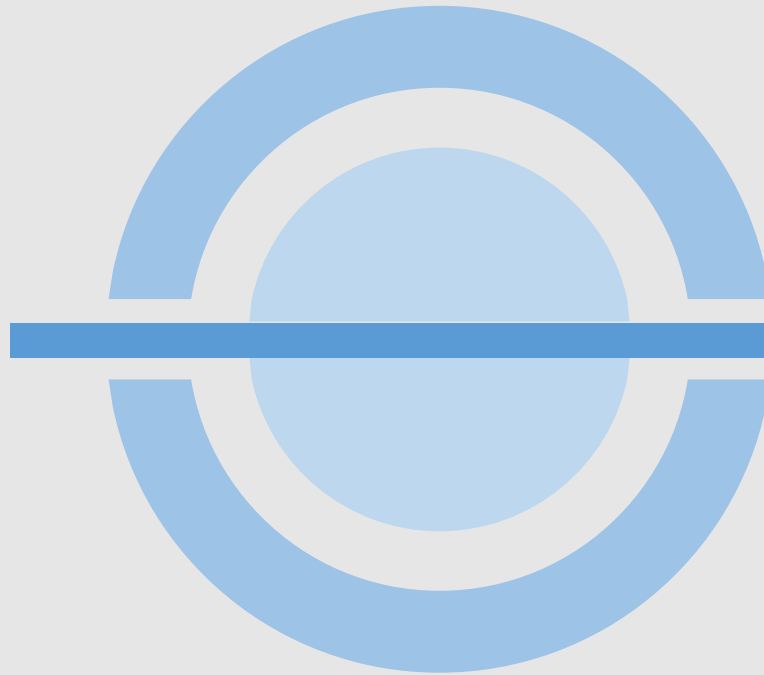


# Section2 Project – PUBG Finish Placement Prediction



AIB\_12th\_이우림

데이터 소개

데이터 전처리

모델 학습

모델 해석



# 데이터 소개

0 1    PUBG게임의 매치안에서의 각 플레이어들을 나타낸 데이터(row: 446966, column: 29)

	Id	groupId	matchId	assists	boosts	damageDealt	DBNOs	headshotKills	heals	killPlace
0	7f96b2f878858a	4d4b580de459be	a10357fd1a4a91	0	0	0.00	0	0	0	60
1	eef90569b9d03c	684d5656442f9e	aeb375fc57110c	0	0	91.47	0	0	0	57
2	1eaf90ac73de72	6a4a42c3245a74	110163d8bb94ae	1	0	68.00	0	0	0	47
3	4616d365dd2853	a930a9c79cd721	f1f1f4ef412d7e	0	0	32.90	0	0	0	75
4	315c96c26c9aac	de04010b3458dd	6dc8ff871e21e6	0	0	100.00	0	0	0	45

•  
•  
•

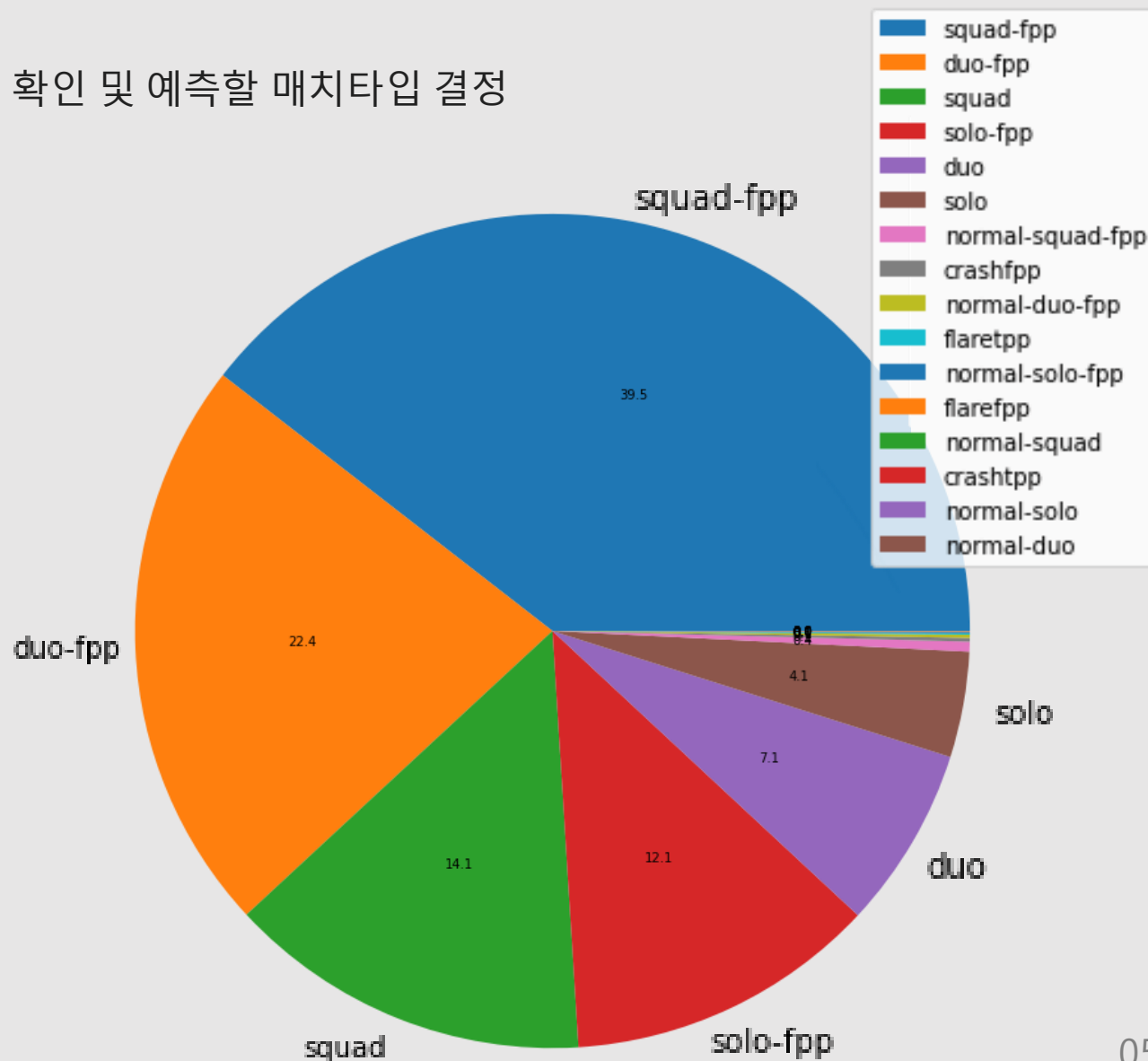
revives	rideDistance	roadKills	swimDistance	teamKills	vehicleDestroys	walkDistance	weaponsAcquired	winPoints	winPlacePerc
0	0.0000	0	0.00	0	0	244.80	1	1466	0.4444
0	0.0045	0	11.04	0	0	1434.00	5	0	0.6400
0	0.0000	0	0.00	0	0	161.80	2	0	0.7755
0	0.0000	0	0.00	0	0	202.70	3	0	0.1667
0	0.0000	0	0.00	0	0	49.75	2	0	0.1875

## 02 Column 소개

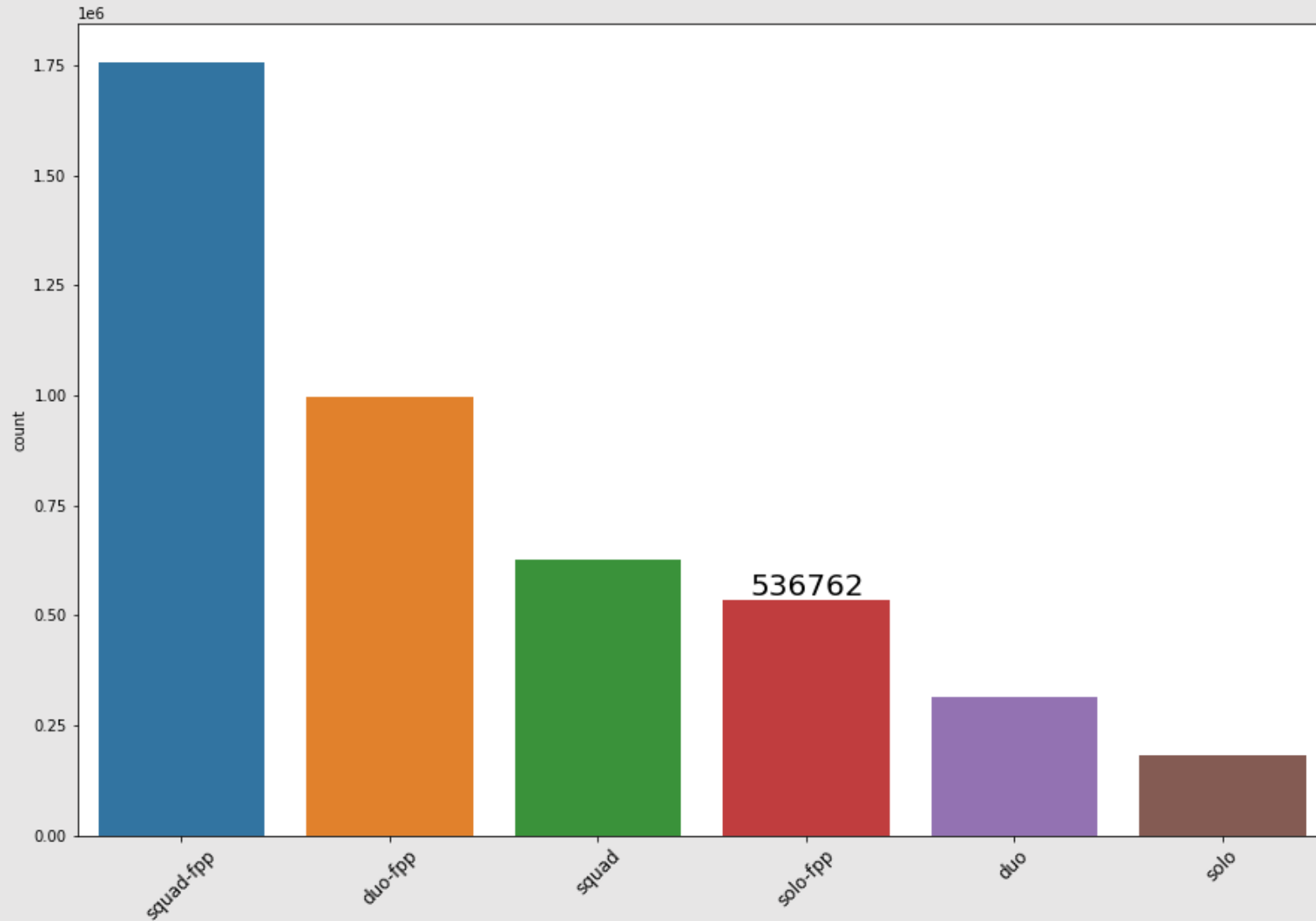
- **boosts**: 매치 내에서 사용한 부스팅 아이템의 수
- **damageDealt**: 팀킬 제외한 총 피해량
- **heals**: 매치 내에서 사용한 치유 아이템의 수
- **killPlace**: 이 매치 안에서의 킬 순위
- **kills**: 매치 내에서 적 플레이어를 죽인 횟수
- **headshotKills**: 매치 내에서 헤드샷으로 적 플레이어를 죽인 횟수
- **matchDuration**: 매치가 진행된 시간
- **matchType**: 게임모드(solo, duo, squad, solo-fpp, duo-fpp, squad-fpp 등)
- **rideDistance**: 차량으로 이동한 총 이동 거리(m 단위)
- **swimDistance**: 수영으로 이동한 총 이동 거리 (m 단위)
- **walkDistance**: 걸어서 이동한 총 이동 거리 (m 단위)
- **maxPlace**: 매치 내 인원
- **winPlacePerc**: 예측의 타겟으로 현재는 순위 백분위(1등이면 1, 꼴찌는 0)

## 01 매치 타입 별 데이터 수 확인 및 예측할 매치타입 결정

	count
squad-fpp	1756186
duo-fpp	996691
squad	626526
solo-fpp	536761
duo	313591
solo	181943
normal-squad-fpp	17174
crashfpp	6287
normal-duo-fpp	5489
flaretp	2505
normal-solo-fpp	1682
flarefpp	718
normal-squad	516
crashtp	371
normal-solo	326
normal-duo	199



## 01 매치 타입 별 데이터 수 확인 및 예측할 매치타입 결정

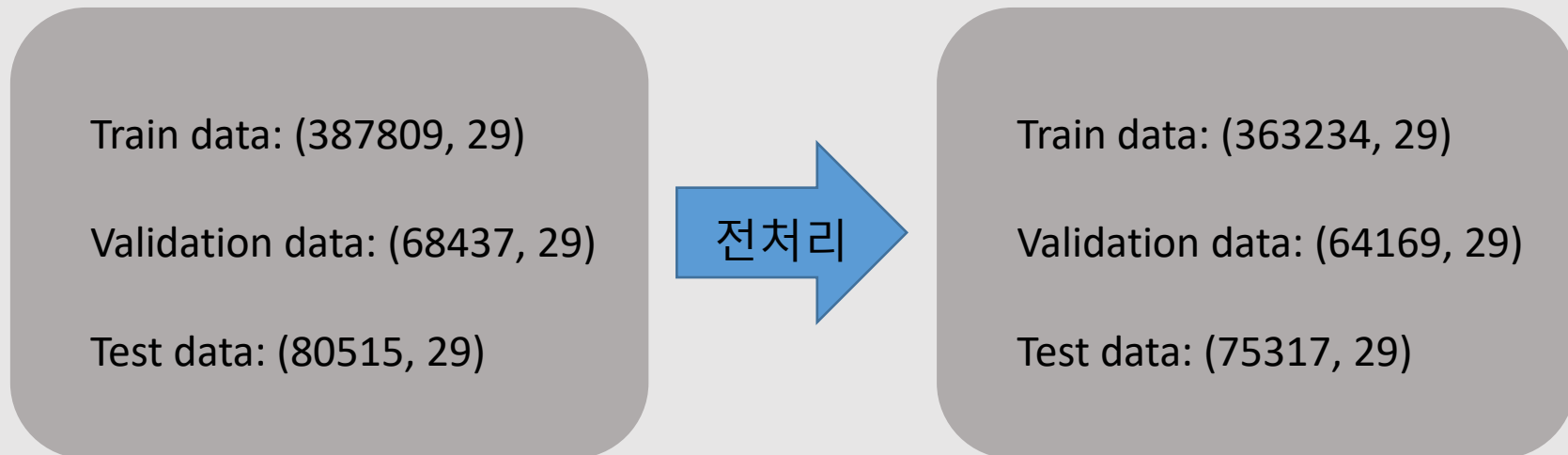


## 02 EDA

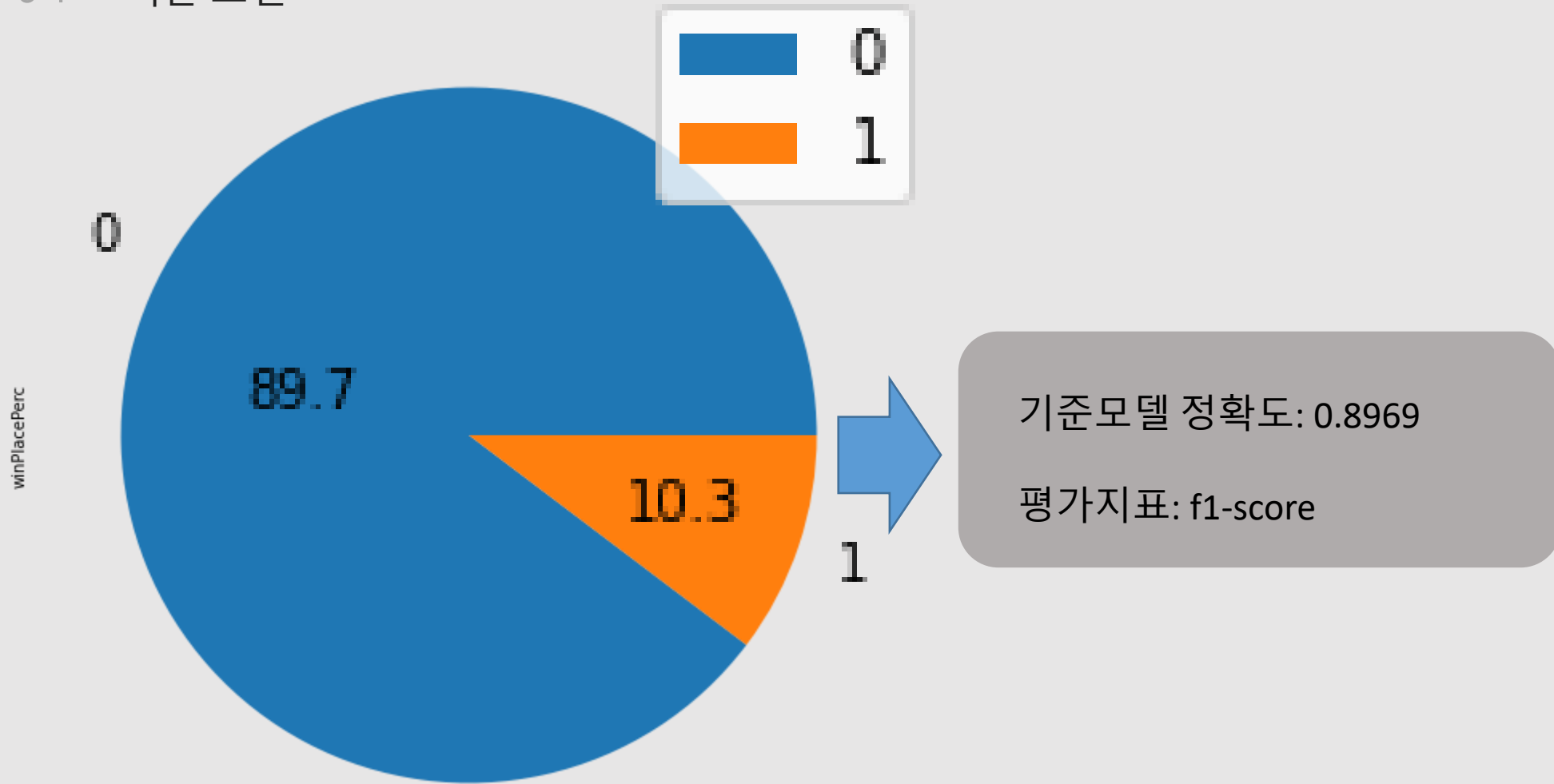
- 01 결측치, 중복값 확인 및 제거
- 02 solo 매치에서 무의미한 컬럼 삭제
- 03 총 이동거리 column 추가
  - **totalDistance**: rideDistance + swimDistance + walkDistance
- 04 킬중 헤드샷 킬의 비율을 나타낸 column 추가
  - **headshotKillsPerKills**: headshotKills/kills
- 05 차 사용 여부를 나타낸 column 추가
  - **carUse**: rideDistance가 0이면 0, 아니면 1
- 06 maxPlace가 90이하인 row 삭제
- 07 winPlacePerc가 0.9보다 작으면 0, 0.9보다 크면 1



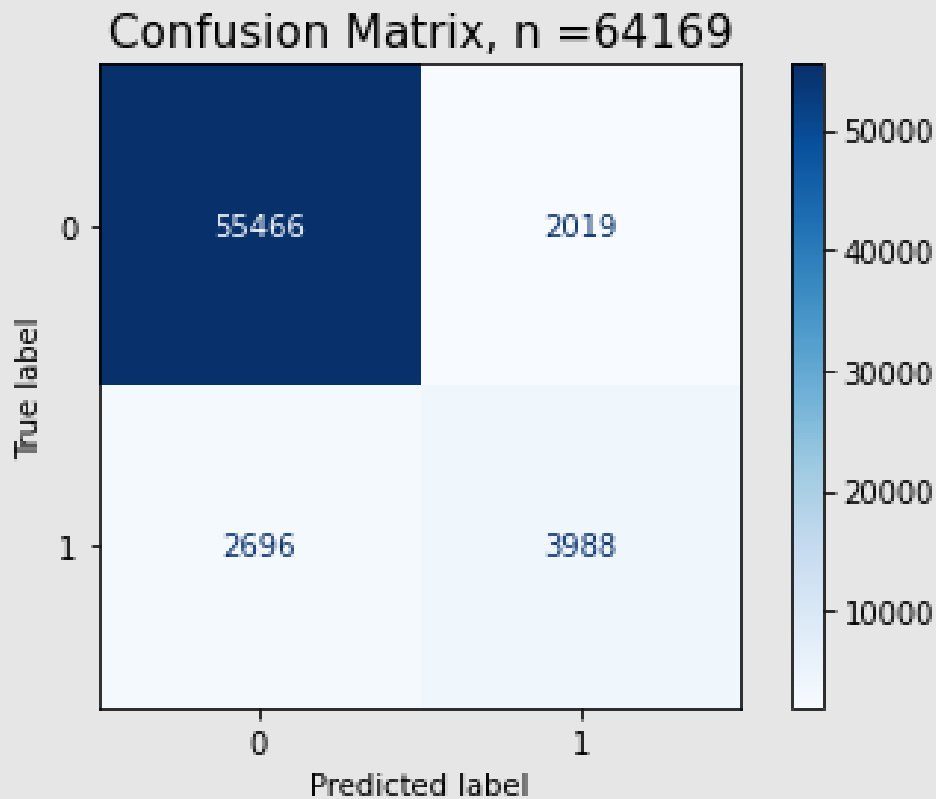
## 03 train data / validation data/ test data



## 01 기준 모델

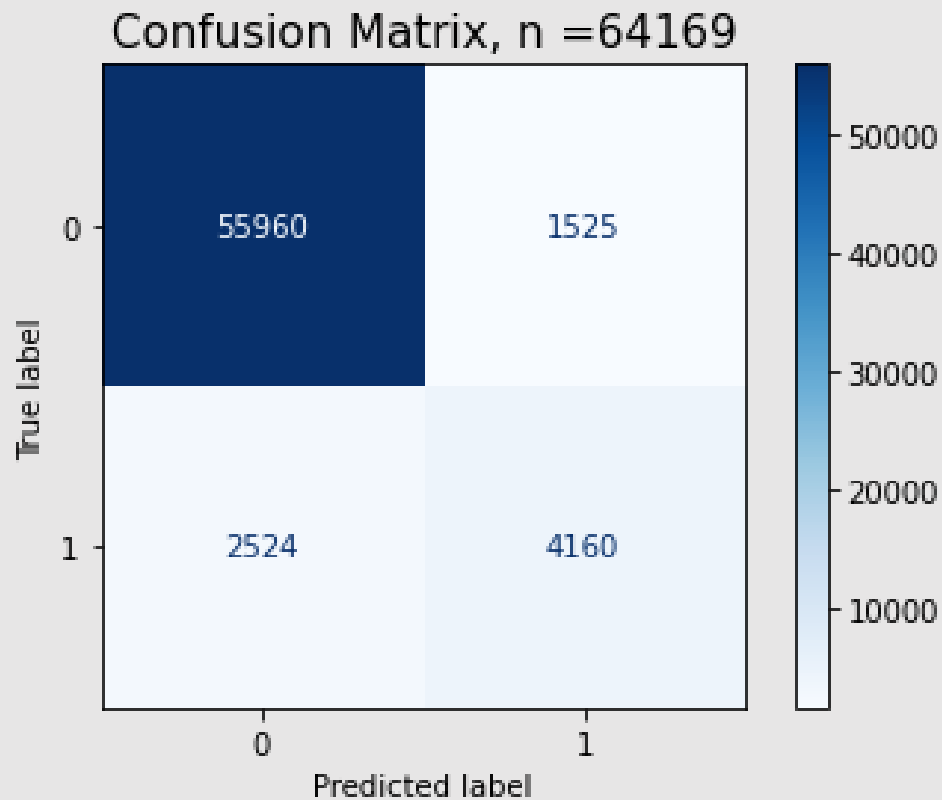


## 02 Decision Tree



검증데이터 f1-score: 0.6285  
검증데이터 recall: 0.5964  
검증데이터 accuracy: 0.9265

## 03 Random Forest

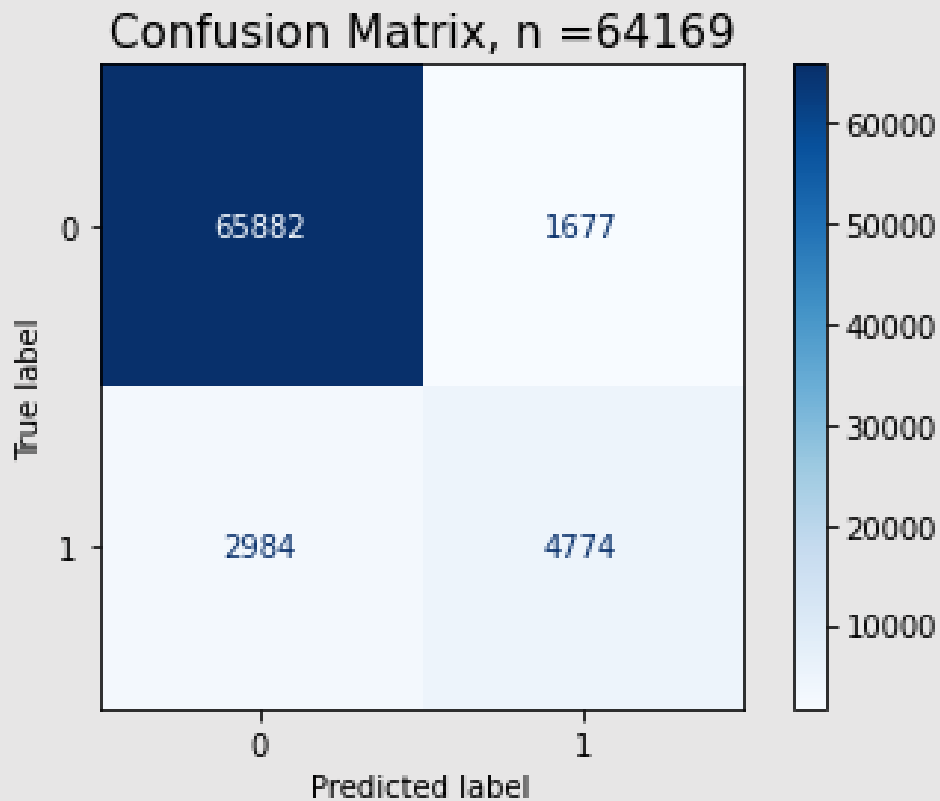


검증데이터 f1-score: 0.6726

검증데이터 recall: 0.6224

검증데이터 accuracy : 0.9369

## 04 모델 선택 및 테스트



테스트 데이터 f1-score: 0.672

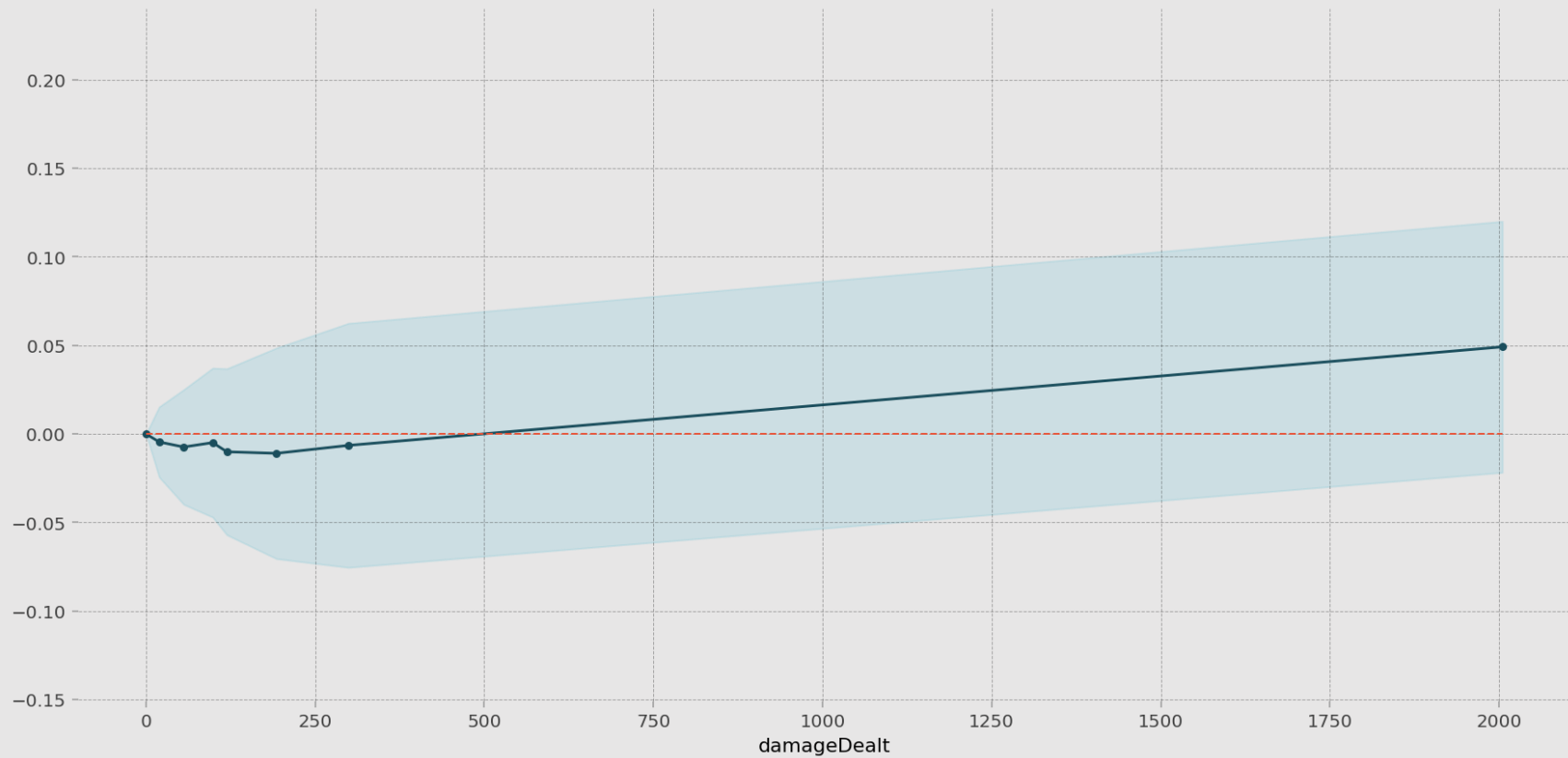
테스트 데이터 recall: 0.6153

테스트 데이터 accuracy : 0.9381

## 01 총 피해량과 target값과의 관계

PDP for feature "damageDealt"

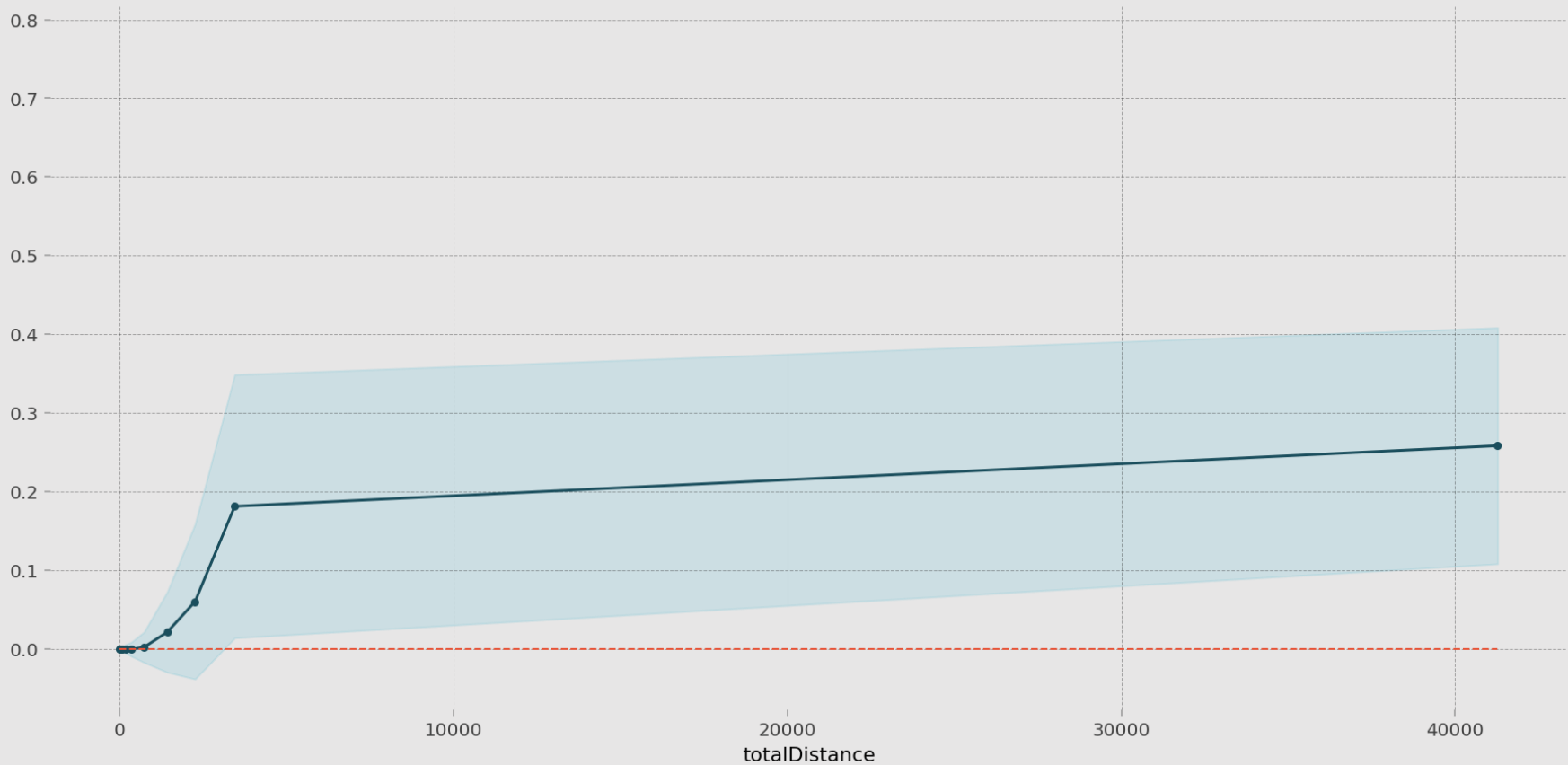
Number of unique grid points: 8



## 02 총 이동거리 target값과의 관계

PDP for feature "totalDistance"

Number of unique grid points: 10



Thank you!