



Python

**주식 예측 프로그램
파이썬으로 만들어보기**

“처음에 계획했던 방법”

처음 계획 (작동 원리)

YOLOv5

YOLO v5

“You Only Look Once”

CNN 딥러닝 모델을 이용한

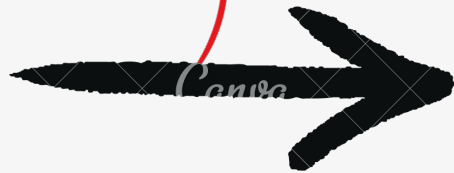
이미지 기반 실시간 객체 검출 시스템

처음 계획 (작동 원리)

YOLOv5

YOLO v5

PyTorch



ONNX

ONNX

“You Only Look Once”

CNN 딥러닝 모델을 이용한

이미지 기반 실시간 객체 검출 시스템

“Open Neural Network Exchange”

다양한 딥러닝 프레임워크 간에 모델

을 호환할 수 있게 해주는 표준 형식

처음 계획 (작동 원리)

YOLOv5

YOLO v5



ONNX

ONNX



AI 추론하기

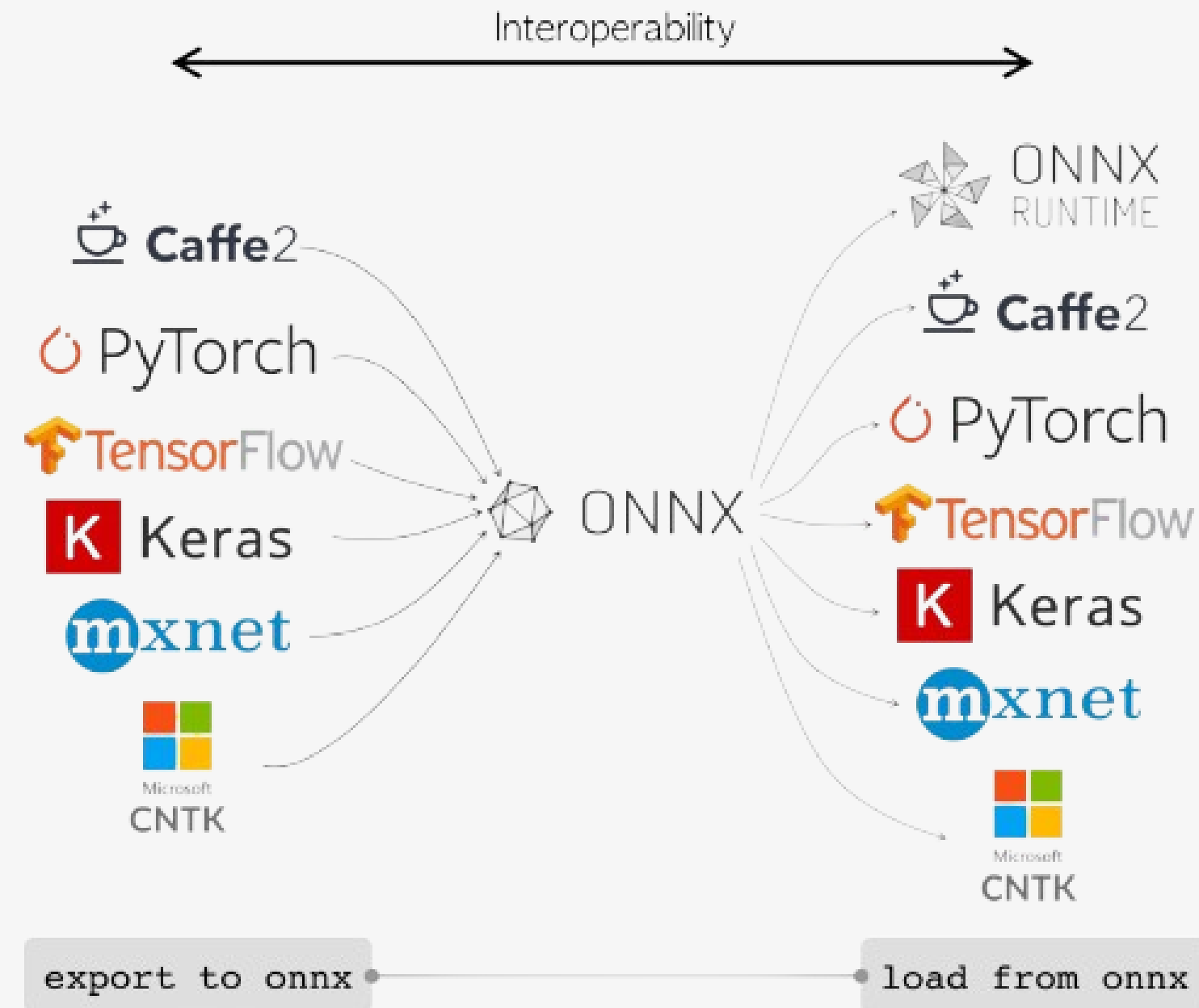
“You Only Look Once”
CNN 딥러닝 모델을 이용한
이미지 기반 실시간 객체 검출 시스템

“Open Neural Network Exchange”
다양한 딥러닝 프레임워크 간에 모델
을 호환할 수 있게 해주는 표준 형식

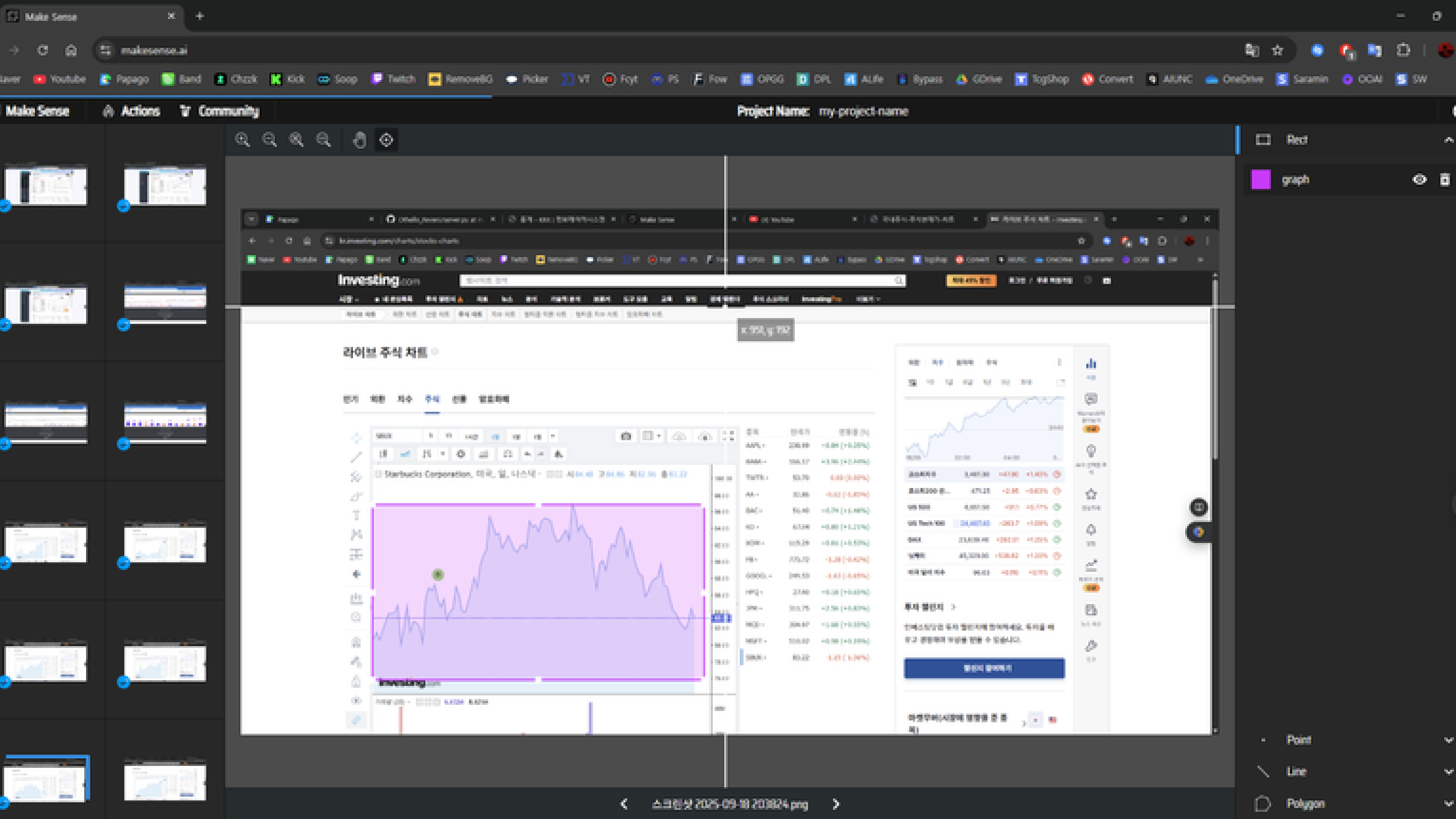
주가 그래프 이미지 분석해
시각적 패턴에서 규칙을 찾기

왜 ONNX 파일로 변환할까?

왜 ONNX 파일로 변환할까?



- 어디서 학습했든 다른 환경에서도 쉽게 사용할 수 있음
- 배포 및 통합 관리 용이성



```

1 import sys
2 import numpy as np
3 import onnxruntime as ort
4 import cv2
5 from pathlib import Path
6 import time
7 import os
8 import math
9
10 labels_folder = Path("labels")
11 models_folder = Path("models")
12
13 # 폴더 존재 여부 확인
14 if not labels_folder.exists():
15     sys.exit(f"Labels folder not found: {labels_folder}")
16 if not models_folder.exists():
17     sys.exit(f"Models folder not found: {models_folder}")
18
19 features_list = []
20
21 for file in labels_folder.iterdir():
22     if file.suffix == ".txt":
23         with open(file, "r") as f:
24             for line in f:
25                 parts = list(map(float, line.strip().split()))
26                 features = parts[1:] # 첫번째 값은 경계, 나머지가 feature
27                 features_list.append(features)
28
29 features_array = np.array(features_list, dtype=np.float32) # shape (N, feature_size)
30 batch_size, feature_size = features_array.shape
31 print(f"Loaded features: {features_array.shape}")
32
33 onnx_files = [f for f in models_folder.iterdir() if f.suffix == ".onnx"]
34 if not onnx_files:
35     sys.exit("No ONNX models found in the models folder.")
36
37 onnx_sessions = [ort.InferenceSession(f, providers=["CPUExecutionProvider"]) for f in onnx_files]
38 print(f"Loaded {len(onnx_sessions)} ONNX models.")
39
40 def pulsar_avg(values, cut_ratio=0.15):
41     sorted_values = np.sort(values)
42     n = len(sorted_values)
43     k = int(n * cut_ratio)
44     if n - 2 * k <= 0:
45         return np.mean(sorted_values)
46     return np.mean(sorted_values[k:n - k])
47
48 prediction_table = []
49 max_table = 100
50 print("Press ESC to exit.")
51
52 while True:
53     if cv2.waitKey(100) & 0xFF == 27: # ESC 키
54         break
55
56     model_preds = []
57     for session in onnx_sessions:
58         input_name = session.get_inputs()[0].name
59         pred = session.run(None, {input_name: features_array})[0] # shape (N,1)
60         model_preds.append(pred.flatten())
61
62     # 오답률 평균
63     model_preds = np.array(model_preds)
64     avg_pred = model_preds.mean(axis=0) # shape (N,)
65
66     # 결사평균
67     smooth_pred = pulsar_avg(avg_pred)
68     prediction_table.append(smooth_pred)
69     if len(prediction_table) > max_table:
70         prediction_table.pop(0)
71
72     img = np.zeros((600, 800, 3), dtype=np.uint8)
73     for i in range(1, len(prediction_table)):
74         x1, y1 = int((i-1)*(800/max_table)), int(600 - prediction_table[i-1]*600)
75         x2, y2 = int(i*(800/max_table)), int(600 - prediction_table[i]*600)
76         cv2.line(img, (x1, y1), (x2, y2), (0, 0, 255), 2)
77     cv2.putText(img, f"Average: {smooth_pred:.3f}", (10, 30),
78                 cv2.FONT_HERSHEY_SIMPLEX, 1, (0, 255, 0), 2)
79
80     cv2.imshow("Stock Prediction", img)
81
82 cv2.destroyAllWindows()

```

```

1 onnx_sessions = [ort.InferenceSession(f, providers=["CPUExecutionProvider"]) for f in onnx_files]
2 print(f"Loaded {len(onnx_sessions)} ONNX models.")
3
4 def pulsar_avg(values, cut_ratio=0.15):
5     sorted_values = np.sort(values)
6     n = len(sorted_values)
7     k = int(n * cut_ratio)
8     if n - 2 * k <= 0:
9         return np.mean(sorted_values)
10    return np.mean(sorted_values[k:n - k])
11
12 prediction_table = []
13 max_table = 100
14 print("Press ESC to exit.")
15
16 while True:
17     if cv2.waitKey(100) & 0xFF == 27: # ESC 키
18         break
19
20     model_preds = []
21     for session in onnx_sessions:
22         input_name = session.get_inputs()[0].name
23         pred = session.run(None, {input_name: features_array})[0] # shape (N,1)
24         model_preds.append(pred.flatten())
25
26     # 오답률 평균
27     model_preds = np.array(model_preds)
28     avg_pred = model_preds.mean(axis=0) # shape (N,)
29
30     # 결사평균
31     smooth_pred = pulsar_avg(avg_pred)
32     prediction_table.append(smooth_pred)
33     if len(prediction_table) > max_table:
34         prediction_table.pop(0)
35
36     img = np.zeros((600, 800, 3), dtype=np.uint8)
37     for i in range(1, len(prediction_table)):
38         x1, y1 = int((i-1)*(800/max_table)), int(600 - prediction_table[i-1]*600)
39         x2, y2 = int(i*(800/max_table)), int(600 - prediction_table[i]*600)
40         cv2.line(img, (x1, y1), (x2, y2), (0, 0, 255), 2)
41     cv2.putText(img, f"Average: {smooth_pred:.3f}", (10, 30),
42                 cv2.FONT_HERSHEY_SIMPLEX, 1, (0, 255, 0), 2)
43
44     cv2.imshow("Stock Prediction", img)
45
46 cv2.destroyAllWindows()

```

C:\WINDOWS\py.exe



Loaded features: (47, 4)
Loaded 32 ONNX models.
Press ESC to exit.

Stock Prediction



Average: 0.086

왜 잘 작동하지 않았을까?

데이터 특성과 모델 불일치

- YOLO는 이미지 객체 탐지에 최적화되어 있음
- 주식 데이터는 시계열 데이터로, 연속적 변화와 패턴의 값을 구해야함
- 이미지 기반 접근은 시간적 흐름을 반영할 수 없음 → 정확한 예측 어려움

정보 손실 문제

- 작은 변동이나 추세를 모델이 제대로 학습하기 어려움
- 그래프를 이미지로 변환하면, 숫자 값 자체의 정밀한 변화가 손실됨

“개선한 방법”

개선된 방법 (작동 원리)



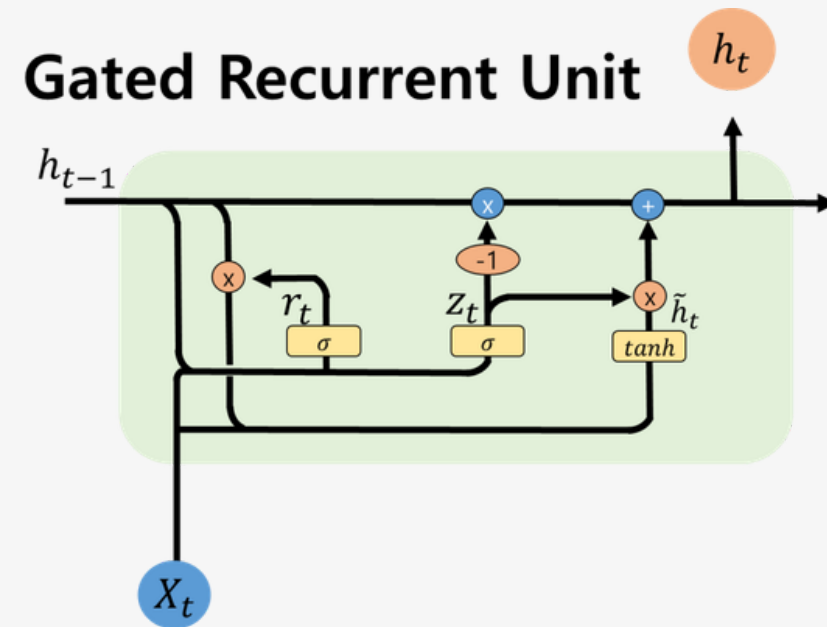
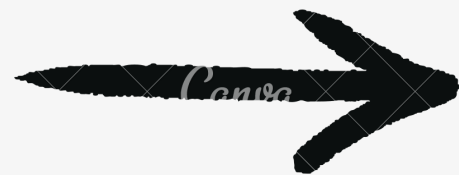
엑셀(CSV) 파일

주식 사이트에서 제공하는 주가
데이터를 담은 엑셀 파일에서
날짜별 시세, 거래량 등의 정보 추출

개선된 방법 (작동 원리)



엑셀(CSV) 파일



GRU

주식 사이트에서 제공하는 주가
데이터를 담은 엑셀 파일에서
날짜별 시세, 거래량 등의 정보 추출

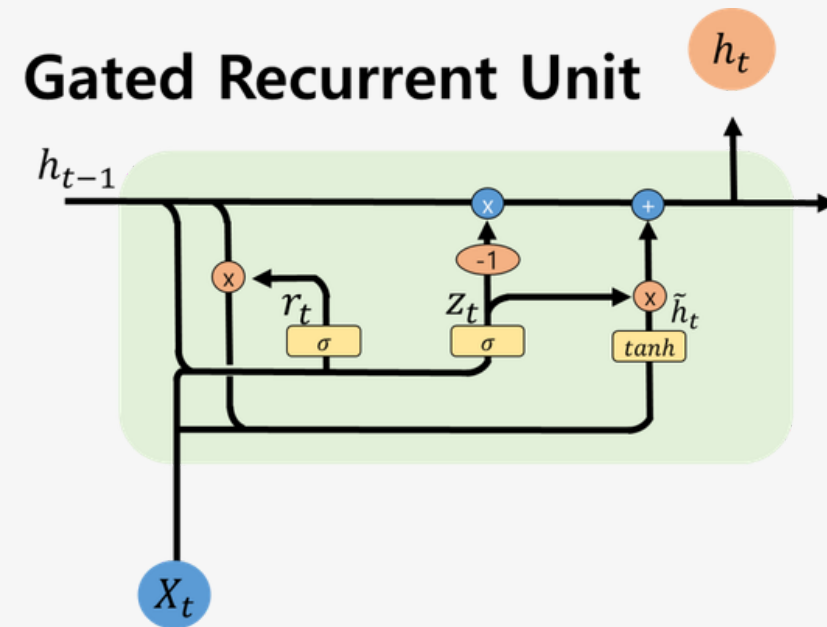
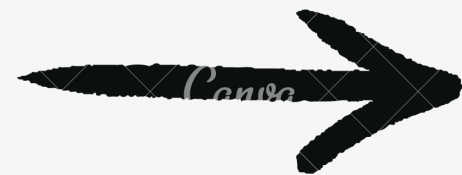
“Gated Recurrent Unit”
시계열 데이터 예측에 강한 순환
신경망(RNN) 모델

개선된 방법 (작동 원리)



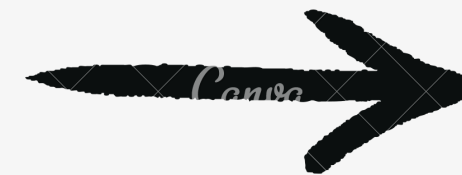
엑셀(CSV) 파일

주식 사이트에서 제공하는 주가 데이터를 담은 엑셀 파일에서 날짜별 시세, 거래량 등의 정보 추출



GRU

“Gated Recurrent Unit” 시계열 데이터 예측에 강한 순환 신경망(RNN) 모델



AI 추론하기

Numpy, Pandas, GRU를 이용해 주가 데이터로 미래 시세 추론

```
1 import sys
2 import msvcrt
3 import math
4 import numpy as np
5 import onnxruntime as ort
6 import cv2
7 from PIL import Image, ImageGrab
8 import os
9 from pathlib import Path
10 import time
11 import torch
12 import torch.nn as nn
13 from sklearn.preprocessing import MinMaxScaler
14 import matplotlib.pyplot as plt
15 import pandas as pd
16 import plotext as plot
17 import json
18
19 stock_ai_ascii_art = r"""
20 /$$$$$$ /$$ /$$ /$$$$$$ /$$$$$ /$$$$$$
21 /$$_ $$ | $$ /$$_ $$ /$$_ $$ /$$_ $$ /
22 | $$ \_//$$$$$$ /$$$$$$ /$$$$$$ | $$ \ $$ | $$
23 | $$$$$$|_ $$ / /$$_ $$ /$$_ / | $$ /$$/ | $$$$$$ | $$
24 \_ $$ | $$ | $$ \ $$ | $$ | $$$$$$/ | $$_ $$ | $$
25 /$$ \ $$ | $$ /$$ | $$ | $$ | $$ | $$_ $$ | $$ | $$ | $$
26 | $$$$$$/ | $$$$/ | $$$$$$/ | $$$$$$ | $$ \ $$ | $$ /$$$$$$
27 \_ / \_ / \_ / \_ / \_ / \_ /
28 ---
29
30 o_ascii_art = r"""
31 /$$$$$
32 /$$_ $$
33 | $$ \ $$
34 | $$ | $$
35 | $$ | $$
36 | $$ | $$
37 | $$$$$/
38 \_ /
39 ---
40
41 x_ascii_art = r"""
42 /$$ /$$
43 | $$ / $$
44 | $$/ $$/
45 \ $$$$/
46 >$$ $$
47 /$$/\ $$
48 | $$ \ $$
49 ---
50
51 sys.stdout.write(stock_ai_ascii_art + "\n")
52
53 # ex)나 실행 디렉토리로 실행할때 오류 나지 않게 하기
54 def resource_path(relative_path):
55     # ex로 실행할때
56     if getattr(sys, 'frozen', False):
57         base_path = Path(sys.executable).parent
58     else:
59         # 일반 Python
60         base_path = Path(__file__).parent
61     return base_path / relative_path
62
63 # 폴더 경로 설정
64 csv_folder_parent_folder_path = resource_path("csv")
65 models_folder_path = resource_path("models")
66 config_folder_path = resource_path("config")
67 settings_file_path = config_folder_path / "settings.json"
68
69 # 'settings.json' 기존 설정 파일
70 settings_data = {
71     "graph": True,
72     "graph_visualization": "matplotlib.pyplot",
73     "trimmed_mean_percentage": 1,
74     "matplotlib_pyplot_asynchronous": False,
75     "br": "br",
76     "kor_comment_graph": "예상 그래프를 보여줄지",
77     "kor_comment_graph_visualization": "plotext: cmd 안에서 줄이나 선으로, matplotlib.pyplot: 외부 창에서",
78     "kor_comment_trimmed_mean_percentage": "공시평균값 (10이면 상위/하위 10% 제거)",
79     "kor_comment_matplotlib_pyplot_asynchronous": "비동기 사용여부 (사용안함: False/ 사용함: (몇 분 후에 창이 꺼질지))",
80     "bl": "bl",
81     "eng_comment_graph": "whether to show the predicted graph",
82     "eng_comment_graph_visualization": "plotext: dots and lines in the CMD, matplotlib.pyplot: in an another external window",
83     "eng_comment_trimmed_mean_percentage": "Trimmed mean value (10 means removing the top and bottom 10%)",
84     "eng_comment_matplotlib_pyplot_asynchronous": "Use asynchronous mode (False: disabled / [minutes]: enabled, window will close after"
85 }
86
87 # 'config_folder', 'settings.json' 없을시 생성, 있으면 계속하기
88 if not config_folder_path.exists():
89     sys.stdout.write("'config folder' does not exist.\n")
90     sys.stdout.write(f"Creating 'config folder' into: {config_folder_path}\n")
91     config_folder_path.mkdir(parents=True)
92     with open(settings_file_path, "w", encoding="utf-8") as f:
93         json.dump(settings_data, f, ensure_ascii=False, indent=4)
94     sys.stdout.write(f"Automatically creating 'settings folder' into: {settings_file_path}\n")
95 elif not settings_file_path.exists():
96     sys.stdout.write("'config folder' found!\n")
97     sys.stdout.write(f"'config folder' location: {config_folder_path}\n")
98     with open(settings_file_path, "w", encoding="utf-8") as f:
99         json.dump(settings_data, f, ensure_ascii=False, indent=4)
100
101 # Launchpad
102 Run Testcases
103 0 0 0
104 Discord RPC
105 Connected to Discord
```

 University of Cambridge

C:\Users\Unknown User\Desktop> Stock_All > server.py > ...

```

106 # 'config_folder', 'settings.json' 없으면 생성, 있으면 계속하기
107 if not config_folder_path.exists():
108     sys.stdout.write("'config folder' does not exist.\n")
109     sys.stdout.write(f"Creating 'config folder' into: {config_folder_path}\n")
110     config_folder_path.mkdir(parents=True)
111     with open(settings_file_path, "w", encoding="utf-8") as f:
112         json.dump(settings_data, f, ensure_ascii=False, indent=4)
113     sys.stdout.write(f"Automatically creating 'settings folder' into: {settings_file_path}\n")
114 elif not settings_file_path.exists():
115     sys.stdout.write("'config folder' found!\n")
116     sys.stdout.write(f"'config folder' location: {config_folder_path}\n")
117     with open(settings_file_path, "w", encoding="utf-8") as f:
118         json.dump(settings_data, f, ensure_ascii=False, indent=4)
119     sys.stdout.write("'settings.json' does not exist.\n")
120     sys.stdout.write(f"Creating 'settings.json' into: {settings_file_path}")
121 else:
122     sys.stdout.write("'config folder' found!\n")
123     sys.stdout.write(f"'config folder' location: {config_folder_path}\n")
124     sys.stdout.write("'settings.json' found!\n")
125     sys.stdout.write(f"'settings.json' location: {settings_file_path}\n")
126     sys.stdout.write("You can change settings in the 'settings.json'.\n")
127
128 # 'settings.json' 안에 설정된 값을 불러오기
129 with open(settings_file_path, "r", encoding="utf-8") as f:
130     settings = json.load(f)
131 graph = settings["graph"]
132 graph_visualization = settings["graph_visualization"]
133 trimmed_mean_percentage = settings["trimmed_mean_percentage"]
134 asynchronous = settings["matplotlib_pyplot_asynchronous"]
135 sys.stdout.write(f"Graph is set to '{graph}'.\n")
136 sys.stdout.write(f"Graph visualization is set to '{graph_visualization}'.\n")
137 sys.stdout.write(f"Trimmed mean value is set to '{trimmed_mean_percentage}%'.\n")
138 sys.stdout.write(f"Asynchronous is set to '{asynchronous}'.\n")
139
140 # 사용자로부터 받은 'graph' 값에 이상한 값인지 확인하고, 이상한 값이면 자동으로 기본값으로 설정
141 if not isinstance(graph, bool):
142     sys.stdout.write("Invalid graph value!\n")
143     graph = True
144     sys.stdout.write(f"Automatically set 'graph' value to '{graph}'!\n")
145     sys.stdout.write("Please open 'settings.json' and set the values as described in the comments.\n")
146
147 # 사용자로부터 받은 'graph_visualization' 값에 이상한 값인지 확인하고, 이상한 값이면 자동으로 기본값으로 설정
148 if graph_visualization not in ["plotext", "matplotlib.pyplot"]:
149     sys.stdout.write("Invalid graph_visualization value!\n")
150     graph_visualization = "matplotlib.pyplot"
151     sys.stdout.write(f"Automatically set 'graph_visualization' value to '{graph_visualization}'!\n")
152     sys.stdout.write("Please open 'settings.json' and set the values as described in the comments.\n")
153
154 # 사용자로부터 받은 'trimmed_mean_percentage' 값에 이상한 값인지 확인하고, 이상한 값이면 자동으로 기본값으로 설정

```

Launchpad Run Emulators Discord 99% Connected to Discord

www.utwente.nl

C:\Users\Unknown User\Desktop\Stock_AI> server.py

```

154 # 사용자로부터 받은 'trimmed_mean_percentage' 값이 이상한 값인지 확인하고, 이상한 값이면 자동으로 기본값으로 설정
155 if not isinstance(trimmed_mean_percentage, int) or trimmed_mean_percentage <= 0:
156     sys.stdout.write(f"Invalid trimmed_mean_percentage value!\n")
157     trimmed_mean_percentage = 1
158     sys.stdout.write(f"Automatically set 'trimmed_mean_percentage' value to '{trimmed_mean_percentage}'!\n")
159     sys.stdout.write("Please open 'settings.json' and set the values as described in the comments.\n")
160
161 # 사용자로부터 받은 'matplotlib_pyplot_asynchronous' 값이 이상한 값인지 확인하고, 이상한 값이면 자동으로 기본값으로 설정
162 if not isinstance(asynchronous, (int, bool)):
163     sys.stdout.write(f"Invalid matplotlib_pyplot_asynchronous value!\n")
164     asynchronous = False
165     sys.stdout.write(f"Automatically set 'matplotlib_pyplot_asynchronous' value to '{asynchronous}'!\n")
166     sys.stdout.write("Please open 'settings.json' and set the values as described in the comments.\n")
167 elif isinstance(asynchronous, int) and asynchronous <= 0:
168     sys.stdout.write(f"Invalid matplotlib_pyplot_asynchronous value!\n")
169     asynchronous = 5
170     sys.stdout.write(f"Automatically set 'matplotlib_pyplot_asynchronous' value to '{asynchronous}'!\n")
171     sys.stdout.write("Please open 'settings.json' and set the values as described in the comments.\n")
172
173 # 'csv_folder_parent_folder' 폴더 있는지 확인, 없으면 생성하기
174 if not csv_folder_parent_folder_path.exists():
175     sys.stdout.write(f"'csv folder's parent folder' does not exist.\n")
176     sys.stdout.write(f"Creating 'csv folder's parent folder' into: {csv_folder_parent_folder_path}\n")
177     csv_folder_parent_folder_path.mkdir(parents=True)
178 else:
179     sys.stdout.write(f"'csv folder's parent folder' found!\n")
180 sys.stdout.write(f"'csv folder's parent folder' location: {csv_folder_parent_folder_path}\n")
181
182 # 'csv_folder' 폴더들을 리스트화
183 csv_folder = [i for i in csv_folder_parent_folder_path.iterdir() if i.is_dir()]
184
185 if csv_folder:
186     # 'csv_folder_parent_folder' 안에 파일 있으면 사용가능한 파일 출력하기
187     sys.stdout.write(f"'csv folders' found!\n")
188     sys.stdout.write("\n")
189     sys.stdout.write(f"Available 'csv folders':\n")
190     count = 1
191     for idx in csv_folder:
192         sys.stdout.write(f"({count}): {idx.name}\n")
193         count += 1
194     sys.stdout.write("\n")
195     sys.stdout.write("Select a 'csv folder' u want to load: \n")
196
197 # 사용자로부터 사용할 'csv_folder' 입력받기
198 while True:
199     if len(csv_folder) == 1:
200         sys.stdout.write("Only one folder available, automatically selected.\n")
201         userinput = 1

```

Launchpad Run Testers Discount NPC Connected to Discount

```

197 #사용자로부터 사용할 'csv_folder' 입력받기
198 while True:
199     if len(csv_folder) == 1:
200         sys.stdout.write("Only one folder available, automatically selected.\n")
201         userInput = 1
202         break
203
204     userInput = sys.stdin.readline().strip()
205     if not userInput.isdigit():
206         sys.stdout.write("Please enter a number!\n")
207         continue
208
209     userInput = int(userInput)
210     if userInput not in range(1, count):
211         sys.stdout.write("Invalid number. Try again: \n")
212         continue
213
214     break
215 else:
216     sys.stdout.write("'csv folders' does not exist.\n")
217     sys.stdout.write("Please make at least one 'csv folder' to proceed!\n")
218     sys.stdout.write("The program will exit in 5 seconds...\n")
219     time.sleep(5)
220     sys.exit()
221
222 # 'csv_folder' 경로 설정
223 csv_folder_path = csv_folder[userInput - 1]
224
225 # 'csv_file' 전부 다 검색
226 csv_file = [i for i in csv_folder_path.iterdir() if i.is_file() and i.suffix == ".csv"]
227
228 # 'csv_folder' 안에 'csv_file'이 있는지 확인, 없으면 종료
229 if csv_file:
230     sys.stdout.write(f"Selected {userInput}th csv file!\n")
231     sys.stdout.write(f"'csv file' location: {csv_folder_path}\n")
232 else:
233     sys.stdout.write("No 'csv files' found in the selected folder!\n")
234     sys.stdout.write("The program will exit in 5 seconds...\n")
235     time.sleep(5)
236     sys.exit()
237
238 # 결사: 상위/하위 trimmed_mean_percentage% 제거
239 def trimmed_array(arr, percentage):
240     n = len(arr)
241     k = int(n * percentage / 100)
242     if k >= n // 2: # 데이터가 너무 적으면 그냥 원본 사용
243         return arr
244     arr_sorted = np.sort(arr)
245     lower = arr_sorted[k]
246     upper = arr_sorted[-k-1]
247     trimmed = np.clip(arr, lower, upper)
248     return trimmed
249
250 # csv에서 거저를 잘 감출
251 required_cols = ['종가', '시가', '고가', '저가', '거래량', '변동 X']
252 all_data = []
253
254 # 'csv_file'들을 모두 읽고 원하는 값 추출해내기
255 for f in csv_file:
256     df = None
257     #utf-8, cp949 코딩에 맞춰서 csv 전체읽히기
258     for enc in ['utf-8', 'cp949']:
259         try:
260             df = pd.read_csv(f, encoding=enc, thousands=',')
261             break
262         except:
263             continue
264     if df is None:
265         sys.stdout.write(f"[Skip] {f.name}: Cannot read\n")
266         continue
267
268     df.columns = df.columns.str.strip()
269     if not all(col in df.columns for col in required_cols):
270         sys.stdout.write(f"[Skip] {f.name}: Missing columns\n")
271         continue
272
273     # 전처리
274     df['거래량'] = pd.to_numeric(df['거래량'].astype(str).str.replace('M', ''), errors='coerce').fillna(0)*1e6
275     df['변동 X'] = pd.to_numeric(df['변동 X'].astype(str).str.replace('X', ''), errors='coerce').fillna(0)/100
276     for col in ['종가', '시가', '고가', '저가']:
277         df[col] = pd.to_numeric(df[col], errors='coerce').fillna(0)
278
279     all_data.append(df[required_cols].values[::-1])
280
281 if not all_data:
282     sys.stdout.write("No valid CSV data found!\n")
283     sys.stdout.write("The program will exit in 5 seconds...\n")
284     time.sleep(5)
285     sys.exit()

```

(F) 편집(F) 맥 영역(C) 보기(V) 이동(M) 실행(R) 터미널(T) 도움말(H) ← → ↺ 검색

server.py X

C:\Users\Unknown User\Desktop\Stock_AI> server.py ...

```

237
238 # 결사: 상위/하위 trimmed_mean_percentage% 제거
239 def trimmed_array(arr, percentage):
240     n = len(arr)
241     k = int(n * percentage / 100)
242     if k >= n // 2: # 데이터가 너무 적으면 그냥 원본 사용
243         return arr
244     arr_sorted = np.sort(arr)
245     lower = arr_sorted[k]
246     upper = arr_sorted[-k-1]
247     trimmed = np.clip(arr, lower, upper)
248     return trimmed
249
250 # csv에서 거저를 잘 감출
251 required_cols = ['종가', '시가', '고가', '저가', '거래량', '변동 X']
252 all_data = []
253
254 # 'csv_file'들을 모두 읽고 원하는 값 추출해내기
255 for f in csv_file:
256     df = None
257     #utf-8, cp949 코딩에 맞춰서 csv 전체읽히기
258     for enc in ['utf-8', 'cp949']:
259         try:
260             df = pd.read_csv(f, encoding=enc, thousands=',')
261             break
262         except:
263             continue
264     if df is None:
265         sys.stdout.write(f"[Skip] {f.name}: Cannot read\n")
266         continue
267
268     df.columns = df.columns.str.strip()
269     if not all(col in df.columns for col in required_cols):
270         sys.stdout.write(f"[Skip] {f.name}: Missing columns\n")
271         continue
272
273     # 전처리
274     df['거래량'] = pd.to_numeric(df['거래량'].astype(str).str.replace('M', ''), errors='coerce').fillna(0)*1e6
275     df['변동 X'] = pd.to_numeric(df['변동 X'].astype(str).str.replace('X', ''), errors='coerce').fillna(0)/100
276     for col in ['종가', '시가', '고가', '저가']:
277         df[col] = pd.to_numeric(df[col], errors='coerce').fillna(0)
278
279     all_data.append(df[required_cols].values[::-1])
280
281 if not all_data:
282     sys.stdout.write("No valid CSV data found!\n")
283     sys.stdout.write("The program will exit in 5 seconds...\n")
284     time.sleep(5)
285     sys.exit()

```

Launched Run Testcases 0.0.0 Discord RPC Connected to Discord

```

381 if not all_data:
382     sys.stdout.write("No valid CSV data found!\n")
383     sys.stdout.write("the program will exit in 5 seconds...\n")
384     time.sleep(5)
385     sys.exit()
386
387 all_data = np.concatenate(all_data, axis=0)
388 idx = all_data.shape[0]
389 sys.stdout.write(f"loaded {idx} rows from CSV files.\n")
390 idx -= idx % 10
391 if 50 > idx // 10:
392     EPOCHS = 50
393 else:
394     EPOCHS = idx // 10
395 sys.stdout.write(f"train the AI {EPOCHS} times.\n")
396
397 # 데이터 준비
398 SEQ_LENGTH = 10
399 scaler = MinMaxScaler()
400 data_scaled = scaler.fit_transform(all_data)
401
402 x_list, y_list = [], []
403 for i in range(len(data_scaled)-SEQ_LENGTH):
404     x_list.append(data_scaled[i:i+SEQ_LENGTH])
405     y_list.append(data_scaled[i+SEQ_LENGTH][0]) # 물가 예측
406
407 x = torch.tensor(np.array(x_list), dtype=torch.float32)
408 y = torch.tensor(np.array(y_list), dtype=torch.float32).unsqueeze(1)
409
410 # GRU 모델 정의
411 class GRUStock(nn.Module):
412     def __init__(self, input_size=6, hidden_size=12, num_layers=1):
413         super().__init__()
414         self.gru = nn.GRU(input_size, hidden_size, num_layers, batch_first=True)
415         self.fc = nn.Linear(hidden_size, 1)
416     def forward(self, x):
417         out, _ = self.gru(x)
418         return self.fc(out[:, -1, :])
419
420 model = GRUStock()
421
422 # 학습하기
423 criterion = nn.MSELoss()
424 optimizer = torch.optim.Adam(model.parameters(), lr=0.001)
425 sys.stdout.write("AI training started...\n")
426 for epoch in range(EPOCHS):
427     model.train()
428     optimizer.zero_grad()

```

```

322 # 학습하기
323 criterion = nn.MSELoss()
324 optimizer = torch.optim.Adam(model.parameters(), lr=0.001)
325 sys.stdout.write("AI training started...\n")
326 for epoch in range(EPOCHS):
327     model.train()
328     optimizer.zero_grad()
329     output = model(x)
330     loss = criterion(output, y)
331     loss.backward()
332     optimizer.step()
333     if (epoch+1)%10==0:
334         sys.stdout.write(f"Success {epoch+1}/{EPOCHS}, Loss: {loss.item():.6f}\n")
335
336 # 예측 및 시각화
337 model.eval()
338 with torch.no_grad():
339     pred = model(x).numpy()
340     pred_prices = scaler.inverse_transform(
341         np.concatenate([pred, np.zeros((pred.shape[0], all_data.shape[1]-1))], axis=1)
342     )[:,0]
343     actual_prices = all_data[SEQ_LENGTH:,0]
344
345 # 예측값에 결사평균 적용
346 pred_prices_trimmed = trimmed_array(pred_prices, trimmed_mean_percentage)
347
348 # 미래 예측값 설정
349 future_days = 20
350 last_known_idx = len(actual_prices) - 1
351 future_pred = pred_prices_trimmed[-future_days:]
352 combined_future_pred = np.concatenate([pred_prices_trimmed[-1]], future_pred))
353 x_future = range(last_known_idx, last_known_idx + len(combined_future_pred))
354
355 # 마지막 실제 값과 미래 예측값 저장하기
356 last_actual_price = actual_prices[-1]
357 last_future_price = combined_future_pred[-1]
358
359 # 화면에 그래프 띄우기
360 if graph and graph_visualization == "matplotlib.pyplot":
361     sys.stdout.write(f"Draw a graph using {graph_visualization}.\n")
362     plt.figure(figsize=(12, 6))
363     plt.plot(actual_prices, label='Actual Price', color='black')
364     plt.plot(pred_prices_trimmed, color='green', label=f'Predicted Price (Trimmed Mean)')
365     plt.axvline(x=last_known_idx, color='red', linestyle='--', label='Future Prediction Start')
366     plt.plot(x_future, combined_future_pred,
367             '--', color='blue', label='Future Prediction')
368     plt.title("Stock Price Prediction with Future Forecast")
369     plt.xlabel("Days")

```

```

359 # 화면에 그래프 띄우기
360 if graph and graph_visualization == "matplotlib.pyplot":
361     sys.stdout.write(f"Draw a graph using {graph_visualization}.\n")
362     plt.figure(figsize=(12, 6))
363     plt.plot(actual_prices, label='Actual Price', color='black')
364     plt.plot(pred_prices_trimmed, color='green', label=f'Predicted Price (Trimmed Mean)')
365     plt.axvline(x=last_known_idx, color='red', linestyle='--', label='Future Prediction Start')
366     plt.plot(x_future, combined_future_pred,
367              '--', color='blue', label='Future Prediction')
368     plt.title("Stock Price Prediction with Future Forecast")
369     plt.xlabel("Days")
370     plt.ylabel("Price")
371     plt.legend()
372     plt.grid(True)
373     if asynchronous == False:
374         plt.show()
375     elif isinstance(asynchronous, int) and asynchronous > 0:
376         plt.show(block=False)
377         plt.pause(asynchronous)
378     else:
379         sys.stdout.write("Invalid asynchronous value!\n")
380         sys.stdout.write("Please open 'settings.json' and set the values as described in the comments.\n")
381         sys.stdout.write("The program will exit in 5 seconds...\n")
382         time.sleep(5)
383         sys.exit()
384
385 elif graph and graph_visualization == "plotext":
386     sys.stdout.write(f"Draw a graph using {graph_visualization}.\n")
387     plot.plot(actual_prices, color="white", label="Actual Price")
388     plot.plot(pred_prices_trimmed, color="green", label="Predicted Price (Trimmed Mean)")
389     plot.vline(last_known_idx, color="red")
390     plot.plot(x_future, combined_future_pred, color="blue", marker="--", label="Future Prediction")
391     plot.title("Stock Price Prediction with Future Forecast")
392     plot.xlabel("Days")
393     plot.ylabel("Price")
394     plot.grid(True)
395     plot.title('Legend:')
396     plot.title('Actual Price: black')
397     plot.title('Predicted Price: green')
398     plot.title('Future Prediction: blue')
399     plot.title('Future Start: red')
400     plot.show()
401
402 elif not graph:
403     sys.stdout.write("Graphs are not drawn according to the settings.\n")
404
405 else:
406     sys.stdout.write("Invailed settings values!\n")

```

```

"graph": true,
"graph_visualization": "matplotlib.pyplot",
"trimmed_mean_percentage": 1,
"matplotlib_pyplot_asynchronous": false,
"br": "br",
"kor_comment_graph": "매일 그래프를 보여줄지",
"kor_comment_graph_visualization": "plotext: cmd 안에서 켜거나 선으로, matplotlib.pyplot: 외부 창에서",
"kor_comment_trimmed_mean_percentage": "절사평균값 (10이면 상위/하위 10% 제거)",
"kor_comment_matplotlib_pyplot_asynchronous": "비동기 사용여부(사용안함: False/ 사용함: (몇 분 후에 값이 재검지))",
"bl": "bl",
"eng_comment_graph": "Whether to show the predicted graph",
"eng_comment_graph_visualization": "plotext: dots and lines in the CMD, matplotlib.pyplot: in an another external window",
"eng_comment_trimmed_mean_percentage": "Trimmed mean value (10 means removing the top and bottom 10%)",
"eng_comment_matplotlib_pyplot_asynchronous": "Use asynchronous mode (False: disabled / [minutes]: enabled, window will clo

```

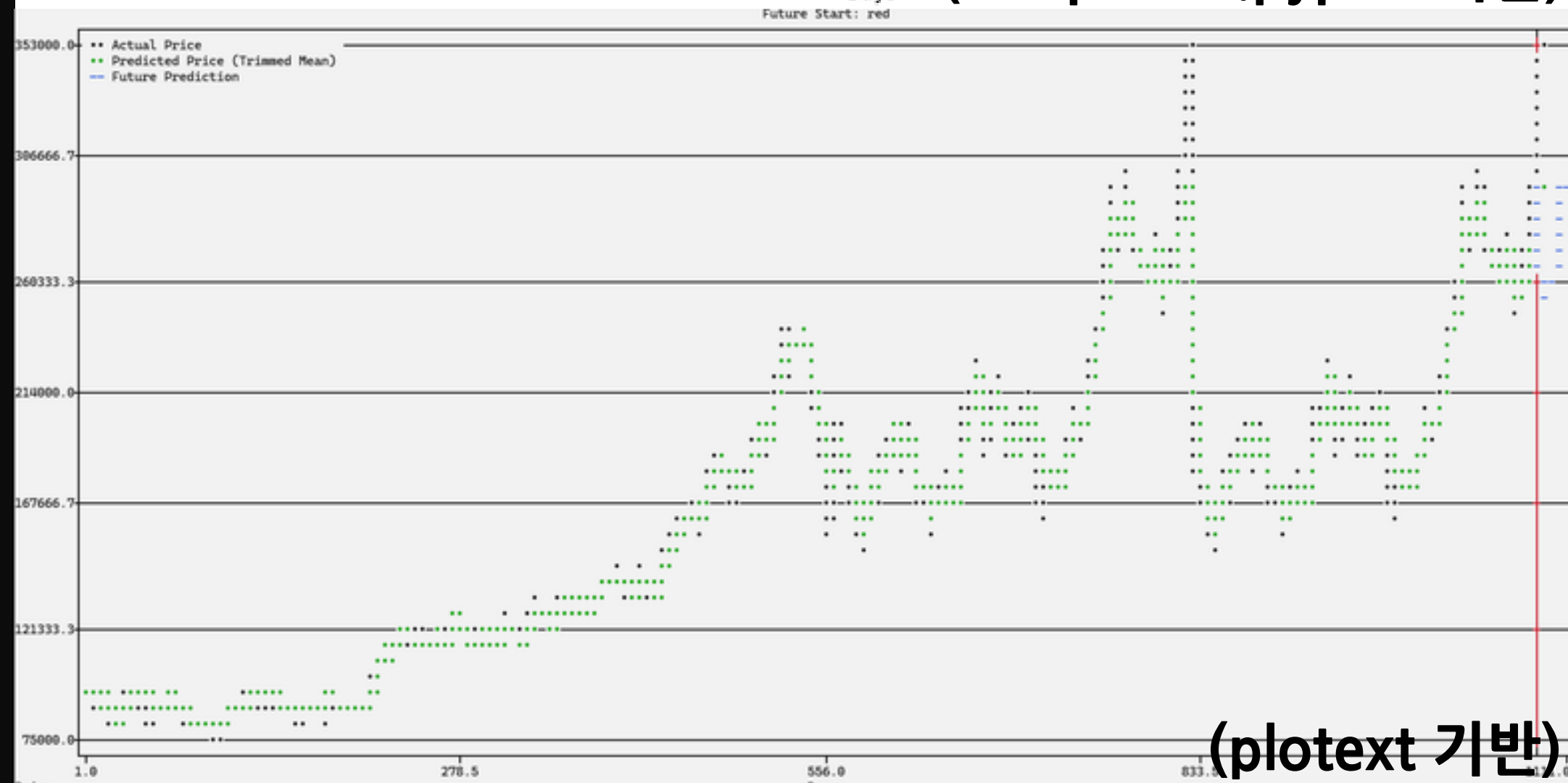
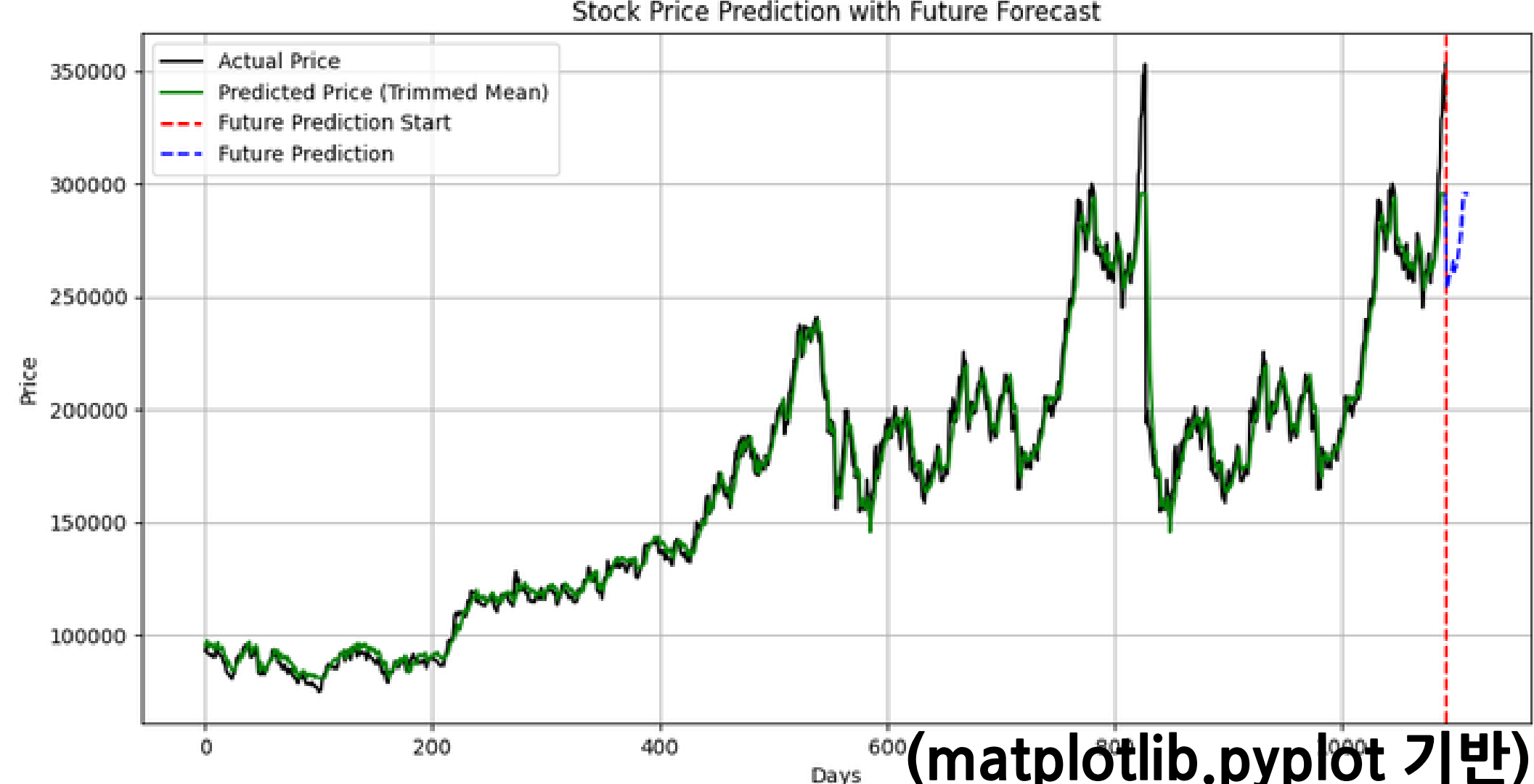
```

394     plot.grid(True)
395     plot.title('Legend:')
396     plot.title('Actual Price: black')
397     plot.title('Predicted Price: green')
398     plot.title('Future Prediction: blue')
399     plot.title('Future Start: red')
400     plot.show()
401
402 elif not graph:
403     sys.stdout.write("Graphs are not drawn according to the settings.\n")
404
405 else:
406     sys.stdout.write("Invailed settings values!\n")
407     sys.stdout.write("Please open 'settings.json' and set the values as described in the comments.\n")
408     sys.stdout.write("The program will exit in 5 seconds...")
409     time.sleep(5)
410     sys.exit()
411
412 # 최종적으로 AI 예측 결과 출력하기
413 sys.stdout.write("\n")
414 if last_actual_price > last_future_price:
415     sys.stdout.write(f"Current value is higher than the value {future_days} days later!\n")
416     sys.stdout.write("AI prediction result: You should NOT buy this stock!\n")
417     sys.stdout.write(o_ascii_art + "\n")
418
419 elif last_actual_price < last_future_price:
420     sys.stdout.write(f"The value {future_days} days later is higher than the current value!\n")
421     sys.stdout.write("AI prediction result: You should buy this stock!\n")
422     sys.stdout.write(o_ascii_art + "\n")
423
424 else:
425     sys.stdout.write(f"The current value and {future_days} days later value are the same!\n")
426     sys.stdout.write("AI prediction result: You can buy this stock if you want.\n")
427     sys.stdout.write(same_ascii_art + "\n")
428
429 # 아무 키 입력하면 프로그램 종료하기
430 sys.stdout.write("Press any key to exit...\n")
431 msvcrt.getch() # 아무 키 입력 대기
432 sys.exit()

```

```
Select a 'csv folder' u want to load:
1
Selected 1th csv file!
'csv file' location: C:\Users\Unknown User\Desktop\Stock_AI\csv\samsung-electroni
Loaded 1102 rows from CSV files.
Train the AI 110 times.
AI training started...
Success 10/110, Loss: 0.015400
Success 20/110, Loss: 0.018306
Success 30/110, Loss: 0.012305
Success 40/110, Loss: 0.005810
Success 50/110, Loss: 0.002232
Success 60/110, Loss: 0.001203
Success 70/110, Loss: 0.001170
Success 80/110, Loss: 0.001071
Success 90/110, Loss: 0.000981
Success 100/110, Loss: 0.000958
Success 110/110, Loss: 0.000934
Draw a graph using matplotlib.pyplot.

Current value is higher than the value 20 days later!
AI prediction result: You should NOT buy this stock!
```



차별점

사용자 정의 설정 가능

- 여러가지 방식으로 그래프를 프린트 할 수 있음 (matplotlib.pyplot, plotext)
- 사용자가 입력한 값대로 절사평균값 처리
- 폴더 내에서 원하는 파일 선택 가능, 폴더가 없으면 자동으로 생성
- 'settings.json' 파일을 생성해 설정값 저장

절사평균 사용

- 주식 데이터에서 갑작스런 급등, 급락 같은 극단값이 전체평균을 왜곡하는 것을 방지

감사합니다.