

## Comparative Analysis of RNN Architectures and Activation Functions with Attention Mechanisms for Mars Weather Prediction

Jaehyeok Jo\*, Yunho Sin\*\*, Bo-Young Kim\*\*\*, Jihoon Moon\*\*\*\*

\*Student, Dept. of AI and Big Data, Soonchunhyang University, Asan, Korea

\*\*Student, Asan Middle School, Asan, Korea

\*\*\*Teacher, Asan Middle School, Asan, Korea

\*\*\*\*Professor, Dept. of AI and Big Data, Soonchunhyang University, Asan, Korea

### [Abstract]

In this paper, we propose a comparative analysis to evaluate the impact of activation functions and attention mechanisms on the performance of time-series models for Mars meteorological data. Mars meteorological data are nonlinear and irregular due to low atmospheric density, rapid temperature variations, and complex terrain. We use long short-term memory (LSTM), bidirectional LSTM (BiLSTM), gated recurrent unit (GRU), and bidirectional GRU (BiGRU) architectures to evaluate the effectiveness of different activation functions and attention mechanisms. The activation functions tested include rectified linear unit (ReLU), leaky ReLU, exponential linear unit (ELU), Gaussian error linear unit (GELU), Swish, and scaled ELU (SELU), and model performance was measured using mean absolute error (MAE) and root mean square error (RMSE) metrics. Our results show that the integration of attentional mechanisms improves both MAE and RMSE, with Swish and ReLU achieving the best performance for minimum temperature prediction. Conversely, GELU and ELU were less effective for pressure prediction. These results highlight the critical role of selecting appropriate activation functions and attention mechanisms in improving model accuracy for complex time-series forecasting.

▶ **Key words:** Mars Weather Prediction, Time-Series Analysis, Attention Mechanism, Activation Functions, Deep Learning, Predictive Modeling

- 
- First Author: Jaehyeok Jo, Corresponding Author: Jihoon Moon
  - \*Jaehyeok Jo (ajdad123@naver.com), Dept. of AI and Big Data, Soonchunhyang University
  - \*\*Yunho Sin (hoya0710@naver.com), Asan Middle School
  - \*\*\*Bo-Young Kim (bboya414@naver.com), Asan Middle School
  - \*\*\*\*Jihoon Moon (jmoon22@sch.ac.kr), Dept. of AI and Big Data, Soonchunhyang University
  - Received: 2024. 09. 27, Revised: 2024. 10. 17, Accepted: 2024. 10. 18.

## [요 약]

본 연구는 화성 기상 데이터를 대상으로 활성화 함수와 어텐션 메커니즘이 시계열 모델의 성능에 미치는 영향을 평가하기 위해 비교 및 분석한다. 화성의 기상 데이터는 대기 밀도가 낮고, 급격한 온도 변동 및 복잡한 지형 등으로 인해 비선형적이고 불규칙적이다. 본 연구에서는 LSTM, BiLSTM, GRU, BiGRU 아키텍처를 사용하여 다양한 활성화 함수와 어텐션 메커니즘의 효과를 평가한다. 실험에 사용된 활성화 함수는 ReLU, Leaky ReLU, ELU, GELU, Swish, SELU이며, 모델 성능은 MAE와 RMSE 지표로 측정된다. 실험 결과, 어텐션 메커니즘을 통합함으로써 MAE와 RMSE가 모두 향상되었으며, Swish와 ReLU는 최저 온도 예측에서 가장 우수한 성능을 보였다. 반면, GELU와 ELU는 기압 예측에서 성능이 저하되었다. 이러한 결과는 복잡한 시계열 예측의 모델 정확도를 향상하기 위해 적절한 활성화 함수와 어텐션 메커니즘을 선택하는 것이 중요함을 보여준다.

▶ **주제어:** 화성 기상 예측, 시계열 데이터 분석, 어텐션 메커니즘, 활성화 함수, 딥러닝, 예측 모델링

## I. Introduction

화성은 태양계에서 네 번째 행성으로서, 지구와의 근접성과 유사성 때문에 수 세기 동안 천문학과 과학 연구의 중심에 있으며, 약 24시간 37분의 하루 길이와 약 25.2도의 자전축 기울기를 갖고 있어 지구와 유사한 네 개의 뚜렷한 계절을 경험한다[1]. 극지방의 얼음과 과거 물의 흐름 흔적은 화성이 한때 생명체가 존재할 수 있는 조건을 가졌을 가능성을 시사한다[2]. 그러나 이러한 유사성에도 불구하고, 화성은 얇은 대기와 약한 자기장을 특징으로 하며, 이는 낮은 표면 온도, 강력한 먼지 폭풍, 높은 수준의 이산화탄소, 우주 방사선 노출 등 가혹한 환경을 조성한다[1, 2].

이러한 극한 조건에도 불구하고, 화성 탐사 및 식민지화에 관한 관심이 증가하고 있다. Elon Musk가 이끄는 SpaceX와 같은 조직은 화성을 인류의 두 번째 거주지로 만들기 위한 연구와 임무를 적극적으로 수행하고 있다[3]. 이러한 노력은 기후 변화, 자연재해, 자원 고갈, 전염병과 같은 지구에만 의존하는 것에 따른 존재론적 위험을 완화해야 한다는 필요성에서 비롯되었다[4]. 지구 밖 거주지의 확립은 인류의 장기적인 생존을 위해 필수적이다. 태양계에서 지구와 유사한 행성 중에서 화성은 낮은 대기압과 극한의 환경을 가지고 있지만, 접근성 및 기술적 적용 가능성 측면에서 가장 유망한 선택지이다[5]. 인간의 생존에 필요한 물을 공급하거나 산소를 생성하기 위한 얼음 형태의 물 자원 활용, 이산화탄소가 풍부한 화성 대기를 이용한 산소나 연료 생성 기술 MOXIE(Mars Oxygen In-Situ Resource Utilization Experiment) 개발 등이 진행되고 있다[2, 5].

화성의 기상 패턴을 이해하는 것은 모든 탐사 활동과 미래의 식민지화 노력에서 매우 중요하다. 화성의 기후는 지

구와는 다른 자전축, 희박한 대기, 약한 자기장 등으로 인해 독특한 특성을 보인다. 화성 기상 데이터의 수집과 분석은 기후 역학에 대한 귀중한 통찰력을 제공하며, 극한 기상 조건을 시뮬레이션하고 연구하는 데 도움이 된다. 이러한 지식은 임무 계획, 거주지 건설, 태양 복사가 우주선에 미치는 영향, 우주 먼지와 먼지 폭풍의 영향 등 인간이 직면할 수 있는 극한 환경에 대비하는 데 필수적이다.

화성의 비선형적이고 동적인 기상 패턴을 효과적으로 분석하고 예측하기 위해서는 복잡한 시간적 의존성을 포착할 수 있는 모델이 필요하다[6]. RNN(Recurrent Neural Network) 아키텍처인 LSTM(Long Short-Term Memory), BiLSTM(Bidirectional LSTM), GRU(Gated Recurrent Unit), BiGRU(Bidirectional GRU)는 이러한 과제에 적합하다[7]. 이러한 모델의 예측 성능은 배치 크기, 레이어 깊이와 같은 초매개변수(Hyperparameter), 학습률(Learning Rate), Attention Mechanism, Activation Function 등 다양한 요인에 의해 크게 영향을 받는다[8]. 특히 Attention Mechanism과 Activation Function은 입력 정보를 처리하고 출력을 생성하는 방식을 결정하여 데이터의 복잡한 패턴을 포착하는 모델의 능력에 영향을 미친다[9].

그러나 현재까지 화성 기상 예측에 사용되는 다양한 RNN 아키텍처에서 Attention Mechanism과 Activation Function의 선택이 성능에 어떻게 영향을 미치는지에 대한 종합적인 가이드라인은 부족한 실정이다. 이에 따라 본 연구에서는 이러한 공백을 메우기 위해, 다양한 RNN 모델에 여러 Activation Function을 적용하고, Attention

Mechanism의 적용 여부에 따른 성능 변화를 평가하여 효과적인 모델링 가이드라인을 제공하고자 한다.

본 연구에서는 LSTM, BiLSTM, GRU, BiGRU 아키텍처를 사용하여 예측 모델을 구축하고, ReLU(Rectified Linear Unit), Leaky ReLU, ELU(Exponential Linear Unit), GELU(Gaussian Error Linear Unit), Swish(SiLU) SELU(Scaled ELU) 등 여섯 가지 Activation Function을 적용한다. 또한, 각 모델에서 Attention Mechanism의 적용 여부에 따라 성능을 평가한다. 모델은 NASA의 MSL(Mars Science Laboratory) Curiosity rover에 탑재된 REMS(Rover Environmental Monitoring Station)이 2012년부터 2022년까지 수집한 화성 기상 데이터를 사용하여 학습 및 평가한다.

본 논문의 나머지 구성은 다음과 같다. 2장에서는 화성 기상 예측에 관한 선행 연구를 검토하고, 3장에서는 사용된 데이터의 전처리 과정과 다양한 RNN 모델의 구성 및 실험 설정을 자세히 설명한다. 4장에서는 실험 결과를 제시하고 분석하며, 마지막으로 5장에서는 결론을 내리고 향후 연구 방향을 제시한다.

## II. Related Work

Pla-García 등[10]은 2020년에 MRAMS(Mars Regional Atmospheric Modeling System)와 MarsWRF(Mars Weather Research and Forecasting model)을 사용하여 화성 2020 퍼서비어런스 로버의 착륙 지점인 예제로 크레이터의 지역 기상 조건을 예측하였다. 이 연구의 목적은 MEDA(Mars Environmental Dynamics Analyzer) 장비로부터 얻어진 데이터를 해석하는 데 도움을 주는 것이다. 연구에서는 연중, 네 계절(태양 적경  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$ ,  $270^\circ$ ) 동안의 일주기를 조사하였으며, 대기 및 지면 온도, 기압, 풍속과 풍향, 지표 복사 플럭스, 수분 데이터를 모델링하였다. 이 연구는 MRAMS와 MarsWRF 모델이 기상 조건을 재현하는 데 효과적임을 보여주었으며, 이를 통해 예제로 크레이터 내의 기상 조건을 예측하는 이들 모델의 신뢰성을 높였다.

Priyadarshini와 Puri[11]는 2021년에 CNN(Convolutional Neural Network), GRU, LSTM, Stacked LSTM, CNN-LSTM 모델 등 다양한 딥러닝 모델을 사용하여 화성 기상 데이터를 종합적으로 분석하였다. 이 연구는 미래의 화성 식민지화 노력을 촉진하기 위해 화성 기상을 이해하는 것이 중요하다는 점을 강조하였

다. 각 모델의 성능은 MAE(Mean Absolute Error)와 MSE(Mean Squared Error), RMSE(Root MSE), 결정 계수( $R^2$ )와 같은 통계적 지표를 통해 평가되었으며, LSTM 모델이 가장 높은 예측 정확도를 도출하였다.

Pant 등[12]은 2023년에 기계학습 기법이 화성 기상 데이터 분석에 미치는 잠재력을 조사하였다. 이 연구는 MSL 로버의 REMS이 제공한 데이터를 활용하였다. 선형 회귀, 의사결정 나무, 서포트 벡터 머신, 랜덤 포레스트, 신경망 등 다양한 기계학습 알고리즘을 사용하여 데이터 내의 패턴과 관계를 파악하고자 하였다. 분석 결과, 화성의 온도는 변동성을 보이며, 최소 온도는 비교적 좁은 범위의 변동을, 최대 온도는 더 큰 변동을 나타내었다. 또한, 모델의 정확도와 해석력을 향상하기 위해 특징 공학과 차원 축소의 중요성을 강조하였다. 엘보우 방법을 사용하여 최적의 클러스터 수를 결정한 결과, 세 개의 클러스터가 최적임을 확인하였다. 선형 회귀 모델은 화성 기상 패턴 예측에서 85%의 정확도를 보였다.

앞서 언급한 연구들은 화성의 기상 조건을 예측하거나 분석하기 위해 다양한 물리 기반 모델과 기계학습 기법을 적용하였다. Pla-García 등[10]은 물리 기반의 대기 모델인 MRAMS와 MarsWRF를 활용하여 특정 지역의 기상 조건을 상세히 예측하였으며, Priyadarshini와 Puri[11], Pant 등[12]은 딥러닝과 기계학습 알고리즘을 사용하여 화성 기상 데이터를 분석하고 예측하였다. 그러나 이들 연구는 주로 특정 모델 또는 알고리즘의 적용과 그 성능 평가에 초점을 맞추고 있으며, 다양한 RNN 아키텍처와 활성화 함수, 어텐션 메커니즘의 조합이 모델 성능에 미치는 영향을 종합적으로 비교해 분석하지는 않았다.

본 연구는 화성 기상 예측에 있어 다양한 RNN 모델(LSTM, BiLSTM, GRU, BiGRU)과 여러 활성화 함수(ReLU, Leaky ReLU, ELU, GELU, SiLU, SELU)의 적용 효과를 체계적으로 비교하고, 어텐션 메커니즘의 활용이 모델 성능에 미치는 영향을 평가한다. 이를 통해 비선형적인 시계열 데이터를 다루는 데 있어 어떤 모델 구성과 활성화 함수가 가장 효과적인지에 대한 구체적인 가이드라인을 제공한다[13, 14]. 특히, 어텐션 메커니즘의 유무에 따른 모델 성능 변화를 분석함으로써, 복잡한 패턴을 가진 화성 기상 데이터를 정확하게 예측하기 위한 최적의 모델링 전략을 제시한다[15, 16].

따라서 본 논문의 기여도는 다음과 같다.

- 다양한 RNN 아키텍처와 활성화 함수의 비교 분석: 기존 연구와 달리 여러 RNN 모델과 활성화 함수를 조합하여 그 성능을 종합적으로 평가한다.

- 어텐션 메커니즘의 영향 평가: 어텐션 메커니즘의 적용 여부에 따른 모델 성능의 변화를 분석하여, 복잡한 시계열 데이터 예측에서 중요성을 확인한다.
- 모델링 가이드라인 제시: 화성 기상 예측을 위한 최적의 모델 구성에 대한 구체적인 지침을 제공함으로써, 향후 연구 및 실무 적용에 유용한 참고 자료를 제공한다.

이러한 기여를 통해 본 연구는 화성 기상 예측 모델링 분야에서 기존 연구의 한계를 보완하고, 더 정확하고 효율적인 예측 모델 개발에 도움을 주고자 한다.

### III. Methodology

#### 1. Dataset and Preprocessing

본 연구에서는 NASA의 MSL 미션의 일환으로 2012년부터 2022년까지 Curiosity 로버에 탑재된 REMS이 수집한 화성 기상 데이터를 사용하였다[17]. 해당 데이터는 화성의 기상 패턴을 이해하고 예측하기 위한 풍부한 정보를 제공한다.

##### 1.1 Dataset Composition

데이터 셋에는 다음과 같은 변수들이 포함되어 있다:

- Terrestrial Date: 지구 날짜로, 데이터가 수집된 날짜
- Sol: 화성 일(Martian Solar Day)로, 화성에서의 날짜
- Ls (Solar Longitude): 태양 경도는 화성의 계절적 변화를 나타내는 변수로, 화성의 공전 궤도 상에서의 위치를 0°부터 360°까지 의미
- 최소 온도(min\_temp): 해당 솔(Sol)에서 측정된 최저 온도(섭씨)
- 최대 온도(max\_temp): 해당 솔에서 측정된 최고 온도(섭씨)
- 대기압(pressure): 해당 솔에서 측정된 대기압(파스칼)

데이터 셋은 총 3,253개의 관측치로 구성되어 있으며, 일부 변수에는 결측치가 존재한다.

##### 1.2 Data Preprocessing

분석의 정확성을 높이기 위해 결측치 처리, 불필요한 변수 제거, 주기성 반영, 데이터 정규화 등과 같은 전처리 과정을 수행하였다. 결측치 처리에서는 최소 온도, 최대 온도, 대기압 변수에서 결측치가 발견되었으며, 이는 선형 보간법을 사용하여 보완하였다[18]. 이를 통해 데이터의 연속성을 유지하고 모델의 학습에 필요한 충분한 데이터를 확보하였다. 불필요한 변수 제거에서는 데이터 식별 번호와 기상 상태 변수는 결측치가 많고 분석에 유의미한 영

향을 주지 않는다고 판단하여 제거하였다. 주기성 반영에서는 날짜 관련 변수의 주기적 특성을 반영하기 위해 지구 날짜(Terrestrial Date), 태양 경도(Ls), 월(month) 데이터를 사인(sin)과 코사인(cos) 함수로 변환하였으며, 이는 주기적 패턴을 모델이 학습할 수 있도록 도와준다[14, 16]. 변환된 변수는 다음과 같다.

- Day\_sin, Day\_cos: 지구 날짜의 주기성 반영
- Ls\_sin, Ls\_cos: 태양 경도의 주기성 반영
- Month\_sin, Month\_cos: 월의 주기성 반영

데이터 정규화에서는 모델 학습의 효율성을 높이고 각 변수의 스케일 차이를 줄이기 위해 최소-최대 정규화를 적용하여 모든 변수를 [0, 1] 범위로 스케일링하였다[19]. 또한, 분석 대상의 종속 변수는 기후 판단에 중요한 최소 온도, 최대 온도, 대기압으로 설정하였다.

전처리된 데이터를 모델의 학습과 평가를 위해 훈련 집합과 평가 집합으로 분할하였다. 일반적인 데이터 분할 방식에 따라 전체 데이터의 80%를 훈련 집합으로 사용하고, 나머지 20%를 평가 집합으로 활용하였으며, 시계열 데이터의 특성을 고려하여 시간 순서를 유지하면서 분할하였다.

#### 2. Models and Activation Functions

화성 기상 데이터의 비선형적이고 시계열적인 특성을 효과적으로 학습하기 위해 다양한 RNN 아키텍처와 활성화 함수를 적용하였다[13, 14]. 또한, 모델의 성능 향상을 위해 어텐션 메커니즘을 도입하였다[15, 16].

##### 2.1 Model Architecture

본 연구는 다음의 네 가지 RNN 기반 모델을 사용하였다.

- LSTM: 장기 의존성 문제를 해결하기 위해 설계된 RNN의 한 종류로, 기억 셀과 게이트 구조를 통해 시계열 데이터의 장기 패턴을 효과적으로 학습한다[20, 21].
- BiLSTM: LSTM의 확장 버전으로, 순방향과 역방향으로 데이터를 처리하여 과거와 미래의 정보를 모두 활용할 수 있어, 양방향의 맥락 정보를 통합하여 예측 성능을 향상한다[22, 23].
- GRU: LSTM의 간소화된 버전으로, 업데이트 게이트와 리셋 게이트를 사용하여 계산 복잡도를 줄이면서도 장기 의존성을 학습할 수 있다[13, 16].
- BiGRU: GRU의 확장 버전으로, BiLSTM과 마찬가지로 양방향으로 데이터를 처리하여 시퀀스의 전후 맥락을 모두 학습한다[24].

### 2.2 Activation Functions

활성화 함수는 신경망에서 입력 신호를 출력으로 변환하는 비선형 함수로, 모델의 학습 능력과 성능에 큰 영향을 미친다[14]. 본 연구에서는 다음의 여섯 가지 활성화 함수를 비교하였다. 각 활성화 함수의 수식은 표 1에 제시되어 있다.

- ReLU: 입력이 양수이면 그대로 출력하고, 음수이면 0을 출력한다. 계산이 효율적이며 딥러닝에서 널리 사용된다.
- Leaky ReLU: ReLU의 변형으로, 음수 입력에 작은 기울기를 부여하여 죽은 뉴런 문제를 완화한다.
- ELU: ReLU의 단점을 보완하기 위해 설계되었으며, 음수 입력에 대해 지속적으로 감소하는 출력을 제공한다.
- GELU: 입력값의 확률 분포를 고려하여 부드러운 비선형성을 제공하며, Transformer 기반 모델에서 자주 사용된다.
- Swish (SiLU): 입력값과 시그모이드 함수의 곱으로 정의되며, 깊은 신경망에서 우수한 성능을 보인다.
- SELU: 입력값의 분산과 평균을 유지하는 특성이 있어, SNN(Self-Normalizing Neural Network)에 적합하다.

Table 1. Formulas of Activation Functions

Activation Functions	Formulas
ReLU	$f(x) = \max(0, x)$
Leaky ReLU	$f(x) = \begin{cases} x & \text{if } x > 0 \\ ax & \text{otherwise} \end{cases}$
ELU	$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha(e^x - 1) & \text{otherwise} \end{cases}$
GELU	$f(x) = 0.5x(1 + \tanh(\sqrt{2/\pi} * (x + 0.044715x^3)))$
Swish (SiLU)	$f(x) = x * \text{sigmoid}(x)$
SELU	$f(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha(e^x - 1) & \text{otherwise} \end{cases}$

### 2.3 Attention Mechanism

어텐션 메커니즘은 시퀀스 데이터에서 중요한 부분에 가중치를 부여하여 모델이 중요한 정보에 집중할 수 있도록 도와준다[25]. 이는 RNN 모델의 성능 향상에 효과적이며, 특히 긴 시퀀스나 복잡한 패턴을 보인 데이터에서 유용하다. 본 연구에서는 각 RNN 모델에 어텐션 메커니즘을 적용하여 성능 변화를 평가하였다. 어텐션 메커니즘은 RNN의 은닉 상태 출력에 기반하여 각 시점(Time Step)의 중요도를 계산하고, 이를 가중하여 최종 출력에 반영한다[16, 24]. 어텐션 메커니즘의 적용 여부에 따라 모델을 두 그룹으로 나누어 실험하였다.

- 어텐션 적용 모델 (With attention, w/ att) : RNN 모델의 출력에 어텐션 레이어를 추가하여 중요한 시점의 정보를 강조한다.
- 어텐션 미적용 모델 (Without attention, w/o att) : 기본 RNN 모델로서 어텐션 메커니즘 없이 시계열 데이터를 처리한다.

## IV. Results

### 1. Experimental Setup

각 모델(LSTM, BiLSTM, GRU, BiGRU)에 대해 여섯 가지 활성화 함수(ReLU, Leaky ReLU, ELU, GELU, SiLU, SELU)를 적용하였으며, 어텐션 메커니즘의 적용 여부에 따라 총 48개의 모델 구성을 실험하였다. 이를 통해 모델 아키텍처, 활성화 함수, 어텐션 메커니즘이 모델 성능에 미치는 영향을 종합적으로 평가하였다.

모델 학습을 위한 공통 설정은 다음과 같다.

- 손실 함수: 이상치에 강인한 후버 손실 함수를 사용하였다[26, 27].
  - 최적화 알고리즘: Adam 옵티마이저를 사용하였으며, 학습률은 0.001로 설정하였다[28].
  - 조기 종료: 과적합을 방지하기 위해 Early Stopping 기법을 적용하였으며, 검증 손실이 10회 연속으로 개선되지 않을 때 학습을 종료하였다[29].
  - 배치 크기 및 에포크: 배치 크기는 12로, 최대 에포크 수는 100으로 설정하였다.
- 모델의 성능은 회귀 문제에서 일반적으로 사용되는 MAE와 RMSE를 사용하여 평가하였다[15, 30].
- MAE: 예측값과 실제값 차이의 절댓값 평균으로, 오차의 크기를 직접적으로 나타낸다.
  - RMSE: 오차의 제곱 평균에 제곱근을 취한 값으로, 큰 오차에 더 큰 페널티를 부여한다.

### 2. Results and Analysis

본 연구에서는 LSTM, BiLSTM, GRU, BiGRU 모델을 각각 사용하여, ReLU, Leaky ReLU, ELU, GELU, SiLU, SELU와 같은 다양한 활성화 함수의 영향을 분석하였다. Table 2부터 Table 7까지 각각 최소 온도(min\_temp), 최대 온도(max\_temp), 기압(pressure)에 대한 모델의 예측 성능을 나타내는 MAE와 RMSE 결과를 제시하였다. 구체적으로, Table 2는 최소 온도 예측에 대한 MAE 결과를, Table 3은 최소 온도 예측에 대한 RMSE 결과를, Table 4

는 최대 온도 예측에 대한 MAE 결과를, Table 5는 최대 온도 예측에 대한 RMSE 결과를, Table 6은 기압 예측에 대한 MAE 결과를, Table 7은 기압 예측에 대한 RMSE 결과를 보여준다.

Table 2. MAE Results for Minimum Temperature (min\_temp)

Model	ReLU	Leaky ReLU	ELU	GELU	Swish	SELU
LSTM (w/ att)	2.652	3.091	3.669	2.462	2.657	3.498
LSTM (w/o att)	3.595	3.408	3.280	2.714	3.132	3.060
BiLSTM (w/ att)	3.446	2.944	3.809	2.581	2.636	3.741
BiLSTM (w/o att)	3.274	3.182	3.010	2.343	3.241	2.848
GRU (w/ att)	2.969	2.418	3.910	3.414	2.422	2.975
GRU (w/o att)	2.728	3.333	3.367	2.761	1.967	3.433
BiGRU (w/ att)	2.972	2.963	3.539	3.422	3.100	3.139
BiGRU (w/o att)	3.120	3.268	3.445	3.612	2.540	3.753

Table 3. RMSE Results for Minimum Temperature (min\_temp)

Model	ReLU	Leaky ReLU	ELU	GELU	Swish	SELU
LSTM (w/ att)	3.809	4.155	4.598	3.682	3.780	4.520
LSTM (w/o att)	4.880	4.587	4.363	3.784	4.142	4.208
BiLSTM (w/ att)	4.656	4.074	4.865	3.815	3.831	4.819
BiLSTM (w/o att)	4.350	4.328	4.093	3.650	4.505	3.979
GRU (w/ att)	4.054	3.623	5.215	4.779	3.602	4.063
GRU (w/o att)	3.779	4.414	4.531	3.953	3.247	4.541
BiGRU (w/ att)	4.815	4.028	4.616	4.548	4.217	4.235
BiGRU (w/o att)	4.234	4.370	4.486	4.779	3.697	4.748

각 표에서는 어텐션 메커니즘을 적용한 모델(w/ att)과 미적용한 모델(w/o att)의 성능을 비교하여, 다양한 활성화 함수(ReLU, Leaky ReLU, ELU, GELU, Swish, SELU)가 모델의 예측 정확성에 미치는 영향을 분석하였다. 이를 통해 어텐션 메커니즘과 활성화 함수 선택이 시계열 데이터 예측 모델의 성능에 큰 영향을 미친다는 것을 확인하였다. 모든 종속 변수 중 최저 기온(min\_temp)이 모델에 따라 가장 높은 예측 성능을 보였으며, 대기압(pressure)은 가장 큰 예

측 오차를 보였다. 이는 대기압 변수의 높은 변동성이 예측 결과에 크게 영향을 미쳤음을 시사한다.

Table 4. MAE Results for Maximum Temperature (max\_temp)

Model	ReLU	Leaky ReLU	ELU	GELU	Swish	SELU
LSTM (w/ att)	4.967	6.479	5.573	5.972	5.755	6.407
LSTM (w/o att)	5.831	6.177	5.918	5.783	6.511	6.767
BiLSTM (w/ att)	6.287	5.515	6.029	5.646	4.928	5.778
BiLSTM (w/o att)	5.482	5.765	6.076	5.431	5.861	5.165
GRU (w/ att)	5.247	5.585	7.248	6.445	5.371	6.521
GRU (w/o att)	5.367	5.233	6.253	6.193	5.316	6.383
BiGRU (w/ att)	5.571	5.077	4.915	6.925	5.435	5.985
BiGRU (w/o att)	5.100	5.512	5.348	5.854	6.419	5.387

Table 5. RMSE Results for Maximum Temperature (max\_temp)

Model	ReLU	Leaky ReLU	ELU	GELU	Swish	SELU
LSTM (w/ att)	7.393	8.824	7.503	8.209	7.948	8.681
LSTM (w/o att)	8.118	8.263	8.313	8.138	8.721	9.290
BiLSTM (w/ att)	9.018	8.037	8.464	8.171	7.258	8.269
BiLSTM (w/o att)	8.010	8.362	8.623	7.770	8.242	7.557
GRU (w/ att)	7.668	7.994	9.335	8.399	7.547	8.860
GRU (w/o att)	7.854	7.541	8.578	8.400	7.421	8.506
BiGRU (w/ att)	8.750	7.510	7.209	9.346	7.532	8.287
BiGRU (w/o att)	7.449	8.037	7.800	8.107	8.849	7.765

LSTM 모델에서는 Attention이 활성화된 상태에서 ReLU와 SELU를 사용했을 때 가장 좋은 결과를 얻었다. 반면, Attention이 비활성화된 상태에서는 일부 변수를 제외하고 성능이 다소 저하되었다는 점을 확인했다. 특히 SELU는 성능 저하가 두드러졌다. BiLSTM에서는 Attention이 활성화된 상태에서 GELU와 ELU를 사용했을 때 높은 성능을 보였으며, Attention이 비활성화된 상태에서는 ReLU가 상대적으로 우수한 성능을 나타내었다.

GRU 모델은 Attention이 활성화된 상태에서 ReLU와 SiLU가 전반적으로 우수한 성능을 보였으며, GELU는 다

른 활성화 함수에 비해 대기압 예측에서 낮은 성능을 보였다. BiGRU에서는 Attention이 활성화된 상태에서 Leaky ReLU가 최저 기온 예측에서 탁월한 성능(MAE: 2.963, RMSE: 4.028)을 보였고, 대기압에서는 GELU가 가장 높은 오차를 나타내었다.

Table 6. MAE Results for Pressure (pressure)

Model	ReLU	Leaky ReLU	ELU	GELU	Swish	SELU
LSTM (w/ att)	19.434	23.995	25.529	27.332	22.675	16.999
LSTM (w/o att)	19.034	24.494	23.319	30.550	18.741	20.414
BiLSTM (w/ att)	20.678	19.956	18.809	18.463	20.866	27.509
BiLSTM (w/o att)	16.934	25.050	22.350	21.686	28.624	24.499
GRU (w/ att)	15.215	21.616	15.615	37.536	15.827	22.893
GRU (w/o att)	14.304	16.040	19.221	31.179	24.216	14.317
BiGRU (w/ att)	20.676	15.631	16.661	25.378	20.488	19.101
BiGRU (w/o att)	17.860	15.661	17.511	33.691	23.458	20.092

Table 7. RMSE Results for Pressure (pressure)

Model	ReLU	Leaky ReLU	ELU	GELU	Swish	SELU
LSTM (w/ att)	24.371	31.024	34.633	37.298	28.293	21.057
LSTM (w/o att)	23.666	31.241	30.925	42.272	23.059	25.576
BiLSTM (w/ att)	26.358	24.226	23.956	23.557	26.201	34.812
BiLSTM (w/o att)	21.006	31.297	29.728	27.878	36.909	36.026
GRU (w/ att)	19.167	30.255	19.659	53.934	20.161	29.773
GRU (w/o att)	17.639	19.552	24.797	41.667	33.211	16.963
BiGRU (w/ att)	25.516	18.913	21.157	35.597	26.216	25.594
BiGRU (w/o att)	22.898	19.689	23.137	51.600	31.256	25.489

최저 기온은 상대적으로 작은 변동 폭과 일정한 주기성을 가지고 있어 모델이 더 효과적으로 학습할 수 있었다. 반면, 대기압은 다양한 외부 요인에 의해 큰 변동을 겪는 변수로, 모델이 이를 학습하기에 한계가 있었다. 또한, 선형 보간법을 사용한 데이터 전처리 과정이 비선형 데이터 특성에 부정적인 영향을 미쳤다는 점을 확인했다. 이를 통해 복잡한 변수에 대한 처리 능력을 향상할 수 있는 모델 선택과 추가적인 데이터 정규화와 전처리 방법이 모델 성능 향상에 이바지할 수 있을 것으로 예상된다.

## V. Conclusions

본 연구에서는 다양한 활성화 함수와 RNN 모델을 적용하여 화성의 날씨 패턴을 예측하는 실험을 수행하였다. 그 결과, GRU와 BiGRU 모델이 탁월한 성능을 보이며, 특히 최저 기온(min\_temp) 예측에서 가장 낮은 오차를 달성함으로써 날씨 패턴 예측에 효과적임을 입증하였다. 이들 모델은 LSTM 및 BiLSTM에 비해 연산 비용이 적고 학습 효율이 높아, 데이터의 비선형성과 시계열의 복잡한 패턴을 효과적으로 처리할 수 있었다.

선정된 활성화 함수 중에서 SiLU와 Leaky ReLU는 모든 모델에 걸쳐 일관되게 우수한 성능을 보여주었다. 특히 SiLU는 최저 기온 예측에서 눈에 띄는 결과를 제공하였으며, Leaky ReLU는 그 견고함을 통해 과적합을 방지하고 모델의 일반화 능력을 향상하였다. 이는 효율적인 활성화 함수의 선택이 모델의 예측 정확도와 성능을 유지하는 데 결정적인 역할을 하였다는 것을 보여준다.

본 연구를 통해 화성 날씨 데이터의 복잡성과 변동성을 효과적으로 모델링할 수 있는 최적의 방법을 모색할 수 있었으며, 이는 향후 화성 탐사 임무의 성공적인 수행에 이바지할 수 있는 중요한 기술적 진보로 평가된다. 향후 GRU 모델을 추가로 개선하고 다양한 하이브리드 모델 또는 고급 신경망 아키텍처를 통합하여 더욱 정밀하고 신뢰할 수 있는 날씨 예측 모델을 개발할 계획이다. 이를 통해 화성과 같은 외계 행성의 극한 환경에서의 생존 및 탐사 활동을 더욱 효과적으로 지원하는 기술적 기반을 마련하고자 한다.

이와 더불어 GPR(Gaussian Process Regression) 등 비선형 보간법을 활용하여 결측치를 정교하게 보완하고, LD(Local Discrepancy) 기반의 이상치 처리 기법을 적용하여 시계열 데이터의 국지적인 불일치를 측정할 계획이다. 또한, 비선형적 데이터 변동성을 효과적으로 처리하는 방법을 도입하여 모델의 안정성과 정확성을 향상하고, 복잡한 변수의 예측 성능을 개선할 예정이다.

## ACKNOWLEDGEMENT

This study was supported by MSIT (Ministry of Science, ICT), Korea, under the National Program for Excellence in SW, supervised by IITP (Institute of Information & Communications Technology Planning & Evaluation) in 2024 (2021-0-01399).

## REFERENCES

- [1] K. Crosswell, "Magnificent Mars," Simon and Schuster, 2003.
- [2] C. P. McKay, "The search for life on Mars," *Origins of Life and Evolution of the Biosphere*, vol. 27, no. 1, pp. 263–289, Jun. 1997. DOI: 10.1023/A:1006500116990
- [3] R. Pyle, "Space 2.0: How private spaceflight, a resurgent NASA, and international partners are creating a new space age," BenBella Books, 2019.
- [4] S. Singh, P. Singh, S. Rangabhashiyam, and K. K. Srivastava, "Global Climate Change," Elsevier, 2021.
- [5] B. L. Ehlmann et al., "The sustainability of habitability on terrestrial planets: Insights, questions, and needed measurements from Mars for understanding the evolution of Earth-like worlds," *Journal of Geophysical Research: Planets*, vol. 121, no. 10, pp. 1927–1961, Sep. 2016. DOI: 10.1002/2016JE005134
- [6] P. L. Read, S. R. Lewis, and D. P. Mulholland, "The physics of Martian weather and climate: a review," *Reports on Progress in Physics*, vol. 78, no. 12, p. 125901, Nov. 2015. DOI: 10.1088/0034-4885/78/12/125901
- [7] A. Barjasteh, S. H. Ghafouri, and M. Hashemi, "A hybrid model based on discrete wavelet transform (DWT) and bidirectional recurrent neural networks for wind speed prediction," *Engineering Applications of Artificial Intelligence*, vol. 127, p. 107340, Jan. 2024. DOI: 10.1016/j.engappai.2023.107340
- [8] Z. Yuan, Z. Yang, Y. Ling, C. Wu, and C. Li, "Spatiotemporal attention mechanism-based deep network for critical parameters prediction in chemical process," *Process Safety and Environmental Protection*, vol. 155, pp. 401–414, Nov. 2021. DOI: 10.1016/j.psep.2021.09.024
- [9] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, Sep. 2021. DOI: 10.1016/j.neucom.2021.03.091
- [10] J. Pla-García et al., "Meteorological predictions for Mars 2020 Perseverance Rover landing site at Jezero crater," *Space Science Reviews*, vol. 216, p. 148, Dec. 2020. DOI: 10.1007/s11214-020-00763-x
- [11] I. Priyadarshini and V. Puri, "Mars weather data analysis using machine learning techniques," *Earth Science Informatics*, vol. 14, pp. 1885–1898, Dec. 2021. DOI: 10.1007/s12145-021-00643-0
- [12] P. Pant et al., "Machine Learning Techniques for Analysis of Mars Weather Data," *Proceedings of the 15th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pp. 1–7, Bucharest, Romania, Jun. 2023. DOI: 10.1109/ECAI58194.2023.10194233.
- [13] J. Moon, Y. Han, H. Chang, and S. Rho, "Multistep-ahead solar irradiance forecasting for smart cities based on LSTM, Bi-LSTM, and GRU neural networks," *The Journal of Society for e-Business Studies*, vol. 27, no. 4, pp. 27–52, Nov. 2022. DOI: 10.7838/jsebs.2022.27.4.027
- [14] J. Moon, S. Park, S. Rho, and E. Hwang, "A comparative analysis of artificial neural network architectures for building energy consumption forecasting," *International Journal of Distributed Sensor Networks*, vol. 15, no. 9, p. 1550147719877616, Sep. 2019. DOI: 10.1177/1550147719877616
- [15] S. Jung, J. Moon, S. Park, and E. Hwang, "A probabilistic short-term solar radiation prediction scheme based on attention mechanism for smart island," *KIISE Transactions on Computing Practices*, vol. 25, no. 12, pp. 602–609, Dec. 2019. DOI: 10.5626/KTCP.2019.25.12.602
- [16] S. Jung, J. Moon, S. Park, and E. Hwang, "An attention-based multilayer GRU model for multistep-ahead short-term load forecasting," *Sensors*, vol. 21, no. 5, p. 1639, Feb. 2021. DOI: 10.3390/s21051639
- [17] D. Atri, N. Abdelmoneim, D. B. Dhuri, and M. Simoni, "Diurnal variation of the surface temperature of Mars with the Emirates Mars Mission: a comparison with Curiosity and Perseverance rover measurements," *Monthly Notices of the Royal Astronomical Society: Letters*, vol. 518, no. 1, pp. L1–L6, Oct. 2022. DOI: 10.1093/mnrasl/slac094
- [18] C. D. Xu, J. F. Wang, M. G. Hu, and Q. X. Li, "Interpolation of missing temperature data at meteorological stations using P-BSHADE," *Journal of Climate*, vol. 26, no. 19, pp. 7452–7463, Oct. 2013. DOI: 10.1175/JCLI-D-12-00633.1
- [19] A. Gökhan, C. O. Güzeller, and M. T. Eser, "The effect of the normalization method used in different sample sizes on the success of artificial neural network model," *International Journal of Assessment Tools in Education*, vol. 6, no. 2, pp. 170–192, Jul. 2019. DOI: 10.21449/ijate.479404
- [20] M. Alizamir et al., "Improving the accuracy of daily solar radiation prediction by climatic data using an efficient hybrid deep learning model: Long short-term memory (LSTM) network coupled with wavelet transform," *Engineering Applications of Artificial Intelligence*, vol. 123, p. 106199, Aug. 2023. DOI: 10.1016/j.engappai.2023.106199
- [21] M. Bukhari, S. Yasmin, S. Naz, M. Y. Durrani, M. Javaid, J. Moon, and S. Rho, "A smart heart disease diagnostic system using deep vanilla LSTM," *Computers, Materials & Continua*, vol. 77, no. 1, pp. 1251–1279, Oct. 2023. DOI: 10.32604/cmc.2023.040329
- [22] M. J. Gul, G. M. Urfa, A. Paul, J. Moon, S. Rho, and E. Hwang, "Mid-term electricity load prediction using CNN and Bi-LSTM," *The Journal of Supercomputing*, vol. 77, pp. 10942–10958, Oct. 2021. DOI: 10.1007/s11227-021-03686-8
- [23] B. Lee, S. Kim, M. Maqsood, J. Moon, and S. Rho, "Advancing autoencoder architectures for enhanced anomaly detection in multivariate industrial time series," *Computers, Materials & Continua*, vol. 81, no. 1, pp. 1275–1300, Oct. 2024. DOI: 10.32604/cmc.2024.040329



10.32604/cmc.2024.054826

- [24] D. So, J. Oh, I. Jeon, J. Moon, M. Lee, and S. Rho, "BiGTA-Net: A hybrid deep learning-based electrical energy forecasting model for building energy management systems," *Systems*, vol. 11, no. 9, p. 456, Sep. 2023. DOI: 10.3390/systems11090456
- [25] D. Soydaner, "Attention mechanism in neural networks: where it comes and where it goes," *Neural Computing and Applications*, vol. 34, pp. 13371-13385, Aug. 2022. DOI: 10.1007/s00521-022-07366-3
- [26] B. Ouyang, Y. Song, Y. Li, G. Sant, and M. Bauchy, "EBOD: An ensemble-based outlier detection algorithm for noisy datasets," *Knowledge-Based Systems*, vol. 231, p. 107400, Nov. 2021. DOI: 10.1016/j.knosys.2021.107400
- [27] J. Kim, J. Moon, E. Hwang, and P. Kang, "Recurrent inception convolution neural network for multi short-term load forecasting," *Energy and Buildings*, vol. 194, pp. 328-341, Jul. 2019. DOI: 10.1016/j.enbuild.2019.04.034
- [28] S. Y. Şen and N. Özkurt, "Convolutional Neural Network Hyperparameter Tuning with Adam Optimizer for ECG Classification," *Proceedings of the 2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, pp. 1-6, Istanbul, Turkey, Oct. 2020. DOI: 10.1109/ASYU50717.2020.9259896.
- [29] T. Miseta, A. Fodor, and Á. Vathy-Fogarassy, "Surpassing early stopping: A novel correlation-based stopping criterion for neural networks," *Neurocomputing*, vol. 567, p. 127028, Jan. 2024. DOI: 10.1016/j.neucom.2023.127028
- [30] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Computer Science*, vol. 7, p. e623, Jul. 2021. DOI: 10.7717/peerj-cs.623

## Authors



Jaehyeok Jo has been a dedicated student in the Department of AI and Big Data at Soonchunhyang University, South Korea, since 2019. He will receive his Bachelor of Science degree in February next year.

His academic interests include time-series forecasting, deep learning, and emerging technologies that integrate these fields to address innovative challenges.



Yunho Sin, a dedicated second grader at Asan Middle School in Asan, South Korea, has a deep passion for software and artificial intelligence and aspires to make a significant contribution to these innovative fields.



Bo-Young Kim, a key information science educator at Asan Middle School, South Korea, since 2002, earned her Bachelor's and Master's degrees in Information Science and Mathematics Education from Hoseo University

in 2002 and 2008, respectively. Ms. Kim is deeply committed to cultivating outstanding students through advanced information education and championing sustainable software training.



Jihoon Moon earned his Ph.D. in Electrical Engineering from Korea University in 2021 and worked as a postdoctoral researcher at Chung-Ang University until 2022. He has been an assistant professor in the Department

of AI and Big Data at Soonchunhyang University since 2022. Dr. Moon's research encompasses a broad spectrum, including advanced data mining, information extraction, and time-series analysis, with a keen interest in applying these methods to industrial innovations. His focus extends to sustainable solutions, advanced machine learning, deep learning applications, and explainable AI, with a special emphasis on optimizing energy management and forecasting.