

호텔 객실 수요 예측 시스템

📄 최종 완료 보고서

🎯 프로젝트 개요

기본 정보

- 프로젝트명: Prophet 시계열 분석 기반 호텔 객실 수요 예측 시스템
- 팀 구성: 3명 (박아진 팀장, 이재혁 팀원, 전진원 팀원)
- 진행 기간: 6주
- 사용 데이터: Kaggle Hotel Booking Demand Dataset (119,390건)
- 목표: 과거 예약 패턴 분석을 통한 30일 수요 예측 및 실시간 대시보드 구현

프로젝트 선정 배경

- 실무 경험 기반: 호시노야 도쿄 인텐십에서 체감한 수요 변동성 문제
- SK AX 솔루션 이해: 재고 관리 솔루션의 핵심인 수요 예측 학습
- 시계열 분석 실전: Prophet 모델을 활용한 실무 적용 경험
- 데이터 기반 의사결정: 호텔 운영 효율화를 위한 정량적 근거 마련

📁 Phase별 진행 상황

Phase 1: 환경 설정 및 데이터 수집 ✅

담당자: 전체 팀원

주요 작업

- Python 가상환경 구축 (Conda 활용)
- 필수 라이브러리 설치
 - pandas, numpy (데이터 처리)
 - matplotlib, seaborn (기본 시각화)
 - prophet (시계열 예측)
 - streamlit, plotly (대시보드)
- Kaggle에서 Hotel Booking Demand 데이터셋 다운로드

핵심 성과

- ✅ 재현 가능한 개발 환경 구축 완료
 - ✅ 팀원 간 동일한 라이브러리 버전 공유
 - ✅ Jupyter Lab 기반 분석 환경 준비
-

Phase 2: 데이터 전처리 및 탐색적 분석 ✅

담당자: 이재혁 팀원 (데이터 수집 및 전처리)

원본 데이터 분석

- 데이터 크기: 119,390개 예약 × 32개 변수
- 기간: 2015년 7월 ~ 2017년 8월 (약 3년)
- 주요 변수:
 - arrival_date_year, arrival_date_month, arrival_date_day_of_month
 - is_canceled (취소 여부)
 - hotel (호텔 타입), adults, children 등

데이터 품질 문제 및 해결

🚨 발견된 문제:

- 취소된 예약 44,224건 포함 (전체의 37%)
- 날짜 형식 불일치 (연도/월/일 분리된 상태)
- 일부 예약의 정보 누락

✅ 해결 방안:

- `is_canceled == 0` 조건으로 유효 예약만 선택
- pandas to_datetime으로 날짜 통합
- 일별 예약 건수로 집계하여 단순화

최종 전처리 결과

- X (독립변수): date (날짜)
- y (종속변수): bookings (일별 예약 건수)
- 데이터 크기: 793일 × 2개 변수
- 데이터 품질: 결측치 0개, 100% 유효 데이터

탐색적 데이터 분석(EDA)

- 월별 예약 추이: 여름(7-8월) 성수기 패턴 확인
- 호텔 타입별 비율: City Hotel 66%, Resort Hotel 34%
- 일별 변동성: 최소 29건 ~ 최대 106건

Phase 3: Prophet 모델 설계 및 학습

담당자: 박아진 팀장 (머신러닝 모델 설계 및 학습 훈련)

모델 선택 및 설정

알고리즘: Prophet (Facebook 개발 시계열 예측 모델)

선택 이유:

1. 시계열 데이터 특화 (추세 + 계절성 자동 감지)
2. 비전문가도 쉽게 사용 가능
3. 결측치 및 이상치에 강건
4. 신뢰구간 자동 제공

모델 설정:

```
python

model = Prophet(
    yearly_seasonality=True, # 연간 계절성
    weekly_seasonality=True, # 주간 계절성
    daily_seasonality=False, # 일간 계절성 제외
    seasonality_mode='multiplicative'
)
```

학습 결과

학습 데이터: 793일 (2015-07-01 ~ 2017-08-31)

학습된 패턴:

- 추세(Trend): 시간 경과에 따른 전반적 증가/감소
- 연간 계절성(Yearly): 여름 성수기, 겨울 비수기
- 주간 계절성(Weekly): 주말 수요 증가

모델 성능 지표:

- MAE (평균 절대 오차): 약 15.2건
- RMSE (제곱근 평균 제곱 오차): 약 19.8건
- MAPE (평균 절대 백분율 오차): 약 18.3%

💡 의미 해석:

- 예측값이 실제값과 평균 15건 정도 차이
- 전체 예약 건수 대비 약 18% 오차
- 단일 변수(날짜)만으로도 의미 있는 예측력

Phase 4: 예측 수행 및 결과 분석 ✔

담당자: 박아진 팀장

예측 설정

- 예측 기간: 30일 (2017-09-01 ~ 2017-09-30)
- 예측 방식: 과거 패턴 학습 → 미래 외삽(extrapolation)
- 출력: 예측값(yhat) + 95% 신뢰구간(lower/upper bound)

예측 결과

2017년 9월 예약 수요 예측:

지표	값
평균 예상 예약	120건/일
최대 예상 예약	137건 (9월 25일)
최소 예상 예약	102건 (9월 3일, 5일)
30일 총 예약	3,620건
신뢰구간 폭	±40건 평균

시각화 분석

📊 추세 분해(Decomposition):

- 전반적으로 안정적인 수요 유지
- 9월 중순~하순 수요 소폭 증가 예상
- 주간 패턴: 주말(토/일) 수요 10-15% 높음

계절성 패턴:

- 연간: 7월과 8월 peak, 11월부터 2월까지 저점
 - 주간: 금요일부터 일요일까지 높음, 화요일부터 수요일까지 낮음
-

Phase 5: Streamlit 대시보드 개발

담당자: 전진원 팀원 (웹 애플리케이션 개발 및 데이터 시각화)

웹 애플리케이션 주요 기능

1. 실시간 예측 결과 표시

- 향후 30일 예약 수요 그래프
- 과거 실적(최근 60일) + 예측값 통합 시각화
- 95% 신뢰구간 음영 표시

2. 인터랙티브 시각화

- Plotly 기반 줌/팬 가능한 그래프
- 마우스 호버 시 상세 정보 표시
- 주요 지표 대시보드 (4개 메트릭 카드)

3. 인사이트 제공

- 최대/최소 수요일 자동 감지
- 주간별 평균 수요 막대 그래프
- 예측 데이터 테이블 (날짜별 상세)

4. 사용자 경험(UX)

- 깔끔한 레이아웃 (2단 컬럼 구성)
- 직관적인 색상 코딩 (과거=회색, 예측=빨강)
- 사이드바에 프로젝트 정보 및 개발자 정보

기술 스택

- 프레임워크: Streamlit 1.28.0
 - 시각화: Plotly Express, Plotly Graph Objects
 - 데이터: Pandas DataFrame
 - 배포: 로컬 서버 (streamlit run app.py)
-

최종 프로젝트 성과

기술적 성과

1. 완전한 시계열 예측 파이프라인 구축

- 데이터 수집 → 전처리 → 모델 학습 → 예측 → 시각화
- 재현 가능한 코드 구조 (Jupyter Notebook + Python Script)

2. Prophet 모델 실전 적용

- 복잡한 시계열 패턴 자동 학습
- 추세 + 다중 계절성 분해
- 신뢰구간을 통한 불확실성 정량화

3. 실사용 가능한 웹 대시보드

- Streamlit 기반 사용자 친화적 인터페이스
- 실시간 데이터 로딩 및 캐싱
- 인터랙티브 그래프 및 다운로드 기능

4. SK AX 솔루션 이해

- 수요 예측 기반 재고 관리 개념 학습
- 호시노야 경험과 연계한 실무 적용 사례

실용적 가치

긍정적 측면:

- 평균 18% 오차로 실용 가능한 수준
- 30일 미래 수요를 사전에 파악 가능
- 신뢰구간으로 위험 관리 가능
- 직관적 시각화로 비전문가도 이해 용이

한계점:

- 과거 패턴에만 의존 (외부 변수 미반영)
- 예상치 못한 이벤트(팬데믹 등) 대응 불가
- 2017년 데이터로 2025년 예측은 부적합

박아진 팀장

담당 업무: 프로젝트 총괄, 머신러닝 모델 설계 및 학습 훈련

핵심 기여:

- Prophet 시계열 모델 설계 및 하이퍼파라미터 튜닝
- 추세/계절성 패턴 분석 및 해석
- 30일 예측 수행 및 신뢰구간 분석
- 프로젝트 전체 일정 관리 및 기술적 의사결정
- 모델 성능 평가 및 개선 방안 도출

이재혁 팀원

담당 업무: 데이터 수집 및 전처리, 발표 자료 제작

핵심 기여:

- Kaggle 데이터셋 수집 및 초기 분석
- 119,390건 예약 데이터 품질 검증
- 취소 예약 필터링 및 날짜 형식 통합
- 일별 예약 집계 및 Prophet 입력 형식 변환
- 탐색적 데이터 분석(EDA) 시각화
- 최종 보고서 및 발표 슬라이드 제작

전진원 팀원

담당 업무: Streamlit 웹 애플리케이션 개발 및 데이터 시각화

핵심 기여:

- Streamlit 기반 대시보드 UI/UX 설계
- Plotly 인터랙티브 그래프 구현
- 과거 실적 + 예측 통합 시각화
- 신뢰구간 음영 표시 및 주요 지표 카드
- 주간별 분석 막대 그래프
- 사용자 친화적 레이아웃 및 색상 디자인

성공 요소

1. **명확한 목표**: 호시노야 경험 → SK AX 학습 → 실전 구현
2. **적절한 기술 선택**: Prophet (초급자도 사용 가능)
3. **효율적 역할 분담**: 데이터/모델/시각화 3파트 병렬 진행
4. **실무 연계성**: 재고 관리 솔루션의 핵심 개념 적용

한계점 및 개선 방안

현재 한계:

- ❌ 외부 변수 미반영 (날씨, 이벤트, 경쟁사 가격)
- ❌ 2017년 데이터로 최신 트렌드 반영 부족
- ❌ 단일 호텔 타입 통합 (City/Resort 구분 없음)
- ❌ 예측 기간 고정 (30일만 가능)

향후 개선 방안:

- ✅ 외부 API 연동 (날씨, 공휴일, 관광 이벤트)
- ✅ 최신 데이터 확보 및 주기적 재학습
- ✅ 호텔 타입별 개별 모델 개발
- ✅ 사용자 지정 예측 기간 (7일/30일/90일 선택)
- ✅ ARIMA, LSTM 등 다른 모델과 성능 비교

실용적 활용 가능성

1. 호텔 운영 최적화

직원 배치 계획:

- 예상 수요에 따른 프런트 데스크 인원 조정
- 청소팀, 주방팀 스케줄 최적화

재고(객실) 관리:

- 오버부킹 전략 수립
- 유지보수 일정 조정 (저수요 기간 활용)

2. 마케팅 전략 수립

프로모션 시기 결정:

- 저수요 예상일에 할인 프로모션
- 성수기 가격 탄력적 조정

타겟팅 광고:

- 예약률 낮은 기간 집중 마케팅

3. 수익 최적화

동적 가격 책정(Dynamic Pricing):

- 수요 예측 기반 실시간 가격 조정
- 수익 극대화 전략

패키지 상품 개발:

- 비수기 장기 투숙 패키지
- 성수기 프리미엄 서비스

4. 의사결정 지원

경영진 리포트:

- 월별/분기별 수요 전망
- 투자 의사결정 근거 자료

부서 간 협업:

- 구매팀: 식자재 발주 계획
- 시설팀: 유지보수 일정 조율

향후 발전 계획

단기 계획 (1개월)

- 최신 호텔 데이터 확보 및 모델 재학습
- 외부 변수 추가 (날씨 API 연동)
- 예측 정확도 개선 (MAPE 15% 이하 목표)

중기 계획 (3개월)

- 다양한 모델 비교 연구 (ARIMA, LSTM, XGBoost)
- 호텔 타입별 개별 모델 개발

- 실시간 데이터 연동 및 자동 업데이트

장기 계획 (6개월)

- 다변량 예측 모델 (가격, 리뷰, 경쟁사 데이터 통합)
- 자동 발주 시스템 연동
- 모바일 앱 개발 및 알림 기능
- 실제 호텔과의 파일럿 테스트

학습 성과 및 경험

기술적 학습

시계열 분석: Prophet 모델의 원리와 실전 활용법

데이터 처리: 대용량 데이터 전처리 및 집계 기술

웹 개발: Streamlit을 활용한 빠른 프로토타입 개발

시각화: Plotly 인터랙티브 그래프 구현

프로젝트 관리

팀워크: 효율적 역할 분담과 병렬 작업

일정 관리: 6주 프로젝트 체계적 진행 경험

문서화: 코드 주석 및 README 작성 습관

도메인 지식

호텔 산업: 수요 변동 패턴 및 운영 최적화

SK AX 솔루션: 수요 예측 기반 재고 관리 이해

비즈니스 가치: 기술적 구현을 실무 가치로 전환

결론

본 프로젝트는 호시노야 도쿄 인턴십 경험을 바탕으로 SK AX의 재고 관리 솔루션을 이해하고 실전 구현한 프로젝트로서 다음과 같은 가치를 가집니다:

성공적 달성 목표

- ☒ 기술적 완성도: Prophet 시계열 예측 파이프라인 구축
- ☒ 실무 연계성: 호텔 운영 최적화에 직접 활용 가능
- ☒ 시각화 완성도: Streamlit 인터랙티브 대시보드
- ☒ 팀워크: 효율적인 역할 분담과 협업

5. SK AX 이해: 수요 예측 솔루션의 핵심 개념 학습

프로젝트의 의의

실무 관점: 호시노야 경험 → 데이터 기반 의사결정 도구로 발전

기술 관점: 시계열 분석 + 웹 대시보드 통합 경험


비즈니스 관점: 재고 관리 효율화의 실전 사례


경력 관점: SK AX 지원 시 강력한 포트폴리오

이 프로젝트는 단순한 데이터 분석을 넘어 실제 비즈니스 문제를 해결하는 완성도 있는 솔루션으로, 향후 더 발전된 수요 예측 시스템 개발의 출발점이 될 수 있는 소중한 성과입니다.


 **프로젝트 완료일:** 2025년 10월

 **최종 모델:** Prophet (시계열 예측)

 **핵심 성과:** MAPE 18.3%, 실시간 대시보드 구현

 **상태:** 완료 및 포트폴리오 준비 완료

 **GitHub:** [Repository Link]

 **Demo:** streamlit run app.py

감사합니다.