CHAPTER 5

DESIGN AND DEVELOPMENT OF FUZZY LOGIC CLASSIFIER (FLC) FOR SENTIMENT ANALYSIS

5.1 INTRODUCTION

Fuzzy logic is used to describe the fuzziness of the situation under consideration. In real life most of the situations are described on a sliding scale, for example, “the engine is running really hot”, “hybrid cars are not very fast”, etc. It is almost impossible to distinguish the members of a class from non-members that are on the border line of threshold when such sliding scale is used for description. Hence fuzzy logic becomes a handy choice for the opinion miners to understand the opinion expressed in a customer review document.

Lotfi Zadeh (1965), the proponent of fuzzy logic defined it as follows:

“*Fuzzy logic is determined as a set of mathematical principles for knowledge representation based on degrees of membership rather than on crisp membership of classical binary logic*.”

Fuzzy logic based systems are handy during the real life situations where the decision to be taken are based on multiple criteria with complex interlink among them. It is very true for a sentiment analysis process, in which the system must be able to understand the sentiment expressed by a customer in a review based on the statements about various features of the product or service. For example, in a movie review, the reviewer may praise the director for the usage of technology and blame on the acting and story. Deciding on the overall sentiment as positive or negative depends on the opinion words or phrases used by the reviewer for each of the features. When the number of features is more, the uncertainty in the decision making gets added and hence the decision making becomes tough. In such situations, fuzzy logic can be effectively used.

Fuzzy logic provides a systematic approach to formalize the approximate reasoning (Lotfi Zadeh, 1994). Fuzzy logic can handle the vagueness of human linguistics and thinking mathematically (Tanaka, 1996).

5.2 FUZZY CLASSIFIER

A classifier that uses fuzzy sets either during its training or during its operation (Kuncheva L.I, 2000) is called as a fuzzy classifier. In this thesis, the fuzzy classifier proposed for the sentiment detection is named as Fuzzy Logic Classifier (FLC). Fuzzy classifiers are handy choice for the statistical pattern recognition task when there is insufficient information. When the user needs some information like the desirability of an option or severity of a disease etc. in addition to the class label of the object, fuzzy classifier can do a good job.

Since there is no rigorous theory for specifying the conditions for optimality of a fuzzy classifier, there is no theoretical methodology available to design fuzzy classifier for every instance (Kuncheva L.I, 2000).

5.2.1 Fuzzy Set

Fuzzy logic allows the partial membership of an element ‘x’ unlike the crisp set that can allow only the full membership of the element or no membership at all. If the universe of discourse is denoted by ‘X’ and its elements are denoted by ‘x’, then the crisp set A of X is defined as $f_A(x)$, known as the characteristic function of A.

$$f_A : X \rightarrow \{0,1\}$$

where $f_A(x) = 1$ if x is totally in A

$f_A(x) = 0$ if x is not in A

The fuzzy set A of universe X is defined by function $\mu_A(x)$, known as the membership function of set A.

$$\mu_A : X \rightarrow \{0,1\}$$

where $\mu_A(x) = 1$ if x is totally in A

$\mu_A(x) = 0$ if x is not in A

$0 < \mu_A(x) < 1$ if x is partly in A

The fuzzy set gives the flexibility of expressing the continuum of possible choices. The membership function of an element x represents the degree to which it is an element of set A. This value is called as membership value, which varies between 0 and 1.

5.2.2 Linguistic Variables and Hedges

Fuzzy set theory is based on the idea of linguistic variables. A linguistic variable is a fuzzy variable. For example, in the statement ‘Paul is short’, *Paul* is the linguistic variable that takes the linguistic value *short*. The universe of discourse of a variable is represented by the range of possible values of that variable.

Every linguistic variable carries the idea of fuzzy set qualifiers called *hedges* (Michael Negnevitsky 2005). The shape of the fuzzy set can be modified using the hedges. The adverbs such as very, quite, somewhat, extremely etc. are some of the hedges commonly used to modify the shape of the fuzzy set. Table 5.1 shows the procedure followed for the calculation of the linguistic value (Fuzzy weight) for each linguistic variable considered in the sentiment analysis process.

Table 5.1 Procedure for calculation of fuzzy weight of linguistic variables

Pattern of the opinion phrase	Fuzzy weight calculation
"Adverb" followed by "Adjective" (Very good, extremely bad) (Fuzzy weight of "good" is taken as 0.6 and fuzzy weight of bad is taken as 0.4 based on the expert's opinion)	(fuzzy weight of adjective)$^{1/3}$ if fuzzy weight of adjective ≥ 0.5 (e.g.) Fuzzy weight of very good = $(0.6)^{1/3} = 0.8434$ (fuzzy weight of adjective)^{3} if fuzzy weight of adjective < 0.5 (e.g.) Fuzzy weight of very bad = $(0.4)^{3} = 0.064$
"Not" followed by "adjective"	(1- fuzzy weight of adjective) (e.g.) Fuzzy weight of not good = 1- 0.6 = 0.4

5.3 GENERAL STEPS INVOLVED IN THE FUZZY INFERENCE

Step 1: Fuzzification of the crisp inputs and determining the degree to which each of these inputs belong to the appropriate fuzzy sets. This step involves assigning the membership values for each crisp value and deciding on the membership function to decide the degree of association. Triangular, Trapezoidal, and Gaussian membership functions are commonly used and choosing the membership function is based on the number of parameters that define the linguistic variable.

Step 2: The fuzzified inputs are used to find the single number that represents the result of the antecedent evaluation.

Step 3: Aggregation of the output of all the rules and combine them into a single fuzzy set.

Step 4: The aggregate output of the fuzzy set calculated in the previous step is transformed into a crisp number. The centroid technique is used to convert the fuzzy value into a crisp number.

Note: For the sentiment analysis process, since an intuitive method has been used for the fuzzy score calculation and final classification, the details of the membership functions like Triangular, Trapezoidal, Gaussian and the defuzzification methods like centroid technique etc. are not discussed in this thesis.

5.4 SENTIMENT CLASSIFICATION USING FUZZY LOGIC

5.4.1 Introduction

In a document level sentiment classification process, the overall sentiment expressed is decided based on the opinion words and phrases used by the reviewer. For example, in a movie review, the reviewer may praise the director for the usage of technology for bringing out the special effects and blame on the acting and story. The reviewer would have expressed a very high satisfaction on the special effects and a moderate dissatisfaction over the story and a strong dissatisfaction over the acting. Accordingly, the choice of words differs. For example, the reviewer would have used “excellent” to describe the very high satisfaction and “boring” to describe the dissatisfaction over the story.

Generally, adjectives and adverbs are the most important parts of a sentence to decide on the opinion. Table 5.1 shows the procedure followed to assign the fuzzy weights for the linguistic variables.

5.4.2 Linguistic Variables for the Sentiment Analysis Process

A special method has been adopted to decide on the linguistic variables for the sentiment classification at document level. While reading a document, though lot of words and phrases are used by the reviewer, the human mind consolidates them under of the following categories viz. Good, Very good, Excellent, Recommended, Bad, Very bad, Disgusting, Never recommended. Though the list can include some more categories, the research in this thesis has been carried out using only these eight categories. These categories have been denoted as *Representative Terms (RT)*.

5.4.3 RTDM Creation

A program written using the PERL programming language, has been used to capture the opinion words and phrases and assign them to any one of the categories mentioned above. The detailed procedure for converting a given text document into a feature vector consisting of RT was discussed in chapter 3. The entries in the RTDM represent the frequency of the corresponding RT in that document. For the purpose of description, the RTDM created for LDS403 is used.

5.4.4 Creation of RTDM with Fuzzy Scores (RTDM-FS)

For the fuzzy logic based classifier, the entries in the RTDM is expressed in the form of a fuzzy number. The procedure for calculating the fuzzy weight of each RT in the RTDM is explained in the subsequent sections. Table 5.2 shows the linguistic values that have been assigned to each of the RT. The procedure given in the Table 5.1 has been followed for few categories and the remaining are decided based on the expert judgments. For example, the fuzzy weight (FW) for “good” is decided as 0.6 using the experts’ opinion and the fuzzy weight for “very good” is calculated using the procedure given in Table 5.1. The fuzzy weight of other categories like Excellent, Recommended, Disgusting and Never recommended was decided based on experts’ judgment only. Table 5.2 shows the fuzzy weight assigned for the various linguistic variables.

Table 5.2 Fuzzy weight of linguistic variables

RT/Linguistic variable	Fuzzy weight (Linguistic value)
Good (Not bad)	0.6.
Very good	0.8434
Excellent	0.9
Recommended	0.95
Bad (Not good)	0.4
Very bad	0.064
Disgusting	0.05
Never recommended	0.04

5.4.5 Fuzzy Score (FS) Calculation

Step 1: Calculate the total frequency (TF) of RT for the given document from the RTDM. To accomplish this, all the entries found in the RTDM for a corresponding document are added. For example, consider the review and the corresponding feature vector using RT shown in Figure 5.1, the TF for this review is 14.

"I wasted enough time actually WATCHING this chore of a movie, I don't want to waste more writing a big review. Not once did I so much as crack a smile. ALL the jokes were boring, forced and lacked any kind of wit. I kept saying, "wheres the punchline?" Almost every single character was an obnoxious stereotype and all the situations were clichéd and half the time there wasn't even any kind of solution. Things just happened to get to the next scene. For the life of me I can't understand how this got as many good reviews as it did. If you like clunky acting and poorly composed film making Fat Girls is the movie for you."

Feature Vector of the above review using RT

File name	Good	Very good	Excellent	Recommended	Bad	Very bad	Disgusting	Never recommended
neg04_2.txt	2	0	0	0	3	5	2	2

Total frequency (TF) = 2+3+5+2+2 = 14

Frequency of "good" (F_{good}) = 2; F_{Bad} = 3; $F_{very\ bad}$ = 5; $F_{disgusting}$ = 2; $F_{Never\ recommended}$ = 2

Number of positive RT captured = 1 (Among the four positive RT, only for 'Good" the entries are found). This is referred as 'CP' in the Equation (5.3)

Number of negative RT captured = 4 (For all the four RT, the entries are found). This is referred as 'CN' in the Equation (5.4)

Figure 5.1 A sample review and its feature vector

Step 2: Calculate the fuzzy score (FS) of each RT based on its linguistic value and its frequency. For example, in the example shown in the Figure 5.1, the RT "good" appears 2 times. The linguistic value of "good" is 0.6 as shown in Table 5.2. The fuzzy weight for the RT "good" can be calculated using the Equation (5.1).

$$FS_{good} = FW_{good} + ((1\text{- } FW_{good})/ TF) * F_{good} \quad (5.1)$$

$$FS_{good} = 0.6 + ((1\text{-}0.6)/14) * 2 = 0.6571$$

By substituting the fuzzy weight and frequency of the corresponding RT, the FS can be calculated as follows:

$FS_{Very\ good} = 0$; $FS_{Excellent} = 0$; $FS_{Recommended} = 0$;

For the review document shown in Figure 5.1, the FS for the other positive RT is zero as they have zero entries in the RTDM. For the RT "bad", the FS is calculated using Equation (5.2)

$$FS_{bad} = FW_{bad} – ((FW_{bad})/ TF) * F_{bad} \quad (5.2)$$

$$FS_{bad} = 0.4 – (0.4/14) * 3 = 0.3143$$

By substituting the fuzzy weight and frequency of the corresponding FS, the FS can be calculated as follows:

$FS_{Very\ bad} = 0.0411$; $FS_{Disgusting} = 0.0429$; $FS_{Never\ recommended} = 0.0343$

Step 3: Calculate the average FS for the positive RT and average RT for negative RT using Equation (5.3) and Equation (5.4) respectively.

$$FS_{Ave}(Pos.) = (FS_{good} + FS_{Very\ good} + FS_{Excellent} + FS_{Recommended})/CP \quad (5.3)$$

$$FS_{Ave}(Pos.) = (0.6571+0+0+0)/1 = 0.6571$$

$$FS_{Ave}(Neg.) = (FS_{bad} + FS_{Very\ bad} + FS_{Dissgusting} + FS_{Never\ Recommended})/CN \quad (5.4)$$

$$FS_{Ave}(Neg.) = (0.3143+ 0.0411+ 0.0429 +0.0343) /4 = 0.1081$$

The $FS_{Ave}(Neg.)$ is subtracted from 0.5 to avoid the misclassification in case of documents where, $FS_{Ave}(Pos.)$ is close to 1 and

FS_{Ave}(Neg.) is close to zero. For example, if a document has entries in all the positive RT and in the negative side, the entry is found only in "Never recommended", then the FS_{Ave}(Neg.) will be very small as the linguistic value assigned to it is 0.04. But when a reviewer comments a product into "never recommended" category, it should be classified as negative. In order to achieve this, in general the FS_{Ave}(Neg.) is subtracted from 0.5 and named as "Effective FS_{Ave}(Neg.)".

Effective FS_{Ave}(Neg.) = 0.5 - 0.1081 = 0.39185

Step 4: Calculate the FS of the review using the Equation (5.5).

$$FS_{Review} = FS_{Ave}(Pos.) - \text{Effective } FS_{Ave}(Neg.) \qquad (5.5)$$

IF $FS_{Review} > 0.5$, the document is classified as positive

If $FS_{Review} \leq 0.5$, the document is classified as negative

For the review shown in Figure 5.1

FS_{Review} = 0.6571- 0.39185 = 0.2653

Hence the review is classified as "Negative". Table 5.3 shows the RTDM-FS for the review shown in Figure 5.1.

Table 5.3 Feature vector with fuzzy score for the review shown in Figure 5.1

File name	Good	Very good	Excellent	Recommended	Bad	Very bad	Disgusting	Never recommended	FS_{Review}
neg04_2.txt	0.6571	0	0	0	0.3143	0.0411	0.0429	0.0343	0.2653

Table 5.4 shows the abridged RTDM-FS for LDS403 with the final fuzzy score for the review.

Table 5.4 Abridged RTDM-FS for LDS403 with fuzzy score and assigned class label

Document No.	RT									
	Good	Very Good	Excellent	Recommended	Bad	Very Bad	Disgusting	Never Recommended	Fuzzy Score (for the review)	Class Label
1	0.66	0.00	0.00	0.00	0.31	0.04	0.04	0.03	0.27	n
2	0.62	0.87	0.91	0.00	0.33	0.05	0.04	0.04	0.42	n
3	0.64	0.86	0.92	0.95	0.27	0.06	0.04	0.00	0.47	n
4	0.70	0.86	0.00	0.00	0.23	0.06	0.05	0.04	0.37	n
5	0.76	0.00	0.91	0.00	0.37	0.06	0.04	0.03	0.46	n
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.
399	0.00	0.89	0.94	0.00	0.28	0.00	0.00	0.00	0.69	p
400	0.75	0.86	0.91	0.95	0.31	0.06	0.05	0.00	0.51	p
401	0.00	0.00	0.95	0.00	0.20	0.00	0.00	0.00	0.65	p
402	0.00	0.86	0.95	0.95	0.36	0.00	0.05	0.04	0.57	p
403	0.71	0.86	0.93	0.96	0.35	0.00	0.00	0.04	0.54	p

5.5 RESULTS AND DISCUSSION

Table 5.5, Table 5.6, Table 5.7, Table 5.8, Table 5.9 and Table 5.10 show the confusion matrix of FLC for Camera reviews, Cell phone reviews, LDS403, LDS2000, LDS11000 and LDS25000 respectively. Table 5.11 shows the performance of FLC on various datasets. The precision score of FLC for camera and Cell phone is 0.95 and 0.94 respectively. This is due to the less *false positive* i.e. the number of negative documents classified as positive. However the recall is badly affected as the *false negative* i.e. the number of positive documents classified as negative is more. But the overall accuracy is considerably good for all the sizes of dataset.

Table 5.5 Confusion matrix for camera reviews (FLC)

	Classified Negative	**Classified Positive**
Actual Negative	104	18
Actual Positive	94	322

Table 5.6 Confusion matrix for cell phone reviews (FLC)

	Classified Negative	**Classified Positive**
Actual Negative	103	19
Actual Positive	109	307

Table 5.7 Confusion matrix for LDS403 (FLC)

	Classified Negative	**Classified Positive**
Actual Negative	183	19
Actual Positive	64	137

Table 5.8 Confusion matrix for LDS2000 (FLC)

	Classified Negative	Classified Positive
Actual Negative	836	164
Actual Positive	448	552

Table 5.9 Confusion matrix for LDS11000 (FLC)

	Classified Negative	Classified Positive
Actual Negative	4545	955
Actual Positive	2736	2764

Table 5.10 Confusion matrix for LDS25000 (FLC)

	Classified Negative	Classified Positive
Actual Negative	10418	2082
Actual Positive	6462	6038

Table 5.11 Classification performance of FLC on various datasets

S.No.	Dataset	P	R	F-Measure	Accuracy
1	Camera	0.95	0.77	0.85	0.791
2	Cell phone	0.94	0.74	0.83	0.762
3	LDS403	0.88	0.68	0.77	0.794
4	LDS2000	0.77	0.55	0.64	0.694
5	LDS11000	0.74	0.50	0.6	0.664
6	LDS25000	0.74	0.48	0.59	0.658

Performance of the FLC depends on how accurately the RTDM of the given dataset is created. RTDM-FS used by the FLC is created from

RTDM and hence if RTDM is created properly, the classification performance would considerably increase.

For the Camera and Cell phone the total negative documents chosen for analysis were 122 and total positive documents were 416. The opinion words and phrases from 200 reviews were captured and hence the higher performance compared to the performance for LDS. LDS contains 25000 reviews, of which only 400 reviews were read for capturing the opinion words and phrases.

FLC can be used to find out the sentiment expressed on each of the features of the product/service. An appropriate feature selection algorithm along with fuzzy logic, the sentiment detection at feature level can be carried out. In this research, since the objective is to detect the sentiment expressed by a reviewer at the document level, the feature level sentiment classification is not discussed.