Drought Prediction Progress Report

Mohamed Abdi

[1] University of Washington Bothell, 18115 Campus Way NE, Bothell, WA 98011
moabdi21@uw.edu

Abstract. This project progress report shows an overview of drought prediction and the changes I have made for this project and the reason behind it. It will cover upcoming project plan of how I will be achieving a drought prediction program and what resources I plan on using to get the best predictive model.

Keywords: Standard Precipitation Index, Palmer drought severity index

1 First Project Report

1.1 Project Progress

My project was original drought prediction for East Africa, but after trying to preprocess the data I have realized the data was very inconsistent & incomplete. I wanted to do the same topic, so I changed the entire location to USA. To narrow down on the location, I decided to focus on five cities in California from 1995 to 2018 (might add more depending on the progress of the project).

Ethopia City Raw Dataset

STATION	NAME	DATE	PRCP	TAVG
ET000063403	GORE, ET	3/28/2002	0.47	70
ET000063403	GORE, ET	4/2/2002		75
ET000063403	GORE, ET	12/26/2002		68
ET000063403	GORE, ET	1/25/2006		72
ET000063403	GORE, ET	2/1/2006		70
ET000063403	GORE, ET	2/5/2006		73
ET000063403	GORE, ET	2/11/2006	0.03	71
ET000063403	GORE, ET	2/12/2006	0.01	76

California City Raw Dataset

STATION	NAME	DATE	PRCP	SNOW	SNWD	TAVG
USW00023259	MONTEREY PENINSUL AIRPORT, CA US	4/1/1998	0.49			51
USW00023259	MONTEREY PENINSUL AIRPORT, CA US	4/2/1998	0.12			54
USW00023259	MONTEREY PENINSUL AIRPORT, CA US	4/3/1998	0.6			52
USW00023259	MONTEREY PENINSUL AIRPORT, CA US	4/4/1998	0.14			52
USW00023259	MONTEREY PENINSUL AIRPORT, CA US	4/5/1998	0			54
USW00023259	MONTEREY PENINSUL AIRPORT, CA US	4/6/1998	0.26			53
USW00023259	MONTEREY PENINSUL AIRPORT, CA US	4/7/1998	0.04			53
USW00023259	MONTEREY PENINSUL AIRPORT, CA US	4/8/1998	0			50

Fig. 1. This figure shows the Raw data set of a city in Ethiopia vs. a city in California.

The figure above show that raw data set for Ethiopia is very inconstant as you see the date of second and third row skips a few days and precipitation is incomplete as it's empty in five slots. The California data is much more organized and will be easier to process and create predictive model with it. Since I had to reevaluate my raw data

set, I took the opportunity to research on creating a drought prediction program, since I'm not a metrologist. I started on a pseudocode of how to preprocess data show in figure 2.

```
'''
Each file has Station name, city, preciptation, temperature
Preprocess Data to have:
    There should be file for each station & year.
    Each file will contain each month with 3 attributes.
    The three atributes is:
        -average month preciptattion
        -Standard Preciptation Index(SPI)
        -Palmer Drought Severity Index(PDSI)
'''
import pandas as pd

raw_data_set = pd.read_csv('rawData/1995-2018Monterey,CA.csv')
#print(targetFile)
arr = raw_data_set

stationDict = {}
#1) iterate throw through raw data csv file
for i, row in raw_data_set.iterrows(): #i=index, j=row
    #2)get raw data for each column data.
    station = row[0]
    location = row[1]
    date = row[2]
    prcp = row[3]
    #3) Get & store a months preciptation.
    #4) If we ecounter a new month. Calulcated the average preciptation,
    #SPI, PDSI, & its class(wet, severe drought, extreme drought, etc). with the all the stored preciptation for that month.
    #5) store calculation for that month.
    #6) repeat step 2 to 4 if you don't encounter a new year.
    #7) When the 12 months averages are recieved for that year.
    #store its own individual file for that year.
    #8) Then repeat step 2 to 7 till it reaches to its last year.
```

Fig. 2. This figure shows pseudocode of how I will be preprocessing the California raw data.

Once the prepossessing stage is complete, I should be able to begin making the drought prediction which will be talked about in the next section.

1.2 Project plan

On the drought prediction I have decided on the type of machine learning to be supervised multiclass classification with the program. In my preprocessing stage I plan on use creating two new attributes, which are Standard Precipitation Index (SPI) and Palmer Drought Severity Index (PDSI) [2].

SPI Values	
2.0+	extremely wet
1.5 to 1.99	very wet
1.0 to 1.49	moderately wet
-.99 to .99	near normal
-1.0 to -1.49	moderately dry
-1.5 to -1.99	severely dry
-2 and less	extremely dry

PDSI Classification			
4.0 or more	extremely wet		
3.0 to 3.99	very wet	-1.0 to -1.99	mild drought
2.0 to 2.99	moderately wet		
1.0 to 1.99	slightly wet	-2.0 to -2.99	moderate drought
0.5 to 0.99	incipient wet spell		
0.49 to -0.49	near normal	-3.0 to -3.99	severe drought
-0.5 to -0.99	incipient dry spell	-4.0 or less	extreme drought

Fig. 3. This figure shows SPI & PDSI values.

Since the project aim is figure if a drought is going to occur or not, I plan on having 5 classes which are normal, mild drought, moderate drought, severe drought, and extreme drought. In my preprocessing data each city will be

having data for every year which will contain it the average for precipitation, SPI, and PDSI for every month. Each month will be classed by the results of SPI & PDSI gives.

In my prediction model I plan on using years 1995-2017 as the training set and 2018 as the test set. Which I then plan on printing 12 scatter plots and maps of California showing each month for how the upcoming year drought will look like. As shown in figure 4 but the map will visualize the surrounding 5 cities I have chosen for California.

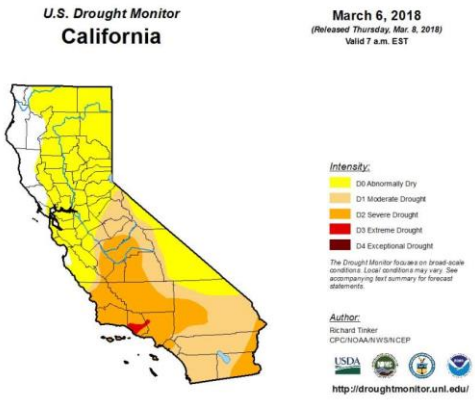

Fig. 4. This figure shows an example how to represent the prediction for an upcoming year.

References

1. Wood, Eric F., et al. “Prospects for Advancing Drought Understanding, Monitoring, and Prediction.” Journal of Hydrometeorology, vol. 16, no. 4, 29 July 2015, pp. 1636–1657., doi:10.1175/jhm-d-14-0164.1. “https://journals.ametsoc.org/doi/full/10.1175/JHM-D-14-0164.1”
2. “Drought Basics.” National Drought Mitigation Center, “drought.unl.edu/ranchplan/DroughtBasics/WeatherandDrought/MeasuringDrought.aspx”
3. Belayneh A. Adamowski J. “Standard Precipitation Index Drought Forecasting Using Neural Networks, Wavelet Neural Networks, and Support Vector Regression”. 18 July 2012