



스마트 온실 환경 데이터 인공지능 예측 및 시각화 기술 개발

팀 리트

컴퓨터공학부 김채훈

컴퓨터공학부 김광일

컴퓨터공학부 이주형



전북대학교
SW중심대학사업단



농촌진흥청
국립농업과학원

목차

01 서론

02 연구 과정

03 연구 결과

04 결론 및 제언

목차

01 서론

02 연구 과정

03 연구 결과

04 결론 및 제언

서론

서론

| 연구 과정

| 연구 결과

| 결론 및 제언

현황

01

작물의 수확량과 노동력 감소를 위해 확대되는 스마트 온실

02

스마트 온실은 구동기 설정 편차에 따른 큰 수확량 차이

03

초보, 청년농업인에게는 구동기 설정이 큰 진입장벽

과제 제안

04

농장주의 경험에 의존하는 의사결정을 보완하기 위해 인공지능을 활용해 스마트 온실의 완전 자동화를 실현

리트의 목표

05

선도 농가의 데이터를 분석하여 **최적의 내부온도를 예측**하고, 이를 유지하기 위한 **구동기 설정값을 도출**

06

예측값과 환경 데이터를 시각적으로 표현해 농장주가 더 쉽게 이해하고 활용이 가능



서론

서론

연구 과정

연구 결과

결론 및 제언

현황

01

작물의 수확량과 노동력 감소를 위해 확대되는 스마트 온실

02

스마트 온실은 구동기 설정 편차에 따른 큰 수확량 차이

03

초보, 청년농업인에게는 구동기 설정이 큰 진입장벽

과제 제안

04

농장주의 경험에 의존하는 의사결정을 보완하기 위해 인공지능을 활용해 스마트 온실의 완전 자동화를 실현

리트의 목표

05

선도 농가의 데이터를 분석하여 **최적의 내부온도를 예측**하고, 이를 유지하기 위한 **구동기 설정값을 도출**

06

예측값과 환경 데이터를 시각적으로 표현해 농장주가 더 쉽게 이해하고 활용이 가능



서론

서론

| 연구 과정

| 연구 결과

| 결론 및 제언

현황

01

작물의 수확량과 노동력 감소를 위해 확대되는 스마트 온실

02

스마트 온실은 구동기 설정 편차에 따른 큰 수확량 차이

03

초보, 청년농업인에게는 구동기 설정이 큰 진입장벽

과제 제안

04

농장주의 경험에 의존하는 의사결정을 보완하기 위해 인공지능을 활용해 스마트 온실의 완전 자동화를 실현

리트의 목표

05

선도 농가의 데이터를 분석하여 **최적의 내부온도를 예측**하고, 이를 유지하기 위한 **구동기 설정값을 도출**

06

예측값과 환경 데이터를 시각적으로 표현해 농장주가 더 쉽게 이해하고 활용이 가능



서론

서론

| 연구 과정

| 연구 결과

| 결론 및 제언

현황

01

작물의 수확량과 노동력 감소를 위해 확대되는 스마트 온실

02

스마트 온실은 구동기 설정 편차에 따른 큰 수확량 차이

03

초보, 청년농업인에게는 구동기 설정이 큰 진입장벽

과제 제안

04

농장주의 경험에 의존하는 의사결정을 보완하기 위해 인공지능을 활용해 스마트 온실의 완전 자동화를 실현

리트의 목표

05

선도 농가의 데이터를 분석하여 **최적의 내부온도를 예측**하고, 이를 유지하기 위한 **구동기 설정값을 도출**

06

예측값과 환경 데이터를 시각적으로 표현해 농장주가 더 쉽게 이해하고 활용이 가능



목차

01 서론

02 연구 과정

03 연구 결과

04 결론 및 제언

연구 과정

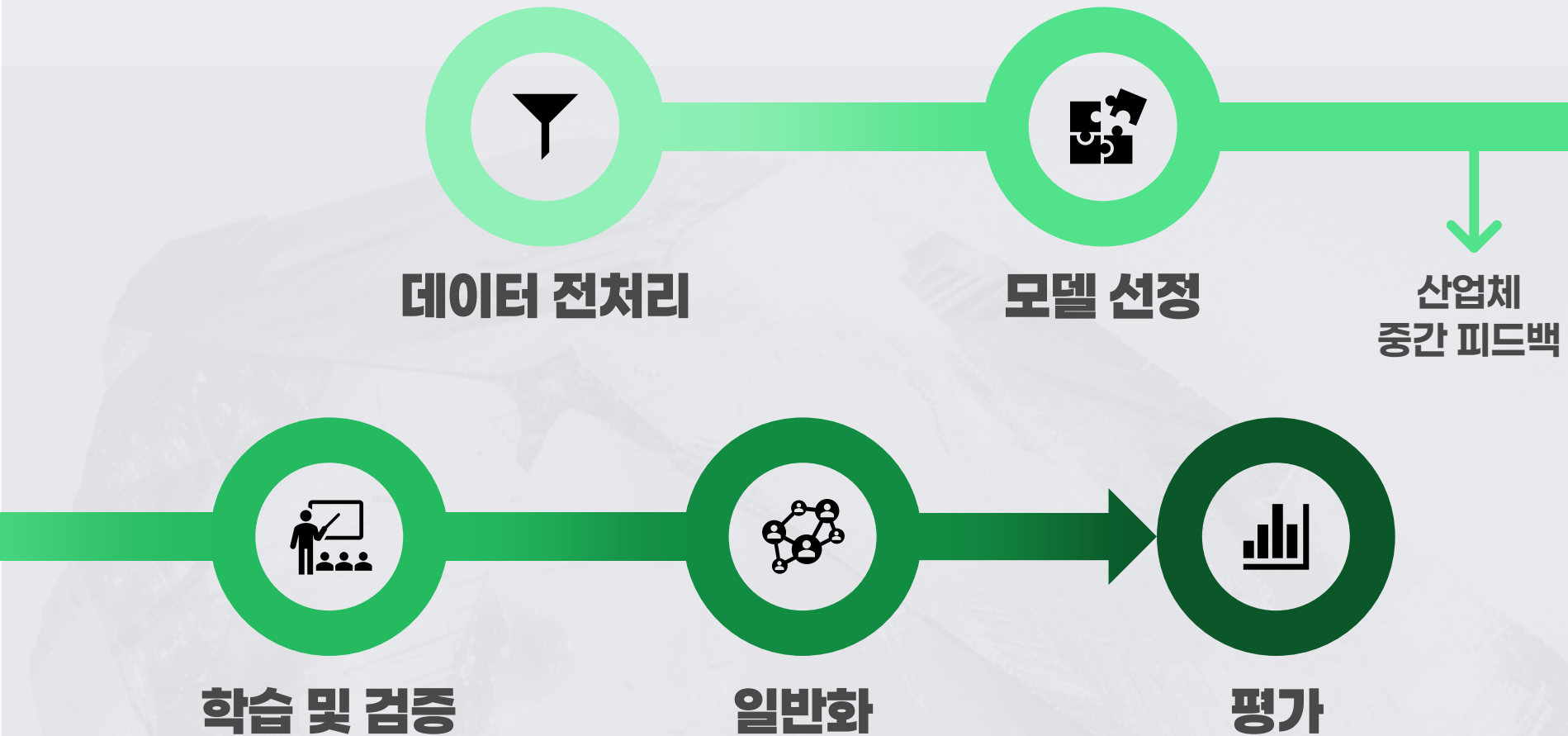
서론

연구 과정

연구 결과

결론 및 제언

연구 목표. 스마트 온실 환경 데이터의 인공지능 예측 제어 모델 개발



연구 과정

서론 | 연구 과정 | 연구 결과 | 결론 및 제언

학습 환경 및 활용 라이브러리



연구 과정

서론

연구 과정

연구 결과

결론 및 제언

데이터 출처 및 특성



토마토 선도농가 A,B,C,D의
데이터셋을 제공받아 사용
각 1분 단위로 5년간 기록된
83개 칼럼의 시계열 데이터
(1개년 데이터는 약 53만건)

- ✓ 내부 데이터 내부온도, 내부습도
- ✓ 제어 데이터 적응난방온도, 적응환기온도
- ✓ 외부 데이터 외부온도, 풍향, 풍속, 일사량, 누적일사량, 감우
- ✓ 설정 데이터 제어환기온도, 제어난방온도, 환기조절 영향값, 난방조절 영향값
- ✓ 구동기데이터 공급온도, 천창, 커튼(상/하/측/외부), 3Way 밸브개도, 냉난방 작동상태, 보일러 작동상태, 순환펌프 작동상태, 유동팬 작동상태, 배기팬 작동상태

최초 미팅 시 전문가의 조언에 따라 83개 칼럼 중 필수 환경 데이터 선정

데이터 전처리

01

데이터 인코딩

- | 구동기 상태 등 범주형 데이터를 수치형 벡터로 변환
- | 원핫 인코딩을 통해 각 범주를 독립된 차원으로 표현

02

이상치 처리, 결측치 보간

- | 차트 시각화를 통해 이상치 확인 후 불필요한 이상치를 제거
- | 결측값을 이전 값과 이후 값을 기반으로 선형보간법을 사용

03

데이터 정규화

- | Min-Max 스케일링 기법을 사용해 모든 피처의 값이 (0~1)내에 있도록 조정
- | 특정 변수의 스케일 차이로 인한 왜곡을 방지해 모델의 안정성을 향상

04

데이터셋 분할

- | 전체 데이터를 학습, 검증, 테스트 데이터셋으로 나누어 모델 학습 및 평가 준비
- | 다양한 분할 방법을 실험한 결과 데이터를 연도에 따라 6:2:2 비율로 분할

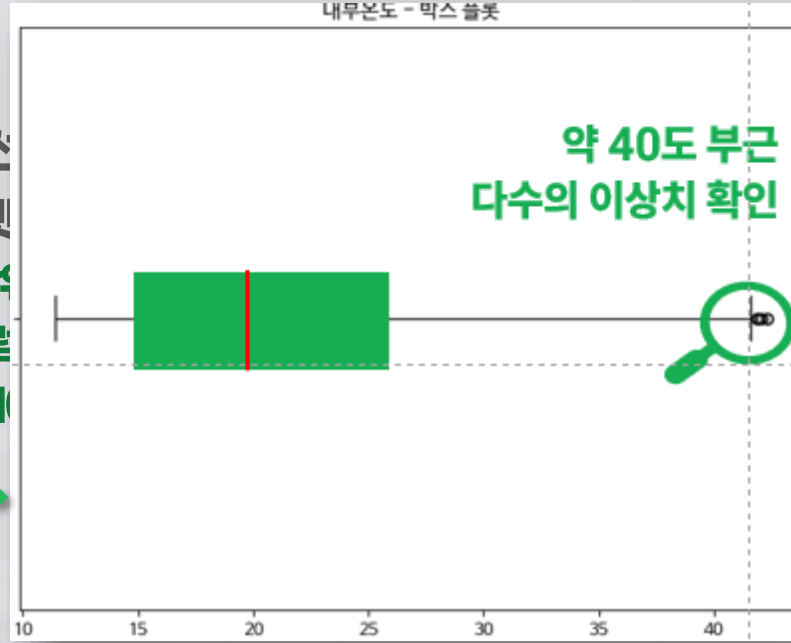
연구 과정

서론 | **연구 과정** | 연구 결과 | 결론 및 제언

데이터 출처 및 특성



토마토 산
데이터셋
각 1분 단위
83개 칼럼
(1개년 데이터)



내부온도, 내부습도
적용난방온도, 적용환기온도
외부온도, 풍향, 풍속, 일사량, 누적일사량, 감우
제어환기온도, 제어난방온도, 환기조절 영향값, 난방조절 영향값
공급온도, 천창, 커튼(상/하/측/외부), **3Way 밸브개도**, 냉난방 작동상태,
보일러 작동상태, 순환펌프 작동상태, 유동팬 작동상태, 배기팬 작동상태

연에 따라 83개 칼럼 중 필수 환경 데이터 선정

데이터 전처리

01

데이터 인코딩

- | 구동기 상태 등 범주형 데이터를 수치형 벡터로 변환
- | 원핫 인코딩을 통해 각 범주를 독립된 차원으로 표현

02

이상치 처리, 결측치 보간

- | 차트 시각화를 통해 이상치 확인 후 불필요한 이상치를 제거
- | 결측값을 이전 값과 이후 값을 기반으로 선형보간법을 사용

03

데이터 정규화

- | Min-Max 스케일링 기법을 사용해 모든 피쳐의 값이 (0~1)내에 있도록 조정
- | 특정 변수의 스케일 차이로 인한 왜곡을 방지해 모델의 안정성을 향상

04

데이터셋 분할

- | 전체 데이터를 학습, 검증, 테스트 데이터셋으로 나누어 모델 학습 및 평가 준비
- | 다양한 분할 방법을 실험한 결과 데이터를 연도에 따라 6:2:2 비율로 분할

연구 과정

서론

연구 과정

연구 결과

결론 및 제언

데이터 출처 및 특성



토마토 선도농가 A,B,C,D의
데이터셋을 제공받아 사용
각 1분 단위로 5년간 기록된
83개 칼럼의 시계열 데이터
(1개년 데이터는 약 53만건)

- ✓ 내부 데이터 내부온도, 내부습도
- ✓ 제어 데이터 적응난방온도, 적응환기온도
- ✓ 외부 데이터 외부온도, 풍향, 풍속, 일사량, 누적일사량, 감우
- ✓ 설정 데이터 제어환기온도, 제어난방온도, 환기조절 영향값, 난방조절 영향값
- ✓ 구동기데이터 공급온도, 천창, 커튼(상/하/측/외부), 3Way 밸브개도, 냉난방 작동상태, 보일러 작동상태, 순환펌프 작동상태, 유동팬 작동상태, 배기팬 작동상태

최초 미팅 시 전문가의 조언에 따라 83개 칼럼 중 필수 환경 데이터 선정

데이터 전처리

01

데이터 인코딩

- | 구동기 상태 등 범주형 데이터를 수치형 벡터로 변환
- | 원핫 인코딩을 통해 각 범주를 독립된 차원으로 표현

02

이상치 처리, 결측치 보간

- | 차트 시각화를 통해 이상치 확인 후 불필요한 이상치를 제거
- | 결측값을 이전 값과 이후 값을 기반으로 선형보간법을 사용

03

데이터 정규화

- | Min-Max 스케일링 기법을 사용해 모든 피처의 값이 (0~1)내에 있도록 조정
- | 특정 변수의 스케일 차이로 인한 왜곡을 방지해 모델의 안정성을 향상

04

데이터셋 분할

- | 전체 데이터를 학습, 검증, 테스트 데이터셋으로 나누어 모델 학습 및 평가 준비
- | 다양한 분할 방법을 실험한 결과 데이터를 연도에 따라 6:2:2 비율로 분할

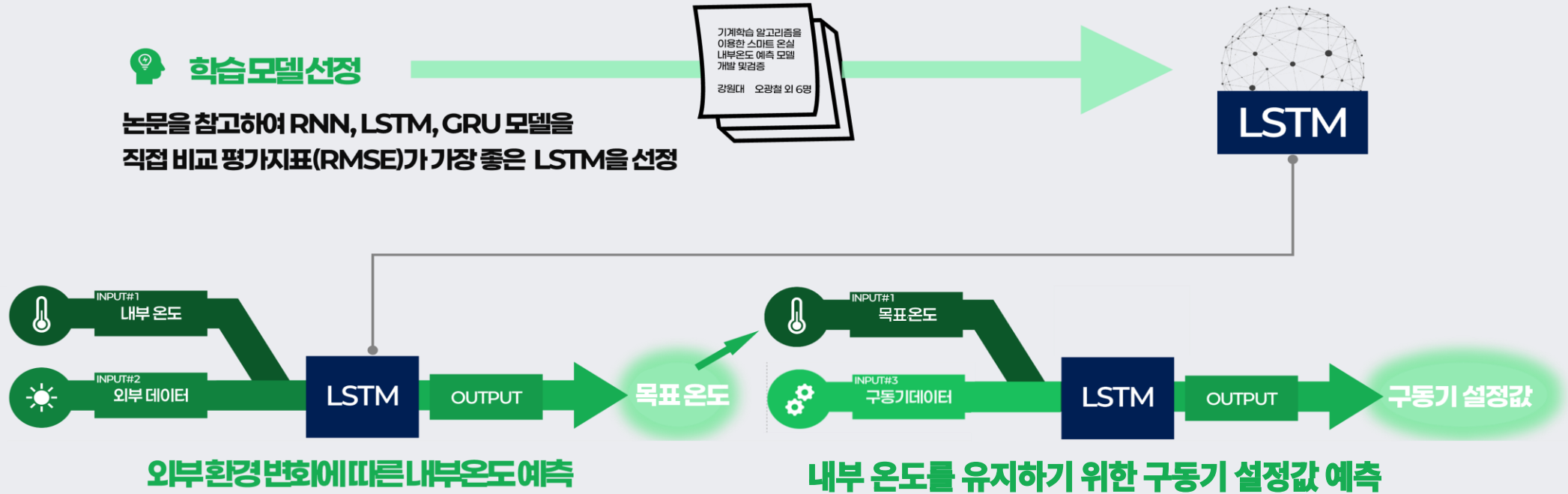
연구 과정

서론

연구 과정

연구 결과

결론 및 제언



점진적 확장 및 일반화를 위해 테스트 케이스를 세 가지로 구분



- 학습, 검증, 테스트(A)
- 학습, 검증, 테스트(A,B,C,D)
- 학습, 검증(A,B,C), 테스트(D)

연구 과정

서론

연구 과정

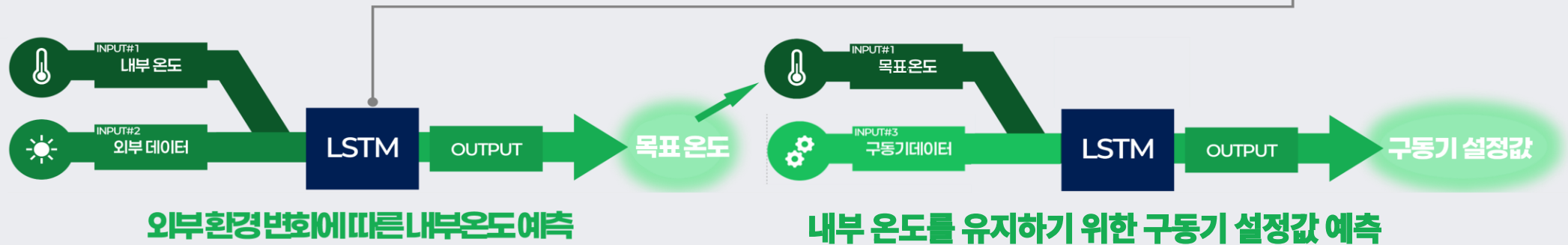
연구 결과

결론 및 제언

학습모델선택

논문을 참고하여 RNN, LSTM, GRU 모델을
직접 비교 평가지표(RMSE)가 가장 좋은 LSTM을 선정

기계학습 알고리즘			
	RNN	LSTM	GRU
RMSE	1.34	0.67	0.82
R ²	0.92	0.98	0.95



점진적 확장 및 일반화를 위해 테스트 케이스를 세 가지로 구분



학습, 검증, 테스트(A)
학습, 검증, 테스트(A,B,C,D)
학습, 검증(A,B,C), 테스트(D)

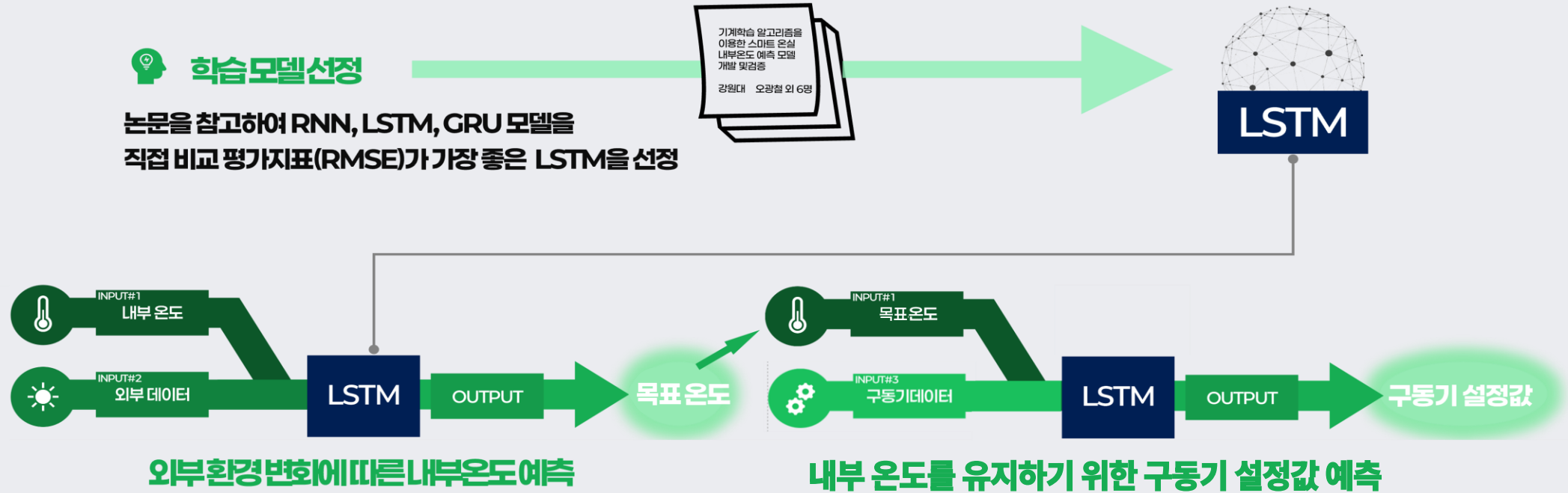
연구 과정

서론

연구 과정

연구 결과

결론 및 제언



점진적 확장 및 일반화를 위해 테스트 케이스를 세 가지로 구분



학습, 검증, 테스트(A)

학습, 검증, 테스트(A,B,C,D)

학습, 검증(A,B,C), 테스트(D)

연구 과정

서론

연구 과정

연구 결과

결론 및 제언

모델 평가 지표



RMSE

평균 오차의 크기
예측값이 실제값과 얼마나 근접한지 보여주는 수치
값이 작을수록 모델의 예측 오차가 작음



R^2

예측모델이 데이터를
얼마나 잘 설명하는지 나타내는 비율
값이 1에 가까울수록 모델의 예측 정확도가 높음

EX) RMSE가 0.1°C 이고 R^2 이 0.95이면 모델이 내부 온도의 변동성을 95% 설명하며, 예측 오차는 평균 $\pm 0.1^{\circ}\text{C}$

산업체 피드백 내용

온도 예측시 일반적으로 $\text{RMSE} < 0.2$ 이면 좋은 모델
실제 환경에서는 오차를 감안하여 0.1도 근처가 이상적

구동기 예측시 일반적으로 $\text{RMSE} < 5$ 이면 좋은 모델
실제 구동기 개도율 5% 이내의 오차는 고려할만한 수준

피드백 당시 온도 예측 모델을 통해 나온 RMSE가 0.3, 이후 RMSE를 더 낮추기 위해 모델을 개선

연구 과정

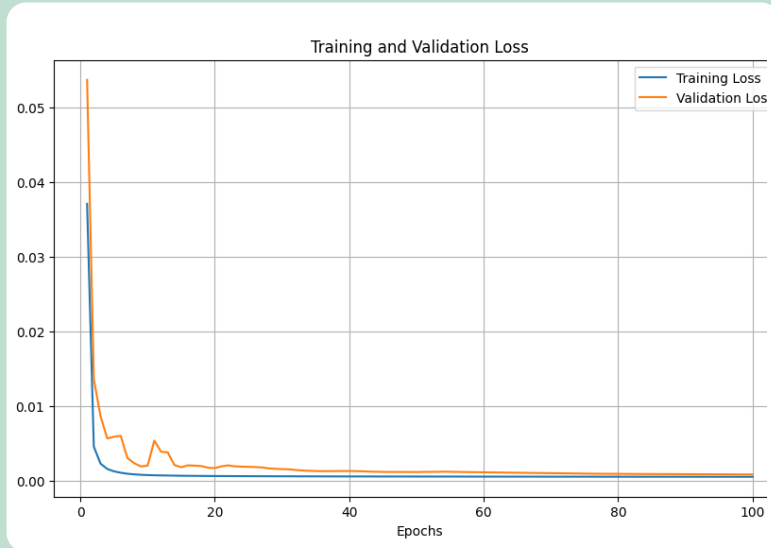
서론

연구 과정

연구 결과

결론 및 제언

모델 학습과정 점검



학습 및 검증 그래프

모델 개발

1 최적화 함수

- ❖ Adam 최적화 알고리즘에서 가중치 감쇠를 개선한 AdamW를 사용
- ❖ 학습률 조정과 적응형 학습 속도의 장점을 유지하면서 가중치를 규제

2 손실 함수

- ❖ 예측값과 실제값의 차이를 제공하여 평균을 계산하는 평균 제곱 오차(MSE) 사용
- ❖ 값의 차이가 클수록 손실이 크게 증가

3 개발 방법

- ❖ 과거 데이터를 학습 데이터로 사용하고, 가까운 시 데이터를 검증 및 테스트 데이터로 활용하여 미래를 예측하는 모델을 개발
- ❖ Seed() 함수를 사용해 항상 일정한 초기값이 입력되도록 설정

연구 과정

서론

연구 과정

연구 결과

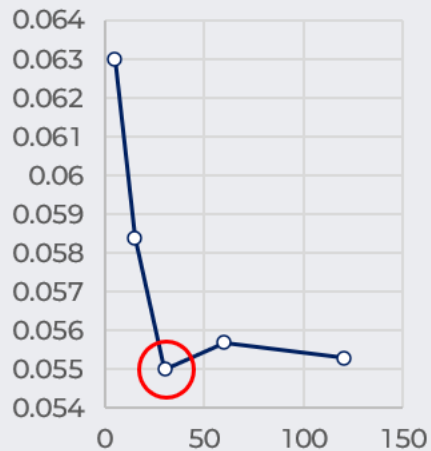
결론 및 제언

하이퍼 파라미터 튜닝

GRID SEARCH

Sequence length

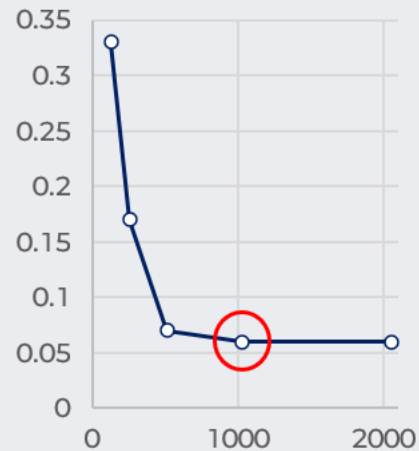
RMSE



범위 5,15,30,60,120

Hidden size

RMSE



128,256,512,1024,2048

layers

RMSE



1,2,3,4,5

Learning rate

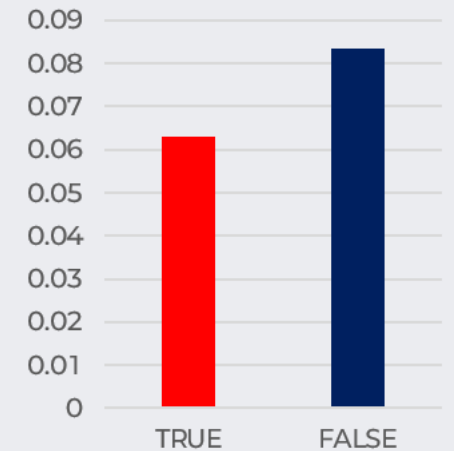
RMSE



0.00001 - 0.001

SHUFFLE

RMSE



셔플 여부

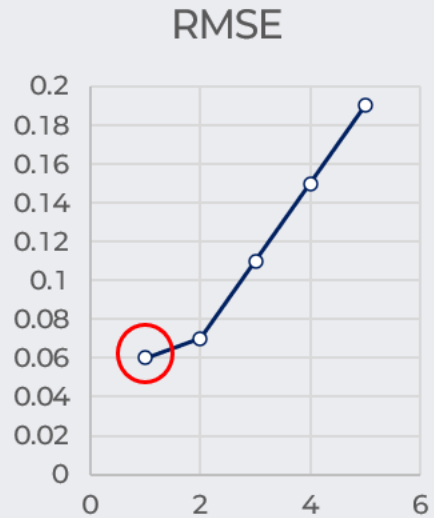
연구 과정

서론 | **연구 과정** | 연구 결과 | 결론 및 제언

GRID SEARCH

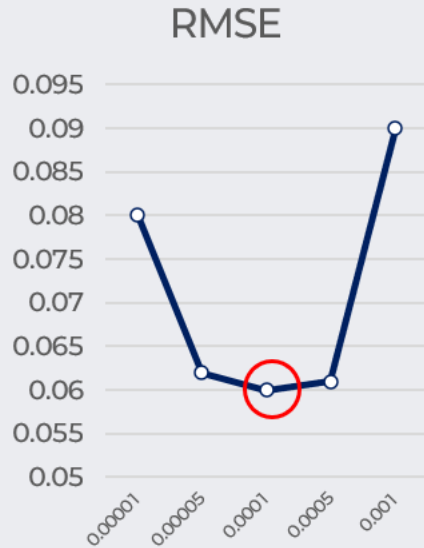
서플을 사용하는 이유

layers



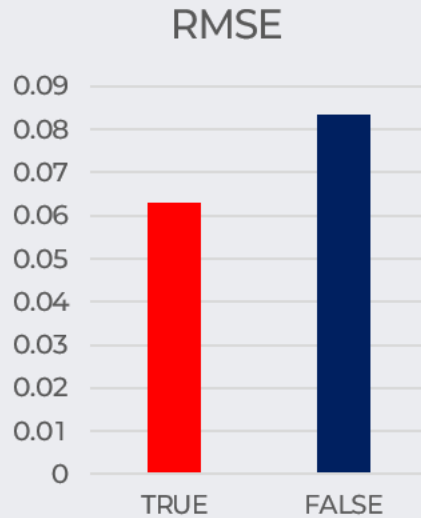
1,2,3,4,5

Learning rate



0.00001 - 0.001

SHUFFLE



서플 여부



시간적 순서의 중요성

시계열 데이터는 시간적 순서를 유지해야 하므로 일반적인 셔플링은 학습 성능을 저하시킬 수 있음



윈도우 방식의 데이터 분할

데이터는 sequence length로 나누어 윈도우 형태로 분할하고, 윈도우 간 셔플링을 적용함



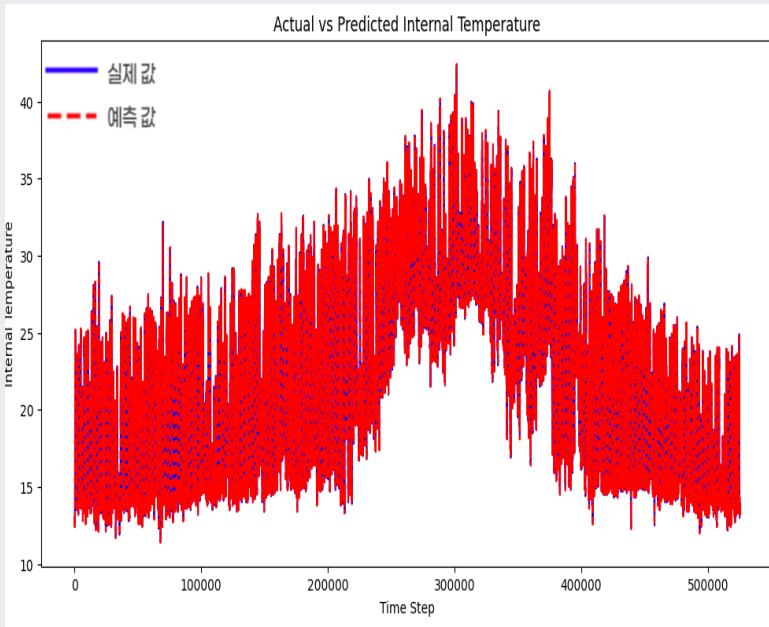
일반화 성능 향상

이를 통해 시간 순서를 유지하면서 데이터 다양성을 확보하고, 과적합 방지 및 일반화 성능을 향상시킬 수 있음

연구 과정

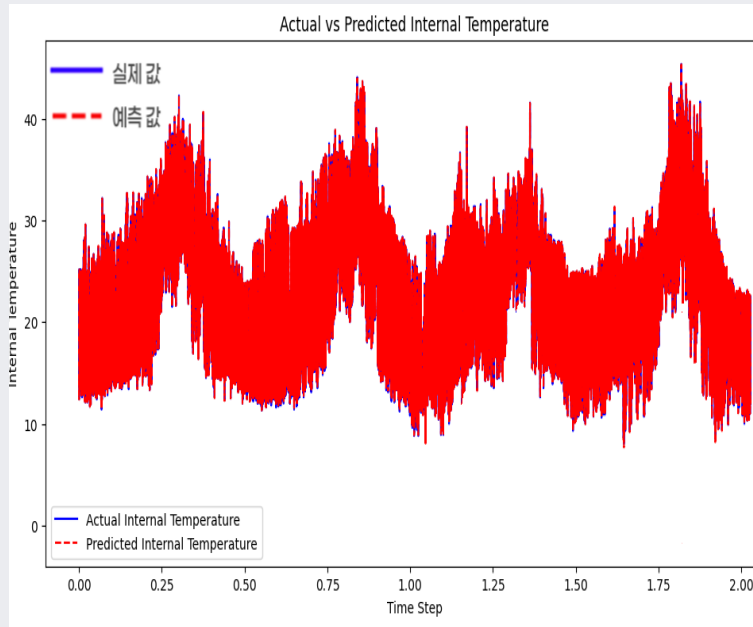
서론 | **연구 과정** | 연구 결과 | 결론 및 제언

22년 데이터와 D 농가의 데이터를 사용하여 테스트한 결과, 실제 값과 예측 값이 전체적으로 잘 일치



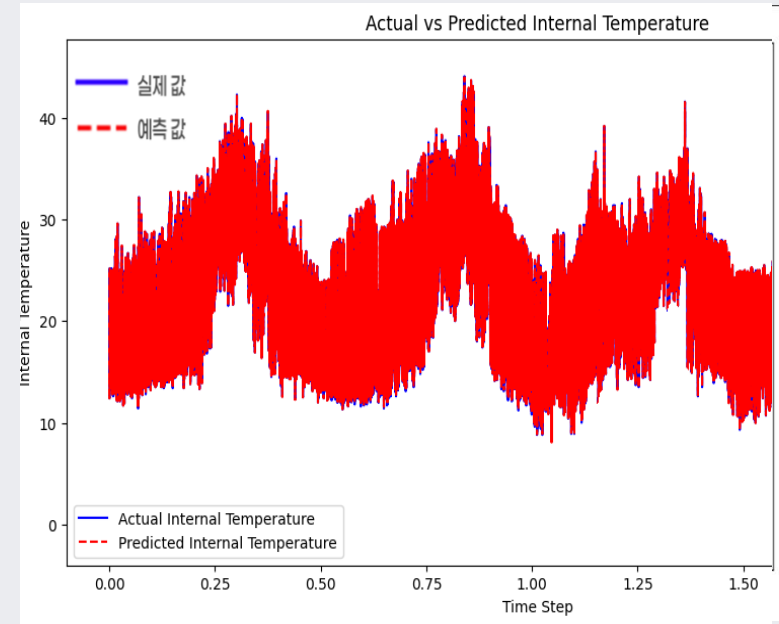
학습,검증 (A)

2022년 데이터 테스트(A)



학습,검증(A,B,C,D)

2022년 데이터 테스트(A,B,C,D)



학습,검증,(A,B,C)

일반화 테스트(D)

목차

01 서론

02 연구 과정

03 연구 결과

04 결론 및 제언

연구 결과

모델 성능

하이퍼 파라미터

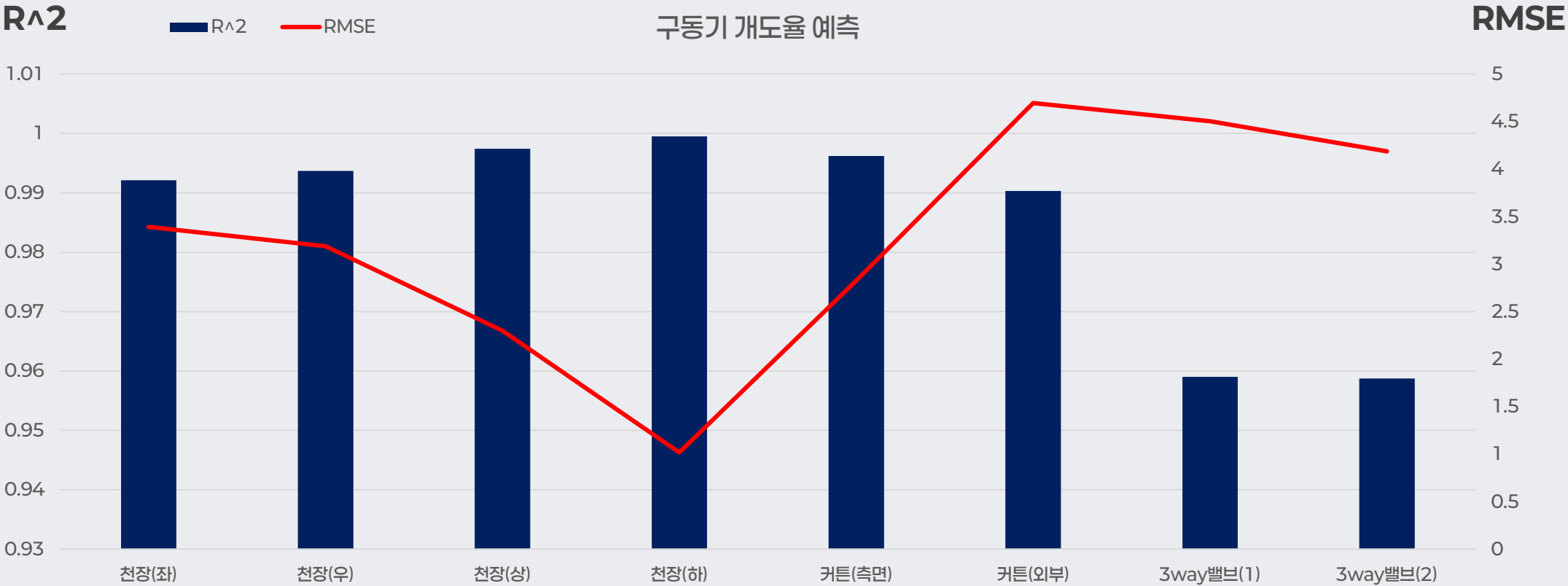
	Sequence length	Batch size	Hidden size	layers	Dropout rate	Learning rate	Epochs	Shuffle
학습,검증,테스트(A)	30	1024	512	1	0	0.0001	100	TRUE
학습,검증,테스트(A,B,C,D)	30	1024	512	1	0	0.0001	100	TRUE
학습,검증,(A,B,C),테스트(D)	30	1024	512	1	0	0.0001	100	TRUE
학습,검증,테스트(A,B,C,D)	30	1024	512	1	0	0.0001	100	TRUE



연구 결과

모델 성능

Sequence length	Batch size	Hidden size	layers	Dropout rate	Learning rate	Epochs	Shuffle
30	1024	512	1	0	0.0001	100	TRUE



연구 결과

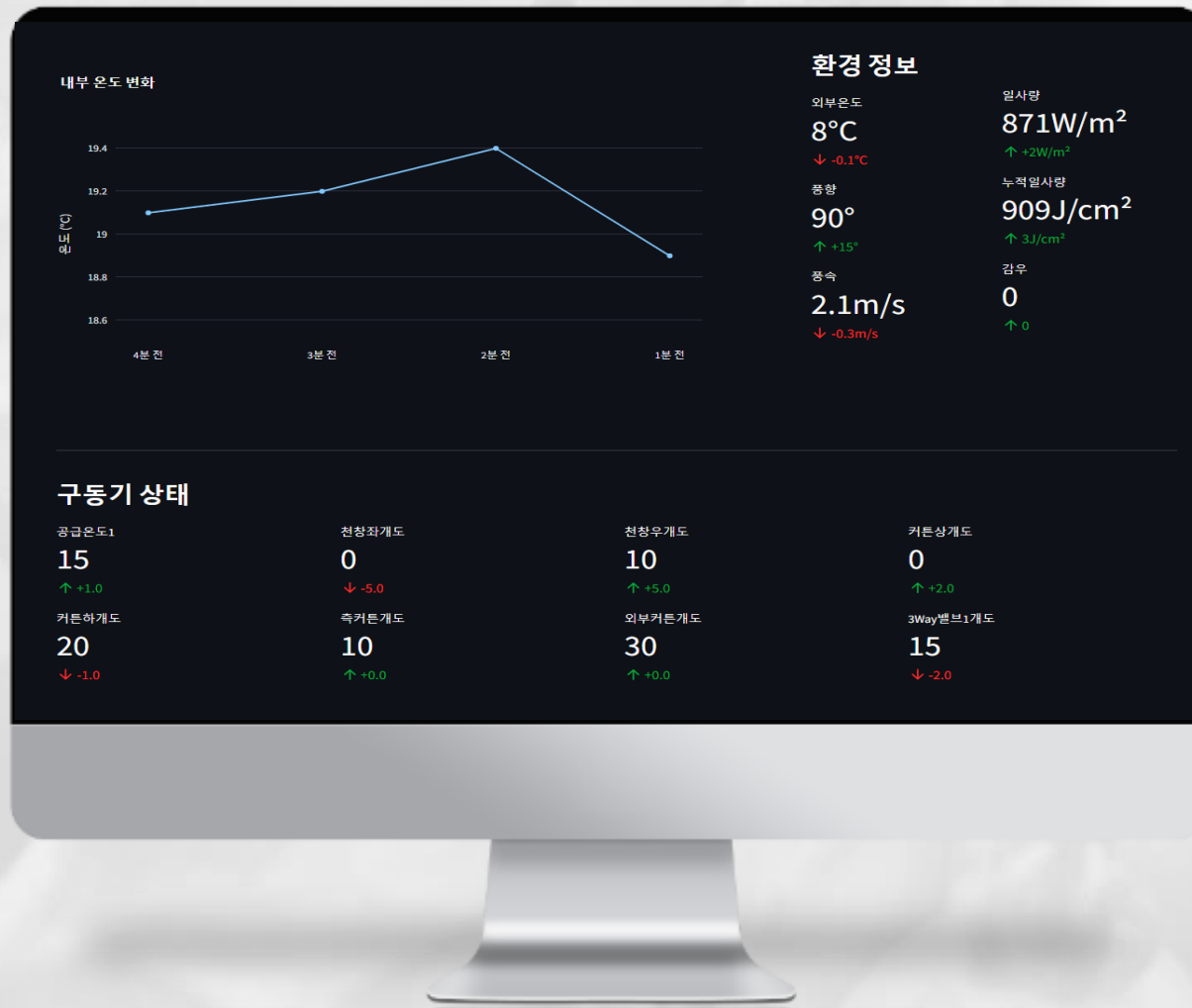
서론

연구 과정

연구 결과

결론 및 제언

내부온도&구동기 제어값 시각화



목차

01 서론

02 연구 과정

03 연구 결과

04 결론 및 제언

결론

서론

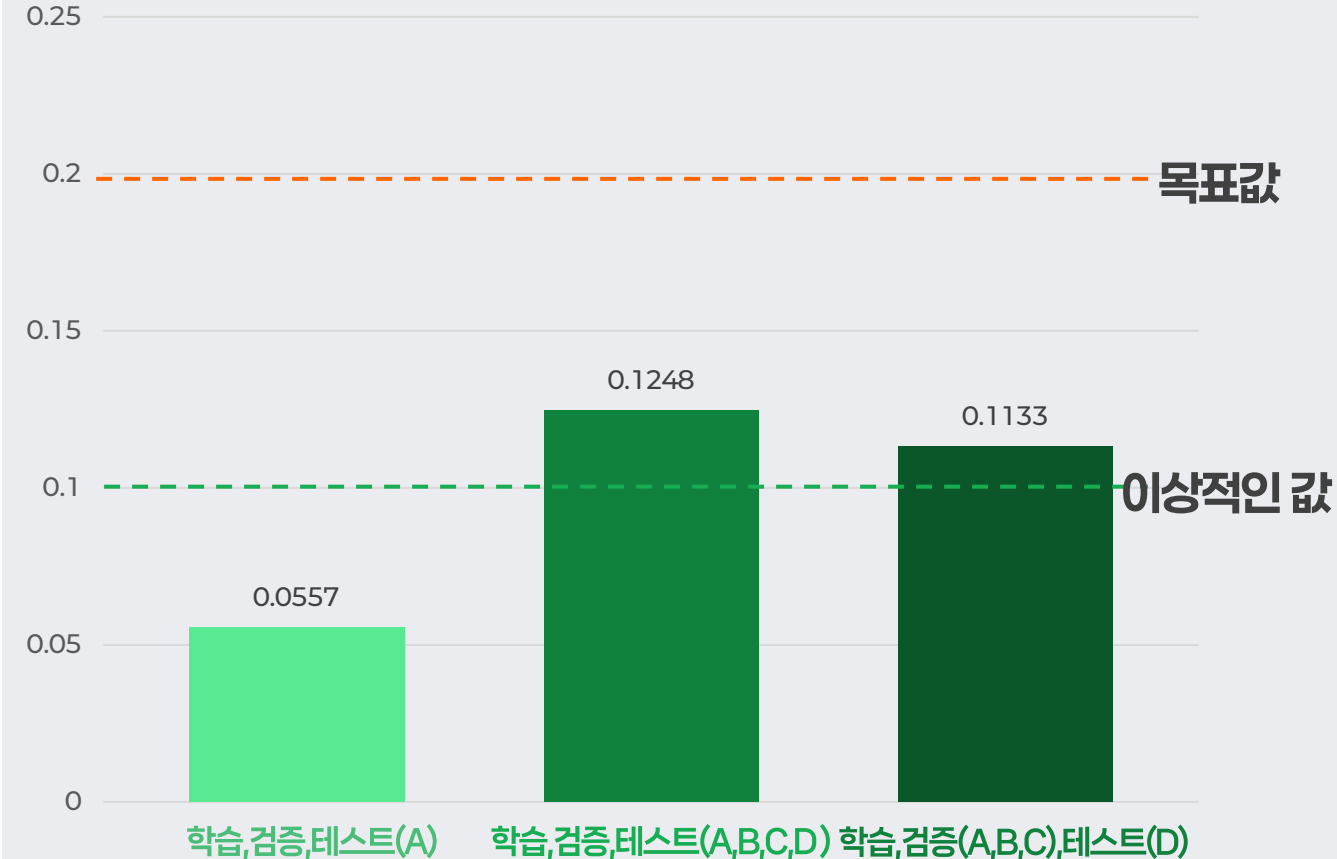
연구 과정

연구 결과

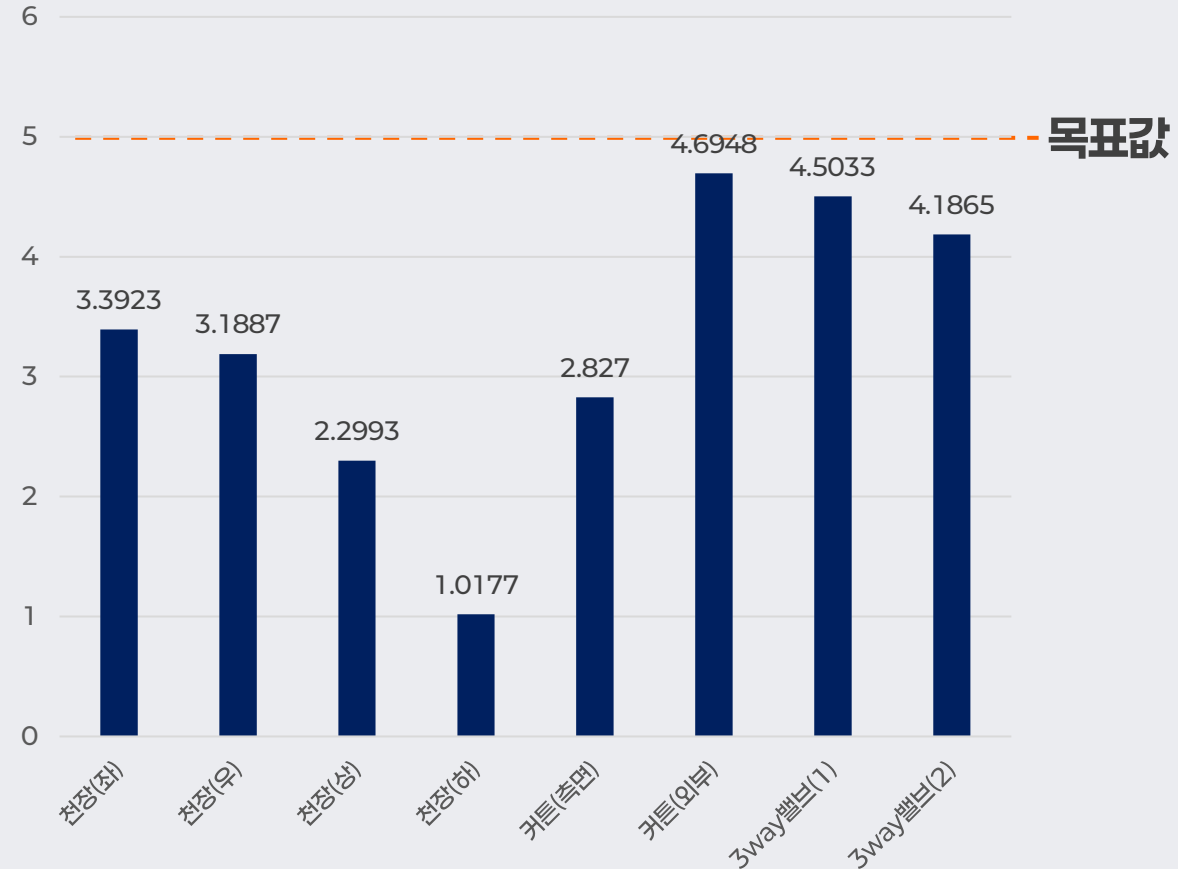
결론 및 제언

본 연구에서 개발한 모델은 목표한 성능을 성공적으로 달성하였으며
이는 농업 현장에서의 자동화 시스템 도입과 운영 효율성을 크게 향상시키는 기반이 될 것으로 기대

그룹별 온도 예측 모델 RMSE



구동기 개도율별 RMSE



제언

📍 다양한 농가 적용 및 일반화

현재 모델이 특정 농가 데이터를 기반으로 학습되었으므로, 다양한 지역과 작물을 대상으로 검증을 확대하여 일반화된 모델로 발전시키는 것이 필수

📍 클라우드 기반 모니터링 및 제어 시스템 도입

농장주가 언제 어디서나 온실 상태를 확인하고 제어할 수 있도록 클라우드와 IoT 기술을 접목한 플랫폼 구축이 필요

📍 확장 가능성 연구

기존 온실 뿐만 아니라 스마트 축산, 스마트 수경재배 등 다양한 분야로 모델 적용 범위를 확대하여 기술의 활용성을 극대화

Q & A

