2016.M3.TQF-ML.ForestFiresPrediction Proposal

Yu Lipeng

Project Goal:

Predict whether there will be forest fires using several factors such as month of the year, day of the week, FFMC (Fine Fuel Moisture Code) index, DMC (Duff Moisture Code) index, DC (Drought Code) index, ISI (Initial Spread) index, temperature, relative humidity, wind speed, outside rain and so on.

Factors:

1. X - x-axis spatial coordinate within the Montesinho park map: 1 to 9
2. Y - y-axis spatial coordinate within the Montesinho park map: 2 to 9
3. month - month of the year: "jan" to "dec"
4. day - day of the week: "mon" to "sun"
5. FFMC - FFMC index from the FWI system: 18.7 to 96.20
6. DMC - DMC index from the FWI system: 1.1 to 291.3
7. DC - DC index from the FWI system: 7.9 to 860.6
8. ISI - ISI index from the FWI system: 0.0 to 56.10
9. temp - temperature in Celsius degrees: 2.2 to 33.30
10. RH - relative humidity in %: 15.0 to 100
11. wind - wind speed in km/h: 0.40 to 9.40
12. rain - outside rain in mm/m2 : 0.0 to 6.4
13. area - the burned area of the forest (in ha): 0.00 to 1090.84

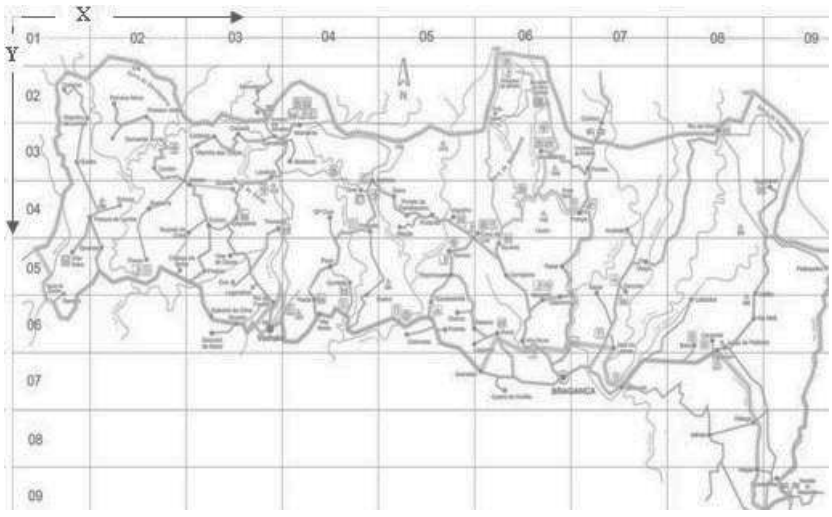

Figure 1　X,Y coordinate graph

Data Sources:

http://archive.ics.uci.edu/ml/datasets/Forest+Fires

Methology:

Mainly use Logistic Regression, SVM, Decision Tree. If possible, I will try random forest and other methods and compare these results.

Expected Results:

Visualize the results. Calculate and compare accuracies of different methods. Then choose the best one.