Preliminary Draft Results

Name: Qinyun (Peter) Yu, 501137007

Date: July 12th, 2022

Supervised by: Dr. Alan Fung

Note: LightGBM results are at the end of the document. To summarize:

Toronto City acts as the “house” and uses Toronto City Centre, Toronto Intl A and Buttonville A as exogenous features.

- I model results in LightGBM, SARIMAX, Prophet, and tried LSTM
- I look into the LightGBM model results having changed the “house” to King City North
- I also look into LightGBM model results having changed the “house” to Vancouver Harbour CS, while having West Vancouver Aut, Vancouver Intl A, and Point Atkinson as exogenous features and also using the RETScreen radiation data
- Feature weighing at the end of report

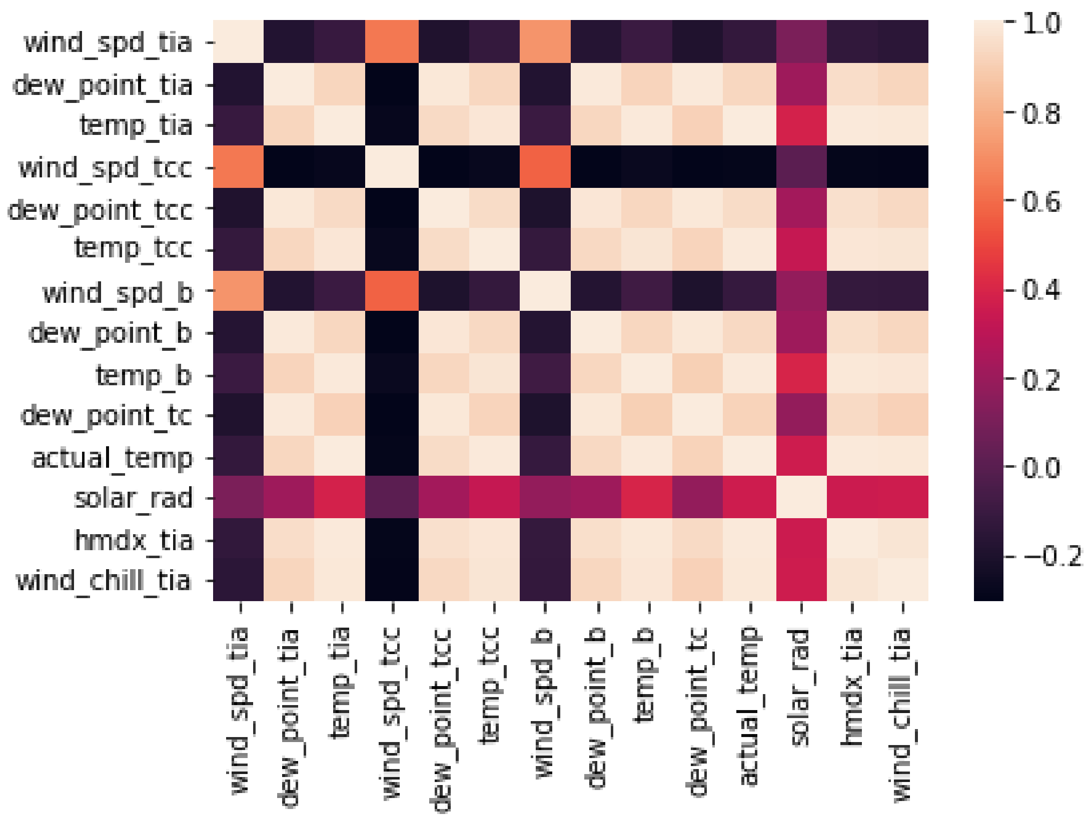
Initial model comparison of all data from 2016-01-01 to 2022-01-01

Model	RMSE	MAE	R2
SARIMAX	2.25	1.77	0.95
Prophet	0.70	0.50	1.00
LightGBM	0.81	0.58	1.00

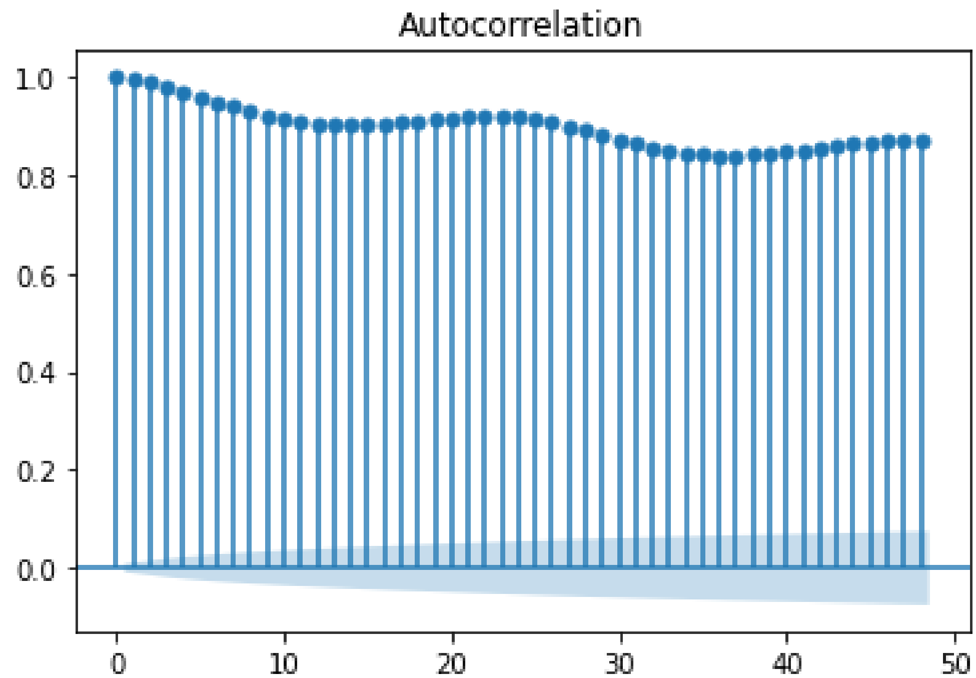
Comparison of Day vs. Night and Winter vs. Summer results across different stations (however it's nice to know) – sectioned off to three groups (the RMSEs, the MAEs, and the R2)

	Summer + Day	Summer + Night	Winter + Day	Winter + Night
RMSE Toronto	0.95	0.61	0.69	0.63
RMSE King City North	1.07	0.75	0.84	0.78
RMSE Vancouver	0.80	0.65	0.65	0.57
MAE Toronto	0.73	0.42	0.52	0.40
MAE King City North	0.80	0.51	0.62	0.50
MAE Vancouver	0.59	0.47	0.46	0.39
R2 Toronto	0.99	0.99	1.00	1.00
R2 King City North	0.99	0.99	0.99	0.99
R2 Vancouver	0.98	0.98	0.98	0.98

Model Feature correlation:



Autocorrelation Function plot of ground truth temperature (Toronto City)



LSTM model results:

- Could not be figured out due to issues in configuring initial model

SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous factors)

```
Performing stepwise search to minimize aic
 ARIMA(2,1,2)(0,0,0)[0] intercept   : AIC=90806.071, Time=116.16 sec
 ARIMA(0,1,0)(0,0,0)[0] intercept   : AIC=92293.407, Time=27.14 sec
 ARIMA(1,1,0)(0,0,0)[0] intercept   : AIC=92200.734, Time=93.43 sec
 ARIMA(0,1,1)(0,0,0)[0] intercept   : AIC=92174.490, Time=112.52 sec
 ARIMA(0,1,0)(0,0,0)[0]             : AIC=92291.408, Time=55.02 sec
 ARIMA(1,1,2)(0,0,0)[0] intercept   : AIC=90843.684, Time=116.84 sec
 ARIMA(2,1,1)(0,0,0)[0] intercept   : AIC=90867.609, Time=112.16 sec
 ARIMA(3,1,2)(0,0,0)[0] intercept   : AIC=90939.971, Time=122.77 sec
 ARIMA(2,1,3)(0,0,0)[0] intercept   : AIC=91022.040, Time=137.27 sec
 ARIMA(1,1,1)(0,0,0)[0] intercept   : AIC=90854.348, Time=108.31 sec
 ARIMA(1,1,3)(0,0,0)[0] intercept   : AIC=91978.095, Time=135.53 sec
 ARIMA(3,1,1)(0,0,0)[0] intercept   : AIC=92006.635, Time=118.10 sec
 ARIMA(3,1,3)(0,0,0)[0] intercept   : AIC=90653.259, Time=143.27 sec
 ARIMA(4,1,3)(0,0,0)[0] intercept   : AIC=90532.676, Time=151.47 sec
 ARIMA(4,1,2)(0,0,0)[0] intercept   : AIC=91094.341, Time=127.03 sec
 ARIMA(5,1,3)(0,0,0)[0] intercept   : AIC=90654.300, Time=129.47 sec
 ARIMA(4,1,4)(0,0,0)[0] intercept   : AIC=90543.519, Time=146.36 sec
 ARIMA(3,1,4)(0,0,0)[0] intercept   : AIC=90574.164, Time=135.36 sec
 ARIMA(5,1,2)(0,0,0)[0] intercept   : AIC=91138.197, Time=149.70 sec
 ARIMA(5,1,4)(0,0,0)[0] intercept   : AIC=90488.940, Time=137.60 sec
 ARIMA(5,1,5)(0,0,0)[0] intercept   : AIC=90508.638, Time=158.78 sec
 ARIMA(4,1,5)(0,0,0)[0] intercept   : AIC=90408.593, Time=158.38 sec
 ARIMA(3,1,5)(0,0,0)[0] intercept   : AIC=90435.542, Time=157.12 sec
 ARIMA(4,1,5)(0,0,0)[0]             : AIC=90396.382, Time=157.09 sec
 ARIMA(3,1,5)(0,0,0)[0]             : AIC=90414.366, Time=149.06 sec
 ARIMA(4,1,4)(0,0,0)[0]             : AIC=90535.914, Time=139.16 sec
 ARIMA(5,1,5)(0,0,0)[0]             : AIC=90499.355, Time=159.32 sec
 ARIMA(3,1,4)(0,0,0)[0]             : AIC=90555.014, Time=131.40 sec
 ARIMA(5,1,4)(0,0,0)[0]             : AIC=90455.948, Time=140.37 sec

Best model:  ARIMA(4,1,5)(0,0,0)[0]
Total fit time: 3726.266 seconds
```

Facebook Prophet

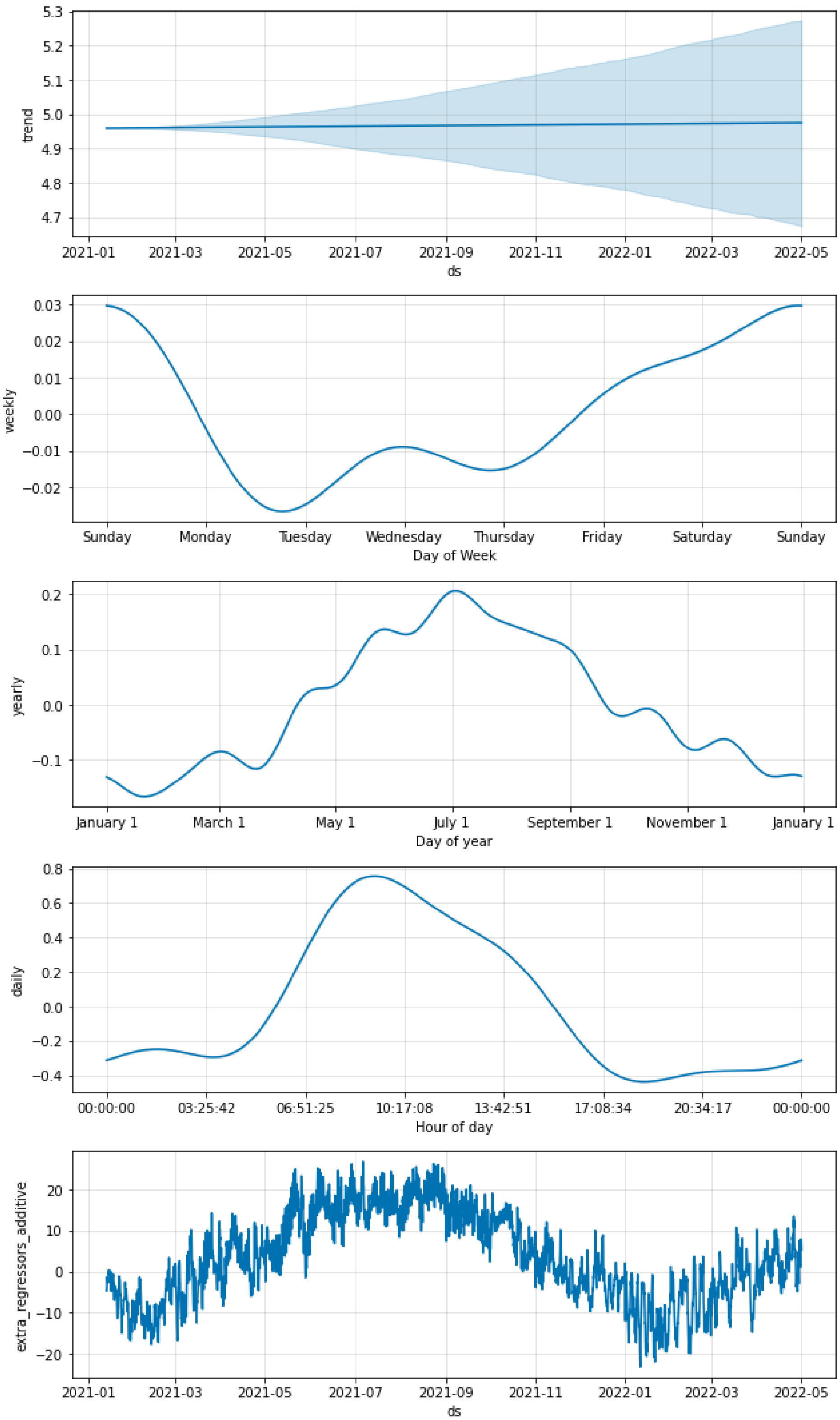
```
Initial log joint probability = -3592.13
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
      99        142404   0.000809671       30362.7           1           1      121   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
     199        142722   0.000253424       19471.8      0.4566           1      231   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
     299        142730   0.000142437       3676.35           1           1      339   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
     399        142732   6.38252e-05       1362.95           1           1      450   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
     499        142732   0.000248158       2108.08           1           1      554   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
     599
```

Out[26]: <prophet.forecaster.Prophet at 0x7f8c2163cf10>

```
   142735   8.60103e-05       2614.31      0.3699      0.3699      663   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
     699        142737   4.58495e-05        864.55           1           1      772   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
     799        142737   0.000158006       1183.53           1           1      886   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
     899        142737   8.79734e-06       1164.91      0.6006      0.6006      992   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
     999        142738   0.000167534       880.975      0.9966      0.9966     1100   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
    1099        142738   0.000449244       1610.29           1           1     1210   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
    1199        142738    0.00015105       1159.69           1           1     1315   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
    1299        142739    4.5134e-05        814.85           1           1     1429   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
    1399        142739   8.59375e-06       578.275      0.2308           1     1541   
    Iter      log prob        ||dx||      ||grad||       alpha      alpha0  # evals  Notes 
    1478        142739   1.37676e-05       345.524           1           1     1624   
Optimization terminated normally: 
  Convergence detected: relative gradient magnitude is below tolerance
```

Decomposition of Facebook Prophet model with respect to ground truth temperatures (Toronto City)

Out[29]:



LightGBM

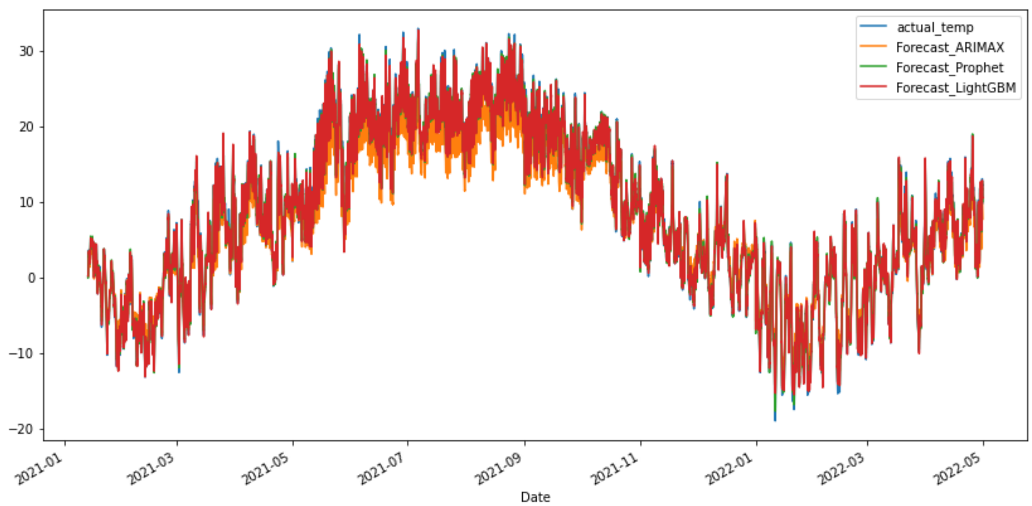
LightGBM

In [34]:
```
params = {"objective": "regression"}

dtrain = lgb.Dataset(train[exogenous_features], label=train.actual_temp.values)
dtest = lgb.Dataset(test[exogenous_features])

model_lgb = lgb.train(params, train_set=dtrain)

forecast = model_lgb.predict(test[exogenous_features])
test["Forecast_LightGBM"] = forecast
```

```
[LightGBM] [Warning] Auto-choosing row-wise multi-threading, the overhead of testing was 0.001100 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 2527
[LightGBM] [Info] Number of data points in the train set: 42288, number of used features: 30
[LightGBM] [Info] Start training from score 10.316399
```

In [35]:
```
test[["actual_temp", "Forecast_ARIMAX", "Forecast_Prophet", "Forecast_LightGBM"]].plot(figsize=(14, 7))
```

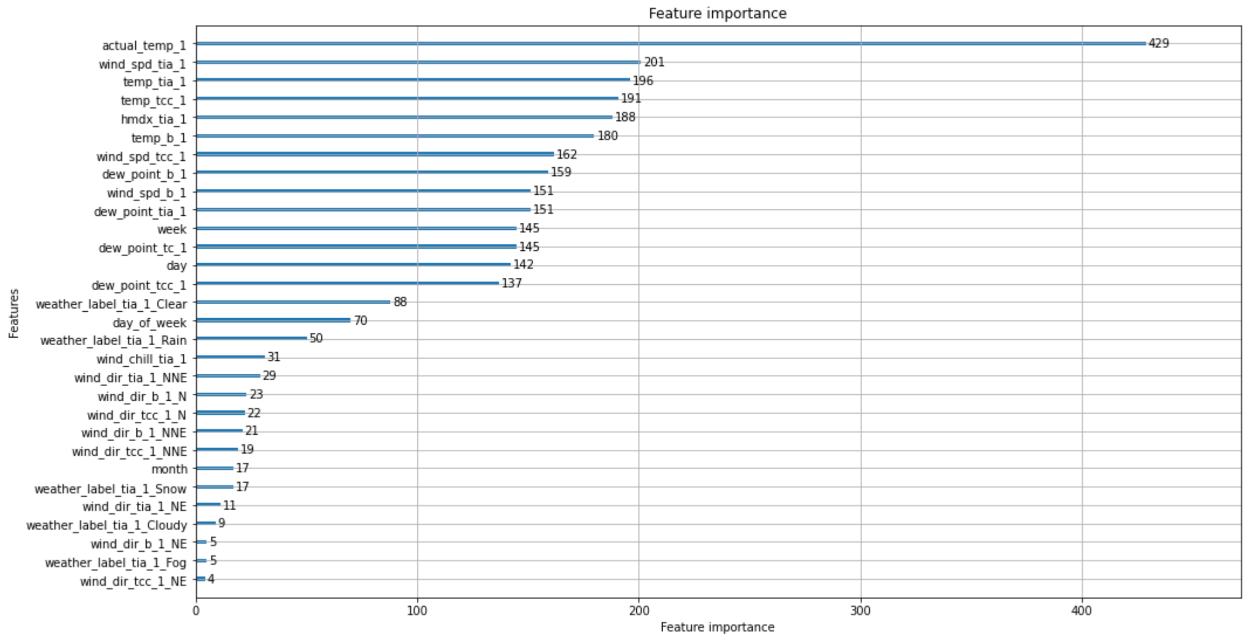
Out[35]:
```
<AxesSubplot:xlabel='Date'>
```



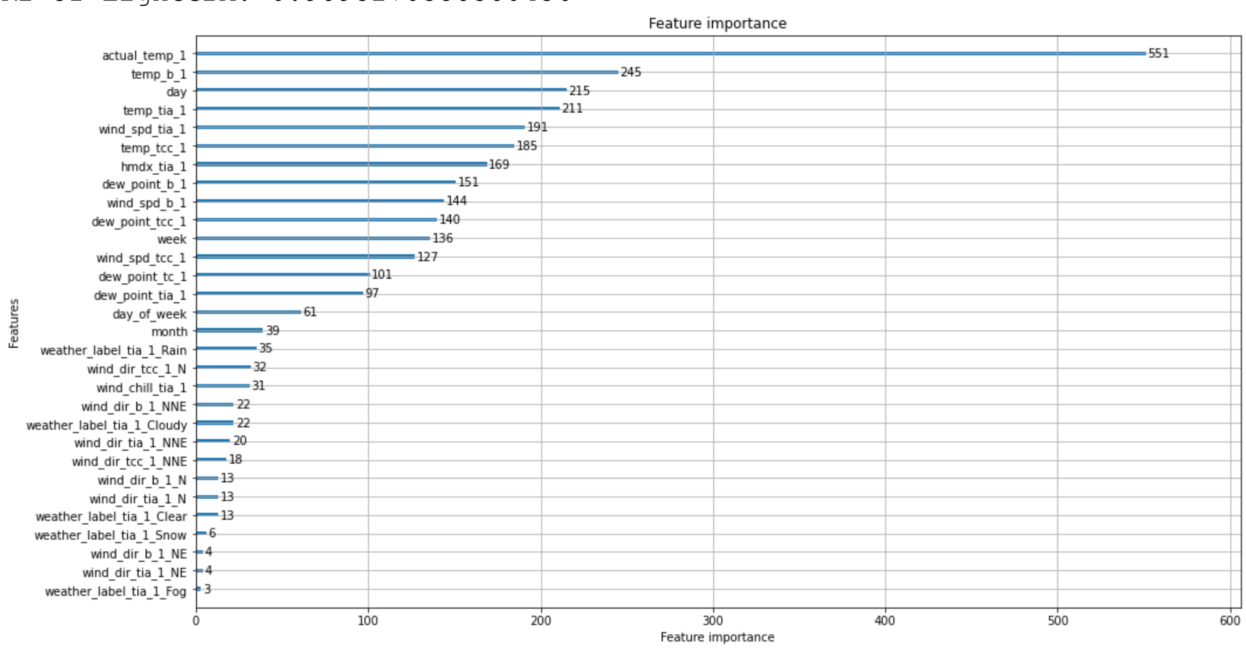
Comparison of results using Toronto City as center, and using Buttonville, Toronto International Airport, and Toronto City Centre as exogenous features

```
RMSE of Auto ARIMAX: 2.2465725197737565
RMSE of Prophet: 0.7022345031468229
RMSE of LightGBM: 0.8081883144061222

MAE of Auto ARIMAX: 1.7742792991644516
MAE of Prophet: 0.4985769353270404
MAE of LightGBM: 0.5753165809303864

R2 of Auto ARIMAX: 0.9549324808296925
R2 of Prophet: 0.995596610569661
R2 of LightGBM: 0.994167591946191
```

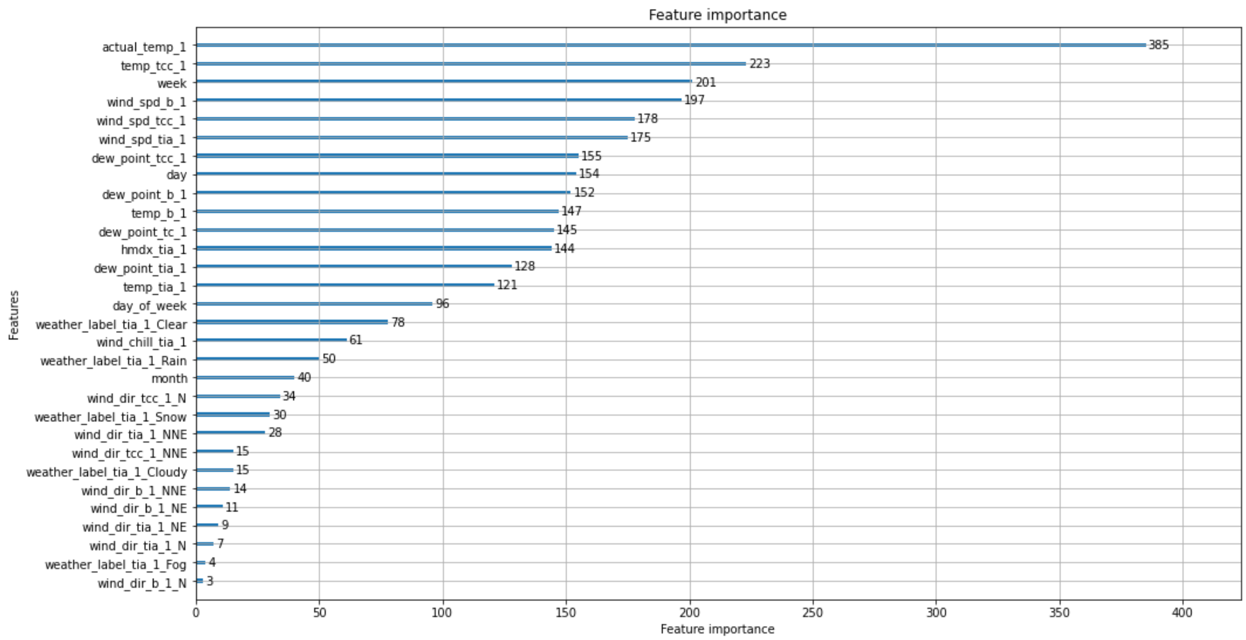
Results of training LightGBM models for Winter vs. Summer, and Day vs. Night – Toronto with Toronto City as center



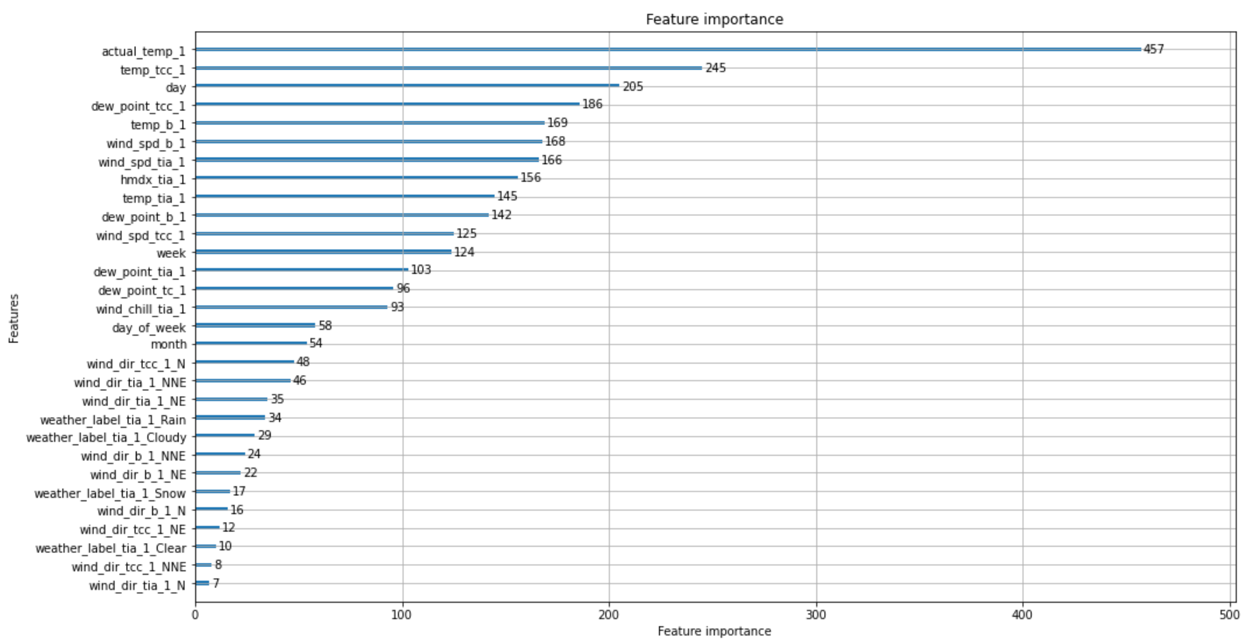

```
summer+day
RMSE of LightGBM: 0.9514951470499308
MAE of LightGBM: 0.7292005655663253
R2 of LightGBM: 0.9895270358380436
```




```
summer+night
RMSE of LightGBM: 0.606674750678484
MAE of LightGBM: 0.41617815053945373
R2 of LightGBM: 0.994825827592962
```

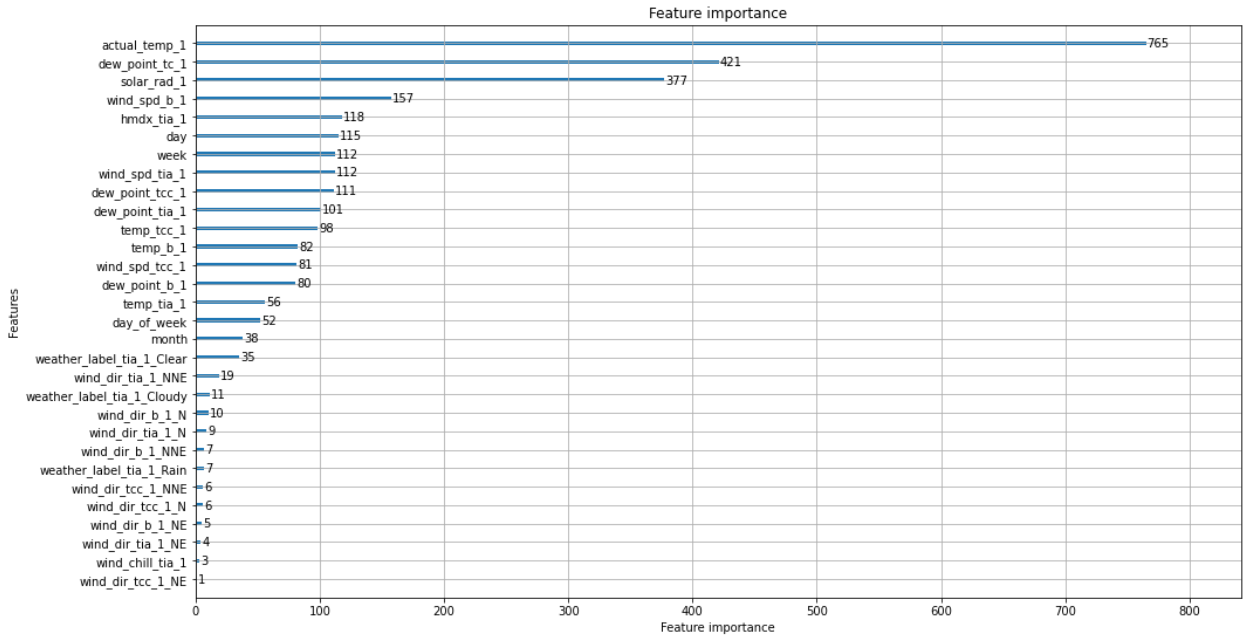


winter+day
RMSE of LightGBM: 0.6867777040983187
MAE of LightGBM: 0.5153370074675717
R2 of LightGBM: 0.9952860250738357

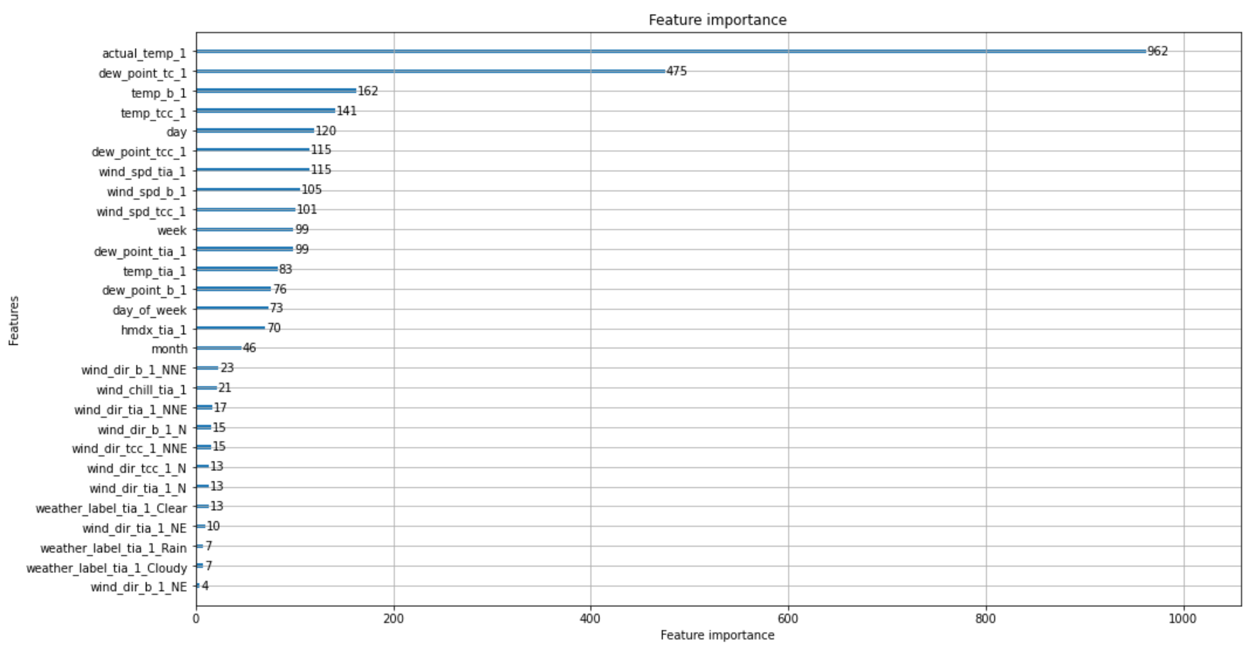


winter+night
RMSE of LightGBM: 0.6297058167375451
MAE of LightGBM: 0.4009749623637006
R2 of LightGBM: 0.9954531382083441

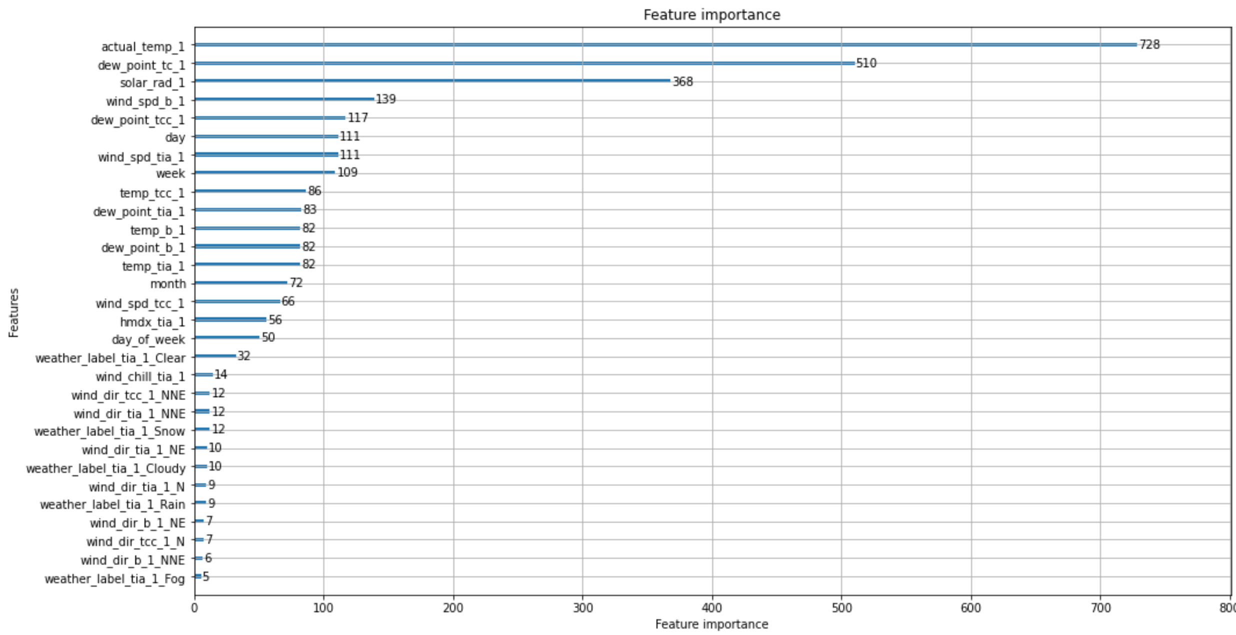
Results of training LightGBM models for Winter vs. Summer, and Day vs. Night with King City North as Center



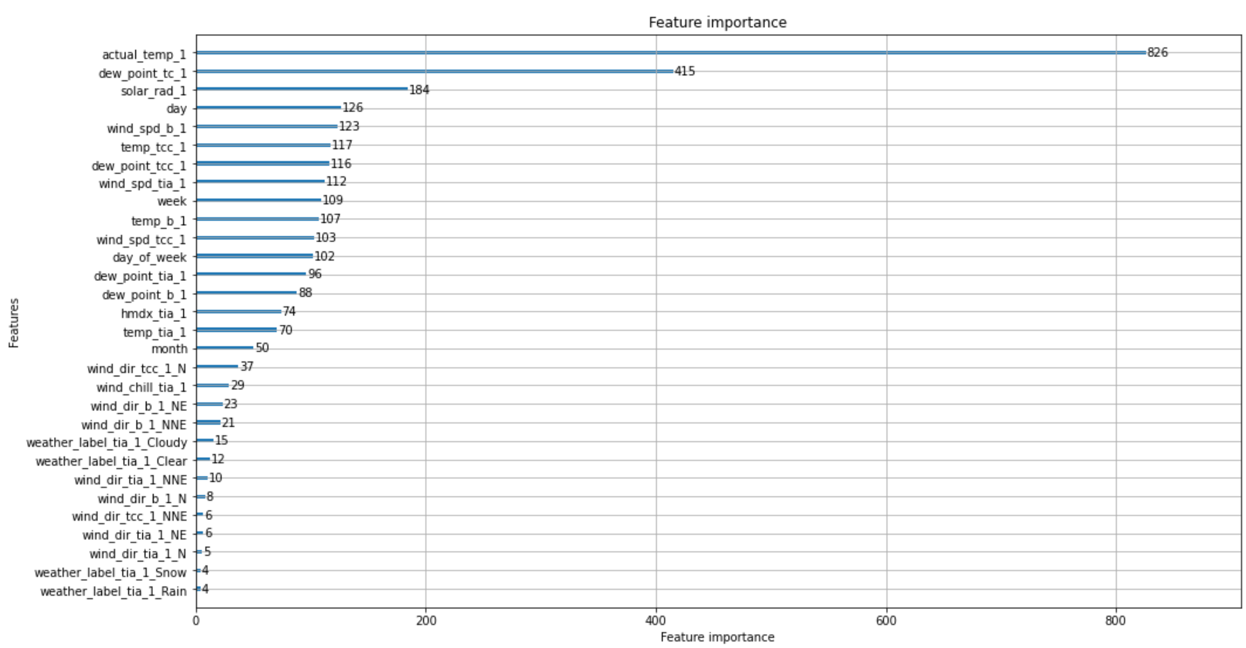
summer+day
RMSE of LightGBM: 1.0748284425722083
MAE of LightGBM: 0.7956344756208281
R2 of LightGBM: 0.988149204188182



summer+night
RMSE of LightGBM: 0.754676539626915
MAE of LightGBM: 0.5060721453669853
R2 of LightGBM: 0.9926396277284432

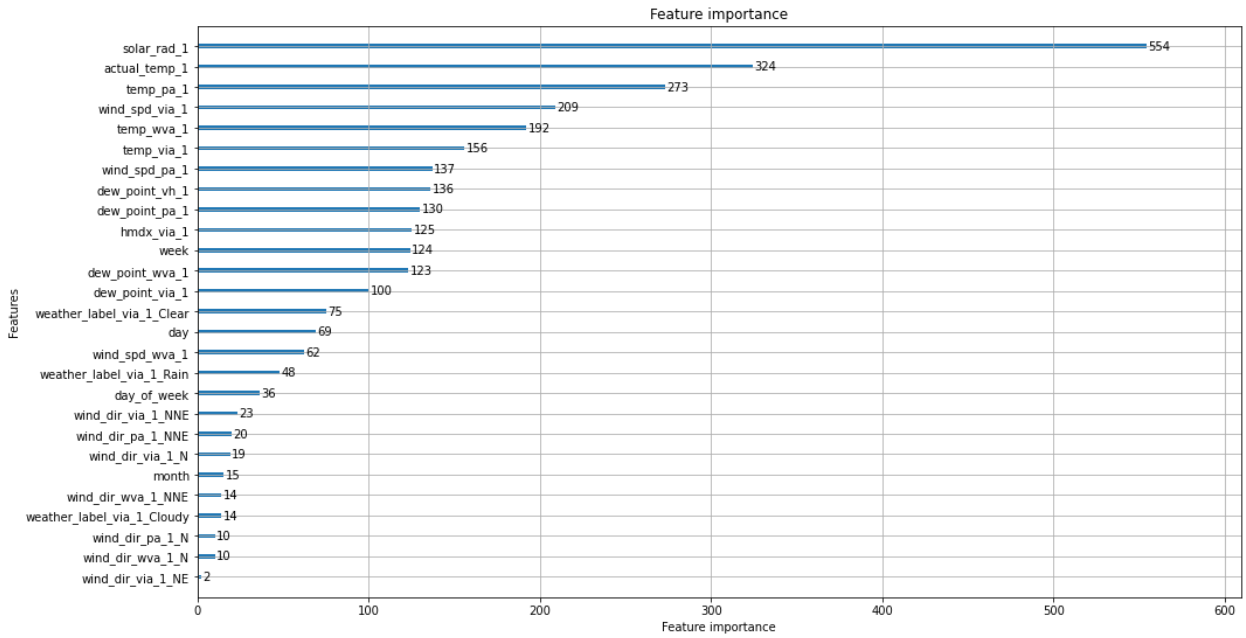


winter+day
RMSE of LightGBM: 0.8448025457852129
MAE of LightGBM: 0.6161362978373904
R2 of LightGBM: 0.9936708273499005

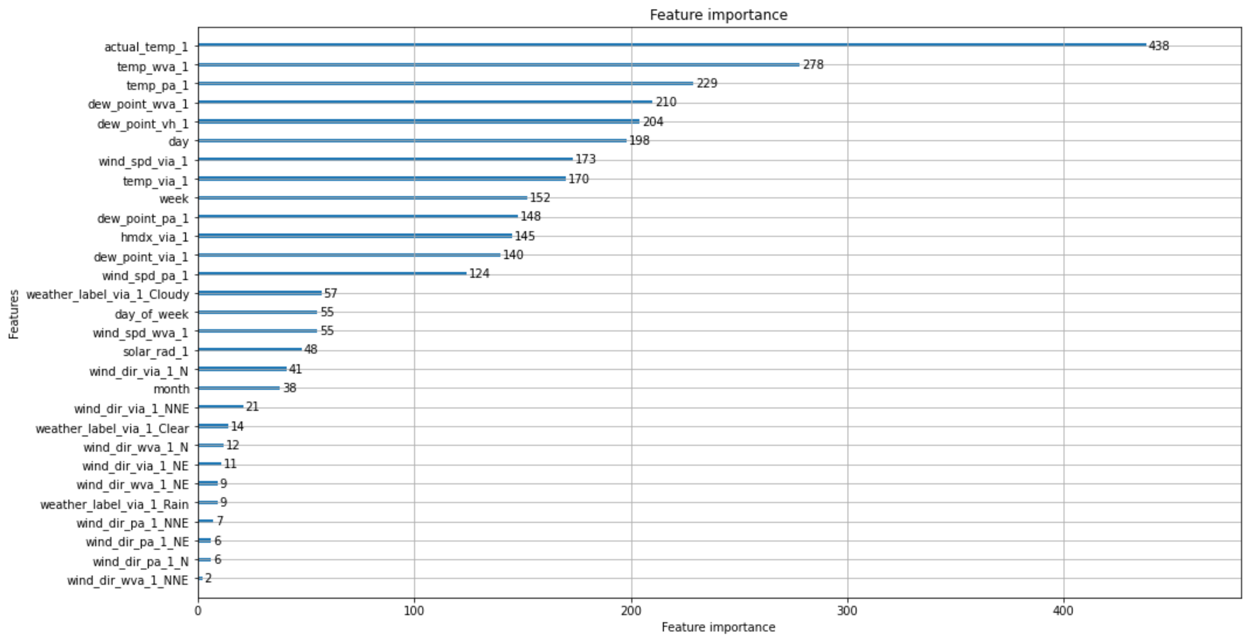


winter+night
RMSE of LightGBM: 0.7779397089068562
MAE of LightGBM: 0.4959331002135865
R2 of LightGBM: 0.9938311538898326

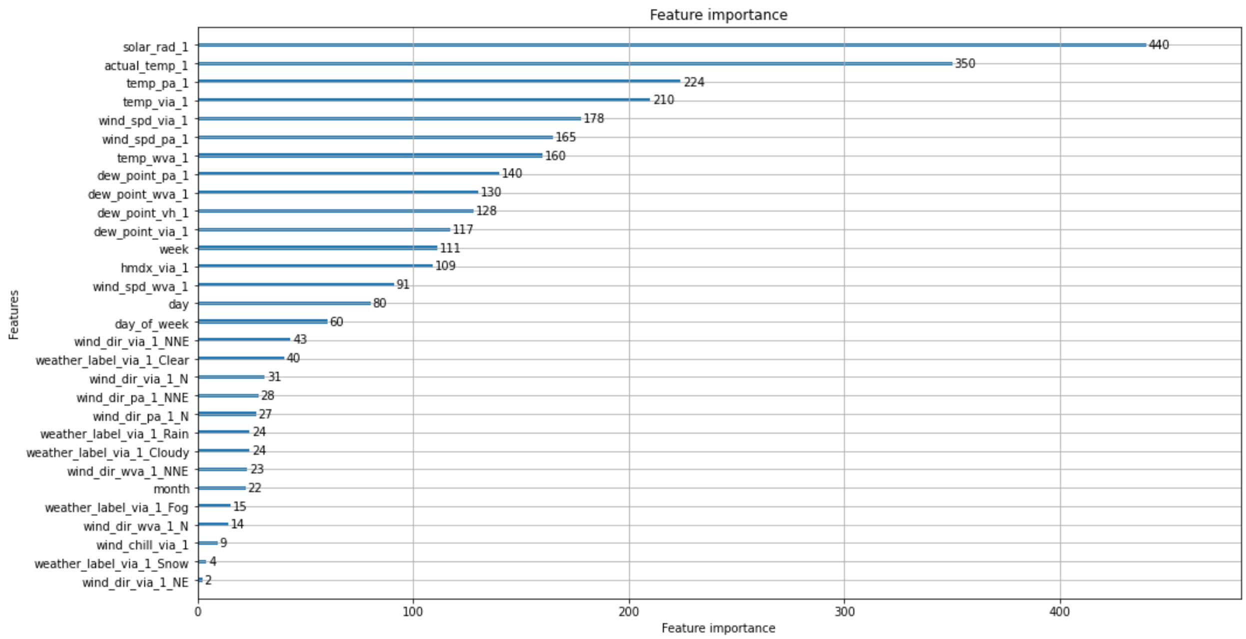
Results of training LightGBM models for Winter vs. Summer, and Day vs. Night in Vancouver with center as Vancouver Harbour CS, exogenous features of West Vancouver Aut, Vancouver Intl Airport, and Point Atkinson



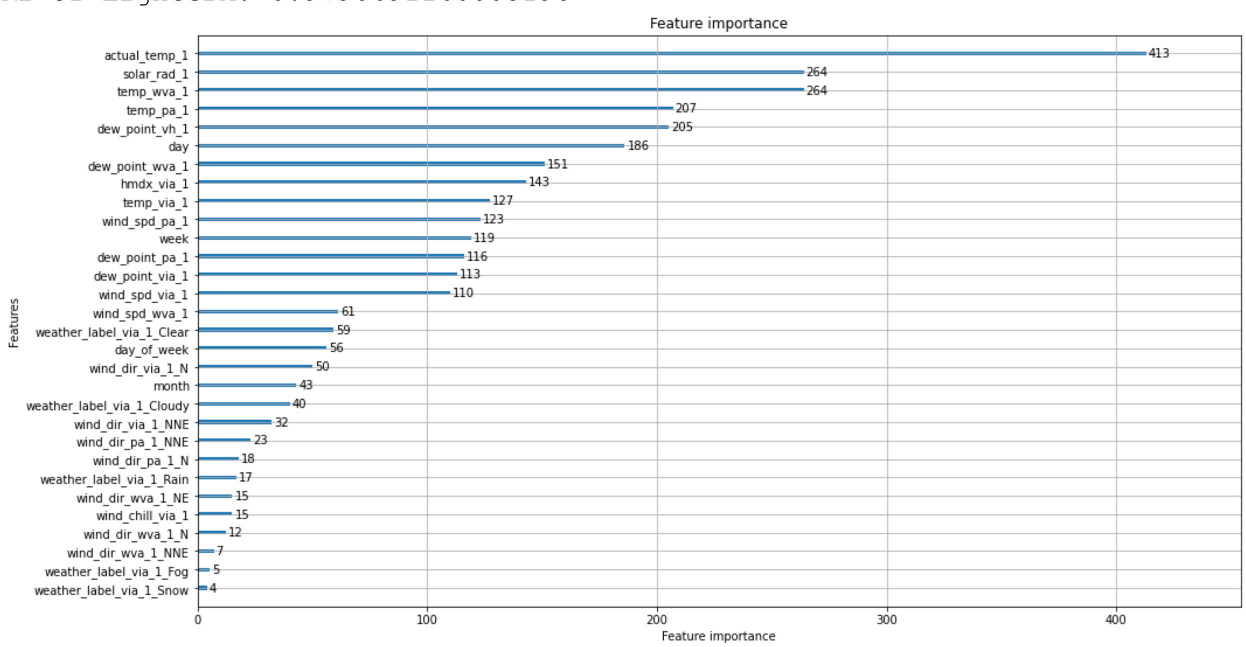
summer+day
RMSE of LightGBM: 0.8014885854789044
MAE of LightGBM: 0.5949593485724824
R2 of LightGBM: 0.9816951047783669



summer+night
RMSE of LightGBM: 0.6469117191366816
MAE of LightGBM: 0.4675209745026675
R2 of LightGBM: 0.9838411211703266



winter+day
RMSE of LightGBM: 0.6540143516556265
MAE of LightGBM: 0.4598819412611143
R2 of LightGBM: 0.9790691160005198



winter+night
RMSE of LightGBM: 0.5735699336463019
MAE of LightGBM: 0.39290662098983825
R2 of LightGBM: 0.980163293814143