[모델]

1. 추천시스템기반 => 물류 예측 시스템
2. 시계열기반 딥러닝 => 일일 물류량 예측 시스템

=>일별 SHPR\_CD + ITEM\_CD = 이게 총 몇번 등장하는지로

=>2개

[제안서]

EDA => Feature 영향력을 검증 + 설명

[문제]

1. ITEM\_QTY가 비정상적으로 높은 데이터가 계속 발견됨 => 오류가 아니라 뭔가 다른 목적으로 분류된 제품 같음

=>ID: sibizi22 data만 따로 빼서 분석

1. ITEM\_AMT 가 영향력이 커보이는데 왜 이것이 영향력이 큰가? => 답을 찾아야함
2. Drop.na => 정해진 column만으로 써보자
3. Ecommerce => 특정 days에 물류가 많이 모일까?

[파트]

1. 머신러닝+Feature 해석 = 지웅,수빈
2. 딥러닝 모델 = 진명,상원,지웅

[task]

1. 지웅: sibizi22빼고 days+month 영향력 확인
2. 상원: column들로 days+month ohe해서 성능 확인 + 개선 확인
3. 진명: sibizi22 data빼서 item\_qty => lstm 모델..? => 만들긴 했는데 loss가 너무 크다!

=>바로 연락(120일분+) => 과적합 (120=>데이터를 오전/오후)

1. 수빈: (1) Heatmap : 히트맵 => column들 기준으로 히트맵을 그려보고 + 히트맵 그려서 => ‘내 생각엔 이런 이런 column들이 관련성이 있는거 같다!’ 라는걸 heatmap을 그려서 설명해보는 것 => heatmap으로 가설 or 가능하면 설명