Time series forecasting

View on TensorFlow.org (https://www.tensorflow.org/tutorials/structured_data/time_series)	Run in Google Colab (https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/tuto

This tutorial is an introduction to time series forecasting using Recurrent Neural Networks (RNNs). This is covered in two parts: first, you will forecast a univariate time series, then you will forecast a multivariate time series.

In [2]:
```
from __future__ import absolute_import, division, print_function, unicode_literals
try:
  # %tensorflow_version only exists in Colab.
  %tensorflow_version 2.x
except Exception:
  pass
import tensorflow as tf

import matplotlib as mpl
import matplotlib.pyplot as plt
import numpy as np
import os
import pandas as pd

mpl.rcParams['figure.figsize'] = (8, 6)
mpl.rcParams['axes.grid'] = False
```

The weather dataset

This tutorial uses a weather time series dataset (https://www.bgc-jena.mpg.de/wetter/) recorded by the Max-Planck-Institute for Biogeochemistry (https://www.bgc-jena.mpg.de/index.php/Main/HomePage).

This dataset contains 14 different features such as air temperature, atmospheric pressure, and humidity. These were collected every 10 minutes, beginning in 2003. For efficiency, you will use only the data collected between 2009 and 2016. This section of the dataset was prepared by François Chollet for his book Deep Learning with Python (https://www.manning.com/books/deep-learning-with-python).

In [3]:
```
zip_path = tf.keras.utils.get_file(
    origin='https://storage.googleapis.com/tensorflow/tf-keras-datasets/jena_climate_2009_2016.csv.z
ip',
    fname='jena_climate_2009_2016.csv.zip',
    extract=True)
csv_path, _ = os.path.splitext(zip_path)
```

```
Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datasets/jena_climate_2009_
2016.csv.zip
13574144/13568290 [==============================] - 1s 0us/step
```

In [4]:
```
df = pd.read_csv(csv_path)
```

Let's take a glance at the data.

In [5]:
```
df.head()
```

Out[5]:

Out[5]:

	Date Time	p (mbar)	T (degC)	Tpot (K)	Tdew (degC)	rh (%)	VPmax (mbar)	VPact (mbar)	VPdef (mbar)	sh (g/kg)	H2OC (mmol/mol)	rho (g/m**3)	wv (m/s)	max wv (m/s
0	01.01.2009 00:10:00	996.52	-8.02	265.40	-8.90	93.3	3.33	3.11	0.22	1.94	3.12	1307.75	1.03	1.75
1	01.01.2009 00:20:00	996.57	-8.41	265.01	-9.28	93.4	3.23	3.02	0.21	1.89	3.03	1309.80	0.72	1.50
2	01.01.2009 00:30:00	996.53	-8.51	264.91	-9.31	93.9	3.21	3.01	0.20	1.88	3.02	1310.24	0.19	0.63
3	01.01.2009 00:40:00	996.51	-8.31	265.12	-9.07	94.2	3.26	3.07	0.19	1.92	3.08	1309.19	0.34	0.50
4	01.01.2009 00:50:00	996.51	-8.27	265.15	-9.04	94.1	3.27	3.08	0.19	1.92	3.09	1309.00	0.32	0.63

As you can see above, an observation is recorded every 10 mintues. This means that, for a single hour, you will have 6 observations. Similarly, a single day will contain 144 (6x24) observations.

Given a specific time, let's say you want to predict the temperature 6 hours in the future. In order to make this prediction, you choose to use 5 days of observations. Thus, you would create a window containing the last 720(5x144) observations to train the model. Many such configurations are possible, making this dataset a good one to experiment with.

The function below returns the above described windows of time for the model to train on. The parameter `history_size` is the size of the past window of information. The `target_size` is how far in the future does the model need to learn to predict. The `target_size` is the label that needs to be predicted.

In [6]:
```
def univariate_data(dataset, start_index, end_index, history_size, target_size):
  data = []
  labels = []

  start_index = start_index + history_size
  if end_index is None:
    end_index = len(dataset) - target_size

  for i in range(start_index, end_index):
    indices = range(i-history_size, i)
    # Reshape data from (history_size,) to (history_size, 1)
    data.append(np.reshape(dataset[indices], (history_size, 1)))
    labels.append(dataset[i+target_size])
  return np.array(data), np.array(labels)
```

In both the following tutorials, the first 300,000 rows of the data will be the training dataset, and there remaining will be the validation dataset. This amounts to ~2100 days worth of training data.

In [7]:
```
TRAIN_SPLIT = 300000
```

Setting seed to ensure reproducibility.

In [8]:
```
tf.random.set_seed(13)
```

Part 1: Forecast a univariate time series

First, you will train a model using only a single feature (temperature), and use it to make predictions for that value in the future.

Let's first extract only the temperature from the dataset.

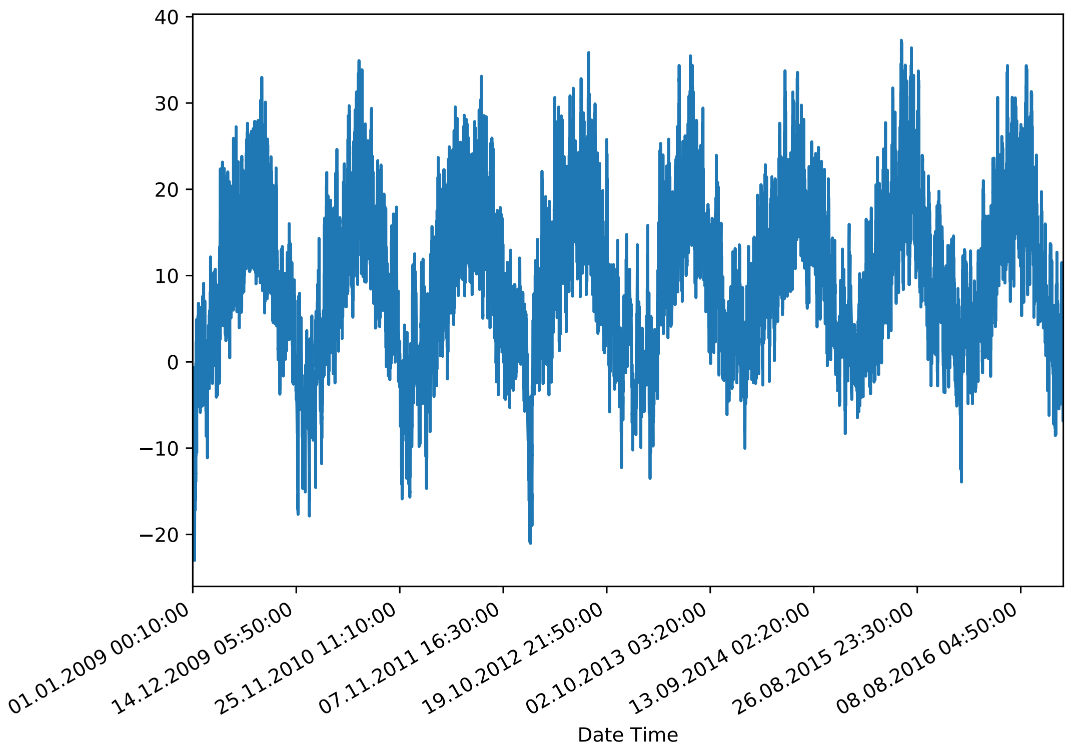
In [10]:
```
uni_data = df['T (degC)']
uni_data.index = df['Date Time']
uni_data.head()
```

Out[10]:
```
Date Time
01.01.2009 00:10:00   -8.02
01.01.2009 00:20:00   -8.41
01.01.2009 00:30:00   -8.51
01.01.2009 00:40:00   -8.31
01.01.2009 00:50:00   -8.27
Name: T (degC), dtype: float64
```

Let's observe how this data looks across time.

In [11]:
```
uni_data.plot(subplots=True)
```

Out[11]:
```
array([<matplotlib.axes._subplots.AxesSubplot object at 0x000002129A76DD48>],
      dtype=object)
```



In [12]:
```
uni_data = uni_data.values
```

It is important to scale features before training a neural network. Standardization is a common way of doing this scaling by subtracting the mean and dividing by the standard deviation of each feature.You could also use a `tf.keras.utils.normalize` method that rescales the values into a range of [0,1].

Note: The mean and standard deviation should only be computed using the training data.

In [13]:
```
uni_train_mean = uni_data[:TRAIN_SPLIT].mean()
uni_train_std = uni_data[:TRAIN_SPLIT].std()
```

Let's standardize the data.

In [14]:
```
uni_data = (uni_data-uni_train_mean)/uni_train_std
```

Let's now create the data for the univariate model. For part 1, the model will be given the last 20 recorded temperature observations, and needs to learn to predict the temperature at the next time step.

In [15]:
```
univariate_past_history = 20
univariate_future_target = 0

x_train_uni, y_train_uni = univariate_data(uni_data, 0, TRAIN_SPLIT,
                                           univariate_past_history,
                                           univariate_future_target)
x_val_uni, y_val_uni = univariate_data(uni_data, TRAIN_SPLIT, None,
                                       univariate_past_history,
                                       univariate_future_target)
```

This is what the `univariate_data` function returns.

In [16]:
```
print ('Single window of past history')
print (x_train_uni[0])
```

```
print ('\n Target temperature to predict')
print (y_train_uni[0])
```

```
Single window of past history
[[-1.99766294]
 [-2.04281897]
 [-2.05439744]
 [-2.0312405 ]
 [-2.02660912]
 [-2.00113649]
 [-1.95134907]
 [-1.95134907]
 [-1.98492663]
 [-2.04513467]
 [-2.08334362]
 [-2.09723778]
 [-2.09376424]
 [-2.09144854]
 [-2.07176515]
 [-2.07176515]
 [-2.07639653]
 [-2.08913285]
 [-2.09260639]
 [-2.10418486]]

 Target temperature to predict
-2.1041848598100876
```

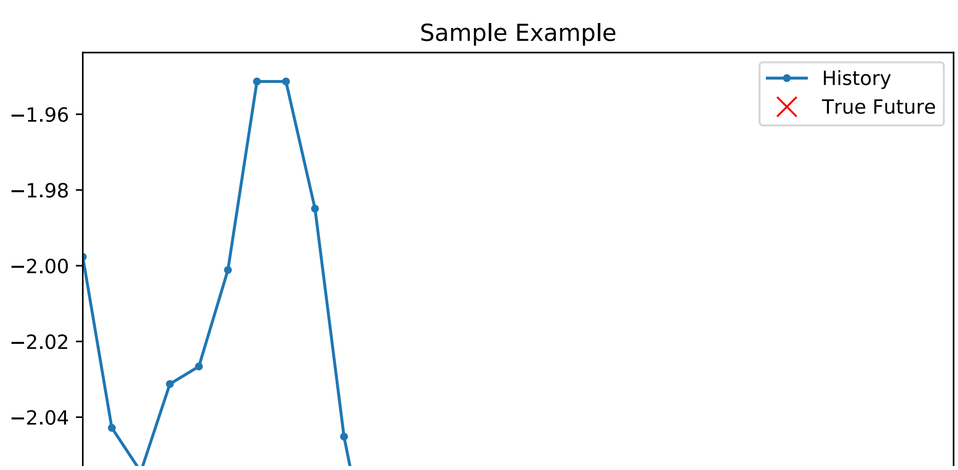
Now that the data has been created, let's take a look at a single example. The information given to the network is given in blue, and it must predict the value at the red cross.

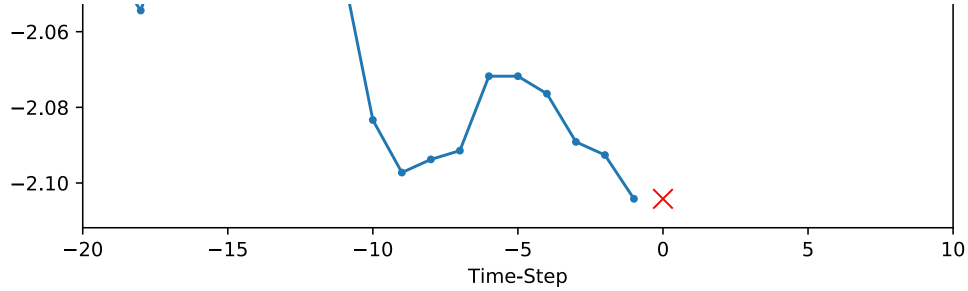
In [17]:
```
def create_time_steps(length):
  return list(range(-length, 0))
```

In [18]:
```
def show_plot(plot_data, delta, title):
  labels = ['History', 'True Future', 'Model Prediction']
  marker = ['.-', 'rx', 'go']
  time_steps = create_time_steps(plot_data[0].shape[0])
  if delta:
    future = delta
  else:
    future = 0

  plt.title(title)
  for i, x in enumerate(plot_data):
    if i:
      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')
  return plt
```

In [19]:
```
show_plot([x_train_uni[0], y_train_uni[0]], 0, 'Sample Example')
```

Out[19]:
```
<module 'matplotlib.pyplot' from 'C:\\Users\\xuesong.wang1\\AppData\\Local\\Programs\\Python\\Python
37\\lib\\site-packages\\matplotlib\\pyplot.py'>
```





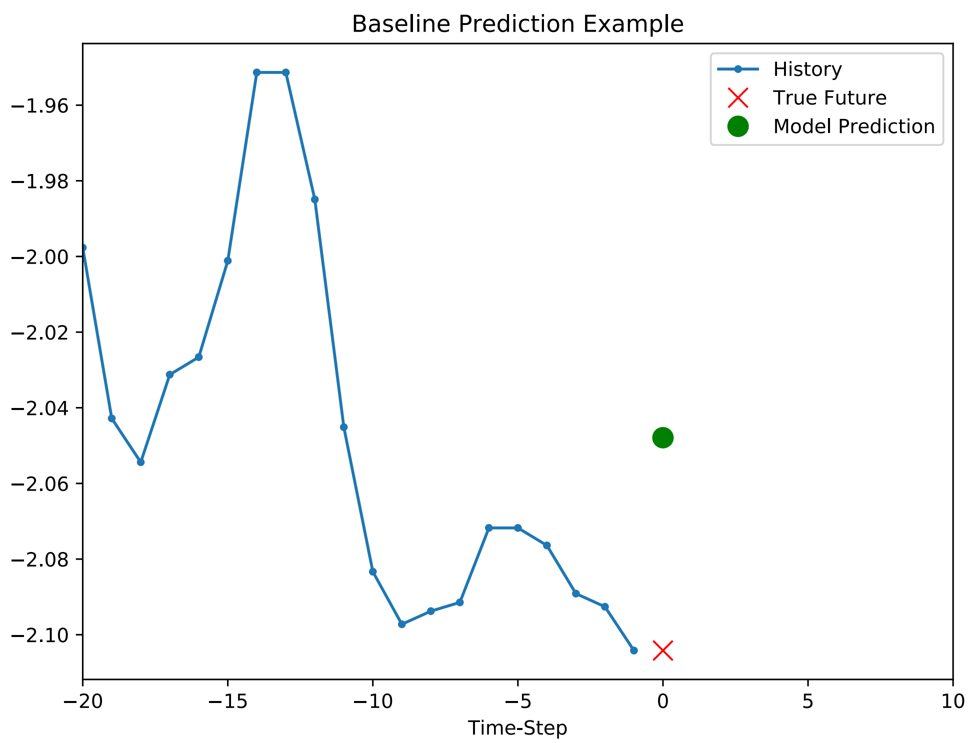
Baseline

Before proceeding to train a model, let's first set a simple baseline. Given an input point, the baseline method looks at all the history and predicts the next point to be the average of the last 20 observations.

In [20]:
```
def baseline(history):
  return np.mean(history)
```

In [21]:
```
show_plot([x_train_uni[0], y_train_uni[0], baseline(x_train_uni[0])], 0,
           'Baseline Prediction Example')
```

Out[21]:
```
<module 'matplotlib.pyplot' from 'C:\\Users\\xuesong.wang1\\AppData\\Local\\Programs\\Python\\Python
37\\lib\\site-packages\\matplotlib\\pyplot.py'>
```



Let's see if you can beat this baseline using a recurrent neural network.

Recurrent neural network

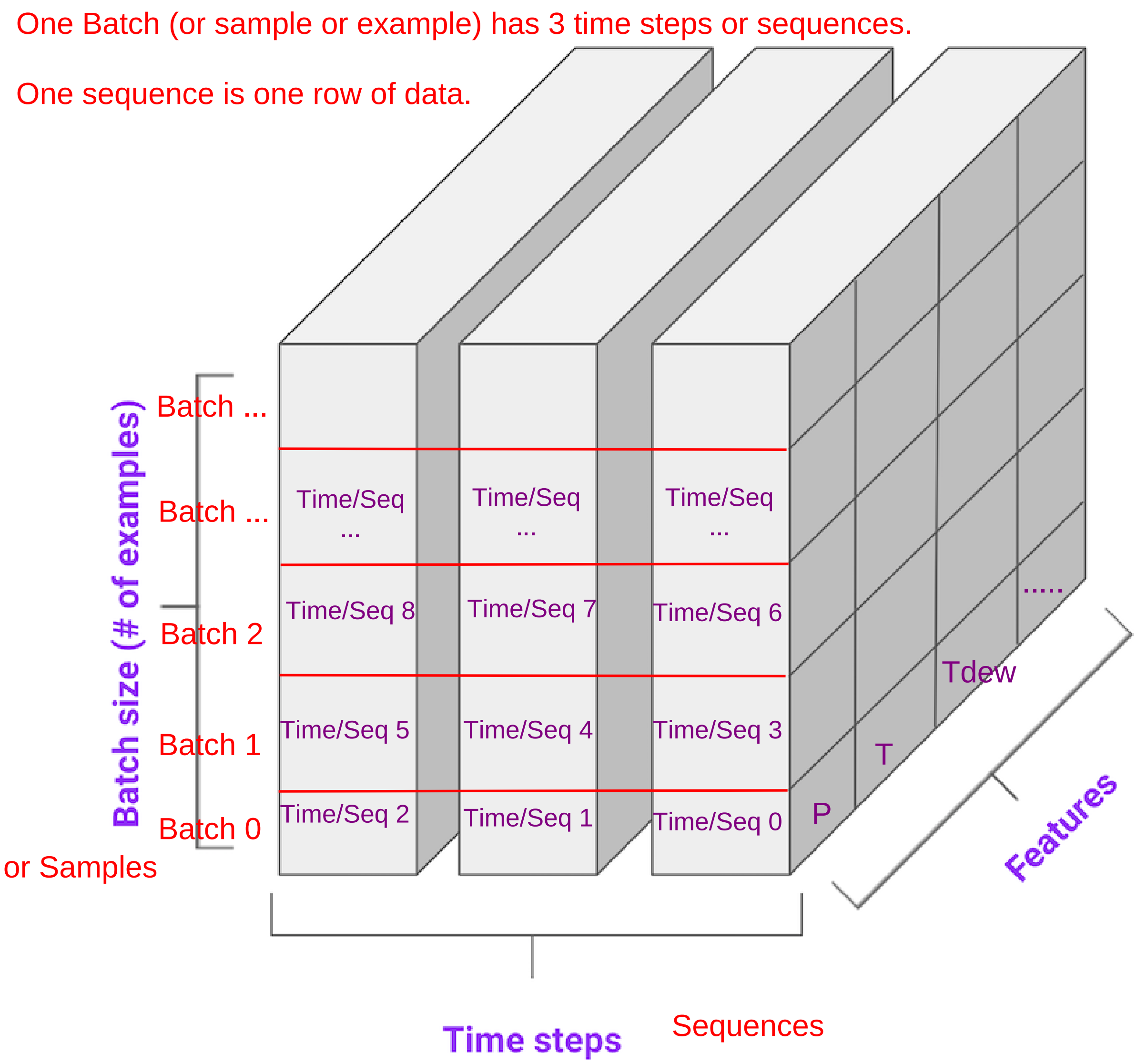
A Recurrent Neural Network (RNN) is a type of neural network well-suited to time series data. RNNs process a time series step-by-step, maintaining an internal state summarizing the information they've seen so far. For more details, read the RNN tutorial (https://www.tensorflow.org/tutorials/sequences/recurrent). In this tutorial, you will use a specialized RNN layer called Long Short Term Memory (LSTM (https://www.tensorflow.org/versions/r2.0/api_docs/python/tf/keras/layers/LSTM))

Let's now use `tf.data` to shuffle, batch, and cache the dataset.

In [22]:
```
BATCH_SIZE = 256
BUFFER_SIZE = 10000

train_univariate = tf.data.Dataset.from_tensor_slices((x_train_uni, y_train_uni))
train_univariate = train_univariate.cache().shuffle(BUFFER_SIZE).batch(BATCH_SIZE).repeat()
```

```
val_univariate = tf.data.Dataset.from_tensor_slices((x_val_uni, y_val_uni))
val_univariate = val_univariate.batch(BATCH_SIZE).repeat()
```

The following visualisation should help you understand how the data is represented after batching.



You will see the LSTM requires the input shape of the data it is being given.

In [23]:
```
simple_lstm_model = tf.keras.models.Sequential([
    tf.keras.layers.LSTM(8, input_shape=x_train_uni.shape[-2:]),
    tf.keras.layers.Dense(1)
])

simple_lstm_model.compile(optimizer='adam', loss='mae')
```

Let's make a sample prediction, to check the output of the model.

In [24]:
```
for x, y in val_univariate.take(1):
    print(simple_lstm_model.predict(x).shape)
```

```
(256, 1)
```

Let's train the model now. Due to the large size of the dataset, in the interest of saving time, each epoch will only run for 200 steps, instead of the complete training data as normally done.

In [25]:
```
EVALUATION_INTERVAL = 200
EPOCHS = 10

simple_lstm_model.fit(train_univariate, epochs=EPOCHS,
```

```
                    steps_per_epoch=EVALUATION_INTERVAL,
                    validation_data=val_univariate, validation_steps=50)
```

```
Train for 200 steps, validate for 50 steps
Epoch 1/10
200/200 [==============================] - 5s 27ms/step - loss: 0.4075 - val_loss: 0.1351
Epoch 2/10
200/200 [==============================] - 3s 14ms/step - loss: 0.1118 - val_loss: 0.0359
Epoch 3/10
200/200 [==============================] - 3s 13ms/step - loss: 0.0489 - val_loss: 0.0290
Epoch 4/10
200/200 [==============================] - 2s 12ms/step - loss: 0.0443 - val_loss: 0.0258
Epoch 5/10
200/200 [==============================] - 3s 13ms/step - loss: 0.0299 - val_loss: 0.0235
Epoch 6/10
200/200 [==============================] - 3s 13ms/step - loss: 0.0317 - val_loss: 0.0224
Epoch 7/10
200/200 [==============================] - 2s 12ms/step - loss: 0.0286 - val_loss: 0.0207
Epoch 8/10
200/200 [==============================] - 2s 12ms/step - loss: 0.0263 - val_loss: 0.0197
Epoch 9/10
200/200 [==============================] - 2s 11ms/step - loss: 0.0253 - val_loss: 0.0181
Epoch 10/10
200/200 [==============================] - 2s 12ms/step - loss: 0.0227 - val_loss: 0.0174
```

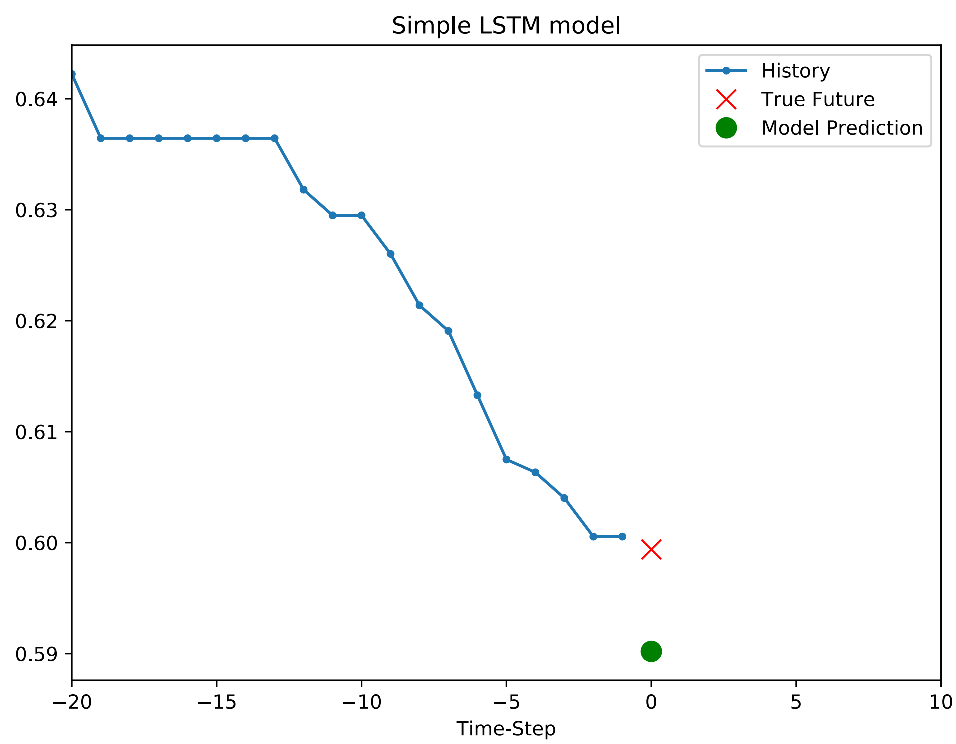
Out[25]: <tensorflow.python.keras.callbacks.History at 0x2129a943b08>

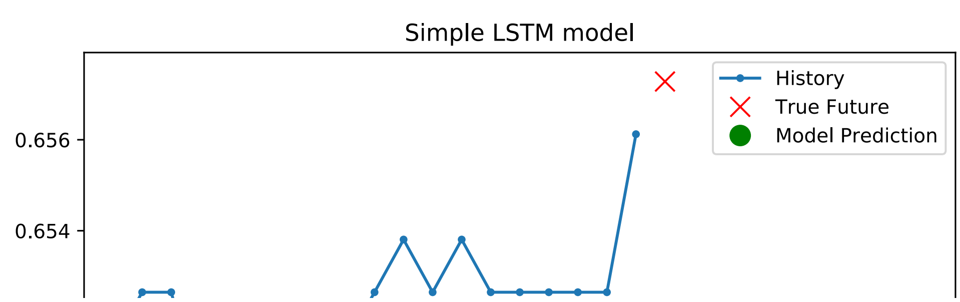
Predict using the simple LSTM model

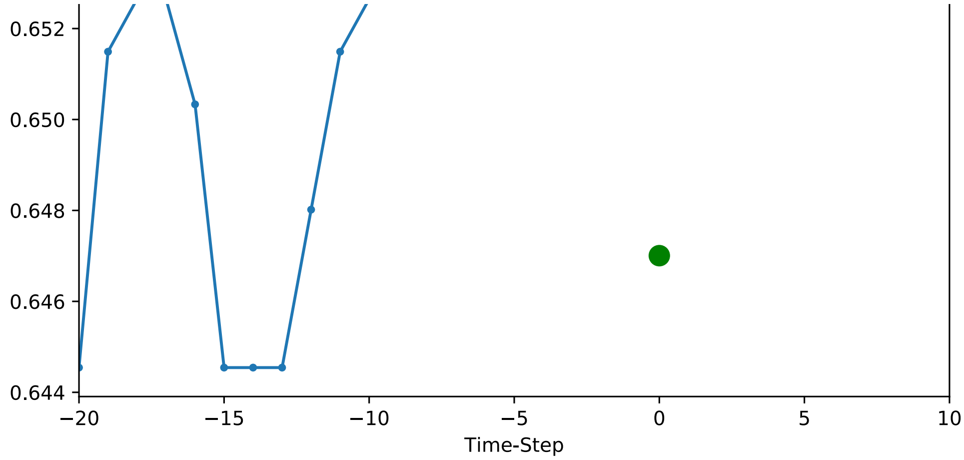
Now that you have trained your simple LSTM, let's try and make a few predictions.

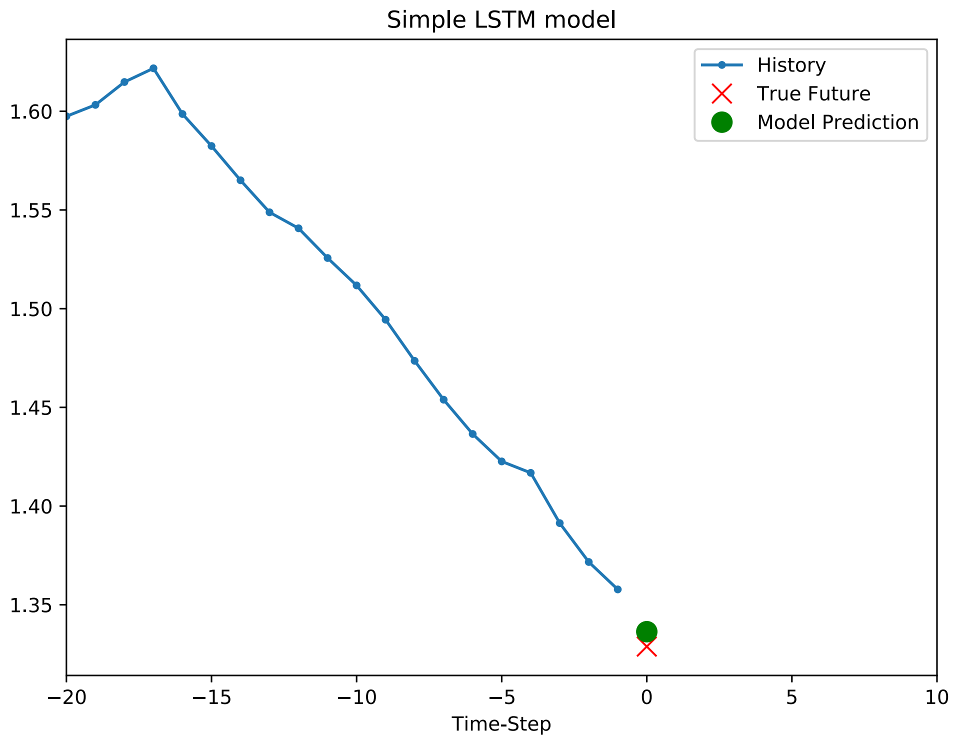
In [26]:

```
for x, y in val_univariate.take(3):
  plot = show_plot([x[0].numpy(), y[0].numpy(),
                    simple_lstm_model.predict(x)[0]], 0, 'Simple LSTM model')
  plot.show()
```









This looks better than the baseline. Now that you have seen the basics, let's move on to part two, where you will work with a multivariate time series.

Part 2: Forecast a multivariate time series

The original dataset contains fourteen features. For simplicity, this section considers only three of the original fourteen. The features used are air temperature, atmospheric pressure, and air density.

To use more features, add their names to this list.

In [27]:
```
features_considered = ['p (mbar)', 'T (degC)', 'rho (g/m**3)']
```

In [28]:
```
features = df[features_considered]
features.index = df['Date Time']
features.head()
```

Out[28]:

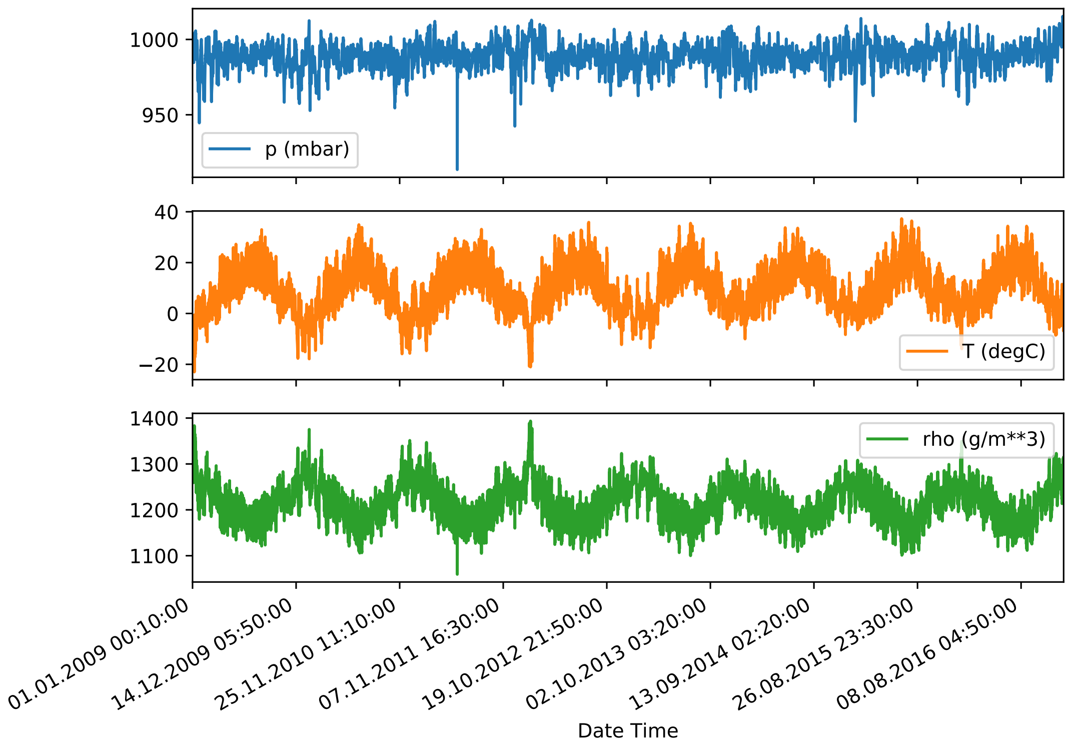
	p (mbar)	T (degC)	rho (g/m**3)
Date Time			
01.01.2009 00:10:00	996.52	-8.02	1307.75
01.01.2009 00:20:00	996.57	-8.41	1309.80
01.01.2009 00:30:00	996.53	-8.51	1310.24

01.01.2009 00:40:00	996.51	-8.31	1309.19
01.01.2009 00:50:00	996.51	-8.27	1309.00

Let's have a look at how each of these features vary across time.

In [29]:
```
features.plot(subplots=True)
```

Out[29]:
```
array([<matplotlib.axes._subplots.AxesSubplot object at 0x00000212AA6475C8>,
       <matplotlib.axes._subplots.AxesSubplot object at 0x00000212ABF0E208>,
       <matplotlib.axes._subplots.AxesSubplot object at 0x00000212AAC1EF08>],
      dtype=object)
```



As mentioned, the first step will be to standardize the dataset using the mean and standard deviation of the training data.

In [30]:
```
dataset = features.values
data_mean = dataset[:TRAIN_SPLIT].mean(axis=0)
data_std = dataset[:TRAIN_SPLIT].std(axis=0)
```

In [31]:
```
dataset = (dataset-data_mean)/data_std
```

Single step model

In a single step setup, the model learns to predict a single point in the future based on some history provided.

The below function performs the same windowing task as below, however, here it samples the past observation based on the step size given.

In [32]:
```
def multivariate_data(dataset, target, start_index, end_index, history_size,
                      target_size, step, single_step=False):
  data = []
  labels = []

  start_index = start_index + history_size
  if end_index is None:
    end_index = len(dataset) - target_size

  for i in range(start_index, end_index):
    indices = range(i-history_size, i, step)
    data.append(dataset[indices])

    if single_step:
```

```
      labels.append(target[i+target_size])
    else:
      labels.append(target[i:i+target_size])

  return np.array(data), np.array(labels)
```

In this tutorial, the network is shown data from the last five (5) days, i.e. 720 observations that are sampled every hour. The sampling is done every one hour since a drastic change is not expected within 60 minutes. Thus, 120 observation represent history of the last five days. For the single step prediction model, the label for a datapoint is the temperature 12 hours into the future. In order to create a label for this, the temperature after 72(12*6) observations is used.

In [34]:
```
past_history = 720
future_target = 72
STEP = 6

x_train_single, y_train_single = multivariate_data(dataset, dataset[:, 1], 0,
                                                   TRAIN_SPLIT, past_history,
                                                   future_target, STEP,
                                                   single_step=True)
x_val_single, y_val_single = multivariate_data(dataset, dataset[:, 1],
                                               TRAIN_SPLIT, None, past_history,
                                               future_target, STEP,
                                               single_step=True)
```

Let's look at a single data-point.

In [35]:
```
print ('Single window of past history : {}'.format(x_train_single[0].shape))
```

```
Single window of past history : (120, 3)
```

In [36]:
```
train_data_single = tf.data.Dataset.from_tensor_slices((x_train_single, y_train_single))
train_data_single = train_data_single.cache().shuffle(BUFFER_SIZE).batch(BATCH_SIZE).repeat()

val_data_single = tf.data.Dataset.from_tensor_slices((x_val_single, y_val_single))
val_data_single = val_data_single.batch(BATCH_SIZE).repeat()
```

In [41]:
```
x_train_single.shape
```

Out[41]:
```
(299280, 120, 3)
```

In [37]:
```
single_step_model = tf.keras.models.Sequential()
single_step_model.add(tf.keras.layers.LSTM(32,
                                           input_shape=x_train_single.shape[-2:]))
single_step_model.add(tf.keras.layers.Dense(1))

single_step_model.compile(optimizer=tf.keras.optimizers.RMSprop(), loss='mae')
```

Let's check out a sample prediction.

In [38]:
```
for x, y in val_data_single.take(1):
  print(single_step_model.predict(x).shape)
```

```
(256, 1)
```

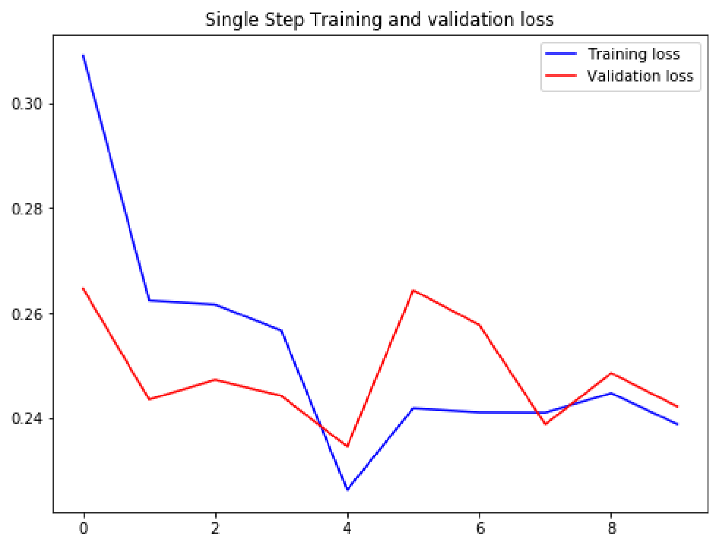
In [39]:
```
single_step_history = single_step_model.fit(train_data_single, epochs=EPOCHS,
                                            steps_per_epoch=EVALUATION_INTERVAL,
                                            validation_data=val_data_single,
                                            validation_steps=50)
```

```
Train for 200 steps, validate for 50 steps
Epoch 1/10
200/200 [==============================] - 31s 155ms/step - loss: 0.3090 - val_loss: 0.2647
Epoch 2/10
200/200 [==============================] - 31s 157ms/step - loss: 0.2625 - val_loss: 0.2432
Epoch 3/10
200/200 [==============================] - 38s 189ms/step - loss: 0.2614 - val_loss: 0.2476
Epoch 4/10
200/200 [==============================] - 45s 226ms/step - loss: 0.2566 - val_loss: 0.2447
Epoch 5/10
200/200 [==============================] - 49s 246ms/step - loss: 0.2267 - val_loss: 0.2360
Epoch 6/10
200/200 [==============================] - 56s 278ms/step - loss: 0.2413 - val_loss: 0.2667
Epoch 7/10
200/200 [==============================] - 57s 283ms/step - loss: 0.2414 - val_loss: 0.2577
Epoch 8/10
```

```
Epoch 8/10
200/200 [==============================] - 53s 267ms/step - loss: 0.2407 - val_loss: 0.2371
Epoch 9/10
200/200 [==============================] - 52s 262ms/step - loss: 0.2447 - val_loss: 0.2486
Epoch 10/10
200/200 [==============================] - 54s 269ms/step - loss: 0.2385 - val_loss: 0.2445
```

In [42]:
```
def plot_train_history(history, title):
  loss = history.history['loss']
  val_loss = history.history['val_loss']

  epochs = range(len(loss))

  plt.figure()

  plt.plot(epochs, loss, 'b', label='Training loss')
  plt.plot(epochs, val_loss, 'r', label='Validation loss')
  plt.title(title)
  plt.legend()

  plt.show()
```

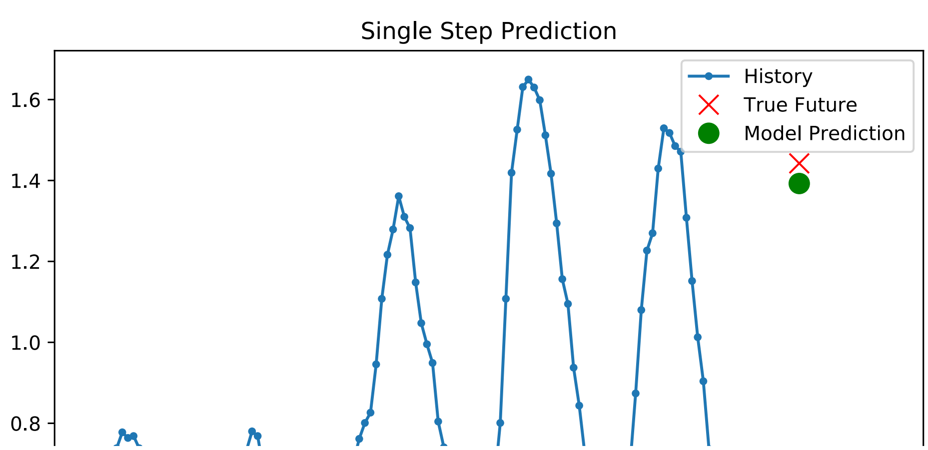
In [39]:
```
plot_train_history(single_step_history,
                   'Single Step Training and validation loss')
```

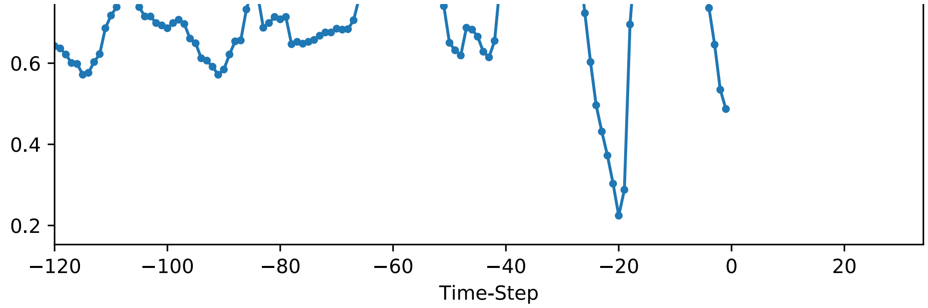


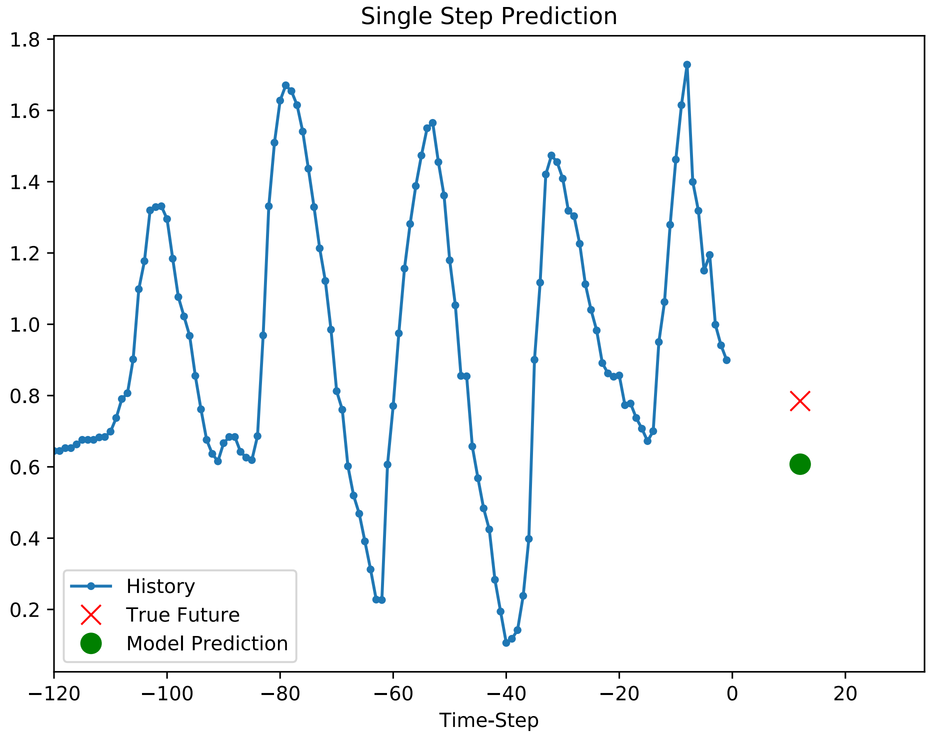
Predict a single step future

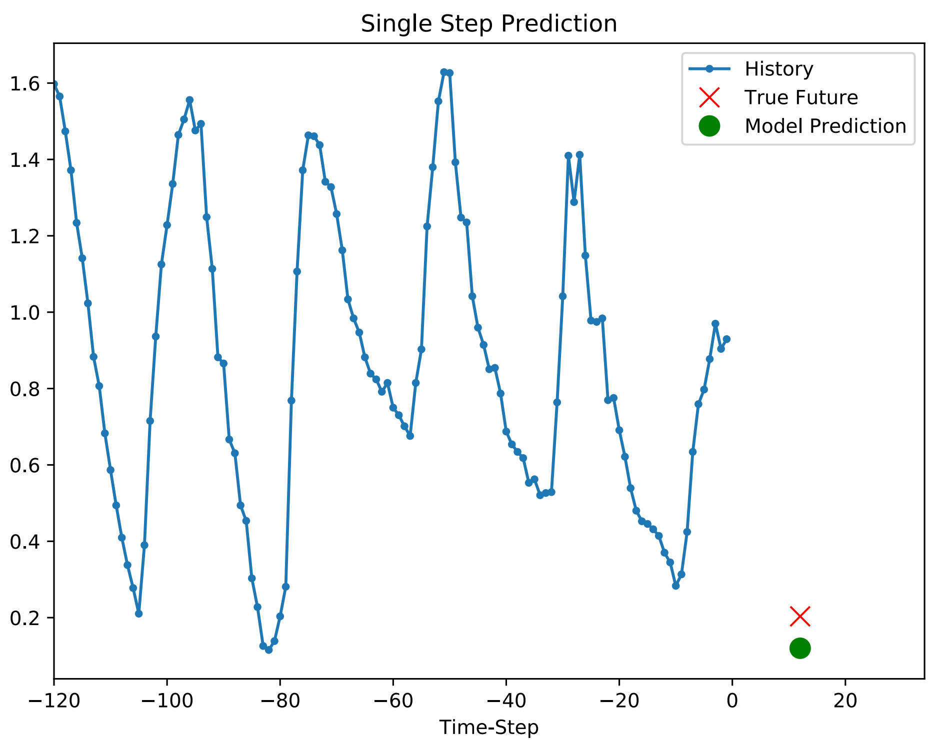
Now that the model is trained, let's make a few sample predictions. The model is given the history of three features over the past five days sampled every hour (120 data-points), since the goal is to predict the temperature, the plot only displays the past temperature. The prediction is made one day into the future (hence the gap between the history and prediction).

In [43]:
```
for x, y in val_data_single.take(3):
  plot = show_plot([x[0][:, 1].numpy(), y[0].numpy(),
                    single_step_model.predict(x)[0]], 12,
                   'Single Step Prediction')
  plot.show()
```









Multi-Step model

In a multi-step prediction model, given a past history, the model needs to learn to predict a range of future values. Thus, unlike a single step model, where only a single future point is predicted, a multi-step model predict a sequence of the future.

For the multi-step model, the training data again consists of recordings over the past five days sampled every hour. However, here, the model needs to learn to predict the temperature for the next 12 hours. Since an obversation is taken every 10 minutes, the output is 72 predictions. For this task, the dataset needs to be prepared accordingly, thus the first step is just to create it again, but with a different target window.

In [44]:
```
future_target = 72
x_train_multi, y_train_multi = multivariate_data(dataset, dataset[:, 1], 0,
                                                 TRAIN_SPLIT, past_history,
                                                 future_target, STEP)
x_val_multi, y_val_multi = multivariate_data(dataset, dataset[:, 1],
                                             TRAIN_SPLIT, None, past_history,
                                             future_target, STEP)
```

Let's check out a sample data-point.

In [45]:
```
print ('Single window of past history : {}'.format(x_train_multi[0].shape))
print ('\n Target temperature to predict : {}'.format(y_train_multi[0].shape))
```

```
Single window of past history : (120, 3)

 Target temperature to predict : (72,)
```

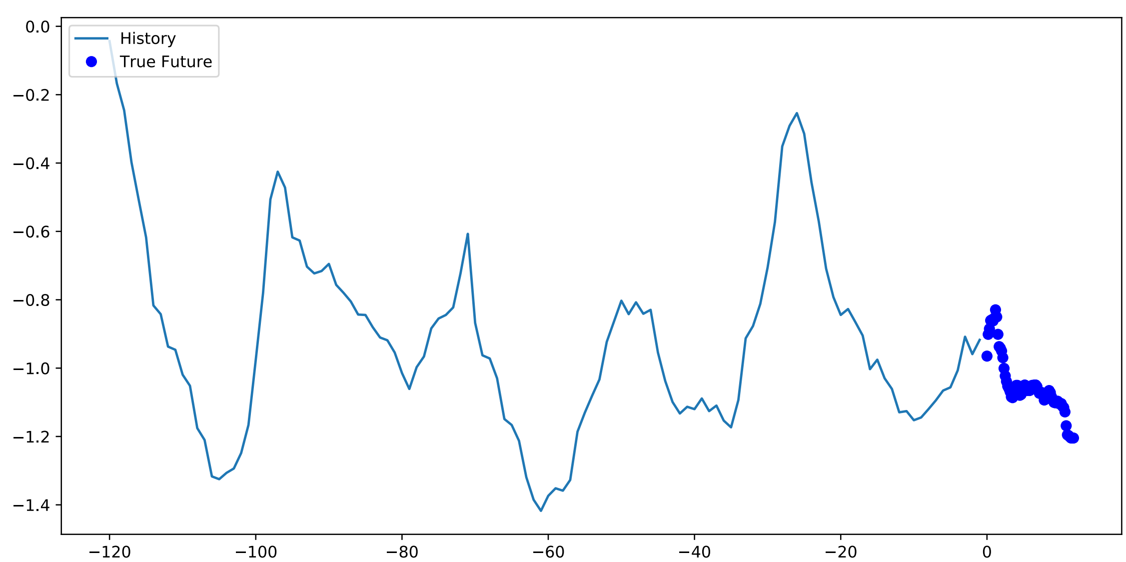
In [46]:
```
train_data_multi = tf.data.Dataset.from_tensor_slices((x_train_multi, y_train_multi))
train_data_multi = train_data_multi.cache().shuffle(BUFFER_SIZE).batch(BATCH_SIZE).repeat()

val_data_multi = tf.data.Dataset.from_tensor_slices((x_val_multi, y_val_multi))
val_data_multi = val_data_multi.batch(BATCH_SIZE).repeat()
```

Plotting a sample data-point.

In [47]:
```
def multi_step_plot(history, true_future, prediction):
  plt.figure(figsize=(12, 6))
  num_in = create_time_steps(len(history))
  num_out = len(true_future)

  plt.plot(num_in, np.array(history[:, 1]), label='History')
  plt.plot(np.arange(num_out)/STEP, np.array(true_future), 'bo',
           label='True Future')
  if prediction.any():
    plt.plot(np.arange(num_out)/STEP, np.array(prediction), 'ro',
             label='Predicted Future')
  plt.legend(loc='upper left')
  plt.show()
```

In this plot and subsequent similar plots, the history and the future data are sampled every hour.

In [48]:
```
for x, y in train_data_multi.take(1):
  multi_step_plot(x[0], y[0], np.array([0]))
```



Since the task here is a bit more complicated than the previous task, the model now consists of two LSTM layers. Finally, since 72 predictions are made, the dense layer outputs 72 predictions.

In [49]:
```
multi_step_model = tf.keras.models.Sequential()
multi_step_model.add(tf.keras.layers.LSTM(32,
                                          return_sequences=True,
                                          input_shape=x_train_multi.shape[-2:]))
multi_step_model.add(tf.keras.layers.LSTM(16, activation='relu'))
multi_step_model.add(tf.keras.layers.Dense(72))

multi_step_model.compile(optimizer=tf.keras.optimizers.RMSprop(clipvalue=1.0), loss='mae')
```

Let's see how the model predicts before it trains.

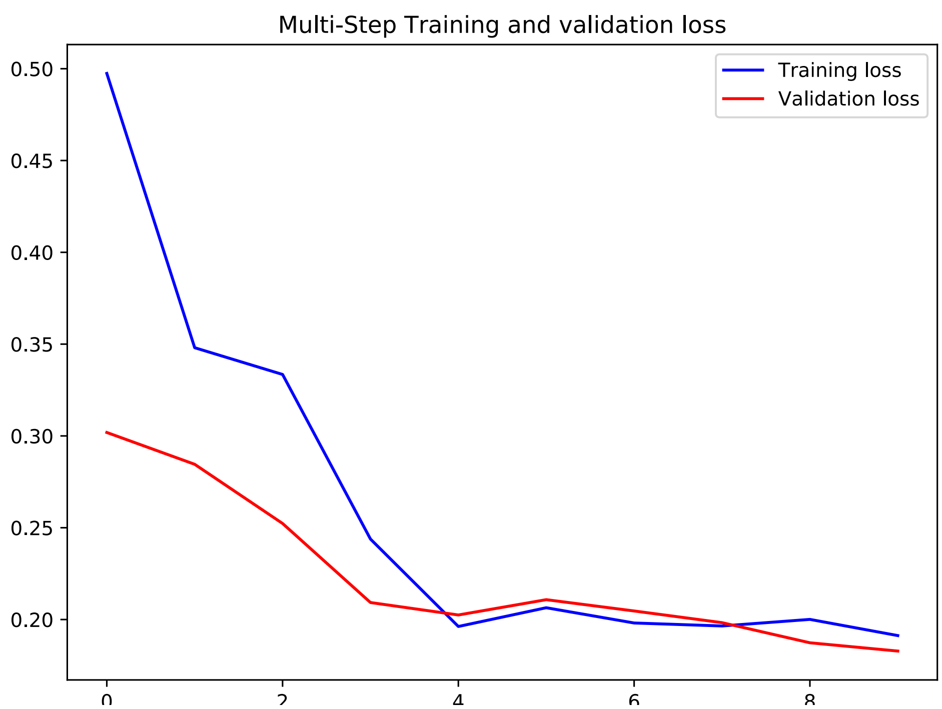
In [50]:
```
for x, y in val_data_multi.take(1):
  print (multi_step_model.predict(x).shape)
```

```
(256, 72)
```

In [51]:
```
multi_step_history = multi_step_model.fit(train_data_multi, epochs=EPOCHS,
                                          steps_per_epoch=EVALUATION_INTERVAL,
                                          validation_data=val_data_multi,
                                          validation_steps=50)
```

```
Train for 200 steps, validate for 50 steps
Epoch 1/10
200/200 [==============================] - 84s 422ms/step - loss: 0.4974 - val_loss: 0.3019
Epoch 2/10
200/200 [==============================] - 87s 437ms/step - loss: 0.3480 - val_loss: 0.2845
Epoch 3/10
200/200 [==============================] - 94s 471ms/step - loss: 0.3335 - val_loss: 0.2523
Epoch 4/10
200/200 [==============================] - 109s 546ms/step - loss: 0.2438 - val_loss: 0.2093
Epoch 5/10
200/200 [==============================] - 120s 599ms/step - loss: 0.1962 - val_loss: 0.2025
Epoch 6/10
200/200 [==============================] - 126s 632ms/step - loss: 0.2062 - val_loss: 0.2108
Epoch 7/10
200/200 [==============================] - 146s 732ms/step - loss: 0.1981 - val_loss: 0.2047
Epoch 8/10
200/200 [==============================] - 129s 647ms/step - loss: 0.1965 - val_loss: 0.1983
Epoch 9/10
200/200 [==============================] - 143s 713ms/step - loss: 0.2001 - val_loss: 0.1873
Epoch 10/10
200/200 [==============================] - 173s 864ms/step - loss: 0.1913 - val_loss: 0.1828
```

In [52]:
```
plot_train_history(multi_step_history, 'Multi-Step Training and validation loss')
```



Predict a multi-step future

Let's now have a look at how well your network has learnt to predict the future.

In [53]:
```
for x, y in val_data_multi.take(3):
  multi_step_plot(x[0], y[0], multi_step_model.predict(x)[0])
```