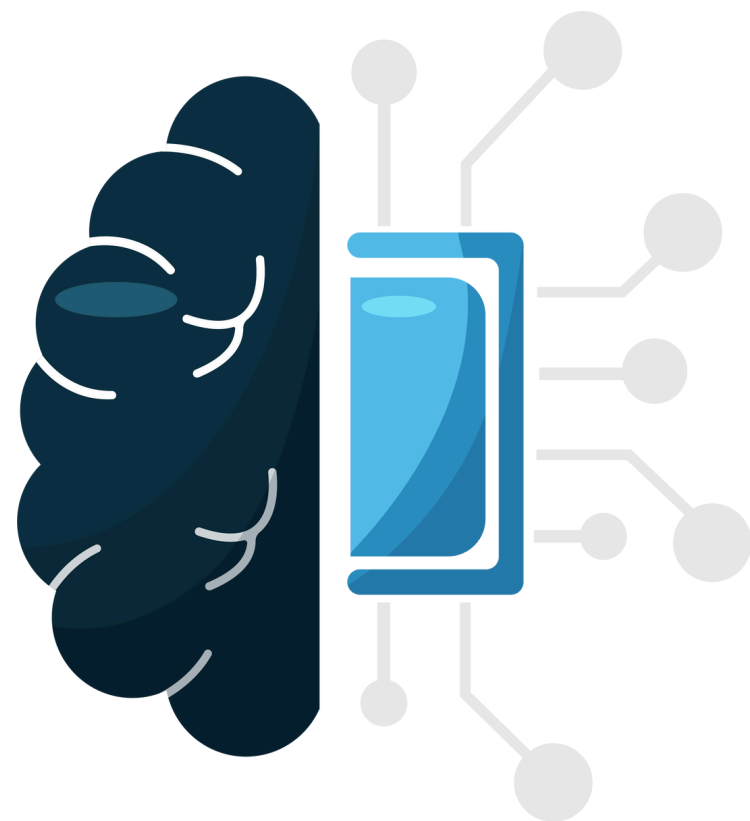
Mr. A. Ajit

Mr. Mihir Ranjan Aich










# Problem Statements

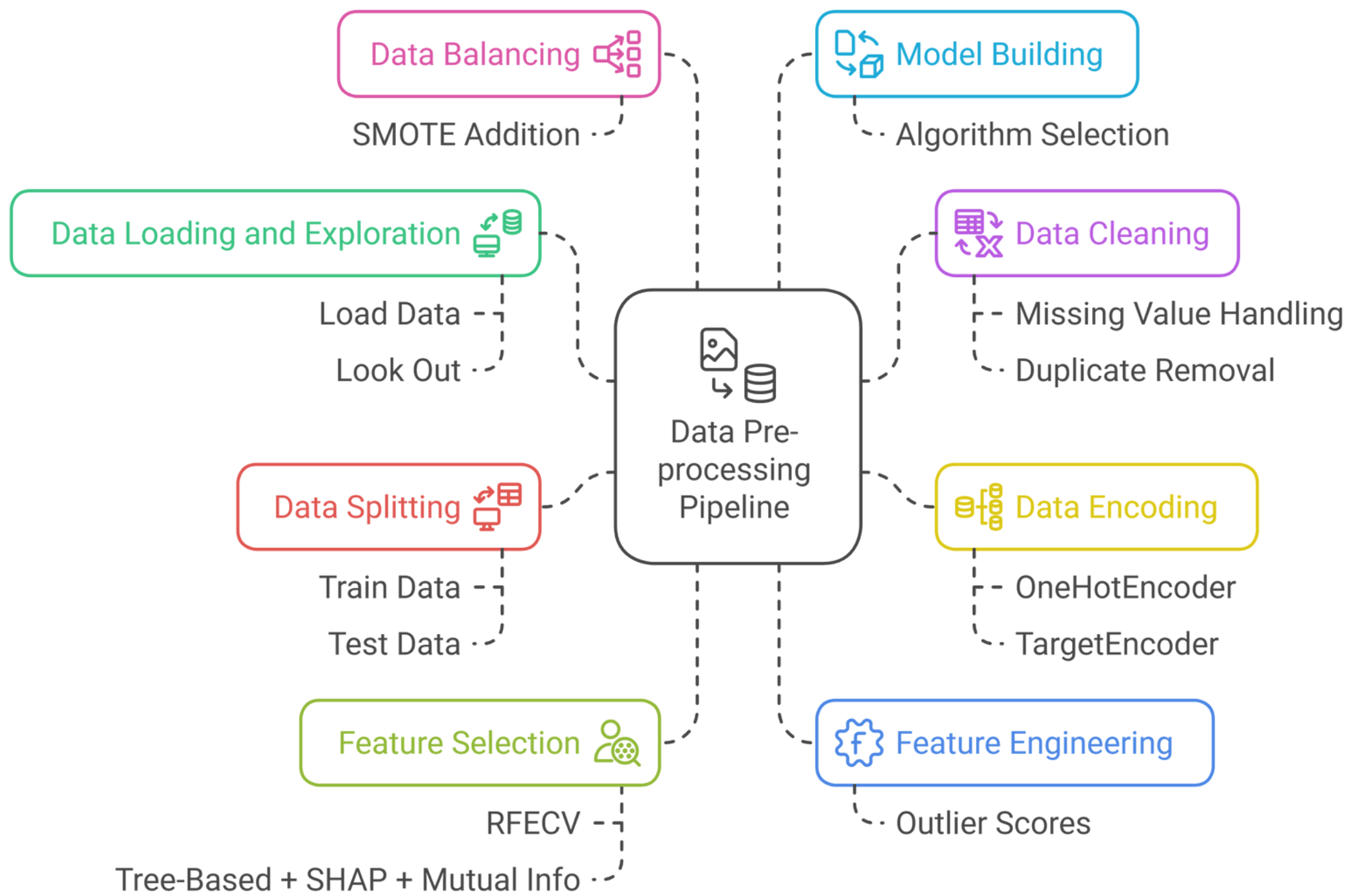
## AI-Powered Fraud Detection in Auto Insurance: Predictive Modeling for Smarter Claims Management



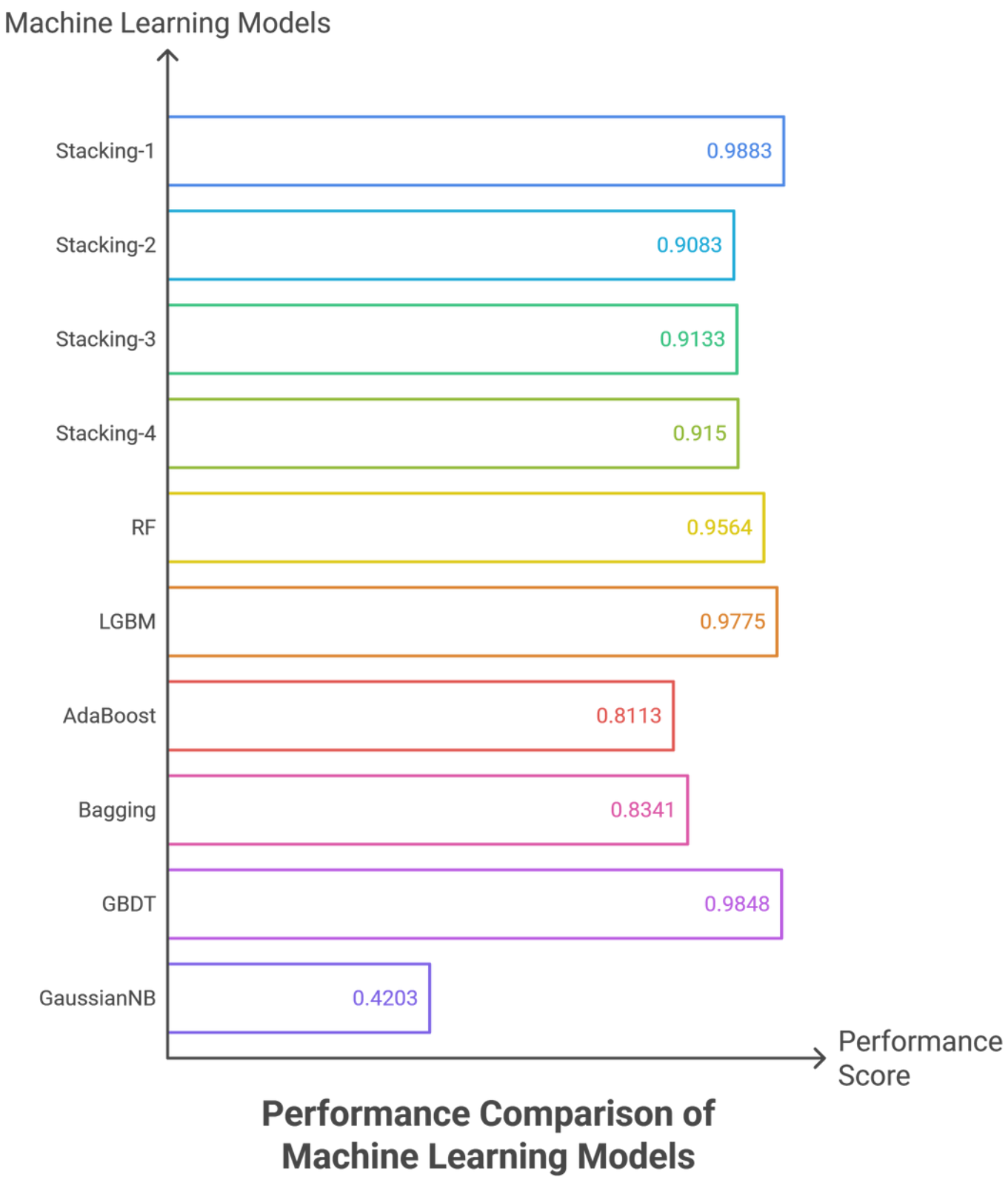
### What We Are Solving (Our AI-Based Approach)

-  Automatically predicts the likelihood of a claim being fraudulent using machine learning models.
-  Analyzes critical features such as customer age, policy number, vehicle price, accident area, and more.
-  Provides a real-time fraud prediction system for smarter and faster decision-making.
-  Helps to reduce manual effort, enabling investigators to focus only on high-risk claims.
-  Supports insurers in minimizing financial losses by flagging fraud early in the process.
-  Delivers interactive dashboards and visual analytics for deeper insight into fraud patterns.
-  Improves operational efficiency, risk management, and overall claims processing.

# DATA PER-PROCESSING PIPRLINE OF OUR APPROACH



# DIFFERNT MODEL USED IN OUR APPROACH



# COMPREHENSIVE EVALUATION OF MULTIPLE MACHINE LEARNING MODELS

Metric	Stacking-1 (RF+GBDT+LGBM)	Stacking-2 (GBDT+LGBM+RF)	Stacking-3 (RF+GBDT+LGBM)	Stacking-4 (RF+LGBM+GBDT)	RF	LGBM	AdaBoost	Bagging	GBDT	GaussianNB
Accuracy	0.9883	0.9083	0.9133	0.915	0.9564	0.9775	0.8113	0.8341	0.9848	0.4203
Precision	0.9916	0.8972	0.8874	0.8986	0.85	0.9165	0.6068	0.627	0.9422	0.2907
Recall (Sensitivity)	0.914	0.9068	0.9319	0.9211	1	1	0.6705	0.8102	1	0.9349
Specificity	0.9034	0.9097	0.8972	0.9097	0.9421	0.9701	0.8574	0.8419	0.9799	0.2515
F1 Score	0.9027	0.902	0.9091	0.9097	0.9189	0.9564	0.637	0.707	0.9702	0.4434
ROC AUC	0.9687	0.9725	0.9708	0.9756	0.9991	0.9998	0.766	0.8992	0.9999	0.7897
Balanced Accuracy	0.9087	0.9082	0.9145	0.9154	0.9711	0.9851	0.764	0.8261	0.9899	0.5932
Cohen's Kappa	0.8161	0.8159	0.8264	0.8294	0.8894	0.9413	0.5099	0.5938	0.9601	0.1068
Log Loss	0.3433	0.2309	0.2312	0.2305	0.2113	0.1271	0.6107	0.4058	0.1124	1.6994
MSE	0.0917	0.0917	0.0867	0.085	0.0436	0.0225	0.1888	0.1659	0.0152	0.5797
RMSE	0.0028	0.3028	0.2944	0.2915	0.2088	0.15	0.4345	0.4074	0.1231	0.7614
MAE	0.0917	0.0917	0.0867	0.085	0.0436	0.0225	0.1888	0.1659	0.0152	0.5797
R²	0.9315	0.6315	0.6516	0.6583	0.7656	0.879	0.0147	0.1079	0.9185	-2.1165



# DATA ANALYSIS

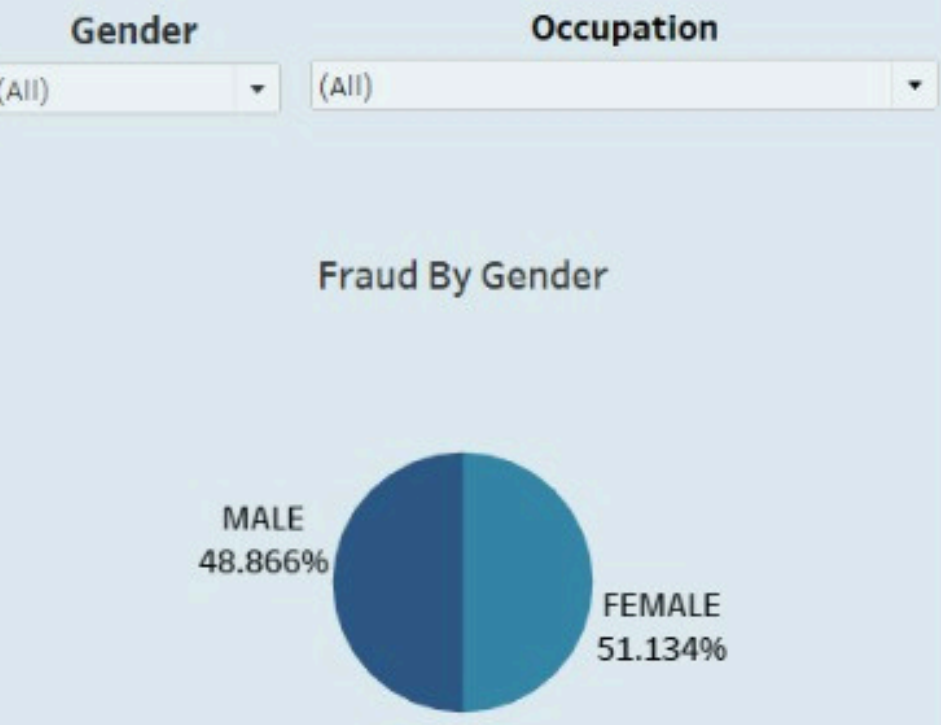
## AUTO INSURANCE FRAUD ANALYSIS

2.576  
Claim Delay Days

₹ 558,051,339.21  
Total Claim

₹ 13,710.21  
Avg Vehicle Cost

Policy\_State  
(All)



Auto Model and Fraud

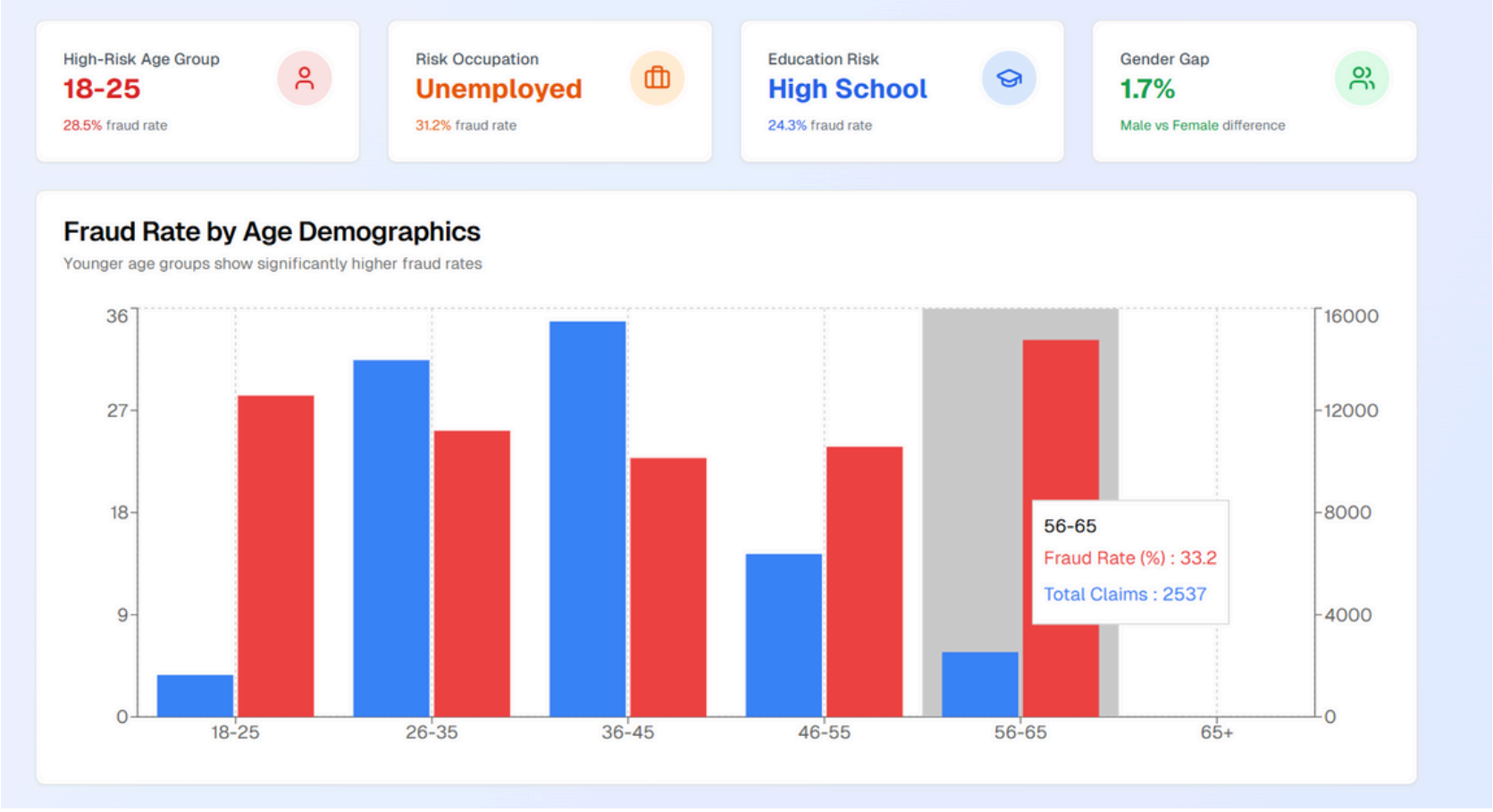
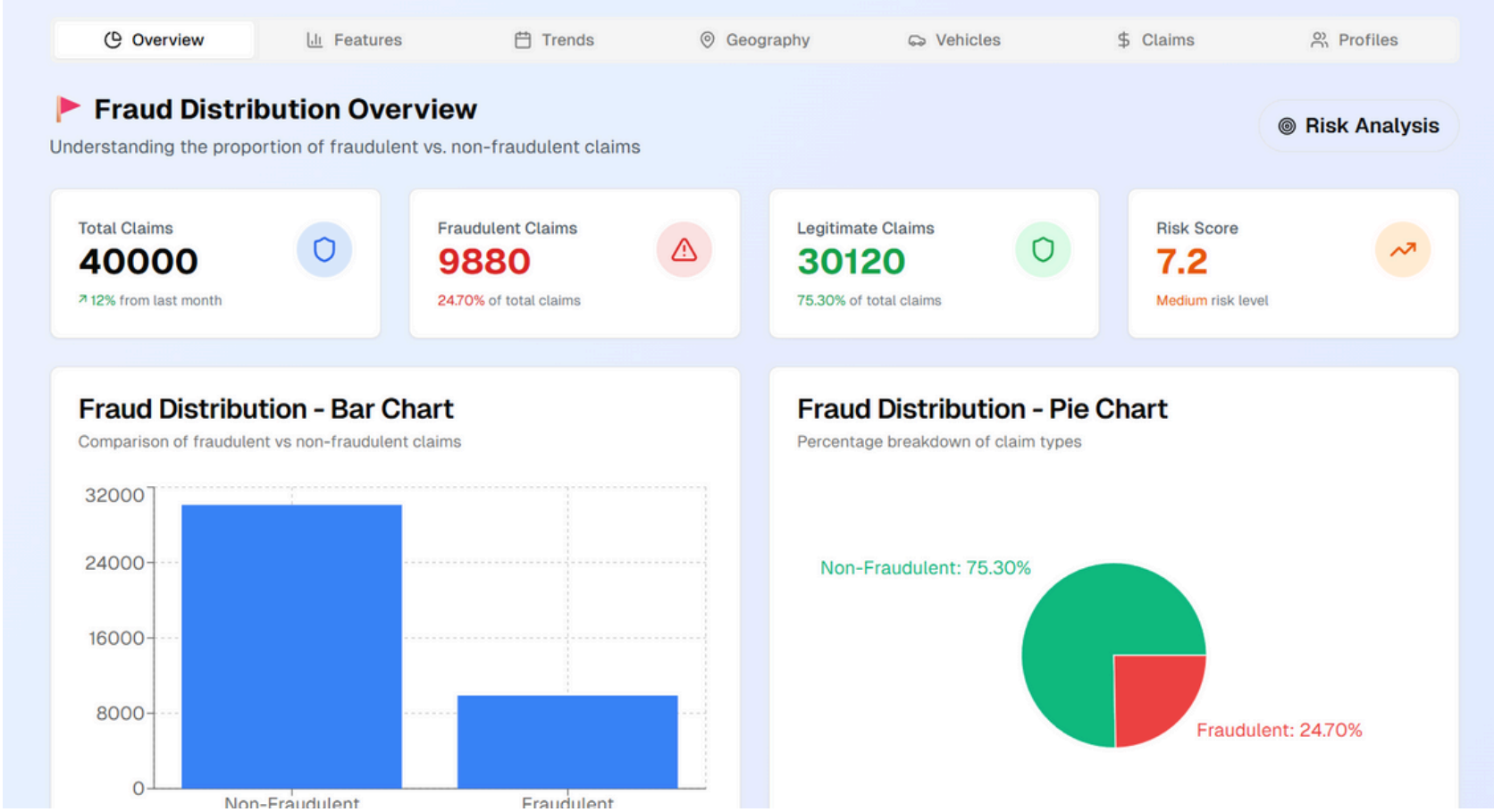
RAM 1,751 31.01%	Jetta 1,399 29.16%	E400 1,158	95 1,064	92x 1,047	93 1,007			
Wrangler 1,663	Passat 1,291	Malibu 1,151	Maxima 966	X5 885	TL 828			
Neon 1,504	Legacy 1,286	Forrester 1,121	Ultima 954		Fusion 625	X6 625		
A3 1,489	A5 1,212	F150 1,109	Silverado 947	Impreza 840	C300		M5	
MDX 1,443	Pathfinder 1,188	Camry 1,088	Civic 889	CRV 836	3 Series			

Collison Vs Severity

Accident_Seve..	Front Colli..	Rear Collis..	Side Collis..	Unknown
Major Damage	2,075	2,418	2,202	
Minor Damage	378	496	273	387
Total Loss	319	695	403	
Trivial Damage				234



# WEBSITE FOR REALWOLD PROBLEM SOLUTION



The dashboard is organized into several tabs for streamlined fraud analysis. The Overview tab summarizes total claims, fraud percentages, and key indicators through dynamic KPIs and charts. The Features tab displays the most influential variables affecting fraud predictions, using importance plots and correlation visuals. In Trends, users can analyze how fraud patterns evolve over time, with timelines and distribution plots. The Geography tab visualizes the regional distribution of fraud cases using interactive maps and location filters. The Vehicles tab breaks down fraud statistics by vehicle type, age, and price to highlight high-risk categories. The Claims section provides insights into fraud based on claim reasons, amounts, and durations, helping spot abnormal behaviors. Lastly, the Profiles tab presents customer demographics like age, occupation, and income level, showing how different profiles correlate with fraud likelihood.

**Thank you  
very much!**