Model Development Phase Template

|  |  |
| --- | --- |
| Date | 18 June 2024 |
| Team ID | 739990 |
| Project Title | Auto Insurance Fraud Detection |
| Maximum Marks | 5 Marks |

**Feature Selection Report Template**

This report provides an overview of the feature selection process for the [Auto Insurance Fraud Detection]. The goal of feature selection is to identify the most relevant and impactful features that contribute to the model's performance.

|  |  |  |  |
| --- | --- | --- | --- |
| **Feature** | **Description** | **Selection**  **(Yes/No)** | **Reason** |
| months\_as\_  customer | Number of months the customer has been with the company | Yes | Important for understanding customer loyalty and tenure. | |
| age | Age of the insured | Yes | Often correlates with risk profiles. |
| policy\_number | Unique identifier for each policy | Yes | Unique Identification number different for each person | |
| policy\_bind\_date | Date when the policy was initiated | No | May not directly impact fraud prediction. |
| policy\_state | State where the policy is issued | No | Geographic location can affect claims. |
| policy\_csl | Policy coverage limit | | No | Determines claim amount and severity. |
| policy\_deductable | Deductible amount on the policy | No | Affects claim amounts and patterns. |
| policy\_annual\_premium | Annual premium paid | Yes | Directly impacts financial risk assessment. |
| umbrella\_limit | Additional liability coverage limit | No | Indicates extended coverage and risk. |
| insured\_zip | ZIP code of the insured | Yes | Geographic location covered by policy\_state. |
| insured\_sex | Gender of the insured | No | Often not a significant factor in claims. |
| insured\_education\_level | Education level of the insured | No | May not directly correlate with claims |
| insured\_occupation | Occupation of the insured | No | Occupation can influence risk but may be complex to model. |
| insured\_hobbies | Hobbies of the insured | No | Unlikely to directly impact claims. |
| insured\_relationship | Relationship of the insured with policyholder | No | Typically not a significant factor in claims. |
| capital-loss | Capital losses recorded | No | Financial data not directly related to insurance claims. |
| capital-gains | Capital gains recorded | Yes | Financial data directly related to insurance claims. |
| incident\_date | Date of the incident | No | Important for understanding claim timing. |
| incident\_type | Type of incident | No | Affects claim type and severity. |
| collision\_type | Type of collision | No | Affects claim severity and cost. |
| incident\_severity | Severity of the incident | No | Crucial for assessing claim severity. |
| authorities\_contacted | How authorities were contacted | No | Process around claim filing. |
| incident\_state | State where the incident occurred | No | Geographic impact on claims. |
| incident\_city | City where the incident occurred Covered by | No |  |
| incident\_hour\_of\_the\_day | Hour of the day when the incident occurred | Yes | Timing can influence claims processing. |
| number\_of\_vehicles\_involved | Number of vehicles involved in the incident. | Yes | Affects claim severity and complexity |
| property\_damage | Whether property damage was reported | No | Indicates claim type and cost. |
| bodily\_injuries | Number of bodily injuries reported. | Yes | Affects claim severity and cost. |
| witnesses | Number of witnesses to the incident. | No | Indicates claim complexity and support. |
| police\_report\_available | Whether a police report was available. | No | Legal documentation for claims processing. |
| total\_claim\_amount | Total amount claimed | Yes | Target variable for prediction. | |
| injury\_claim | Amount claimed for injuries | Yes | Component of total claim amount. | |
| property\_claim | Amount claimed for property damage | No | Component of total claim amount. |
| vehicle\_claim | Amount claimed for vehicle damage | No | Component of total claim amount. |
| auto\_make | Make of the insured vehicle | No | Vehicle details impact claim severity. |
| auto\_model | Model of the insured vehicle | No | Vehicle details impact claim severity. |
| auto\_year | Year of the insured vehicle | Yes | Vehicle age can affect claims. |
| fraud\_reported | Whether fraud was reported. | No | Target variable for model building |