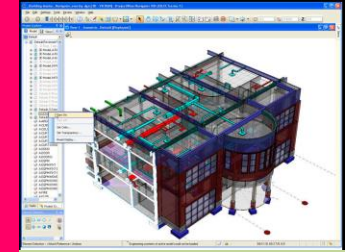


수요 예측 기법

문일경 교수

서울대학교 산업공학과

Construction

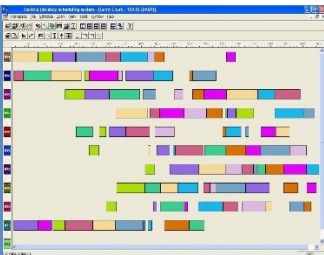


Operations
Research



Simulation

Scheduling



Analysis

Copyright © 2025 Supply Chain Management Lab
무단 복제 엄금



ENGINEERING
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

Contents

1. 수요 예측 개요
2. 전통적 통계 기반 예측 기법
3. ARIMA 계열 예측 기법
4. 머신러닝 예측 기법
5. 하이브리드 예측 기법

1. 수요 예측 개요

1.1. 수요 예측의 0원칙

1. 수요 예측 개요

2. 전통적 통계 기반 예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

“Forecasts are always wrong”
예측은 항상 틀릴 수 밖에 없다

1. 수요 예측 개요

1.2. 수요 예측의 1원칙

1. 수요 예측 개요

2. 전통적 통계 기반 예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

"Yet, forecasting is still essential"
그럼에도 예측은 필요하다

1. 수요 예측 개요

1.3. 수요 예측의 필요성

- 재고 최적화
 - 수요 예측 실패
 - 재고 과잉으로 인한 비용 증가
 - 재고 부족으로 인한 매출 손실
 - 예측 기반 재고 운영: 적정 재고 수준 유지, 서비스 수준 향상
- 생산 및 운영 계획 수립
 - 설비, 인력, 원자재 확보를 사전에 계획할 수 있음
 - 급박한 수급 상황 방지, 리드 타임 안정화
 - 특히 설비 전환이 까다로운 산업에서 필수적 (ex. 제조업)



1. 수요 예측 개요

1.3. 수요 예측의 필요성

- 고객 만족도 향상
 - 고객이 원하는 시점에 원하는 제품 제공 → 고객 기대 충족
 - 고객군별, 지역별, 채널별 선호 예측 → 개인화된 경험 제공
 - 브랜드 신뢰도 향상



1. 수요 예측 개요

1.4. 수요 예측 필수 요소

미국 Gartner사가 제시한 성숙한 수요 예측 체계의 5가지 핵심 능력

1. 협업 기반 예측 (Collaborative Forecasting)

- 여러 부서가 함께 예측에 참여

2. 다양한 계층에서의 예측 (Multi-level Forecasting)

- 제품, 지역, 고객, 유통채널 등에 따라 각기 다른 예측

3. 예측 정확도 측정 (Forecast Accuracy Measurement)

- 예측의 정확성을 지속적으로 측정하고 피드백 루프 운영

4. 이벤트 기반 수요 반영 (Event Planning Integration)

- 프로모션, 캠페인 등 특별 이벤트를 예측에 통합

5. 시나리오 기반 계획 (Scenario-Based Planning)

- 다양한 수요 상황에 대비하는 예측 시나리오 구성

The logo for Gartner REPORT, featuring the word "Gartner" in a large, white, sans-serif font with a registered trademark symbol, and the word "REPORT" in a smaller, blue, sans-serif font below it.

출처: https://www.gartner.com/en/articles/demand-planning?utm_source=chatgpt.com

1. 수요 예측 개요

1.5. 수요 예측 사례

사례1) Zara의 실시간 수요 예측 기반 공급망 전략

- 스페인 패션 기업 Zara
- 산업적 특성: 빠르게 변화하는 트렌드 속 단기 제품 수명
→ 예측의 어려움
- 예측 전략
 - 매장별 실시간 판매 데이터를 본사에 매일 보고
 - 현장 직원의 정성적 피드백을 분석 과정에 포함
 - 예측 결과 기반으로 디자인-생산-물류를 2주 내에 완결하는 초단기 공급망 구축



출처: <https://supplychain360.io/zaras-supply-chain-mastery-an-analysis-of-strategy-and-execution/>

1. 수요 예측 개요

1.5. 수요 예측 사례

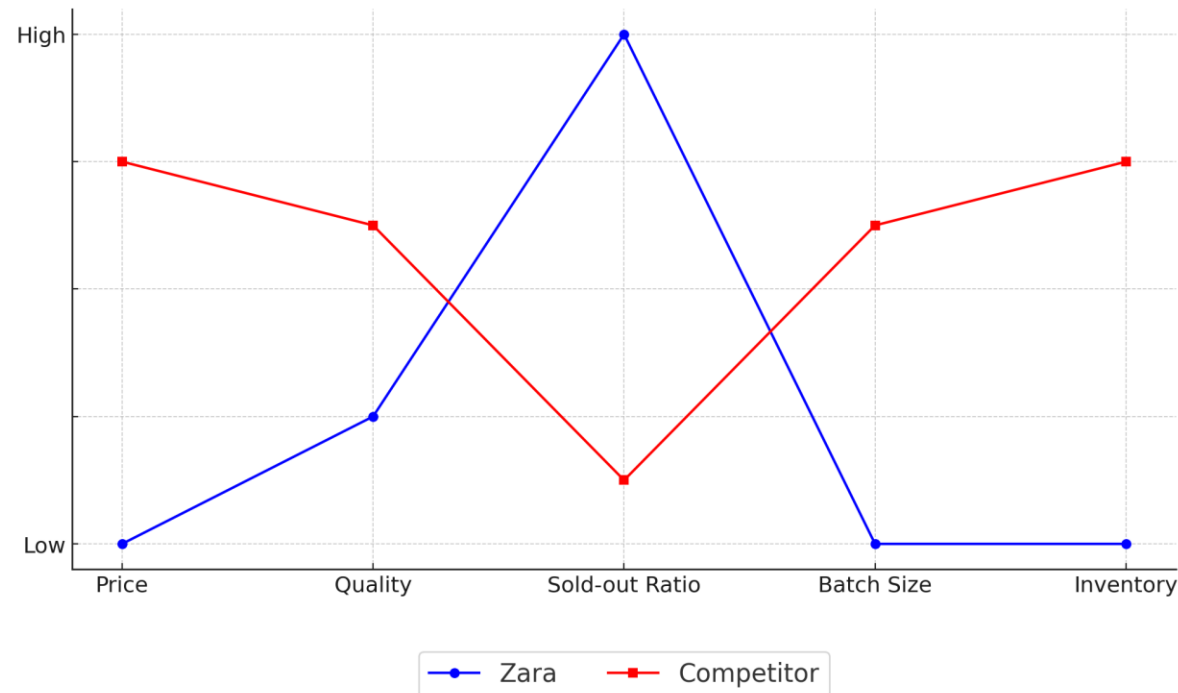
사례1) Zara의 실시간 수요 예측 기반 공급망 전략

- 결과

- 매진율 85% 이상
 - 업계 평균(60~70%) 대비 우수한 성과
- 재고 비용 최소화
- 소비자 만족도와 브랜드 충성도 증가

- 시사점

- 예측 정확도가 높아질수록 재고 리스크는 줄고 고객 반응 속도는 빨라진다



1. 수요 예측 개요

1.5. 수요 예측 사례

사례2) Target Canada의 과잉 예측 사례

- 미국 유통업체 Target이 2013년 캐나다 시장에 진출
- 새로운 시장에 빠른 확장 전략: 133개 매장을 동시에 오픈
- 수요 예측 실패 요인
 - 미국 내 데이터를 그대로 적용 → 현지화 부족
 - 재고 데이터 입력 시스템이 준비되지 않은 상태에서 오픈
 - ERP 시스템 안정화 없이 매장 오픈과 동시에 적용
 - 수요 예측 결과가 각 매장에 즉시 반영되지 않음
 - 피드백 루프의 부재



1. 수요 예측 개요

1.5. 수요 예측 사례

사례2) Target Canada의 과잉 예측 사례

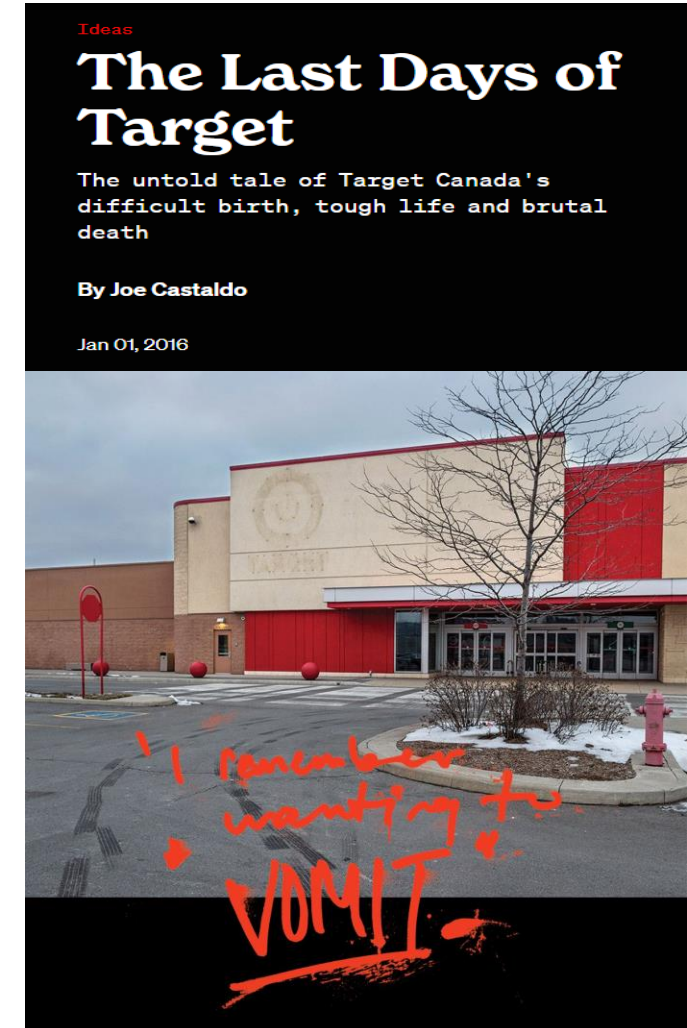
- 결과

- 오픈 2년 만에 전 매장 철수
- 재무적 손실: 약 20억 달러의 손실 발생
- 철수 발표 이후 큰 폭의 주가 하락, CEO 사임
- 고객 불만 폭주 → 브랜드 이미지에 큰 타격

- 시사점

- 시장 특성과 데이터 품질을 무시한 예측은 큰 실패를 초래

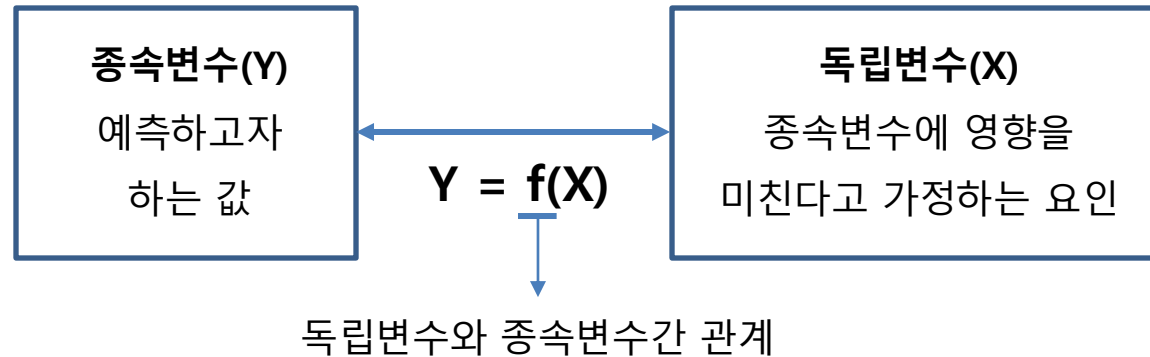
출처: <https://canadianbusiness.com/ideas/the-last-days-of-target-canada/>



2. 전통적 통계 기반 예측 기법

2.1. 회귀분석

- 하나 이상의 독립변수와 종속변수 간의 관계를 **함수로 정의**하고, 해당 함수의 **계수**를 실제 데이터를 바탕으로 **추정**하는 통계적 기법
- 추정된 계수를 통해 독립변수와 종속변수 간 **관계를 파악**할 수 있으며, 독립변수 값으로 종속변수 값을 **예측**할 수도 있음



회귀분석 예시

문제상황: 광고비가 많을수록 매출이 증가할까? 혹은 광고비를 통해 매출을 예측할 수 있을까?



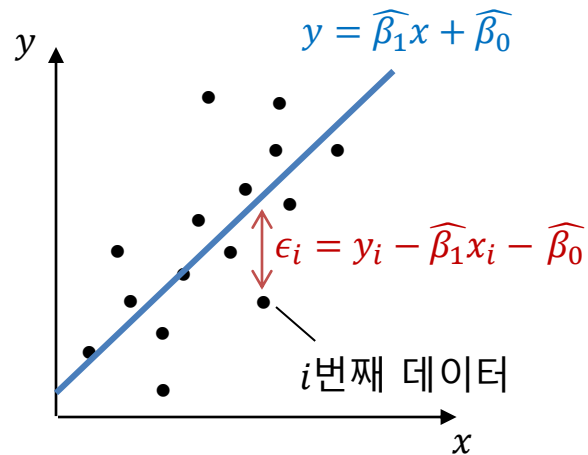
1. 광고비(X)를 독립변수, 매출(Y)을 종속변수로 한 회귀분석 모형 정의: $Y = \alpha X + \beta$
2. 지난 10년간 X, Y 데이터를 통해 α 및 β 를 최소제곱법 등을 통해 추정
3. 올해 X 값과 추정된 α, β 를 통해 올해 Y 값을 계산(예측)

2. 전통적 통계 기반 예측 기법

2.1. 회귀분석: 단순 선형회귀 (Simple linear regression)

- 독립변수가 하나일 때, 독립변수와 종속변수의 관계를 **1차 함수**(직선)로 모델링하는 회귀분석 기법
- 기본 형태: $y_t = \beta_1 x_t + \beta_0 + \epsilon_t, t = 1, 2, \dots, N$

β_0 : 절편, β_1 : 기울기, ϵ : 오차



단순 선형 회귀선 그래프

N 개의 데이터 쌍 (x, y) 를 가장 잘 설명하는 계수 β_0 및 β_1 의 추정치 $\hat{\beta}_0$ 와 $\hat{\beta}_1$ 를 구하는 것이 단순 선형회귀의 목적

가장 잘 설명한다는 것? \longrightarrow 오차 ϵ_i 의 제곱합을 최소화한다는 것

\longrightarrow **최소제곱법 (least squares method)**

1. 수요 예측 개요

2. 전통적 통계 기반 예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

2. 전통적 통계 기반 예측 기법

2.1. 회귀분석: 단순 선형회귀 - 최소제곱법

1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

- 오차 ϵ_i 의 제곱합을 최소화하는 계수 추정치 $\widehat{\beta}_0$ 와 $\widehat{\beta}_1$ 를 구하는 방법
- $\epsilon_i = y_i - \widehat{y}_i = y_i - (\widehat{\beta}_0 + \widehat{\beta}_1 x_i)$

$$\epsilon_i = y_i - \widehat{y}_i = y_i - (\widehat{\beta}_0 + \widehat{\beta}_1 x_i)$$

$$\sum_{i=1}^N \epsilon_i^2 = \sum_{i=1}^N [y_i - (\widehat{\beta}_0 + \widehat{\beta}_1 x_i)]^2$$

Q



$$\frac{\partial Q}{\partial \beta_0} = \sum_{i=1}^N (-2)[y_i - (\widehat{\beta}_0 + \widehat{\beta}_1 x_i)] = 0$$

$$\frac{\partial Q}{\partial \beta_1} = \sum_{i=1}^N (-2)x_i[y_i - (\widehat{\beta}_0 + \widehat{\beta}_1 x_i)] = 0$$

$$\widehat{\beta}_1 = \frac{N(\sum_{i=1}^N x_i y_i) - (\sum_{i=1}^N x_i)(\sum_{i=1}^N y_i)}{N(\sum_{i=1}^N x_i^2) - (\sum_{i=1}^N x_i)^2}$$

$$\widehat{\beta}_0 = \frac{\sum_{i=1}^N y_i - \widehat{\beta}_1 \sum_{i=1}^N x_i}{N}$$

2. 전통적 통계 기반 예측 기법

2.1. 회귀분석: 다중 선형회귀 (Multiple linear regression)

- 독립변수가 **두 개 이상**일 때, 독립변수와 종속변수의 관계를 1차 함수로 모델링하는 회귀분석 기법
- 기본 형태: $y_t = \beta_k x_k^t + \dots + \beta_1 x_1^t + \beta_0 + \epsilon_t, t = 1, 2, \dots, N$
- 단순 선형회귀와 동일하게 최소제곱법을 통해 계수들의 추정치 $\widehat{\beta}_0, \widehat{\beta}_1, \dots, \widehat{\beta}_k$ 을 계산할 수 있음

선형회귀분석이 잘 작동되기 위한 조건

- 선형성: 독립변수와 종속변수 간 관계는 선형적
- 독립성: 오차항들은 서로 독립적
- 등분산성: 오차의 분산은 일정. 즉 오차가 종속변수의 모든 값에서 비슷하게 나타남
- 정규성: 오차항들은 정규 분포를 따름
- 다중공선성: 다중 선형회귀의 경우, 독립 변수들 간에 강한 상관관계가 없음

2. 전통적 통계 기반 예측 기법

2.1. 회귀분석: 선형회귀분석이 적합한 데이터

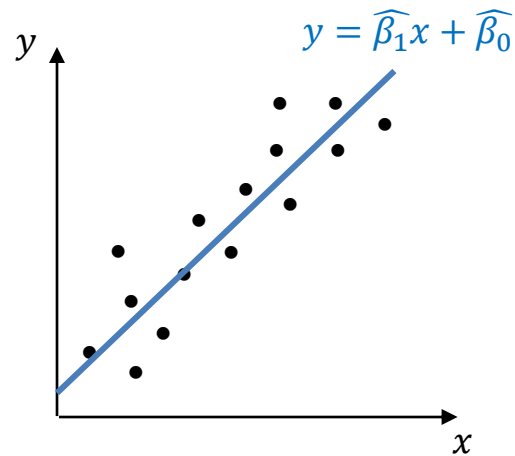
1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

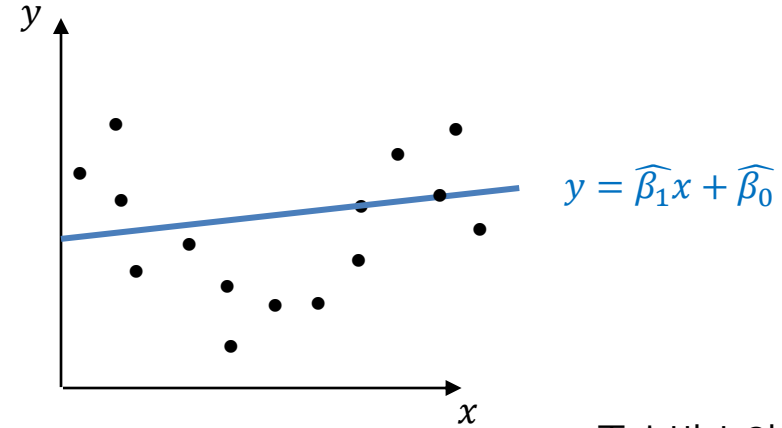
4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

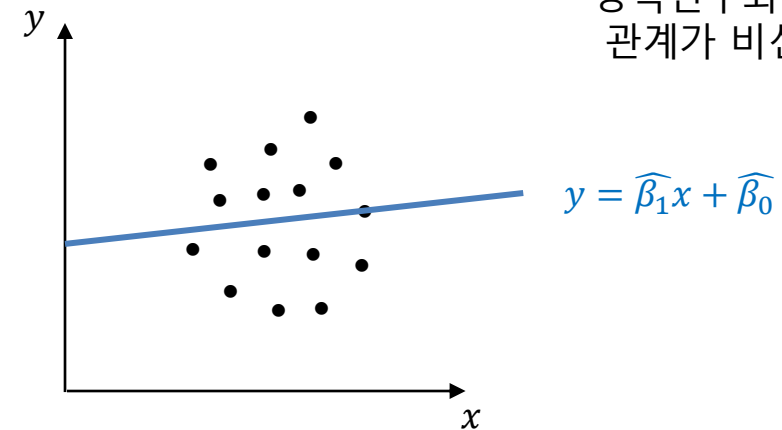


종속변수와 독립변수간
선형 관계가 뚜렷한 경우

적합



종속변수와 독립변수간
관계가 비선형인 경우



부적합

2. 전통적 통계 기반 예측 기법

2.1. 회귀분석: 단순 선형회귀 사용 예시

- 2025년 광고비 금액을 통해 2025년 수요를 예측하고자 하는 상황

	A	B	C
1	연도	광고비(천만원)	수요(10만개)
2	2020	10	2.1
3	2021	15	2.5
4	2022	20	3.2
5	2023	25	3.8
6	2024	30	4.3
7	2025	28	

데이터: 어느정도 선형
관계가 있는 것으로 파악됨

예측하고자 하는 값

엑셀 또는 파이썬 등의 프로그래밍
언어를 활용하여 최소제곱법으로
계수 추정치 계산

$$\hat{\beta}_1 = 0.114$$

$$\hat{\beta}_0 = 0.9$$

$$2025년\ 수요\ 예측치 = \underbrace{0.114}_{\hat{\beta}_1} \times \underbrace{28}_{2025년\ 광고비} + \underbrace{0.9}_{\hat{\beta}_0} = 4.092\ (10만개)$$

2025년 광고비

1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

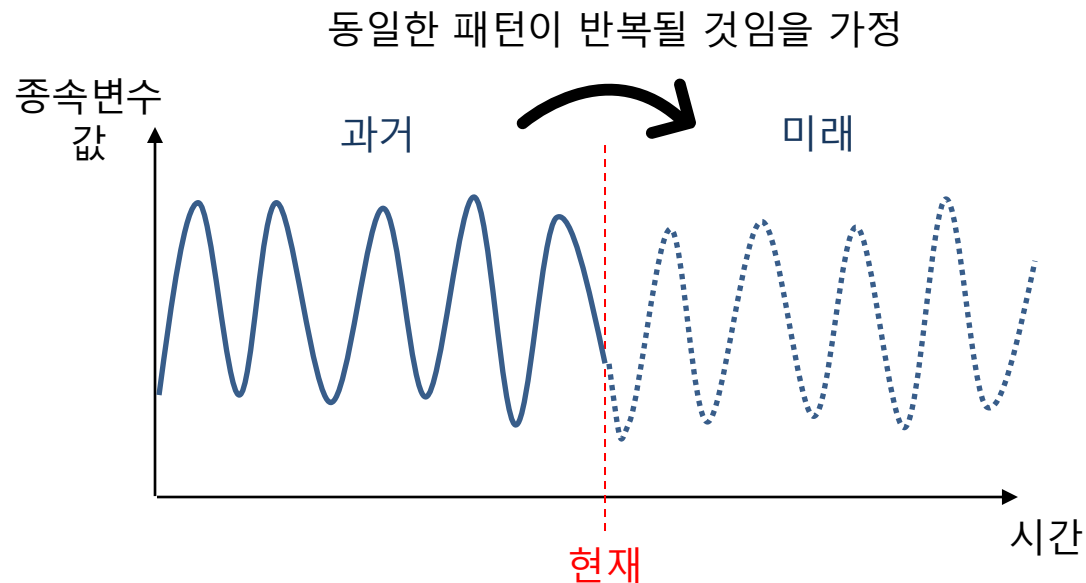
4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

2. 전통적 통계 기반 예측 기법

2.2. 시계열 분석

- 회귀분석과 달리 독립변수를 사용하지 않고, 과거 데이터 중 종속변수만을 고려하여 종속변수의 패턴을 분석하고 이를 바탕으로 미래의 값을 예측하는 통계적 기법
- 기본 가정: 종속변수의 과거의 행태가 미래에도 유사하게 지속된다



1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

2. 전통적 통계 기반 예측 기법

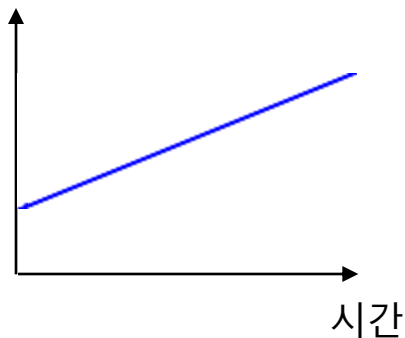
2.2. 시계열 분석: 변동요인

- 시계열 데이터에서는 규칙적인 변동, 즉 패턴이 발견될 수 있음
- 우연적이고 불규칙적인 변동은 예측하기 쉽지 않음

추세(Trend)

장기적으로 나타나는
상승 또는 하락

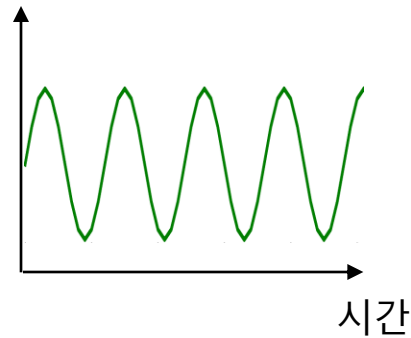
예시: 인구 증가, 기술
발전에 따른 생산성 향상



계절성(Seasonality)

일정한 주기를 가지고
반복되는 패턴

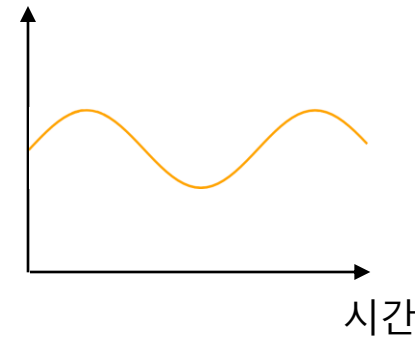
예시: 여름철 에어컨 판매
증가, 연말 쇼핑 수요



순환성(Cycle)

정치적, 경제적 변화
등으로 인해 긴 주기를
갖고 반복되는 패턴

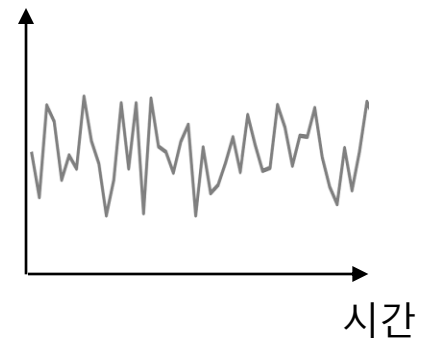
예시: 경기 호황 및 불황
주기, 부동산 사이클



불규칙성(Random component)

설명 불가능한 우연적 요인으로
인해 발생하는 변동

예시: 자연재해, 공급망 붕괴
등으로 인한 수요 변화



1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

2. 전통적 통계 기반 예측 기법

2.2. 시계열 분석: 변동요인

- 두 가지 이상의 변동요인이 모두 존재하는 경우도 있음

1. 수요 예측 개요

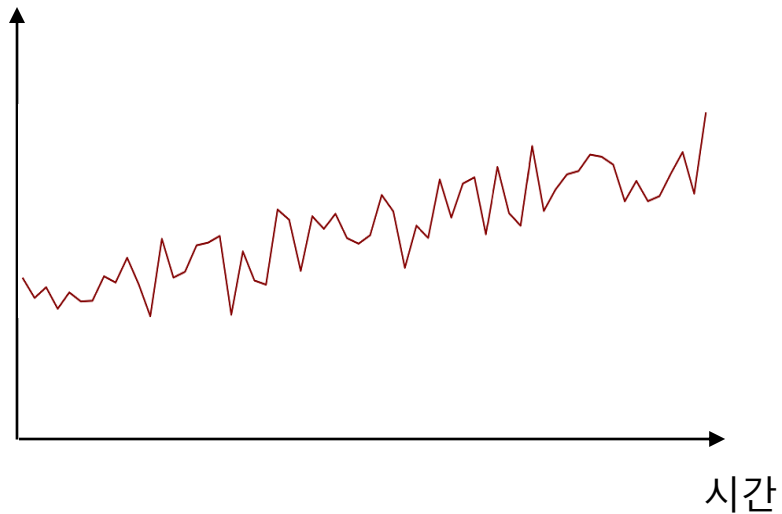
2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

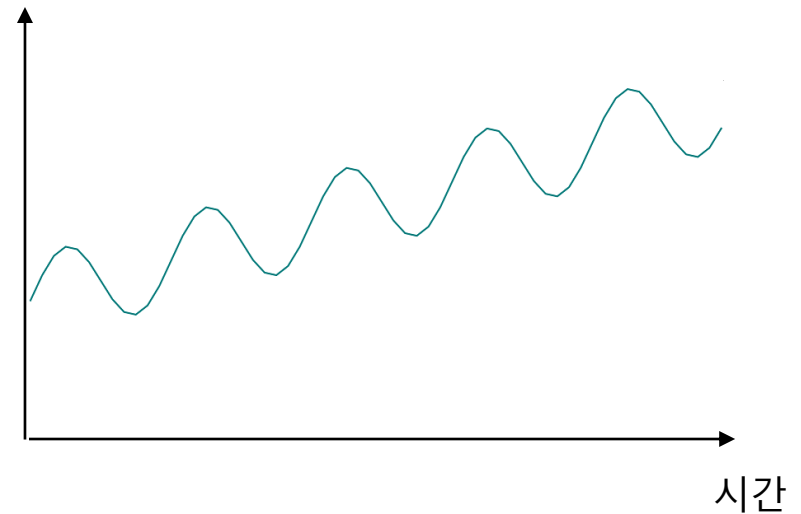
추세 + 불규칙성



예시: 소셜 미디어 트래픽

사용자 수는 점차 증가하면서 예상치 못한
뉴스 이슈/사건에 따라 트래픽이 변동함

계절성 + 추세



예시: 전력 소비량

매년 전력 소비량은 증가하는
추세이지만 계절별 변동 또한 존재함

2. 전통적 통계 기반 예측 기법

2.2. 시계열 분석: 이동평균 (Moving average)

- 과거의 일정 구간 데이터를 평균 내어 미래를 예측하는 방식
- 단순하고 직관적이며, 노이즈 제거와 패턴 파악에 유용함
- 급격한 데이터 변화에 대해 반응이 느림

단순 이동평균 (Simple moving average)

- 최근 N 개의 데이터 평균으로 예측

$$\hat{y}_t = \frac{y_{t-1} + y_{t-2} + \cdots + y_{t-N}}{N}$$

\hat{y}_t : t 시점 예측치
최근 N 개 데이터 평균

2. 전통적 통계 기반 예측 기법

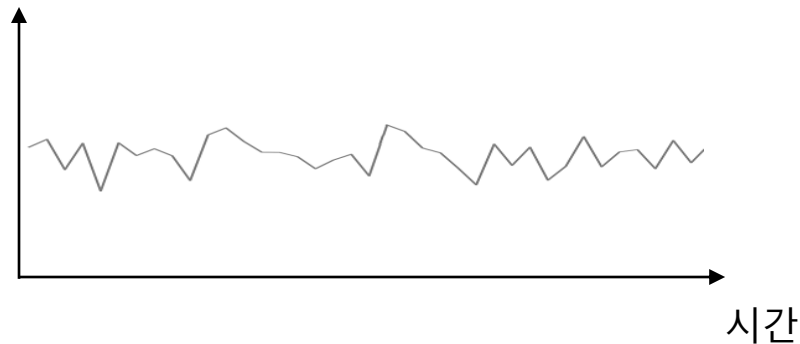
2.2. 시계열 분석: 가중 이동평균 (Weighted moving average)

- 최근 데이터에 더 큰 가중치를 둠

$$\hat{y}_t = w_1 y_{t-1} + w_2 y_{t-2} + \cdots + w_N y_{t-N}$$

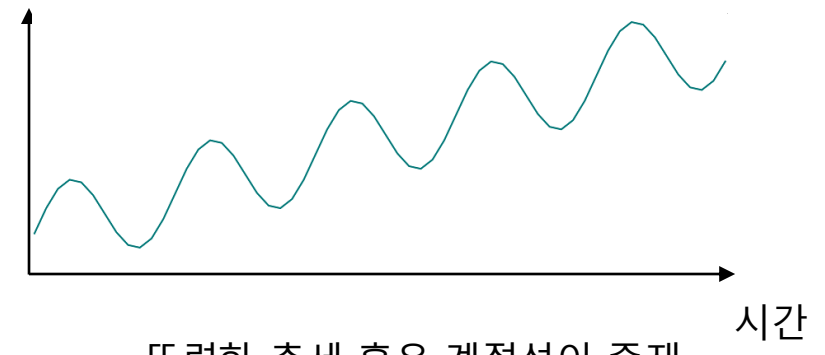
$$w_1 + w_2 + \cdots + w_N = 1$$

이동평균 기법이 적합한 데이터



패턴이 크지 않고, 불규칙성만 존재
→ 예측값이 안정적이며 유의미함

적합



뚜렷한 추세 혹은 계절성이 존재
→ 예측이 실제 패턴을 따라가지 못함

부적합

1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

2. 전통적 통계 기반 예측 기법

2.2. 시계열 분석: 지수평활법 (Exponential smoothing)

- 과거 데이터에 지수적으로 감소하는 가중치를 적용하여 예측
- 최근 데이터에 더 큰 비중을 두어 변화에 민감하게 반응

단순 지수평활 (Simple exponential smoothing)

- 이동평균법과 유사하게 추세 및 계절성이 없는 데이터에 사용하기 적합

$$\hat{y}_t = \alpha y_{t-1} + (1 - \alpha) \widehat{y}_{t-1}$$

\hat{y}_t : t 시점 예측치

\widehat{y}_{t-1} : $t-1$ 시점 예측치

α : 평활 상수 ($0 < \alpha < 1$)

α 값이 클수록 최근 데이터에 민감하게 반응
 α 값이 작을수록 오래된 데이터까지 반영

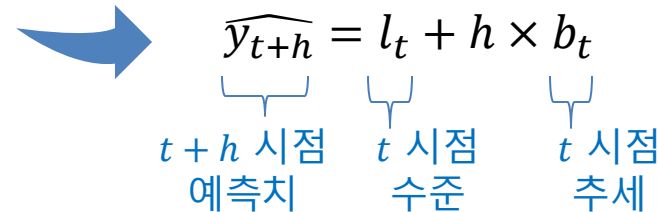
2. 전통적 통계 기반 예측 기법

2.2. 시계열 분석: 홀트 선형 평활법 (Holt's linear exponential smoothing)

- 추세(추세)가 있는 시계열 데이터에 적합
- 추세가 반영되어 장기 예측이 가능
- 계절성까지는 다루지 못함

$$l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1}): \text{수준(level)}$$

$$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}: \text{추세(trend)}$$



A blue arrow points from the left towards the forecast equation. The equation is $\widehat{y_{t+h}} = l_t + h \times b_t$. Below the equation, three curly braces are used to group and label the terms: the first brace under $\widehat{y_{t+h}}$ is labeled "t + h 시점 예측치" (Forecast at time t+h); the second brace under l_t is labeled "t 시점 수준" (Level at time t); and the third brace under b_t is labeled "t 시점 추세" (Trend at time t).

$$\widehat{y_{t+h}} = l_t + h \times b_t$$

t + h 시점
예측치

t 시점
수준

t 시점
추세

2. 전통적 통계 기반 예측 기법

2.2. 시계열 분석: 계절지수 적용 예시

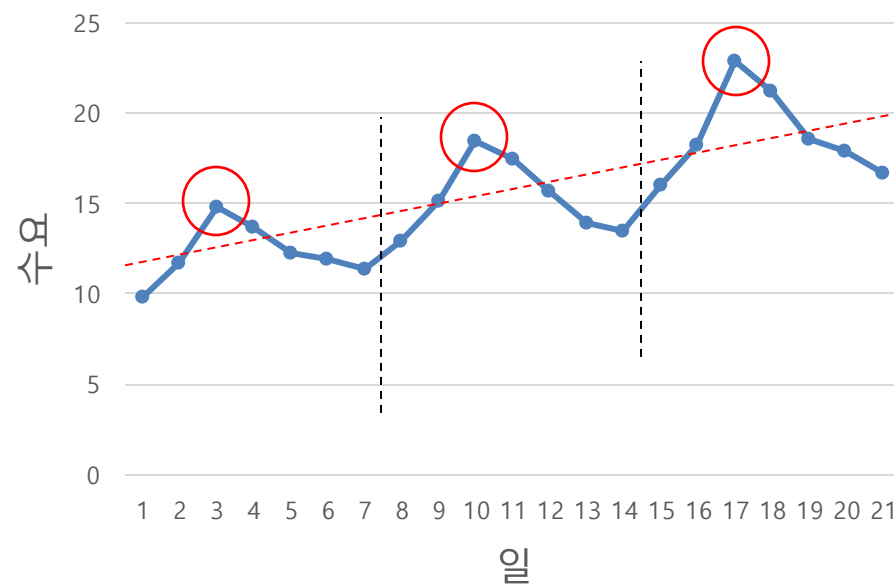
1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법



주기가 7인 계절성을 가지는 데이터

세번째 계절(수요일)은 평균적인 추세보다 큰
경향을 보임

계절지수 α_3 를 1보다 큰 값으로 설정

$$\text{예시) } \alpha_3 = \frac{(14.86 + 18.46 + 22.89) / 3}{15.44} = 1.21$$

세번째 계절(수요일)의 평균

전체 수요 평균

2. 전통적 통계 기반 예측 기법

2.2. 시계열 분석: 홀트-윈터스 평활법 (Holt-Winters exponential smoothing)

- 추세와 계절성을 모두 고려하는 지수평활법
- 가법형(additive) 방식과 승법형(multiplicative) 방식이 있음

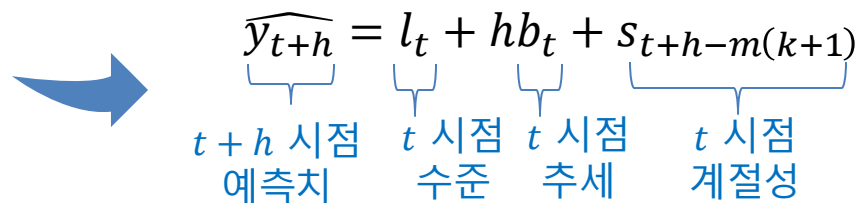
가법형

$$l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1}): \text{수준(level)}$$

$$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}: \text{추세(trend)}$$

$$s_t = \gamma(y_t - l_t) + (1 - \gamma)s_{t-m}: \text{계절성(seasonality)}$$

$$m: \text{계절성 주기}, k = \lceil \frac{h-1}{m} \rceil$$


$$\widehat{y_{t+h}} = \underbrace{l_t}_{t+h \text{ 시점 예측치}} + \underbrace{h}_{t \text{ 시점 수준}} \underbrace{b_t}_{t \text{ 시점 추세}} + \underbrace{s_{t+h-m(k+1)}}_{t \text{ 시점 계절성}}$$

2. 전통적 통계 기반 예측 기법


2.2. 시계열 분석: 홀트-윈터스 평활법 (Holt-Winters exponential smoothing)

승법형

$$l_t = \alpha \left(\frac{y_t}{s_{t-m}} \right) + (1 - \alpha)(l_{t-1} + b_{t-1})$$

$$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$$

$$s_t = \gamma \left(\frac{y_t}{l_t} \right) + (1 - \gamma)s_{t-m}$$


$$\widehat{y_{t+h}} = \underbrace{(l_t + hb_t)}_{\substack{t+h \text{ 시점} \\ \text{예측치}}} \times \underbrace{s_{t+h-m(k+1)}}_{\substack{t \text{ 시점} \\ \text{계절성}}}$$

t 시점 수준 t 시점 추세



2. 전통적 통계 기반 예측 기법

2.2. 시계열 분석: 홀트-윈터스 평활법 (Holt-Winters exponential smoothing)

1. 수요 예측 개요

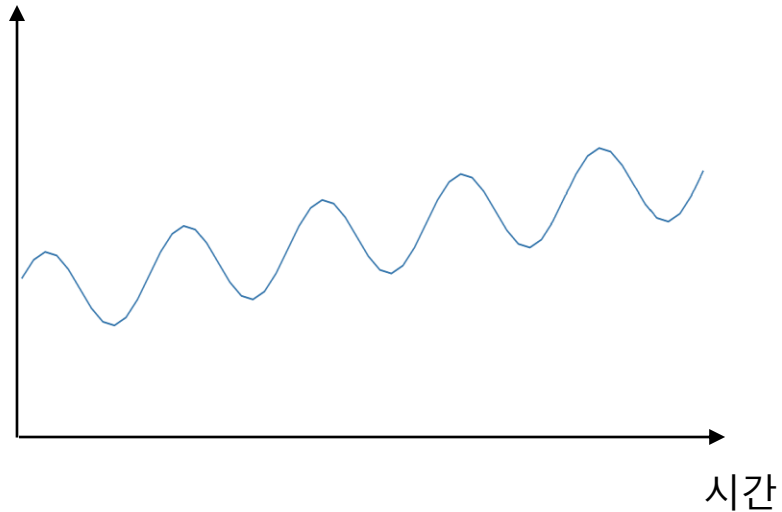
2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

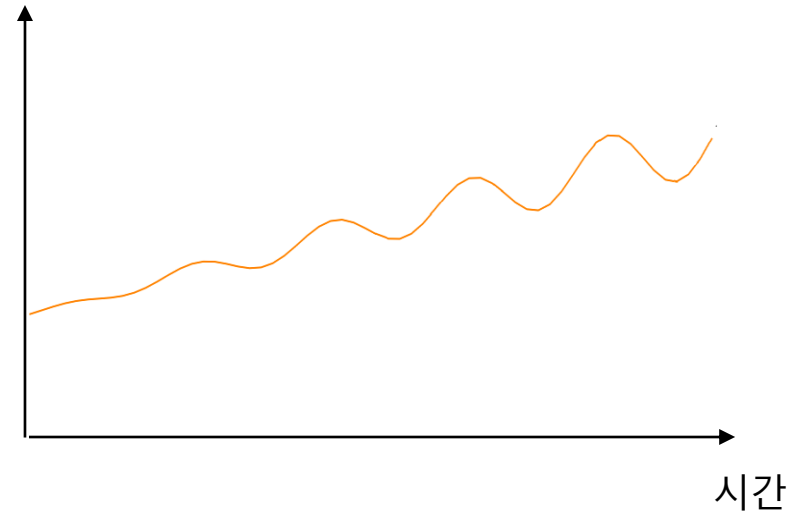
5. 하이브리드 예측 기법

계절성의 크기가 일정한 경우



가법형 방식이 적절

계절성이 추세와 함께 비례하여 증가



승법형 방식이 적절

2. 전통적 통계 기반 예측 기법

2.2. 시계열 분석: Holt-윈터스 평활법 (승법형) 활용 예시

1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법



과거 12개월의 데이터를 통해
13개월차의 수요를 예측하고자 함

② 보유한 마지막 데이터로부터 1 시점 뒤의
값을 예측하고자 하므로 h 를 1로 설정

③ 수준, 추세, 계절성의 평활 계수 α, β, γ 를 각각 0과 1사이의 값으로 설정
(1에 가까운 값으로 설정할수록 최근값을 중요시한다는 뜻)

2. 전통적 통계 기반 예측 기법

2.2. 시계열 분석: 홀트-윈터스 평활법 (승법형) 활용 예시

1. 수요 예측 개요

2. 전통적 통계 기반 예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

④ 앞서 설명된 수식을 통해 l_0, b_0, s_0 부터 l_{12}, b_{12}, s_{12} 까지 순차적으로 계산 (엑셀 혹은 파이썬 활용)

※ l_0 (초기 수준), b_0 (초기 추세), $s_0 \sim s_2$ (초기 계절성)은 아래와 같은 방식으로 계산

시점	수요
0	10.21
1	23.01
2	10.97
3	14.59
4	29.44
5	16.80
6	18.86
7	38.90
8	18.61
9	24.20
10	48.90
11	22.78

주기(m)=3

평균 14.73

평균 20.28

l_0 : 첫번째 주기의 수요 평균 $\rightarrow 14.73$

b_0 : 첫번째 주기와 두번째 주기의 수요 평균 차이 \div 주기 $\rightarrow 5.55 \div 3 = 1.85$

$s_0 \sim s_3$: (각 계절 시점의 수요를 해당 주기 평균으로 나눈 값)의 평균

$$\rightarrow s_0 = \frac{1}{2} \times [10.21 \div 14.73 + 14.59 \div 20.28] = 0.71$$

두번째 주기 평균

첫번째 주기 평균

$$s_1 = \frac{1}{2} \times [23.01 \div 14.73 + 29.44 \div 20.28] = 1.51$$

$$s_2 = \frac{1}{2} \times [10.97 \div 14.73 + 16.80 \div 20.28] = 0.79$$



첫번째 주기의
세번째 계절의 수요

두번째 주기의
세번째 계절의 수요

⑤ 13번째 수요의 예측치 계산: $\widehat{y}_{12+1} = (l_{12} + 1 \times b_{12}) \times s_{12+1-3(0+1)}$

3. ARIMA 계열 예측 기법

3.1. ARIMA 개요

- ARIMA (Autoregressive Integrated Moving Average) 란?

과거(p)·추세(d)·오차(q)를 한 번에 고려해 미래 값을 예측하는 시계열 모델!!

p, d, q 최적값을 지정하는 라이브러리 사용 (AutoARIMA)

$$\underbrace{\phi(L)}_p \underbrace{(1-L)^d}_d X_t = \underbrace{\theta(L)}_q \varepsilon_t$$

AR (p) - $\phi(L)$: 과거 p시점 관측치를 선형 결합하여 현재 값 예측

I (d) - $(1-L)^d$: 시계열을 d회 차분해 정상성 확보

MA (q) - $\theta(L)$: 과거 q시점 오차를 이동 평균하여 잡음 제거

3. ARIMA 계열 예측 기법

3.1. ARIMA 개요

- 자동차 비유로 쉽게 풀어보는 ARIMA

"AR" + "I" + "MA"

$ARIMA(p, d, q)$

AR 모형 차수

차분

MA 모형 차수

→ AR : 과거 신호 참조하여 속도 조절하기

 I : 언덕길은 엑셀 · 브레이크로 평지처럼 운전하기

 MA : 요철 등은 서스펜션으로 충격 흡수하듯 잡음 정리하기

과거(AR) + 추세제거(I) + 잡음정리(MA) 한 번에 고려하여 값을 찾는 예측

3. ARIMA 계열 예측 기법

3.2. ARIMA 분석 절차

1. 식별 (Identification)

- 과거(p)·오차(q) 차수 식별 후 결정!
- ACF(자동 상관) → MA(q) 아이디어
- PACF(부분 상관) → AR(p) 아이디어

3. 진단 (Diagnostic)

- 잔차(Residual)가 랜덤인지 확인!
- 잔차 ACF 그려서 백색소음 확인

2. 추정 (Estimation)

- ϕ, θ 계수(AR·MA 계수) 계산!
- ML(최대우도) or LS(최소제곱) 사용

4. 예측 (Forecasting)

- 미래값 + $\pm 95\%$ 신뢰구간 출력
- fitted 모델에 steps 지정 후 forecast

식별→추정→진단→예측 이 4단계를 반복하며, 최적의 p, d, q 모델을 완성

1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

3. ARIMA 계열 예측 기법

3.3. ARIMA 주요 특징

- **파라미터형 모델**

p,d,q 만 지정하면 간단 적용 가능

- **해석 용이성**

각 계수(φ , θ)의 의미가 명확해 커뮤니케이션에 강점

- **선형성 기반**

복잡한 비선형성 없이 빠른 연산 가능

- **확장성**

SARIMA(계절성), ARIMAX(외생 변수)로 쉽게 확장 가능

1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

3. ARIMA 계열 예측 기법

3.4. ARIMA 추천 활용 분야

단기 수요·매출 예측

- 프로모션 전후의 매출 변화 추적
- 재고관리 최적화

금융 시계열 분석

- 주가·환율 동향 예측
- VaR(Value at Risk) 계산 보조

에너지·교통량 예측

- 전력 수요 피크 예측
- 도로/대중교통 이용량 예측

공정 제어 모니터링

- 생산 라인 센서 값 이상 탐지
- 설비 유지보수 시점 예측

계절성 강하면 SARIMA, 외부 변수 넣으려면 ARIMAX로 확장!!!

1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

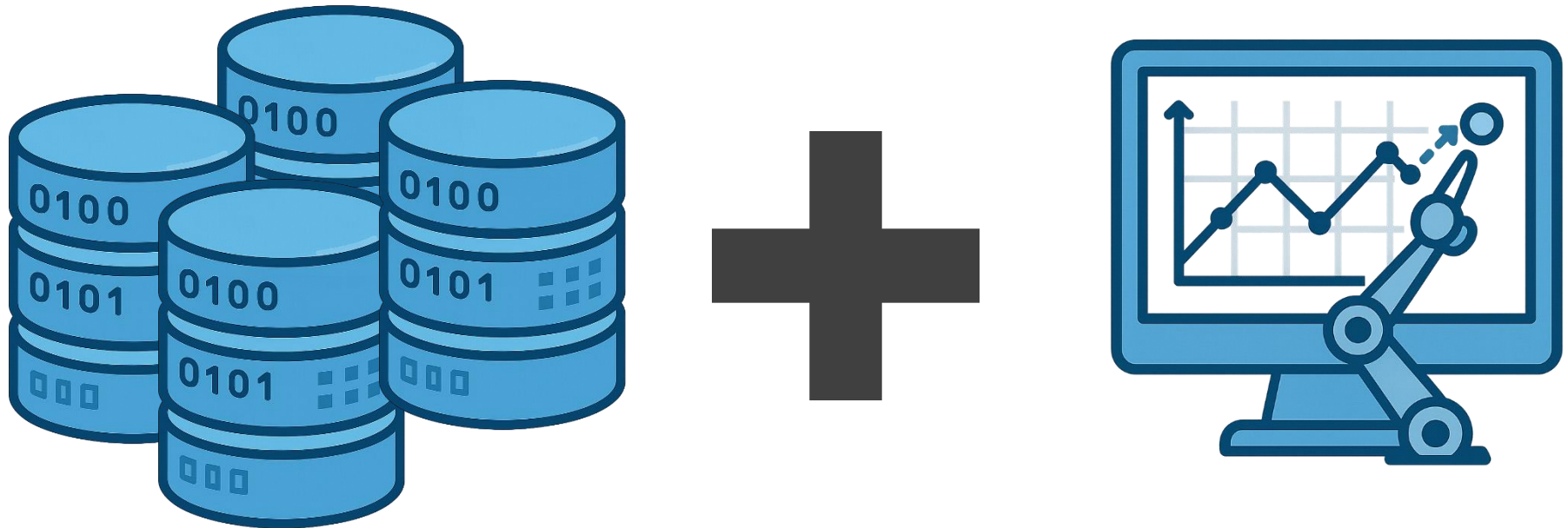
5. 하이브리드 예측 기법

4. 머신러닝 예측 기법

4.1. 머신러닝 개요

- 수요 예측에 활발히 쓰이는 머신러닝

- 엄청난 데이터 양 + 빠른 컴퓨터 → 수요 예측에 머신러닝 기법 활용
- 간단한 패턴, 공식 기반의 전통적 기법 → 컴퓨터가 자동으로 패턴을 찾아 예측



1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

4. 머신러닝 예측 기법

4.1. 머신러닝 개요

- 머신러닝 기법의 장단점

- 장점

- ✓ 복잡한 패턴(비선형)도 잘 찾아냄
 - ✓ 여러 정보(요일, 계절, 이벤트 등)를 한번에 반영 가능
 - ✓ 다수의 독립변수를 모델에 넣을 수 있음

- 단점

- ✓ 충분한 데이터 확보 필요
 - ✓ 컴퓨팅 자원의 확보 필요

1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

4. 머신러닝 예측 기법

4.2. 머신러닝 수요 예측 기법 분류

- 전통적 회귀 모형 기반의 머신러닝 기법
 - ✓ 데이터 학습을 통해 예측 공식을 만들어 수요를 예측하는 방식
 - ✓ 주어진 데이터 외 다양한 정보(계절, 이벤트, 요일 등)를 추가 가능
- 의사결정나무 기반의 머신러닝 기법
 - ✓ 질문을 반복해서 수요를 예측하는 방식
 - ✓ 컴퓨팅 자원을 활용하여 더 복잡한 의사결정나무들을 만들고, 융합할 수 있음

4. 머신러닝 예측 기법

4.3. 전통적 회귀 모형 기반의 머신러닝 기법

- 전통적 회귀 모형에 다양한 정보 추가

- [전통적 회귀 모형] $y = \beta_0 + \beta_1 x$
- '요일 변수' z 를 추가
 - ✓ 금요일이면 $z = 1$, 그 외의 요일이면 $z = 0$
- 예시 데이터

1(월)	2(화)	3(수)	4(목)	5(금)	8(월)	9(화)	10(수)	11(목)	12(금)	15(월)	16(화)	17(수)	18(목)	19(금)
100	102	99	105	87	108	114	112	117	98	117	118	123	122	?

- [전통적 회귀 모형] $y = 95.38 + 1.43x \rightarrow 122.55$
- [요일 변수 추가 모형] $y = 98.54 + 1.36x - 17.56z \rightarrow 106.82$

머신러닝 모델에서는 데이터를 보고 설명력을 높이는 변수를 자동으로 찾아줌

1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

4. 머신러닝 예측 기법

4.3. 전통적 회귀 모형 기반의 머신러닝 기법

- 시계열 데이터 예측 머신러닝 모델 Prophet

- Meta(前 Facebook)사에서 개발한 오픈소스 머신러닝 모델
 - 시계열 데이터(수요) 예측에 탁월한 성과를 보임
 - Python, R 등 프로그래밍 언어로 불러와 무료로 사용 가능
 - 추세, 계절성, 휴일 이벤트 등 변수를 자동으로 추가
-
- 학습을 위해 충분한 데이터 양(1년치 이상)이 있어야 함



Facebook Open Source



Prophet 홈페이지 <https://facebook.github.io/prophet/>

- 1. 수요 예측 개요
- 2. 전통적 통계 기반 예측 기법
- 3. ARIMA 계열 예측 기법
- 4. 머신러닝 예측 기법
- 5. 하이브리드 예측 기법

4. 머신러닝 예측 기법

4.3. 전통적 회귀 모형 기반의 머신러닝 기법

Python으로 Prophet 모델 불러오기

```
1 # Python
2 import pandas as pd
3 from prophet import Prophet
```

데이터 불러오기

```
1 # Python
2 df = pd.read_csv('https://raw.githubusercontent.com/facebook/prophet/main/examples/example_wp_log_peyton_manning.csv')
3 df.head()
```

모델 Fitting

```
1 # Python
2 m = Prophet()
3 m.fit(df)
```

향후 1년간 예측

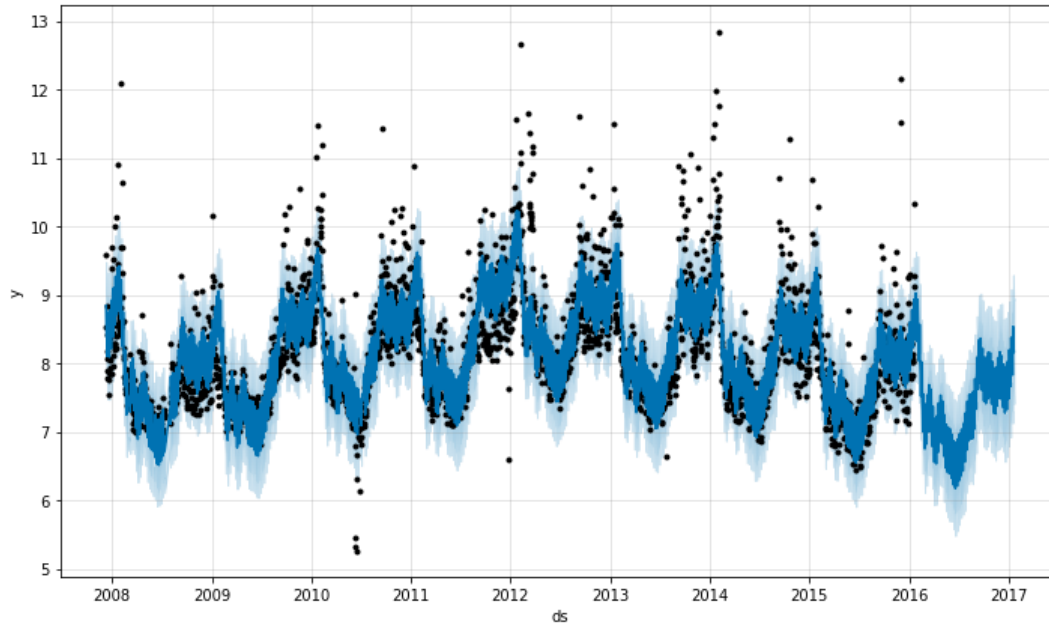
```
1 # Python
2 future = m.make_future_dataframe(periods=365)
3 future.tail()
```

Prophet 홈페이지 https://facebook.github.io/prophet/docs/quick_start.html#python-api

4. 머신러닝 예측 기법

4.3. 전통적 회귀 모형 기반의 머신러닝 기법

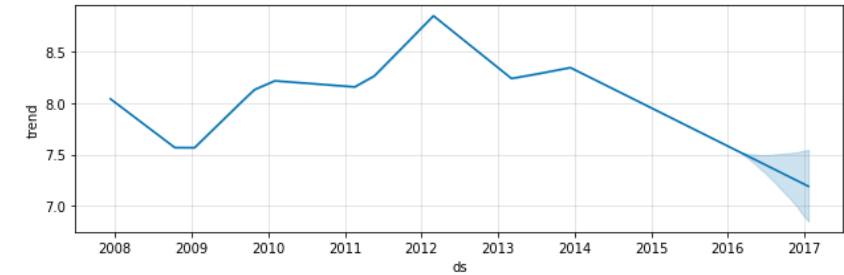
예측 결과



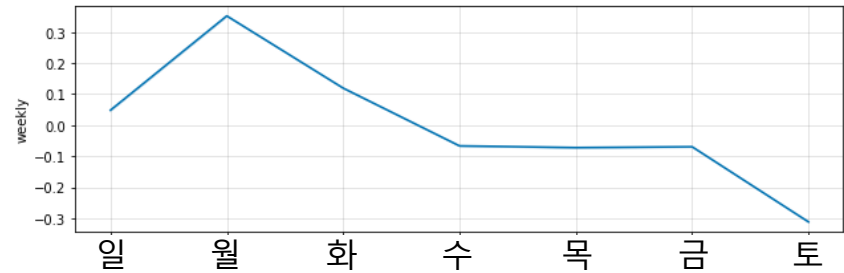
- 검은 점 : 실제 데이터
- 파란 선 : 점 추정치
- 하늘색 범위 : 구간 추정치

추가한 변수들

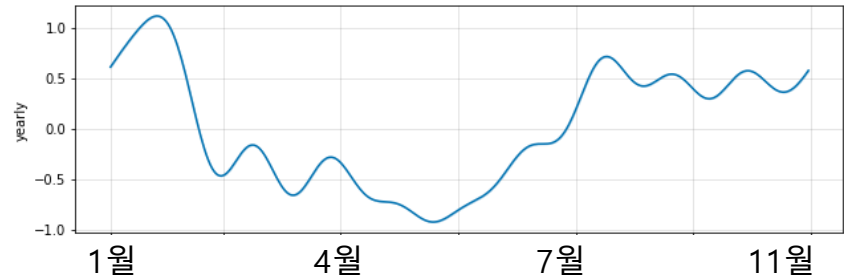
추세



요일



계절성



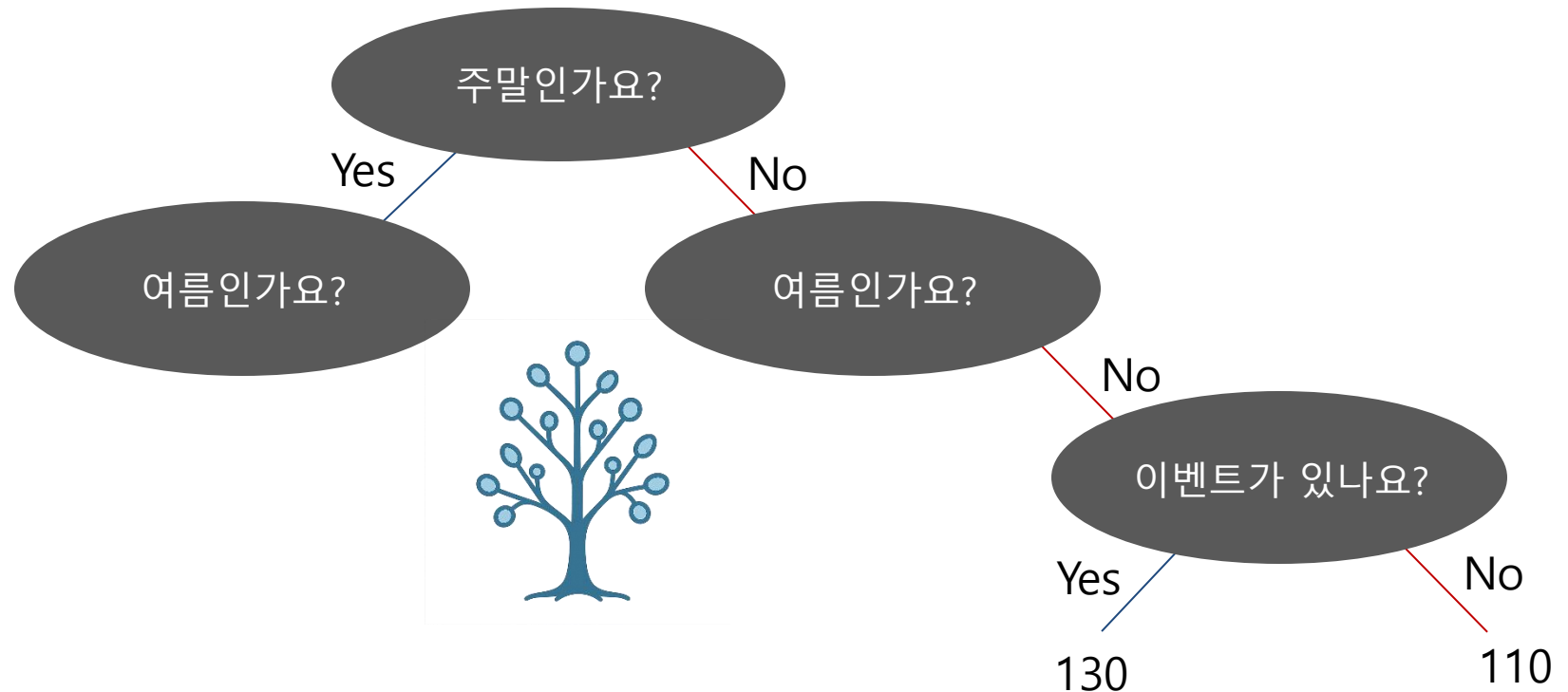
Prophet 홈페이지 https://facebook.github.io/prophet/docs/quick_start.html#python-api

4. 머신러닝 예측 기법

4.4. 의사결정나무 기반의 머신러닝 기법

- 의사 결정 나무(Decision Tree)기반 수요 예측

- 질문을 반복해서 수요를 예측하는 방식
- "오늘이 주말인가요?" → "여름인가요?" → "이벤트가 있나요?"
- 실제로는 모형이 데이터를 보고 자동으로 질문을 만들어 냄



1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

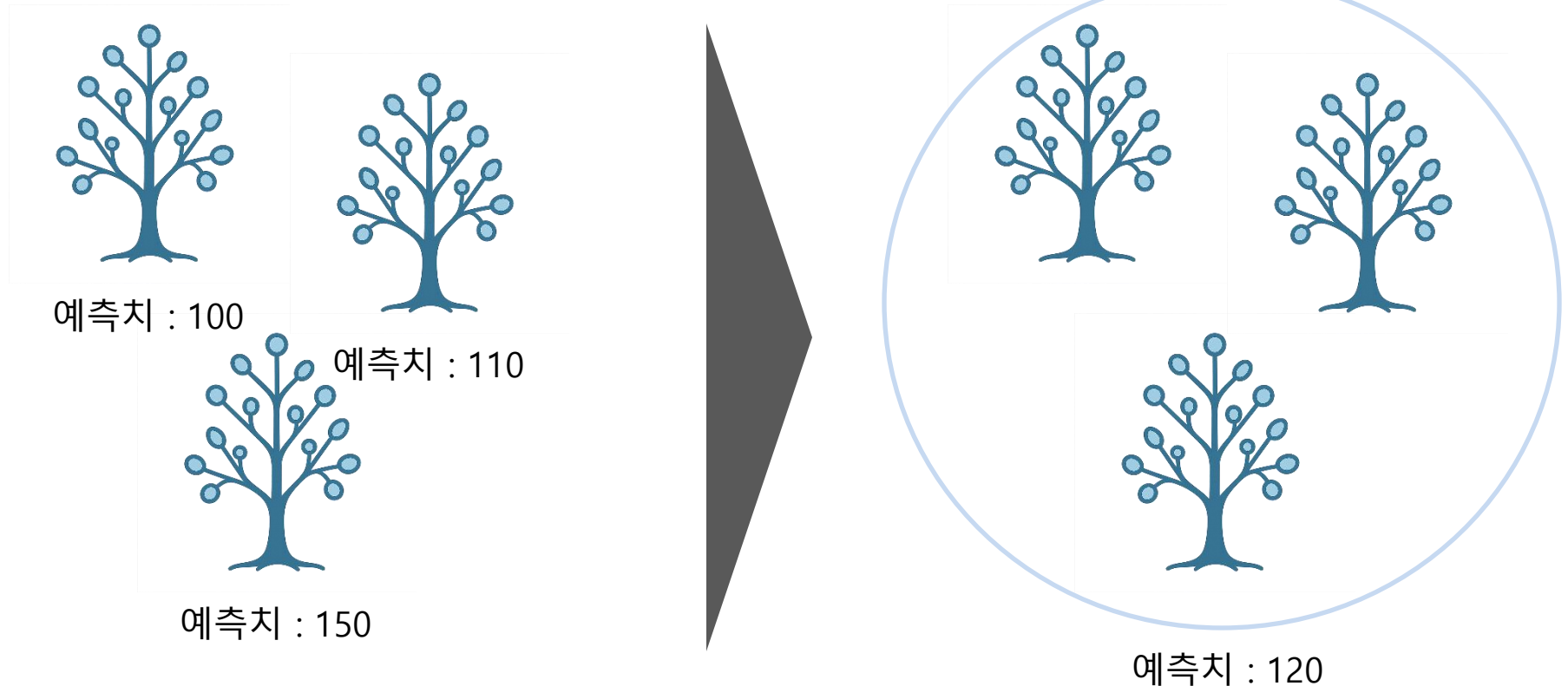
5. 하이브리드 예측 기법

4. 머신러닝 예측 기법

4.4. 의사결정나무 기반의 머신러닝 기법

- 랜덤 포레스트(Random Forest) – 트리의 집단 지성

- 여러 결정 트리를 모아서 예측 정확도를 높임
- “백짓장도 맞들면 낫다”처럼, 여러 트리가 투표해서 더 좋은 답을 찾음

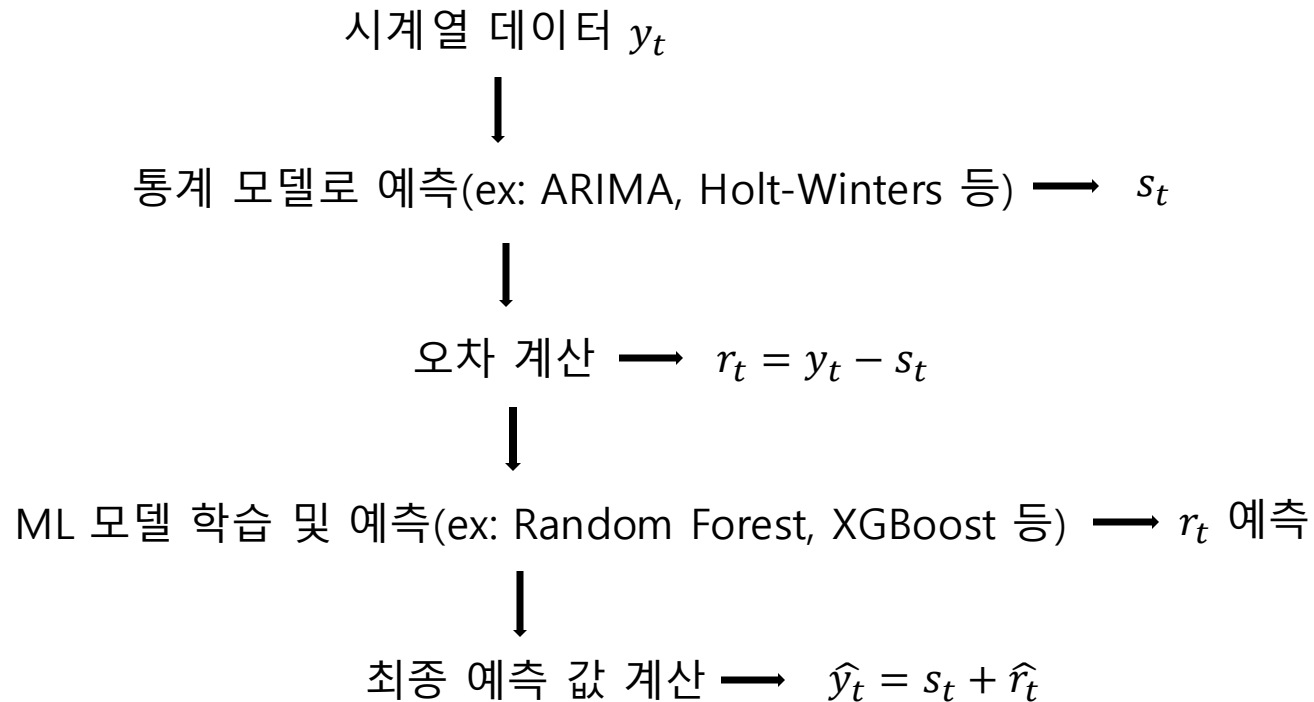


5. 하이브리드 예측 기법

5.1. 하이브리드 예측 기법

- 하이브리드 예측 기법

- 종속변수 y 를 통계 모델로 1차 예측하고, 오차를 머신러닝 모델로 보정하는 방식



- 효과

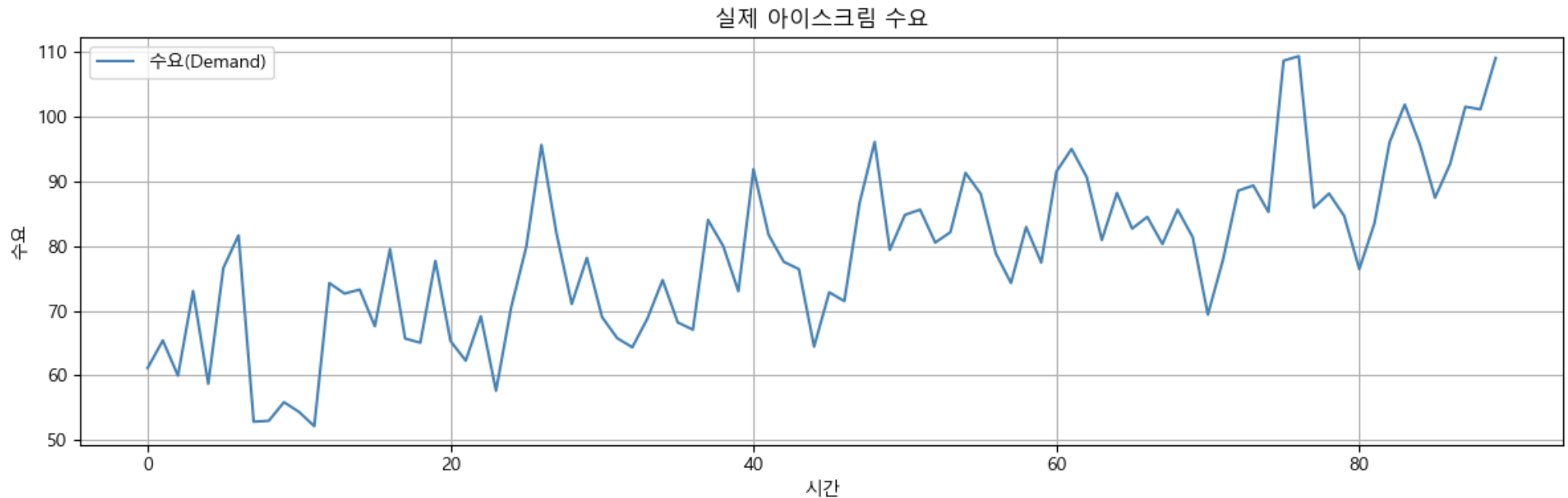
- 통계 모델이 시계열의 기본 패턴을 효과적으로 잡아내고, 머신러닝 모델이 오차의 복잡한 비선형성을 보완하여 예측 정확도와 유연성을 동시에 향상

5. 하이브리드 예측 기법

5.2. 하이브리드 예측 기법 사용 예시

• 예시 문제 – 여름철 아이스크림 수요 예측

- GS THE FRESH는 5월, 6월, 7월 동안의 아이스크림의 수요, 아이스크림 할인율, 프로모션 여부, 일평균 기온 데이터를 가지고 있다. 이를 바탕으로 8월의 아이스크림 수요를 예측하고자 한다. 분석가 A씨는 예측 모델로 전통적인 통계 모형을 사용할지, 머신러닝 모델을 사용할지 고민 중이다. 어떤 모델을 사용하는 것이 좋을까?

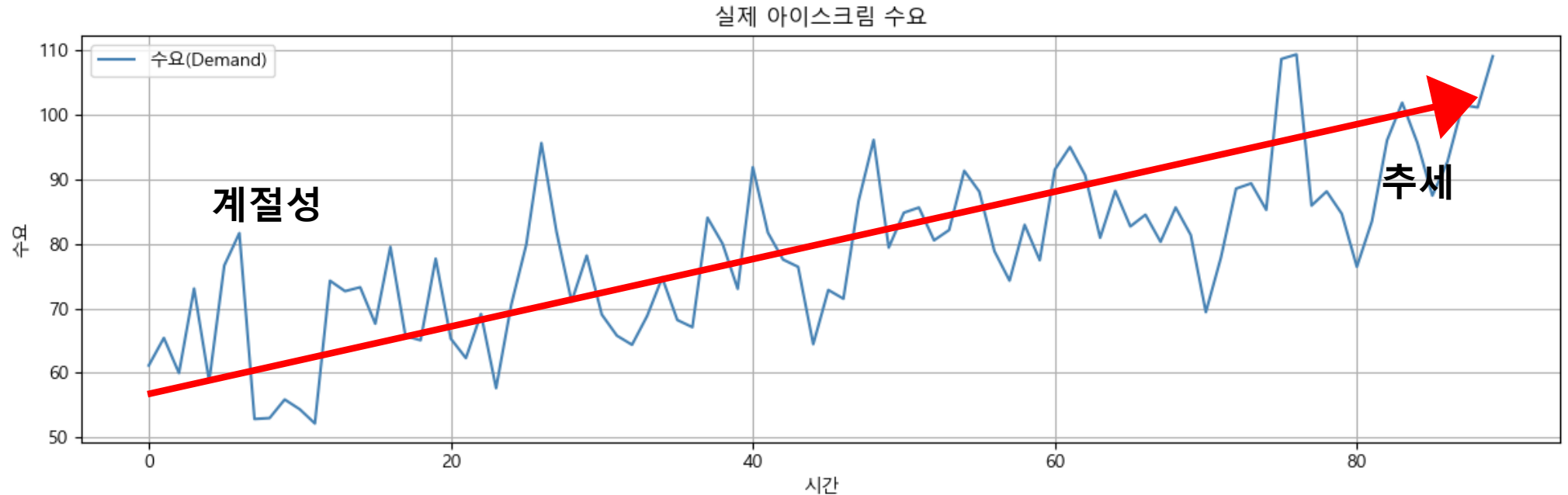


5. 하이브리드 예측 기법

5.2. 하이브리드 예측 기법 사용 예시

- Holt-Winters 모형

- 5월, 6월, 7월 데이터를 보면 시간이 갈수록 수요가 증가하는 추세가 있음
- 요일 별로 (월요일 vs 토요일) 수요가 다른 패턴을 보이는 계절성을 보임
- 추세와 계절성을 반영하는 Holt-Winters 모형을 사용



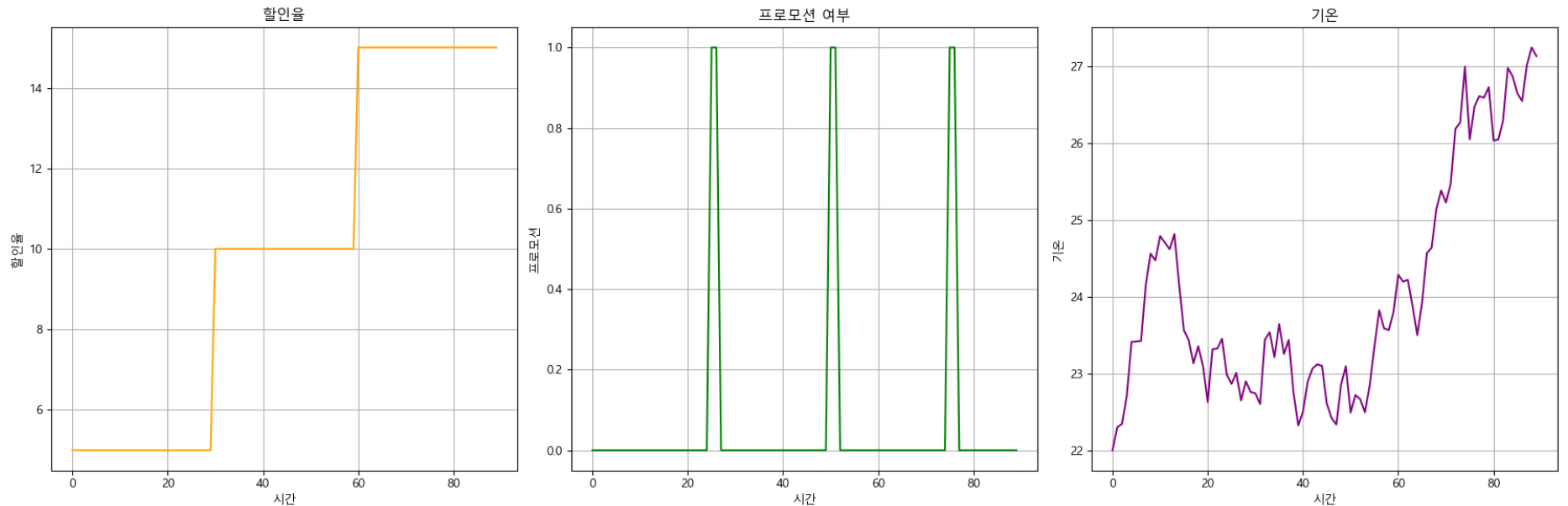
5. 하이브리드 예측 기법

5.2. 하이브리드 예측 기법 사용 예시

- Random Forest 모델

- 할인율 : GS THE FRESH는 아이스크림의 할인율을 주기적으로 올리고 있음
- 프로모션 진행 여부 : GS THE FRESH는 주기적으로 1+1 행사 등의 프로모션을 이틀간 진행함
- 일평균 기온 : 5월, 6월, 7월의 일평균 기온 데이터를 가지고 있음

이 변수들을 이용해 Random Forest 모델을 학습



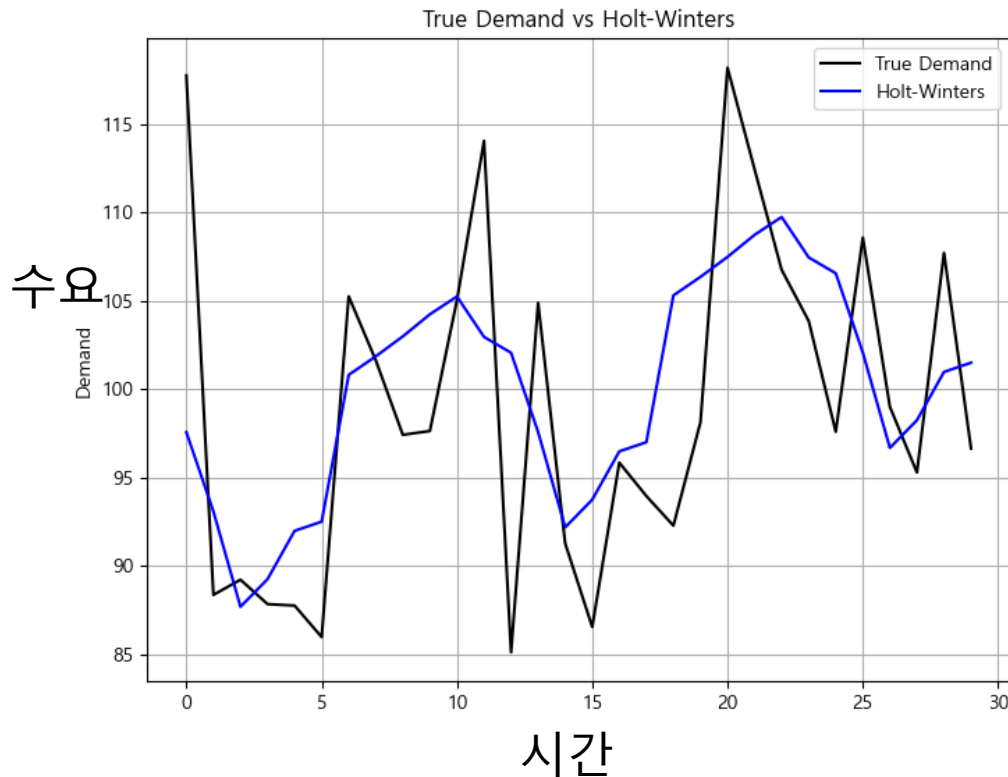
5. 하이브리드 예측 기법

5.2. 하이브리드 예측 기법 사용 예시

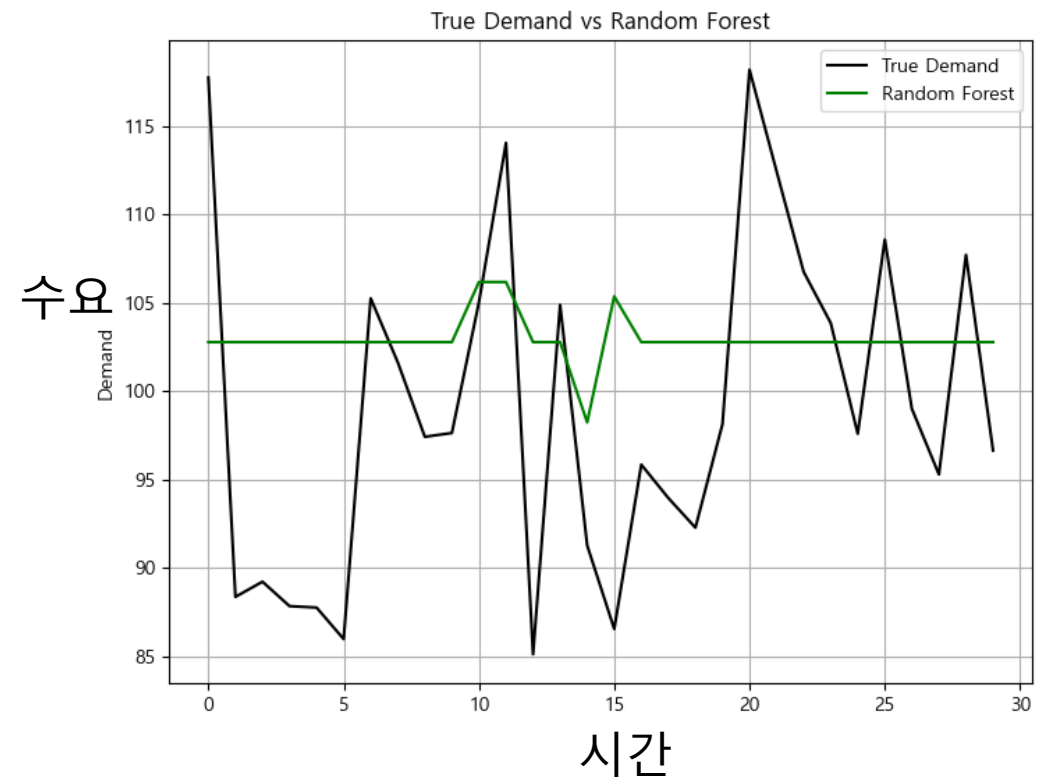
- 통계 모델과 인공지능 모델 비교

- 30일 동안의 아이스크림 수요를 예측한 결과 Holt-Winters 모델이 Random Forest 모델보다 성능이 좋음

Holt-Winters 로 예측



Random Forest 로 예측



1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

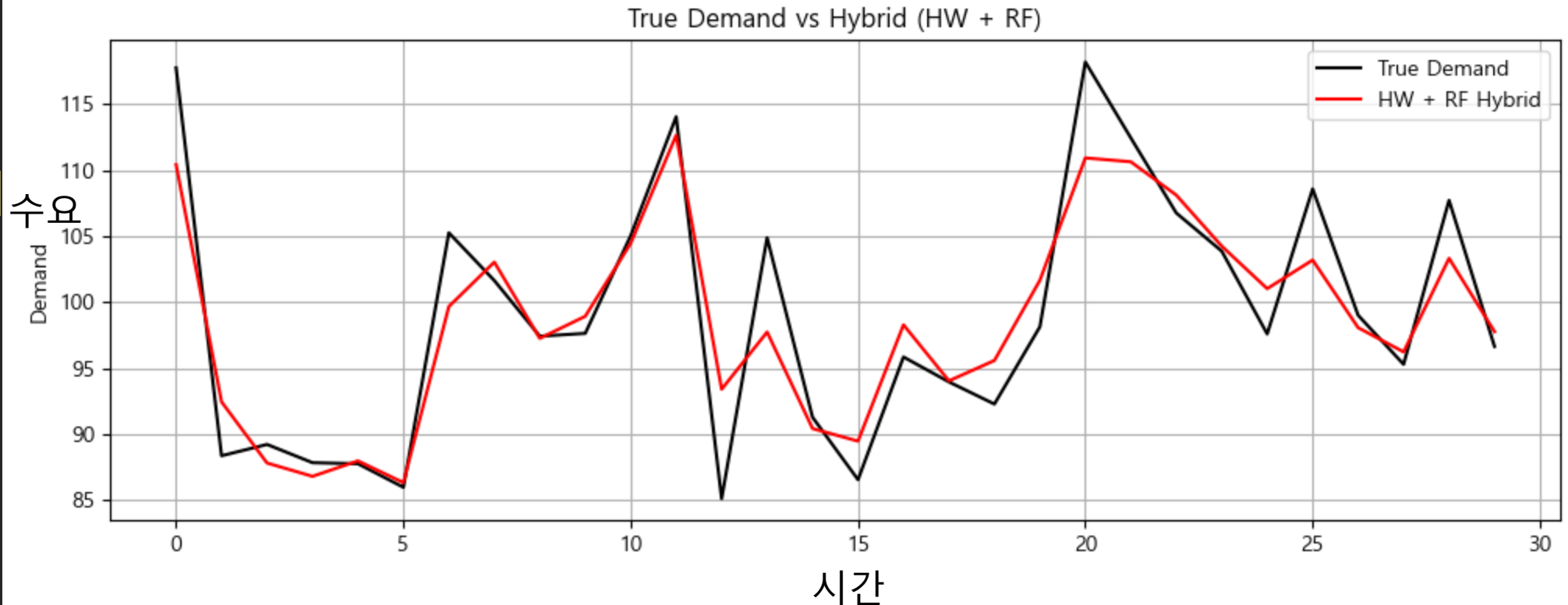
5. 하이브리드 예측 기법

5. 하이브리드 예측 기법

5.2. 하이브리드 예측 기법 사용 예시

- 하이브리드 모델

- Holt-Winters가 학습하지 못한 오차를 Random Forest로 학습시키는 방법
- 이전보다 모델의 예측력이 훨씬 좋아짐



5. 하이브리드 예측 기법

5.2. 하이브리드 예측 기법 사용 예시

• 모델 성능 비교

▪ MSE(Mean Squared Error)

✓ Holt-winters : 56.91

✓ Random Forest : 99.60

✓ Hybrid model : 13.00

▪ MAE(Mean Absolute Error)

✓ Holt-winters : 5.91

✓ Random Forest : 8.42

✓ Hybrid model : 2.68

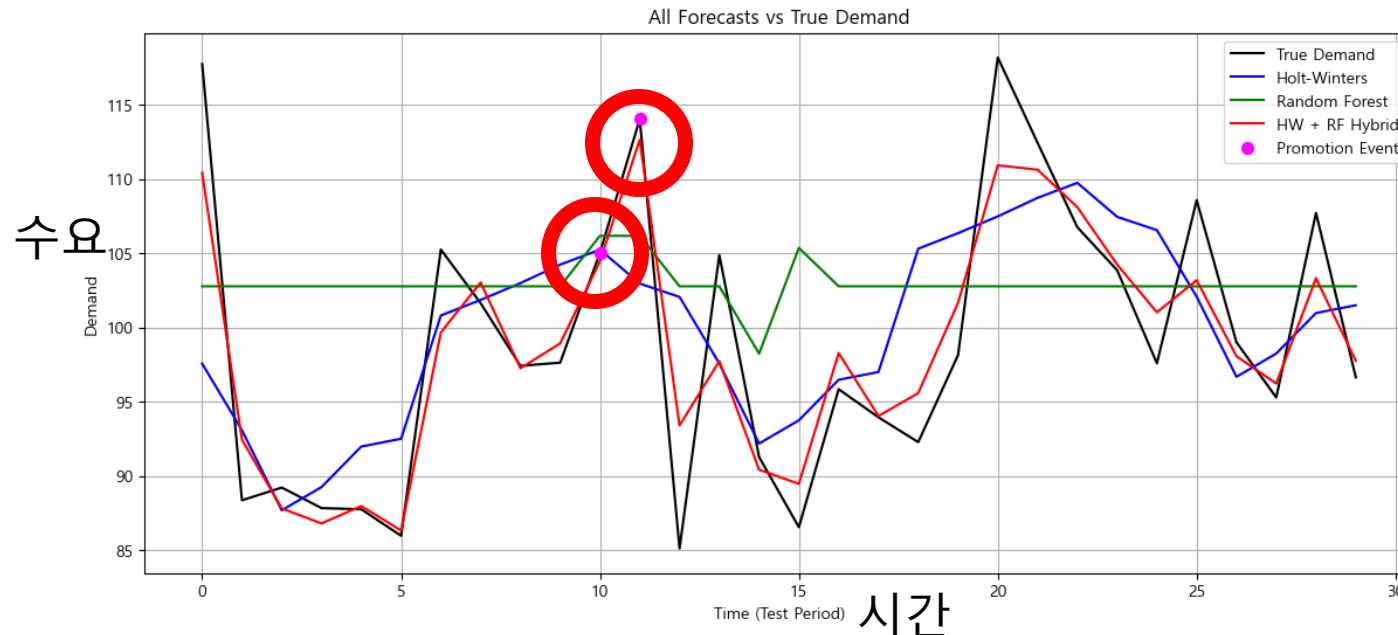
▪ MAPE(Mean Absolute Percentage Error)

✓ Holt-winters : 0.06

✓ Random Forest : 0.08

✓ Hybrid model : 0.03

▪ Promotion이 있을 때도 Hybrid모델이 가장 성능이 좋음



1. 수요 예측 개요

2. 전통적 통계 기반
예측 기법

3. ARIMA 계열 예측 기법

4. 머신러닝 예측 기법

5. 하이브리드 예측 기법

5. 하이브리드 예측 기법

5.2. 하이브리드 예측 기법 사용 예시

- 결론

- 통계 모델과 머신러닝 모델을 합친 하이브리드 모델이 단일 모델보다 성능이 좋음
- 시계열 이외에 여러가지 변수들까지 학습에 활용 가능
- 프로모션 등의 이벤트가 있을 때의 예측력을 높일 수 있어 기업의 의사결정에 도움을 줄 수 있음

Thank you for your attention.



Professor ILKYEONG MOON
Department of Industrial Engineering
Seoul National University
Seoul 08826, KOREA
ikmoon@snu.ac.kr
<http://scm.snu.ac.kr>