경기도 부동산 가격 예측 GO or NOT

5조 link to korea -링코



TEAM MEMBER



김민재

프로젝트 기획 데이터수집 Clustering Q-GIS Figma

ENFJ



안다미로

프로젝트 기획 데이터수집 데이터 분석 ARIMA 발표

ESTJ



^{조장} 정은원

프로젝트 기획 데이터수집 LSTM PPT

ESTJ



정현태

프로젝트 기획 데이터수집 FB Prophet 평가지표 분석

INTJ





- 1. 기획의도
- 2. 모델 선정
- 3. LSTM 변수 선정
- 4. 다변량 LSTM
- 5. 서비스 방향



거주 VS 투자?

1.기획의도



<아파트사이클연구소 이현철 소장>

1.기획의도

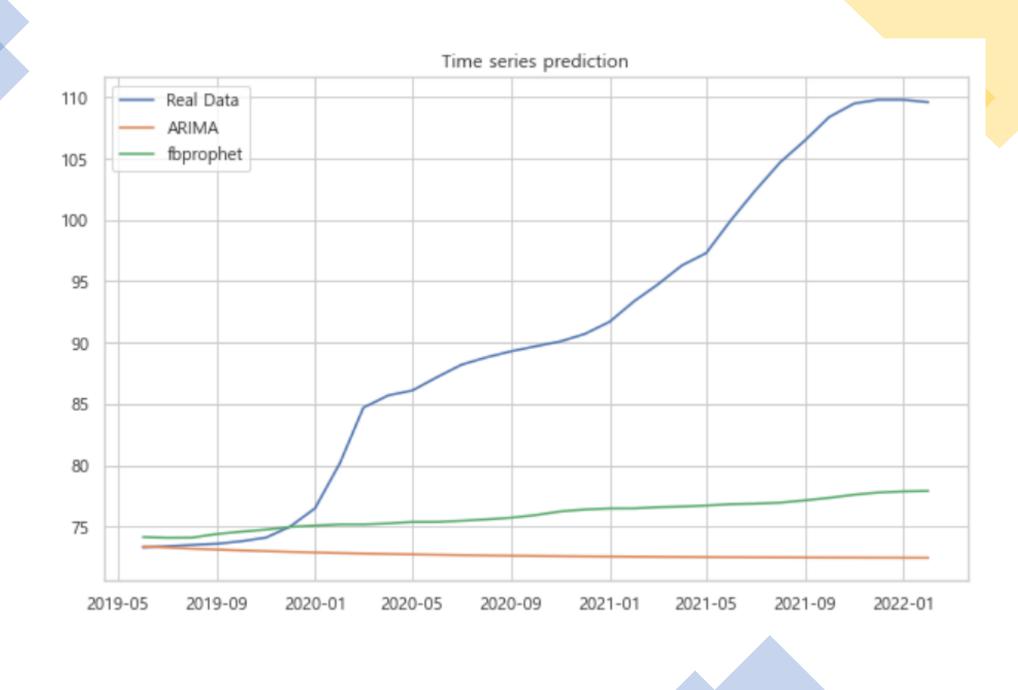


실거주자를 위한 정보제공 서비스 구현



ARIMA VS FBprophet ?





3.LSTM



https://facebook.github.io/prophet/docs/quick_start.html#python-api

3.LSTM-변수



거시경제 변수

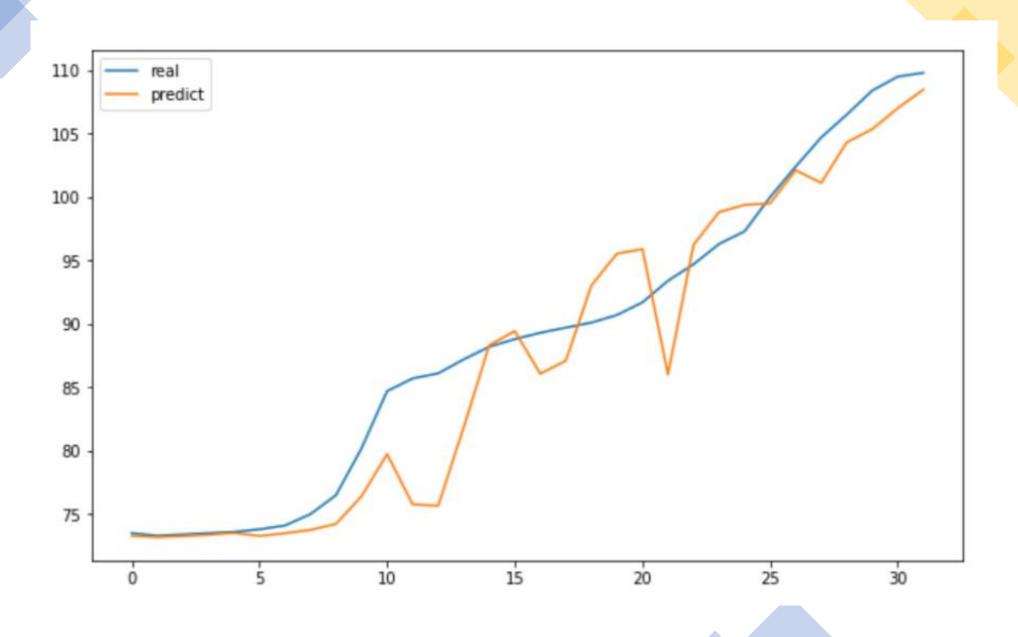
盟 3.LSTM-거시변수



MMI(광공업지수) CPI(소비자물가지수) 대출금리 환율 M2(통화량) 주가

전세가격지수 건축허용면적 거래량

	ммі	CPI	대출금 리	전세가격지수 전국	전세가격지수 경기도	환율	M2	코스피	코스닥	건축허용면적	거래량 전 국	거래량 경기
date												
2013- 01-01	102.86800	92.72800	5.78000	82.68933	76.00764	1089.00000	1841.10000	1961.90000	503.67000	8413721.00000	54632	13631
2013- 02-01	90.00900	93.03800	5.66000	82.87118	76.10804	1083.00000	1857.10000	2026.50000	535.87000	7268443.00000	65640	14925
2013- 03-01	101.38800	92.95200	5.54000	83.16272	76.37669	1111.00000	1862.40000	2004.90000	555.02000	10160177.00000	87563	17718
2013- 04-01	101.29500	92.82300	5.42000	83.55404	76.85819	1101.00000	1867.70000	1964.00000	563.87000	9757932.00000	106498	22790
2013- 05-01	102.31300	92.82300	5.24000	83.79708	77.09053	1129.70000	1870.30000	2001.05000	577.87000	10837979.00000	120393	25910
				***	***	***	***	***	***	***	***	***
2021- 10-01	114.80000	103.35000	3.47000	102.45093	103.90395	1168.60000	3543.40000	2970.68000	992.33000	14689855.00000	119686	29187



盟 3.LSTM-변수



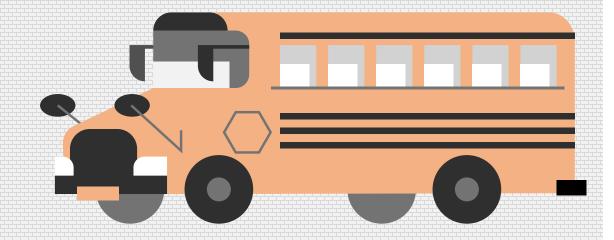
부동산 지역성



3. LSTM-클러스터링



경기도 : 서울의 영향이 큼



3. LSTM-클러스터링

시계열데이터 군집화 방법



DTW(Dynamic Time Warping)

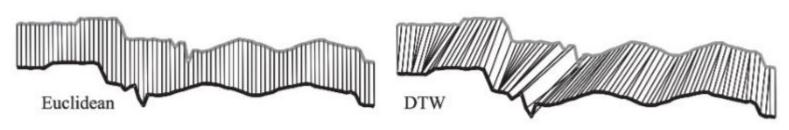


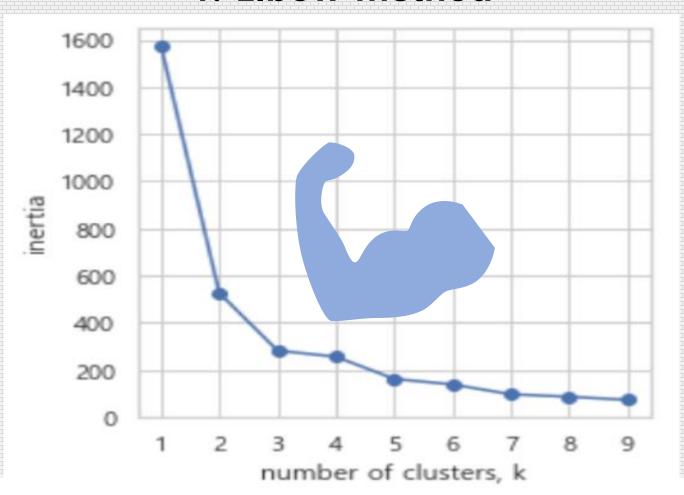
Fig. 1. Note that, while the two time series have an overall similar shape, they are not aligned in the time axis. Euclidean distance, which assumes the i^{th} point in one sequence is aligned with the i^{th} point in the other, will produce a pessimistic dissimilarity measure. The nonlinear dynamic time warped alignment allows a more intuitive distance measure to be calculated



3.LSTM-클러스터링

군집 개수

1. Elbow method

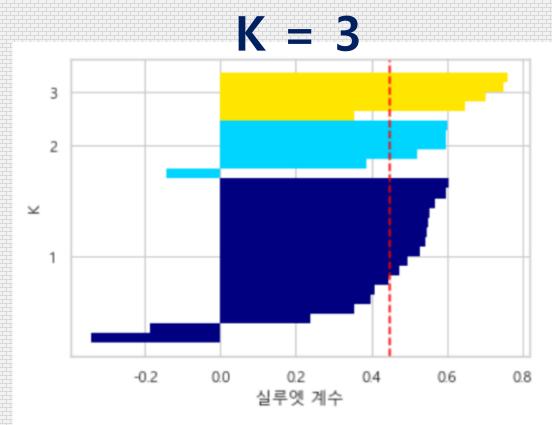


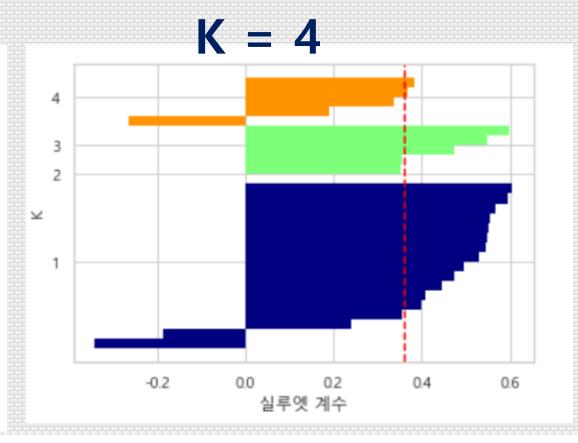


3. LSTM-클러스터링



2. 실루엣 계수







囲 3. LSTM-클러스터링

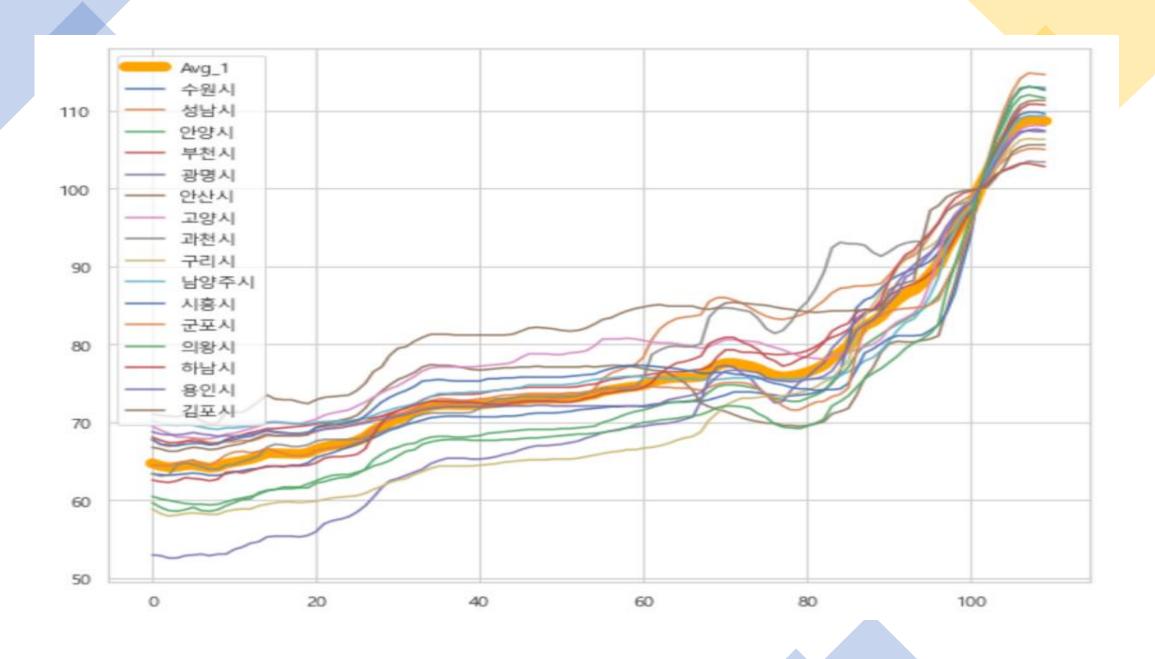
from tslearn.clustering import TimeSeriesKMeans,

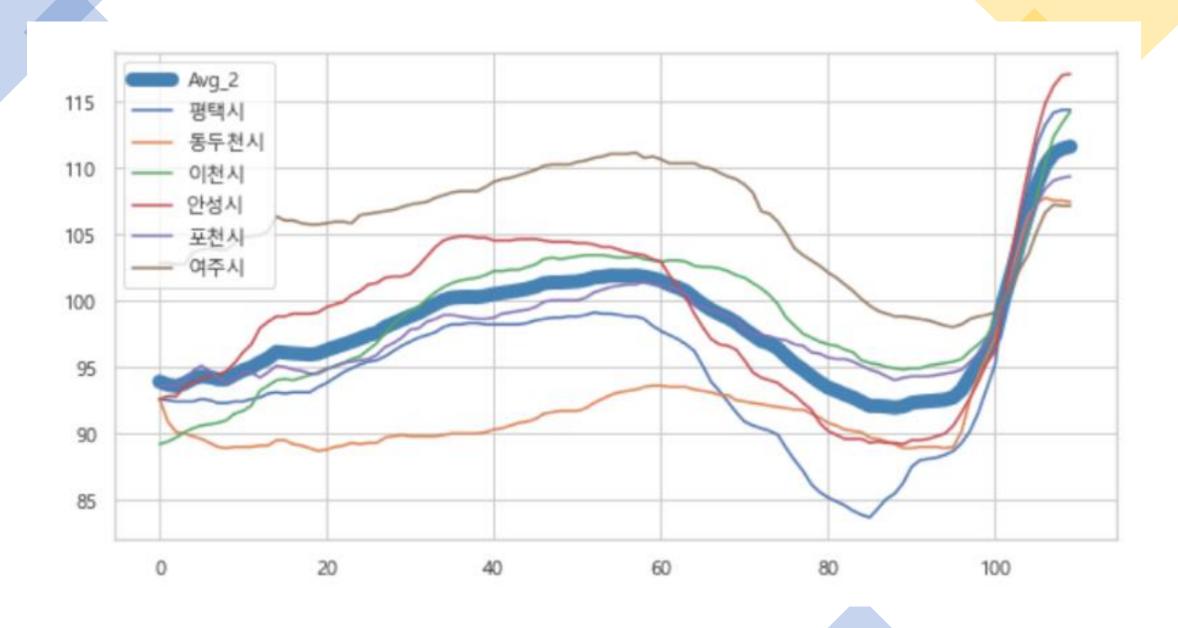


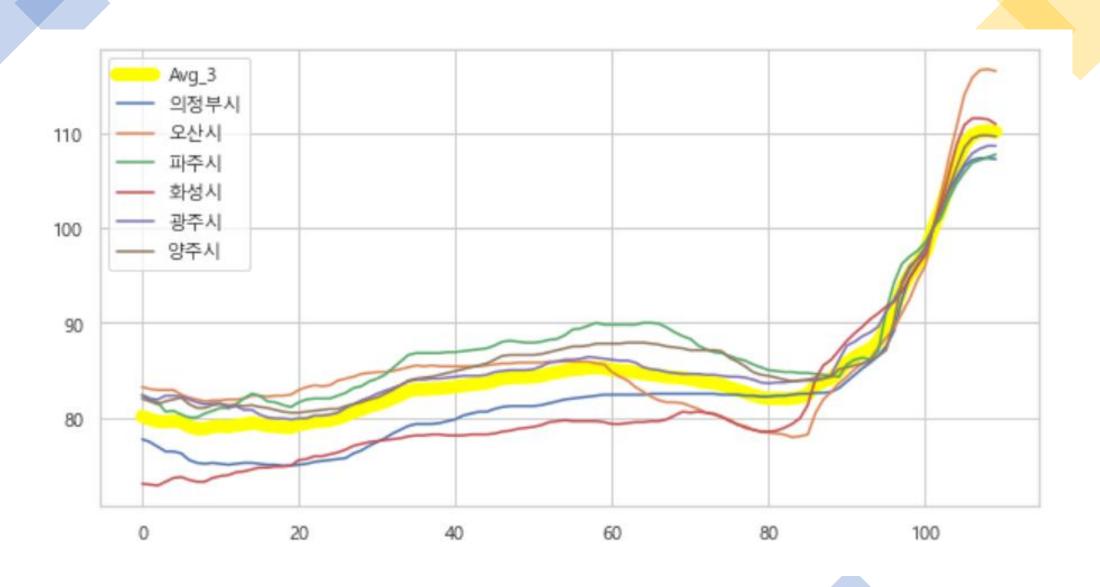
$$K = 3$$

클러스터링 실행











3.LSTM-클러스터링





서울과의 거리에 따라 가격지수의 변화가 비슷한 경향

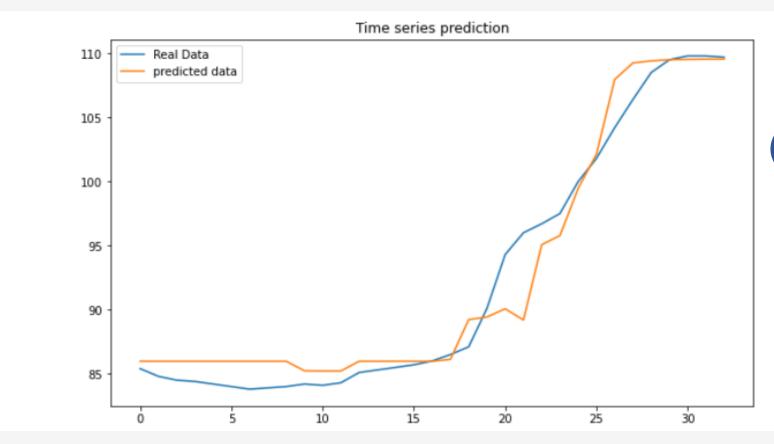
3. LSTM-클러스터링

data_S.corr()

	seoul	Avg_1	Avg_2	Avg_3
seoul	1.000000	0.909367	0.227910	0.773868
Avg_1	0.909367	1.000000	0.430139	0.952320
Avg_2	0.227910	0.430139	1.000000	0.636631
Avg_3	0.773868	0.952320	0.636631	1.000000



3. LSTM-클러스터링





가설 검증

서울시 가격지수, 군집 평균 데이터를 변수로 Lstm 실행

3. LSTM-클러스터링

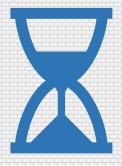
상관성이 있다



변수 추가

3.LSTM-변수 추출

왜?



시계열 데이터 자체의 한계

盟 3.LSTM-변수 추출

변수와 군집별 상관관계>0.7 추출

train.corr()

Pearson 상관계수

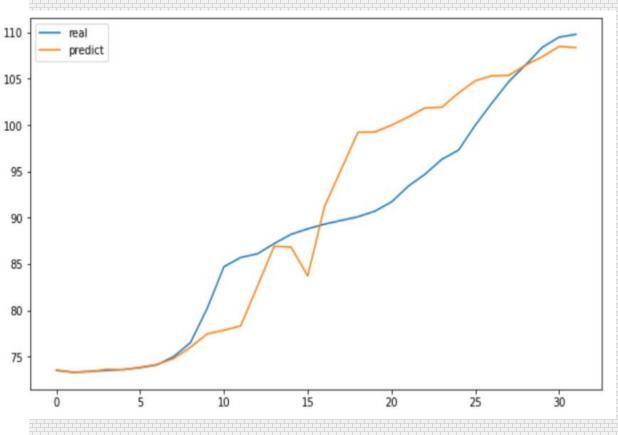
	seoul	Avg_1	Avg_2	Avg_3	ммі	СРІ	대출금리	전세가격 지수 전국	전세가격 지수 경기 도	환율	M2	코스피	코스닥	건축허용 면적	:
seoul	1.000000	0.909367	0.227910	0.773868	0.655650	0.985362	-0.788762	0.821937	0.833499	0.511194	0.983945	0.713424	0.804257	0.120689	
Avg_1	0.909367	1.000000	0.430139	0.952320	0.659035	0.914310	-0.738248	0.859271	0.864163	0.454030	0.950728	0.825427	0.851910	0.213492	
Avg_2	0.227910	0.430139	1.000000	0.636631	0.263814	0.316192	-0.263846	0.629858	0.605018	0.122268	0.258434	0.404361	0.410531	0.467495	-
Avg_3	0.773868	0.952320	0.636631	1.000000	0.617700	0.802606	-0.595459	0.829918	0.825010	0.352483	0.825879	0.833771	0.805576	0.255809	-
MMI	0.655650	0.659035	0.263814	0.617700	1.000000	0.643640	-0.508876	0.590326	0.592166	0.225834	0.660918	0.590472	0.575644	0.370323	
CPI	0.985362	0.914310	0.316192	0.802606	0.643640	1.000000	-0.788664	0.860513	0.867821	0.477902	0.977648	0.736793	0.800200	0.135486	
대출 금리	-0.788762	-0.738248	-0.263846	-0.595459	-0.508876	-0.788664	1.000000	-0.883444	-0.897772	-0.557470	-0.827397	-0.575287	-0.730375	-0.416933	-
전세 가격 지수 전국	0.821937	0.859271	0.629858	0.829918	0.590326	0.860513	-0.883444	1.000000	0.998765	0.444093	0.854004	0.738510	0.832382	0.455081	

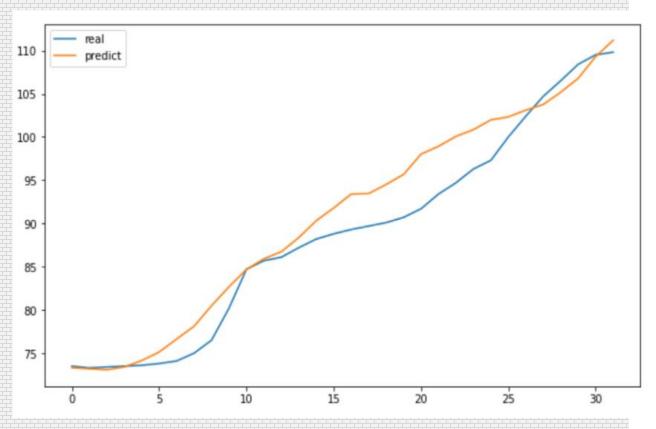


3. LSTM-변수 추출

추출 전









양방향 LSTM



4. 양방향 LSTM

군집2의 평가지수 상대적 Low

data_S.corr()

	seoul	Avg_1	Avg_2	Avg_3
seoul	1.000000	0.909367	0.227910	0.773868
Avg_1	0.909367	1.000000	0.430139	0.952320
Avg_2	0.227910	0.430139	1.000000	0.636631
Avg_3	0.773868	0.952320	0.636631	1.000000

data_S[-48:].corr()

4년전

data_S[-12:].corr()

1년전

	seoul	Avg_1	Avg_2	Avg_3
seoul	1.000000	0.981992	0.600982	0.933081
Avg_1	0.981992	1.000000	0.671144	0.973664
Avg_2	0.600982	0.671144	1.000000	0.816429
Avg_3	0.933081	0.973664	0.816429	1.000000

	seoul	Avg_1	Avg_2	Avg_3
seoul	1.000000	0.994531	0.999155	0.998441
Avg_1	0.994531	1.000000	0.993325	0.997756
Avg_2	0.999155	0.993325	1.000000	0.997365
Avg_3	0.998441	0.997756	0.997365	1.000000

4. 양방향 LSTM

평가지표





	평택시	동두천시	이천시	안성시	포천시
RMSE	0.381192	8.902235	2.272186	11.319	3.646946
MAPE	0.020086	0.092631	0.067622	0.032118	0.02936



	평택시	동두천시	이천시	안성시	포천시
RMSE	2.021074	8.795798	2.212662	3.366094	1.936525
MAPE	0.017677	0.081885	0.019299	0.028728	0.017723



4. 양방향 LSTM - forecasting

6개월 후 미래 가격지수를 예측



검증: 실제 가격지수 있는 22년 1월~3월 비교

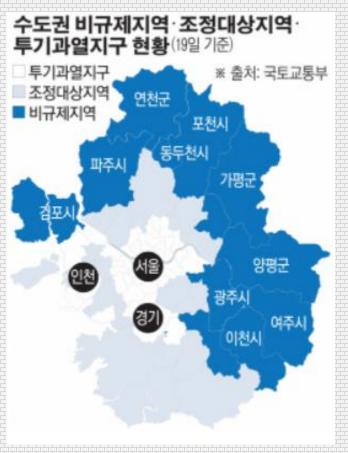
	수원시	성남시	의정부시	만양시	부전시	광명시	평택시	동두전시	안산시	고양시	과전시	구리시	남양주시	시흥시	군포시
2022-01-01 0:00	108.041	104.8988	107.1614	111.0145	110.6732	104.0236	109.0838	107.4137	97.02008	107.0951	99.04395	94.57448	106.7932	109.9355	112.9525
2022-02-01 0:00	100.9836	102.2846	107.207	120.2015	113.1941	99.66396	109.0838	103.4663	80.41757	107.2925	103.2973	95.80846	107.0065	113.0613	116.7516
2022-03-01 0:00	85.15893	106.0991	107.1763	109.8201	109.7421	102.3151	109.0838	94.52822	85.28823	108.2624	101.4519	99.68064	105.3706	116.8944	88.90529
2022-04-01 0:00	80.20289	103.5171	107.182	97.88582	97.5012	82.71621	109.0838	91.0604	83.82309	107.2969	95.89838	94.34965	107.488	115.3204	85.71557
2022-05-01 0:00	78.52642	103.9956	107.2016	90.62004	106.8767	88.90528	109.0838	89.94101	94.11882	101.2438	94.35016	93.67408	83.65395	113.1158	106.7776
2022-06-01 0:00	77.81567	105.0367	107.1695	110.4399	110.552	106.8659	109.0838	91.62721	82.66403	107.3484	103.1682	95.00045	113.6507	115.5253	99.18234



4. 양방향 LSTM - forecasting

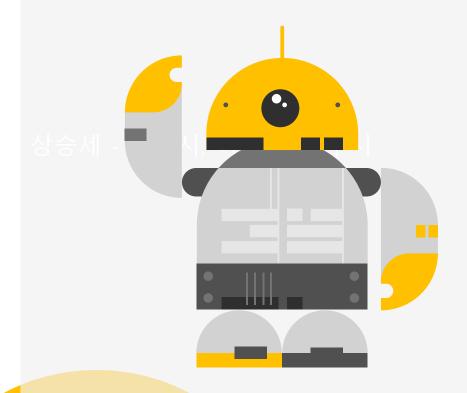


예측	실제
111.9639	113.3
112.1753	114.2
112.266	115.5



비규제 지역선정 ARREST ARE

■ 5.서비스 구현



실수요자를 위한 정보 제공

예측추세 정보 제공

▦ 5.서비스 구현







평택시, 이천시, 파주시

구리시

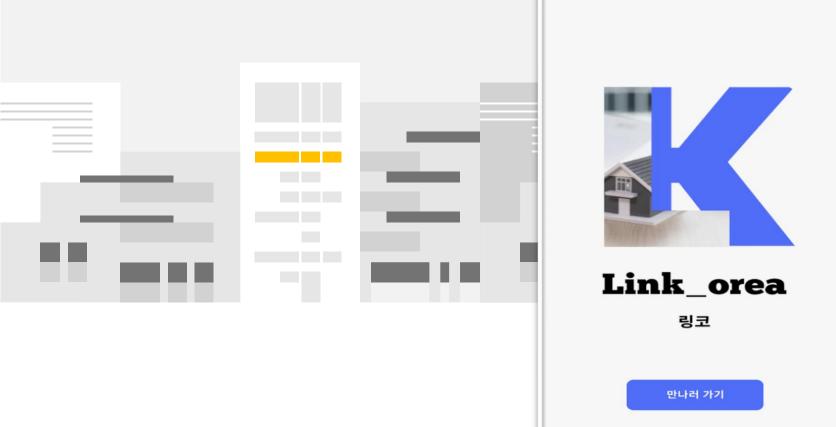
수원시, 성남시, 안양시, 부천시, 광명시, 안산시, 고양시, 과천시, 시흥시, 군포시, 하남시, 동두천시, 오산시, 광주시

상승세

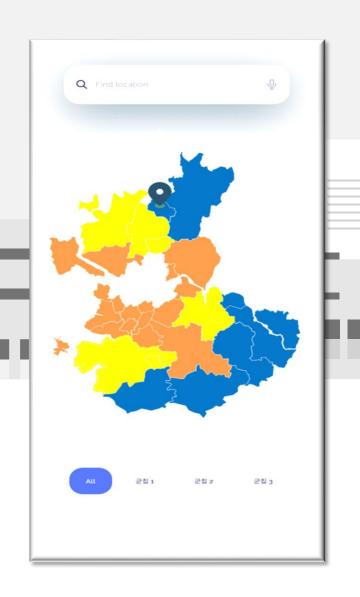
유지세

하락세

■ 5.서비스 구현



Made with team 5



Reference

- 신은경,김은미,홍태호, "SOM과 LSTM을 활용한 지역기반의 부동산 가격 예측"
- 김경외,정성도, "군집 분석을 활용한 아파트 실거래가격지수 연구- 서울시 사례를 중심으로 -"
- 배성완,유정석, "머신 러닝 방법과 시계열 분석 모형을 이용한 부동산 가격지수 예측"
- KOSIS 국가통계포털
- e-나라지표
- 한국부동산원
- https://facebook.github.io/prophet/



감사합니다



