

Deep Learning Model Experiment for Anomaly Detection of Heat Supply Chain based on Big Data

학번, 이름 : 21110690 오운서

Email : 21110690@sju.ac.kr

Deep Learning Project Proposal

목차

- I. Goal
- II. System Environment
- III. System Architecture
- IV. System Result
- V. Reference

I. Goal

- 열 공급망은 중앙의 난방 기관에서 한 지역 내의 수용가에 온수나 증기를 보내는 열 공급 방식
- 규모가 매우 크므로, 한 번 이상상황이 발생하면 그 피해에 따른 경제적 손실이 막대함
- 혹한의 날씨, 수격 작용(water hammer), 부품 노후화(밸브의 부식, 노후 배관) 등의 이유로 많은 고장이 발생함
- 최근 대한민국 수도권의 20년 넘은 노후 배관 비율이 26%를 넘음
- 그러나 빅데이터 기반 열 공급망의 고장을 사전에 감지하는 방법이 부족
- 빅데이터 기반 열 공급망에서의 이상을 딥러닝 모델을 통해 사전에 감지하여 고장을 방지하는 것을 목표로 함

II. System Environment

- **Dataset**

- 한국지역난방공사 열판매량(2014_2018 시간별 청주 데이터)

- **Preference**

- Intel® Core™ i5-10400 CPU 2.90GHz
- NVIDIA GeForce RTX 2080 Ti

Date	Day	Time	Cheongju
01/01/2014	수	1	136
01/01/2014	수	2	127
01/01/2014	수	3	123
01/01/2014	수	4	122
01/01/2014	수	5	123

▲ Dataset의 형태

II. System Environment

- **Library**

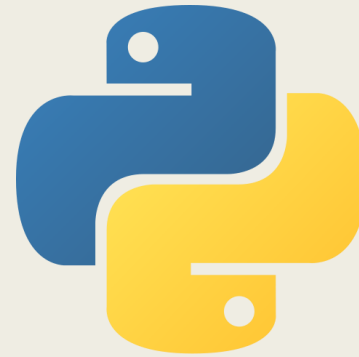
- keras
- Tensorflow
- Sklearn
- ...

- **Programming Language**

- Python 3.x

- **IDE**

- Pycharm
- Colab-google



III. System Architecture

1) 한국지역난방공사 전체 데이터의 5%를 인공 이상데이터로 대체

	Date	Cheongju	Anomaly	Label
89	2014-01-04 18:00	99	99	0
90	2014-01-04 19:00	133	133	0
91	2014-01-04 20:00	142	142	0
92	2014-01-04 21:00	156	156	0
93	2014-01-04 22:00	159	238.5	1

- a. Cheongju != Anomaly, 이상이므로 '1'
- b. Cheongju == Anomaly, 정상이므로 '0'

2) 24시간의 주기와 12시간의 간격으로 [평균, 최대값, 최소값, 분산, 중앙값]을 생성 → 모델의 Input Data로 사용

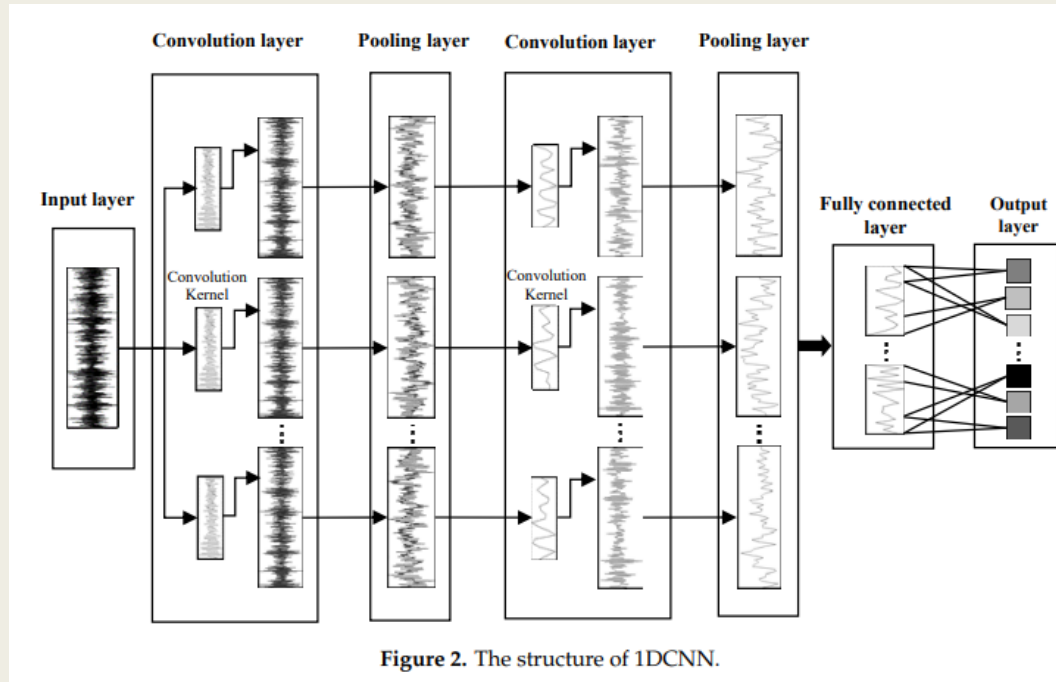
Input Data

	Mean	Min	Max	Variance	Median	Label
43	146.4167	84	207	1509.993	150	0
44	142.1667	81	201	1555.971	150	0
45	137.0833	81	201	1414.08	136	0
46	130.2917	87	164	739.5199	135	0
47	115.856	24.544	163	929.2914	112	1

Target Data

III. System Architecture

3.1) 1DCNN



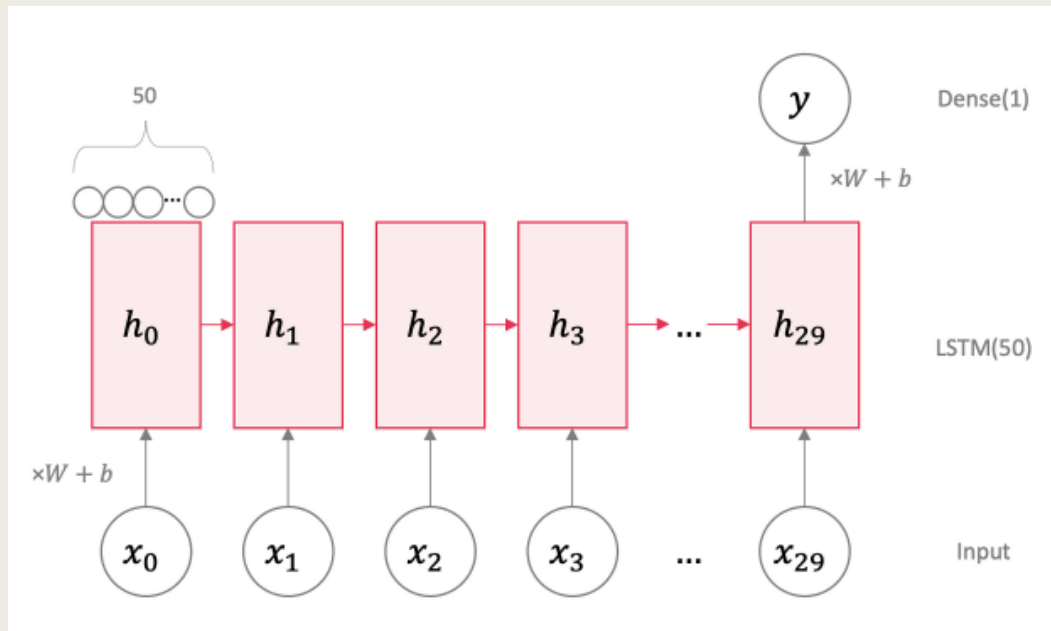
▲ 1DCNN 구조도

- 시계열 데이터를 사용하였기 때문에 일반적으로 많이 사용하는 2D CNN이 아닌 1D-kernel CNN 사용
- 1DCNN 모델 Architecture

```
model_cnn = Sequential()
model_cnn.add(Conv1D(filters = 64, kernel_size = 2,
                    activation = 'relu', input_shape =
                    (x_train1_cnn.shape[1],
                     x_train1_cnn.shape[2])))
model_cnn.add(MaxPooling1D(pool_size = 2))
model_cnn.add(Flatten())
model_cnn.add(Dense(50, activation = 'relu'))
model_cnn.add(Dense(2, activation = "relu"))
model_cnn.compile(loss = 'mse', optimizer = adam)
```

III. System Architecture

3.2) LSTM(Long Short-Term Memory)



▲ LSTM(Long Short-Term Memory) 구조도

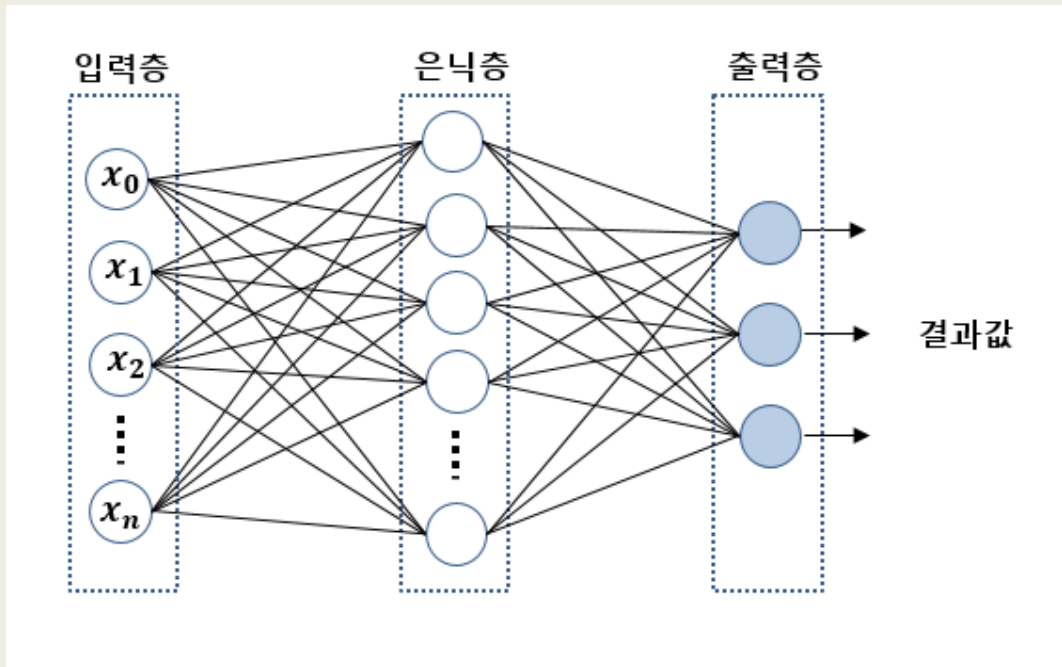
- LSTM은 RNN의 한 종류로, 시계열 데이터에서 가지고 있는 특유의 긴 sequence에 걸쳐 있는 패턴을 추출하는 데에 잘 작동하므로 사용

LSTM 모델 Architecture

```
model_lstm = Sequential()  
model_lstm.add(LSTM(50, activation = "relu",  
                    input_shape = (x_train1_lstm.shape[1],  
                                    x_train1_lstm.shape[2])))  
model_lstm.add(Dense(1, activation = "relu"))  
Model_lstm.compile(loss = 'mse', optimizer = adam)
```


III. System Architecture

3.3) MLP(다층 퍼셉트론)



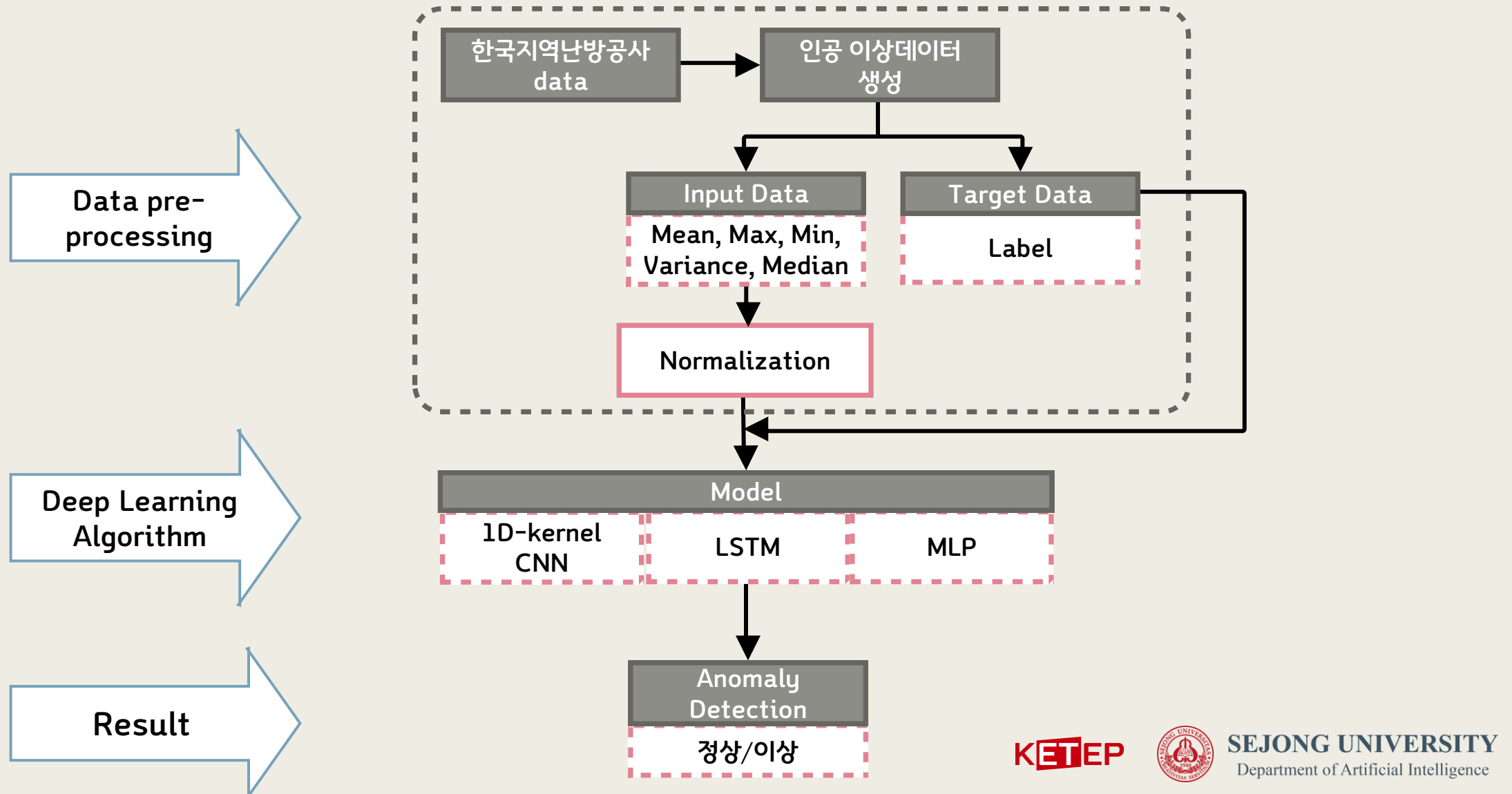
▲ MLP(다층 퍼셉트론) 구조도

- 일반적 딥러닝 방법을 사용하여 다른 모델과의 결과 비교 위해 사용

- MLP 모델 Architecture

```
model_mlp = Sequential()  
model_mlp.add(Dense(100, activation = 'relu',  
                    input_shape = (x_train1_lstm.shape[1],  
                                    x_train1_lstm.shape[2])))  
model_mlp.add(Dense(1))  
model_mlp.compile(loss = 'mse', optimizer = adam)
```

System Workflow



IV. System Result

- 3가지 모델의 결과 비교
 - 총 100번의 테스트 시행
 - 인공 이상데이터를 생성한 후 여러가지 딥러닝 모델을 통해 이상 감지의 정확도를 확인함
 - 3가지 모델이 비슷한 결과를 냄

Accuracy \ Model		1D-kernel CNN	LSTM	MLP
Test Accuracy(%)	Average	78.695	80.018	79.125
	Variance	4.961	2.367	0.6945

V. Reference

[한국지역난방 열생산량 정보 데이터]

(<https://www.data.go.kr/data/15044420/openapi.do>)

[Time series CNN Anomaly Detection img]

(<https://www.semanticscholar.org/paper/Time-Series-Anomaly-Detection-Using-Convolutional-Wen-Keyes/0017aeb0c049a383d962399d26100ec2bf5cc7c7>)

[Deep Learning for Time Series]

(<https://leedakyeong.tistory.com/entry/Deep-Learning-for-Time-Series-Forecasting-kaggle>)

V. Reference

- [1] 전준서, “정적 구조 해석을 이용한 복합하중 및 경계조건에 따른 열수송관의 안전성 평가,” *The Society of Convergence Knowledge Transactions*, vol. 9, no. 2, pp. 1-10, Apr. 2021.
- [2] P. Lipcák, M. Macak and B. Rossi, “Big Data Platform for Smart Grids Power Consumption Anomaly Detection,” in *Proc. FedCSIS*, Leipzig, Germany, Sep. 2019, pp. 771-780.
- [3] L. Wen, K. Zhou, S. Yang and L. Li, “Compression of smart meter big data: A survey,” *Renewable and Sustainable Energy Reviews*, vol. 91, pp. 59-69, Aug. 2018.
- [4] Z. Sun, H. Jin, J. Gu, Y. Huang, X. Wang and X. Shen, “Gradual fault early-stage diagnosis for air source heat pump system using deep learning techniques,” *International Journal of Refrigeration*, vol. 107, pp. 63-72, Nov. 2019.
- [5] J.L. Casteleiro-Roca, H. Quintián, J.L. Calvo-Rolle, E. Corchado, M.C. Meizoso-López and A. Piñón-Pazos, “An intelligent fault detection system for a heat pump installation based on a geothermal heat exchanger,” *Journal of Applied Logic*, vol. 17, pp. 36-47, Sep. 2016.
- [6] S. Huang, J. Tang, J. Dai and Y. Wang, “Signal Status Recognition Based on 1DCNN and Its Feature Extraction Mechanism Analysis,” *Sensors*, vol.19, no. 9, pp.1-19, April. 2019.
- [7] S. Narmini, N. Tavakoli and A. Namin, “The Performance of LSTM and BiLSTM in Forecasting Time Series,” *IEEE international conference on big data(big data)*, pp.3285-3292, 2019
- [8] I. R. Widiyari, L. E. Nugroho and Widyawan, "Deep learning multilayer perceptron (MLP) for flood prediction model using wireless sensor network-based hydrology time series data mining", *2017 International Conference on Innovative and Creative Information Technology (ICITech)*, pp. 1-5, Nov. 2017.

Thank You