Seismological data Discrimination using Fuzzy Classification

Krishna G. Vadolia [1]
Department of Computer Engineering
1L.J Institute of Engineering and Technology, Ahmedabad, Gujarat, India

ABSTRACT

Earthquake mainly produces P- and S- wave at the point of occurrence. Mining site that mines the geological ma-terial from earth often makes the quarry blasts on the surface of earth to accelerate the mining process instead of drilling. This mining blast also causes the same wave as the earthquake. It is very difficult to differentiate between two of them as the characteristics of waves produced by these two events are similar. Sometimes mining sites are located near to residential area and discrimination of earthquake and quarry blast data is important to analyze the geological activities to create awareness. Data classification is one of the effective techniques that differentiate the data into a number of classes. In this paper , I calculated several parameters from wave form of an event. Fuzzy classification is good idea to differentiate them. Experimental result shows that Rule- based classification method is not useful to differentiate seismic data into mining blast and earthquake. While Fuzzy Classification Method gives 80-85% accurate result to identify class over noisy data.

Keywords – *seismic data classification, data classification applications, seismic data differentiation, application of fuzzy classification algorithm*.

---

1. INTRODUCTION

Modern digital seismographs may record seismic waves of earthquakes and significant explosions occurring sequentially or simultaneously. These sequential or simultaneous occurring characteristics would be harmful to properly explain the recorded seismic waves and might beget some false conclusions. So it is very meaningful to separate earthquake events and explosion events which may occur sequentially or simultaneously from recorded seismic waves. The separation of simultaneously occurring earthquake and explosion events is beyond the scope of our current researches. And should be investigated in further research by some special signal processing techniques such as independent component analysis [1]. A challenge in seismic monitoring is to uniquely discriminate between natural seismicity and anthropogenic events such as mining blasts. Two basic types of elastic waves are generated from seismic events like earthquake and mining blast namely P- and S- waves. The P- waves move in a compression motion and the S – wave move in shear motion perpendicular to the wave direction. These waves result in shaking of earth surface and may cause damages. It is good to focus on P- wave spectra because they have good signal-to-noise (STN) ratio over much wider bandwidth than the S-wave spectra. Magnitude of lower earthquake and quarry blast is same. Sometimes heavily loaded vehicles also cause high magnitude that can't be discriminated by visualizing their signals.

In this paper I applied fuzzy classification methods by finding the several parameters of an events. We show experimental results of fuzzy classification techniques on seismic data discrimination, and conclude that the rule based technique of classification is less appropriate in such discrimination where having overlapping data problem. While on other hand Fuzzy Classification Method has its advantages against simple partitioning methods and seismic data can be classified with better accuracy of 80-85%.

2. DATA AND METHODS

The research area is selected in Saurashtra-Kutchh that is an active mining region with many types of mines and quarries that are also Gujarat's top producer of hard coal, rock and sands. However, the region also has prevalence of natural seismicity due to active faults in Surendranagar, Bhavnagar, Kutchh, Bhuj, Chotila and Junagadh. These factors make the region an ideal test bed for classification techniques examinated to separate data from earthquakes and mining blasts, including time - of – day analysis, spectral ratios, and complexity analysis. In order to assemble a database containing natural and probable mining-related events, before the data selection step, we first determined the site of major mining and the sites of earthquake sources.

Sayla Mining blasts that take place in Surendranagar, Saurashtra are ripple fire blasts in open pit, sand, rock mine. The geographic coordinates of the mining area was derived by satellite image from Google Earth program (http://earth.google.com) and from the seismological map of Saurashtra from ISR (Institute of Seismological Research), Gandhinagar. The information about blasting times on an average of fourty times monthly and the average amount of explosives varying from 20 to 20.000 kg were obtained.

The statistical distribution of an events with Md≤ 5.0, with a 150-km radius of the mining area between 2008 to 2013 that was recorded by ISR, in daytime and night time was investigated and effectively outlines regions of mining activity (where the dominant percentage of daytime events occur in regions of known mine locations). Figure1 reveals that an unusually high number of daytime events are seen in the Surendranagar mining area that is a likely sign of blasting activity. The number of nightly events is very low compared to the one of daytime events.

Figure 1: Location of mining site in Google map near Sayla (Surendranagar) that produce rock and Sands

In this paper all earthquake and explosion seismic wave records are came from ISR data centre. All the events occur from 2008 to 2013 year, magnitudes are between 0.5 and 5.5. Data formats are SEED. 139 seismic waves in vertical direction are used. Each wave correlates to a different event. Of the 139 events, 80 events are earthquakes; another 59 events are explosions. The basic parameters of some these events are listed in table 1 and table 2.

Table 1: The basic parameters of 10 earthquake events

Date	Time	Magnitude	Lat	Long	Depth	Location
21-01-2016	2:16 PM	1.1	21.184	70.577	27	15 km NNE from Talala, Saurashtra
11-06-2015	3:38 PM	1.6	21.02	70.704	8.4	19 km ESE from Talala, Saurashtra
11-05-2015	6:11 AM	1.8	21.192	70.585	9.9	16 Km NNE from Talala, Saurashtra
11-02-2015	2:42 PM	1.8	21.165	70.457	9.1	14 Km NNW from Talala, Saurashtra
20-09-2015	12:59 AM	1.3	21.191	70.616	6.1	18 km NNE from Talala, Saurashtra
09-11-2015	2:44 PM	2.3	21.148	70.603	6.8	13 km NNE from Talala, Saurashtra
08-08-2015	11:02 AM	1.8	21.169	70.604	6.8	14 Km ENE from Talala, Saurashtra
22-06-2015	10:24 PM	1.5	21.048	70.617	13.1	09 km ESE from Talala, Saurashtra
03-03-2015	4:04 PM	2.3	21.073	70.594	9.7	07 km ENE from Talala, Saurashtra
01-05-2015	7:47 AM	1.6	21.246	70.449	15	23 km NNW from Talala, Saurashtra

Table2: The basic parameters of 10 explosion events

Date	Time	Magnitude	Lat	Long	Depth	Location
19-06-2012	11:15 AM	3.2	20.649	70.302	11.1	51 km SSW of Talala,Saurashtra
14-11-2012	3:03 AM	1.2	21.152	70.568	21.7	12 km NNE of Talala, Saurashtra
17-11-2012	2:01 AM	1.7	21.175	70.649	3.1	17 km ENE Talala, Saurashtra
26-11-2012	6:07 AM	1.7	21.17	70.664	1.7	19 km ENE of Talala,Saurashtra
12-04-2012	3:16 PM	2.1	21.086	70.716	8.4	20 km ENE of Talala,Saurashtra
18-12-2012	6:19 PM	1.8	21.121	70.556	11.8	08 km NNE of Talala, Saurashtra
28-12-2012	6:30 AM	1.8	21.105	70.529	6.5	6 km NNW from Talala, Saurashtra
01-04-2013	2:43 AM	1.7	21.183	70.696	11.1	16 km NNE of Talala,Saurashtra
01-11-2013	6:23 AM	1.4	21.135	70.561	6.1	10 km NNE from Talala, Saurashtra
01-11-2013	11:12 PM	0.9	21.159	70.649	6.1	17 km ENE from Talala, Saurashtra

3. EXPERIMENTAL RESULTS

A. Data Preparation

I extracted numbers of events from original data file. The extraction and conversation needed to findings the parameters to apply as an input of classfication techniques. Seisan is an effective seismic analysis tool that I used to prepare the data from waveforms.

I first, extracted the data base from file. Than merge the same station file into one. Finally, using the same tool, seisan, converted into txt format from ASCII. It gives amplitude of each sensor that measured independently from each other. Applied these amplitude values as input of Matlab graph to generate constant wave form.

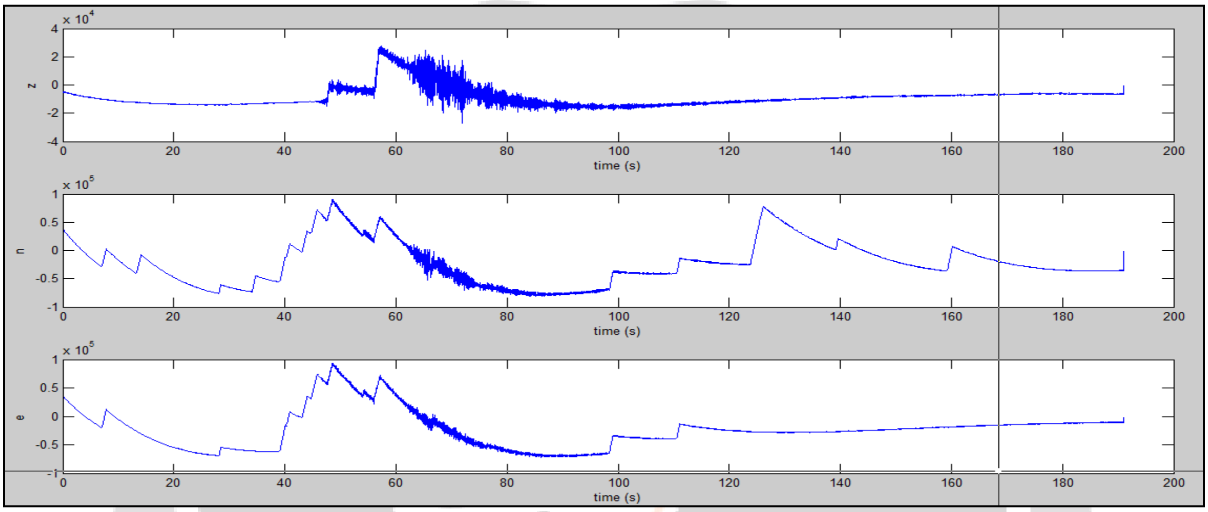

Figure 2 : Generation of wave forms according to the amplitude of event

Only the variation in waveform is important, not all records. So it is advisable to capture and select portions of waveform that clearly indicates the variations in graph. Numbers formula given in Matlab built in function, I calculated several parameters such as starting point of captured waveform, end of line between starting and ending point, spectral ratio, and complexity.

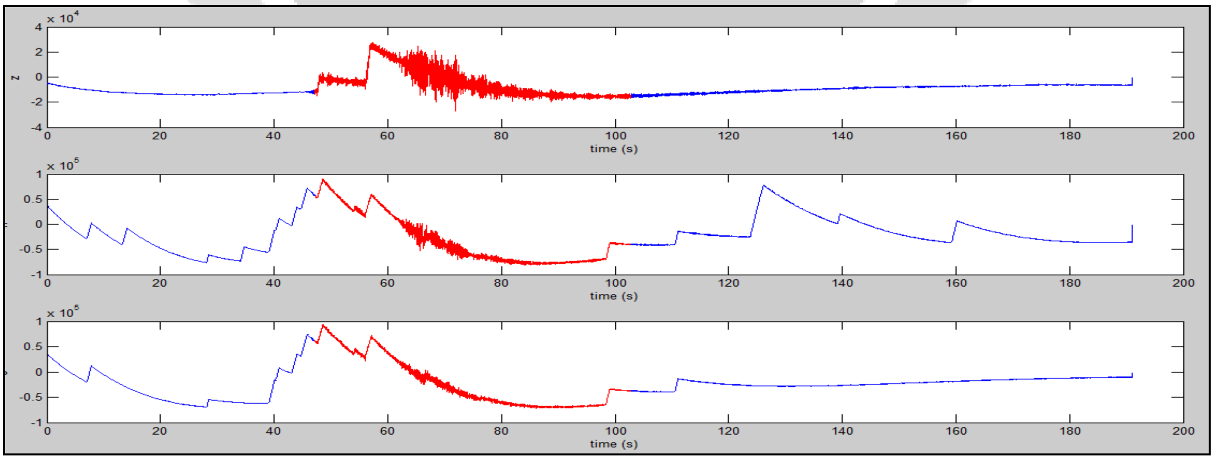

Figure 3: selecting waveform of an event having variation in it

Complexity played key role. The calculations of complexity, spectral ratio and other parameters are as follow:

X = starting of waveform

Y = end point of selection

NumLine = Number of data between X and Y
T0 = starting data point of waveform (same as X)
T2 = ending data point of waveform (same as Y)
T1 = middle point data of selected waveform
Y1 = $\int_{t1}^{t2} s2 \ dt$

Y2 = $\int_{t0}^{t1} s2 \ dt$

Complexity = y1 / y2

Spectral Ratio = $\int_{h1}^{h2}(a) \ dt \ / \ \int_{h0}^{h1}(a) \ dt$

B. K- nearest neighbor classification

K- nearest neighbor classification aims to partition n observations into no. of classes in which each observation belongs to the class with the nearest neighbour, serving as a prototype of the claas. This results is a partitioning of the data space into Voronoi cells. In Matlab I prepared program for K- nearest neighbor classification algorithm. I set the number of iterations are 5, numbers of classes were 3, Matlab gives the functionality to use distance in some vary form like cosine; I used distance, and the input parameters for X axes I gave Complexity and for Y axes I left it empty, so by default it took time as second parameter. The result gave 3 separate classes. By analyzing numbers of records class in K- nearest neighbor class, 4.75 be a threshold figure to differentiate both event nicely.

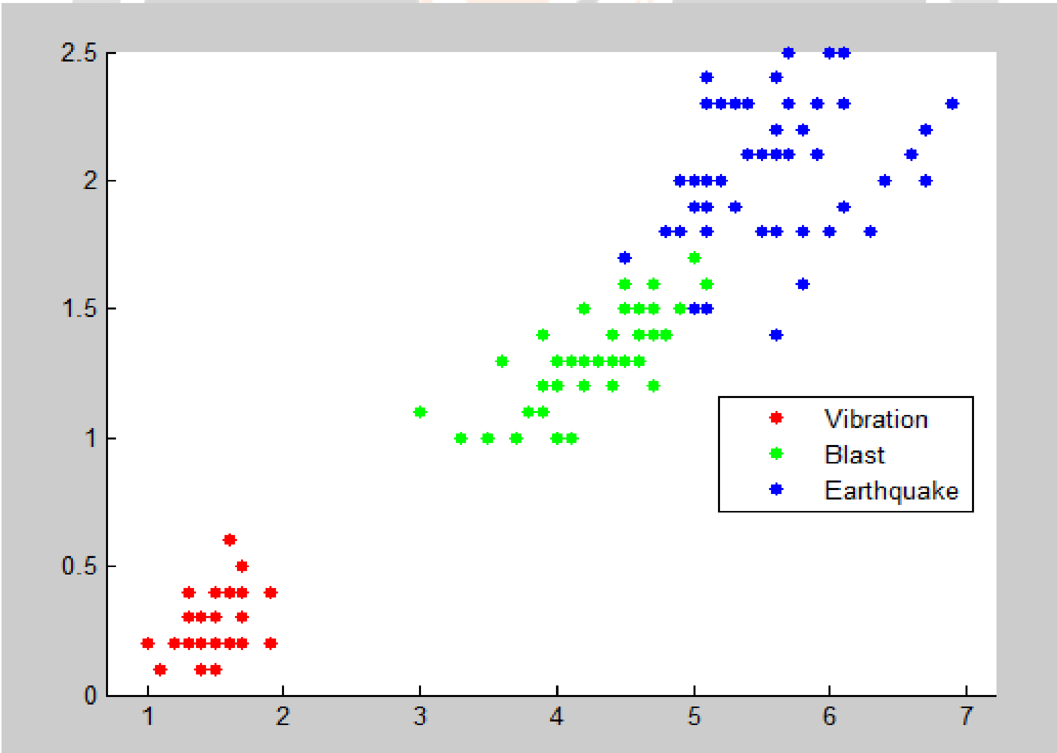

Figure 4 : K- nearest neighbor classification

C. Fuzzy K- nearest neighbor classification

In above method I set only complexity as a parameter of classification. Spectral ratio remains same for any amplitude value, complexity vary accordingly to the amplitude and spectral ratio. I used both as an input to subtractive classification algorithm as shown in figure 5

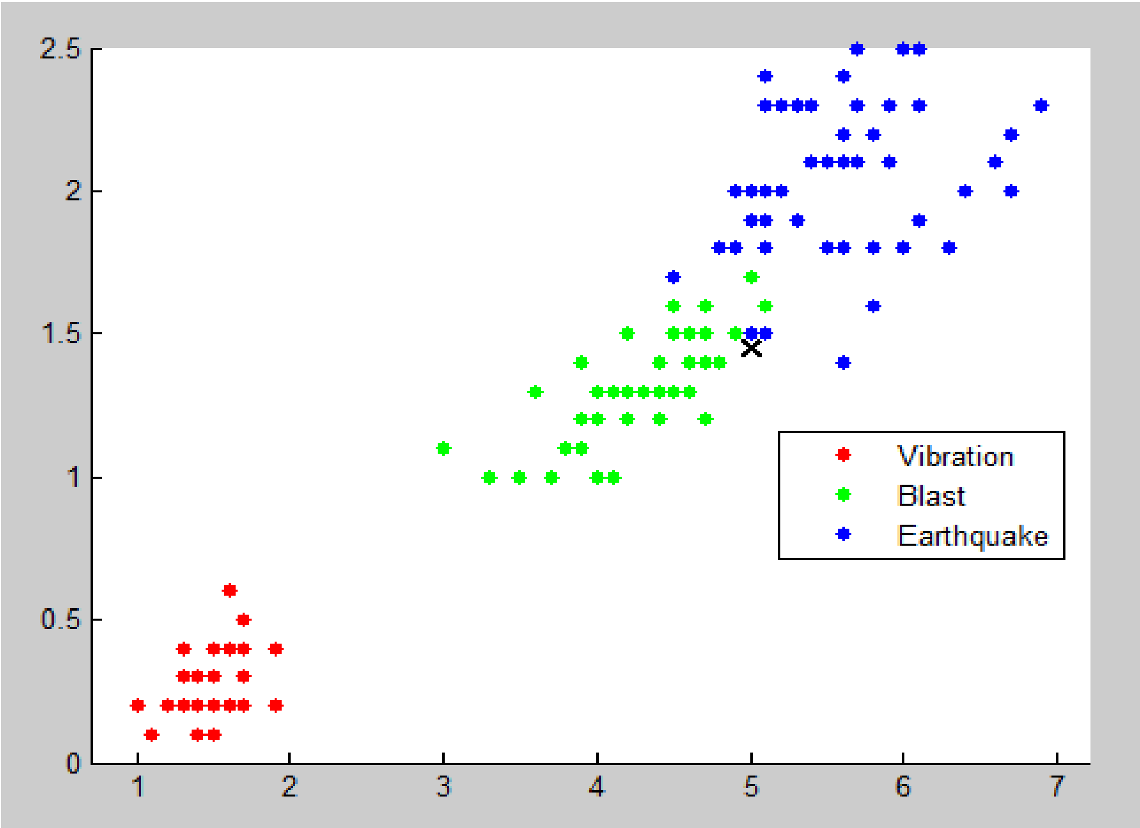

Figure 5 : Class center- k-nearest neighbor classification

As shown in fig. 5, I applied complexity as an input of k-nearest neighbor algorithm in matlab to find out class centers only. And then found nearest neighbor using distance metric. It appears that KNN search has found only the nearest eight neighbors. It is shown in fig. 6

In above method I set only complexity as a parameter of classification. Spectral ratio remains same for any amplitude value, complexity vary accordingly to the amplitude and spectral ratio. I used both as an input to subtractive classification algorithm as shown in figure. As discussed early, spectral ratio remains same for any complexity value. I classified data within two values of frequency domain; 0-3HZ - 3-5HZ and 0-5HZ – 5-10HZ. It gives spectral ratio 3.1250 and 3.3030 respectively. For better discrimination of frequency, I set same frequency throughout whole experimental. Here, centers may vary by adding more number of data into it and reduce the accuracy from near about 75-80% to 60-75%.

D. Nearest Neighbors scale zoom in

Early experimental results shows that Fuzzy classification Method is effective to locatie class over noisy data. I had 139 files having magnitude recorded by ISR. I write the exponential valu of fcm is 1.02 to 1.20. exponentail value decides the distance of object to its center while it iterate. And I took 1 to 100 iteration to classify data. I wrote a code for fcm technique. Thus the input paramenters were exponential value, numbers of iterations and obvisouly complexity and spectral ratio. The KNN classes the data as shown bellow in figure 6.

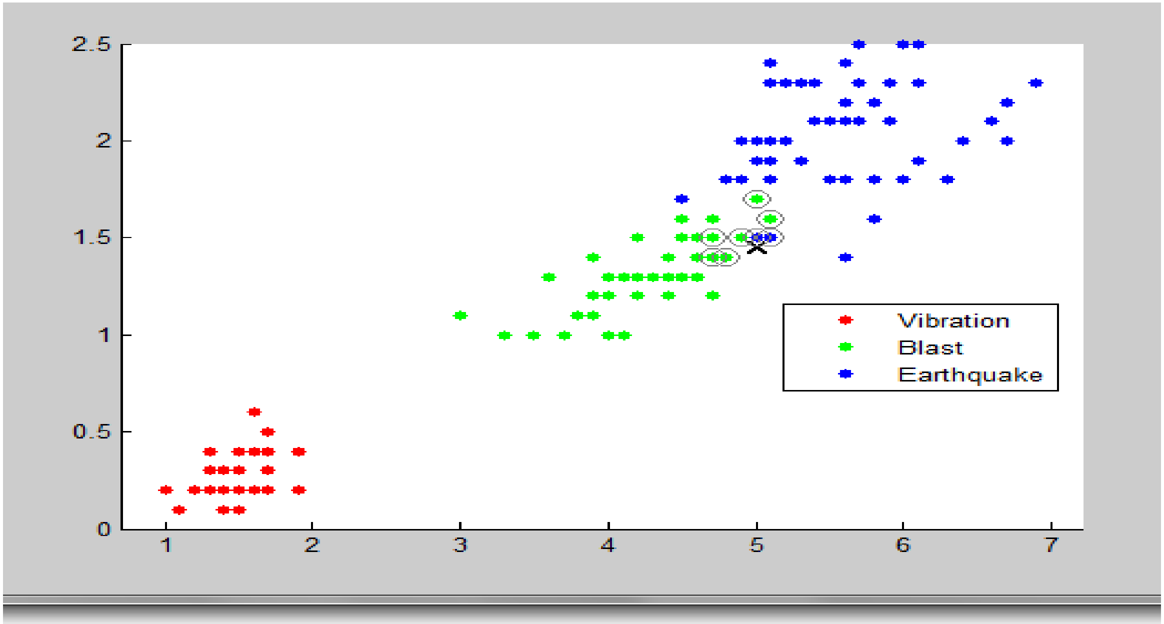

Figure 6 : Nearest Neighbours

E. Nearest Neighbors scale zoom in

There were last chance to apply the same method using GUI. The best part of Interface is that I just worried about only data, I could tune up the number of iterations and exponential values. By performing same methods several time I found the optimal parameters of fcm method to differentiate seismic data. I increased exponential value 1.02 to 1.50 and gave its maximum height of 2.10. Also numbers of iterations are important for locating classes over fuzzy data. Now, Iterations are raised to 0-200 from 0-100. Completing the set up of parameters I just applied both input parameters as I applied in coding, and mentioned the numbers of classes equal to three, surprised results came by this experiment, now accuracy of classify the objects were 80-85%. Which was optimal against all discussed above.

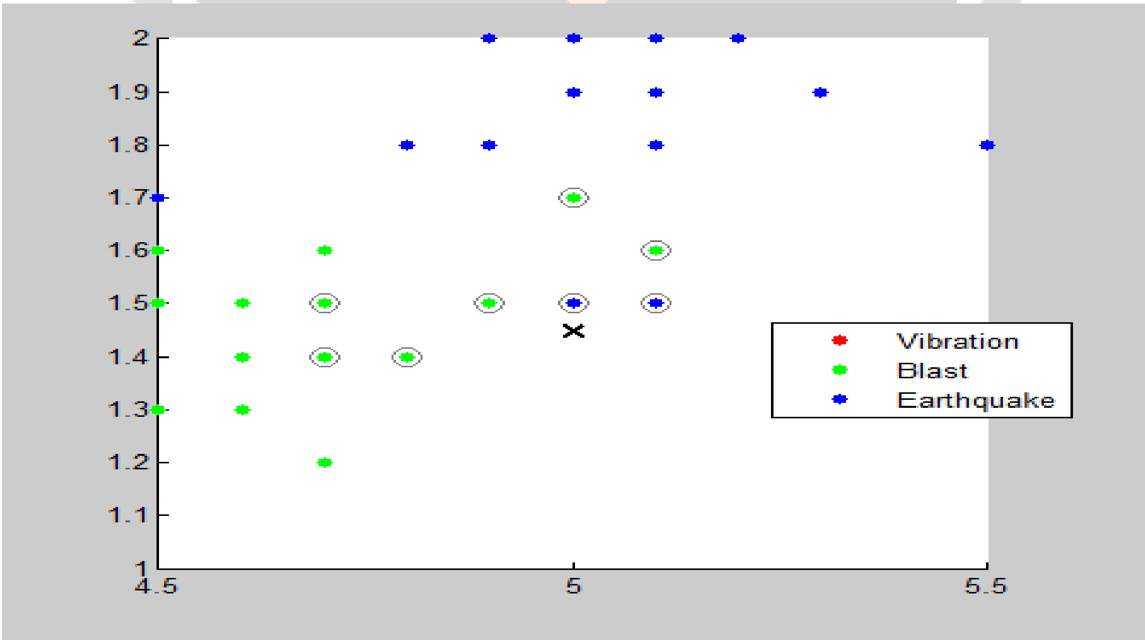

Figure 7 : Nearest Neighbors scale zoom in

F. Nearest neighbors and Group

Identifed the neighbors by drawing a circle around the group of them. Defineed the center and diameter of a circle, based on the location of the new point. From below figure it is clearly indicate that classification again proved is important role to differentiate data.

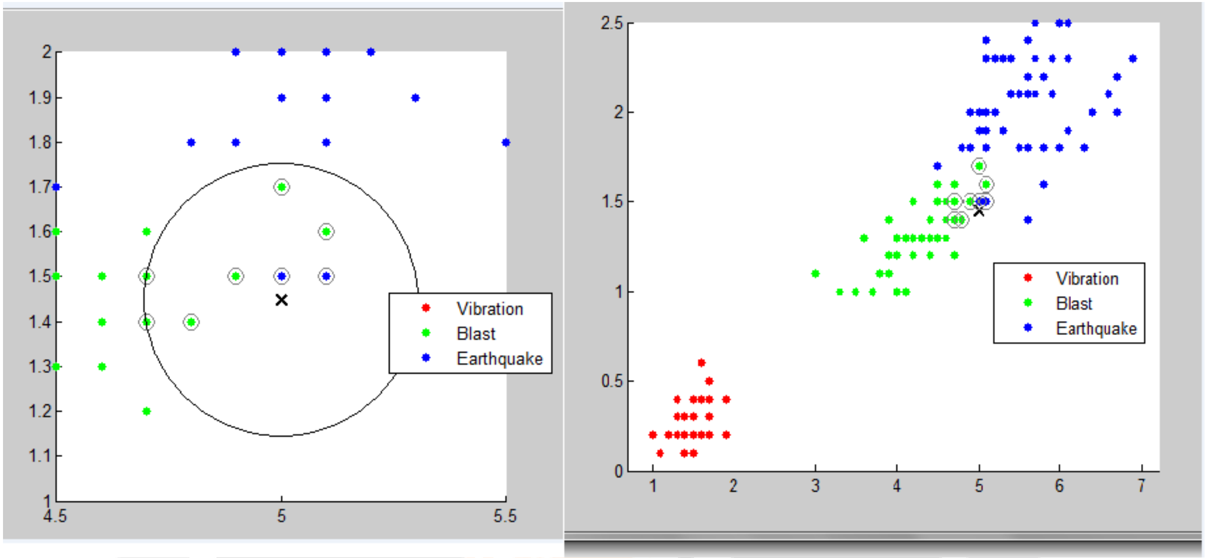

Figure 8 : Nearest neighbors and Group

4. CONCLUSION

There is no doubt that fuzzy classification is best technique to differentiate data without consulting its class label. Experimental results show that Fuzzy classification gives the. Also the accuracy of fuzzy classification is higher, it may vary according to exponent, iteration and noise data. The outlier analysis can be applicable to detect outlier data as noise

5. REFERENCES

[1] Han, Jiawei, and Micheline Kamber. Data Mining. Amsterdam: Elsevier, 2006. Print

[2] El Hassan Ait Laasri, Es-Saïd Akhouayri, Dris Agliz, and Abderrahman Atmani, "Seismic Signal Discrimination between Earthquakes and Quarry Blasts Using Fuzzy Logic Approach" , 5th International Conference, ICISP 2012, Agadir, Morocco, June 28-30, Springer 2012, pp 166-174

[3] Iftikhar U. Sikder , Toshinori Munakata ''Application of rough set and decision tree for characterization of premonitory factors of low seismic activity'' ,Expert System with applications 36(1), Springer,2009, pp. 102–110

[4] Anxu Wu "The Preliminary Review and Application Software on Methods of Seismic Identification", 2010 3rd International Congress on Image and Signal Processing (CISP2010),IEEE,2010 pp. 3721- 3728

[5] Eray Yıldırım , Ali G ulbag , Gunduz Horasanan, Emrah Dogan , "Discrimination of quarry blasts and earthquakes in the vicinity of Istanbul using soft computing techniques" Computers & Geosciences, Volume 37, Issue 9, September 2011, Elsevier,2010, pp. 1209–1217

[6] [Anthony J. Konstantaras, "Classification of Distinct Seismic Regions and Regional Temporal

Modelling of Seismicity in the Vicinity of the Hellenic Seismic Arc", IEEE Journal Of selected topics in Applied Easth Observation and Remote Sensig, VOL. 6, NO. 4, AUGUST ,IEEE, 2012, pp. 1857- 1863

[7] Hisao Ishibuchi, Member, Tomoharu Nakashima, Student Member, IEEE, Tadahiko Murata, Member, IEEE , "Performance Evaluation of Fuzzy Classifier Systems for Multidimensional Pattern Classification Problems" , IEEE Transactions On Systems, Man, And Cybernetics—Part B: CYBERNETICS, VOL. 29, NO. 5, OCTOBER 1999, IEEE, 1999, pp. 601- 618

[8] Kekovalı K*, Kalafat D and Deniz P, "Spectral discrimination between mining blasts and natural earthquakes: Application to the vicinity of Tunçbilek mining area, Western Turkey" International Journal of the Physical Sciences Vol. 7(35), pp. 5339-5352, 13 September, 2012, pp 5339-5352