

## Model Development Phase Template

Date	20-Jun-25
Team ID	SWTID1751617613
Project Title	Auto Fraud Detection
Marks	5 marks

Feature	Category	Justification for Inclusion
months_as_customer	Customer Profile	New customers may have a higher likelihood of fraudulent claims.
age	Customer Profile	Age can influence risk behavior; useful for profiling.
policy_bind_date	Policy Info	Claims made soon after policy start may indicate fraud.
policy_state	Policy Info	Fraud patterns vary geographically.
policy_csl	Policy Info	High coverage limits may encourage exaggerated claims.
policy_deductable	Policy Info	Low deductibles might lead to more frequent and potentially fraudulent claims.
policy_annual_premium	Policy Info	High premiums may correlate with large or fraudulent claims.
umbrella_limit	Policy Info	Higher umbrella limits may be targeted for high-value fraud.
insured_sex	Demographics	Could contribute marginally to risk profiling.
insured_education_level	Demographics	Education may correlate with claim behavior and honesty.
insured_occupation	Demographics	Certain occupations might have fraud risk patterns.
insured_relationship	Demographics	May affect motivations or likelihood for fraudulent behavior.
capital-gains	Financials	Large gains may affect claim patterns or provide insights into financial motivation.

Feature	Category	Justification for Inclusion
capital-loss	Financials	High losses may motivate fraudulent claims for recovery.
incident_date	Incident Info	Helps detect patterns or clusters in time (e.g., month-end fraud).
incident_type	Incident Info	Certain types (e.g., theft, collision) are more commonly faked.
collision_type	Incident Info	Specific collision types (e.g., sideswipe) have higher fraud tendencies.
incident_severity	Incident Info	Exaggerated severity is a red flag for fraud.
authorities_contacted	Incident Info	No contact with police despite severe incident is suspicious.
incident_state	Incident Info	Regional trends and fraud hotspots vary.
incident_city	Incident Info	Localized clusters of fraud can exist.
incident_hour_of_the_day	Incident Info	Fraud patterns may show correlation with odd hours.
number_of_vehicles_involved	Incident Info	Unusual numbers in claims may indicate staging.
property_damage	Claim Details	Missing or inconsistent property damage is a fraud indicator.
bodily_injuries	Claim Details	Fake injuries are common in fraud cases.
witnesses	Claim Details	Lack of witnesses in major incidents may be suspicious.
police_report_available	Claim Details	Absence of police report in serious cases is a red flag.
total_claim_amount	Claim Amounts	Core indicator to assess fraud likelihood based on value.
injury_claim	Claim Amounts	Often inflated in fraudulent claims.
property_claim	Claim Amounts	Discrepancy between damage and claim value indicates fraud.

Feature	Category	Justification for Inclusion
vehicle_claim	Claim Amounts	Helps in identifying vehicle-related claim exaggerations.
auto_make	Vehicle Info	Some brands/models are fraud-prone or high-value targets.
auto_model	Vehicle Info	Specific models might have consistent fraud patterns.
auto_year	Vehicle Info	Older vehicles might be more frequently involved in fraudulent claims.