**음식점 추천 시스템 보고서**

**1. 프로젝트 개요**

**프로젝트 이름**: 음식점 추천 시스템

**목표**: 네이버 플레이스에서 크롤링한 음식점 데이터를 기반으로 사용자의 위치, 취향, 평점 등을 분석하여 맞춤형 음식점을 추천하는 시스템을 개발합니다. 사용자의 음식 선호도, 과거 방문 기록, 현재 위치 등을 바탕으로 최적의 음식점을 추천하는 것이 목표입니다.

**2. 시스템 구조 및 데이터 수집**

**데이터 수집 방법**:

• **크롤링 대상**: 네이버 플레이스의 음식점 정보(위치, 카테고리, 리뷰, 평점 등)를 웹 크롤링을 통해 수집합니다.

• **크롤링 기술**: Python의 **BeautifulSoup**과 **Selenium**을 사용하여 웹 페이지에서 데이터를 수집합니다.

• **데이터 수집 범위**:

• **음식점 정보**: 이름, 위치(주소), 카테고리(예: 한식, 중식, 일식 등), 운영 시간, 전화번호.

• **리뷰 및 평점**: 사용자들이 남긴 리뷰와 평점.

• **음식점 특징**: 분위기, 인기 메뉴, 가격대 등 추가적인 속성.

**3. 데이터 셋 구축**

수집된 데이터는 정리 및 전처리 과정을 통해 분석에 적합한 형태로 변환됩니다.

1. **데이터 필드**:

• **음식점 이름**: 음식점의 고유한 이름.

• **카테고리**: 음식점의 유형(예: 한식, 양식, 중식 등).

• **위치(주소 및 좌표)**: 주소와 지도 상의 좌표 정보.

• **리뷰 및 평점**: 사용자들이 남긴 리뷰 텍스트와 평균 평점(1점~5점).

• **가격대**: 대략적인 메뉴의 가격 범위.

• **운영 시간**: 음식점의 영업 시간.

• **기타 특징**: 인기 메뉴, 분위기, 주차 여부 등.

2. **데이터 전처리**:

• **결측값 처리**: 일부 필드가 비어 있는 데이터를 채우거나 제거합니다.

• **텍스트 데이터 처리**: 리뷰 텍스트를 분석하여 긍정적/부정적 의견을 분류할 수 있도록 텍스트 정규화 및 토큰화 작업을 진행합니다.

• **좌표 변환**: 주소를 좌표 정보로 변환하여 지도 기반 추천에 활용할 수 있도록 합니다.

**4. 사용 알고리즘**

추천 시스템 구축에 적합한 다양한 알고리즘을 사용하여 사용자의 선호도에 맞춘 음식을 추천합니다.

1. **협업 필터링(Collaborative Filtering)**:

• **개념**: 사용자가 과거에 평점이나 리뷰를 남긴 음식점 데이터를 바탕으로 비슷한 취향을 가진 사용자들이 좋아할 만한 음식을 추천하는 방식입니다.

• **알고리즘**: **사용자 기반 협업 필터링**(User-Based Collaborative Filtering)을 사용하여 비슷한 음식을 선호하는 사용자들의 행동을 분석합니다. \*\*코사인 유사도(Cosine Similarity)\*\*를 통해 사용자 간의 유사성을 계산합니다.

• **장점**: 사용자 간 유사도를 바탕으로 추천할 수 있어, 정확도가 높습니다.

2. **콘텐츠 기반 필터링(Content-Based Filtering)**:

• **개념**: 음식점의 특성(예: 메뉴, 카테고리, 분위기)을 바탕으로 사용자가 선호할 가능성이 있는 음식을 추천하는 방식입니다.

• **알고리즘**:

• **TF-IDF**(Term Frequency-Inverse Document Frequency)를 사용하여 리뷰 및 음식점 설명 텍스트를 분석합니다.

• 음식점의 속성(가격대, 분위기 등)을 벡터화한 후 **코사인 유사도**로 유사성을 계산하여 비슷한 특성의 음식점을 추천합니다.

• **장점**: 사용자의 명시적인 선호도를 반영한 맞춤형 추천이 가능합니다.

3. **하이브리드 필터링(Hybrid Filtering)**:

• **개념**: 협업 필터링과 콘텐츠 기반 필터링을 결합하여, 두 방식의 장점을 모두 활용한 추천 방식입니다.

• **알고리즘**: 협업 필터링의 결과와 콘텐츠 기반 필터링의 결과를 혼합하거나 가중치를 부여하여 최종 추천을 제공합니다.

• **장점**: 추천의 다양성과 정확성을 동시에 높일 수 있습니다.

**5. 시스템 설계**

1. **데이터 전처리**:

• 크롤링한 데이터에서 불필요한 정보를 제거하고, 텍스트 데이터를 정규화하여 추천 모델에 적합한 형태로 변환합니다.

• 리뷰 텍스트를 바탕으로 음식점의 긍정/부정 평가를 분류하여 추천 알고리즘에 반영합니다.

2. **모델 학습 및 추천**:

• **협업 필터링**: 사용자 간의 상호작용(리뷰, 평점)을 바탕으로 유사한 사용자들의 선호를 분석하고, 이를 기반으로 음식점을 추천합니다.

• **콘텐츠 기반 필터링**: 음식점의 특징을 분석하여 유사한 속성을 가진 음식점을 추천합니다.

• **하이브리드 모델**: 두 가지 방식의 결과를 결합하여 사용자에게 최적의 음식점을 추천합니다.

3. **추천 방식**:

• **위치 기반 추천**: 사용자의 현재 위치를 기반으로 가까운 음식점을 우선 추천합니다.

• **실시간 리뷰 반영**: 최근 리뷰와 평점을 실시간으로 반영하여 트렌디한 음식점을 추천합니다.

• **카테고리 필터링**: 사용자가 선호하는 음식 종류(한식, 양식 등)에 따라 추천 결과를 필터링합니다.

**6. 성능 평가**

1. **평가 지표**:

• **RMSE(평균 제곱근 오차)**: 추천된 음식점의 평점 예측 정확도를 측정합니다.

• **Precision(정밀도)**: 추천된 음식점 중 실제로 사용자가 선호한 비율을 측정합니다.

• **Recall(재현율)**: 실제로 사용자가 선호한 음식점 중 추천된 음식점의 비율을 측정합니다.

2. **사용자 피드백**:

• 사용자로부터의 피드백을 바탕으로 추천 알고리즘을 지속적으로 개선합니다.

• 추천된 음식점에 대한 평가를 통해 시스템의 성능을 향상시킵니다.

**7. 기대 효과**

• **맞춤형 음식점 추천 제공**: 사용자 개인의 취향과 위치를 기반으로 맞춤형 추천을 제공하여 음식점 선택의 어려움을 해결합니다.

• **사용자 만족도 향상**: 사용자가 선호하는 음식점을 더 쉽게 찾을 수 있도록 도와줘서 만족도를 높입니다.

• **반복 사용 유도**: 매번 새로운 음식점 추천을 통해 사용자가 계속해서 시스템을 사용할 수 있도록 유도합니다.

**8. 결론**

이 음식점 추천 시스템은 네이버 플레이스에서 크롤링한 데이터를 활용하여 사용자의 위치, 선호도, 리뷰 데이터를 분석하고, 협업 필터링과 콘텐츠 기반 필터링을 결합한 하이브리드 추천 알고리즘을 사용하여 맞춤형 음식점을 추천합니다. 이 시스템은 사용자의 편의성을 높이고, 만족스러운 음식점 선택 경험을 제공하는 데 기여할 수 있습니다.