UNIVERSITY of BRADFORD

Faculty of Engineering and Informatics
Department of Computer Science

Mars Weather Prediction Model

Machine Learning

Yusuf Furkan Baki

19001214

25 May 2023

Declaration

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others. The candidate agrees that this report can be electronically checked for plagiarism.

Yusuf Furkan Baki

Table of Contents

Table of Figures

1. Introduction

Exploration of Mars has been an area of increasing interest for over 2 decades now, and a significant scientific endeavour in recent times. In order for there to be manned missions to Mars, the understanding of the weather patterns on the planet must be developed fully and is crucial for these missions to be successful. The aim is to potentially colonize Mars by establishing sustainable habitats on the planet's surface; and analysing and predicting temperatures and weather patterns would make this easier. This can be achieved by utilising machine learning techniques which can analyse previously read weather data from Mars and create a model which can be used to predict weather phenomena on the planet. This report aims to document the use of some of these machine learning techniques which can predict the temperature of Mars depending on the day, month, and year.

2. Background

Currently, there is an ongoing mission on Mars courtesy of NASA. Curiosity, which is a car-sized Mars rover, was launched on the 26th of November 2011. It landed on Mars on the 5th of August 2012 and has been on the planet ever since. It was sent to Mars with the purpose of accomplishing four main goals:

1. Characterizing the climate of Mars.
2. Studying the geology of Mars.
3. Determining whether life ever existed on Mars.
4. Preparing Mars for potential human exploration. (Howell & Dobrijevic, 2022)

It is able to do this with the use of specialized cameras and instruments that can give the rover information about the environment on Mars. The data that will be used for training the machine learning model is data that has been collected by Curiosity, via the Rover Environmental Monitoring Station (REMS) on-board the rover. (Kannan.K.R, 2020)

3. Methodology and Data

This section will outline the methodology that was employed when training the machine learning model, and the pre-processing that was done on the data so that it was in a good state for the machine learning model to efficiently use.

3.1 Methodology

For the purpose of this project, an LSTM model will be used to analyse and predict weather phenomena on Mars. LSTM stands for long short-term memory networks, which are used in the field of deep learning (Intellipaat, 2023). It is a type of recurrent neural network (RNN).

3.2 Data

The dataset used will be Mars weather dataset that, as aforementioned, has been obtained from Curiosity. The dataset contains 1894 instances and 10 attributes. These 10 attributes are the following:

1. id, which is the identification number of a single transmission.
2. terrestrial_date, which is the date on Earth.
3. sol, which is the date on Mars (e.g. Sol 1 is the 7th of August 2012 on Earth).
4. ls, which is the solar longitude of the Mars-Sun angle.
5. month, which is the Martian month.
6. min_temp, which is the minimum temperature.
7. max_temp, which is the maximum temperature.
8. pressure, which is the atmospheric pressure.
9. wind_speed, which is the average wind speed.
10. atmo_opacity, which is the overall weather condition (e.g. Sunny).

As has been outlined in the Jupyter notebook which holds the model, a lot of these attributes will be disregarded due to the information they hold. As an example, atmo_opacity has the overall weather condition listed as Sunny for every instance in the dataset, making it unnecessary to include when training the model.

Explanations of which attributes were chosen to train the model and why they were chosen are outlined in detail in the Jupyter notebook.

4. Analysis and Discussions

In this section, the data the model has generated will be analysed and the significance of the data will be discussed in detail.

4.1 Analysis

Before the processed data was fed through to the machine learning model, it was plotted onto a graph so that the data could be visualized.

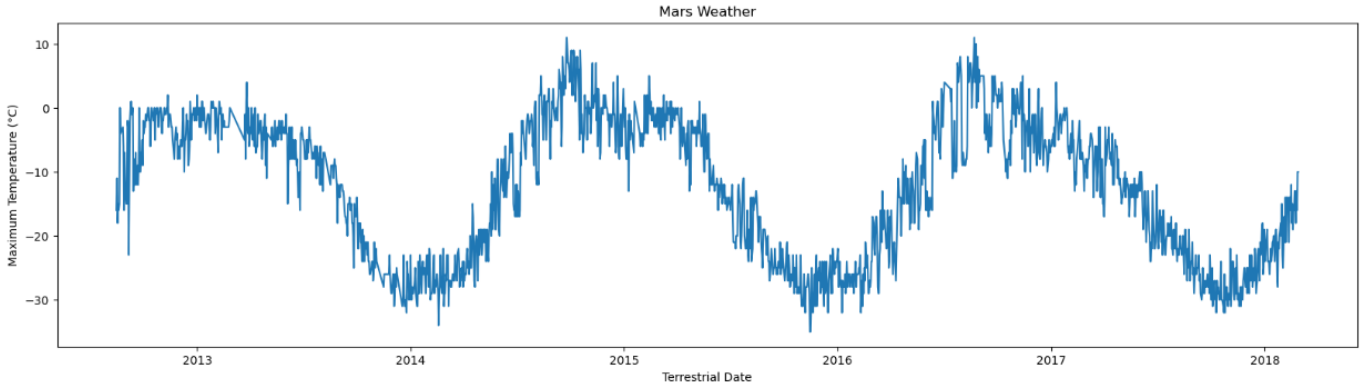

Figure 1: Mars Weather Data Visualized

The graph above shows the maximum temperature of Mars depending on the date. As you can see, the data follows a pattern which is similar to what we see on Earth in the form of seasons. On Earth, it is generally observed that the mid-point of the year (June, July etc.) is when the temperature is at its highest (if you are in the northern hemisphere), while the

end and beginning points of the year (December, January etc.) are when temperatures are of the lowest. Mars seems to experience a pattern that is similar to the one that is observed on Earth but over a longer period of time. In 2013, temperatures are high, while in 2014 temperatures are low, and then in 2015 the temperatures are high again, which is the pattern that the temperature on Mars seems to follow. This is consistent to the fact that a solar year on Mars is 687 Earth days, nearly double the length of a year on Earth.

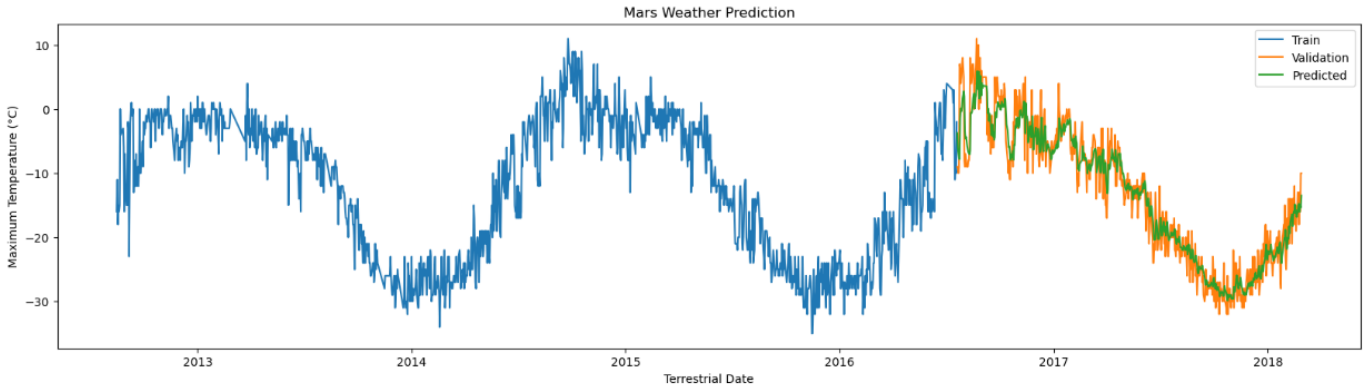

Figure 2: LSTM Model Prediction Data

The figure above shows the values that were predicted by the model. The line in green indicates the data that the model has predicted while the line in orange is the data that was observed and was used for validation to determine the accuracy of the model. As you can see, the model is very accurate in its prediction of the maximum temperature on Mars, as the green line follows the orange line very closely.

4.2 Discussions

The model has recognized the pattern that could be seen when the initial data was plot and has used this pattern accurately when predicting values for the maximum temperature on Mars. It may look like that the model has overfit the predicted values, but this does not seem to be the case if the entire data that the model has been trained on is considered as a whole. The maximum temperature on Mars is generally very sporadic. The model has incorporated this sporadic nature of the data in its predictions which gives the line its characteristic appearance that one may compare to an overfitted line.

5. Conclusions and Future Work

In this section, the conclusions drawn from the model will be outlined and discussed in detail. I will also personally talk about what I have learnt from doing this assignment and any future methods I plan to incorporate into my projects going forward.

5.1 Conclusion

As could be seen in the analysis section, the accuracy of the model can be seen in the graph which had been plotted for that purpose. Therefore, the accuracy of the model is very high with its ability to predict the maximum temperature on Mars depending on the date. In terms of metrics, the model was given an AUC score of 0.9871 after 10 epochs, which demonstrates its ability to discriminate between the different classes at an outstanding level.

```
248/248 [==============================] - 3s 10ms/step - loss: 0.0053 - mean_squared_error: 0.0053 - auc: 0.9806
Epoch 8/10
248/248 [==============================] - 3s 10ms/step - loss: 0.0053 - mean_squared_error: 0.0053 - auc: 0.9647
Epoch 9/10
248/248 [==============================] - 3s 10ms/step - loss: 0.0053 - mean_squared_error: 0.0053 - auc: 0.9842
Epoch 10/10
248/248 [==============================] - 3s 10ms/step - loss: 0.0050 - mean_squared_error: 0.0050 - auc: 0.9871
18/18 [==============================] - 0s 7ms/step
```

Figure 3: Accuracy of the Model

In terms of cross-validation, the initial data had been divided into training and test sets which the model has been trained on and has been used to validate the predictions that the model would make. The model had been tested on the test set and trained on the training set.

```
# Splitting the data into training and testing data, and then scaling the data using a min-max scaler.
data.index = data.terrestrial_date
data.drop("terrestrial_date", axis = 1, inplace = True)

final_data = data.values
train_data = final_data[0:1300, :]
valid_data = final_data[1300:, :]

scaler = MinMaxScaler(feature_range = (0, 1))

scaled_data = scaler.fit_transform(final_data)
x_train, y_train = [], []
for i in range(60, len(train_data)):
    x_train.append(scaled_data[i - 60:i, 0])
    y_train.append(scaled_data[i, 0])

x_train = np.asarray(x_train)
y_train = np.asarray(y_train)
```

Figure 4: Train-Test Split

The metrics that have been discussed demonstrate the feasibility of the model that has been designed for the purpose of predicting weather phenomena on Mars.

5.2 Future Work

In order for the implementation to be improved upon, different features could be analysed, and the attributes of the model could have been adjusted to obtain different results. It may also be argued that the obtained model has overfit the data, due to the nature of the line it has produced, but this may be due to the amount of data as well as the sporadic nature of the data it has been trained on. Therefore, one improvement, should this be deemed a problem, could be to better represent the predicted values that the model has produced, or to adjust the values of the model to produce a more generalized curve (if it is considered to be overfitted). For future projects, more metrics could be used to better determine the performance of the model, but for the purpose of this project the AUC score should suffice.

In terms of what I have personally learned when developing this model and working on this project, I believe I have gained a deeper understanding of neural networks under the blanket term of deep learning. I believe I understand now how the KSTM model works specifically, and I will now be able to implement this model in future machine learning projects going forward. I am also more confident when it comes to using the machine learning libraries in Python, such as keras and matplotlibs, which will be beneficial should I do work with machine learning in the future. This project has been a valuable learning experience.

6. Bibliography and Citations

Howell, E., & Dobrijevic, D. (2022, August 2). *Curiosity rover: The ultimate guide*. Space.com. https://www.space.com/17963-mars-curiosity.html

Kannan.K.R. (2020, July 30). *Mars Weather Data*. Kaggle. https://www.kaggle.com/datasets/imkrkannan/mars-weather-data

NASA. (2019, August 1). *Summary*. NASA. https://mars.nasa.gov/msl/timeline/summary/

Intellipaat (2023, March 4). What is LSTM - introduction to long short term memory. Intellipaat Blog. https://intellipaat.com/blog/what-is-lstm/?US