

(세종시) 주택 시장 특성분석

최종 분석 보고서

주제 : 세종시 부동산 가격 변화에 미치는 변수 분석

팀 : 전병팀_전병용

1. 문제 정의 및 주제 선정의 이유

• 문제 정의 및 주제 선정 이유

- 세종시는 2005년 행정중심복합도시 건설을 위한 특별법 제정
- 이후 2012년 특별자치시로 출범, 행정수도로써 급격한 도시화가 진행
- 지속적인 인구 증가와 함께 부동산 가격이 상승
- 이에 제공해준 데이터를 통해 지속적인 부동산 가격 상승 원인 파악

목표 : 변수가 부동산 거래가격변화에 미치는 영향 분석

가설 : 상권(특히 Q(음식점)) 밀집도가 부동산 거래가격변화에 큰 영향을 미친다.

2. 분석 방법론 – 데이터 수집 및 추출 방안

세종시 상권정보

	D	F	L	N	O	P	Q	R
행정동명								
고운동	61.0	74.0	25.0	7.0	NaN	3.0	136.0	47.0
금남면	184.0	102.0	98.0	13.0	2.0	1.0	284.0	15.0
대평동	25.0	32.0	2.0	2.0	NaN	1.0	38.0	11.0
도담동	217.0	83.0	51.0	13.0	NaN	3.0	388.0	48.0
보람동	61.0	69.0	19.0	5.0	NaN	5.0	230.0	31.0
부강면	135.0	69.0	14.0	11.0	8.0	NaN	189.0	10.0
새롬동	313.0	159.0	157.0	21.0	NaN	4.0	349.0	115.0
소담동	38.0	44.0	11.0	NaN	NaN	NaN	81.0	38.0
소정면	43.0	22.0	4.0	1.0	1.0	NaN	87.0	7.0
아름동	45.0	38.0	13.0	3.0	NaN	6.0	115.0	57.0
연기면	56.0	37.0	28.0	1.0	3.0	NaN	67.0	2.0
연동면	46.0	14.0	15.0	NaN	3.0	NaN	71.0	7.0
연서면	162.0	50.0	40.0	5.0	7.0	NaN	164.0	16.0
장군면	128.0	40.0	78.0	10.0	3.0	1.0	203.0	2.0
전동면	53.0	22.0	5.0	2.0	6.0	NaN	51.0	4.0
전의면	139.0	61.0	9.0	7.0	5.0	NaN	182.0	14.0
조치원읍	1133.0	574.0	105.0	97.0	36.0	3.0	1402.0	203.0
종촌동	87.0	48.0	20.0	1.0	NaN	3.0	131.0	56.0
한솔동	77.0	49.0	89.0	1.0	NaN	NaN	174.0	71.0

세종시 세대수

	읍면동	1인	2인	3인	4인	5인	6인	7인	8인	9인	10인이상
0	조치원읍	8295	4165	3042	2729	944	220	63	9	4	1
1	연기면	855	299	133	101	36	11	7	1	2	0
2	연동면	834	461	206	114	44	14	5	1	0	1
3	부강면	1603	821	420	273	93	34	3	0	0	2
4	금남면	2128	1283	589	328	122	41	12	6	0	0
5	장군면	1956	806	332	200	106	35	11	6	1	0
6	연서면	1863	1022	484	276	129	38	16	3	2	1
7	전의면	1551	753	352	244	107	29	8	0	1	0
8	전동면	1224	491	213	124	57	15	6	0	0	0
9	소정면	565	308	160	127	58	9	1	2	1	0
10	한솔동	1069	1134	1468	1944	584	101	25	5	1	2
11	새롬동	6019	3980	4283	4754	1201	201	39	12	4	2
12	도담동	5002	2117	2383	2825	667	106	29	3	4	0
13	아름동	1058	1262	1858	2656	597	98	20	7	1	3
14	종촌동	2525	2049	2420	2847	675	120	34	17	2	2
15	고운동	2025	2341	2839	3240	814	145	19	7	4	2
16	소담동	2754	2268	2597	2545	671	106	27	1	1	3
17	보람동	1536	1363	1544	1802	471	69	20	2	2	1
18	대평동	857	977	932	886	228	33	9	4	0	0

2. 분석 방법론 – 데이터 수집 및 추출 방안

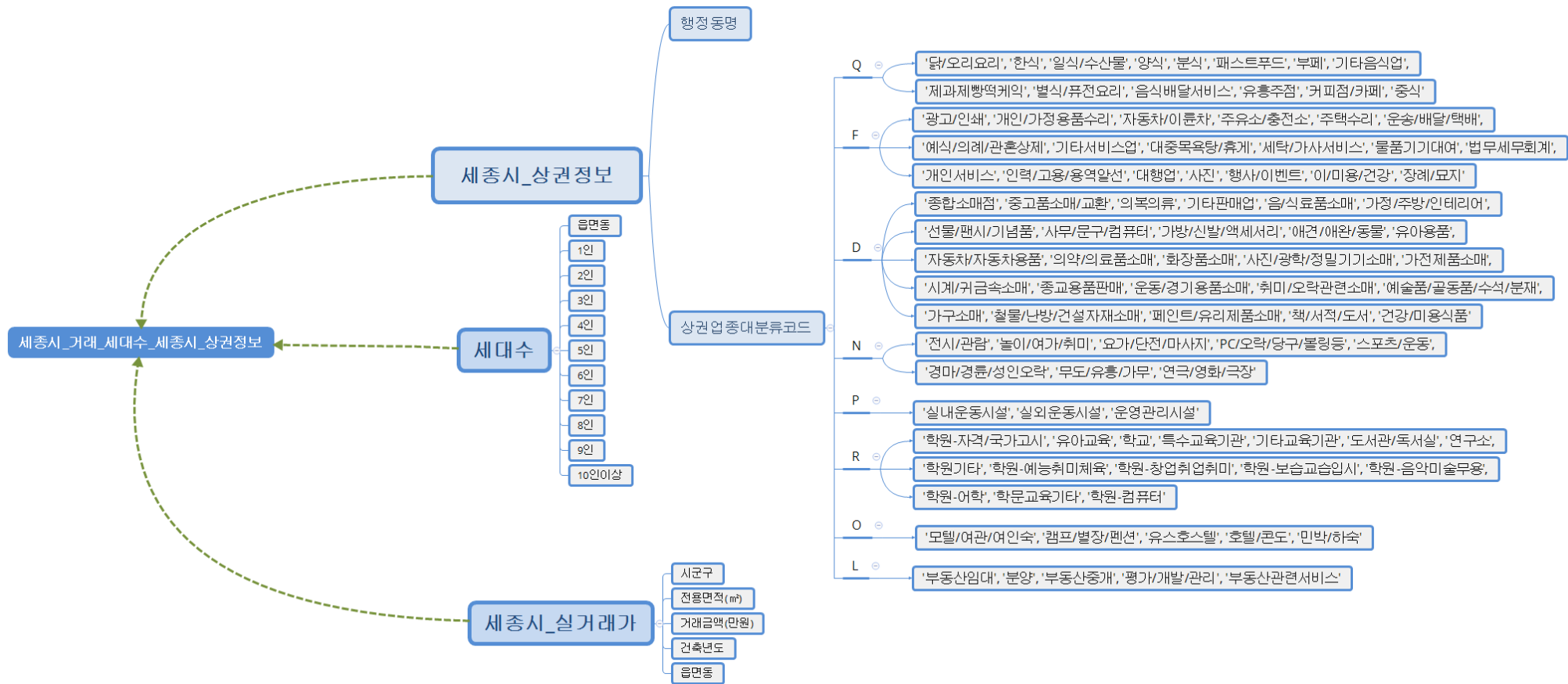
세종시 실거래가

아파트(매매, 전월세,), 연립다세대(매매, 전월세), 단독다가구(매매, 전월세), 오피스텔(매매, 전월세)

	계약년월	계약일	시군구	본번	부번	단지명	전용면적(m ²)	거래금액(만원)	층	건축년도
0	201701	1	세종특별자치시 보람동	628	2	스타힐타워 I (C2-3)	16.9849	8600	7	2016
1	201701	1	세종특별자치시 종촌동	683	0	가재마을2단지(호반베르디움3차)	59.8652	27300	7	2014
2	201701	2	세종특별자치시 고운동	1393	0	가락마을15단지(중흥파크뷰)	84.9894	32000	3	2015
3	201701	2	세종특별자치시 고운동	1706	0	가락마을8단지(고운뜰아파트)	84.9800	30500	20	2015
4	201701	2	세종특별자치시 아름동	1284	0	범지기마을7단지 호반에코시티	84.9280	31300	6	2014
...
21135	202010	27	세종특별자치시 조치원읍 신흥리	399	0	조치원 신흥 e편한세상(아)	176.3430	50000	10	2008
21136	202010	27	세종특별자치시 조치원읍 침산리	89	0	옥일아파트	84.9900	23500	20	1996
21137	202010	28	세종특별자치시 도담동	639	0	도람마을17단지(모아미래도)	84.9690	68000	9	2015
21138	202010	28	세종특별자치시 부강면 부강리	401	48	주연	59.8200	6500	11	1992
21139	202010	28	세종특별자치시 조치원읍 죽림리	396	0	조치원죽림자이	84.6641	37000	6	2008

∴ 건물,계약 종류 에 따라 컬럼 변경

2. 분석 방법론 – 분석 대상 및 타겟 변수 정의



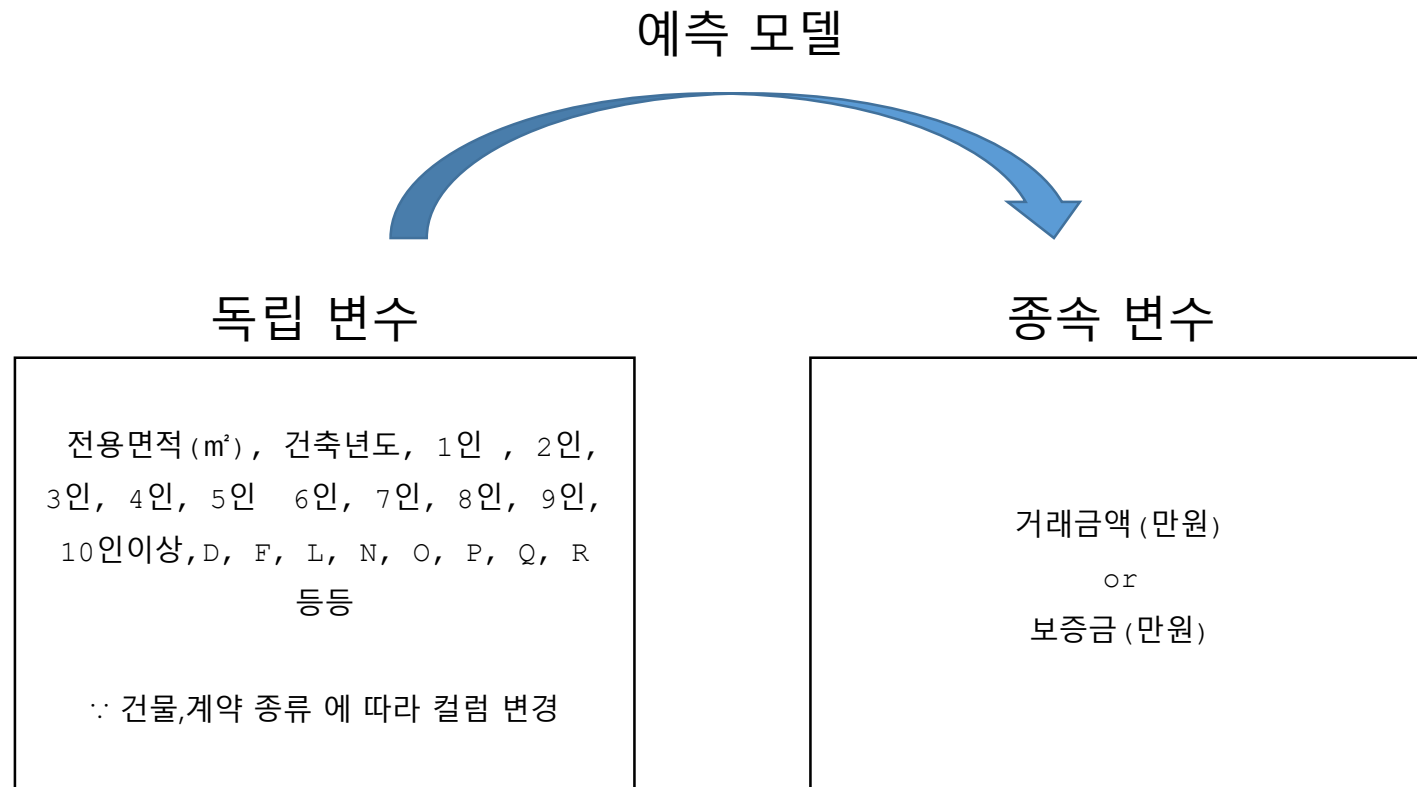
2. 분석 방법론 – 분석 대상 및 타겟 변수 정의

세종시_거래_세대수_세종시_상권정보

	시군구	전용면적 (㎡)	거래금액 (만원)	건축 년도	1인	2인	3인	4인	5인	6인	...	9 인	10 인 이상	D	F	L	N	O	P	Q	R
0	세종특별자치시 보람동	16.9849	8600	5	1536.0	1363.0	1544.0	1802.0	471.0	69.0	...	2.0	1.0	61.0	69.0	19.0	5.0	0.0	5.0	230.0	31.0
1	세종특별자치시 중촌동	59.8652	27300	7	2525.0	2049.0	2420.0	2847.0	675.0	120.0	...	2.0	2.0	87.0	48.0	20.0	1.0	0.0	3.0	131.0	56.0
2	세종특별자치시 고운동	84.9894	32000	6	2025.0	2341.0	2839.0	3240.0	814.0	145.0	...	4.0	2.0	61.0	74.0	25.0	7.0	0.0	3.0	136.0	47.0
3	세종특별자치시 고운동	84.9800	30500	6	2025.0	2341.0	2839.0	3240.0	814.0	145.0	...	4.0	2.0	61.0	74.0	25.0	7.0	0.0	3.0	136.0	47.0
4	세종특별자치시 아름동	84.9280	31300	7	1058.0	1262.0	1858.0	2656.0	597.0	98.0	...	1.0	3.0	45.0	38.0	13.0	3.0	0.0	6.0	115.0	57.0
...
21135	세종특별자치시 조치원읍 신흥리	176.3430	50000	13	8295.0	4165.0	3042.0	2729.0	944.0	220.0	...	4.0	1.0	1133.0	574.0	105.0	97.0	36.0	3.0	1402.0	203.0
21136	세종특별자치시 조치원읍 침산리	84.9900	23500	25	8295.0	4165.0	3042.0	2729.0	944.0	220.0	...	4.0	1.0	1133.0	574.0	105.0	97.0	36.0	3.0	1402.0	203.0
21137	세종특별자치시 도담동	84.9690	68000	6	5002.0	2117.0	2383.0	2825.0	667.0	106.0	...	4.0	0.0	217.0	83.0	51.0	13.0	0.0	3.0	388.0	48.0
21138	세종특별자치시 부강면 부강리	59.8200	6500	29	1603.0	821.0	420.0	273.0	93.0	34.0	...	0.0	2.0	135.0	69.0	14.0	11.0	8.0	0.0	189.0	10.0
21139	세종특별자치시 조치원읍 죽림리	84.6641	37000	13	8295.0	4165.0	3042.0	2729.0	944.0	220.0	...	4.0	1.0	1133.0	574.0	105.0	97.0	36.0	3.0	1402.0	203.0

21140 rows × 22 columns

3. 이슈사항 해결 – 어떤 모델을 어떻게 사용할 것인가?



랜덤포레스트, 라쏘를 통해 예측 진행

4. 최종모델 및 결과해석

랜덤서치 + 랜덤 포레스트 (세종시_아파트(매매)_실거래가)

Best model: {'max_depth': 11, 'max_features': 13, 'n_estimators': 143}

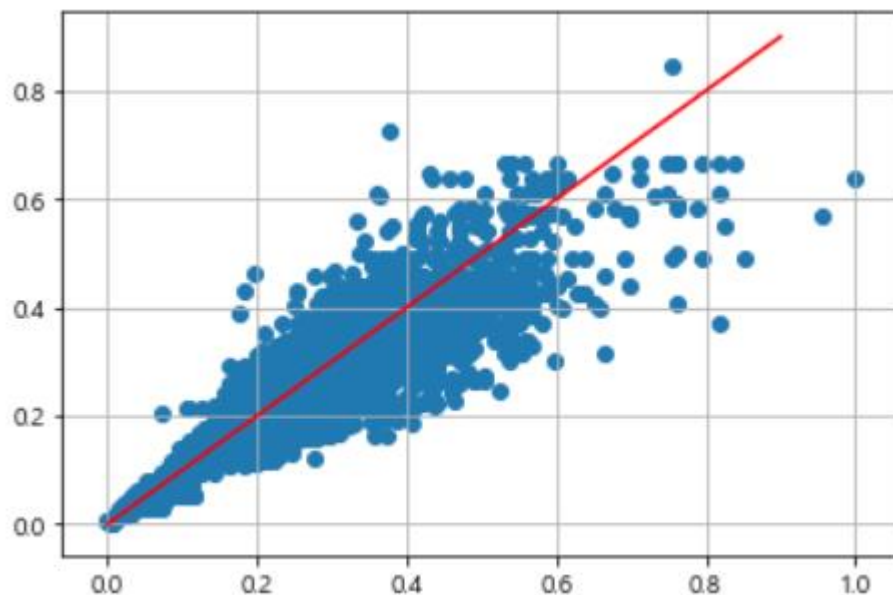
X_train,y_train Accuracy: 84.66%

X_test,y_test Accuracy: 83.16%

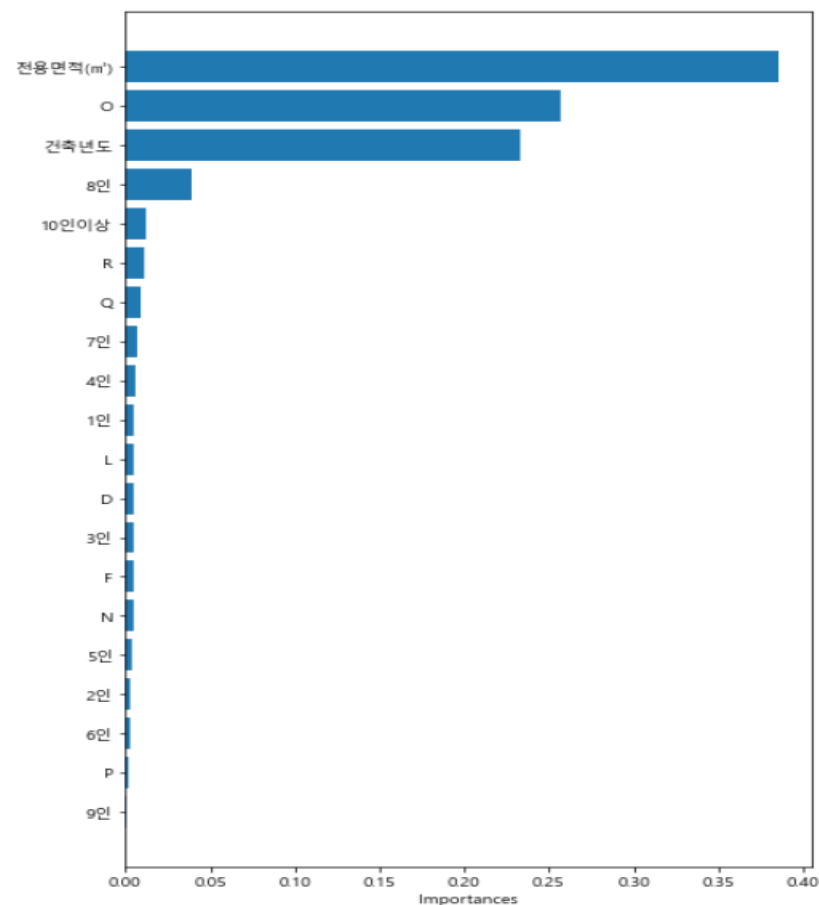
y_test,pred Mse 0.034043549248270036

y_test,pred SQR Mse 0.05190358350492426

정확도 산점도



변수 중요도

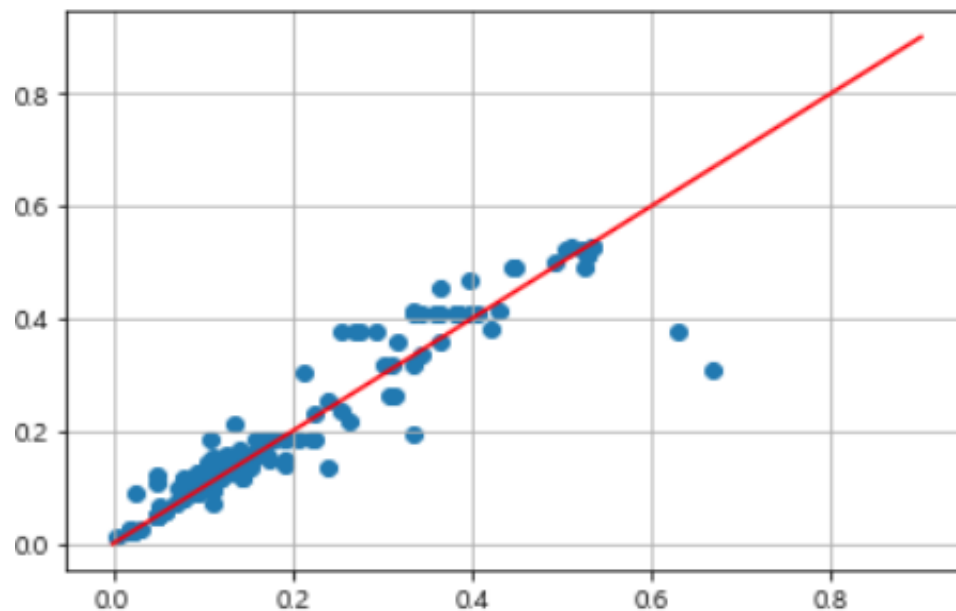


4. 최종모델 및 결과해석

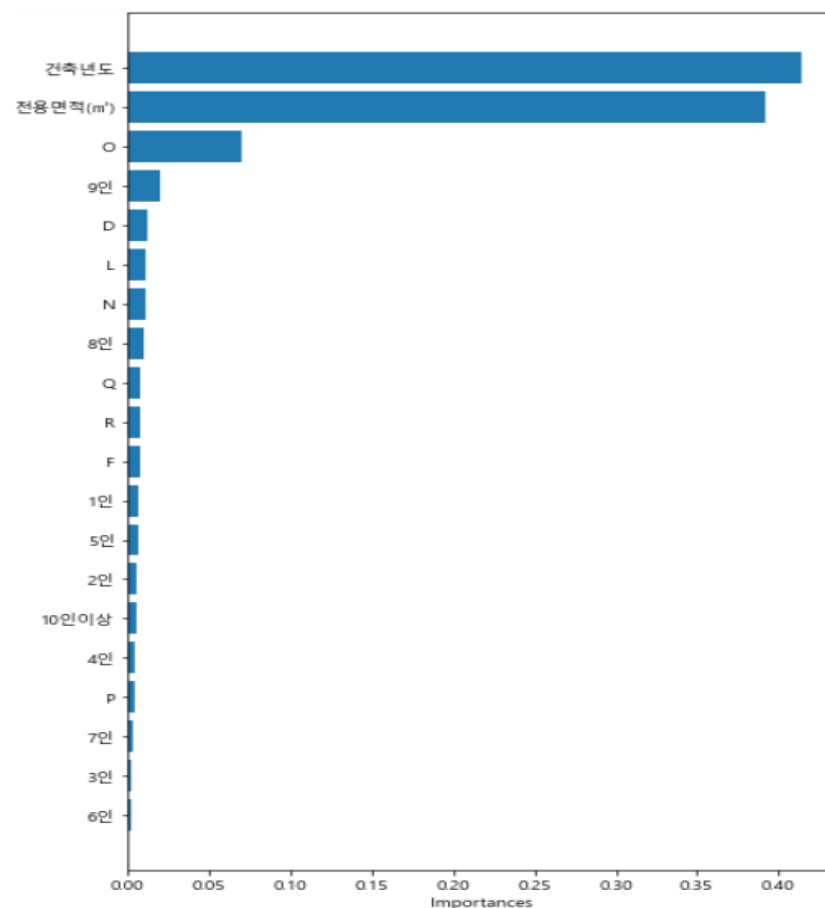
랜덤서치 + 랜덤 포레스트(세종시_연립다세대(매매)_실거래가)

Best model: {'max_depth': 19, 'max_features': 11, 'n_estimators': 72}
X_train,y_train Accuracy: 94.73%
X_test,y_test Accuracy: 89.22%
y_test,pred Mse 0.025255882803287968
y_test,pred SQRT Mse 0.049121774342569736

정확도 산점도



변수 중요도



4. 최종모델 및 결과해석

랜덤서치 + 랜덤 포레스트(세종시_단독다가구(매매)_실거래가)

Best model: {'max_depth': 12, 'max_features': 16, 'n_estimators': 82}

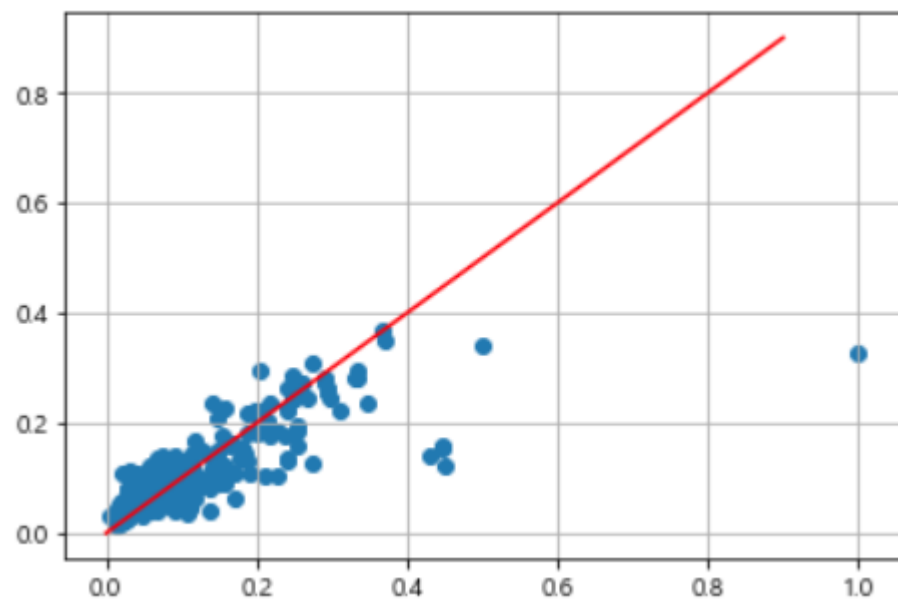
X_train,y_train Accuracy: 96.13%

X_test,y_test Accuracy: 66.12%

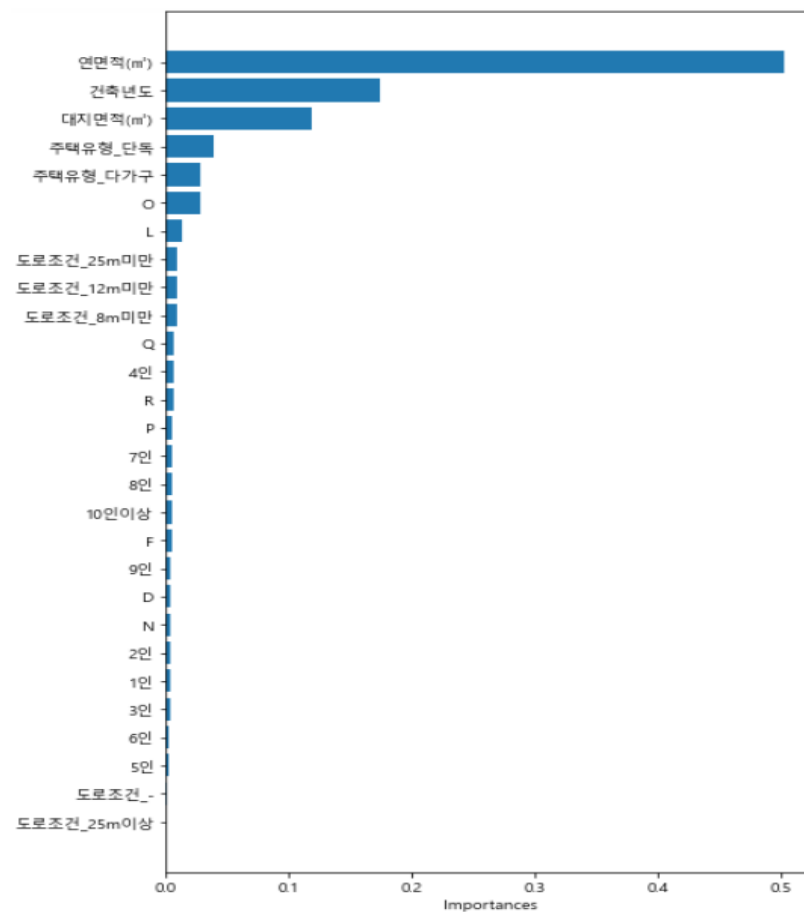
y_test,pred Mse 0.03135912804845423

y_test,pred SQRT Mse 0.061491775814701576

정확도 산점도



변수 중요도



4. 최종모델 및 결과해석

랜덤서치 + 랜덤 포레스트(세종시_오피스텔(매매)_실거래가)

Best model: {'max_depth': 8, 'max_features': 18, 'n_estimators': 104}

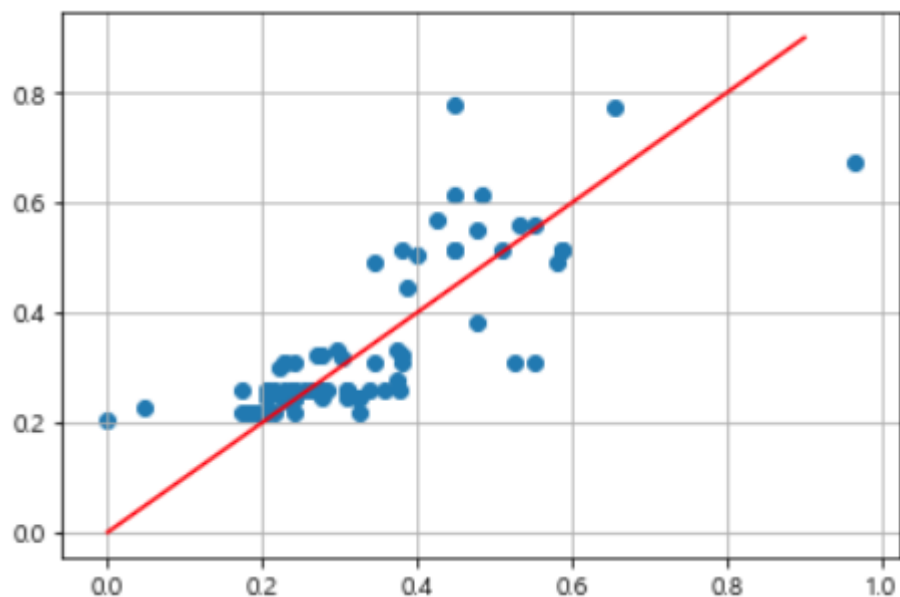
X_train,y_train Accuracy: 87.46%

X_test,y_test Accuracy: 58.5%

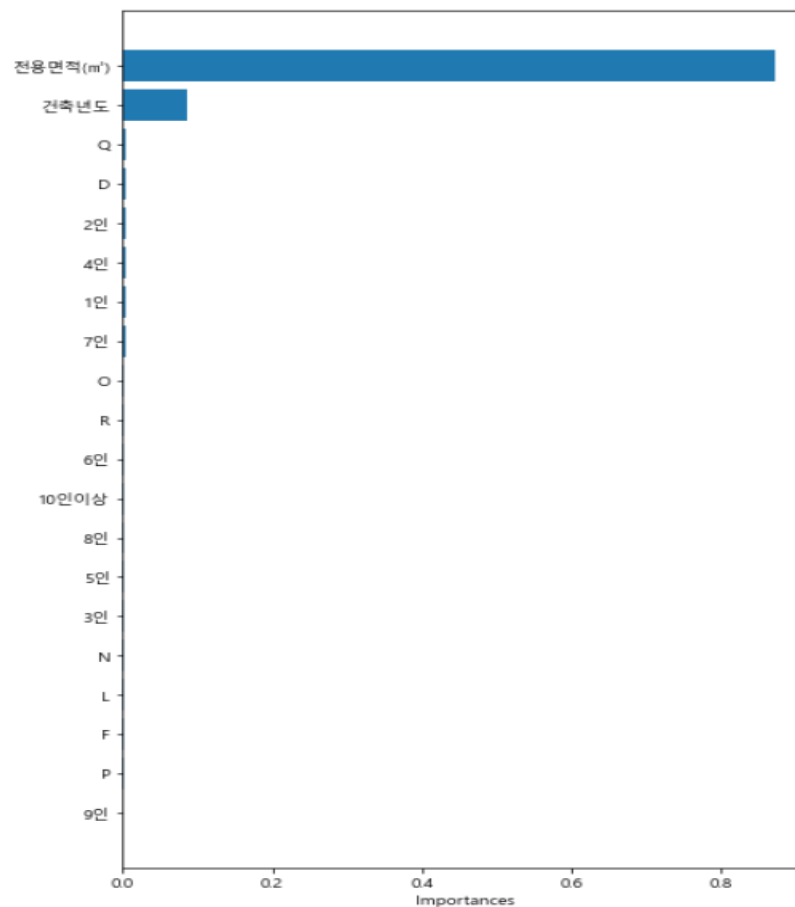
y_test,pred Mse 0.06563561729094496

y_test,pred SQRT Mse 0.09220300043699634

정확도 산점도



변수 중요도



4. 최종모델 및 결과해석

랜덤서치 + 랜덤 포레스트(세종시_아파트(전월세)_실거래가)

Best model: {'max_depth': 17, 'max_features': 19, 'n_estimators': 63}

X_train,y_train Accuracy: 88.28%

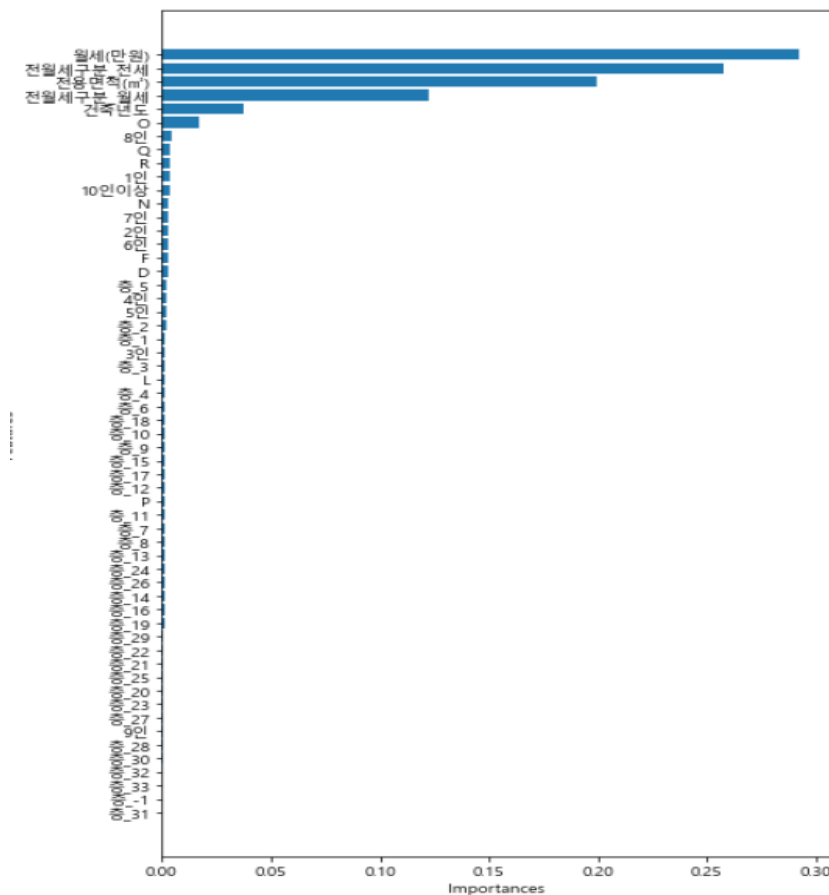
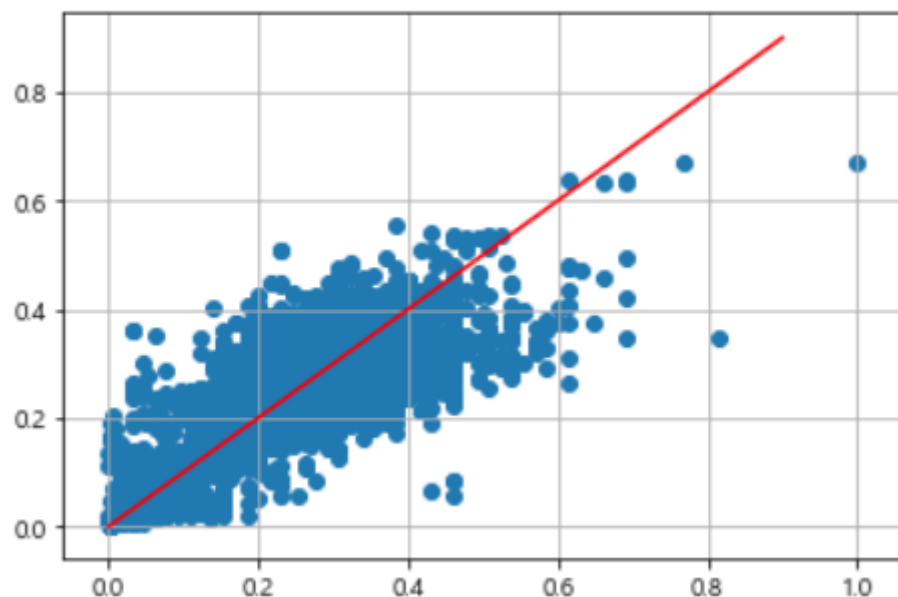
X_test,y_test Accuracy: 83.02%

y_test,pred Mse 0.03257029059776794

y_test,pred SQRT Mse 0.05032289610634758

변수 중요도

정확도 산점도



4. 최종모델 및 결과해석

랜덤서치 + 랜덤 포레스트(세종시_연립다세대(전월세)_실거래가)

Best model: {'max_depth': 19, 'max_features': 15, 'n_estimators': 190}

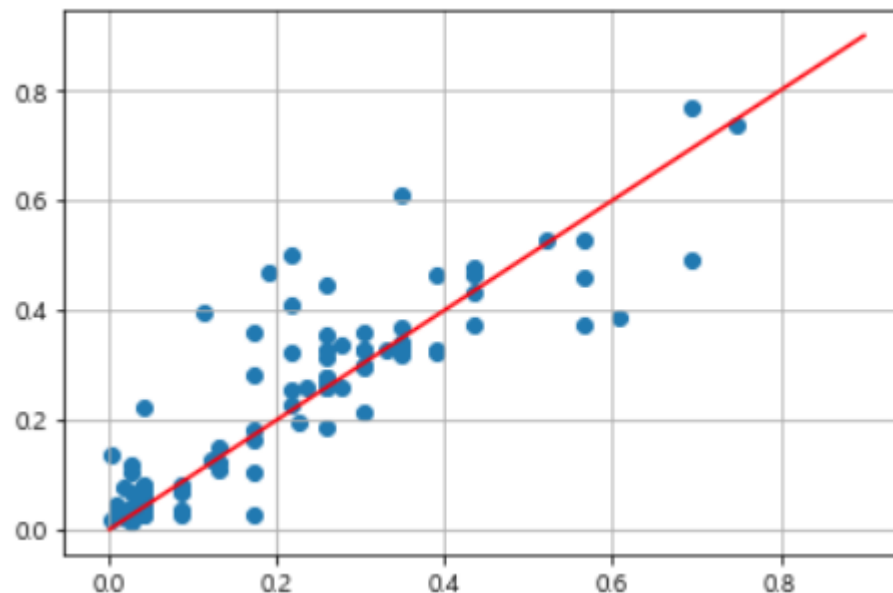
X_train,y_train Accuracy: 97.11%

X_test,y_test Accuracy: 76.28%

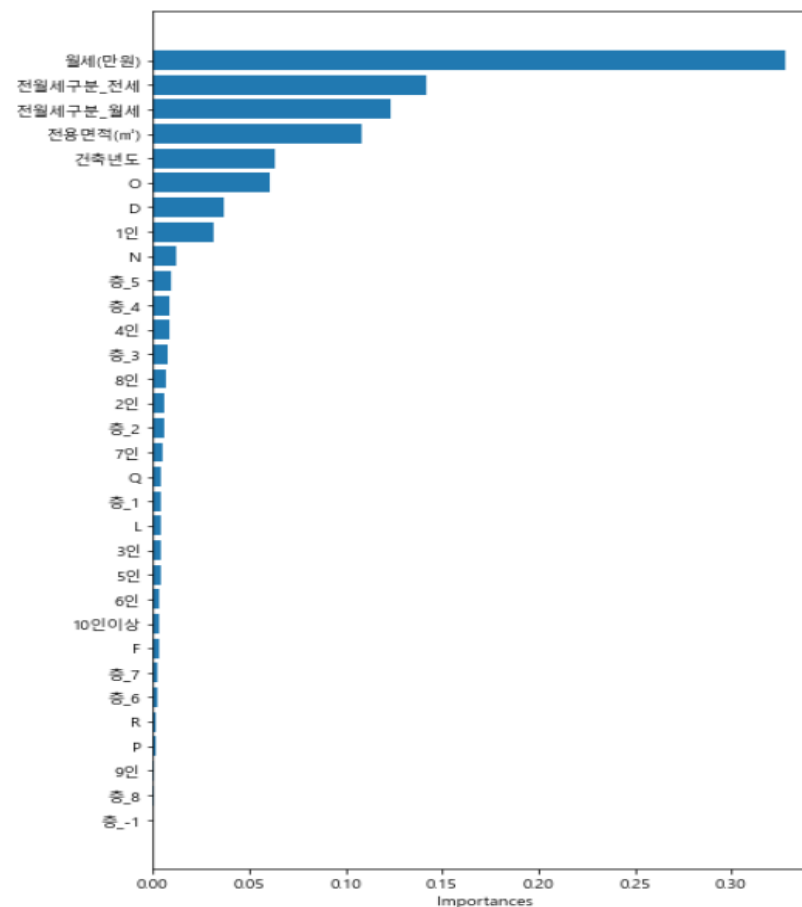
y_test,pred Mse 0.0496608704049944

y_test,pred SQRT Mse 0.08537384894025656

정확도 산점도



변수 중요도



4. 최종모델 및 결과해석

랜덤서치 + 랜덤 포레스트(세종시_단독다가구(전월세)_실거래가)

Best model: {'max_depth': 10, 'max_features': 12, 'n_estimators': 89}

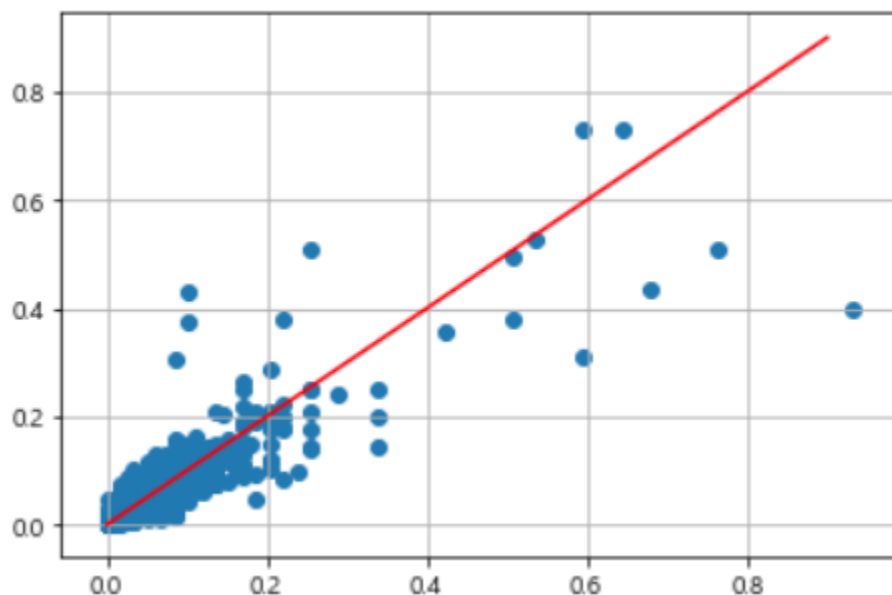
X_train,y_train Accuracy: 90.71%

X_test,y_test Accuracy: 77.09%

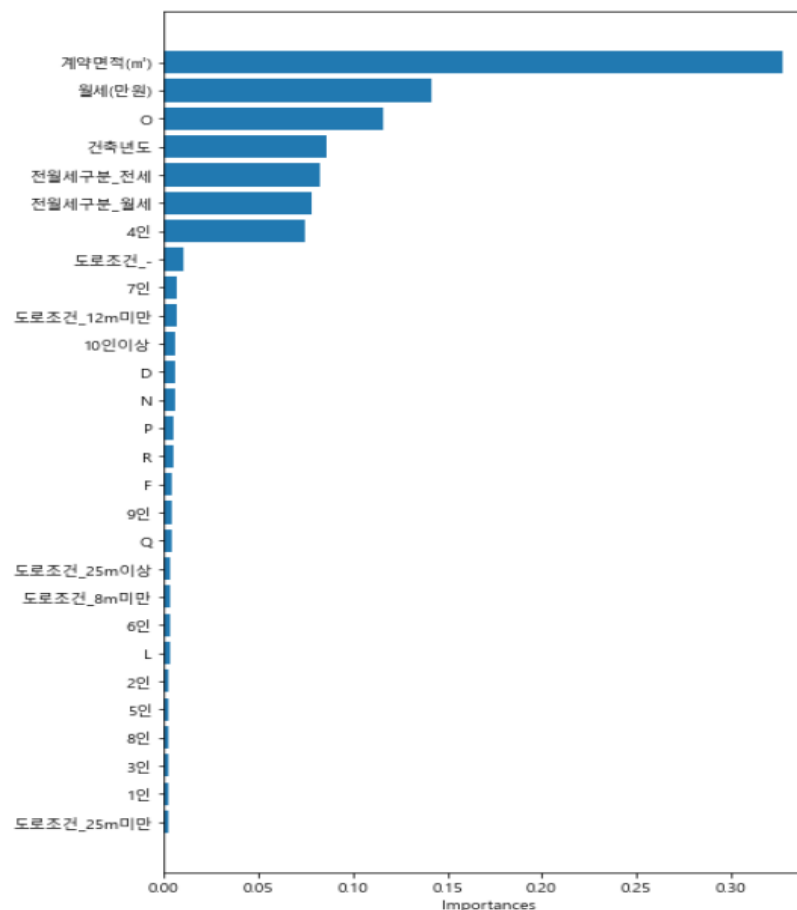
y_test,pred Mse 0.01576042607837237

y_test,pred SQRT Mse 0.033132577820334935

정확도 산점도



변수 중요도



4. 최종모델 및 결과해석

랜덤서치 + 랜덤 포레스트(세종시_오피스텔(전월세)_실거래가)

Best model: {'max_depth': 11, 'max_features': 9, 'n_estimators': 15}

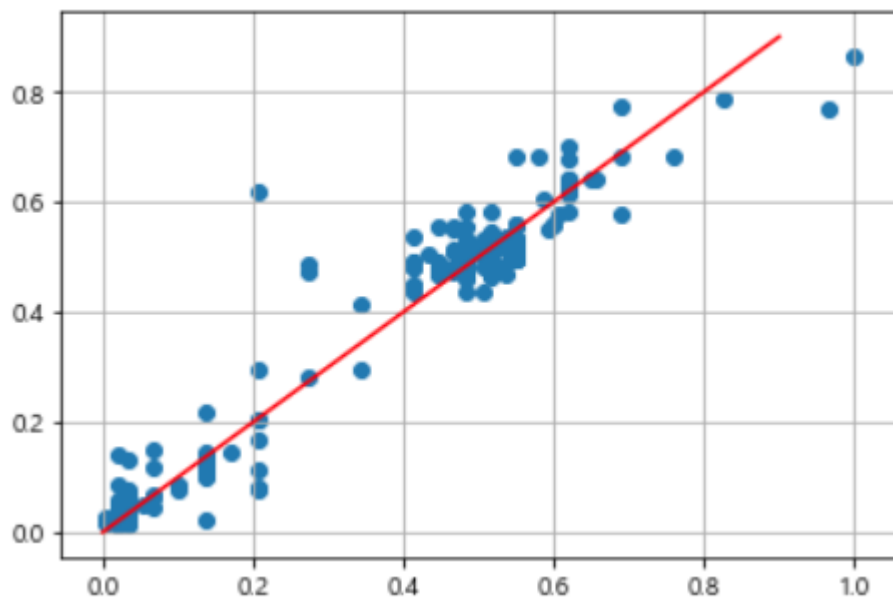
X_train,y_train Accuracy: 98.26%

X_test,y_test Accuracy: 97.05%

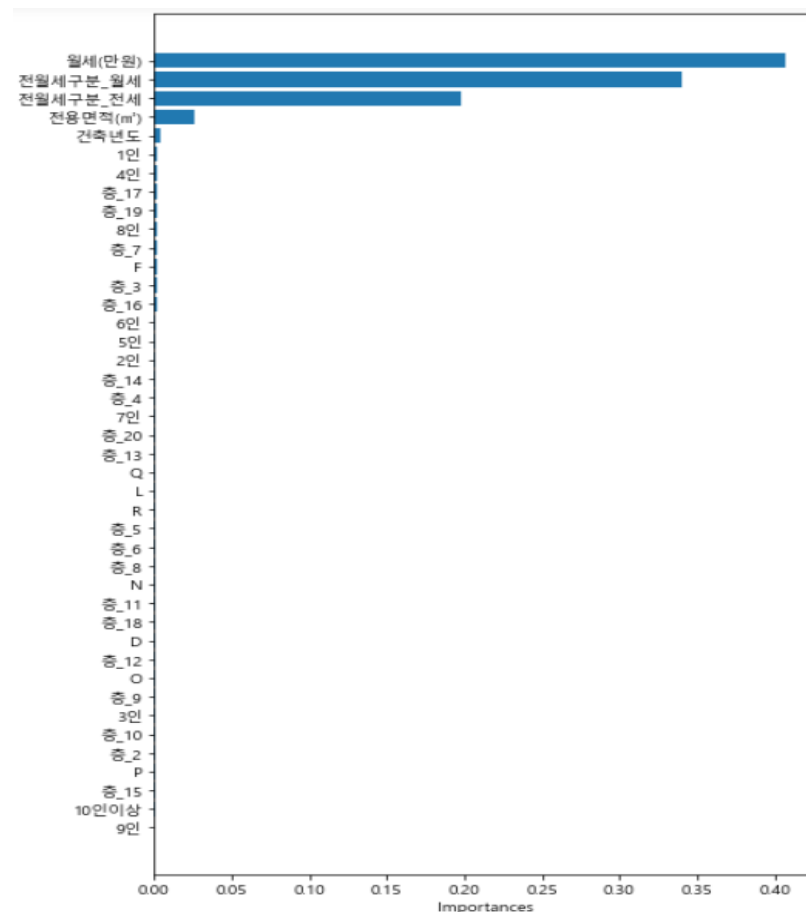
y_test,pred Mse 0.024856891751340524

y_test,pred SQRT Mse 0.042408026825100587

정확도 산점도



변수 중요도



4. 최종모델 및 결과해석

랜덤 포레스트 중요변수 결과 및 한계

아파트(매매)

순위	변수
1	전용면적
2	분류 O
3	건축년도

연립다세대(매매)

순위	변수
1	건축년도
2	전용면적
3	분류 O

단독다가구(매매)

순위	변수
1	연면적
2	건축년도
3	대지면적

오피스텔(매매)

순위	변수
1	전용면적
2	건축년도
3	-

아파트(전월세)

순위	변수
1	월세(만원)
2	전월세구분 _전세
3	전용면적

연립다세대(전월세)

순위	변수
1	월세(만원)
2	전월세구분 _전세
3	전월세구분 _월세

단독다가구(전월세)

순위	변수
1	계약면적
2	월세(만원)
3	분류 O

오피스텔(전월세)

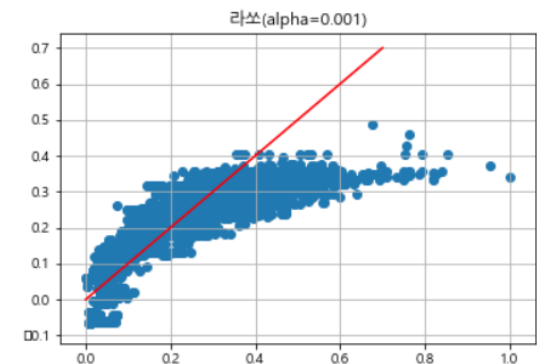
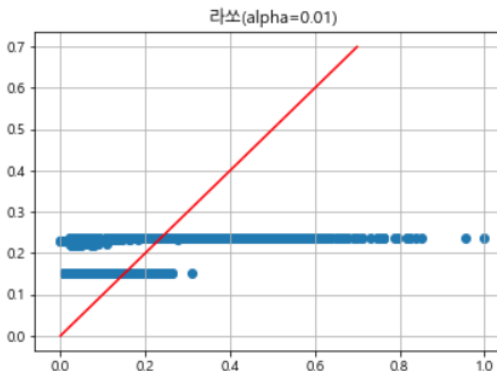
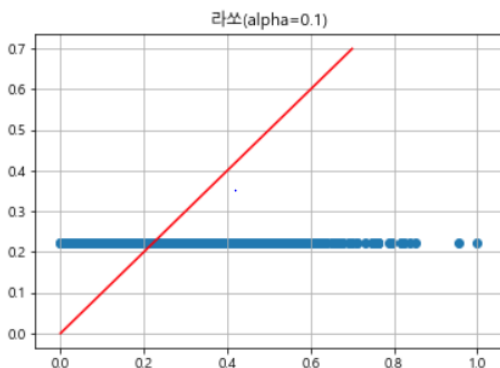
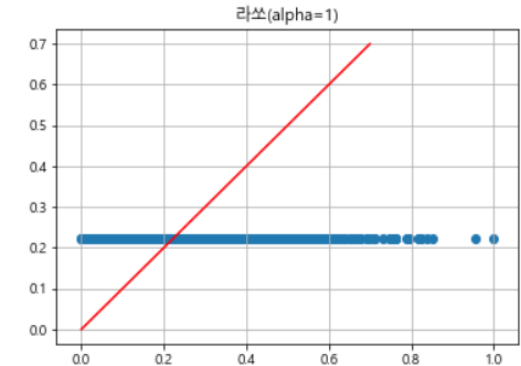
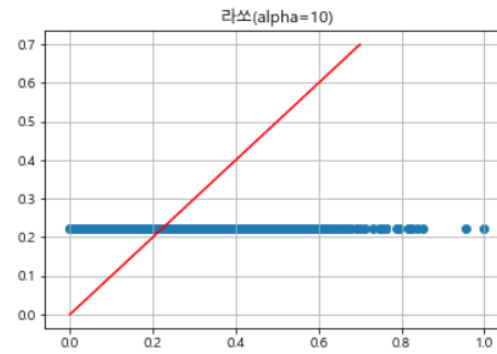
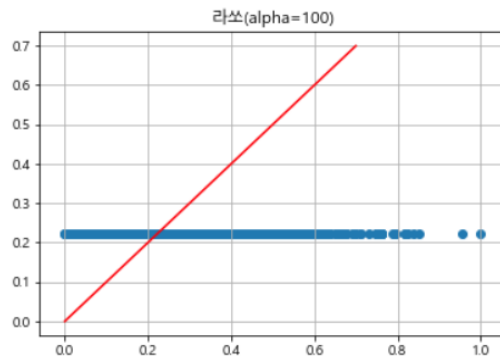
순위	변수
1	월세(만원)
2	전용면적
3	전월세구분 _전세

한계 : 변수 영향이 음인지 양인지 알 수 없다.

4. 최종모델 및 결과해석

라쏘 모델

알파(컬럼 컨트롤)값 을 100, 10, 1, 0.1, 0.01, 0.001, 0.0001을 통해 결과를 확인 하였고 알파값 0.001이 변수 중요성을 잘 설명 해주어 알파값 0.001을 기반으로 모델 결과 해석



알파값을 낮추면 규제의 효과가 없어져 과대적합이 되므로 LinearRegression의 결과와 비슷해진다

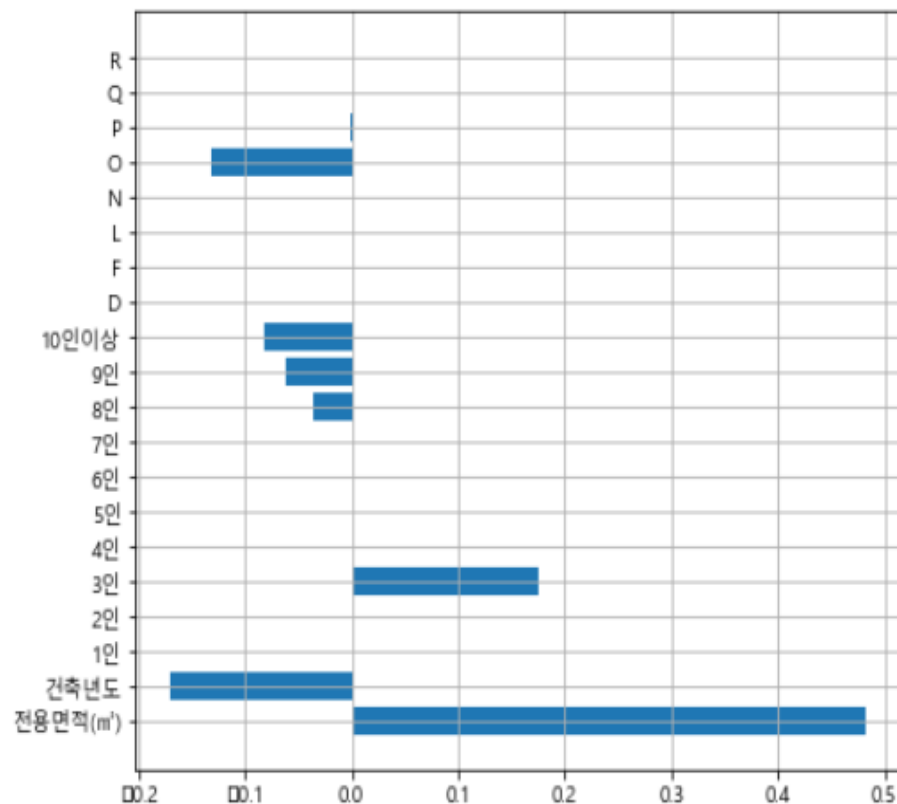
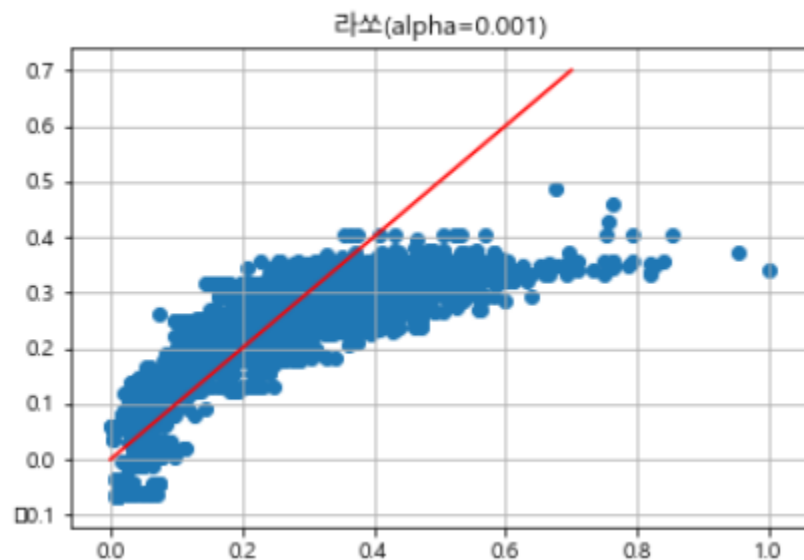
4. 최종모델 및 결과해석

라쏘 (세종시_아파트(매매)_실거래가)

X_train,y_train Accuracy: 60.86%
X_test,y_test Accuracy: 60.11%
lasso alpha0.001 y_test,pred Mse 0.059497946129468814
lasso alpha0.001 y_test,pred SQRT Mse 0.07988539857730388

변수 중요도

정확도 산점도



4. 최종모델 및 결과해석

라쏘(세종시_연립다세대(매매)_실거래가)

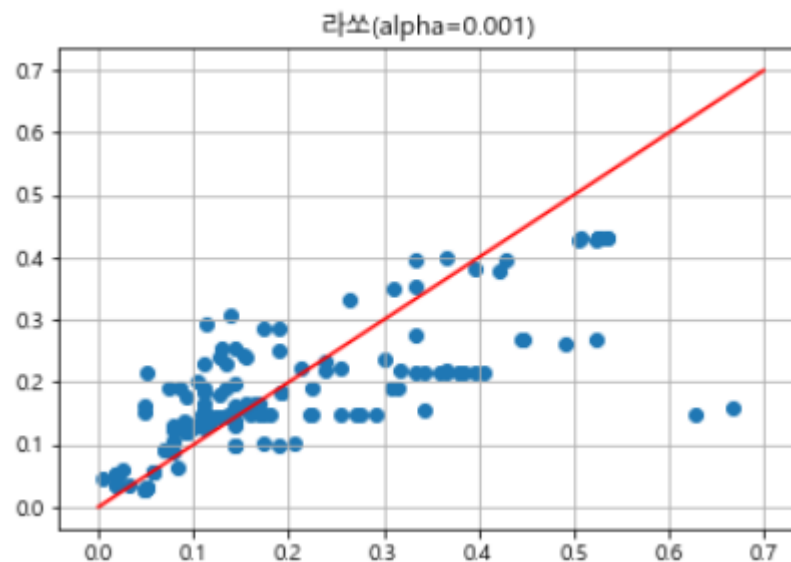
X_train,y_train Accuracy: 54.63%

X_test,y_test Accuracy: 56.24%

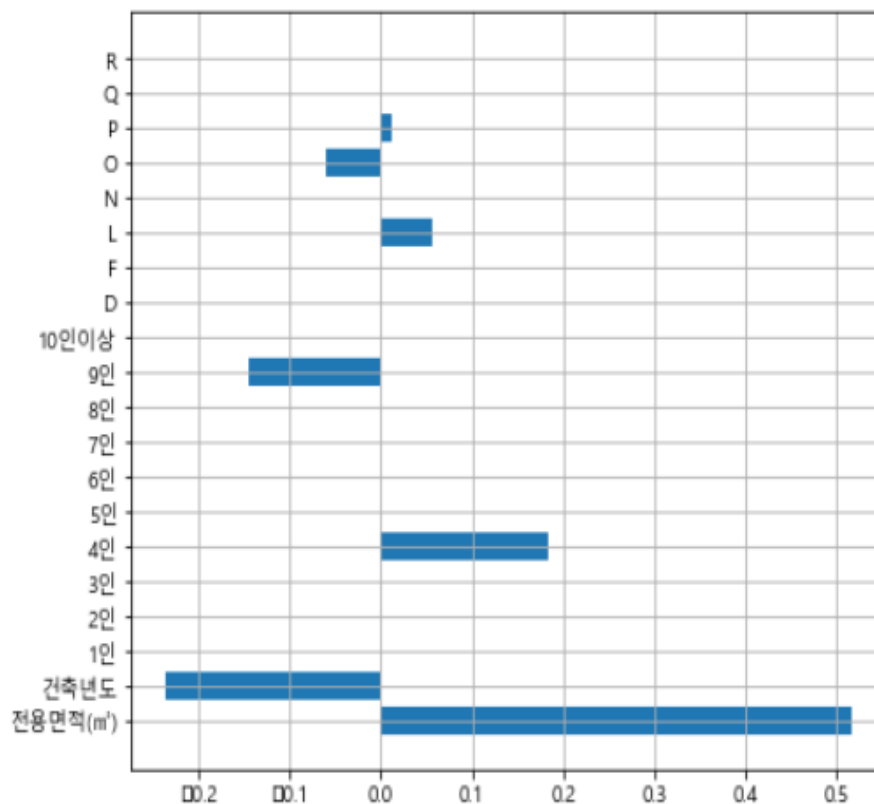
lasso alpha0.001 y_test,pred Mse 0.06763123582615367

lasso alpha0.001 y_test,pred SQR T Mse 0.0989741637958161

정확도 산점도



변수 중요도



4. 최종모델 및 결과해석

라쏘(세종시_단독다가구(매매)_실거래가)

X_train,y_train Accuracy: 65.16%

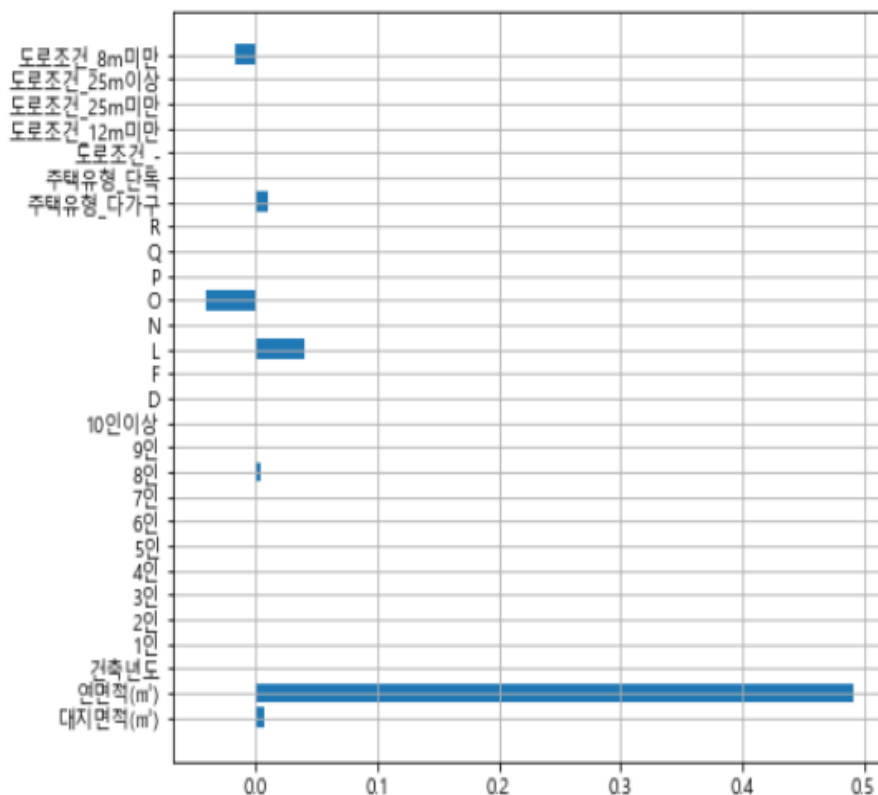
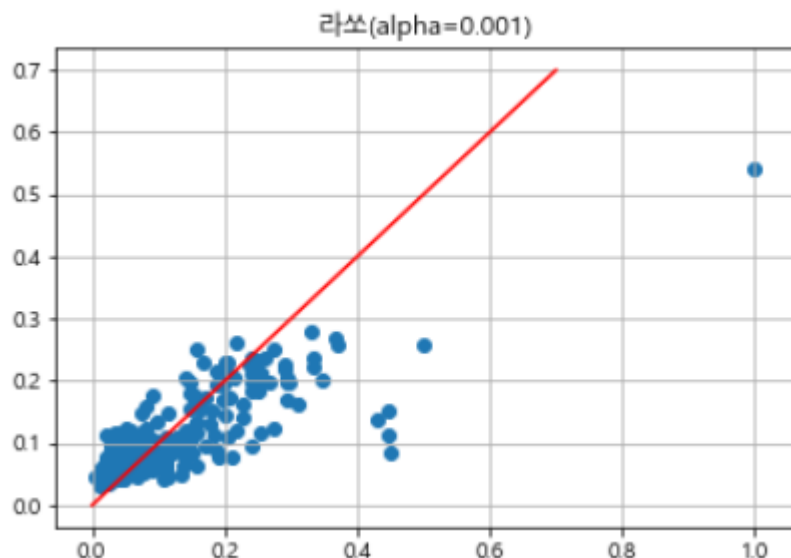
X_test,y_test Accuracy: 61.31%

lasso alpha0.001 y_test,pred Mse 0.04049445360593398

lasso alpha0.001 y_test,pred SQRT Mse 0.06570674399216708

변수 중요도

정확도 산점도



4. 최종모델 및 결과해석

라쏘(세종시_오피스텔(매매)_실거래가)

X_train,y_train Accuracy: 68.68%

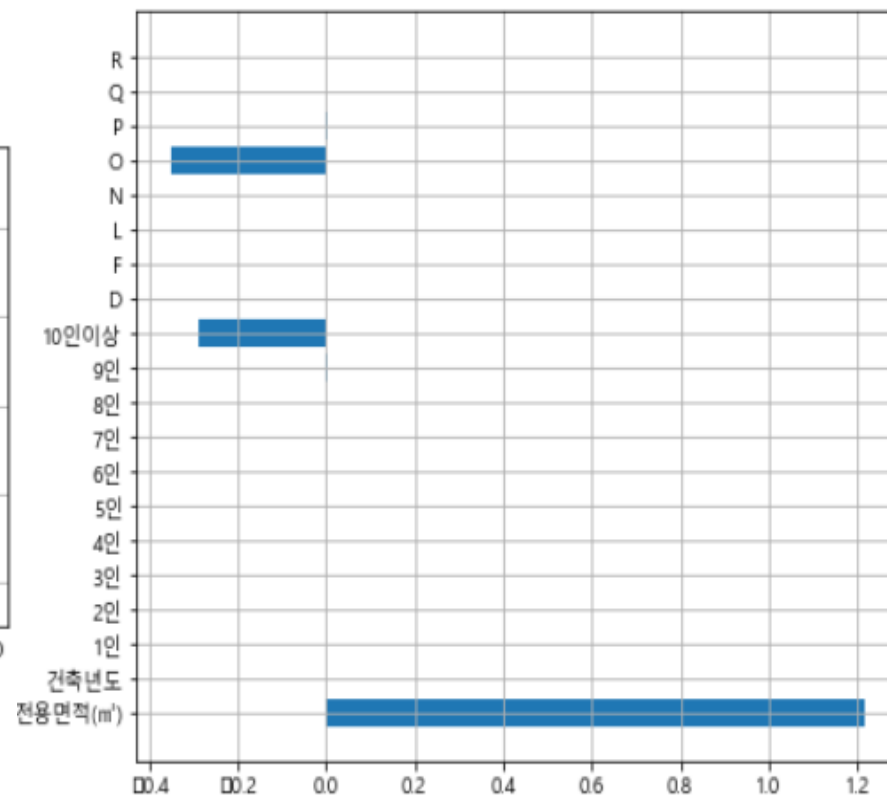
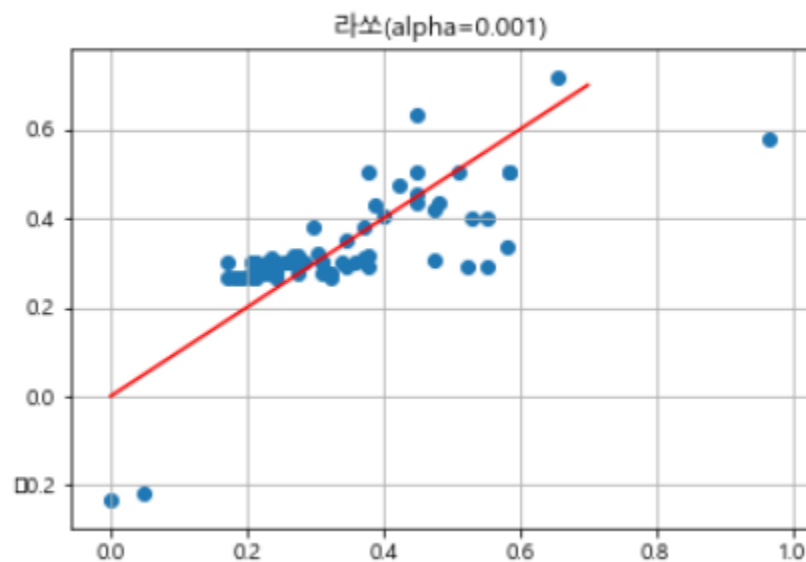
X_test,y_test Accuracy: 51.77%

lasso alpha0.001 y_test,pred Mse 0.07262588594932184

lasso alpha0.001 y_test,pred SQRT Mse 0.09938986599572813

변수 중요도

정확도 산점도



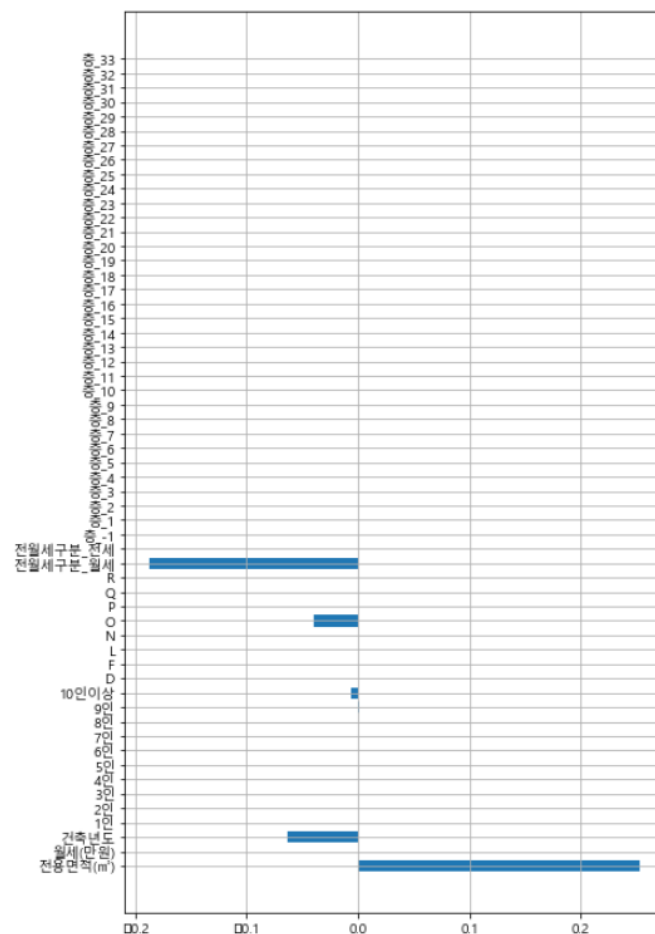
라쏘(세종시_아파트(전월세)_실거래가)

X_test,y_test Accuracy: 69.94%

```
lasso alpha0.001 y_test,pred Mse 0.0510622556462008
```

```
lasso alpha0.001 y_test,pred SQRT Mse 0.06694635708515383
```

변수 중요도



4. 최종모델 및 결과해석

라쏘(세종시_연립다세대(전월세)_실거래가)

X_train,y_train Accuracy: 74.74%

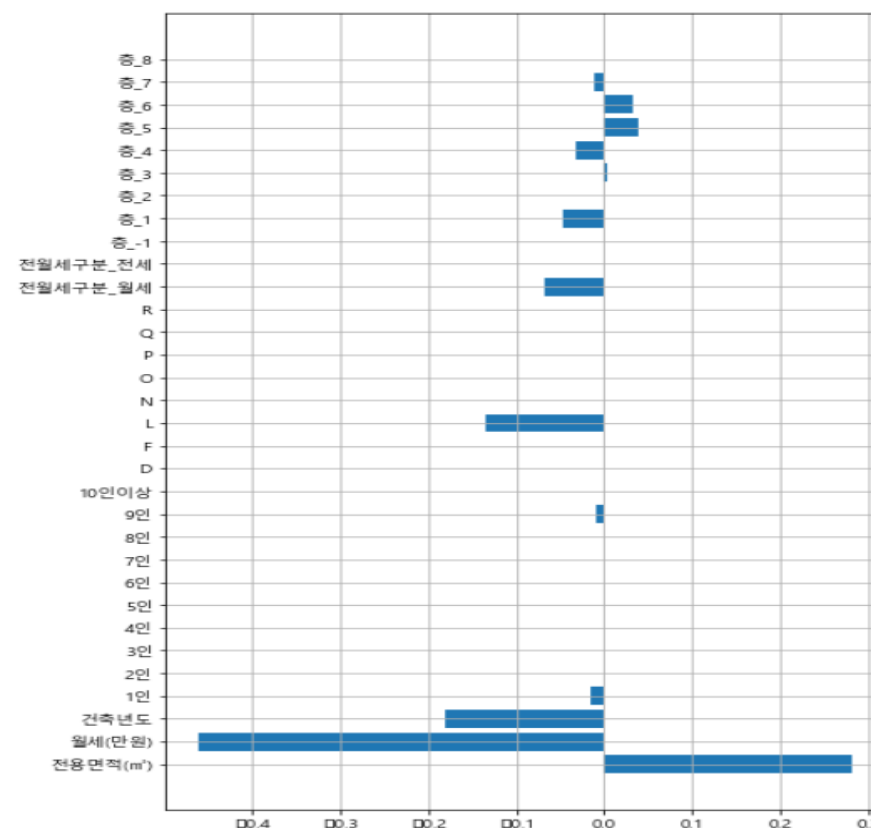
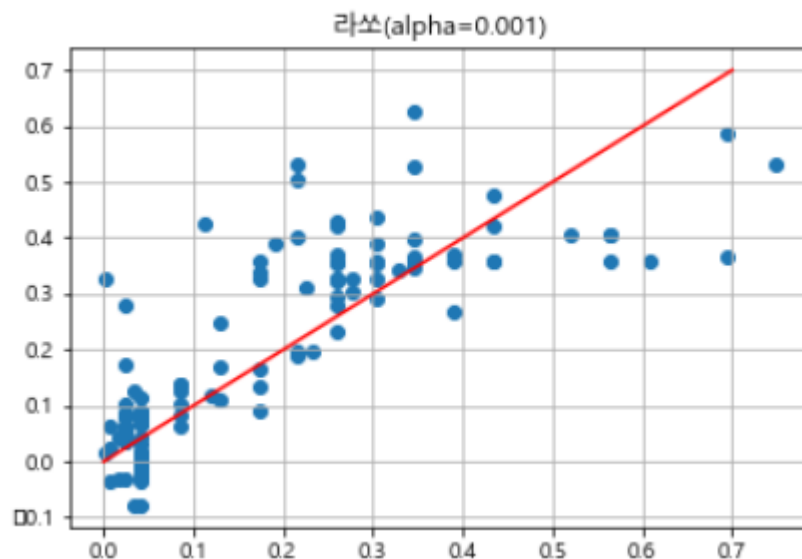
X_test,y_test Accuracy: 60.77%

lasso alpha0.001 y_test,pred Mse 0.07719118092952089

lasso alpha0.001 y_test,pred SQRT Mse 0.10979456029394626

변수 중요도

정확도 산점도



4. 최종모델 및 결과해석

라쏘(세종시_단독다가구(전월세)_실거래가)

X_train,y_train Accuracy: 51.83%

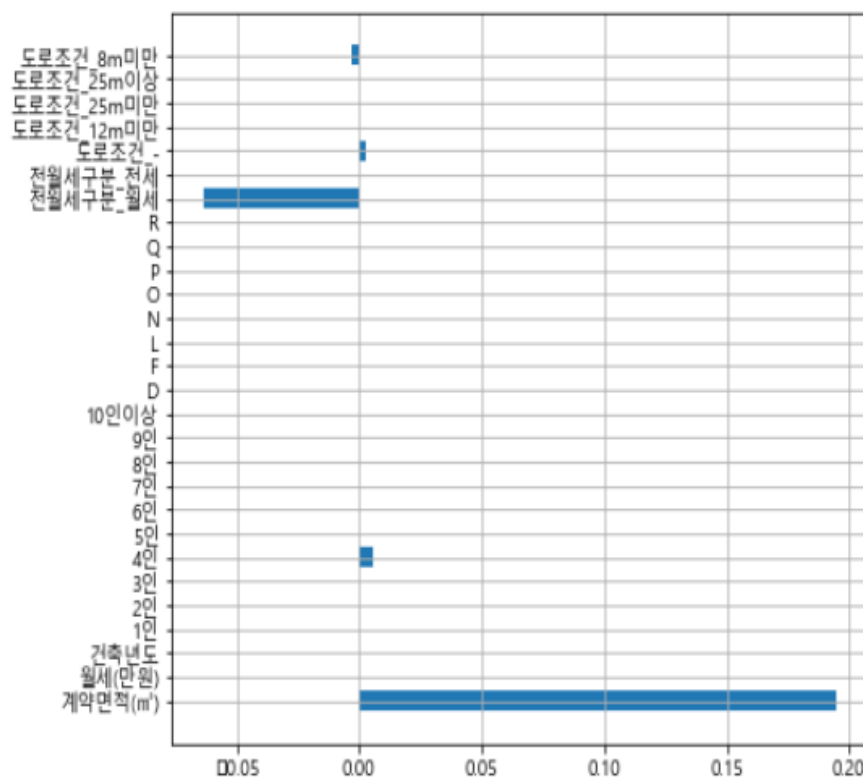
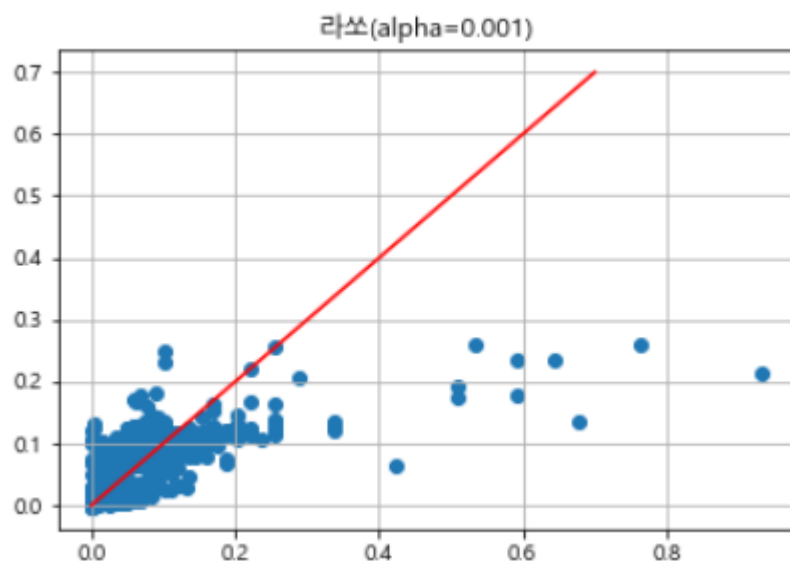
X_test,y_test Accuracy: 47.16%

lasso alpha0.001 y_test,pred Mse 0.023873462380675378

lasso alpha0.001 y_test,pred SQRT Mse 0.0503162219781248

변수 중요도

정확도 산점도



4. 최종모델 및 결과해석

라쏘(세종시_오피스텔(전월세)_실거래가)

X_train,y_train Accuracy: 94.17%

X_test,y_test Accuracy: 93.11%

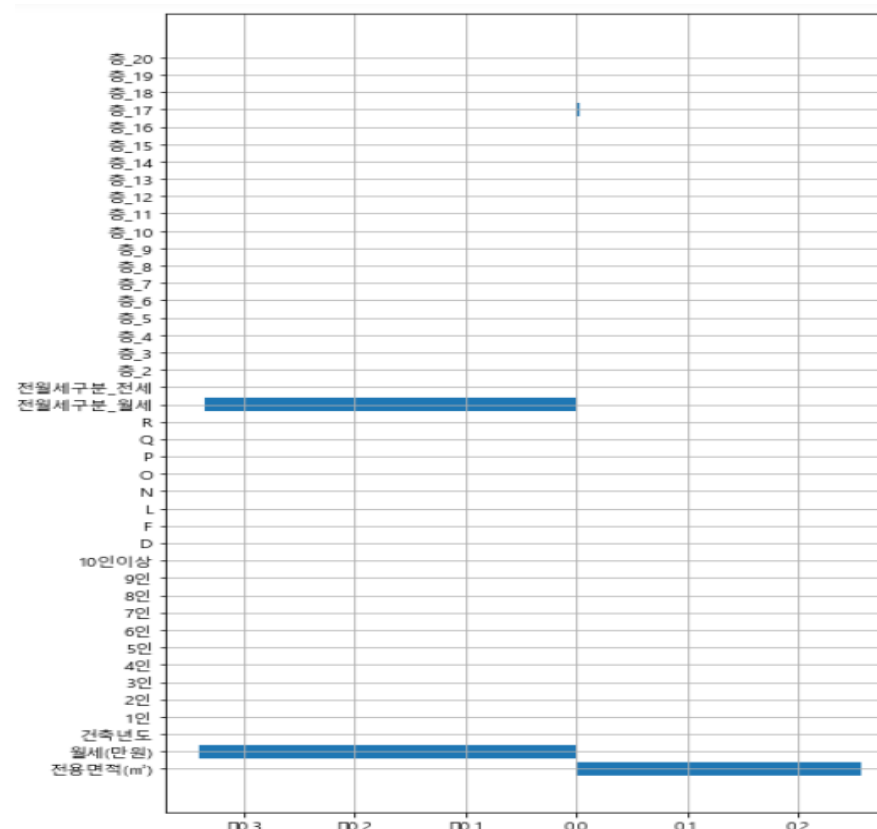
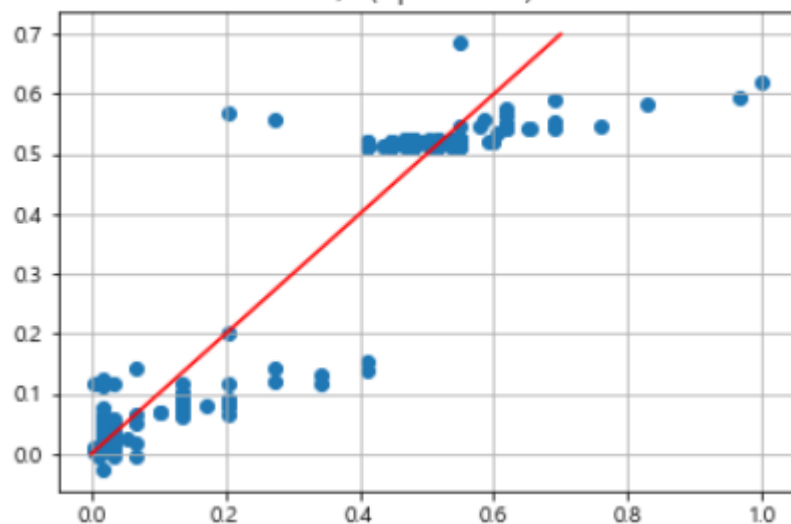
lasso alpha0.001 y_test,pred Mse 0.039180075399792395

lasso alpha0.001 y_test,pred SQRT Mse 0.06487116243057316

변수 중요도

정확도 산점도

라쏘(alpha=0.001)



4. 최종모델 및 결과해석

라쏘 중요변수 결과 (양)

아파트(매매)

순위	변수
1	전용면적
2	3인
3	-

연립다세대(매매)

순위	변수
1	전용면적
2	4인
3	분류 L

단독다가구(매매)

순위	변수
1	연면적
2	분류 L
3	주택유형 다가구

오피스텔(매매)

순위	변수
1	전용면적
2	-
3	-

아파트(전월세)

순위	변수
1	전용면적
2	-
3	-

연립다세대(전월세)

순위	변수
1	전용면적
2	5층
3	6층

단독다가구(전월세)

순위	변수
1	계약면적
2	4인
3	도로조건 _

오피스텔(전월세)

순위	변수
1	전용면적
2	-
3	-

4. 최종모델 및 결과해석

라쏘 중요변수 결과(음) 및 한계

아파트(매매)

순위	변수
1	건축년도
2	분류 O
3	10인 이상

연립다세대(매매)

순위	변수
1	건축년도
2	9인
3	분류 O

단독다가구(매매)

순위	변수
1	분류 O
2	도로조건 8m 미만
3	-

오피스텔(매매)

순위	변수
1	분류 O
2	10인 이상
3	-

아파트(전월세)

순위	변수
1	전월세구분 _월세
2	건축년도
3	분류 O

연립다세대(전월세)

순위	변수
1	월세(만원)
2	건축년도
3	분류 L

단독다가구(전월세)

순위	변수
1	전월세구분 _월세
2	-
3	-

오피스텔(전월세)

순위	변수
1	전월세구분 _월세
2	월세(만원)
3	-

한계 : 낮은 정확도

5. 현업에서 활용 가능한 부분 및 시사점

- 정확도 : 랜덤 서치 + 랜덤 포레스트
- 독립변수 양/음 확인 : 라쏘(Lasso)
- 부동산종류(매매,전월세 등)에 따른 변수 변화 및 순위 파악
- 부동산 거래시 가격 결정에 영향을 미치는지 변수 파악
- 예측 하지 못한 변수를 기점으로 원인 파악 진행
- 상권이 생각보다 가격결정에 큰 영향을 못 미침(분류 0는 가격결정에 음의 관계를 가짐)