Fuzzy Soft Set for Rock Igneous Clasification

Iwan Tri Riyadi Yanto [1]
[1] Department of Information Systems
Universitas Ahmad Dahlan
Yogyakarta, Indonesia, 55161
yanto.itr@is.uad.ac.id

Edi Sutoyo[2]
[2] Department of Information Systems,
Telkom University,
Bandung, West Java, Indonesia, 40257
edisutoyo@telkomuniversity.ac.id

Ani Apriani [3], Okki Verdiansyah [4]
[3] Sekolah Tinggi Teknologi Nasional
Sleman, Yogyakarta, Indonesia 55281
aniapriani@sttnas.ac.id, okki.verdiansyah@sttnas.ac.id

Abstract—**Rock classification is one of the fundamental tasks in geological studies. This process normally requires a human expert to examine a sample the rocks. In this research, we employ machine learning algorithm, called Fuzzy Soft Set Classifier (FSSC) to classify igneous rock which based on their chemical composition. This algorithm is hybridization of soft set theory and fuzzy for classifying numerical data. The results showed that the Fuzzy Soft Set Classifier is capable of precise classification of igneous rocks and achieved satisfactory result in terms of accuracy, precision and recall, respectively.**

Keywords—soft set; fuzzy soft set; classification; igneous rocks;

I. INTRODUCTION

One of the fundamental branch in geology is the study of igneous rocks [1]. Igneous rocks are one of the three major groups of rocks along with metamorphic rocks and sedimentary rocks [2]. Igneous rock is formed from solidified molten material [1]. Although the deposit of igneous rocks in some areas is not abundant, all rocks on the surface of the earth should have igneous process in their past history. Therefore, the study of igneous rocks is important to understand the composition of the earth interior [1].

Igneous rocks are not homogenous even within or between the rocks association, there might be difference on minerals and rocks composition. This diversity of igneous rocks sometimes related to the time and place where the rocks are formed [3]. Therefore, the elements composition of igneous rocks in different places might be different because of the different origin. The diversity is expressed by the varieties of mineral and chemical composition of igneous rocks. The chemical analyses of rocks are expressed as weight percent of oxides (wt %) for major elements (SiO_2, TiO_2, Al_2O_3, FeO, Fe_2O_3, MnO, MgO, CaO, Na_2O, K_2O, and P_2O_5) and parts per million (ppm) for trace elements [1].

Igneous rocks can be classified based on their mineralogical or chemical composition, which belong in quantitative classification. Based on the mineral composition, igneous rocks are classified into felsic or silicic rocks, intermediate rocks, mafic rocks and ultramafic rocks categories [2]. While based on the chemical composition, igneous rocks are classified into acid rocks, intermediate rocks, basic rocks, and ultrabasic rocks categories [2]. The classification task of igneous rocks will become challenging because of the diversity composition of igneous rocks.

Some researchers have developed the igneous rock classification method. Peacock proposed the extension of two-fold classification on igneous rock series [4]. The previous two-fold classification categorized rock series as alkalic or sub-alkalic was extended into four-fold division as alkalic, alkali-calcic, calc-alkalic, and calcic because some rock series cannot be properly classified as alkalic or sub-alkalic. The four-fold classification used both chemical and mineralogical basis.

Le Bas *et al.* proposed the standard classification method of igneous rocks based on their chemical composition, called Total Alkali-Silica (TAS) diagram [5], [6], which shown in Fig. 1. From Fig. 1, igneous rocks are classified as acid rocks if have more than 63 wt % of SiO_2, classified as intermediate rocks if have 52 wt % - 63 wt % of SiO_2, classified as basic rocks if have 45 wt % - 52 wt % of SiO_2, and classified as ultrabasic rocks if have less than 45 wt % of SiO_2 [5], [6]. Despite of its simplicity, the TAS diagram classification method only uses SiO_2 as classification parameter, whereas some chemical elements also found inside the igneous rocks that worth to be considered.

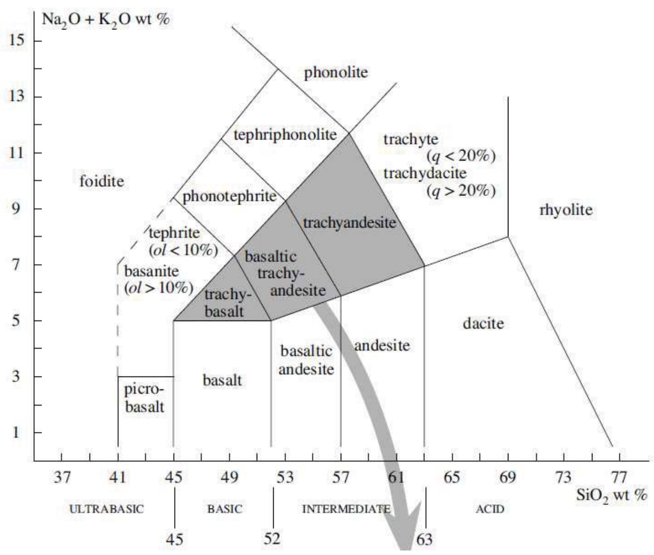

Fig. 1. Total Alkali-Silica (TAS) diagram [5], [6]

978-1-5386-5280-0/18/$31.00 ©2018 IEEE

Encinas used cluster analysis method on large dataset of igneous rock chemical analyses [7]. The purpose of the study was to create procedure to divide a large dataset of rock chemical analyses into homogenous groups regardless their quantitative characteristic. The procedure was done by (a) normalizing the variables' variances (b) using Principal Component Analysis (PCA) to reduce the variables and (c) applying cluster analysis to the moderate number of groups containing large numbers of samples. The result was six differentiated groups with discriminant functions to assign new samples to the groups.

This research proposes fuzzy based classification of igneous rocks based on their chemical analyses because chemical composition represents fundamental characteristic of igneous rocks and it will become a quantitative classification. In engineering geology, fuzzy based method has been used by some researchers to classify rock masses [8], [9], to classify rock facies [10], or to classify rocks strength [11].

The rest of this paper is organized as follow: Section II presents the proposed igneous rocks classification method. Section III presents the results and discussion. Finally, the conclusion of this work is described in Section IV.

II. FUZZY SOFT SET

In this section, the essential notions of soft set theory and fuzzy soft set theory is described. Let U be a non-empty universe of objects, E is a set of parameters in relation in objects of U, $P(U)$ is the power set of U and $A \subseteq E$. The soft set is defined as parameterized family of subsets of the universe U [12], [13], [14]. A soft set of $\underline{\mathbf{U}}$ also can be defined as a *pair* (F, E) which mapping of $F : E \rightarrow P(U)$. For $\varepsilon \in E$, $F(\varepsilon)$ may be considered as the set of ε-elements of the soft set (F, E) or as the set of ε-approximate elements of the soft set, instead of a (crisp) set. Meanwhile, in Fuzzy Soft Set Theory, $\underline{P}(U)$ denotes the power of set of all fuzzy subsets of U and $A \subseteq E$. Then, A pair $(\underline{F}, E)$ is called a fuzzy soft set over U, where $\underline{F}$ is mapping given by $\underline{F} : A \rightarrow \underline{P}(U)$.

In the other words, fuzzy subsets in the universe U are used as substitutes for the crisp subsets of U. Hence, it is easy to see that every standard soft sets may be considered as fuzzy soft sets. Generally speaking, $\underline{F}(\varepsilon)$ is a fuzzy subset in U and. It is called the fuzzy approximate value set of parameter ε.

It is well-known that the idea of fuzzy sets provides a convenient tool to represent the concepts uncertainty by using partial membership. In the definition of soft fuzzy sets, fuzzy subsets are used as substitutes for crisp subsets. Therefore, each soft set can be regarded as a fuzzy soft set. Moreover, by analogy with the soft set, one can easily see that every fuzzy soft set can be seen as a fuzzy information system and is represented by a data table with entries belonging to the unit interval [0,1]. For illustration, we consider the following example.

Example 1 (See [15])**.** Let given a fuzzy soft set (F, E) that describes attractiveness of the shirts with respect to the given parameters, which are going to buy. $U = \{x_1, x_2, x_3, x_4, x_5\}$ that is the set of all shirts under consideration. Let $\underline{P}(U)$ is the collection of all fuzzy subsets of U. And, let $E = \{e_1, e_2, e_3, e_4\}$ means representing the parameter, e.g. colorful, bright, cheap, warm, respectively.

Let

$\underline{F}(e_1) = \{x_1/0.5, x_2/0.9, x_3/0.0, x_4/0.0, x_5/0.0\}$

$\underline{F}(e_2) = \{x_1/1.0, x_2/0.8, x_3/0.7, x_4/0.0, x_5/0.0\}$

$\underline{F}(e_3) = \{x_1/0.0, x_2/0.0, x_3/0.0, x_4/0.6, x_5/0.0\}$

$\underline{F}(e_4) = \{x_1/0.0, x_2/1.0, x_3/0.0, x_4/0.0, x_5/0.0\}$

Then the family $E(e_i)$ where $i = \{1,2,3,4\}$ of $\underline{P}(U)$. The tabular representation of fuzzy soft set $(\underline{F}, E)$ is shown in Table I.

TABLE I. REPRESENTATION OF FUZZY SOFT SET

(U, E)	e_1	e_2	e_3	e_4
x_1	0.5	1.0	0	0
x_2	0.9	0.8	0	1.0
x_3	1	0.7	0	0
x_4	1	0	0.6	0
x_5	0	0	0	0.3

In the next section, the idea of soft set-based classification is presented.

III. CLASSIFICATION BASED ON SOFT SET THEORY

Soft set classifier which learns by computing the average value of every parameter (attribute or feature) from all objects or instant with the same class label, to build a model of the soft set model with the universe comprising of all class label has been proposed Mushrif *et al.* [16]. The algorithm is divided into two phases, e.g. training phase and classification phase. The complete algorithm is as shown in algorithm 1.

Algorithm 1. Soft Set Classifier (SSC)

Training phase

a. Given N samples obtained from the texture w, decompose each sample with wavelet transform
b. Compute the L_1 norm of each channel of the wavelet decomposition and obtain a feature vector E_{wi}, for $i = 1,2,\cdots,N$
c. Calculate the cluster center vector E_w using equation given below

$$E_w = \frac{1}{N}\sum_{i=1}^{N} E_{wi}$$

d. Repeat the process for all W classes.
e. Obtain a soft set (F, E) which is basically a $W \times D$ table cluster centers in which an element of the table is g_{wd},

$w = 1,2,\cdots,W$, $d = 1,2,\cdots,D$ and a row gw is a cluster center vector for class w having D features.

Classification phase

a. Decompose an unknown texture with the wavelet transform
b. Compute the L_1 norm of each channel of the wavelet decomposition and obtain a feature vector E_f
c. Obtain a soft set (F,E) in which an element g_{wd}, $w = 1,2,\cdots,W$ and $d = 1,2,\cdots,D$ is calculated using as follows.

$$p_{wd} = 1 - \frac{|g_{wd} - E_{fd}|}{\max_w [g_{wd}]}$$

d. Compute a comparison table of soft set (F,E)
e. Compute the score vector s.
f. Assign the unknown class data to class w if $w = \arg[\max_{w=1}^{W} \{S\}]$

However, the high complexity is still the main issue in the phase of classification. Therefore, Handaga *et al.* [17] proposed Fuzzy Soft Set Classifier (FSSC) as an algorithm for classifying numerical data which is a modification of the SSC algorithm. To classify general numerical data features, he replaced second step in both phases of train and classification of SSC by taking fuzzy number, so that all parameters have a value in an interval [0,1]. The complete algorithm is as follows.

Algorithm 2. Fuzzy Soft Set Classifier (FSSC)

Pre-processing phase

a. Feature fuzzification to obtain a feature vector E_{wi}, for $i = 1,2,\cdots,N$ for all data, training and testing dataset.

Training phase

b. Given N samples obtained from the data class w
c. Calculate the cluster vector E_w using equation below.

$$E_w = \frac{1}{N}\sum_{i=1}^{N} E_{wi}$$

d. Obtain a fuzzy soft set model for class w, where $(\underline{F}_w, E)$, is a cluster center vector for class w having D features.
e. Repeat steps (b), (c), and (d) for all W classes.

Classification phase

f. Get an unknown class data
g. Obtain a fuzzy soft set model for unknown class data, $(\underline{G},E)$
h. Compute similarity between $(\underline{G},E)$ and $(\underline{F}_w, E)$ for each w using equation below.

$$S(F_\rho, G_\delta) = M_i(\underline{F},\underline{G}) = 1 - \frac{\sum_{j=1}^{n} |\underline{F}_{ij} - \underline{G}_{ij}|}{\sum_{j=1}^{n} |\underline{F}_{ij} + \underline{G}_{ij}|}$$

i. Assign the unknown data to class w if similarity it reaches maximum

$$w = \arg[\max_{w=1}^{W} S(\underline{G}, \underline{F}_w)]$$

IV. Rock Clasification Using FSSC

To evaluate the performance of this classification algorithm, the research used geochemical of igneous rock dataset. This real-world dataset contains 11 features, namely Silicon dioxide (SiO_2), Titanium dioxide (TiO_2), Aluminium oxide (Al_2O_3), Iron(II) oxide+Iron(III) oxide (FeO+Fe_2O3), Manganese(II) oxide (MnO), Magnesium oxide (MgO), Calcium oxide (CaO), Sodium oxide (Na_2O), Potassium oxide (K_2O), Phosphorus pentoxide (P_2O_5), and Class Label, respectively. This dataset was collected from on Mount Wungkal, Godean, Yogyakarta, Indonesia and some bencmarks dataset from Pet_DB.

The objective of this research is to classify igneous rocks based on their chemical analyses because chemical composition represents fundamental characteristic of igneous rocks and it will become a quantitative classification. there are 4 target output or class label, namely Andesite, Basalt, Basaltic andesite, Dacite. The 3 instances of the used dataset are shown in the Table II.

Fuzzification can be done by dividing each attribute values with the largest value of each attributes. Afterwards, the dataset is split into two datasets, one used for training and the other used for testing. The split of the dataset is done randomly selected in each experiment. The experiments are performed 9 times, with 9 different percentages of training and testing dataset for each experiment. Composition comparison of training and testing datasets are as follows, 60% training and 10% testing, 60% training and 20% testing, 60% training and 30% testing, 60% training and 40% testing, 70% training and 30% testing, 70% training and 20% testing, 70% training and 10% testing, 80% training and 20% testing, and 80% training and 20% testing, respectively.

Table II. Sample of Igneous Rock Dataset

(U,E)	SiO_2	TiO_2	Al_2O_3	Fe_2O_3T	MnO	MgO	CaO	Na_2O	K_2O	P_2O_5	Class Label
x_1	57.10	0.61	17.32	7.05	0.14	6.26	6.97	3.39	0.79	0.37	Andesite
x_2	47.67	0.88	17.64	11.53	0.19	6.48	11.33	1.83	1.26	0.24	Basalt
x_3	54.38	0.63	20.67	7.00	0.12	4.53	8.58	2.96	0.80	0.33	Basaltic andesite
x_3	64.48	0.10	15.77	4.93	0.19	5.54	3.02	3.73	1.57	0.67	Dacite

In order to test the proposed algorithm, the experiment is developed using MATLAB version 7.14.0.334 (R2012a). The algorithm is implemented on a processor Intel Core i3-3217U CPU @ 1.80Ghz, with total main memory 8G of RAM and the operating system is Windows 10.

Experiments are carried out on algorithm Fuzzy Soft Set Classifier by [17] which focuses on calculating accuracy, precision, recall, specificity, and MEAN_TIME. Accuracy is calculated using total Overall Classifier Accuracy (OCA) and *F* measure (micro average and macro average). The experiment result is summarized as in Table III. The results show that classification of fuzzy soft set for rock igneous have good performance. It can be seen that the technique is rising up to 1, 0.9979, 0.9854 in average in terms of accuracy, Precision and specificity, respectively. Moreover, the time response is quite of 0.00059 second.

TABLE III. THE EXPERIMENT RESULTS OF FUZZY SOFT SET CLASSIFIER

Training (%)	Testing (%)	Accuracy	Precision	Recall	Mean Time (second)
60	**10**	1	1.0000	1.0000	0.0043
60	**20**	1	1.0000	1.0000	0.0039
60	**30**	1	1.0000	0.9667	0.0025
60	**40**	1	1.0000	1.0000	0.0075
70	**30**	1	1.0000	1.0000	0.0085
70	**20**	1	1.0000	1.0000	0.0086
70	**10**	1	1.0000	0.9667	0.0060
80	**20**	1	0.9833	0.9500	0.0056
80	**10**	1	1.0000	1.0000	0.0043
Average		1	0.9979	0.9854	0.0059

V. CONCLUSION

This paper presented the usage of Fuzzy Soft Set Classifier (FSSC) for classifying igneous rock which focuses on their chemical composition. The real-world dataset was used to examine the FSSC. The experiments were carried out 9 times, with 9 different percentages of training and testing dataset. The results showed that the Fuzzy Soft Set Classifier is capable of precise classification of igneous rocks and achieved satisfactory result in terms of accuracy, precision and recall, respectively.

REFERENCES

[1] K. G. Cox, *The interpretation of igneous rocks*. Springer Science & Business Media, 2013.

[2] J. Schön, “Rocks—Their Classification and General Properties,” *Handbook of Petroleum Exploration and Production*, vol. 8, pp. 1–16, 2011.

[3] G. M. BROWN, “The problem of the diversity of igneous rocks,” *Evolution of the Igneous Rocks: Fiftieth Anniversary Perspectives*, p. 1, 2015.

[4] M. A. Peacock, “Classification of igneous rock series,” *The Journal of Geology*, vol. 39, no. 1, pp. 54–67, 1931.

[5] M. J. Le Bas, R. Le Maitre, A. Streckeisen, B. Zanettin, and others, “A chemical classification of volcanic rocks based on the total alkali-silica diagram,” *Journal of petrology*, vol. 27, no. 3, pp. 745–750, 1986.

[6] R. W. Le Maitre, A. Streckeisen, B. Zanettin, M. Le Bas, B. Bonin, and P. Bateman, *Igneous rocks: a classification and glossary of terms: recommendations of the International Union of Geological Sciences Subcommission on the Systematics of Igneous Rocks*. Cambridge University Press, 2005.

[7] L. H. Encinas, “Partitioning large data sets: Use of statistical methods applied to a set of Russian igneous-rock chemical analyses,” *Computers & Geosciences*, vol. 20, no. 10, pp. 1405–1414, 1994.

[8] A. Aydin, “Fuzzy set approaches to classification of rock masses,” *Engineering Geology*, vol. 74, no. 3, pp. 227–245, 2004.

[9] H. Jalalifar, S. Mojedifar, and A. Sahebi, “Prediction of rock mass rating using fuzzy logic and multi-variable RMR regression model,” *International Journal of Mining Science and Technology*, vol. 24, no. 2, pp. 237–244, 2014.

[10] M. K. Dubois, G. C. Bohling, and S. Chakrabarti, “Comparison of four approaches to a rock facies classification problem,” *Computers & Geosciences*, vol. 33, no. 5, pp. 599–617, 2007.

[11] C. Ozturk and E. Nasuf, “Strength classification of rock material based on textural properties,” *Tunnelling and Underground Space Technology*, vol. 37, pp. 45–54, 2013.

[12] D. Molodtsov, “Soft set theory—first results,” *Computers & Mathematics with Applications*, vol. 37, no. 4–5, pp. 19–31, 1999.

[13] P. Maji, R. Biswas, and A. Roy, “Soft set theory,” *Computers & Mathematics with Applications*, vol. 45, no. 4–5, pp. 555–562, 2003.

[14] P. K. Maji, R. Biswas, and A. Roy, “Fuzzy soft sets,” 2001.

[15] P. Majumdar and S. K. Samanta, “Generalised fuzzy soft sets,” *Computers & Mathematics with Applications*, vol. 59, no. 4, pp. 1425–1432, 2010.

[16] M. M. Mushrif, S. Sengupta, and A. K. Ray, “Texture classification using a novel, soft-set theory based classification

algorithm," in *Asian Conference on Computer Vision*, 2006, pp. 246–254.

[17] B. Handaga, T. Herawan, and M. M. Deris, "FSSC: an algorithm for classifying numerical data using fuzzy soft set theory," *International Journal of Fuzzy System Applications (IJFSA)*, vol. 2, no. 4, pp. 29–46, 2012.