

「제6회 대구 빅데이터 분석 경진대회」

분석 결과 보고서

접수번호 ※ 작성하지 않음

성명(팀명)	SPADA
분석과제명	대구로 어플 이상거래탐지 모델을 통한 사용자 불만 사전예측

I. 분석개요

□ 분석목적

본 프로젝트는 대구로 앱에서 발생하는 리뷰, 주문, 결제 데이터를 분석하여 부정한 거래 및 허위 리뷰를 사전에 탐지하는 것입니다. 이를 통해 허위 리뷰로 인한 소비자 피해를 방지하고, 앱 내 거래 및 리뷰 시스템의 공정성을 유지하고자 합니다. 머신러닝 및 딥러닝 기법을 활용하여 이상 데이터를 조기에 탐지함으로써 불법적인 리뷰 조작이나 악성 거래를 미연에 방지할 수 있습니다.

□ 배경 및 필요성

현대 소비자들은 앱을 통해 가게나 메뉴를 선택할 때, 리뷰와 평점을 중요한 참고 자료로 활용합니다. 긍정적인 리뷰가 많으면 매출이 증가하고, 부정적인 리뷰는 매출에 타격을 줄 수 있습니다. 그러나 일부 가게는 리뷰 대행 업체나 다수의 계정을 통해 허위 리뷰를 작성함으로써 소비자를 속이는 경우가 많습니다. 이는 소비자가 리뷰를 신뢰하고 구매를 결정하게 만들지만, 실제 경험이 리뷰와 일치하지 않으면 사용자 불만이 발생하고, 이는 앱에 대한 부정적인 평가로 이어지며, 최종적으로 사용자 이탈을 초래할 수 있습니다.

또한, 특정 경쟁 상점을 악의적으로 평가절하 하는 부정 리뷰는 정직한 가게에 피해를 입힐 수 있습니다. 이러한 문제를 해결하기 위해 머신러닝 및 딥러닝 기반의 이상 탐지 모델을 활용한 데이터 분석이 필수적입니다. 타 플랫폼에서는 이미 딥러닝을 기반으로 리뷰의 신뢰성을 제고하기 위한 방안을 도입하고 있으며, 이를 대구로 앱에서도 적용할 필요가 있습니다.

□ 분석요약

◦ 활용데이터

- 대구로 리뷰, 주문, 결제 데이터(대구 빅데이터 활용센터 제공): 가맹점 ID, 고객 ID, 리뷰 작성 시간, 결제 금액, 주문 상태 등 다양한 변수 포함

◦ 분석도구

- **Python** 및 **Google Colab**: 데이터 전처리, 모델링, 시각화 등 분석 전반에 사용

◦ 분석기법

이상 거래 및 허위 리뷰 탐지를 위해 여러 모델을 앙상블 방식으로 결합하여 분석을 수행하였습니다.

- **Isolation Forest**: F1 데이터를 분리하는 데 특화된 트리 기반 모델로, 점수가 최대화되는 임계값을 사용하여 허위 리뷰 탐지
- **Keras AutoEncoder**: 리뷰 데이터를 압축 및 복원하는 신경망, 리뷰시간과 유클리드 거리를 결합하여 허위 리뷰 탐지
- **Pytorch AutoEncoder**: 심층 신경망 기반의 오토인코더로, 세부적인 설정을 통해 리뷰 데이터의 재구성 오류를 기반으로 허위 리뷰 탐지
- **Random Forest**: 다수의 결정 트리를 이용한 앙상블 방식으로 허위 리뷰 탐지

◦ 분석결과

- 허위 리뷰로 의심되는 리뷰어들을 1차 필터링하는 데 성공. 이상 탐지 모델을 통해 리뷰 및 거래 데이터에서 비정상적인 패턴을 보이는 사용자를 효과적으로 탐지할 수 있었습니다.

□ 독창성 및 차별성

1. 리뷰 본문을 제외한 리뷰 이상 탐지

- 대구로 앱은 리뷰 본문을 매트로 방식으로 자동 생성할 수 있는 특성상, 텍스트 기반 탐지는 효과적이지 않습니다. 이에 따라 본 프로젝트에서는 리뷰 본문을 제외하고 가맹점 **ID**, 고객 **ID**, 리뷰 작성 시간, 리뷰 별점과 같은 메타데이터를 기반으로 허위 리뷰와 이상 거래를 탐지하였습니다.

2. 크롤링을 사용하지 않은 자사 데이터 활용

- 실무 적용 가능성을 고려하여 다른 데이터를 크롤링하지 않고, 대구로 앱 내부 데이터를 활용하여 분석을 진행하였습니다. 이를 통해 데이터베이스 침해와 같은 법적 문제를 방지하고, 실제 운영 환경에서의 활용성을 높였습니다.

3. 앙상블 기법 적용

- 다양한 이상 탐지 모델들이 각각의 장단점을 가지고 있어, 하나의 모델만으로는 탐지 사각지대가 발생할 수 있습니다. 본 프로젝트에서는 여러 모델을 결합하여 앙상블 방식으로 종합적인 분석을 진행하였으며, 투표 방식을 통해 각각의 모델이 도출한 결과를 종합적으로 판단함으로써 허위 리뷰 탐지의 정확성을 높였습니다.

4. 유클리드 거리 기반 이상 탐지

- 기존의 코사인 유사도는 주기적인 시계열 데이터를 탐지하는 데 유리하지만, 본 프로젝트에서는 **AutoEncoder** 모델에서 유클리드 거리를 사용하여 재구성 오류를 측정하였습니다. 이를 통해 허위 리뷰 탐지에서 더 높은 성능을 달성하였습니다.

II. 분석방법

▣ 활용데이터

데이터명	형식	대상기간	사용변수	출처	비고
리뷰 데이터	csv	2022년.1월 ~ 2023년.12월	SHOP_ID, CUST_ID, REVIEW_TIME, REVIEW_RANK	대구 빅데이터 활용 센터	https://dipbigdata.kr/ /
결제 데이터	csv	2022년.1월 ~ 2023년.12월	SHOP_ID, CUST_ID, PAY_GBN, TOT_AMT, DISC_AMT, DEST_PAY_GBN, CARD_PAY_GBN, CARD_APP_TIME, CARD_APP_AMT, PACK_DISC_AMT	대구 빅데이터 활용 센터	https://dipbigdata.kr/ /
주문 데이터	csv	2022년.1월 ~ 2023년.12월	SHOP_ID, ORDER_DT, CUST_ID, ORDER_ST, COMP_TIME, SHOP_CONF_TIME, CANCEL_DT, PACK_YN	대구 빅데이터 활용 센터	https://dipbigdata.kr/ /

▣ 분석 도구 및 알고리즘

◦ 분석 도구

- **Python:** 데이터 전처리, 모델 학습 및 성능 평가, 시각화 등 분석 전반에 걸쳐 사용
- **Google Colab:** 클라우드 환경에서 대규모 데이터를 처리하고 모델을 실험하기 위한 개발 환경으로 사용

◦ 데이터 전처리

1. 개인정보 암호화 처리

개인정보 보호를 위해 ‘가맹점 ID’ 및 ‘고객 ID’는 암호화 처리하였으며, 동일한 리뷰 작성자와 가맹점을 파악할 수 있도록 일관성 있는 암호화 처리 방식을 사용하였습니다. 이로 인해 고객과 가맹점 간의 리뷰 및 거래 패턴을 분석할 수 있습니다.

2. 데이터 병합

‘리뷰’, ‘결제’, ‘주문 데이터’를 통합하여 종합적인 분석을 수행하였습니다. **inner join**을 통해 세 개의 데이터셋을 연결하여 관련 있는 정보만을 결합하였으며, 이를 통해 각 고객의 리뷰, 결제, 주문 간의 상호작용을 분석하였습니다.

3. 이상 데이터 생성

허위 리뷰 탐지의 성능을 높이기 위해 특정한 이상 데이터 시나리오를 가정하고 가상 데이터를 추가 생성하였습니다.

- **상황 1:** 특정 리뷰 작성자가 특정 가맹점에만 별점 **5점**의 리뷰를 반복해서 작성하며, 다른 가맹점에는 리뷰를 작성하지 않는 경우.
- **상황 2:** 특정 리뷰 작성자가 특정 가맹점에만 별점 **1점**의 리뷰를 반복해서 작성하며, 다른 가맹점에는 리뷰를 작성하지 않는 경우.

◦ 분석 알고리즘

- **Isolation Forest:** 트리 기반 비지도 학습 알고리즘으로, 주어진 데이터를 분리하는 방식으로 이상치(허위 리뷰)를 탐지. **F1** 점수를 최대화할 수 있는 임계값을 설정하여 허위 리뷰로 의심되는 데이터 탐지.
- **Keras AutoEncoder:** 딥러닝 기반의 **AutoEncoder** 모델을 사용하여, 리뷰 데이터를 압축 및 복원하고 리뷰 작성 시간과 유클리드 거리를 결합하여 재구성 오류를 분석. 재구성 오류가 큰 리뷰를 허위 리뷰로 판단.
- **Pytorch AutoEncoder:** 심층 신경망 기반의 **AutoEncoder**로, 리뷰 데이터를 학습한 후 재구성 오류를 계산하여 허위 리뷰 탐지. 모델의 하이퍼파라미터를 세부적으로 조정하여 최적의 성능을 달성.
- **Random Forest:** 다수의 결정 트리를 결합하여 가맹점과 고객 간의 거래 및 리뷰 패턴을 분석하고, 허위 리뷰 여부를 분류하는 데 사용. 랜덤 포레스트는 각 트리의 예측을 종합하여 허위 리뷰 탐지 성능을 강화.

◦ 분석 결과 및 검증 방법

◦ 학습 데이터 성능 평가

- **Accuracy:** 모델이 정상 거래와 허위 리뷰를 얼마나 정확하게 분류하는지 평가.
- **Recall:** 실제 이상 거래 중에서 모델이 성공적으로 탐지한 비율.
- **F1 Score:** 정밀도와 재현도를 균형 있게 반영한 지표로, 허위 리뷰 탐지의 종합적인 성능을 평가.

◦ 테스트 데이터 검증 방법

- **상위 리뷰어 분석:** 대구 전체 데이터에서 리뷰 작성 횟수가 많은 상위 **30명**을 추출한 후, 그들의 결제 및 주문 데이터를 기반으로 허위 리뷰 여부를 분석.
- **정성적 평가:** 상위 리뷰어들의 결제 및 주문 데이터를 검토하여, 비정상적인 거래 패턴이나 의심스러운 리뷰가 존재하는지 정성적으로 평가.

▣ 시각화 방법

◦ 학습 데이터

- **Scatter Plot:** 리뷰어별 데이터 분포 및 허위 리뷰 여부를 시각적으로 확인하기 위한 산점도. 리뷰 작성 시간, 결제 금액, 리뷰 별점 등 다양한 변수에 따른 분포를 시각화하여 이상 패턴을 탐지.
- **Confusion Matrix:** 학습 데이터 내에서 모델이 정상 데이터를 정상으로 이상데이터를 이상 데이터로 몇개를 정확히 맞췄는지 확인하는 행렬로, 실제 정확성을 한눈에 볼수 있도록 함.
- 성능 평가 지표 표: 학습 데이터에서의 **Accuracy, Recall, F1 Score**를 표로 요약하여 각 모델의 성능을 비교하고 분석 결과를 쉽게 확인할 수 있도록 함.

◦ 테스트 데이터

- **Scatter Plot:** 상위 리뷰어들의 결제 및 주문 데이터를 시각적으로 분석하여 이상 거래 여부를 확인. 데이터 간의 관계를 직관적으로 확인할 수 있도록 다양한 변수 조합에 대한 산점도 시각화를 제공.

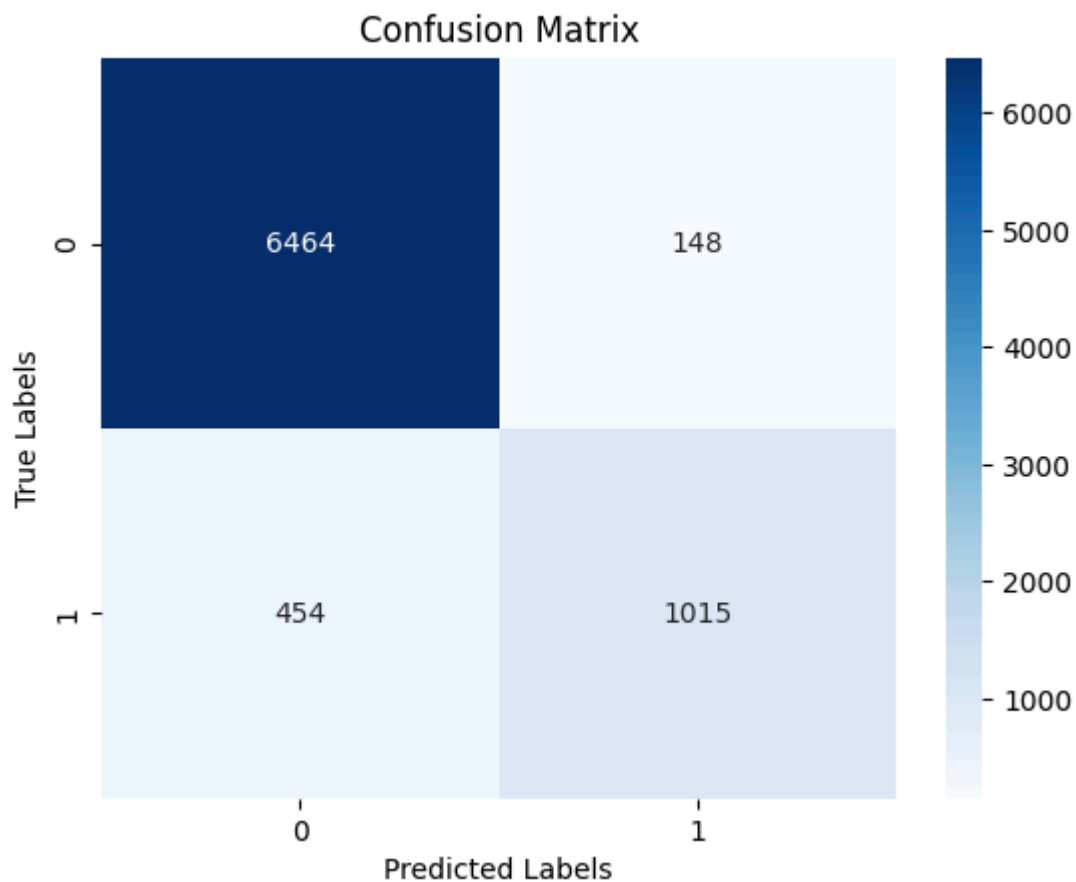
Ⅲ. 분석결과

학습데이터 시각화

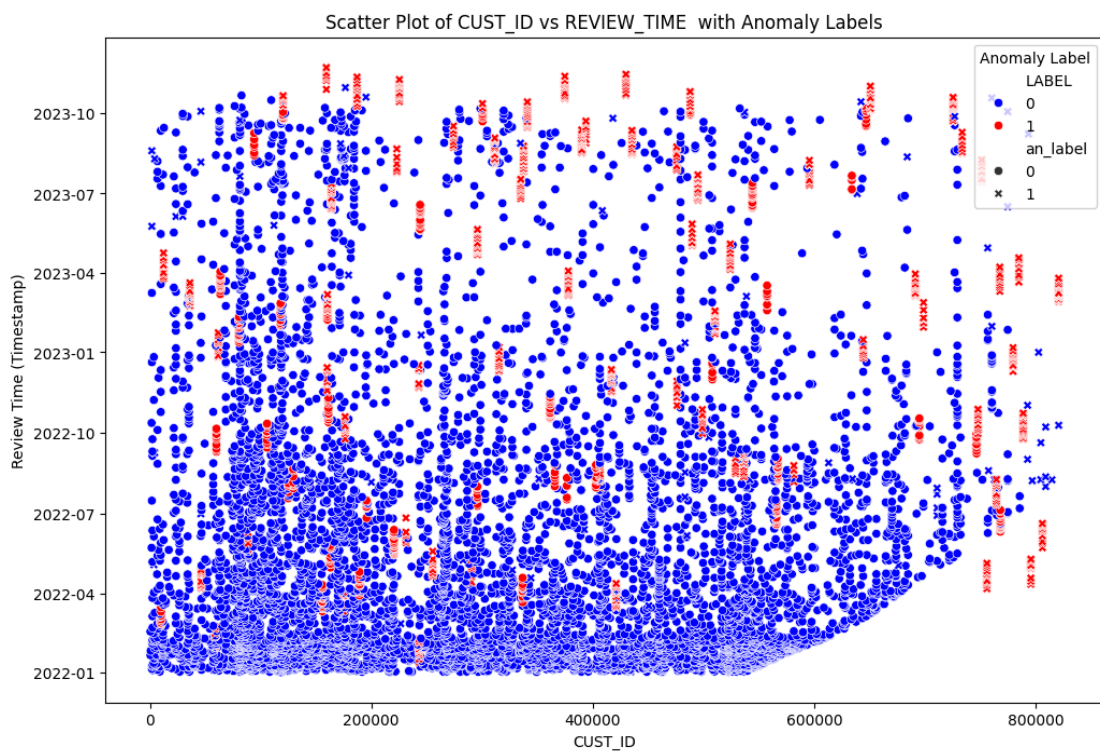
	Isolation Forest	Random Forest	Keras AutoEncoder	Pytorch AutoEncoder	Ensemble
Accuracy	86.29%	94.43%	91.26%	81.70%	92.55%
Recall	75.56%	73.16%	68.41%	50.24%	69.09%
F1 Score	66.71%	83.58%	74.01%	49.95%	77.13%

각각의 모델 성능과 최종 모델(Ensemble) 성능:

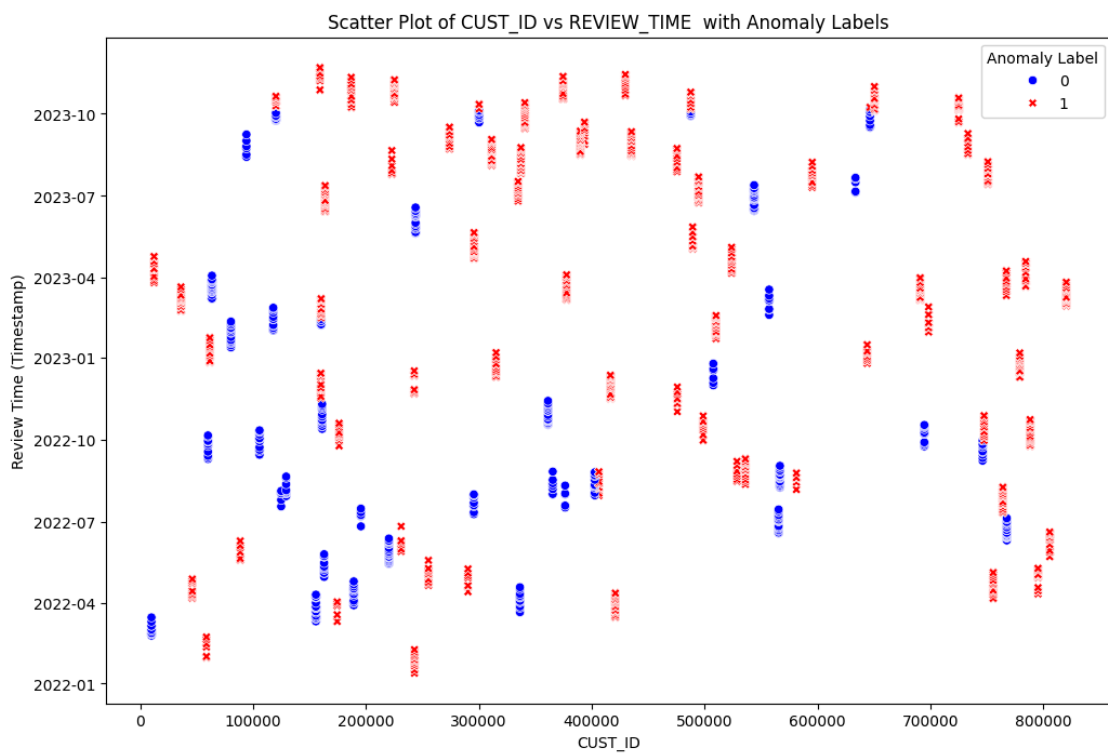
앙상블 모델이 각각의 모델의 단점을 보완하며 준수한 성능을 가짐



혼동행렬: 정상과 이상 데이터를 높은 확률로 분류하고 있음을
확인할수 있음

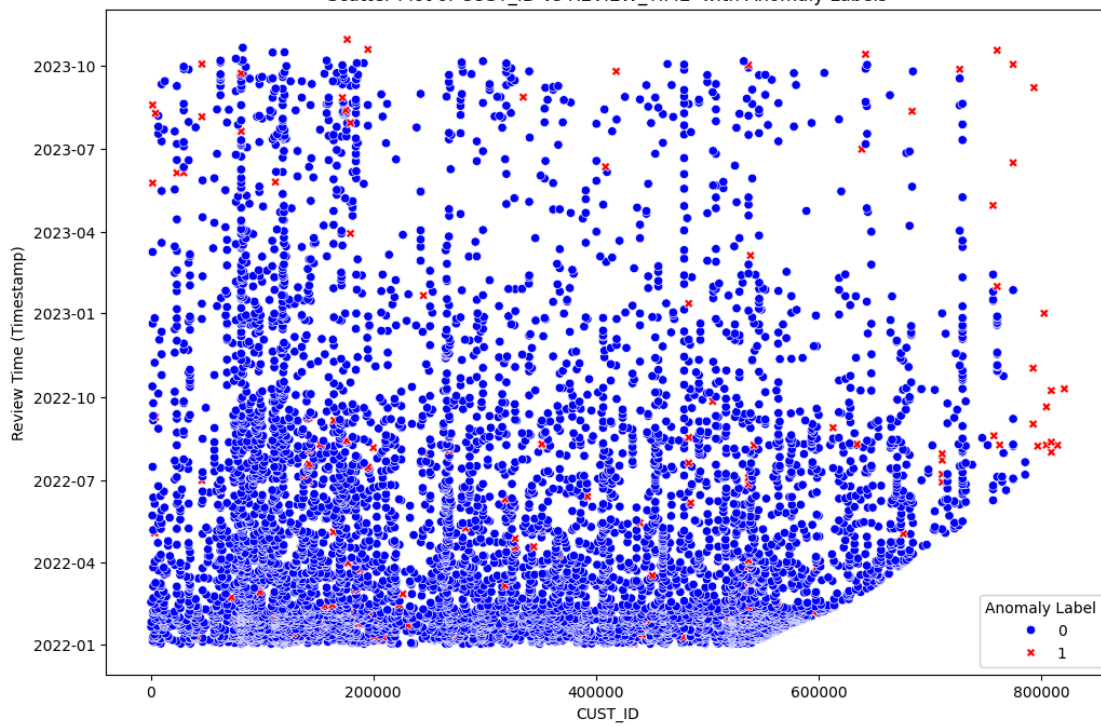


종합 데이터 확인



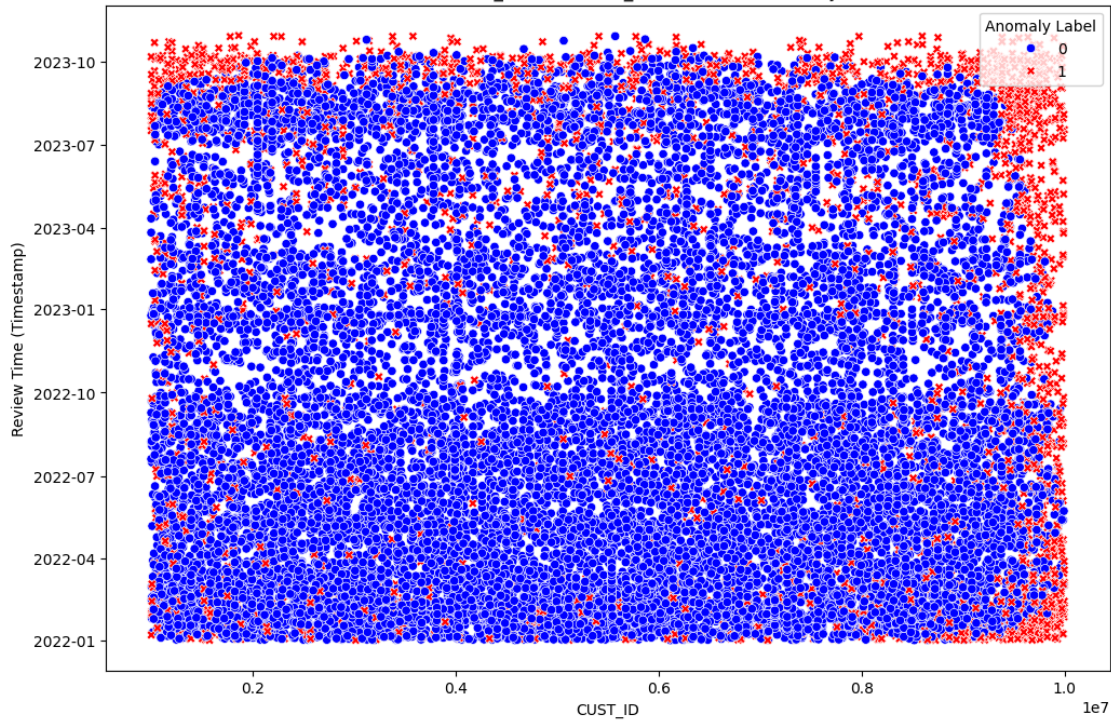
이상 데이터

Scatter Plot of CUST_ID vs REVIEW_TIME with Anomaly Labels



정상데이터

Scatter Plot of CUST_ID vs REVIEW_TIME without Anomaly Labels



라벨없는 대구시 전체 대구로 리뷰 데이터:

이상치들이 전체적으로 높은 CUST_ID에서 많이 발견됨

※ 검증 수행했을시 결과

케이스: CUST_ID = 9485582



1 con_df.loc[con_df["CUST_ID"]== 9485582]



SHOP_ID

CUST_ID

REVIEW_TIME

REVIEW_RANK

11709

9380485

9485582

2022-01-03 11:56:00

5.0

11710

9380485

9485582

2022-02-14 08:58:20

5.0

11711

9380485

9485582

2022-05-06 11:23:29

5.0

11712

9380485

9485582

2022-06-01 10:49:39

5.0

11713

9380485

9485582

2022-08-27 12:11:08

4.0

검증-리뷰데이터: 하나의음식점에 높은 리뷰를 남김

1 order_df.loc[order_df["CUST_ID"]== 9485582]

SHOP_ID

ORDER_DT

CUST_ID

ORDER_ST

COMP_TIME

SHOP_CONF_TIME

CANCEL_DT

PACK_YN

13255	9380485	2022-01-16 10:01:42	9485582	취소	NaN	2022-01-16 10:01:42	2022-01-16 10:01:03	N
13256	9380485	2022-01-16 10:01:29	9485582	취소	NaN	2022-01-16 10:01:56	2022-01-16 11:01:43	N
13257	9380485	2022-01-03 09:01:40	9485582	완료	2022-01-03 09:01:02	2022-01-03 09:01:06	NaN	N
18406	9602334	2022-02-01 06:02:01	9485582	완료	2022-02-01 09:02:03	2022-02-01 06:02:22	NaN	N
85627	8745119	2022-11-18 11:11:04	9485582	취소	NaN	2022-11-18 11:11:04	2022-11-18 11:11:09	N
226172	5315396	2022-01-14 06:01:11	9485582	완료	2022-01-14 09:01:25	2022-01-14 06:01:53	NaN	N
226642	2315489	2022-12-27 05:12:51	9485582	완료	2022-12-27 05:12:34	2022-12-27 05:12:26	NaN	N
226643	2315489	2023-09-22 05:09:08	9485582	완료	2023-09-22 05:09:15	2023-09-22 05:09:53	NaN	N
244103	1005172	2022-01-04 04:01:31	9485582	완료	2022-01-04 05:01:09	2022-01-04 04:01:57	NaN	N
244104	1005172	2022-01-16 06:01:26	9485582	완료	2022-01-16 07:01:57	2022-01-16 06:01:48	NaN	N
244105	1005172	2022-08-12 05:08:09	9485582	완료	2022-08-12 08:08:16	2022-08-12 05:08:56	NaN	N
244106	1005172	2022-03-28 12:03:40	9485582	완료	2022-03-28 03:03:09	2022-03-28 12:03:35	NaN	Y

검증-주문데이터:주문 취소가 잦고,몇몇 음식점에 반복적인주문



1 payment_df.loc[payment_df["CUST_ID"]== 9485582]



SHOP_ID

CUST_ID

PAY_GBN

TOT_AMT

DISC_AMT

DEST_PAY_GBN

CARD_PAY_GBN

CARD_APP_TIME

CARD_APP_AMT

PACK_DISC_AMT

16422	9380485	9485582	4.0	0.035070	0.000000	2.0	0.0	2022-02-12 08:34:06	0.035070	0.0
16423	9380485	9485582	4.0	0.035070	0.069926	2.0	0.0	2022-01-03 09:33:41	0.029409	0.0
16424	9380485	9485582	4.0	0.035070	0.069926	2.0	1.0	2022-01-16 10:39:43	0.029409	0.0
16425	9380485	9485582	4.0	0.035070	0.069926	2.0	1.0	2022-01-16 10:42:30	0.029409	0.0
16426	9380485	9485582	4.0	0.035070	0.022772	2.0	0.0	2022-06-01 08:38:06	0.033226	0.0
22361	9602334	9485582	4.0	0.038076	0.049505	2.0	0.0	2022-02-01 18:35:01	0.034068	0.0

검증-결제데이터:낮은 금액의 결제, 반복적

IV. 활용방안

□ 적용부문

- 실무에서 초기 필터링으로써 사용하여 운영의 효율성 증가
- 리뷰가 작성되는 대구로 택시, 전통시장 등에서 리뷰 건정성 확보
- 리뷰 탐지로 더욱 건전성이 강화된 착한 매장
- 대구 지하철, 대구 버스등 리뷰가 작성되는 공공 어플에서의 악의적 비난 탐지
- 정상 작성자들을 통한 음식 트렌드, 배달 집중도 확인

□ 정책 활용 방안

- 악의적 리뷰 작성자에 대한 법률적 문제 지원 정책
- 리뷰 인증제도 도입
- 사용자들의 리뷰 작성 동기부여를 위한 지원 정책

V. 기타

배경 및 동기:

- <https://news.mt.co.kr/mtview.php?no=2023103008582937721> : 허위리뷰 후속처리
- https://m.joseilbo.com/news/view.htm?newsid=451438#_enliple : 허위리뷰 탐지
- https://blog.naver.com/nhnccloud_official/223266309674 : NHN 이상탐지
- <https://dacon.io/competitions/official/235930/codeshare/5947> : 데이콘 이상탐지

데이터:

- <https://kmong.com/gig/430029> : 리뷰 작성 크몽
- <https://chatgpt.com/> : 이상데이터 생성 및 검수
- <https://claude.ai/new> : 이상데이터 생성 및 검수

모델링:

- <https://techblog.woowahan.com/11829/#toc-5> : 배달의 민족 모델링 방법