

Feature Selecting

대표적인 보험사기 징후 (Fraud Indicator: FI)

보험계약이 월 마감일에 근접하여 체결

계약자의 월 납입보험료의 수준이 가입자들의 월 평균 납입 보험료를 상회

보험계약이 손해가 발생하기 전에 단기간으로 가입

사고목격자가 전혀 없음

고물차에 고보장형 상품을 가입

경미한 교통사고에 비해 과다한 치료비 청구

[표 4-3] 보험사기 징후의 데이터 변환

보험사기 징후	사용데이터
보험계약이 월 마감일에 근접하여 체결	보험계약일, 업무 마감일
계약자의 월 납입보험료의 수준이 가입자들의 월 평균 납입 보험료를 상회	계약자 월 납입보험료, 가입자 평균 납입 보험료
보험계약이 손해가 발생하기 전에 단기간으로 가입	클레임 청구 전 각 월별 보험가입 여부
사고목격자가 전혀 없다	사고 목격자 존재 여부, 목격자수, 목격자와 사고자와의 관계
고물차에 고보장형 상품을 가입	보장금액/차량가액 차량연식별 동급차량 보장금액
경미한 교통사고에 비해 과도한 진료비 청구	진료비/수리금액, 사고규모별 평균 진료비
서로 연관 없던 탑승객이 같은 변호사를 선임	차량 탑승객별 관련 제3자

표 2: 보험사기 패턴을 반영하여 생성된 입력변수

변수명	변수 설명
N_OF_INSUR	보험 가입 갯수
N_OF_KIND	보험 종류 갯수
N_OF_CLAIM	보험 청구 횟수
MAIN_GOOD	주 보험 종류 코드
CAUS_CODE_YN	사고원인에서 사기 빈도가 높은 코드 이면 “Y”
DMND_RESN_CODE_YN	청구사유가 입원 또는 장애이면 “N”
VALI_HOSP_MAX	보험 청구 중에서 입원/통원 최대 일수
HOSP_DVSN_YN	병원 구분에서 요양병원 또는 한방병원이면 “Y”
DIFF_AMOUNT_MAX	보험 청구에서 청구액과 지급액 차이의 최대값
DIFF_AMOUNT_MIN	보험 청구에서 청구액과 지급액 차이의 최소값
FP_RISK	보험설계사 리스크(추정 값)
SALE_CHNL_YN	보험판매 채널에서 설계사, 법인이 아니면 “Y”
PAYM_TERM_YN	보험료 납입 기간이 25개월 미만이면 “Y”
MAX_PAY_DF	최대보험료 납입일과 청구일 간의 기간 차이
CHANG_FP_YN	보험 청구 중 설계사가 변경된 청구가 있으면 “Y”

표 3: 변수 선택과정에서 선정된 입력변수

변수명	변수 설명
CUST_FRAUD_YN	Target
SEX	성별
AGE	나이
RESL_COST	주택 가격
RESL_TYPE_CODE	거주지 형태
FP_CAREER	보험가입자의 FP 경력 “Y” = yes, “N” = no.
CUST_RGST	최초 고객 등록일 부터 현재까지의 월 수
OCCP_GRP	직업코드 8개
TOTALPREM	납입 총액
MINCRDT	신용등급 최소값
MAXCRDT	신용등급 최대값
WEDD_YN	결혼 여부
MATE_OCCP_GRP	배우자 직업코드
CHLD_CNT	자녀수
LTBN_CHLD_AGE	막내 자녀 나이
MAX_PRM	추정 소득
CUST_INCM	고객 소득
RCBASE_HSHD_INCM	고객 가구 추정소득
RESL_CD1	사고원인 코드
RESL_NM1	사고결과 코드
PMMI_DLNG_YN	실손 처리여부
HEED_HOSP_YN	금감원 유의 병원 여부

<표 5-3> 변수선정

구분	변수명	비고
전체변수	보험사기 적발 여부 ⁷³⁾ , 사기유형, 사기형태 연령, 성별, 거주지, 직업, 월 평균 보험료, 보험가입 건수, 일당담보금액, 환수여부, 환수유형, 처벌유형, 계약상태, 모집채널	
특성변수	사고유형, 사기형태, 환수여부, 환수유형, 처벌유형	보험사기자 정보