

# 라이징 상권 예측 머신러닝 분석 결과

## 1. 프로젝트 개요

이 프로젝트는 서울시 상권 데이터를 활용하여 "다음에 뜰 상권"을 머신러닝으로 예측하는 분석입니다.

## 분석 목표

- 이미 성공한 상권(성수, 을지로 등)의 패턴을 학습
- 비슷한 패턴을 보이는 "라이징 후보" 상권 발굴
- 어떤 요소가 상권 성공에 중요한지 파악

## 분석 데이터

22 개 상권의 검색량, 매출, 유동인구 등 9 개 변수를 분석했습니다.

## 3. 사용한 모델과 성능

3 가지 머신러닝 모델을 비교했습니다:

모델명	정확도	특징
Decision Tree	95.5%	가장 높은 정확도!
Random Forest	90.9%	변수 중요도 분석에 활용
Logistic Regression	86.4%	확률 예측에 활용

**해석:** Decision Tree 모델이 22 개 상권 중 21 개를 정확히 분류했습니다. 이는 우리 데이터의 패턴이 명확하다는 의미입니다.

## 4. 어떤 변수가 중요할까? (Feature Importance)

"라이징 상권이 되는 데 어떤 요소가 가장 중요한가?"를 분석한 결과입니다.

순위	변수	중요도	의미
1	Model2_점수 (서울 전체 모델)	35.9%	가장 중요!
2	핫플_유사도_점수	19.0%	성공상권 유사도
3	blog_post (블로그 포스팅)	14.1%	온라인 관심도
4	주말_매출_비중	11.0%	핫플 특성
5	avg_naver (네이버 검색량)	10.3%	검색 관심도

**핵심 인사이트:** "서울 전체 모델 점수"와 "성공 상권 유사도"가 전체의 55%를 차지합니다. 즉, 기존 성공 상권과 비슷한 특성을 가진 곳이 다음 라이징 상권이 될 가능성이 높습니다.

## 5. 라이징 확률 예측 결과

각 상권이 "라이징 상권"이 될 확률을 0~100%로 예측했습니다.

### 성공 상권 (이미 라이징함)

상권	라이징 확률	상태
가산동	99.5%	이미 성공
서교동 (홍대)	98.9%	이미 성공
압구정동	85.4%	이미 성공
여의동	76.7%	이미 성공

### 라이징 후보 Top 5 (주목!)

상권	라이징 확률	추천 이유
삼성 1동	42.5%	성공 클러스터

서초 2 동	42.0%	성공 클러스터
청담동	18.5%	MZ 비중 높음
구로 3 동	14.8%	MZ 비중 높음
상암동	12.1%	성공 클러스터

## 6. 클러스터링 분석

22 개 상권을 특성이 비슷한 3 개 그룹으로 분류했습니다.

**클러스터링이란?** 비슷한 특성을 가진 것들끼리 자동으로 묶어주는 기술입니다.

예: 과일을 색깔과 크기로 분류하면 "빨간+작은(체리)", "노란+큰(바나나)" 같은 그룹이 생깁니다.

### 분석 결과

성공 상권과 같은 클러스터에 속한 후보 상권:

- 삼성 1 동, 상암동, 구로 3 동, 효창동, 서초 2 동
- 이 상권들은 성공 상권과 유사한 특성을 보유

## 7. 결론 및 시사점

### 핵심 발견

- 서울 전체 모델 점수와 핫플 유사도가 라이징 예측에 가장 중요 (55%)
- 삼성 1 동, 서초 2 동이 다음 라이징 후보로 가장 유력
- 대치 4 동은 검색량 급증(CAGR 34.5%)으로 4 사분면 라이징 후보

### 한계점

- 22 개 상권으로 데이터 수가 적음
- 성공 상권 라벨링이 하이브리드 점수 상위 4 개 기준
- SNS, 트렌드 등 정성적 요인 미반영