**데이터 분석을 위한 주요 지표 및 메트릭 선정**

Aiden.ji(지윤수)

Luna.jung(정지연)

**1. 데이터 품질 관련 메트릭**

* **결측치 비율 (Missing Value Ratio)**: 데이터셋에서 결측치가 차지하는 비율을 확인합니다. 결측치는 데이터 분석 결과에 큰 영향을 미치므로, 이를 처리하기 위한 방법론을 함께 포함해야 합니다.
* **중복 데이터 비율 (Duplicate Data Ratio)**: 동일한 데이터가 중복 입력되었는지 확인하는 지표입니다. 중복된 데이터를 제거함으로써 데이터셋의 신뢰성을 높일 수 있습니다.
* **오류 데이터 비율 (Error Rate)**: 입력된 데이터에 잘못된 값이 포함되어 있는지 확인하는 지표입니다. 예를 들어, 잘못된 칼로리 값이나 논리적으로 일관성이 없는 데이터가 이에 해당할 수 있습니다.

**2. 사용자 입력 데이터 관련 메트릭**

* **일일 데이터 입력 빈도 (Daily Input Frequency)**: 사용자들이 얼마나 자주 데이터를 입력하는지 측정합니다. 이는 데이터 수집의 완전성과 지속성을 평가하는 데 유용한 지표입니다.
* **평균 입력 시간 (Average Input Time)**: 사용자가 데이터를 입력하는 데 걸리는 평균 시간을 측정하여, 사용성 향상 및 데이터 입력 과정을 최적화하는 데 도움을 줍니다.

**3. 음식 사진 분석 관련 메트릭**

* **이미지 분류 정확도 (Image Classification Accuracy)**: 음식 사진에서 음식 종류를 분류하는 AI 모델의 정확도를 평가하는 지표입니다. 머신러닝 모델을 사용하여 음식 사진을 분류할 때 얼마나 정확한지 측정합니다.
* **칼로리 예측 정확도 (Kcal Prediction Accuracy)**: AI가 사진을 분석해 예측한 칼로리 값이 실제 값과 얼마나 가까운지를 평가하는 지표입니다. 예측 모델의 성능을 평가하는 데 중요한 역할을 합니다.
* **처리 시간 (Processing Time)**: 음식 사진을 분석하고 결과를 도출하는 데 걸리는 시간을 측정합니다. 빠른 분석은 실시간 서비스에 중요한 요소입니다.

**4. Kcal 데이터 관련 메트릭**

* **평균 칼로리 섭취량 (Average Kcal Intake)**: 사용자가 입력한 Kcal 데이터를 기반으로 하루 또는 특정 기간 동안의 평균 섭취 칼로리를 측정
* **권장 칼로리 대비 초과/부족 비율 (Excess/Deficit Rate Compared to Recommended Kcal)**: 사용자가 하루 권장 섭취량보다 얼마나 초과하거나 부족한 칼로리를 섭취하는지 비율을 계산하는 지표
* **Kcal 데이터 일관성 (Kcal Data Consistency)**: 같은 음식을 여러 번 입력했을 때, Kcal 데이터가 일관되게 입력되었는지 확인하는 지표

**5. 모델 성능 관련 메트릭 (Azure에서 확인 후 입력)**

* **정확도 (Accuracy)**: 모델이 얼마나 정확하게 예측 및 분류를 했는지 평가하는 기본적인 성능 지표
* **정밀도 (Precision)**: 예측한 값 중 실제로 정답인 데이터의 비율을 측정하는 지표 음식 사진 분류나 칼로리 예측에서 잘못된 데이터를 줄이는 데 용이
* **재현율 (Recall)**: 실제 데이터 중에서 얼마나 많이 예측했는지를 평가하는 지표

**6. 사용자 행동 분석 메트릭**

* **유지율 (Retention Rate)**: 사용자가 지속적으로 데이터를 입력하는지, 사용을 중단했는지를 측정하는 지표입니다. 서비스의 성공 여부를 평가할 수 있는 중요한 요소입니다.
* **사용자 참여도 (User Engagement)**: 사용자가 얼마나 자주 앱이나 서비스에 접속하여 데이터를 입력하는지, 또는 음식을 분석하는지 평가하는 지표입니다.