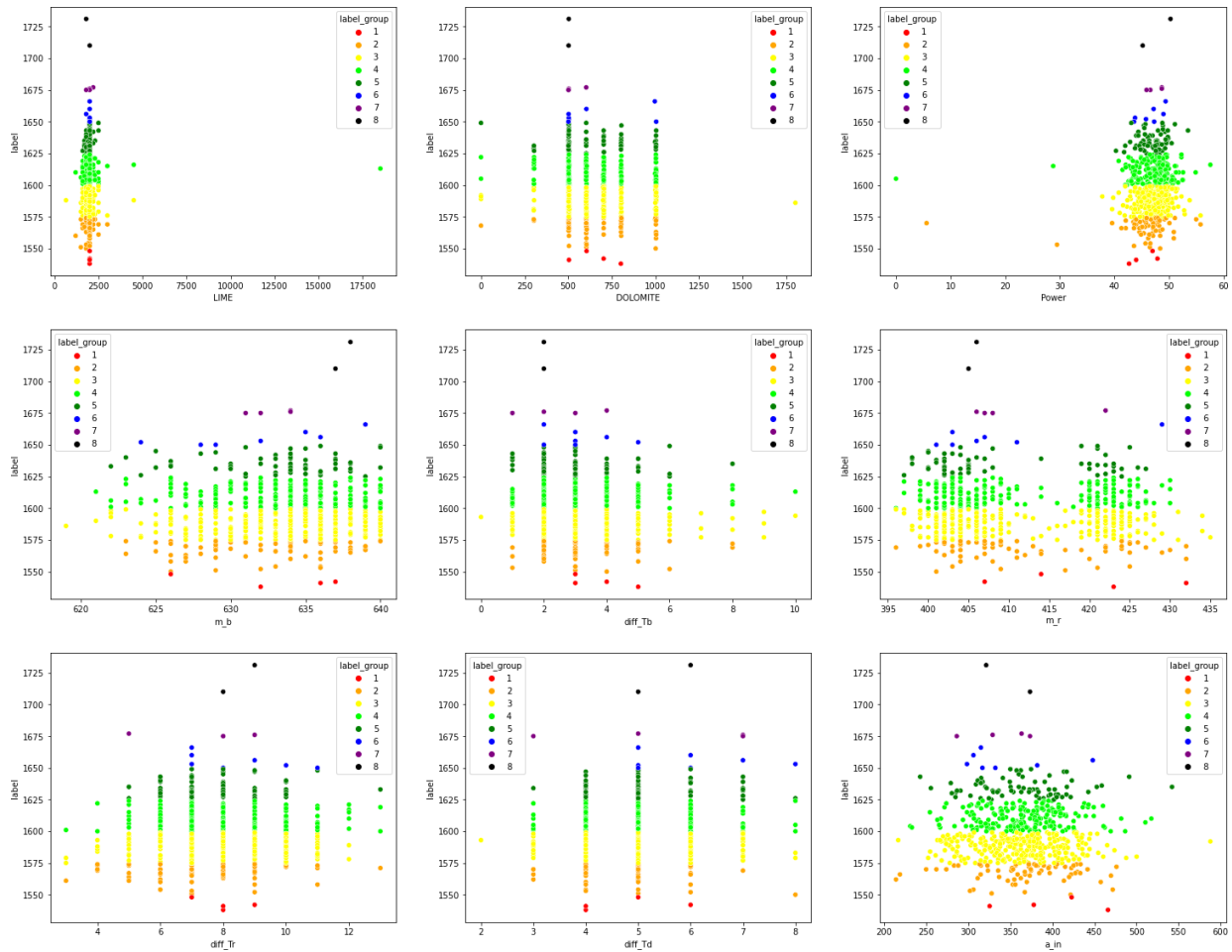


출강온도 예측 오차 분석 결과

RandomForest 모델의 출강온도 예측 성능 : **RMSE**

23



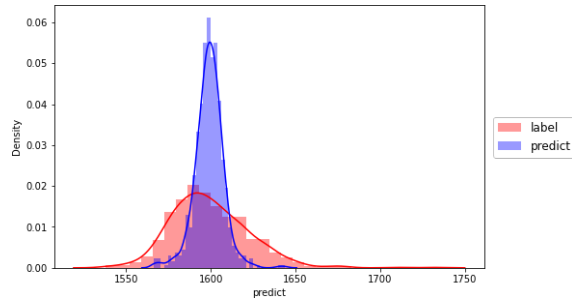
먼저 출강온도와 각 변수 간 산점도를 그렸습니다. 1~8로 출강온도를 아래의 기준으로 그룹화했습니다.

- min ~ 1550 : 1
- 1550 ~ 1575 : 2
- 1575 ~ 1600 : 3
- 1600 ~ 1625 : 4
- 1625 ~ 1650 : 5
- 1650 ~ 1675 : 6
- 1675 ~ 1700 : 7
- 1700 ~ max : 8

위 산점도를 보시면, 각 변수와 출강온도 간의 상관관계가 크게 없다는 것을 확인할 수 있었습니다

.

모델의 예측값과 실제 출강 온도의 차이를 prediction_diff라는 새로운 칼럼에 추가하여 분석해본 결과, 모델이 대부분의 데이터에 대해 1600도 근처로 예측하고 있음을 확인했습니다. 아래 이미지는 예측값(파란색)과 실제 출강온도(빨간색)의 분포 그래프입니다.



실제로 아래 결과 이미지를 보면, 대부분 predict 값이 1600 전후인 것을 확인하실 수 있습니다.

```
In [118]: result[result["label"]<1575].sort_values(by="label")
```

	HEAT	LIME	DOLOMITE	Power	m_b	diff_Th	m_r	diff_Td	a_in	label	predict	prediction_diff	label_group	
11	4F3851	2000.0	802.0	42.7	632	5	423	8	4	466.12054	1538	1595.87	57.87	1
138	4F3806	2001.0	504.0	44.0	636	3	432	8	4	325.277884	1541	1567.17	26.17	1
12	4F4023	2000.0	703.0	47.9	637	4	407	9	8	377.781350	1542	1589.54	57.54	1
48	4F4041	2002.0	605.0	47.0	626	3	414	7	5	422.851727	1548	1586.80	38.80	1
17	4F4763	1802.0	1002.0	48.4	626	4	401	7	8	421.649664	1550	1604.21	54.21	2
23	4F3805	1901.0	604.0	46.6	629	3	417	7	5	327.600111	1551	1598.30	47.30	2
70	4F4832	2004.0	504.0	43.6	632	6	409	9	8	387.086721	1552	1587.56	35.56	2
40	4F3731	1804.0	606.0	39.5	636	1	421	7	4	301.803203	1553	1593.88	40.88	2
36	4F4096	1802.0	603.0	46.7	626	3	426	8	4	346.404448	1554	1596.19	42.19	2
18	4F4078	2005.0	606.0	46.6	633	3	403	8	5	471.758827	1554	1596.46	42.46	2
53	4F4217	2003.0	604.0	46.2	626	2	408	9	4	355.819125	1558	1595.47	37.47	2
27	4F4445	2003.0	1004.0	50.9	627	5	404	11	6	352.803868	1558	1603.50	45.50	2
47	4F3847	1204.0	804.0	47.1	634	3	420	6	4	457.024082	1560	1599.16	39.16	2
422	4F3808	2002.0	504.0	46.4	636	2	432	7	5	389.458858	1560	1598.92	9.92	2
74	4F4244	2002.0	703.0	46.4	629	2	406	5	5	413.623757	1561	1599.65	34.65	2
159	4F4204	2004.0	1002.0	44.9	637	2	411	3	5	346.520467	1561	1595.93	24.93	2
34	4F4299	1901.0	1003.0	45.4	631	3	404	6	5	354.633001	1562	1605.21	43.21	2
13	4F4876	2004.0	603.0	47.0	635	1	404	5	3	313.787891	1562	1619.02	57.02	2
61	4F4445	2003.0	1004.0	50.9	627	2	402	8	5	411.045219	1563	1599.58	36.58	2
68	4F4109	1802.0	803.0	44.6	626	3	412	9	5	382.342478	1563	1598.70	35.70	2
198	4F4271	2002.0	502.0	47.6	639	3	414	8	5	419.889898	1564	1593.93	21.93	2
91	4F4086	2003.0	504.0	48.4	633	4	407	5	5	356.455856	1564	1600.38	44.38	2
81	4F3896	2002.0	604.0	47.3	636	4	425	7	4	373.017050	1565	1596.88	31.88	2

실제 출강온도 값이 1575 이하인 경우입니다. 1600과 비교적 차이가 있으므로, 예측값과의 차이인 prediction_diff 역시 높습니다.

```
result[(result["label"]>1595) & (result["label"]<1605)].sort_values(by="label")
```

	HEAT	LIME	DOLOMITE	Power	m_b	diff_Th	m_r	diff_Td	a_in	label	predict	prediction_diff	label_group	
559	4F3483	2002.0	1000.0	48.8	639	3	428	8	6	301.803205	1596	1599.09	3.09	3
684	4F4242	1801.0	803.0	48.1	626	3	387	7	4	282.580715	1598	1593.50	5.50	3
688	4F4202	2003.0	502.0	49.7	632	2	399	11	6	320.911787	1598	1598.04	0.04	3
655	4F4146	2001.0	302.0	48.0	629	3	409	9	6	303.803886	1598	1599.29	3.29	3
594	4F4658	2003.0	503.0	49.5	638	4	406	9	6	370.348406	1598	1597.76	1.76	3
604	4F4873	2004.0	503.0	49.5	633	5	402	9	5	376.517041	1598	1601.78	5.78	3
519	4F4781	2005.0	1005.0	48.7	632	2	403	6	5	338.910749	1598	1600.92	4.92	3
616	4F3932	2003.0	503.0	48.4	635	1	421	8	4	295.603080	1598	1598.45	0.45	3
485	4F3902	1702.0	701.0	46.7	630	2	425	5	3	389.890302	1598	1603.14	7.14	3
590	4F3809	2002.0	502.0	44.8	632	7	419	10	8	429.350678	1598	1590.52	5.48	3
315	4F3631	2003.0	506.0	48.5	632	3	423	10	5	372.424018	1598	1590.14	15.86	3
487	4F4886	2005.0	505.0	50.3	640	3	404	7	5	435.581912	1597	1603.30	6.30	3
371	4F4202	1902.0	604.0	48.8	633	2	404	7	7	275.483206	1597	1606.47	12.47	3
673	4F4808	2005.0	704.0	51.2	634	2	406	7	6	287.027428	1597	1594.48	2.54	3
522	4F3917	1804.0	504.0	48.6	635	2	423	8	4	340.881972	1597	1601.25	4.25	3
602	4F3953	2003.0	603.0	47.8	636	9	421	7	4	404.744255	1597	1593.57	3.43	3
305	4F3514	1901.0	502.0	51.5	640	3	431	9	5	381.616012	1597	1590.79	16.21	3
630	4F3914	2003.0	703.0	44.4	632	4	421	7	5	301.558709	1597	1592.14	4.86	3
625	4F3586	1903.0	705.0	50.6	637	2	424	6	4	375.389177	1597	1597.20	0.20	3
658	4F3710	1702.0	805.0	43.9	636	3	421	7	8	386.140209	1598	1594.86	3.14	3
585	4F4098	2003.0	704.0	50.7	635	4	404	5	5	375.690593	1598	1595.11	2.89	3
229	4F3617	2004.0	302.0	46.1	639	4	422	8	4	405.040771	1598	1619.36	20.36	3
467	4F4005	2002.0	703.0	48.0	632	2	417	6	6	389.880587	1598	1605.80	7.80	3

```
result[result["label"]>1650].sort_values(by="label")
```

	HEAT	LIME	DOLOMITE	Power	m_b	diff_Th	m_r	diff_Td	a_in	label	predict	prediction_diff	label_group	
7	4F4091	2005.0	504.0	45.8	624	5	411	10	5	381.912528	1602	1583.63	68.37	6
10	4F4896	2005.0	503.0	43.8	632	3	408	7	8	298.290233	1603	1592.22	65.78	6
14	4F4602	1903.0	502.0	46.0	636	4	407	9	7	448.020382	1606	1598.20	56.80	6
16	4F4898	2002.0	604.0	47.2	635	3	403	7	6	306.024448	1600	1604.54	55.46	6
4	4F3482	2002.0	987.0	49.4	639	2	439	7	5	314.889826	1606	1590.51	75.49	6
6	4F4180	2001.0	503.0	46.6	631	1	408	8	3	373.316386	1675	1605.39	69.61	7
8	4F4641	1803.0	501.0	45.9	632	3	407	8	7	286.137880	1675	1600.87	65.33	7
9	4F4603	2003.0	503.0	48.7	634	2	406	9	7	328.836175	1676	1600.17	75.83	7
2	4F3648	2202.0	603.0	48.7	634	4	422	5	5	383.232024	1677	1594.66	82.34	7
1	4F4672	2004.0	502.0	45.2	637	2	405	8	5	373.313586	1770	1601.52	108.48	8
0	4F4642	1803.0	503.0	50.3	638	2	406	9	6	321.126761	1771	1608.71	121.29	8

실제 출강온도 값이 1650 이상인 경우입니다. 높은 prediction_diff 인 것을 확인할 수 있습니다.