CS4701_Regional_Climate_Forecast_with_ANN_final

May 22, 2021

1 CS4701: Regional Weather Forcast with ANN

by Yifan Zhao(yz348) and Kristina Tian (yt523)

1.1 Abstract

In this project, we are developing a weather forecasting tool that aims to predict key regional weather parameters based on historical data. Various Artificial Neural Network methods are used in comparison to extract trends and predict patterns of the regional weather sequences. Multiple data visualization and analysis techniques are employed to aid feature engineering, feature selection, and model selection processes. Validation and test methods are employed to evaluate and compare the performance of the tool.

1.1.1 External Libraries

```
[485]: import numpy as np
      import pandas as pd
      import matplotlib.pyplot as plt
      from netCDF4 import Dataset
      import xarray as xr
       # for ML
      from sklearn.preprocessing import StandardScaler, MinMaxScaler, RobustScaler,
       →PolynomialFeatures
      from keras.models import Sequential
      from keras.layers import LSTM, Dense , Dropout, Bidirectional
      from sklearn.metrics import mean_squared_error
       ## for plotting
      import matplotlib.pyplot as plt
      import seaborn as sns
      import statsmodels.api as sm
      from statsmodels.tsa.stattools import adfuller
      import warnings
      warnings.filterwarnings("ignore")
```

1.2 1. Data

1.2.1 1.1 Basic data

This is a data gateway for downloading customized GCMs (Global Climate Models)datasets, downsacled using MACA method (Multivariate Adaptive Constructed Analogs). This method enables us to retrive historical metrological datasets as far as 1950 and with a resolution up to 4km. *This gateway is currently not in use by this project, where we instead directly import MACA datasets about the Finger Lakes AVA.

```
[]: lat_target=45.0
     lon_target=360-117.0
[]: pathname = 'http://thredds.northwestknowledge.net:8080/thredds/dodsC/
      -agg_macav2metdata_huss_BNU-ESM_r1i1p1_historical_1950_2005_CONUS_daily.nc'
[]: filehandle=Dataset(pathname, 'r', format="NETCDF4")
     lathandle=filehandle.variables['lat']
     lonhandle=filehandle.variables['lon']
     timehandle=filehandle.variables['time']
     datahandle=filehandle.variables['specific_humidity']
[]: time num=365
     timeindex=range(0,time_num,1)
     time=timehandle[timeindex]
     lat = lathandle[:]
     lon = lonhandle[:]
[]: #find indices of target lat/lon/day
     lat_index = (np.abs(lat-lat_target)).argmin()
     lon_index = (np.abs(lon-lon_target)).argmin()
     #check final is in right bounds
     if(lat[lat_index]>lat_target):
             if(lat_index!=0):
                     lat_index = lat_index - 1
     if(lat[lat_index]<lat_target):</pre>
             if(lat_index!=len(lat)):
                     lat_index =lat_index +1
     if(lon[lon_index]>lon_target):
             if(lon index!=0):
                     lon_index = lon_index - 1
     if(lon[lon_index]<lon_target):</pre>
             if(lon_index!=len(lon)):
                     lon_index = lon_index + 1
     lat=lat[lat_index]
     lon=lon[lon_index]
```

[]: data = datahandle[timeindex,lat_index,lon_index]

1.2.2 1.2 Import Data

Import MACA datasets about the Finger Lakes AVA regional GCMs.

```
[3]: # Fucntion for reading datasets
def read_file(file):
    df = pd.read_csv(file, sep=',',header=0, encoding='unicode_escape')
    return df

# Encode date to the number of days from 1/1/2020
def encode_date(df):
    return (pd.to_datetime(df['date']).rsub(pd.Timestamp('2006/1/1')).dt.
    days)*(-1)
```

```
[489]: from google.colab import drive import os drive.mount('/content/drive', force_remount=True)

path = os.path.join(os.getcwd(), "drive", "My Drive", □

→"Regional-Climate-Forecast-with-ANN")
```

```
[]: df_his_daily = read_file(os.path.join(path, 'MACA_41.9795 Latitude, -76.9813_\( \to \text{Longitude}_2006-2019.csv'))

df_cur_daily = read_file(os.path.join(path, 'MACA_41.9795 Latitude, -76.9813_\( \text{U} \to \text{Longitude}_2019-2020.csv'))

df_future_daily = read_file(os.path.join(path, 'MACA_41.9795 Latitude, -76.9813_\( \text{U} \to \text{Longitude}_2020-2021.csv'))

df_his_daily['date'] = encode_date(df_his_daily)
```

```
[478]: #Without Google Colab

df_his_daily = read_file('MACA_41.9795 Latitude, -76.9813 Longitude_2006-2019.

→csv')

df_cur_daily = read_file('MACA_41.9795 Latitude, -76.9813 Longitude_2019-2020.

→csv')
```

1.3 2. Feature Engineering

In this section, we analyze and visualize different features of the data to get a sense of what we are working with. These visualization show the distinct patterns of different features in the training set from 2006 to 2019. It also shows where anomalies are present, which will be addressed during normalization.

```
[311]: # functions for visualizing numerical data distributions
       def check_distribution_conti(dtf, x):
           fig, ax = plt.subplots(nrows=1, ncols=2, sharex=False, sharey=False)
           fig.suptitle(x, fontsize=20)
           ### distribution
           ax[0].title.set_text('distribution')
           variable = dtf[x].fillna(dtf[x].mean())
           breaks = np.quantile(variable, q=np.linspace(0, 1, 11))
           variable = variable[ (variable > breaks[0]) & (variable <</pre>
                                breaks[10]) ]
           sns.distplot(variable, hist=True, kde=True, kde_kws={"shade": True},__
        \Rightarrowax=ax[0])
           des = dtf[x].describe()
           ax[0].axvline(des["25%"], ls='--')
           ax[0].axvline(des["mean"], ls='--')
           ax[0].axvline(des["75%"], ls='--')
           ax[0].grid(True)
           des = round(des, 2).apply(lambda x: str(x))
```

```
ax[0].text(0.95, 0.95, box, transform=ax[0].transAxes, fontsize=10, __
        →va='top', ha="right", bbox=dict(boxstyle='round', facecolor='white', alpha=1))
           ### boxplot
           ax[1].title.set_text('outliers (log scale)')
           tmp_dtf = pd.DataFrame(dtf[x])
           tmp_dtf[x] = np.log(tmp_dtf[x])
           tmp_dtf.boxplot(column=x, ax=ax[1])
           plt.show()
       # functions for visualizing bivariant distribution between a categorical feature,
        →and a numerical feature
       def check bivariant cat(dtf, cat, num):
           fig, ax = plt.subplots(nrows=1, ncols=3, sharex=False, sharey=False)
           fig.suptitle(cat+" vs "+num, fontsize=20)
           ### distribution
           ax[0].title.set_text('density')
           for i in dtf[cat].unique():
               sns.distplot(dtf[dtf[cat]==i][num], hist=False, label=i, ax=ax[0])
           ax[0].grid(True)
           ### stacked
           ax[1].title.set_text('bins')
           breaks = np.quantile(dtf[num], q=np.linspace(0,1,11))
           tmp = dtf.groupby([cat, pd.cut(dtf[num], breaks, duplicates='drop')]).size().
        →unstack().T
           tmp = tmp[dtf[cat].unique()]
           tmp["tot"] = tmp.sum(axis=1)
           for col in tmp.drop("tot", axis=1).columns:
                tmp[col] = tmp[col] / tmp["tot"]
           tmp.drop("tot", axis=1).plot(kind='bar', stacked=True, ax=ax[1],__
        →legend=False, grid=True)
           ### boxplot
           ax[2].title.set text('outliers')
           sns.catplot(x=cat, y=num, data=dtf, kind="box", ax=ax[2])
           ax[2].grid(True)
           plt.show()
\lceil 471 \rceil: titles = \lceil
           "Daily Minimum Near-Surface Air Temperature(K)",
           "Daily Maximum Near-Surface Air Temperature(K)",
           "Daily Mean Near-Surface Specific Humidity(kgkg-1)",
           "vpd(kPa)".
           "Precipitation(mm)",
       ]
```

box = '\n'.join(("min: "+des["min"], "25%: "+des["25%"], "mean:

→"+des["mean"], "75%: "+des["75%"], "max: "+des["max"]))

```
feature_keys = [
    "Daily Minimum Near-Surface Air Temperature(K)",
    "Daily Maximum Near-Surface Air Temperature(K)",
    "Daily Mean Near-Surface Specific Humidity(kgkg-1)",
    "vpd(kPa)",
    "Precipitation(mm)",
]
colors = \Gamma
    "blue",
    "orange",
    "green",
    "red",
    "purple",
1
# Function for visualizing data distribution
def show_raw_visualization(data):
    time_data = data["date"]
    fig, axes = plt.subplots(
        nrows=3, ncols=2, figsize=(15, 20), dpi=80, facecolor="w", edgecolor="k"
    for i in range(len(feature_keys)):
        key = feature_keys[i]
        c = colors[i % (len(colors))]
        t_data = data[key]
        t_data.index = time_data
        t_data.head()
        ax = t_data.plot(
            ax=axes[i // 2, i % 2],
            color=c,
            title="{} - {}".format(titles[i], key),
            rot=25,
        ax.legend([titles[i]])
    plt.tight_layout()
# Function for visualizing correlation heatmap
def show_heatmap(data):
    plt.matshow(data.corr())
    plt.xticks(range(data.shape[1]), data.columns, fontsize=14, rotation=90)
    plt.gca().xaxis.tick_bottom()
    plt.yticks(range(data.shape[1]), data.columns, fontsize=14)
    cb = plt.colorbar()
    cb.ax.tick_params(labelsize=14)
```

```
plt.title("Feature Correlation Heatmap", fontsize=14)
plt.show()
```

```
[481]: check_distribution_conti(df_his_daily, 'Daily Minimum Near-Surface Air_

→Temperature(K)')

check_distribution_conti(df_his_daily, 'Daily Maximum Near-Surface Air_

→Temperature(K)')

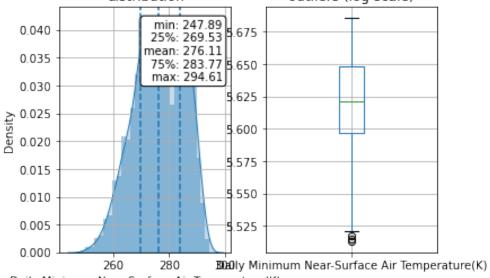
check_distribution_conti(df_his_daily, 'Daily Mean Near-Surface Specific_

→Humidity(kgkg-1)')

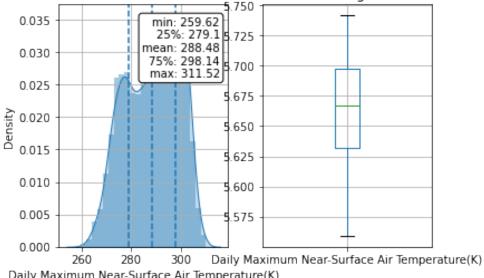
check_distribution_conti(df_his_daily, 'vpd(kPa)')

check_distribution_conti(df_his_daily, 'Precipitation(mm)')
```

Daily Minimum Near-Surface Air Temperature (K)

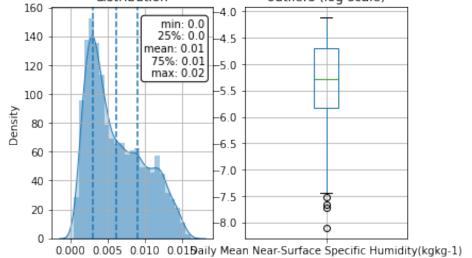


Daily Maximum Near-Surface Air Temperature (K)

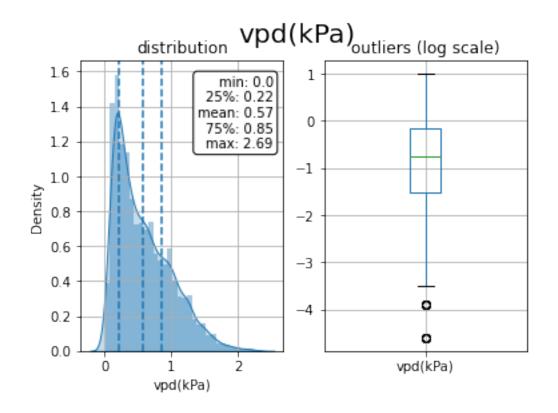


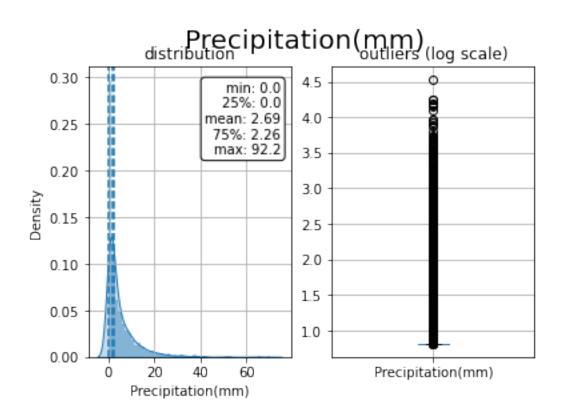
Daily Maximum Near-Surface Air Temperature(K)

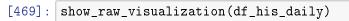
Daily Mean Near-Surface Specific Humidity(kgkg-1)

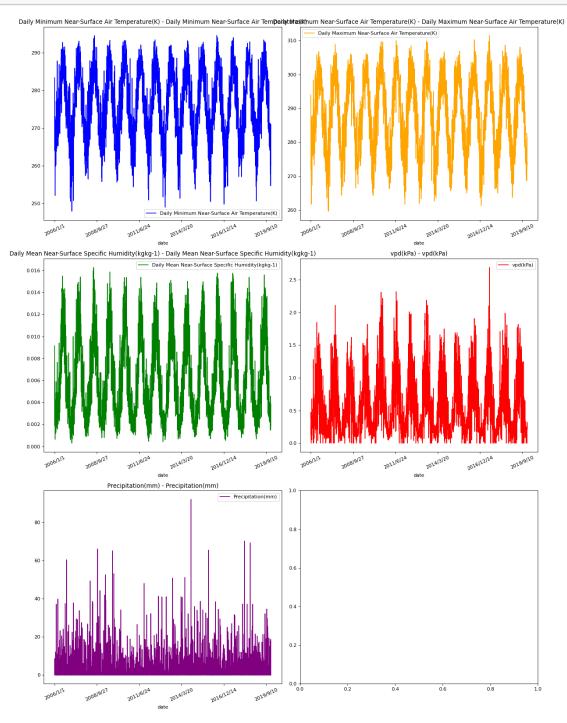


Daily Mean Near-Surface Specific Humidity(kgkg-1)

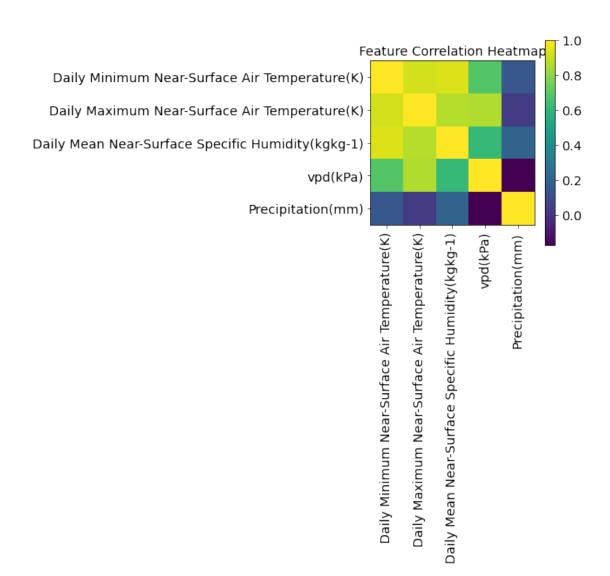








[480]: show_heatmap(df_his_daily.drop('date', 1))



```
[486]: # Check if the timeseries in the data is staionary

fig = plt.figure(figsize=(12,8))

ax1 = fig.add_subplot(211)

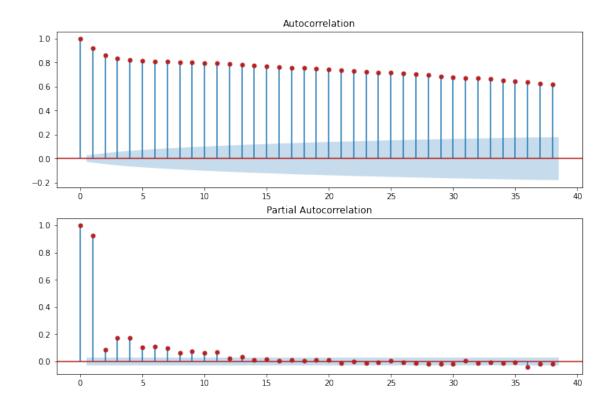
fig = sm.graphics.tsa.plot_acf(df_his_daily['Daily Maximum Near-Surface Air_

→Temperature(K)'], ax=ax1,color ='firebrick')

ax2 = fig.add_subplot(212)

fig = sm.graphics.tsa.plot_pacf(df_his_daily['Daily Maximum Near-Surface Air_

→Temperature(K)'], ax=ax2,color='firebrick')
```



```
[488]: # Perform AD Fuller Test on the data to test if the timeseries is stationary result = adfuller(df_his_daily['Daily Maximum Near-Surface Air Temperature(K)']) print('ADF Statistic on the entire dataset: {}'.format(result[0])) print('p-value: {}'.format(result[1])) print('Critical Values:') for key, value in result[4].items(): print('\t{}: {}'.format(key, value))
```

ADF Statistic on the entire dataset: -4.471451655114319 p-value: 0.0002213634266702279

Critical Values:

1%: -3.4316328616522527 5%: -2.862106888575586 10%: -2.5670717550557476

1.4 3. Train RNN for Predicting Tempreture

In this section, we train and evaluate a Pytorch RNN model in predicting Daily Maximum Near-Surface Air Temperature(K)

```
[332]: import torch import torch.nn as nn
```

1.4.1 3.1 Data preprocessing

```
[443]: # Data normalization
       sc = MinMaxScaler(feature_range=(0,1))
[444]: train_dataset = his_max_temp
       testdataset = cur_max_temp
[445]: train_dataset = sc.fit_transform(train_dataset)
       x_train = []
       v train = []
       n future = 1
       n_past = 30 \# Past 30 days
       for i in range(0,len(train_dataset)-n_past-n_future+1):
           x_train.append(train_dataset[i : i + n_past ])
           y_train.append(train_dataset[i + n_past : i + n_past + n_future ])
       x_train , y_train = np.array(x_train), np.array(y_train)
[446]: # Preprocessing test data
       testdataset = sc.transform(testdataset)
       x_test = []
       v_test = []
       n_future = 1
       n_past = 30
       for i in range(0,len(testdataset)-n_past-n_future+1):
           x_test.append(testdataset[i : i + n_past])
           y_test.append(testdataset[i + n_past : i + n_past + n_future, 0])
       x_test , y_test = np.array(x_test), np.array(y_test)
       x_test = np.reshape(x_test, (x_test.shape[0] , x_test.shape[1], 1) )
[447]: x_test_tensor = torch.Tensor(x_test)
       y_test_tensor = torch.Tensor(y_test)
       x_train_tensor = torch.Tensor(x_train)
       y_train_tensor = torch.Tensor(y_train)
[429]: x_train_tensor.shape
[429]: torch.Size([335, 30, 1])
      1.4.2 3.2 Train model
[430]: class RNN(nn.Module):
               def __init__(self, input_dim, hidden_dim, layer_dim, output_dim): # Add_
        \rightarrowrelevant parameters
                       super(RNN, self).__init__()
                       self.input_dim = input_dim
                       self.hidden_dim = hidden_dim
                       self.layer_dim = layer_dim
```

```
self.output_dim = output_dim
                self.rnn = nn.RNN(input_dim, hidden_dim, layer_dim,_
 ⇒batch_first=True, nonlinearity='relu')
                #torch.nn.init.kaiming_normal_(self.rnn.weight_ih_l0,__
 →nonlinearity='relu')
                #torch.nn.init.kaiming_normal_(self.rnn.weight_hh_l0,_
 \rightarrow nonlinearity='relu')
                self.lastlayer = nn.Linear(hidden_dim, output_dim)
                self.softmax = nn.LogSoftmax(dim=1)
                self.loss = nn.MSELoss()
        def compute_Loss(self, predicted_vector, gold_label):
                return self.loss(predicted_vector.float(), gold_label.
 →float())
        def forward(self, inputs):
                # begin code
                #print(inputs)
                output, final_hidden = self.rnn(inputs)
                #output = output[:, -1]
                final_output = self.lastlayer(output).squeeze(-1) #[np.
 \rightarrow arange(inputs.size(0)), seq_length-1,:]
                #final_output = final_output.squeeze(-1)
                #print(final_output)
                #print(final_output.shape)
                predicted_vector = torch.mode(final_output,1).values
                #print(predicted_vector)
                # end code
                return predicted_vector
        def load_model(self, save_path):
                self.load_state_dict(torch.load(save_path))
        def save_model(self, save_path):
                torch.save(self.state_dict(), save_path)
n = pochs = 1000
```

```
[431]: model = RNN(input_dim=1, output_dim=1, hidden_dim=12, layer_dim=1)
    n_epochs = 1000
    lr=0.1
    criterion = nn.MSELoss()
    optimizer = torch.optim.Adam(model.parameters(), lr=lr)
    for epoch in range(1, n_epochs + 1):
        optimizer.zero_grad() # Clears existing gradients from previous epoch
        #input_seq.to(device)
        output = model(x_train_tensor)
        #print(output)
        #print(y_train_tensor.view(-1))
```

```
y_train_tensor = y_train_tensor.view(-1) #.type(torch.LongTensor)
#output = (output.unsqueeze(-1))
loss = criterion(output.float(), y_train_tensor.float())
loss.backward() # Does backpropagation and calculates gradients
optimizer.step() # Updates the weights accordingly

if epoch%10 == 0:
    print('Epoch: {}/{}.....'.format(epoch, n_epochs), end=' ')
    print("Loss: {:.4f}".format(loss.item()))
```

```
Epoch: 10/1000... Loss: 0.1007
Epoch: 20/1000... Loss: 0.0585
Epoch: 30/1000... Loss: 0.0398
Epoch: 40/1000... Loss: 0.0337
Epoch: 50/1000... Loss: 0.0298
Epoch: 60/1000... Loss: 0.0281
Epoch: 70/1000... Loss: 0.0279
Epoch: 80/1000... Loss: 0.0277
Epoch: 90/1000... Loss: 0.0276
Epoch: 100/1000... Loss: 0.0276
Epoch: 110/1000... Loss: 0.0276
Epoch: 120/1000... Loss: 0.0276
Epoch: 130/1000... Loss: 0.0276
Epoch: 140/1000... Loss: 0.0276
Epoch: 150/1000... Loss: 0.0276
Epoch: 160/1000... Loss: 0.0276
Epoch: 170/1000... Loss: 0.0276
Epoch: 180/1000... Loss: 0.0276
Epoch: 190/1000... Loss: 0.0276
Epoch: 200/1000... Loss: 0.0276
Epoch: 210/1000... Loss: 0.0276
Epoch: 220/1000... Loss: 0.0276
Epoch: 230/1000... Loss: 0.0276
Epoch: 240/1000... Loss: 0.0276
Epoch: 250/1000... Loss: 0.0276
Epoch: 260/1000... Loss: 0.0276
Epoch: 270/1000... Loss: 0.0276
Epoch: 280/1000... Loss: 0.0276
Epoch: 290/1000... Loss: 0.0276
Epoch: 300/1000... Loss: 0.0276
Epoch: 310/1000... Loss: 0.0276
Epoch: 320/1000... Loss: 0.0276
Epoch: 330/1000... Loss: 0.0276
Epoch: 340/1000... Loss: 0.0276
Epoch: 350/1000... Loss: 0.0276
Epoch: 360/1000... Loss: 0.0276
Epoch: 370/1000... Loss: 0.0276
Epoch: 380/1000... Loss: 0.0276
```

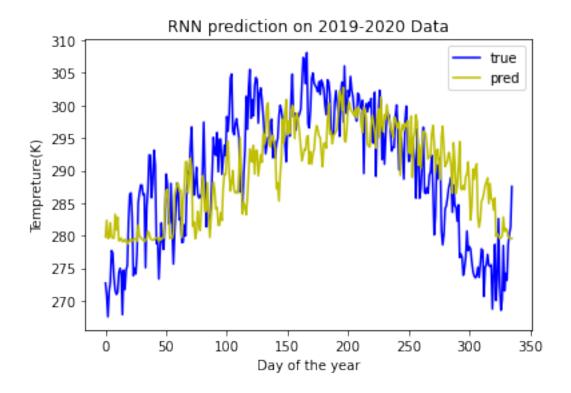
```
Epoch: 390/1000... Loss: 0.0276
Epoch: 400/1000... Loss: 0.0276
Epoch: 410/1000... Loss: 0.0276
Epoch: 420/1000... Loss: 0.0276
Epoch: 430/1000... Loss: 0.0276
Epoch: 440/1000... Loss: 0.0276
Epoch: 450/1000... Loss: 0.0276
Epoch: 460/1000... Loss: 0.0276
Epoch: 470/1000... Loss: 0.0276
Epoch: 480/1000... Loss: 0.0276
Epoch: 490/1000... Loss: 0.0276
Epoch: 500/1000... Loss: 0.0276
Epoch: 510/1000... Loss: 0.0276
Epoch: 520/1000... Loss: 0.0276
Epoch: 530/1000... Loss: 0.0276
Epoch: 540/1000... Loss: 0.0276
Epoch: 550/1000... Loss: 0.0276
Epoch: 560/1000... Loss: 0.0276
Epoch: 570/1000... Loss: 0.0276
Epoch: 580/1000... Loss: 0.0276
Epoch: 590/1000... Loss: 0.0276
Epoch: 600/1000... Loss: 0.0276
Epoch: 610/1000... Loss: 0.0276
Epoch: 620/1000... Loss: 0.0276
Epoch: 630/1000... Loss: 0.0276
Epoch: 640/1000... Loss: 0.0276
Epoch: 650/1000... Loss: 0.0276
Epoch: 660/1000... Loss: 0.0276
Epoch: 670/1000... Loss: 0.0276
Epoch: 680/1000... Loss: 0.0276
Epoch: 690/1000... Loss: 0.0276
Epoch: 700/1000... Loss: 0.0276
Epoch: 710/1000... Loss: 0.0276
Epoch: 720/1000... Loss: 0.0276
Epoch: 730/1000... Loss: 0.0276
Epoch: 740/1000... Loss: 0.0276
Epoch: 750/1000... Loss: 0.0276
Epoch: 760/1000... Loss: 0.0276
Epoch: 770/1000... Loss: 0.0276
Epoch: 780/1000... Loss: 0.0276
Epoch: 790/1000... Loss: 0.0276
Epoch: 800/1000... Loss: 0.0276
Epoch: 810/1000... Loss: 0.0276
Epoch: 820/1000... Loss: 0.0276
Epoch: 830/1000... Loss: 0.0276
Epoch: 840/1000... Loss: 0.0276
Epoch: 850/1000... Loss: 0.0276
Epoch: 860/1000... Loss: 0.0276
```

```
Epoch: 870/1000... Loss: 0.0276
Epoch: 880/1000... Loss: 0.0276
Epoch: 890/1000... Loss: 0.0276
Epoch: 900/1000... Loss: 0.0276
Epoch: 910/1000... Loss: 0.0276
Epoch: 920/1000... Loss: 0.0276
Epoch: 930/1000... Loss: 0.0276
Epoch: 940/1000... Loss: 0.0276
Epoch: 950/1000... Loss: 0.0276
Epoch: 960/1000... Loss: 0.0276
Epoch: 970/1000... Loss: 0.0276
Epoch: 980/1000... Loss: 0.0276
Epoch: 990/1000... Loss: 0.0276
Epoch: 990/1000... Loss: 0.0276
```

1.4.3 3.3 Test and Evaluation

```
[452]: pred = model(x_test_tensor)
    pred = pred.cpu().detach().numpy()
    pred = sc.inverse_transform(pred.reshape(-1, 1))
    y_test = sc.inverse_transform(y_test)
    mse = mean_squared_error(y_test, pred)
    print('MSE for Tempreture Prediction using RNN: ', mse)
    plt.plot(y_test, color = 'b', label = "true")
    plt.plot(pred, color = 'y', label = "pred")
    plt.ylabel('Tempreture(K)')
    plt.xlabel('Day of the year')
    plt.title("RNN prediction on 2019-2020 Data")
    plt.legend()
    plt.show()
```

MSE for Tempreture Prediction using RNN: 54.48186965591475



1.5 4. Train LSTM for Predicting Tempreture

In this section, we train and evaluate a LSTM RNN model in predicting Daily Maximum Near-Surface Air Temperature(K)

1.5.1 4.1 Data preprocessing

1.5.2 4.2 Train Model

```
[236]: # Train LSTM model
   regressor = Sequential()
   regressor.add(Bidirectional(LSTM(units=30, return_sequences=True, input_shape = ___
    \rightarrow(x_train.shape[1],1)))
   regressor.add(Dropout(0.2))
   regressor.add(LSTM(units= 30 , return_sequences=True))
   regressor.add(Dropout(0.2))
   regressor.add(LSTM(units= 30 , return_sequences=True))
   regressor.add(Dropout(0.2))
   regressor.add(LSTM(units= 30))
   regressor.add(Dropout(0.2))
   regressor.add(Dense(units = n_future,activation='linear'))
   regressor.compile(optimizer='adam', loss='mean_squared_error',metrics=['acc'])
   regressor.fit(x_train, y_train, epochs=500,batch_size=32 )
   Epoch 1/500
   0.0030
   Epoch 2/500
   0.0030
   Epoch 3/500
   0.0030
   Epoch 4/500
   0.0030
   Epoch 5/500
   0.0030
   Epoch 6/500
   0.0030
   Epoch 7/500
   0.0030
   Epoch 8/500
   0.0030
   Epoch 9/500
   0.0030
   Epoch 10/500
   0.0030
   Epoch 11/500
```

```
0.0030
Epoch 12/500
0.0030
Epoch 13/500
0.0030
Epoch 14/500
0.0030
Epoch 15/500
0.0030
Epoch 16/500
0.0030
Epoch 17/500
0.0030
Epoch 18/500
0.0030
Epoch 19/500
0.0030
Epoch 20/500
0.0030
Epoch 21/500
0.0030
Epoch 22/500
0.0030
Epoch 23/500
0.0030
Epoch 24/500
0.0030
Epoch 25/500
0.0030
Epoch 26/500
0.0030
Epoch 27/500
```

```
0.0030
Epoch 28/500
0.0030
Epoch 29/500
0.0030
Epoch 30/500
0.0030
Epoch 31/500
0.0030
Epoch 32/500
0.0030
Epoch 33/500
0.0030
Epoch 34/500
0.0030
Epoch 35/500
0.0030
Epoch 36/500
0.0030
Epoch 37/500
0.0030
Epoch 38/500
0.0030
Epoch 39/500
0.0030
Epoch 40/500
0.0030
Epoch 41/500
0.0030
Epoch 42/500
0.0030
Epoch 43/500
```

```
0.0030
Epoch 44/500
0.0030
Epoch 45/500
0.0030
Epoch 46/500
0.0030
Epoch 47/500
0.0030
Epoch 48/500
0.0030
Epoch 49/500
0.0030
Epoch 50/500
0.0030
Epoch 51/500
0.0030
Epoch 52/500
0.0030
Epoch 53/500
0.0030
Epoch 54/500
0.0030
Epoch 55/500
0.0030
Epoch 56/500
0.0030
Epoch 57/500
0.0030
Epoch 58/500
0.0030
Epoch 59/500
```

```
0.0030
Epoch 60/500
0.0030
Epoch 61/500
0.0030
Epoch 62/500
0.0030
Epoch 63/500
0.0030
Epoch 64/500
0.0030
Epoch 65/500
0.0030
Epoch 66/500
0.0030
Epoch 67/500
0.0030
Epoch 68/500
0.0030
Epoch 69/500
0.0030
Epoch 70/500
0.0030
Epoch 71/500
0.0030
Epoch 72/500
0.0030
Epoch 73/500
0.0030
Epoch 74/500
0.0030
Epoch 75/500
```

```
0.0030
Epoch 76/500
0.0030
Epoch 77/500
0.0030
Epoch 78/500
0.0030
Epoch 79/500
0.0030
Epoch 80/500
0.0030
Epoch 81/500
0.0030
Epoch 82/500
0.0030
Epoch 83/500
0.0030
Epoch 84/500
0.0030
Epoch 85/500
0.0030
Epoch 86/500
0.0030
Epoch 87/500
0.0030
Epoch 88/500
0.0030
Epoch 89/500
0.0030
Epoch 90/500
0.0030
Epoch 91/500
```

```
0.0030
Epoch 92/500
0.0030
Epoch 93/500
0.0030
Epoch 94/500
0.0030
Epoch 95/500
0.0030
Epoch 96/500
0.0030
Epoch 97/500
0.0030
Epoch 98/500
0.0030
Epoch 99/500
0.0030
Epoch 100/500
0.0030
Epoch 101/500
0.0030
Epoch 102/500
0.0030
Epoch 103/500
0.0030
Epoch 104/500
0.0030
Epoch 105/500
0.0030
Epoch 106/500
0.0030
Epoch 107/500
```

```
0.0030
Epoch 108/500
0.0030
Epoch 109/500
0.0030
Epoch 110/500
0.0030
Epoch 111/500
0.0030
Epoch 112/500
0.0030
Epoch 113/500
0.0030
Epoch 114/500
0.0030
Epoch 115/500
0.0030
Epoch 116/500
0.0030
Epoch 117/500
0.0030
Epoch 118/500
0.0030
Epoch 119/500
0.0030
Epoch 120/500
0.0030
Epoch 121/500
0.0030
Epoch 122/500
0.0030
Epoch 123/500
```

```
0.0030
Epoch 124/500
0.0030
Epoch 125/500
0.0030
Epoch 126/500
0.0030
Epoch 127/500
0.0030
Epoch 128/500
0.0030
Epoch 129/500
0.0030
Epoch 130/500
0.0030
Epoch 131/500
0.0030
Epoch 132/500
0.0030
Epoch 133/500
0.0030
Epoch 134/500
0.0030
Epoch 135/500
0.0030
Epoch 136/500
0.0030
Epoch 137/500
0.0030
Epoch 138/500
0.0030
Epoch 139/500
```

```
0.0030
Epoch 140/500
0.0030
Epoch 141/500
0.0030
Epoch 142/500
0.0030
Epoch 143/500
0.0030
Epoch 144/500
0.0030
Epoch 145/500
0.0030
Epoch 146/500
0.0030
Epoch 147/500
0.0030
Epoch 148/500
0.0030
Epoch 149/500
0.0030
Epoch 150/500
0.0030
Epoch 151/500
0.0030
Epoch 152/500
0.0030
Epoch 153/500
0.0030
Epoch 154/500
0.0030
Epoch 155/500
```

```
0.0030
Epoch 156/500
0.0030
Epoch 157/500
0.0030
Epoch 158/500
0.0030
Epoch 159/500
0.0030
Epoch 160/500
0.0030
Epoch 161/500
0.0030
Epoch 162/500
0.0030
Epoch 163/500
0.0030
Epoch 164/500
0.0030
Epoch 165/500
0.0030
Epoch 166/500
0.0030
Epoch 167/500
0.0030
Epoch 168/500
0.0030
Epoch 169/500
0.0030
Epoch 170/500
0.0030
Epoch 171/500
```

```
0.0030
Epoch 172/500
0.0030
Epoch 173/500
0.0030
Epoch 174/500
0.0030
Epoch 175/500
0.0030
Epoch 176/500
0.0030
Epoch 177/500
0.0030
Epoch 178/500
0.0030
Epoch 179/500
0.0030
Epoch 180/500
0.0030
Epoch 181/500
0.0030
Epoch 182/500
0.0030
Epoch 183/500
0.0030
Epoch 184/500
0.0030
Epoch 185/500
0.0030
Epoch 186/500
0.0030
Epoch 187/500
```

```
0.0030
Epoch 188/500
0.0030
Epoch 189/500
0.0030
Epoch 190/500
0.0030
Epoch 191/500
0.0030
Epoch 192/500
0.0030
Epoch 193/500
0.0030
Epoch 194/500
0.0030
Epoch 195/500
0.0030
Epoch 196/500
0.0030
Epoch 197/500
0.0030
Epoch 198/500
0.0030
Epoch 199/500
0.0030
Epoch 200/500
0.0030
Epoch 201/500
0.0030
Epoch 202/500
0.0030
Epoch 203/500
```

```
0.0030
Epoch 204/500
0.0030
Epoch 205/500
0.0030
Epoch 206/500
0.0030
Epoch 207/500
0.0030
Epoch 208/500
0.0030
Epoch 209/500
0.0030
Epoch 210/500
0.0030
Epoch 211/500
0.0030
Epoch 212/500
0.0030
Epoch 213/500
0.0030
Epoch 214/500
0.0030
Epoch 215/500
0.0030
Epoch 216/500
0.0030
Epoch 217/500
0.0030
Epoch 218/500
0.0030
Epoch 219/500
```

```
0.0030
Epoch 220/500
0.0030
Epoch 221/500
0.0030
Epoch 222/500
0.0030
Epoch 223/500
0.0030
Epoch 224/500
0.0030
Epoch 225/500
0.0030
Epoch 226/500
0.0030
Epoch 227/500
0.0030
Epoch 228/500
0.0030
Epoch 229/500
0.0030
Epoch 230/500
0.0030
Epoch 231/500
0.0030
Epoch 232/500
0.0030
Epoch 233/500
0.0030
Epoch 234/500
0.0030
Epoch 235/500
```

```
0.0030
Epoch 236/500
0.0030
Epoch 237/500
0.0030
Epoch 238/500
0.0030
Epoch 239/500
0.0030
Epoch 240/500
0.0030
Epoch 241/500
0.0030
Epoch 242/500
0.0030
Epoch 243/500
0.0030
Epoch 244/500
0.0030
Epoch 245/500
0.0030
Epoch 246/500
0.0030
Epoch 247/500
0.0030
Epoch 248/500
0.0030
Epoch 249/500
0.0030
Epoch 250/500
0.0030
Epoch 251/500
```

```
0.0030
Epoch 252/500
0.0030
Epoch 253/500
0.0030
Epoch 254/500
0.0030
Epoch 255/500
0.0030
Epoch 256/500
0.0030
Epoch 257/500
0.0030
Epoch 258/500
0.0030
Epoch 259/500
0.0030
Epoch 260/500
0.0030
Epoch 261/500
0.0030
Epoch 262/500
0.0030
Epoch 263/500
0.0030
Epoch 264/500
0.0030
Epoch 265/500
0.0030
Epoch 266/500
0.0030
Epoch 267/500
```

```
0.0030
Epoch 268/500
0.0030
Epoch 269/500
0.0030
Epoch 270/500
0.0030
Epoch 271/500
0.0030
Epoch 272/500
0.0030
Epoch 273/500
0.0030
Epoch 274/500
0.0030
Epoch 275/500
0.0030
Epoch 276/500
0.0030
Epoch 277/500
0.0030
Epoch 278/500
0.0030
Epoch 279/500
0.0030
Epoch 280/500
0.0030
Epoch 281/500
0.0030
Epoch 282/500
0.0030
Epoch 283/500
```

```
0.0030
Epoch 284/500
0.0030
Epoch 285/500
0.0030
Epoch 286/500
0.0030
Epoch 287/500
0.0030
Epoch 288/500
0.0030
Epoch 289/500
0.0030
Epoch 290/500
0.0030
Epoch 291/500
0.0030
Epoch 292/500
0.0030
Epoch 293/500
0.0030
Epoch 294/500
0.0030
Epoch 295/500
0.0030
Epoch 296/500
0.0030
Epoch 297/500
0.0030
Epoch 298/500
0.0030
Epoch 299/500
```

```
0.0030
Epoch 300/500
0.0030
Epoch 301/500
0.0030
Epoch 302/500
0.0030
Epoch 303/500
0.0030
Epoch 304/500
0.0030
Epoch 305/500
0.0030
Epoch 306/500
0.0030
Epoch 307/500
0.0030
Epoch 308/500
0.0030
Epoch 309/500
0.0030
Epoch 310/500
0.0030
Epoch 311/500
0.0030
Epoch 312/500
0.0030
Epoch 313/500
0.0030
Epoch 314/500
0.0030
Epoch 315/500
```

```
0.0030
Epoch 316/500
0.0030
Epoch 317/500
0.0030
Epoch 318/500
0.0030
Epoch 319/500
0.0030
Epoch 320/500
0.0030
Epoch 321/500
0.0030
Epoch 322/500
0.0030
Epoch 323/500
0.0030
Epoch 324/500
0.0030
Epoch 325/500
0.0030
Epoch 326/500
0.0030
Epoch 327/500
0.0030
Epoch 328/500
0.0030
Epoch 329/500
0.0030
Epoch 330/500
0.0030
Epoch 331/500
```

```
0.0030
Epoch 332/500
0.0030
Epoch 333/500
0.0030
Epoch 334/500
0.0030
Epoch 335/500
0.0030
Epoch 336/500
0.0030
Epoch 337/500
0.0030
Epoch 338/500
0.0030
Epoch 339/500
0.0030
Epoch 340/500
0.0030
Epoch 341/500
0.0030
Epoch 342/500
0.0030
Epoch 343/500
0.0030
Epoch 344/500
0.0030
Epoch 345/500
0.0030
Epoch 346/500
0.0030
Epoch 347/500
```

```
0.0030
Epoch 348/500
0.0030
Epoch 349/500
0.0030
Epoch 350/500
0.0030
Epoch 351/500
0.0030
Epoch 352/500
0.0030
Epoch 353/500
0.0030
Epoch 354/500
0.0030
Epoch 355/500
0.0030
Epoch 356/500
0.0030
Epoch 357/500
0.0030
Epoch 358/500
0.0030
Epoch 359/500
0.0030
Epoch 360/500
0.0030
Epoch 361/500
0.0030
Epoch 362/500
0.0030
Epoch 363/500
```

```
0.0030
Epoch 364/500
0.0030
Epoch 365/500
0.0030
Epoch 366/500
0.0030
Epoch 367/500
0.0030
Epoch 368/500
0.0030
Epoch 369/500
0.0030
Epoch 370/500
0.0030
Epoch 371/500
0.0030
Epoch 372/500
0.0030
Epoch 373/500
0.0030
Epoch 374/500
0.0030
Epoch 375/500
0.0030
Epoch 376/500
0.0030
Epoch 377/500
0.0030
Epoch 378/500
0.0030
Epoch 379/500
```

```
0.0030
Epoch 380/500
0.0030
Epoch 381/500
0.0030
Epoch 382/500
0.0030
Epoch 383/500
0.0030
Epoch 384/500
0.0030
Epoch 385/500
0.0030
Epoch 386/500
0.0030
Epoch 387/500
0.0030
Epoch 388/500
0.0030
Epoch 389/500
0.0030
Epoch 390/500
0.0030
Epoch 391/500
0.0030
Epoch 392/500
0.0030
Epoch 393/500
0.0030
Epoch 394/500
0.0030
Epoch 395/500
```

```
0.0030
Epoch 396/500
0.0030
Epoch 397/500
0.0030
Epoch 398/500
0.0030
Epoch 399/500
0.0030
Epoch 400/500
0.0030
Epoch 401/500
0.0030
Epoch 402/500
0.0030
Epoch 403/500
0.0030
Epoch 404/500
0.0030
Epoch 405/500
0.0030
Epoch 406/500
0.0030
Epoch 407/500
0.0030
Epoch 408/500
0.0030
Epoch 409/500
0.0030
Epoch 410/500
0.0030
Epoch 411/500
```

```
0.0030
Epoch 412/500
0.0030
Epoch 413/500
0.0030
Epoch 414/500
0.0030
Epoch 415/500
0.0030
Epoch 416/500
0.0030
Epoch 417/500
0.0030
Epoch 418/500
0.0030
Epoch 419/500
0.0030
Epoch 420/500
0.0030
Epoch 421/500
0.0030
Epoch 422/500
0.0030
Epoch 423/500
0.0030
Epoch 424/500
0.0030
Epoch 425/500
0.0030
Epoch 426/500
0.0030
Epoch 427/500
```

```
0.0030
Epoch 428/500
0.0030
Epoch 429/500
0.0030
Epoch 430/500
0.0030
Epoch 431/500
0.0030
Epoch 432/500
0.0030
Epoch 433/500
0.0030
Epoch 434/500
0.0030
Epoch 435/500
0.0030
Epoch 436/500
0.0030
Epoch 437/500
0.0030
Epoch 438/500
0.0030
Epoch 439/500
0.0030
Epoch 440/500
0.0030
Epoch 441/500
0.0030
Epoch 442/500
0.0030
Epoch 443/500
```

```
0.0030
Epoch 444/500
0.0030
Epoch 445/500
0.0030
Epoch 446/500
0.0030
Epoch 447/500
0.0030
Epoch 448/500
0.0030
Epoch 449/500
0.0030
Epoch 450/500
0.0030
Epoch 451/500
0.0030
Epoch 452/500
0.0030
Epoch 453/500
0.0030
Epoch 454/500
0.0030
Epoch 455/500
0.0030
Epoch 456/500
0.0030
Epoch 457/500
0.0030
Epoch 458/500
0.0030
Epoch 459/500
```

```
0.0030
Epoch 460/500
0.0030
Epoch 461/500
0.0030
Epoch 462/500
0.0030
Epoch 463/500
0.0030
Epoch 464/500
0.0030
Epoch 465/500
0.0030
Epoch 466/500
0.0030
Epoch 467/500
0.0030
Epoch 468/500
0.0030
Epoch 469/500
0.0030
Epoch 470/500
0.0030
Epoch 471/500
0.0030
Epoch 472/500
0.0030
Epoch 473/500
0.0030
Epoch 474/500
0.0030
Epoch 475/500
```

```
0.0030
Epoch 476/500
0.0030
Epoch 477/500
0.0030
Epoch 478/500
0.0030
Epoch 479/500
0.0030
Epoch 480/500
0.0030
Epoch 481/500
0.0030
Epoch 482/500
0.0030
Epoch 483/500
0.0030
Epoch 484/500
0.0030
Epoch 485/500
0.0030
Epoch 486/500
0.0030
Epoch 487/500
0.0030
Epoch 488/500
0.0030
Epoch 489/500
0.0030
Epoch 490/500
0.0030
Epoch 491/500
```

```
0.0030
Epoch 492/500
0.0030
Epoch 493/500
0.0030
Epoch 494/500
0.0030
Epoch 495/500
0.0030
Epoch 496/500
11/11 [============== ] - Os 26ms/step - loss: 0.0081 - acc:
0.0030
Epoch 497/500
0.0030
Epoch 498/500
0.0030
Epoch 499/500
0.0030
Epoch 500/500
0.0030
```

[236]: <tensorflow.python.keras.callbacks.History at 0x13a2a2dfa60>

1.5.3 4.3 Test and Evaluate

1) Evaluate performance with single step prediction of Tempreture on the 31th day of 2019

```
[276]: # Plot single step prediction

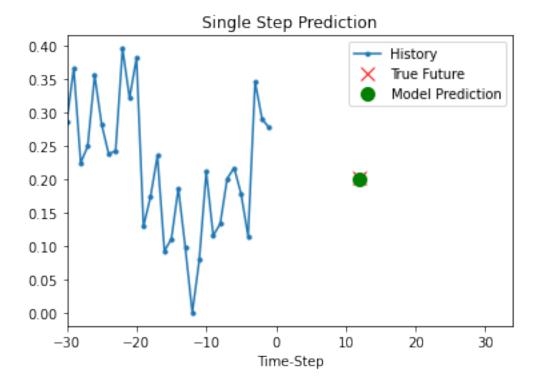
# Modified from: Team, K. (n.d.). Keras documentation: Timeseries forecasting

→ for weather prediction. Keras. https://keras.io/examples/timeseries/

→ timeseries_weather_forecasting/.

def show_plot(plot_data, delta, title):
    labels = ["History", "True Future", "Model Prediction"]
    marker = [".-", "rx", "go"]
    time_steps = list(range(-(plot_data[0].shape[0]), 0))
    if delta:
        future = delta
    else:
        future = 0
```

```
plt.title(title)
           for i, val in enumerate(plot_data):
                   plt.plot(future, plot_data[i], marker[i], markersize=10,__
        →label=labels[i])
               else:
                   plt.plot(time_steps, plot_data[i].flatten(), marker[i],__
        →label=labels[i])
           plt.legend()
           plt.xlim([time_steps[0], (future + 5) * 2])
           plt.xlabel("Time-Step")
           plt.show()
           return
[288]: # Preprocessing test data
       testdataset = cur_max_temp[:31]
       testdataset = sc.transform(testdataset)
       x test = []
       y_test = []
       n_future = 1
       n_past = 30
       for i in range(0,len(testdataset)-n_past-n_future+1):
           x_test.append(testdataset[i : i + n_past, 0])
           y_test.append(testdataset[i + n_past : i + n_past + n_future, 0])
       x_test , y_test = np.array(x_test), np.array(y_test)
       x_test = np.reshape(x_test, (x_test.shape[0] , x_test.shape[1], 1) )
[299]: show_plot(
               [x_test[0], y_test[0], regressor.predict(x_test)[0]],
               "Single Step Prediction",
           )
```

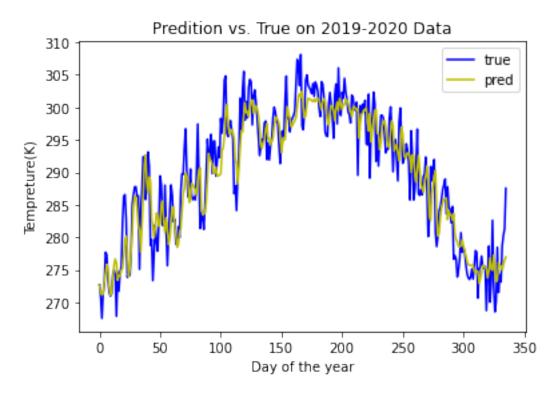


2) Evaluate performance with 2019-2020 test set

```
[301]: # Making prediction and evaluation
pred_30 = regressor.predict(x_test)
pred_30 = sc.inverse_transform(pred_30)
y_test = sc.inverse_transform(y_test)
mse = mean_squared_error(y_test, pred_30)
print('MSE for Tempreture Prediction using LSTM: ', mse)
plt.plot(y_test, color = 'b', label = "true")
plt.plot(pred_30, color = 'y', label = "pred")
plt.ylabel('Tempreture(K)')
```

```
plt.xlabel('Day of the year')
plt.title("LSTM prediction on 2019-2020 Data")
plt.legend()
plt.show()
```

MSE for Tempreture Prediction using LSTM: 14.898463512340081



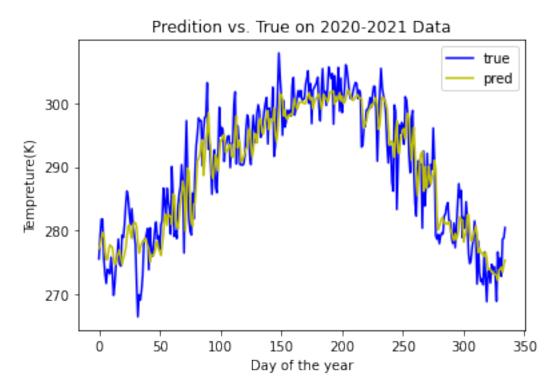
3) Evaluate performance with 2020-2021 test set

pred_2021 = regressor.predict(x_test)

```
[302]: # Preprocessing test set
    testdataset = future_max_temp
    testdataset = sc.transform(testdataset)
    x_test = []
    y_test = []
    n_future = 1
    n_past = 30
    for i in range(0,len(testdataset)-n_past-n_future+1):
        x_test.append(testdataset[i : i + n_past, 0])
        y_test.append(testdataset[i + n_past : i + n_past + n_future, 0])
    x_test, y_test = np.array(x_test), np.array(y_test)
    x_test = np.reshape(x_test, (x_test.shape[0], x_test.shape[1], 1))
[303]: # Making prediction and evaluation
```

```
pred_2021 = sc.inverse_transform(pred_2021)
y_test = sc.inverse_transform(y_test)
mse = mean_squared_error(y_test, pred_2021)
print('MSE for Tempreture Prediction using LSTM: ', mse)
plt.plot(y_test, color = 'b', label = "true")
plt.plot(pred_2021, color = 'y', label = "pred")
plt.ylabel('Tempreture(K)')
plt.xlabel('Day of the year')
plt.title("LSTM prediction on 2020-2021 Data")
plt.legend()
plt.show()
```

MSE for Tempreture Prediction using LSTM: 16.6580981560278



1.6 References:

Kosandal, R. (2020, January 5). Weather forecasting with Recurrent Neural Networks. Medium. https://medium.com/analytics-vidhya/weather-forecasting-with-recurrent-neural-networks-1eaa057d70c3. Team, K. (n.d.). Keras documentation: Timeseries forecasting for weather prediction. Keras. https://keras.io/examples/timeseries/timeseries_weather_forecasting/. Fatmakursun. (2020, January 28). Rain Forecasting with Artificial Neural Network. Kaggle. https://www.kaggle.com/fatmakursun/rain-forecasting-with-artificial-neural-network. Pietro, M. D. (2020, December 8). Machine Learning with Python: Regression. Medium. https://towardsdatascience.com/machine-learning-with-python-regression-complete-tutorial-

 $47268e546cea.\ Paialunga,\ P.\ (2021,\ April\ 18).\ Weather\ forecasting\ with\ Machine\ Learning,\ using\ Python.\ Medium.\ https://towardsdatascience.com/weather-forecasting-with-machine-learning-using-python-55e90c346647.$