

Univariate Time Series Analysis | Weather Prediction | +

localhost:8888/notebooks/Univariate%20Time%20Series%20Analysis%20for%20Weather%20Prediction%20using%20the%20Prophet%20Library.ipynb

jupyter Univariate Time Series Analysis for Weather Prediction using the Prophet Li... Last Checkpoint: 2 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Kernel

Reading the dataset

```
In [2]: weather_data=pd.read_csv("weather_forecast.csv")
```

```
In [3]: weather_data.head()
```

```
Out[3]:
```

	datetime_utc	_conds	_dewptm	_fog	_hail	_heatindexm	_hum	_precipm	_pressurem	_rain	_snow	_tempm	_thunder	_tornado	_vism	_wdird	_wdir
0	19961101-11:00	Smoke	9.0	0	0	NaN	27.0	NaN	1010.0	0	0	30.0	0	0	5.0	280.0	We
1	19961101-12:00	Smoke	10.0	0	0	NaN	32.0	NaN	-9999.0	0	0	28.0	0	0	NaN	0.0	Nort
2	19961101-13:00	Smoke	11.0	0	0	NaN	44.0	NaN	-9999.0	0	0	24.0	0	0	NaN	0.0	Nort
3	19961101-14:00	Smoke	10.0	0	0	NaN	41.0	NaN	1010.0	0	0	24.0	0	0	2.0	0.0	Nort
4	19961101-15:00	Smoke	11.0	0	0	NaN	47.0	NaN	1011.0	0	0	23.0	0	0	1.2	0.0	Nort

```
In [4]: weather_data.tail()
```

```
Out[4]:
```

	datetime_utc	_conds	_dewptm	_fog	_hail	_heatindexm	_hum	_precipm	_pressurem	_rain	_snow	_tempm	_thunder	_tornado	_vism	_wdird	_wdir
100985	20170424-06:00	Haze	17.0	0	0	NaN	25.0	NaN	1005.0	0	0	34.0	0	0	4.0	320.0	
100986	20170424-09:00	Haze	14.0	0	0	NaN	16.0	NaN	1003.0	0	0	38.0	0	0	4.0	320.0	
100987	20170424-12:00	Haze	12.0	0	0	NaN	14.0	NaN	1002.0	0	0	36.0	0	0	4.0	270.0	
100988	20170424-15:00	Haze	15.0	0	0	NaN	27.0	NaN	1004.0	0	0	32.0	0	0	2.0	320.0	

Windows taskbar: Type here to search, 2033, 25-02-2021

Univariate Time Series Analysis | Weather Prediction | +

localhost:8888/notebooks/Univariate%20Time%20Series%20Analysis%20for%20Weather%20Prediction%20using%20the%20Prophet%20Library.ipynb

jupyter Univariate Time Series Analysis for Weather Prediction using the Prophet Li... Last Checkpoint: 2 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Kernel

STEP 1-Importing the necessary libraries

```
In [1]: from fbprophet import Prophet
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
```

In C:\Users\Ifrah Fatima\Anaconda3\lib\site-packages\matplotlib\mpl-data\stylelib_classic_test.mplstyle: The text.latex.preview rcparam was deprecated in Matplotlib 3.3 and will be removed two minor releases later.
In C:\Users\Ifrah Fatima\Anaconda3\lib\site-packages\matplotlib\mpl-data\stylelib_classic_test.mplstyle: The matplotlib.figure.facecolor rcparam was deprecated in Matplotlib 3.3 and will be removed two minor releases later.
In C:\Users\Ifrah Fatima\Anaconda3\lib\site-packages\matplotlib\mpl-data\stylelib_classic_test.mplstyle: Support for setting the 'mathtext.fallback_to_cm' rcparam is deprecated since 3.3 and will be removed two minor releases later; use 'mathtext.fallback_to_cm' instead.
In C:\Users\Ifrah Fatima\Anaconda3\lib\site-packages\matplotlib\mpl-data\stylelib_classic_test.mplstyle: The validate_bool_maybe_none function was deprecated in Matplotlib 3.3 and will be removed two minor releases later.
In C:\Users\Ifrah Fatima\Anaconda3\lib\site-packages\matplotlib\mpl-data\stylelib_classic_test.mplstyle: The savefig.jpeg_quality rcparam was deprecated in Matplotlib 3.3 and will be removed two minor releases later.
In C:\Users\Ifrah Fatima\Anaconda3\lib\site-packages\matplotlib\mpl-data\stylelib_classic_test.mplstyle: The keymap.all_axes rcparam was deprecated in Matplotlib 3.3 and will be removed two minor releases later.
In C:\Users\Ifrah Fatima\Anaconda3\lib\site-packages\matplotlib\mpl-data\stylelib_classic_test.mplstyle: The animation.avconv_path rcparam was deprecated in Matplotlib 3.3 and will be removed two minor releases later.
In C:\Users\Ifrah Fatima\Anaconda3\lib\site-packages\matplotlib\mpl-data\stylelib_classic_test.mplstyle: The animation.avconv_args rcparam was deprecated in Matplotlib 3.3 and will be removed two minor releases later.

Reading the dataset

```
In [2]: weather_data=pd.read_csv("weather_forecast.csv")
```

```
In [3]: weather_data.head()
```

Windows taskbar: Type here to search, 2032, 25-02-2021

In [5]: weather_data.dtypes

```
Out[5]: datetime_utc    object
        _conds          object
        _dewptm         float64
        _fog            int64
        _hail           int64
        _heatindexm     float64
        _hum            float64
        _precipm        float64
        _pressurem      float64
        _rain           int64
        _snow           int64
        _tempm         float64
        _thunder        int64
        _tornado        int64
        _vism           float64
        _wdird          float64
        _wdire          object
        _wgustm         float64
        _windchillm     float64
        _wspdm          float64
        dtype: object
```

In [6]: weather_data['datetime_utc']=pd.to_datetime(weather_data['datetime_utc'])

In [7]: weather_data.set_index('datetime_utc',inplace=True)

In [8]: weather_data

```
Out[8]: _conds _dewptm _fog _hail _heatindexm _hum _precipm _pressurem _rain _snow _tempm _thunder _tornado _vism _wdird _wdire
datetime_utc
```

In [8]: weather_data

```
Out[8]: _conds _dewptm _fog _hail _heatindexm _hum _precipm _pressurem _rain _snow _tempm _thunder _tornado _vism _wdird _wdire
datetime_utc
```

1995-11-01 11:00:00	Smoke	9.0	0	0	NaN	27.0	NaN	1010.0	0	0	30.0	0	0	5.0	280.0	West
1995-11-01 12:00:00	Smoke	10.0	0	0	NaN	32.0	NaN	-9999.0	0	0	28.0	0	0	NaN	0.0	North
1995-11-01 13:00:00	Smoke	11.0	0	0	NaN	44.0	NaN	-9999.0	0	0	24.0	0	0	NaN	0.0	North
1995-11-01 14:00:00	Smoke	10.0	0	0	NaN	41.0	NaN	1010.0	0	0	24.0	0	0	2.0	0.0	North
1995-11-01 16:00:00	Smoke	11.0	0	0	NaN	47.0	NaN	1011.0	0	0	23.0	0	0	1.2	0.0	North
...
2017-04-24 06:00:00	Haze	17.0	0	0	NaN	25.0	NaN	1005.0	0	0	34.0	0	0	4.0	320.0	NW
2017-04-24 09:00:00	Haze	14.0	0	0	NaN	16.0	NaN	1003.0	0	0	38.0	0	0	4.0	320.0	NW
2017-04-24 12:00:00	Haze	12.0	0	0	NaN	14.0	NaN	1002.0	0	0	36.0	0	0	4.0	270.0	West
2017-04-24 15:00:00	Haze	15.0	0	0	NaN	27.0	NaN	1004.0	0	0	32.0	0	0	2.0	320.0	NW
2017-04-24 18:00:00	Haze	15.0	0	0	NaN	30.0	NaN	1005.0	0	0	30.0	0	0	2.0	320.0	NW

100990 rows x 19 columns

In [9]: weather_data=weather_data.resample('D').mean()

In [10]: weather_data

Univariate Time Series Analysis for Weather Prediction using the Prophet Library

```
In [9]: weather_data=weather_data.resample('D').mean()

In [10]: weather_data

Out[10]:
```

	_dewptm	_fog	_hall	_heatindexm	_hum	_precipm	_pressurem	_rain	_snow	_tempm	_thunder	_tornado	_vism	_wdird	_windspeedm
datetime_utc															
1996-11-01	11.666667	0.0	0.0	NaN	52.916667	NaN	-2659.666667	0.0	0.0	22.333333	0.0	0.0	2.250000	23.333333	
1996-11-02	10.458333	0.0	0.0	NaN	48.625000	NaN	1009.833333	0.0	0.0	22.916667	0.0	0.0	3.476190	106.666667	
1996-11-03	12.041667	0.0	0.0	26.80	55.958333	NaN	1010.500000	0.0	0.0	21.791667	0.0	0.0	2.286364	106.666667	
1996-11-04	10.222222	0.0	0.0	27.50	48.055556	NaN	1011.333333	0.0	0.0	22.722222	0.0	0.0	2.326667	55.555556	
1996-11-05	8.200000	0.0	0.0	NaN	29.400000	NaN	1011.800000	0.0	0.0	27.800000	0.0	0.0	3.900000	208.000000	
...
2017-04-20	16.750000	0.0	0.0	NaN	27.500000	NaN	998.625000	0.0	0.0	34.500000	0.0	0.0	2.500000	216.666667	
2017-04-21	21.500000	0.0	0.0	NaN	39.375000	NaN	999.875000	0.0	0.0	34.250000	0.0	0.0	2.000000	92.500000	
2017-04-22	20.400000	0.0	0.0	38.05	40.900000	NaN	1001.600000	0.0	0.0	32.900000	0.2	0.0	1.950000	113.750000	
2017-04-23	15.125000	0.0	0.0	NaN	27.500000	NaN	1002.125000	0.0	0.0	32.875000	0.0	0.0	2.500000	310.000000	
2017-04-24	14.857143	0.0	0.0	NaN	27.142857	NaN	1004.142857	0.0	0.0	32.000000	0.0	0.0	3.142857	262.857143	

7480 rows x 17 columns

```
In [11]: weather_data=weather_data[['_tempm']]

In [12]: type(weather_data[['_tempm']])

Out[12]: pandas.core.frame.DataFrame

In [13]: weather_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 7480 entries, 1996-11-01 to 2017-04-24
Freq: D
Data columns (total 1 columns):
 _tempm      7336 non-null float64
dtypes: float64(1)
memory usage: 116.9 KB
```

Finding missing values

```
In [14]: weather_data.isnull().any()

Out[14]:
_tempm      True
dtype: bool

In [15]: weather_data['_tempm'].fillna(weather_data['_tempm'].mean(), inplace=True)

C:\Users\Ifrah Fatima\Anaconda3\lib\site-packages\pandas\core\generic.py:6287: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
self._update_inplace(new_data)

In [16]: weather_data.head()

Out[16]:
```

	_tempm
datetime_utc	
1996-11-01	22.333333
1996-11-02	22.916667

Univariate Time Series Analysis | xWeather Prediction x+

localhost:8888/notebooks/Univariate%20Time%20Series%20Analysis%20for%20Weather%20Prediction%20using%20the%20Prophet%20Library.ipynb

jupyter Univariate Time Series Analysis for Weather Prediction using the Prophet Li... Last Checkpoint: 2 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Kernel O

Run Code

Out[16]:

_tempm

datetime_utc

1996-11-01	22.333333
1996-11-02	22.916667
1996-11-03	21.791667
1996-11-04	22.722222
1996-11-05	27.800000

In [17]: weather_data.reset_index(inplace=True)

In [18]: weather_data.rename(columns={'datetime_utc':'ds', '_tempm':'y'}, inplace=True)

C:\Users\Ifrah Fatima\Anaconda3\lib\site-packages\pandas\core\frame.py:4223: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
return super().rename(**kwargs)

In [19]: weather_data.head()

Out[19]:

	ds	y
0	1996-11-01	22.333333
1	1996-11-02	22.916667
2	1996-11-03	21.791667
3	1996-11-04	22.722222
4	1996-11-05	27.800000

Type here to search

Univariate Time Series Analysis | xWeather Prediction x+

localhost:8888/notebooks/Univariate%20Time%20Series%20Analysis%20for%20Weather%20Prediction%20using%20the%20Prophet%20Library.ipynb

jupyter Univariate Time Series Analysis for Weather Prediction using the Prophet Li... Last Checkpoint: 2 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Kernel O

Run Code

Visualising the data

In [20]: plt.figure(figsize=(12,8))
plt.plot(weather_data.set_index(["ds"]))

Out[20]: [matplotlib.lines.Line2D at 0x2379d8ea9c8]

Type here to search

Model Fitting

```
In [21]: model=Prophet()
model.fit(weather_data)

INFO:numexpr.utils:NumExpr defaulting to 8 threads.
INFO:fbprophet:Disabling daily seasonality. Run prophet with daily_seasonality=True to override this.
C:\Users\Ibrahim\Anaconda3\lib\site-packages\pystan\misc.py:399: FutureWarning: Conversion of the second argument of issubdtype from 'float' to 'np.floating' is deprecated. In future, it will be treated as 'np.float64 == np.dtype(float).type'.
  elif np.issubdtype(np.asarray(v).dtype, float):

Out[21]: <fbprophet.forecaster.Prophet at 0x23710107c48>
```

Making Future Predictions

```
In [22]: future=model.make_future_dataframe(periods=365, freq="D")
future.tail()

Out[22]:
      ds
7840 2018-04-20
7841 2018-04-21
7842 2018-04-22
7843 2018-04-23
7844 2018-04-24

In [23]: future.head()

Out[23]:
      ds
0 1996-11-01
```

```
Out[23]:
      ds
0 1996-11-01
1 1996-11-02
2 1996-11-03
3 1996-11-04
4 1996-11-05
```

Obtaining the Forecasts

```
In [24]: forecast=model.predict(future)

In [25]: forecast.head()

Out[25]:
      ds    trend  yhat_lower  yhat_upper  trend_lower  trend_upper  additive_terms  additive_terms_lower  additive_terms_upper  weekly  weekly_lower  w
0 1996-11-01  23.521637  17.897922  24.075713  23.521637  23.521637  -2.256154  -2.256154  -2.256154  -0.062900  -0.062900
1 1996-11-02  23.523379  18.168187  24.286779  23.523379  23.523379  -2.360807  -2.360807  -2.360807  0.003322  0.003322
2 1996-11-03  23.525122  18.000908  24.290962  23.525122  23.525122  -2.547109  -2.547109  -2.547109  -0.011382  -0.011382
3 1996-11-04  23.526864  17.658139  23.849415  23.526864  23.526864  -2.715783  -2.715783  -2.715783  -0.007156  -0.007156
4 1996-11-05  23.528606  17.441369  24.048196  23.528606  23.528606  -2.905169  -2.905169  -2.905169  -0.021823  -0.021823

In [26]: forecast[['ds','yhat','yhat_lower','yhat_upper']].tail()

Out[26]:
```

```
In [26]: forecast[['ds', 'yhat', 'yhat_lower', 'yhat_upper']].tail()
Out[26]:
```

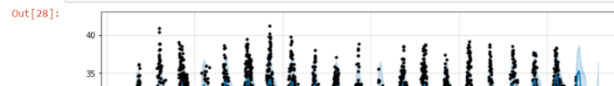
	ds	yhat	yhat_lower	yhat_upper
7840	2019-04-20	32.513276	29.455970	35.366515
7841	2019-04-21	32.724116	29.682652	35.860843
7842	2019-04-22	32.853331	29.701953	35.737704
7843	2019-04-23	33.000876	29.807881	35.802719
7844	2019-04-24	33.128920	29.895047	36.416263

```
In [27]: forecast[['ds', 'yhat', 'yhat_lower', 'yhat_upper']].head()
Out[27]:
```

	ds	yhat	yhat_lower	yhat_upper
0	1996-11-01	21.285483	17.897922	24.075713
1	1996-11-02	21.162573	18.168187	24.286779
2	1996-11-03	20.978013	18.000908	24.290962
3	1996-11-04	20.811090	17.658139	23.849415
4	1996-11-05	20.623437	17.441369	24.048166

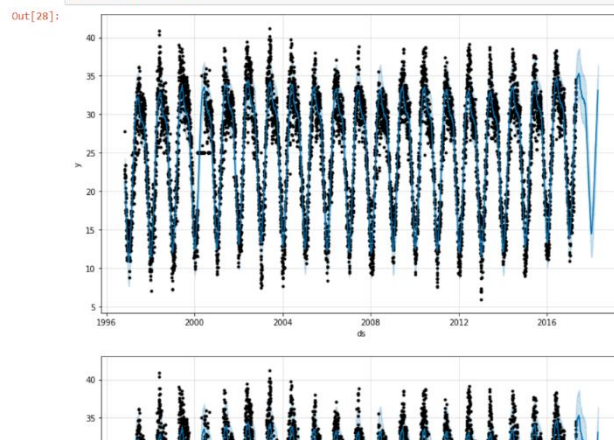
Plotting the Forecasts

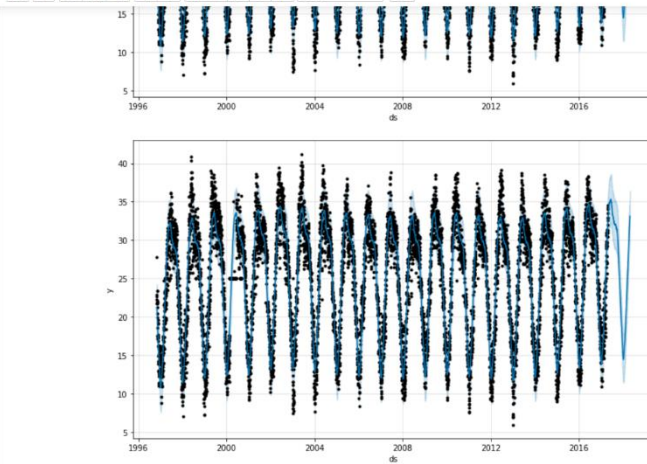
```
In [28]: model.plot(forecast)
```



Plotting the Forecasts

```
In [28]: model.plot(forecast)
```

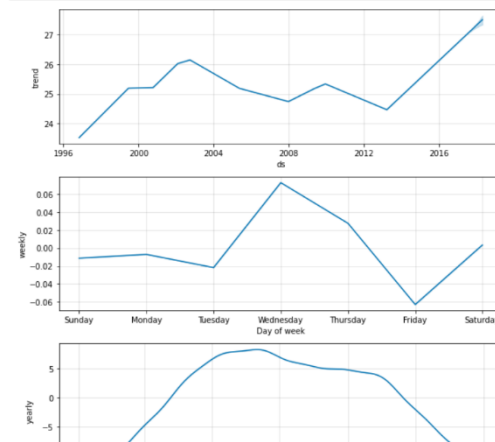




Plotting the Forecast Components

Plotting the Forecast Components

In [29]: `fig = model.plot_components(forecast)`



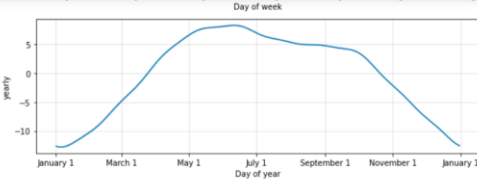
Univariate Time Series Analysis | xWeather Prediction | x+

localhost:8888/notebooks/Univariate%20Time%20Series%20Analysis%20for%20Weather%20Prediction%20using%20the%20Prophet%20Library.ipynb

jupyterUnivariate Time Series Analysis for Weather Prediction using the Prophet Li...Last Checkpoint: 2 hours ago (autosaved)Logout

FileEditViewInsertCellKernelWidgetsHelp

RunCode



Cross Validation

```
In [30]: from fbprophet.diagnostics import cross_validation
df_cv = cross_validation(model, initial='730 days', period='180 days', horizon = '365 days')
df_cv.head()
```

INFO:fbprophet:Making 36 forecasts with cutoffs between 1999-01-24 00:00:00 and 2016-04-24 00:00:00
C:\Users\Ifrat\Anaconda3\lib\site-packages\pystan\misc.py:399: FutureWarning: Conversion of the second argument of issubdtype from 'float' to 'np.floating' is deprecated. In future, it will be treated as 'np.float64 == np.dtype(float).type'.
elif np.issubdtype(np.asarray(v).dtype, float):

```
Out[30]:
```

	ds	yhat	yhat_lower	yhat_upper	y	cutoff
0	1999-01-25	13.719397	10.537031	16.759790	13.20	1999-01-24
1	1999-01-26	13.860526	10.804572	16.921983	10.00	1999-01-24
2	1999-01-27	14.041650	10.802333	17.079861	17.75	1999-01-24
3	1999-01-28	14.115146	10.866196	16.811213	11.75	1999-01-24
4	1999-01-29	13.887185	10.970404	17.121100	11.20	1999-01-24

Univariate Time Series Analysis | xWeather Prediction | x+

localhost:8888/notebooks/Univariate%20Time%20Series%20Analysis%20for%20Weather%20Prediction%20using%20the%20Prophet%20Library.ipynb

jupyterUnivariate Time Series Analysis for Weather Prediction using the Prophet Li...Last Checkpoint: 2 hours ago (autosaved)Logout

FileEditViewInsertCellKernelWidgetsHelp

RunCode

Cross Validation

```
In [30]: from fbprophet.diagnostics import cross_validation
df_cv = cross_validation(model, initial='730 days', period='180 days', horizon = '365 days')
df_cv.head()
```

INFO:fbprophet:Making 36 forecasts with cutoffs between 1999-01-24 00:00:00 and 2016-04-24 00:00:00
C:\Users\Ifrat\Anaconda3\lib\site-packages\pystan\misc.py:399: FutureWarning: Conversion of the second argument of issubdtype from 'float' to 'np.floating' is deprecated. In future, it will be treated as 'np.float64 == np.dtype(float).type'.
elif np.issubdtype(np.asarray(v).dtype, float):

```
Out[30]:
```

	ds	yhat	yhat_lower	yhat_upper	y	cutoff
0	1999-01-25	13.719397	10.537031	16.759790	13.20	1999-01-24
1	1999-01-26	13.860526	10.804572	16.921983	10.00	1999-01-24
2	1999-01-27	14.041650	10.802333	17.079861	17.75	1999-01-24
3	1999-01-28	14.115146	10.866196	16.811213	11.75	1999-01-24
4	1999-01-29	13.887185	10.970404	17.121100	11.20	1999-01-24

Obtaining the Performance Metrics

```
In [31]: from fbprophet.diagnostics import performance_metrics
df_p = performance_metrics(df_cv)
df_p.head()
```

```
Out[31]:
```

	horizon	mse	rmse	mae	mdape	coverage
0	37 days	6.853411	2.617902	2.065823	0.100982	0.073711
1	38 days	6.851475	2.617532	2.071824	0.101032	0.073711
2	39 days	6.882861	2.623521	2.080719	0.101145	0.074131
3	40 days	6.925700	2.631634	2.088331	0.101144	0.074495
4	41 days	6.968539	2.640000	2.095943	0.101144	0.074859

Univariate Time Series Analysis | xWeather Prediction x+

localhost:8888/notebooks/Univariate%20Time%20Series%20Analysis%20for%20Weather%20Prediction%20using%20the%20Prophet%20Library.ipynb

jupyterUnivariate Time Series Analysis for Weather Prediction using the Prophet Li...Last Checkpoint: 2 hours ago (autosaved)Logout

FileEditViewInsertCellKernelWidgetsHelpTrustedKernel

RunCode

Obtaining the Performance Metrics

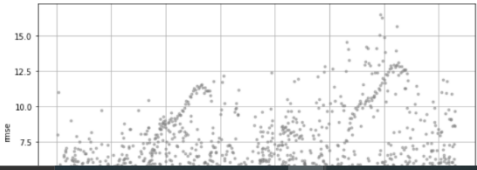
```
In [31]: from fbprophet.diagnostics import performance_metrics
df_p = performance_metrics(df_cv)
df_p.head()
```

```
Out[31]:
```

	horizon	mse	rmse	mae	mape	mdape	coverage
0	37 days	6.853411	2.617902	2.085823	0.100882	0.073711	0.796043
1	38 days	6.851475	2.617532	2.071824	0.101032	0.073711	0.793790
2	39 days	6.882891	2.623521	2.080719	0.101145	0.074131	0.790335
3	40 days	6.925709	2.631674	2.086371	0.101144	0.074186	0.788432
4	41 days	7.032192	2.651828	2.094893	0.100933	0.074186	0.784027

Visualizing Performance Metrics

```
In [32]: from fbprophet.plot import plot_cross_validation_metric
fig = plot_cross_validation_metric(df_cv, metric='rmse')
```



The plot shows RMSE on the y-axis (ranging from 7.5 to 15.0) against Horizon in days on the x-axis (ranging from 0 to 350). The data points are scattered, showing a general upward trend in RMSE as the horizon increases.

Univariate Time Series Analysis | xWeather Prediction x+

localhost:8888/notebooks/Univariate%20Time%20Series%20Analysis%20for%20Weather%20Prediction%20using%20the%20Prophet%20Library.ipynb

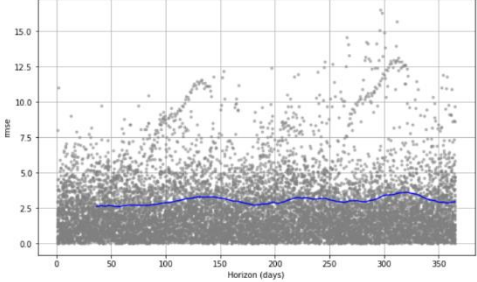
jupyterUnivariate Time Series Analysis for Weather Prediction using the Prophet Li...Last Checkpoint: 2 hours ago (autosaved)Logout

FileEditViewInsertCellKernelWidgetsHelpTrustedKernel

RunCode

Visualizing Performance Metrics

```
In [32]: from fbprophet.plot import plot_cross_validation_metric
fig = plot_cross_validation_metric(df_cv, metric='rmse')
```



The plot shows RMSE on the y-axis (ranging from 0.0 to 15.0) against Horizon in days on the x-axis (ranging from 0 to 350). The data points are scattered, and a blue line represents the trend, showing a slight increase in RMSE as the horizon increases.

Saving the model file

```
In [33]: import pickle
pickle.dump(model, open("weather_prediction.pickle", "wb"))
```