

기상 조건 기반
열수요 예측 정확도 향상을 위한
다중 모델 앙상블 시스템 연구

정환웅, 배성민

목차

Chapter 1

배경 및 분석 목표

Chapter 2

데이터 특징

Chapter 3

EDA 분석

Chapter 4

특성공학

Chapter 5

모델링 및 검증 결과

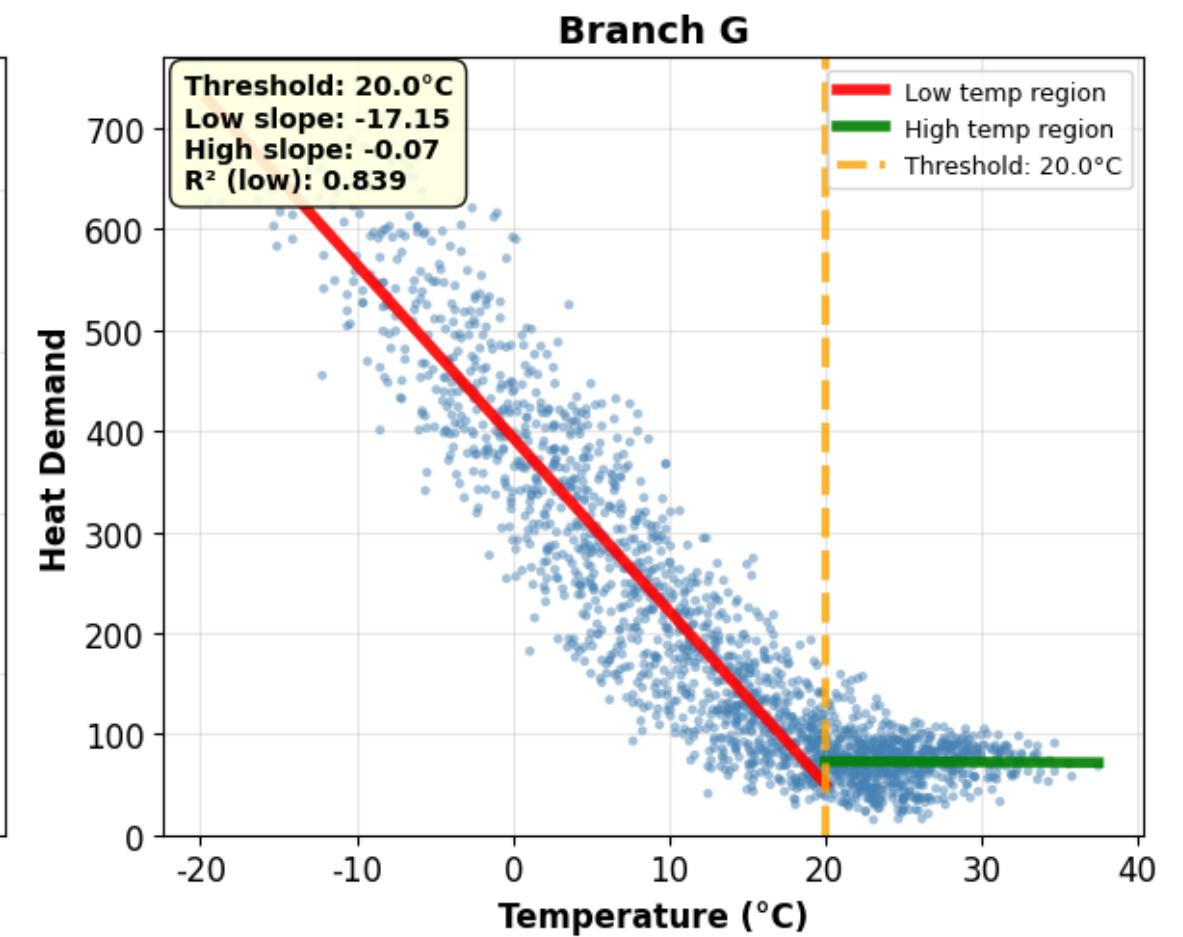
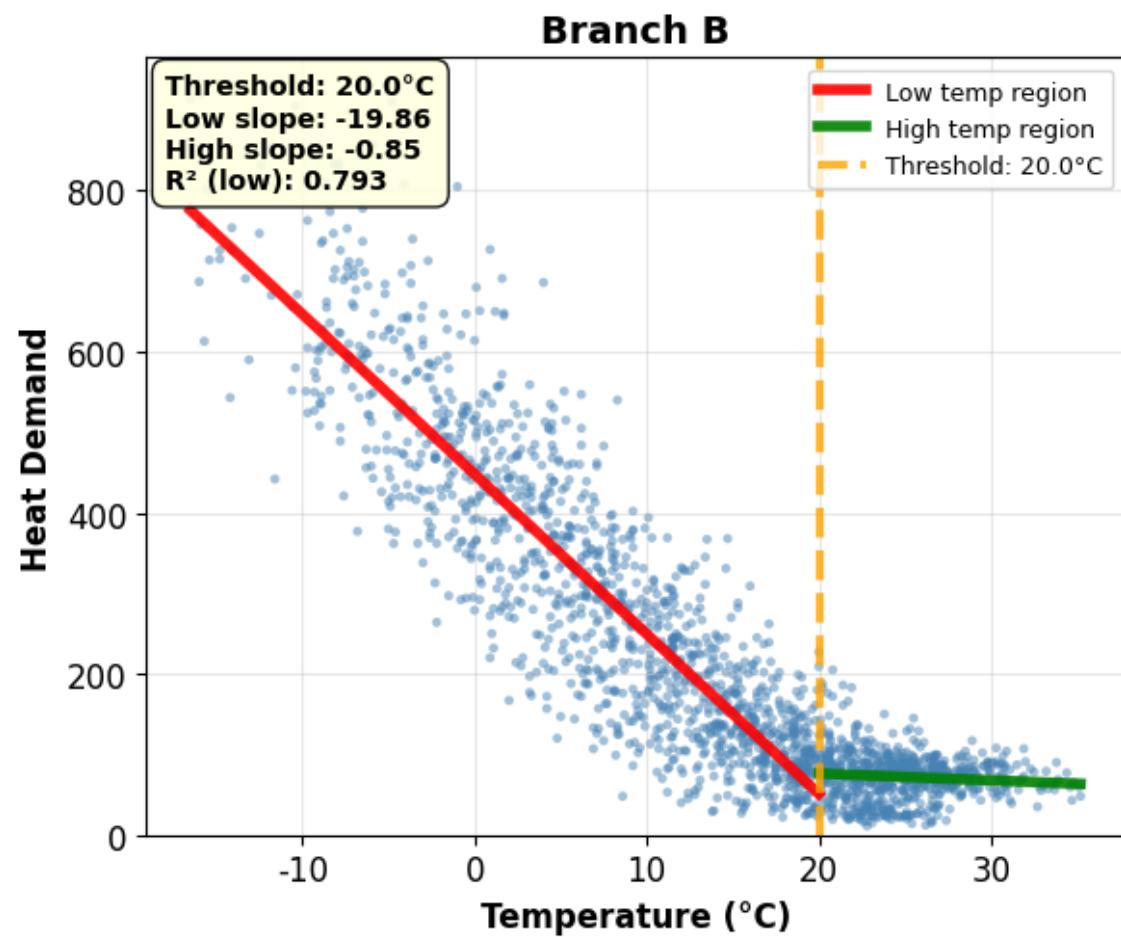
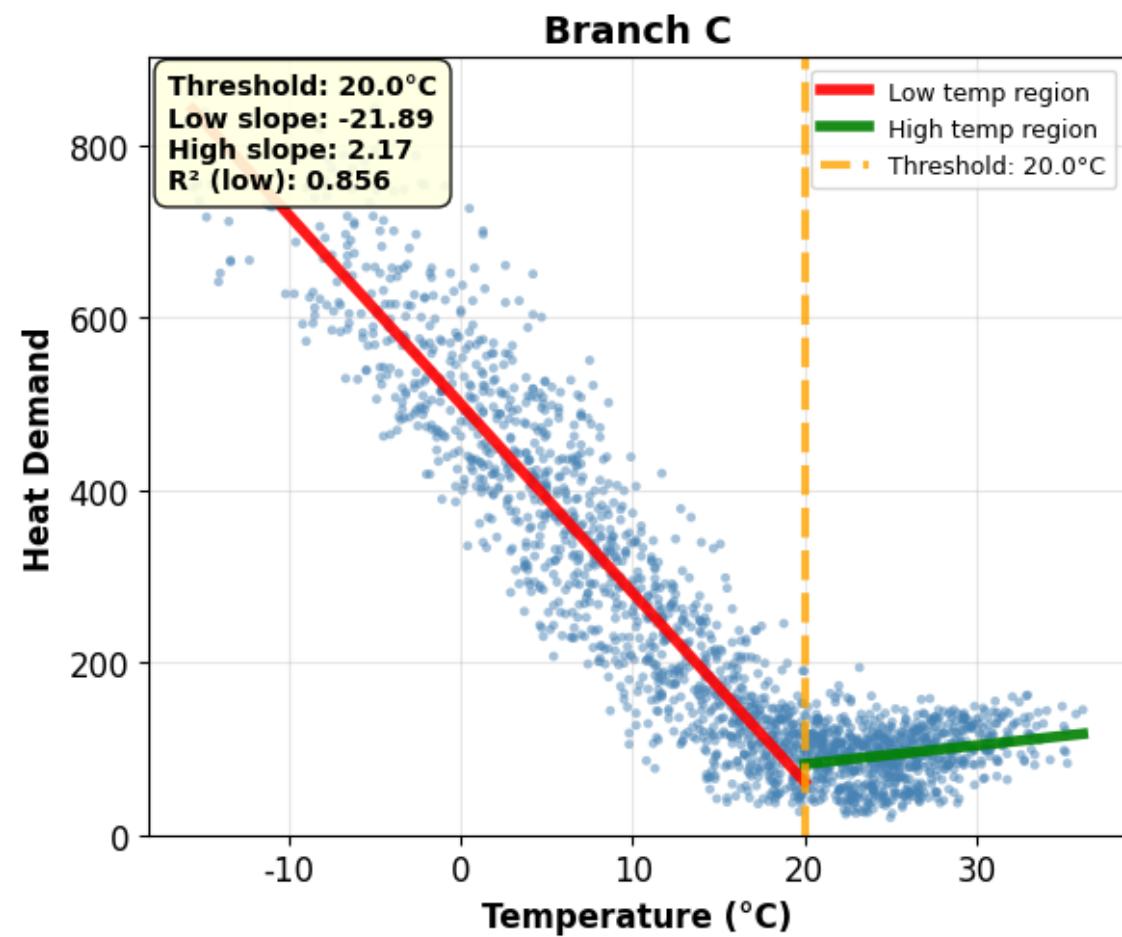
Chapter 6

활용방안

정확한 예측, 효율적 난방, 지속가능한 미래

- 지역난방 시스템의 효율적 운영은 에너지 소비 최적화와 탄소 배출 저감이라는 지속 가능한 에너지 정책 목표 달성에 필수
- 특히 도시화의 가속화와 함께 난방 시스템에 대한 수요가 급증, 정확한 열 수요 예측의 중요성 증가
- 이는 외기온도를 비롯한 기상변수와 사회 문화적 요소를 이해하여 이력현상으로 인한 예측 저해 타파 및 적확한 예측 필요

L-shaped Heat Demand Relationships by Branch



지체현상

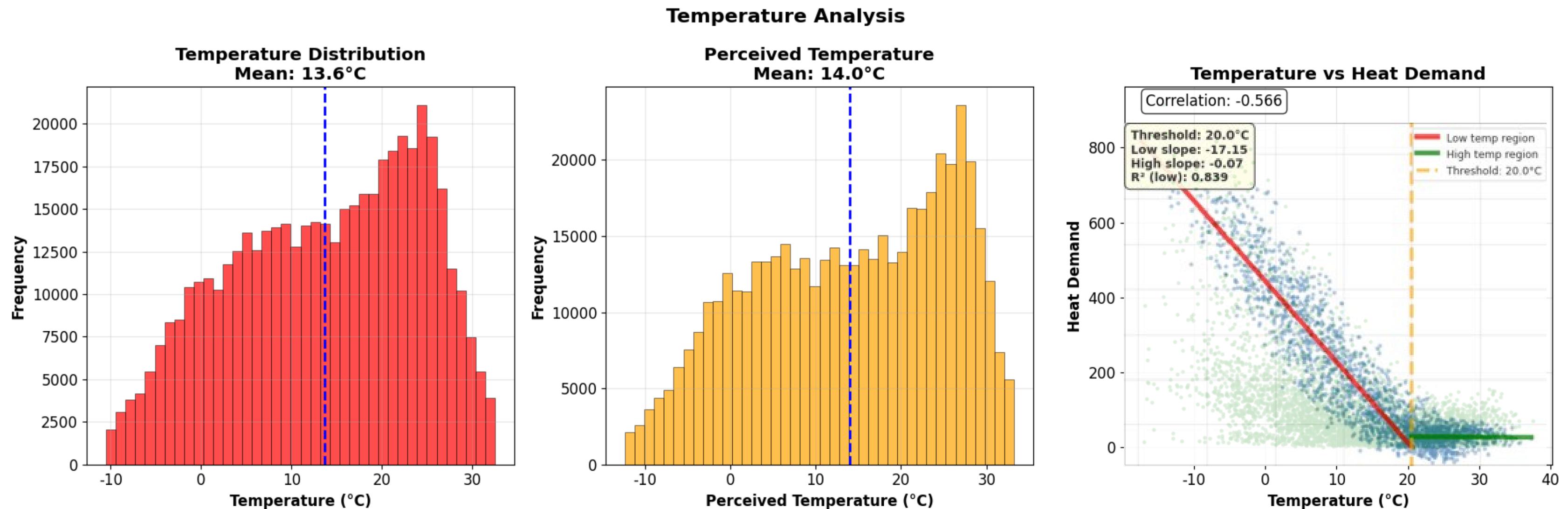
열 에너지는 사용 후 즉시 휘발되지 않고 일시적으로 저장되어 시간차를 두고 영향을 미치는 현상

이력현상

난방 시스템의 반응이 현재 온도뿐만 아니라 과거의 온도 변화 이력에도 영향을 받는 현상

D20-16

일평균 외기온도가 16°C 이하일 때 실내온도를 20°C 로 유지하기 위해 난방이 필요하다는 기준



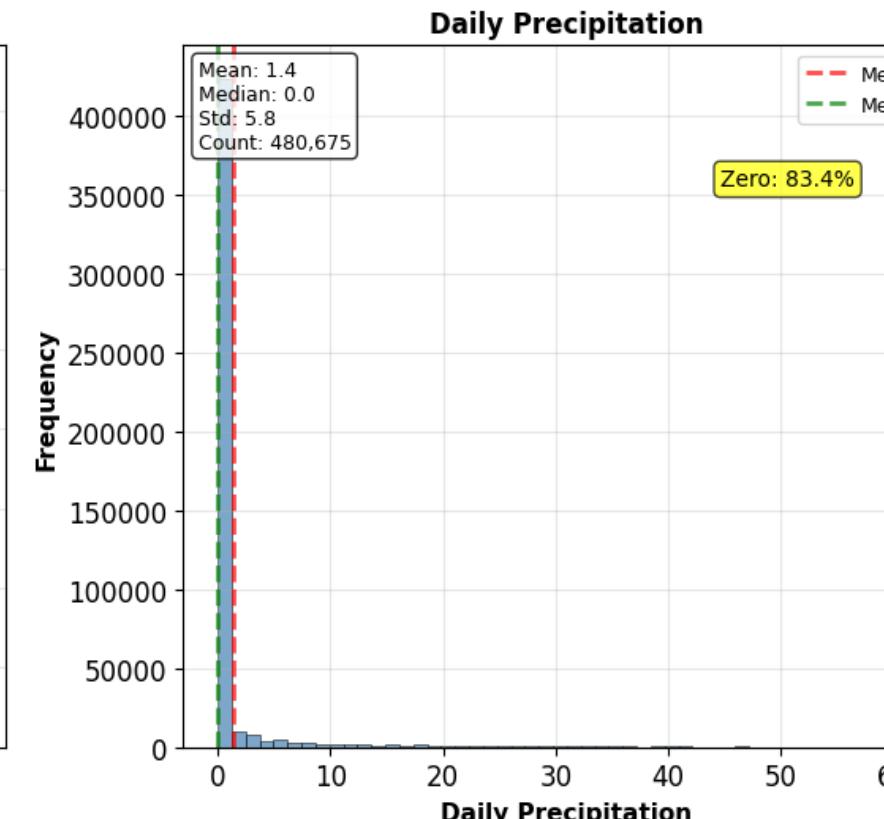
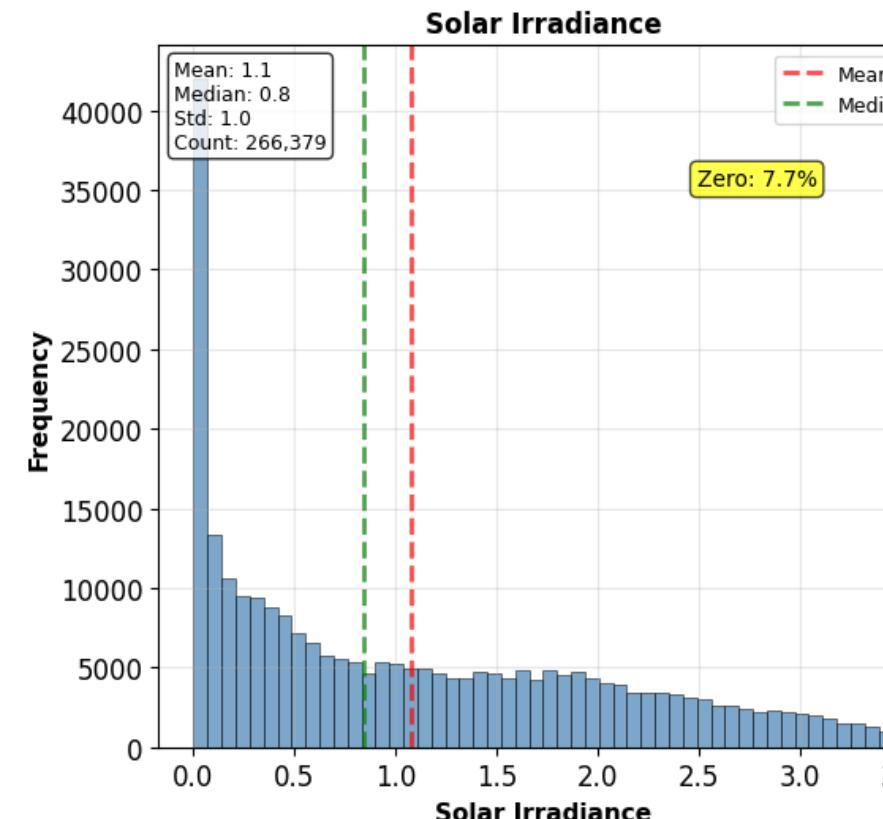
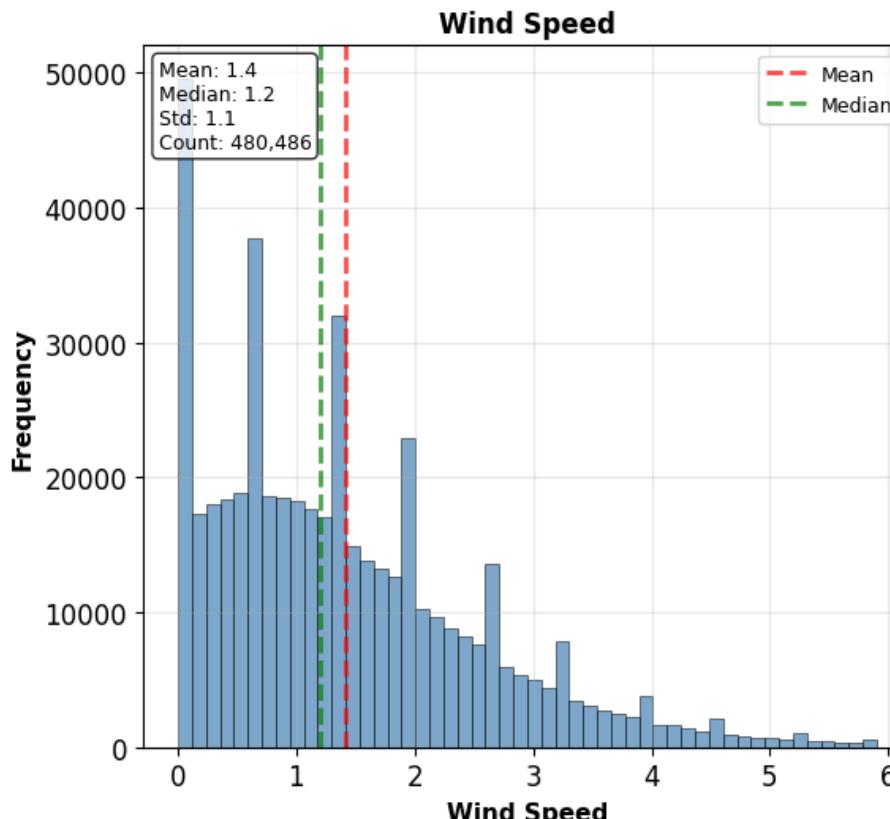
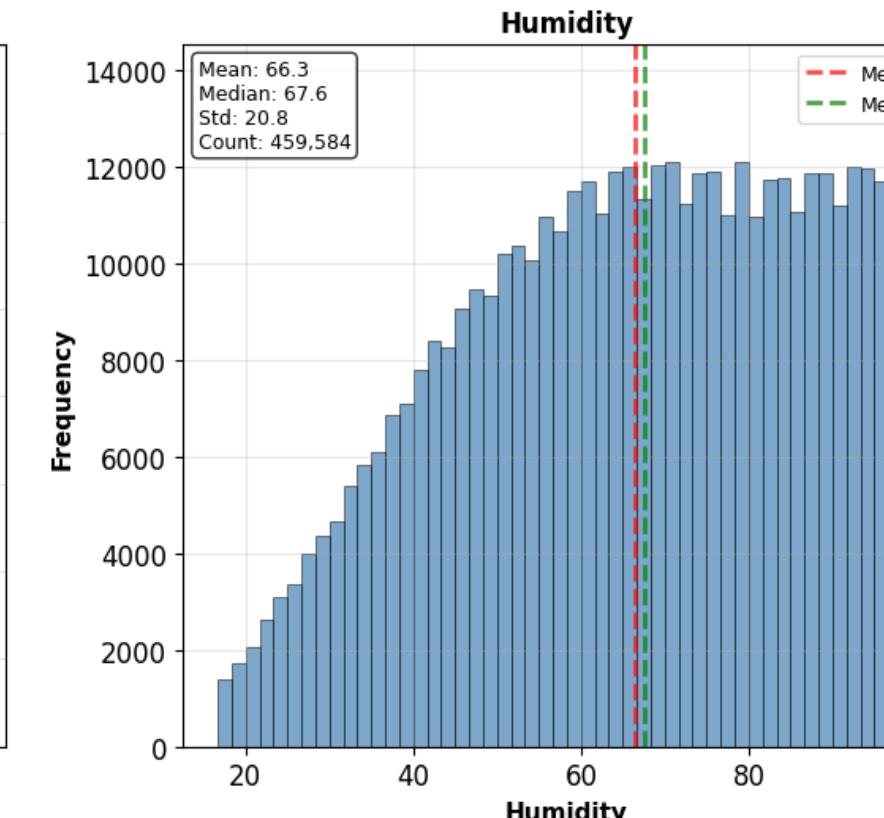
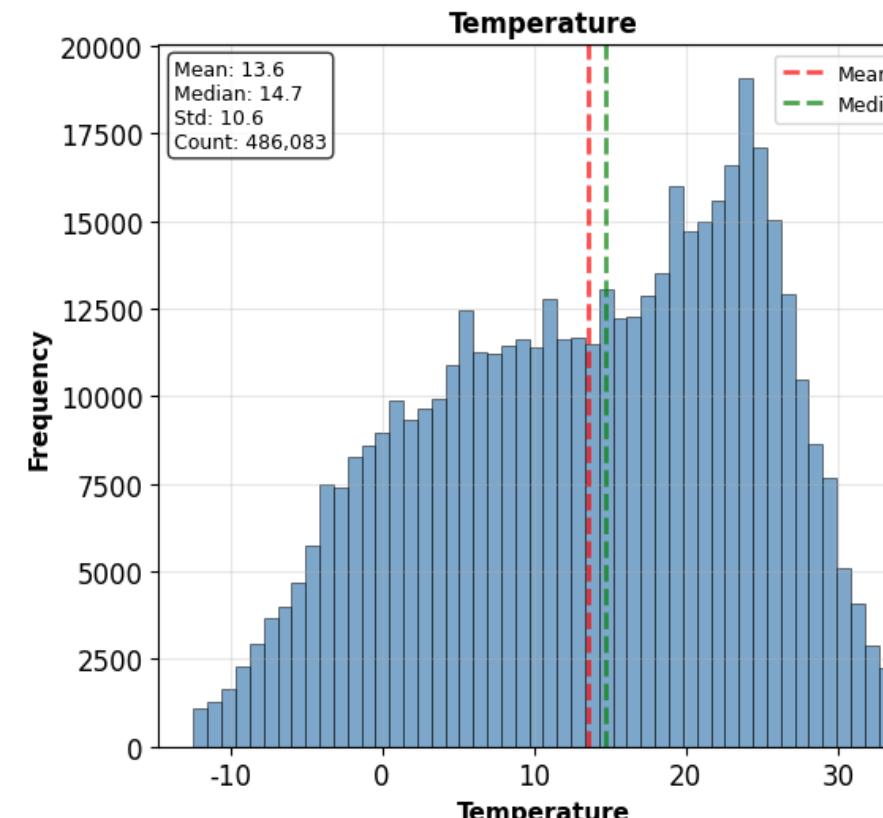
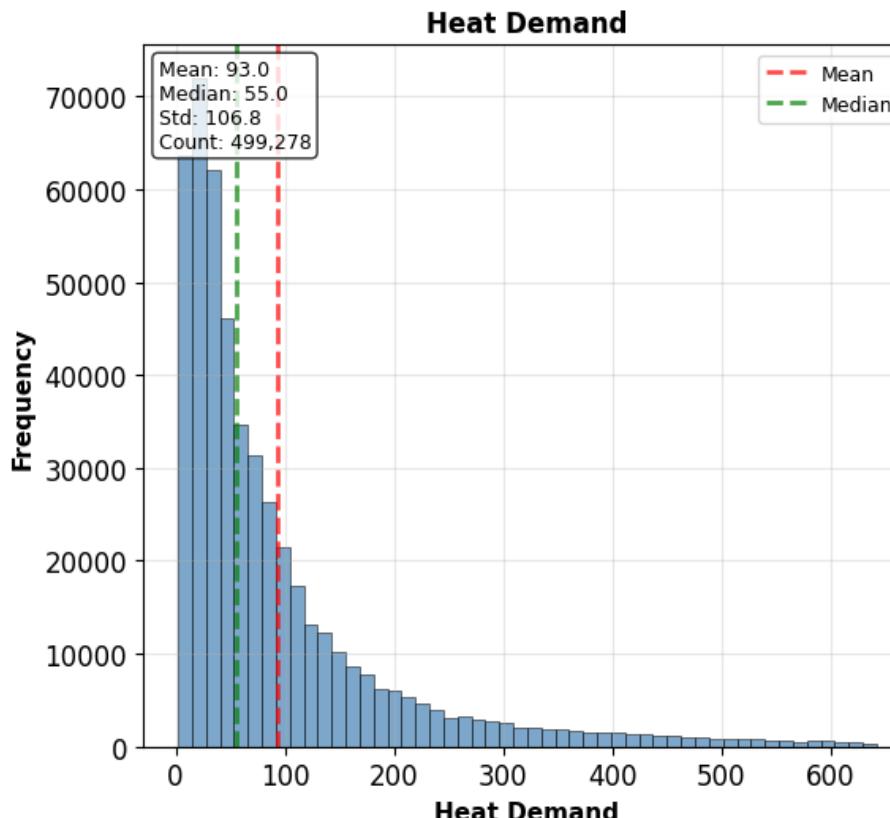
외기온도와 체감온도

전체적인 분포는 동일, 평균은 소폭 상이
체감온도 기반 파생 특성이 외기온도 변수의 정보 확장·보완

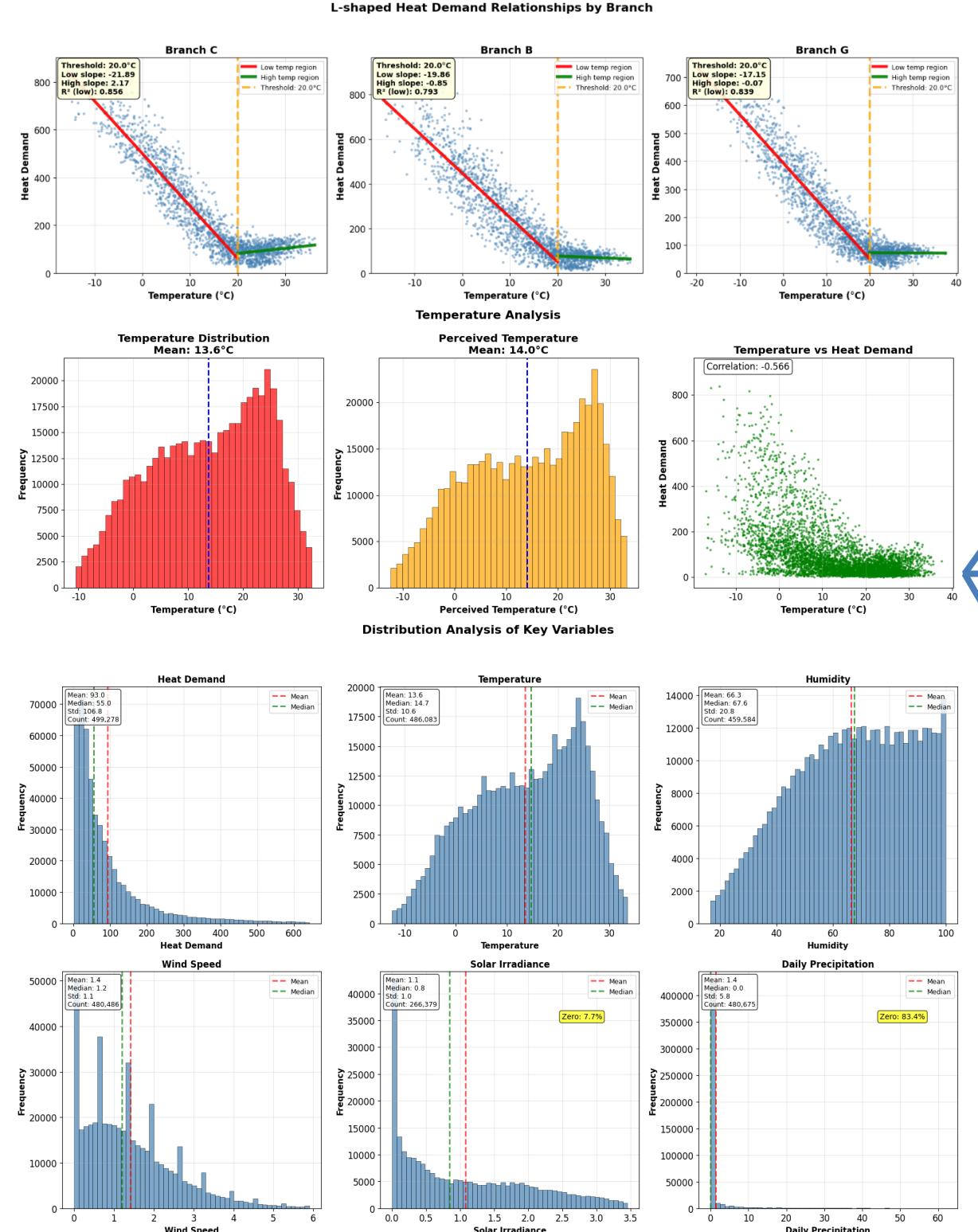
피어슨 상관계수

음의 선형 경향 및 명확한 절곡점
L자형 수요 반응 패턴과 일치

Distribution Analysis of Key Variables



- 열수요: 비대칭 분포
- 외기온도: 이중봉 구조
- 상대습도: 고른 분포
- 풍속: 지수분포
- 일사량: 이원구조
- 일 강수량: 간헐성 분포



- 온도 변화량, 가속도 등을 활용한 특성과 1시간부터 최대 72시간까지 다양한 자연 변수와 롤링 통계 특성을 생성하여 이력현상 포착에 용이
- 체감온도 기반, 계절적 특성을 명확히 반영하기 위하여 여름철(5월~9월)에는 습도의 영향을 크게 반영하는 열지수(heat index)를 사용하고, 겨울철(10월~4월)에는 바람의 냉각 효과를 반영하는 윈드칠(wind chill) 특성 생성
- 각 지역별 branch 유사한 기상 조건을 군집화하기 위하여 K-means 클러스터링 실시, 인간의 일주기 리듬을 반영한 추가 특성 개발, 이 외에도 지점별 열수요와 기상 데이터의 통계, 난방도일과 같은 전문 도메인 지식을 반영한 파생 변수 생성

#지연변수

#체감온도

#난방도일

데이터 특징 및 EDA 분석

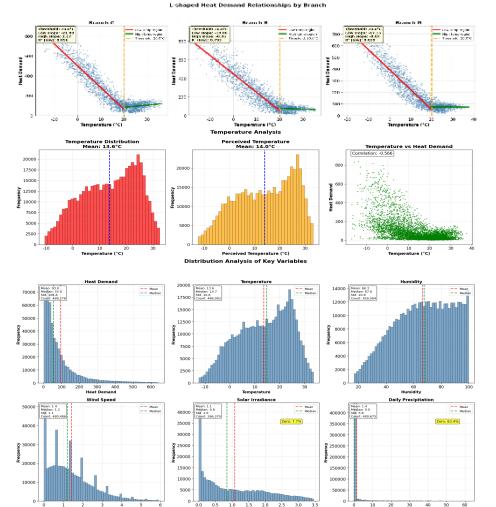


트리 기반 양상별



특성공학

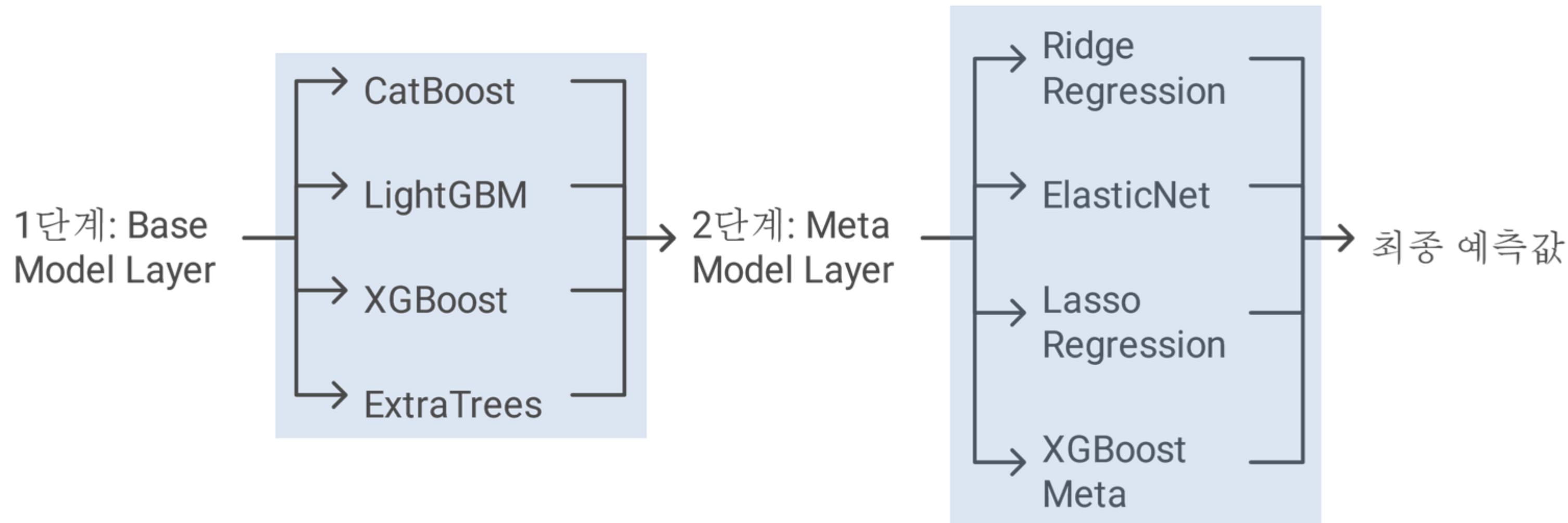
- 지체현상, 이력현상, D20-16, 비선형적 관계, L자형 수요패턴

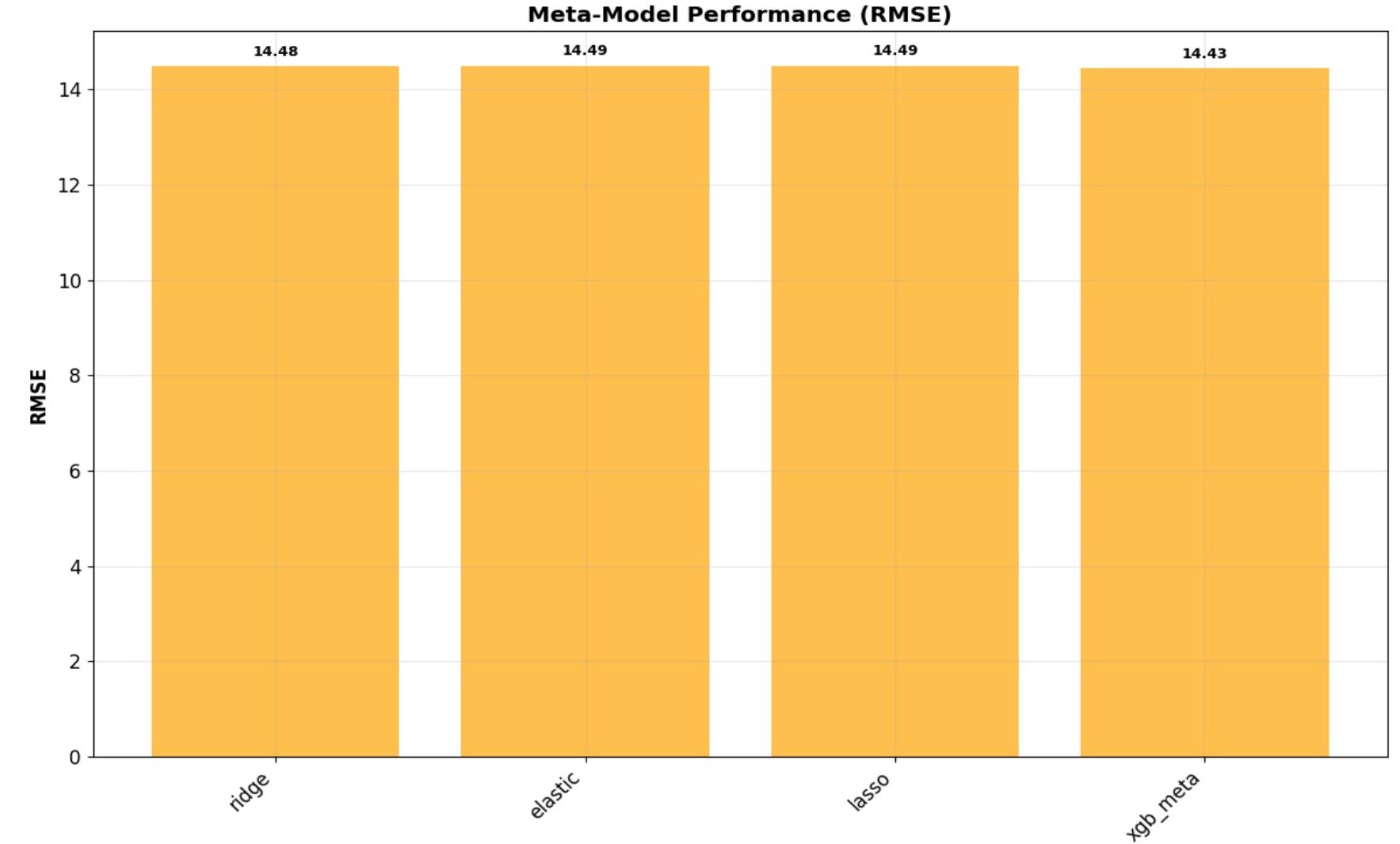
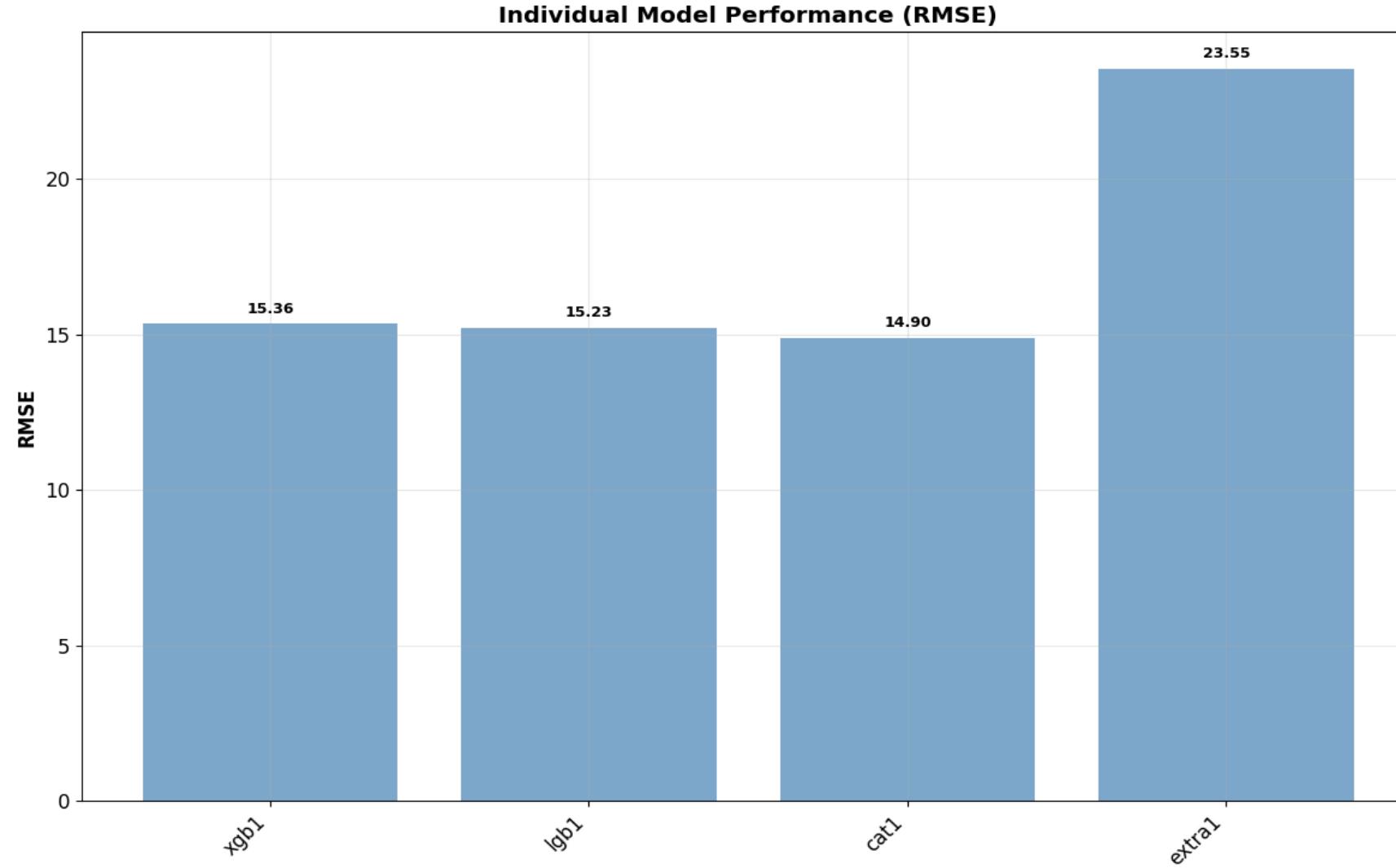


- 비선형 관계 자동 포착
(L자형 패턴을 자연스럽게 학습 가능)
- 특성 간 상호작용 학습
(기상 변수간 복잡한 상호작용을 명시적 정의 없이 포착)
- 강건성과 안정성
(과적합 위험 크게 감소)
- 실무적 해석가능
(특성 중요도 산출로 근거 설명 가능)

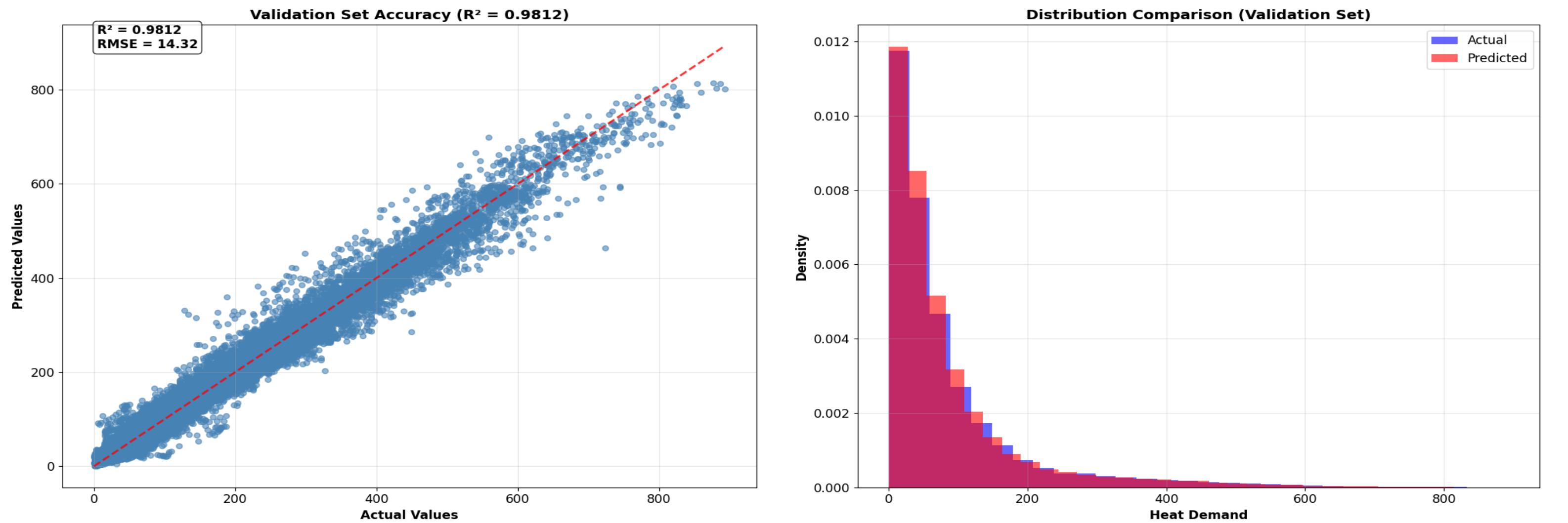
- 지연변수 및 롤링 통계 특성
- 체감온도 기반 특성
- 기상 군집화
- 계절적 특성
- 인문사회적 특성 (일주기, 주중, 주말)
- 도메인 지식 반영한 특성

모델 구조도

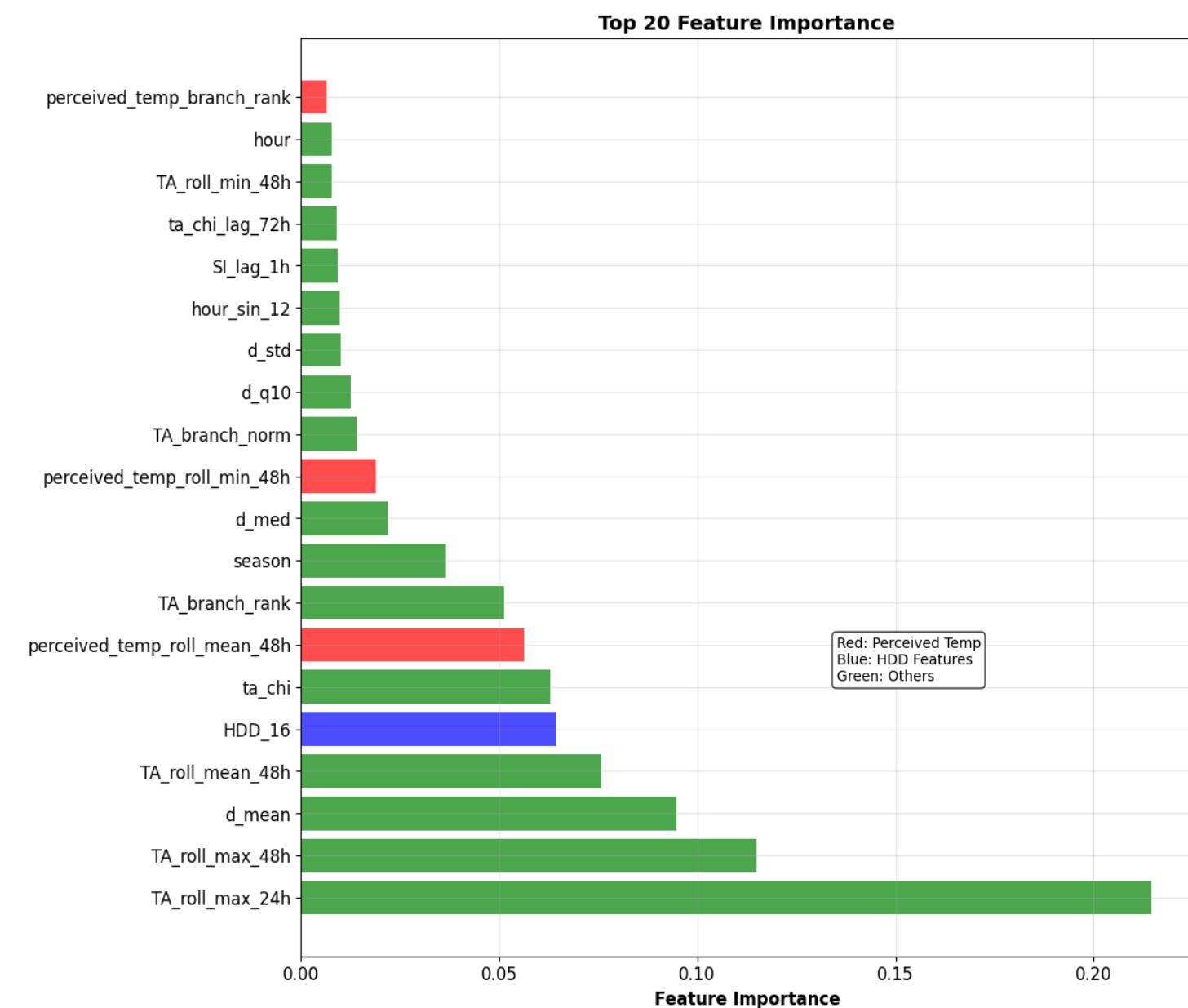
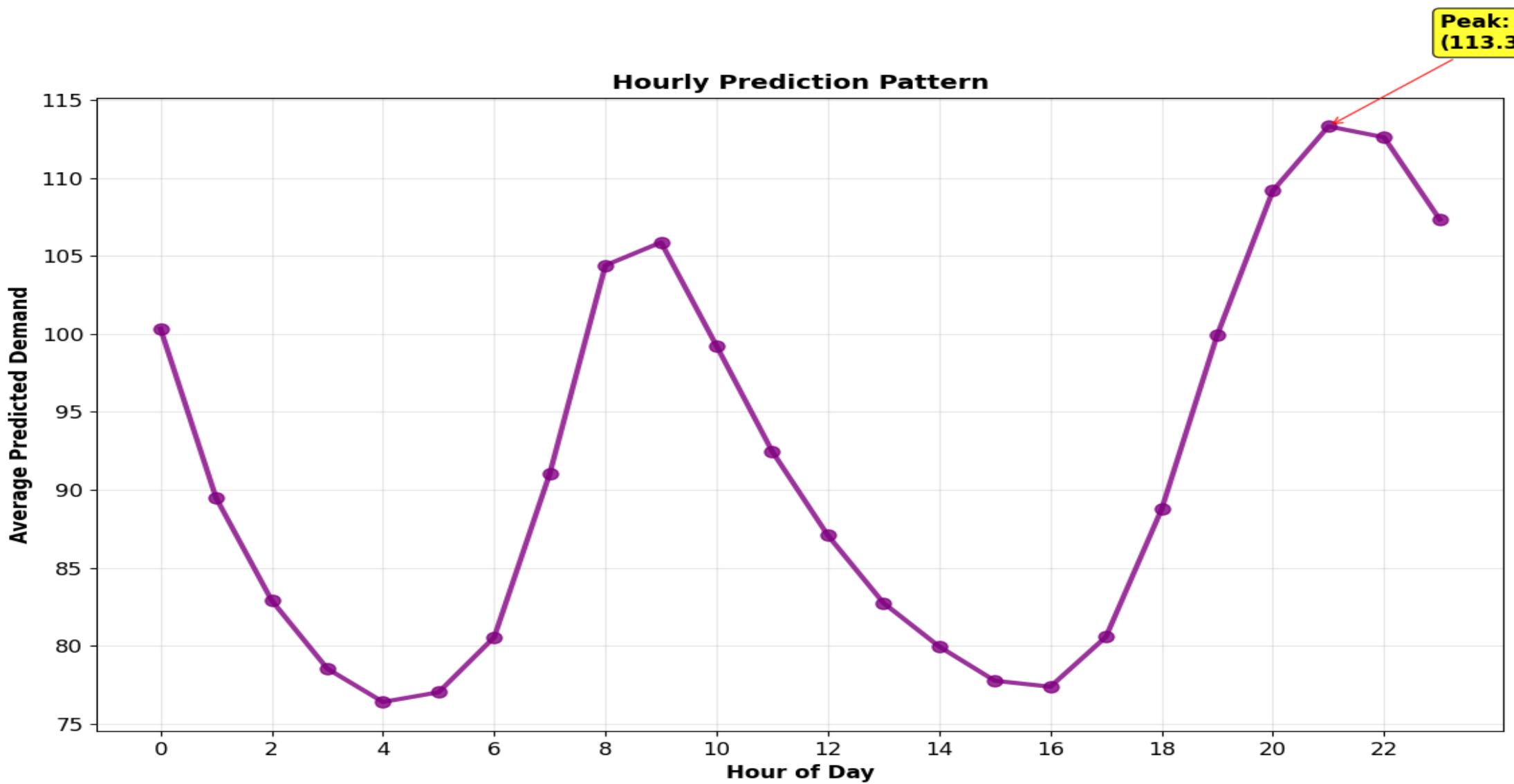




- **CatBoost(14.90), LightGBM(15.23), XGBoost(15.36), ExtraTrees(23.55)**
- 4개의 메타모델(Ridge, ElasticNet, Lasso, XGBoost Meta)은 모두 유사한 성능을 보여 균등 가중치로 최종 결합
- 최종 메타 양상블은 14.32 RMSE를 달성하여 단일 최고 모델(CatBoost) 대비 약 4%의 성능 개선
- 실제 테스트 세트에 대한 최종 검증 평가에서 **RMSE 14.13**을 기록하여, 학습/검증 단계의 14.32 대비 1.3%의 추가 성능 향상



- 시각화 분석을 통한 모델 해석 결과, 실제값 대비 예측값 산점도는 거의 완벽한 대각선 분포를 보여주며 모든 수요 범위에서 일관된 예측 정확도를 보임
- 특히 저수요(0-200) 구간부터 고수요(600-800) 구간까지 편향 없는 예측 성능을 보여, 실제 운영에서 발생할 수 있는 다양한 수요 상황에 대한 신뢰할 만한 대응이 가능함을 시사



- 예측 패턴은 실제 인간의 생활 리듬과 높은 일치성을 보임
- 새벽 3-4시 최저점은 최소 활동 시간대를 반영하며, 아침(7-9시)과 저녁(18-21시)시간대의 이중 정점 구조는 현대 도시 생활의 전형적인 에너지 사용 패턴을 나타냈고 이러한 패턴의 정확한 재현은 모델이 단순한 수치적 관계를 넘어 인문 사회적 패턴을 정확히 포착하고 있다는 것을 알 수 있음

열수요 결정 및 분석 요인

원인

- 여름에는 외기온도와 체감온도가 높고 풍속이 낮은 반면, 겨울에는 외기온도와 체감온도가 낮고 풍속이 높음
- 열수요는 시계열적 특성인 이력현상이 존재하여 현재의 기온과 이전 온도 변화 추이에 영향을 받음

결과

- 지역, 계절별 고유한 열수요 특성이 나타남
- 특정 절곡점 이후에 완만한 열수요 패턴 존재

에너지 효율화와 탄소중립을 위한 실무·정책 활용 전략

항목

세부 내용

설비 운영 최적화

시간대별 예측 패턴의 정확한 재현을 통해 피크 시간대 대비 및 최저 수요 시간대 활용이 가능하여, 설비 운영 스케줄 최적화와
유지, 보수 계획 수립 활용

또한, 모델의 강건한 일반화 성능은 계절, 날씨별 다양한 조건에서도 안정적인 예측을 제공하여 예방적 관리 체계 구축 가능

예측 시스템 구축

환경적 측면에서는 효율적 운영을 통한 탄소 배출 8-12% 감소 효과를 기대할 수 있으며,
과적합이 없는 강건한 모델 특성을 활용하여 다양한 지역과 건물 유형에 적용할 수 있는 범용 예측 시스템 구축이 가능

교육 프로그램 개발

전국 단위의 지역난방 효율화 정책 수립에 활용하여 모델의 해석 가능성을 바탕으로 에너지 효율 개선 지점을 명확히 식별할 수 있으며,
사용자 행동 변화 유도와 에너지 절약 교육 프로그램 개발에도 활용

감사합니다.