Fuzzy Rule Based Systems for Gender Classification from Blog Data

Han Liu
School of Computer Science and Informatics
Cardiff University
Cardiff, United Kingdom
liuh48@cardiff.ac.uk

Mihaela Cocea
School of Computing
University of Portsmouth
Portsmouth, United Kingdom
mihaela.cocea@port.ac.uk

***Abstract*—Gender classification is a popular machine learning task, which has been undertaken in various domains, e.g. business intelligence, access control and cyber security. In the context of information granulation, gender related information can be divided into three types, namely, biological information, vision based information and social network based information. In traditional machine learning, gender identification has been typically treated as a discriminative classification task, i.e. it is aimed at learning a classifier that discriminates between male and female. In this paper, we argue that it is not always appropriate to identify gender in the way of discriminative classification, especially when considering the case that both male and female people are of high diversity and thus individuals of different genders could have high similarity to each other in terms of their characteristics. In order to address the above issue, we propose the use of a fuzzy method for generative classification of gender. In particular, we focus on gender classification based on social network information. We conduct an experiment study by using a blog data set, and compare the fuzzy method with C4.5, Naive Bayes and Support Vector Machine in terms of classification performance. The results show that the fuzzy method outperforms the other methods and is also capable of capturing the diversity of both male and female people and dealing with the fuzziness in terms of gender identification.**

***Keywords*—data mining; machine learning; fuzzy rule based systems; text classification; gender classification**

I. Introduction

Gender classification is aimed at identifying the gender of a person, i.e. it is to determine person is male or female. In practice, gender classification can sever various applications, such as business intelligence [1], access control [2] and security checks [3].

Gender classification can be done through manual classification by using expert knowledge or automatic classification by learning classifiers from real data. As the the size of data has been increased rapidly, machine learning techniques have been used increasingly more popularly for gender identification. Some popular learning approaches reported in [4] include support vector machine (SVM) [5], k nearest neighbour (KNN) [6] and Gaussian mixture models (GMM) [7] .

From granular computing perspectives, gender related information can be decomposed into biological information (e.g. EEG and DNA), vision based information (e.g. height and hair length) and social network information (e.g. Facebook posts, tweets and blogs). From this point of view, gender classification can be achieved by learning classifiers from data obtained from different sources, such as biological data, images and text, i.e. different types of features are extracted for training gender classifiers.

In traditional machine learning, gender identification has been typically treated as a discriminative classification task, due to the case that the two classes (male and female) are considered to be mutually exclusive. However, in reality, both male and female people are of high diversity and can be divided into many different groups, which indicates that individuals of different genders may have high similarity to each other in terms of their characteristics. It is also possible that a person of one gender intentionally shows characteristics of the other gender, e.g. they may try to disguise themselves.

On the basis of the above argumentation, it is not always appropriate to treat gender identification as a discriminative classification task. Instead, generative classification is considered to be more suitable for such classification tasks. In this paper, we propose the use of fuzzy methods for generative classification, and focus the study on gender identification based on features extracted from online text.

The rest of this paper is organized as follows: Section II outlines related work on gender classification, feature extraction from text and fuzzy classification. In section III, we present a fuzzy approach in terms of its key features, and justify why fuzzy approaches would be more suitable for gender identification from textual data. In Section IV, we report an experimental study conducted by using a blog gender data set and discuss the results to highlight the strengths of the fuzzy approach. In Section V, the contributions of this paper will be highlighted and some further directions for this research area will be suggested towards achieving further advances.

II. Related Work

In this section, we provide an overview of gender classification in the context of text mining and review popular methods of feature extraction in the area of text classification. Also, we provide the background and recent developments of fuzzy text classification.

978-1-5386-4362-4/18/$31.00

A. Review of Feature Extraction Methods

Feature extraction from textual data consists of four stages: enrichment, pre-processing, transformation and vectoring [8].

The enrichment stage aims at assigning semantic information by recognizing and tagging named entities (NE) in order to support term filtering in the later stages. Popular taggers include Part of Speech (POS) Tagger, Dictionary Tagger and Abner Tagger. A detailed description of text enrichment is provided in [8].

Pre-processing aims to filter those irrelevant terms such as stop words, numbers, punctuation and N-Char words (each word containing less than n characters) [8]. Also, all words are converted from their upper cases to lower ones and the endings of these words are removed through word stemming [8].

Transformation aims to transform textual data into structured data, i.e. feature extraction, in order to adopt machine learning algorithms directly for training classifiers. In this context, the bag-of-words (BOW) method has been a very popular one used for feature extraction [9], [10] by transforming each word into a feature.

In the above context, each word, which is used as a feature, is viewed as a single-word term. However, a term can also consists of multiple words (i.e. multi-word term), when N-Gram (an extension of BOW) is used for transforming each combination of n sequential into a feature.

Following the above transformation, the frequency of each term is calculated in order to enable feature selection by filtering those less frequently occurring terms. In this way, the data dimensionality can be decreased greatly leading to more efficient processing in later stages.

In the vectoring stage, each feature in an instance is assigned either a binary or numerical value. For a binary feature, the Boolean value indicates the presence or absence of the corresponding term in a specific instance. For a numerical feature, the frequency of the corresponding term is used as the value of the feature in the learning stage.

For BOW, there are four types of frequency, namely, term absolute frequency, term relevant frequency, inverse document frequency and inverse class frequency. For N-Gram, there are three types of frequency, namely, corpus frequency, document frequency and sentence frequency. More details on these types of frequency can be found in [8].

B. Overview of Gender Classification

In the context of text mining, gender classification is typically achieved by learning classifiers from text posted on social networks, such as emails, Facebook posts, tweets and blogs.

In [4], Lin et al listed several representative studies on gender classification through using daily information posted via social network platforms, and reported that the classification accuracy was relatively low, in comparison with using features extracted from biological data and images.

In particular, an investigation was conducted in [11] for mining gender attribution of authorship from emails. In this investigation, SVM was used to learn classifiers from manually extracted features of content-free emails, e.g. style markers, structural characteristics, and gender-preferential language features, and the classification accuracy was about 70% [4].

Another study was conducted in [12] by using a real-life blog data set. In this study, an ensemble feature selection approach was proposed, and SVM and Naive Bayes (NB) were used together for learning classifiers, which led to the classification accuracy of 88.56%.

Overall, gender classification through using social network based information is generally more difficult than using other sources of information. As reported in [4], the number of features extracted from social network data is very high and the number of instances is also massive, which could lead to high computational complexity and affect the learning performance due to the presence of more irrelevant features. Also, by its nature, text is characterized by fuzziness, imprecision and uncertainty, which leads to further difficulty in identifying gender from social network based information.

C. Background of Fuzzy Text Classification

In the area of text classification, a review of fuzzy approaches for natural language processing (NLP) was made in [13] in 2012, which highlighted that there was a very low percentage of papers relating to fuzzy approaches over all the papers published in the NLP area and that there were very few NLP related application papers published in the area of fuzzy systems. Following the publication of the above review paper, a number of fuzzy approaches have been proposed for various applications, since fuzzy approaches are more suitable to deal with the ambiguity and fuzziness of text.

A fuzzy approach was developed in [14] for classification of companies based on fuzzy fingerprint text. The classification results showed that the fuzzy approach outperformed the commonly used non-fuzzy approaches. Another fuzzy approach was used in [15] for automatically building a corpus for comparison of text similarity. The results reported in [15] showed that the fuzzy metrics had a higher correlation with human ratings in comparison with the traditional metrics. An unsupervised fuzzy approach was used in [16] for classification of Twitter users according to their gender.

On the other hand, a fuzzy rule based approach was proposed in [17] for addressing the model complexity issue, and the experimental results showed that the fuzzy approach led to a reduction in computational complexity, while maintaining a similar classification performance, when comparing with other non-fuzzy approaches popularly used for text classification. Based on this work, the fuzzy approach was investigated further in [18] for discussing how the membership degree values can be used for more refined outputs, which could reflect different intensities of sentiment.

III. Fuzzy Rule Based Classification

In this section, we provide theoretical preliminaries relating to fuzzy logic and illustrate how a fuzzy rule based system is used for classifying unseen instances. Also, we justify why fuzzy methods are more suitable for gender classification than those popularly used non-fuzzy approaches.

A. Theoretical Preliminaries

Fuzzy logic is a generalization of deterministic logic. In this context, a fuzzy truth value ranges from 0 to 1 whereas the two values (0 and 1) represent the special cases that form deterministic logic. The theory of fuzzy logic is mainly adopted in the contexts of fuzzy sets and fuzzy rule based systems.

Each element e_i in a fuzzy set S has a membership degree $f_S(e_i)$, where $f_S(e_i) \in [0, 1]$ and $1 \leq i \leq n$. In other words, a fuzzy set employs a soft boundary determining the membership or non-membership of each element to the set.

The main operation of a fuzzy rule based system is to transform each numerical attribute into a number (n) of linguistic attributes for induction of a set of fuzzy rules. In particular, each linguistic attribute transformed from a continuous attribute is essentially a fuzzy set defined with a membership function that maps the crisp value of the numerical attribute into a value of membership degree (the value of the linguistic attribute).

Membership functions could be of different shapes, such as trapezoidal, triangular and rectangular membership functions. Generally speaking, a trapezoidal membership function is viewed as a generalization of the ones of the triangular or rectangular shape. In fact, in order to define a membership function, the essence is at the estimation of four parameters (a, b, c, d), as illustrated in the equation below and in Fig. 1.

$$f_T(x) = \begin{cases} 0, & \text{when } x \leq a \text{ or } x \geq d; \\ (x-a)/(b-a), & \text{when } a < x < b; \\ 1, & \text{when } b \leq x \leq c; \\ (d-x)/(d-c), & \text{when } c < x < d; \end{cases}$$

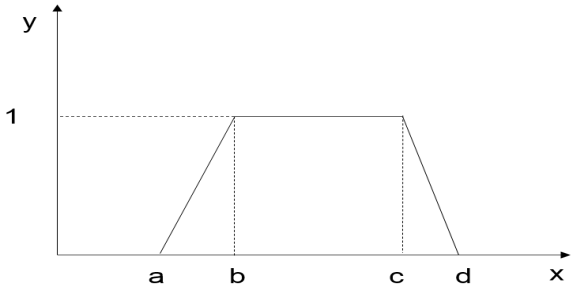

Fig. 1. Trapezoid fuzzy membership function [17]

As shown in Fig 1, if b=c, then the membership function would be shaped as triangular. Similarly, if a=b and c=d, then the membership function would be shaped as rectangular.

In practice, the parameters of a membership function can be estimated according to expert knowledge [19] or through learning statistically from data [20], [21].

B. Procedure

In the classification stage, a fuzzy rule based system involves five main operations: fuzzification, application, implication, aggregation and defuzzification. The whole procedure is illustrated by using the following fuzzy rules as an example:

- Rule 1: if x_1 is *Short* and x_2 is *Cold* then *class=No*;
- Rule 2: if x_1 is *Short* and x_2 is *Warm* then *class=No*;
- Rule 3: if x_1 is *Short* and x_2 is *Hot* then *class=Yes*;
- Rule 4: if x_1 is *Middle* and x_2 is *Cold* then *class=No*;
- Rule 5: if x_1 is *Middle* and x_2 is *Warm* then *class=Yes*;
- Rule 6: if x_1 is *Middle* and x_2 is *Hot* then *class=No*;
- Rule 7: if x_1 is *Long* and x_2 is *Cold* then *class=Yes*;
- Rule 8: if x_1 is *Long* and x_2 is *Warm* then *class=No*;
- Rule 9: if x_1 is *Long* and x_2 is *Hot* then *class=No*;

Each of the two attributes x_1 and x_2 is transformed into three linguistic ones. The corresponding membership functions are illustrated in Fig. 2 and Fig. 3, respectively.

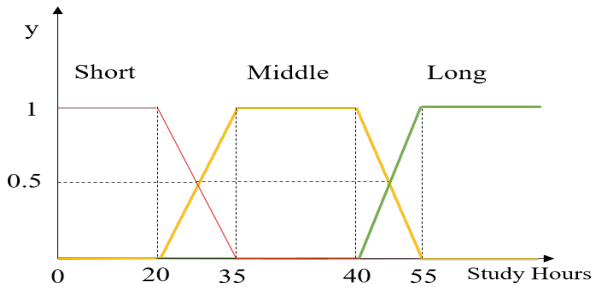

Fig. 2. Fuzzy membership functions for study hours

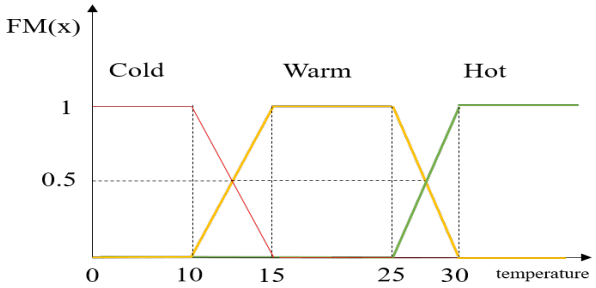

Fig. 3. Fuzzy membership functions for temperature

According to Fig. 2 and Fig. 3, if $x_1 = 45$ and $x_2 = 28$, then the following operations would be done:

Fuzzification:
Rule 1: $f_{Short}(45) = 0$, $f_{Cold}(28) = 0$;
Rule 2: $f_{Short}(45) = 0$, $f_{Warm}(28) = 0.4$;
Rule 3: $f_{Short}(45) = 0$, $f_{Hot}(28) = 0.6$;
Rule 4: $f_{Middle}(45) = 0.67$, $f_{Cold}(28) = 0$;
Rule 5: $f_{Middle}(45) = 0.67$, $f_{Warm}(28) = 0.4$;
Rule 6: $f_{Middle}(45) = 0.67$, $f_{Hot}(28) = 0.6$;
Rule 7: $f_{Long}(45) = 0.33$, $f_{Cold}(28) = 0$;
Rule 8: $f_{Long}(45) = 0.33$, $f_{Warm}(28) = 0.4$;
Rule 9: $f_{Long}(45) = 0.33$, $f_{Hot}(28) = 0.6$;

In the fuzzification stage, the notation $f_{Warm}(28) = 0.4$ represents that the numerical value '28' has a membership

degree of 0.4 to the linguistic attribute 'Warm'. This stage is aimed at mapping a crisp value from a numerical attribute to the value of a linguistic attribute (transformed from the numerical attribute), where the value of the linguistic attribute is essentially the membership degree to the fuzzy set (defined for the the linguistic attribute).

Application:
Rule 1: $f_{Short}(45) \wedge f_{Cold}(28) = Min(0, 0) = 0$;
Rule 2: $f_{Short}(45) \wedge f_{Warm}(28) = Min(0, 0.4) = 0$;
Rule 3: $f_{Short}(45) \wedge f_{Hot}(28) = Min(0, 0.6) = 0$;
Rule 4: $f_{Middle}(45) \wedge f_{Cold}(28) = Min(0.67, 0) = 0$;
Rule 5: $f_{Middle}(45) \wedge f_{Warm}(28) = Min(0.67, 0.4) = 0.67$;
Rule 6: $f_{Middle}(45) \wedge f_{Hot}(28) = Min(0.67, 0.6) = 0.6$;
Rule 7: $f_{Long}(45) \wedge f_{Cold}(28) = Min(0.33, 0) = 0$;
Rule 8: $f_{Long}(45) \wedge f_{Warm}(28) = Min(0.33, 0.4) = 0.33$;
Rule 9: $f_{Long}(45) \wedge f_{Hot}(28) = Min(0.33, 0.6) = 0.33$;

In the application stage, the two membership degree values obtained respectively for the two attributes 'x_1 and 'x_2' are combined through conjunction for inferring the strength to which a fuzzy rule fires. For example, Rule 8 has x_1 is Long and x_2 is Warm as its antecedent, so Rule 8 obtains the firing strength of 0.33, while $f_{Long}(45) = 0.33$ and $f_{Warm}(28) = 0.4$.

Implication:
Rule 1: $f_{Rule1 \rightarrow No}(45, 28) = 0$;
Rule 2: $f_{Rule2 \rightarrow No}(45, 28) = 0$;
Rule 3: $f_{Rule3 \rightarrow Yes}(45, 28) = 0$;
Rule 4: $f_{Rule4 \rightarrow No}(45, 28) = 0$;
Rule 5: $f_{Rule5 \rightarrow Yes}(45, 28) = 0.67$;
Rule 6: $f_{Rule6 \rightarrow No}(45, 28) = 0.6$;
Rule 7: $f_{Rule7 \rightarrow Yes}(45, 28) = 0$;
Rule 8: $f_{Rule8 \rightarrow No}(45, 28) = 0.33$;
Rule 9: $f_{Rule9 \rightarrow No}(45, 28) = 0.33$;

In the implication stage, the aim is at identifying the degree to which an input vector belongs to the class label 'Yes' or 'No' (i.e. the consequent of the fuzzy rule), based on the firing strength of a fuzzy rule identified in the application stage. For example, $f_{Rule6 \rightarrow No}(45, 28) = 0.6$ indicates that the class label 'No' is the consequent of Rule 6 and the input vector '$(45, 28)$' belongs to the class label 'No' to the membership degree of 0.6. In other words, the input vector '$(45, 28)$' obtains the membership degree of 0.6 to the class 'No' according to the inference through Rule 6.

Aggregation:
$f_{Yes}(45, 28) = f_{Rule3 \rightarrow Yes}(45, 28) \vee f_{Rule5 \rightarrow Yes}(45, 28) \vee f_{Rule7 \rightarrow Yes}(45, 28) = Max(0, 0.67, 0) = 0.67$

$f_{No}(45, 28) = f_{Rule1 \rightarrow No}(45, 28) \vee f_{Rule2 \rightarrow No}(45, 28) \vee f_{Rule4 \rightarrow No}(45, 28) \vee f_{Rule6 \rightarrow No}(45, 28) \vee f_{Rule8 \rightarrow No}(45, 28) \vee f_{Rule9 \rightarrow No}(45, 28)$
$= Max(0, 0, 0, 0.6, 0.33, 0.33) = 0.6$

In the aggregation stage, the aim is at deriving the overall degree to which an input vector belongs to the class 'Yes' or 'No', through identifying the maximum among all the membership degree values obtained through the inferences using the rules of each class. For example, the class label 'Yes' is the consequent of Rule 3, Rule 5 and Rule 7 and the input vector '$(45, 28)$' obtains the membership degree values of 0, 0.67 and 0, respectively, to the class label 'Yes', through using the above three rules. Since the maximum among the values of membership degree is 0.67, the inference in this stage indicates that the input vector belongs to the class label 'Yes' to the degree of 0.67.

Defuzzification:
$f_{Yes}(45, 28) > f_{No}(45, 28) \rightarrow class = Yes$;

In the defuzzification stage, the aim is at identifying the class label to which the input vector has the highest value of membership degree. In this example, as the input vector $(45, 28)$ obtains the membership degree of 0.67 to the class 'Yes', which is higher than the membership degree (0.6) to the class label 'No', the unseen instance $(45, 28, ?)$ is finally assigned 'Yes' as its class label.

C. Justification

We propose the adoption of fuzzy rule based systems for gender classification based on social network information, due to the strengths of fuzzy logic and its suitability for text processing, as outlined below.

Firstly, fuzzy logic is highly capable of handling the fuzziness, imprecision and uncertainty of text. In particular, it considers a classification problem to be a 'shade of grey' one rather than a 'black and white' one (currently considered in text classification). This way of defining the classification problem leads to a reduction of bias on both male and female classes. For example, popular probabilistic methods for text classification, such as C4.5, NB and SVM, handle continuous attributes through setting up crisp intervals, each of which is used to judge whether a condition is met through checking the values of the continuous attributes, towards classifying unseen instances. The above way of dealing with continuous attributes has been generally criticized as judgment bias in fuzzy systems literature, which can be replaced with using fuzzy intervals.

Secondly, fuzzy methods work in the strategy of generative learning rather than discriminative learning (typically used for training gender classifiers). In other words, fuzzy methods are designed to train classifiers that consider each class equally, through measuring the degree to which an instance belongs to each class independently, whereas those popularly used non-fuzzy methods are designed to train classifiers that aim to discriminate one class from all other classes, towards uniquely classifying an unseen instance. In the gender classification context, male and female people could have some shared language terms in writing blogs and posts [4]. Also, people of different genders may learn from each other in terms of writing style. Furthermore, it is possible in reality that people

may try to disguise themselves by showing intentionally the characteristics of the other gender in terms of writing style.

Thirdly, both male and female people are of high diversity in the world, i.e. people of each gender can be divided into different groups. From granular computing perspectives, each group of people can be viewed as a subclass of the male or female class. In real applications, it is unlikely that a training set can represent the full population of male and female people. From this point of view, each class (male or female) assigned to a training instance would actually represent a subclass of the male or female class, so an unseen instance may not belong to either one of the two classes, due to the case that the instance belongs to another subclass that is not included in the training set. When the above case arises, fuzzy approaches are capable of capturing it through showing that the instance has no membership (the membership degree of 0) to both classes [18]. In contrast, discriminative approaches cannot capture the above case, due to their nature of training classifiers to discriminate between the two classes.

IV. Experimental Setup and Results

In this section, we report an experimental study conducted by using a blog gender data set [12]. The data set contains 3226 blogs (1551 male and 1672 female).

In terms of classification performance, we compare the fuzzy approach with SVM, NB and C4.5, while three different types of features are extracted, namely, uni-gram (1-word term), bi-gram (2-word term) and tri-gram (3-word term). The results are shown in Table I.

TABLE I. Classification Accuracy

Feature Extraction	C4.5	NB	SVM	Fuzzy
Uni-gram	0.559	0.662	0.508	0.776
Bi-Gram	0.535	0.579	0.573	0.892
Tri-gram	0.641	0.692	0.781	0.806

Table I shows that the fuzzy method outperforms significantly the three non-fuzzy ones in all these cases. The results are likely due to the case that a fuzzy classifier is not biased on one of the two classes but judges independently on each class in terms of the membership degree of an instance to the class. As argued in Section III-C, individuals of different genders may have high similarity to each other in terms of writing style, which indicates that the two classes (male and female) would have overlaps. In this case, the nature of generative learning through fuzzy approaches makes it achievable to capture that highly similar patterns (writing styles) exist in blogs posted by both male and female authors.

In terms of feature extraction, the results show that the extraction of Bi-grams (2-word terms) leads to the best performance of the fuzzy approach, which are likely due to the case that the extraction of 2-word terms results in more features of higher frequency, leading to more useful and confident information (reflecting gender characteristics) being captured. In contrast, the use of 1-word terms could lead to loss of some important information, since multi-word terms are generally more informative than single-word terms. In addition, increasing the number of combined words (for making up a term) generally results in the decrease of term frequency, leading to the extracted features being less useful.

TABLE II. Membership Degree

No	Class	FM(Class=M)	FM(Class=F)	Prediction(Class)
1	M	1	0	M
2	F	0	1	F
3	F	0.33	0.67	F
4	F	0.5	0.5	F
5	M	0.67	0	M
6	M	0.17	0	M
7	F	0	0.5	F
8	M	1	0.43	M
9	M	1	1	M
10	F	1	1	F
11	M	0	0	?
12	F	0	0	?

The membership degree values of instances (selected as representative examples) to the two classes are shown in Table II. The results show diverse cases of gender classification. In particular, the first two cases (row 1 and row 2) indicate that the fuzzy classifier judges that the instance fully belongs to the male or female class, i.e. only characteristics of one gender are captured by the classifier and these characteristics are uniquely originated from people of one gender. The third case indicates that the fuzzy classifier captures both male and female characteristics from a blog posted by a female person, but the majority of the characteristics match the ones of female. The fourth case indicate that the fuzzy classifier captures characteristics that 50% match both male and female.

The above cases show that the membership degree values of an instance to the two classes can be added up to 1. However, the sum of the membership degree values is not necessarily equal to 1, i.e. it could be greater or less than 1. In particular, the fifth and sixth cases indicate that the fuzzy classifier captures characteristics of male only but the characteristics do not fully match the ones of male, i.e. for the fifth case the degree of matching is higher, but for the sixth case the degree is much lower. Also, the seventh case indicates that the fuzzy classifier capture characteristics of female only with the matching degree of 0.5. The above phenomenon can be explained by the commonsense that people of the same gender present different intensities of the characteristics originated from the majority of people of this gender.

The eighth case indicates that the fuzzy classifier captures characteristics that fully match the ones of male but also partially match the ones of female. This could be partially explained by the point (mentioned in Section III-C) that people of different genders have shared language styles in writing blogs. From this point of view, the author of the blog strongly presents the characteristics of male styled writing but the writing style also has some similarity to the one of female. The 9th and 10th cases indicate that the fuzzy classifier captures characteristics that fully match the ones of both male and female. The above phenomenon could be explained by two

points: a) a blog is written fully in shared language terms; b) a person of one gender presents in full the characteristics of the other gender in terms of writing style, which results in discovery of highly similar or even the same pattern from blogs posted by both male and female people.

The last two cases indicate that the fuzzy classifier judges that the instances do not belong to either one of the two classes, i.e. none of the gender characteristics, which are discovered from the training instances (blogs), are captured from the unseen instances. This is likely due to the high diversity of people. As mentioned in Section III-C, both male and female people can be subdivided into different groups, which are viewed as sub-classes of the male or female class. In real applications, it is likely that the training data only represents one or more (not all) groups of male and female people, which leads to the situation that an unseen instance belongs to another sub-class of the male or female class but the sub-class is absent from the training set.

V. Conclusion

In this paper, we proposed the use of fuzzy approaches for gender classification. In particular, we treat gender identification as a task of generative classification instead of discriminative classification. We compared the fuzzy approach with popularly used discriminative approaches (SVM, NB and C4.5), in terms of classification accuracy. The results show that the fuzzy approach outperforms the other three ones.

We also reported the results on fuzzy membership degree values of instances to two classes (male and female). The results show diverse cases of gender classification. In particular, individuals of different genders could have high similarity to each other in terms of their writing style. Also, due to the high diversity of people, it is likely that the training data does not represent a full population of male and female people, which could result in the case that a person does not present any characteristics that match the ones of male or female discovered from the training instances. Furthermore, it is also possible that the writing style captured from a blog matches fully the characteristics of shared language terms rather than any characteristics of a specific gender. In addition, it is also possible in reality that a person of one gender tries to disguise themselves by presenting the characteristics of the other gender. All of the above cases can be captured by using fuzzy approaches through identifying the degrees to which an instance belongs to the male and female classes.

In future, we will investigate how to achieve effective gender identification through using granular computing concepts. For example, due to the high diversity of people, both the male and female classes can be specialized/decomposed into sub-classes through information granulation [22]. Since the classes and sub-classes are located in different levels of granularity, traditional gender classification tasks can thus be extended in the setting of multi-granularity learning.

Acknowledgment

The authors acknowledge support for the research reported in this paper through the Research Development Fund at the University of Portsmouth.

References

[1] G. Guo, "Human age estimation and sex classification," in Video Analytics for Business Intelligence, C. Shan, F. Porikli, T. Xiang, and S. Gong, Eds., vol. 409. Heidelberg: Springer, 2012, pp. 101–131.

[2] ——, "Gender classification," in Encyclopedia of Biometrics, S. Z. Li and A. K. Jain, Eds. Boston, MA: Springer, 2014, pp. 1–6.

[3] N. Ali and L. Xavier, "Person identification and gender classification using gabor filters and fuzzy logic," Int. J. Electr. Electr, vol. 2, pp. 20–23, April 2014.

[4] F. Lin, Y. Wu, Y. Zhuang, X. Long, and W. Xu, "Human gender classification: a review," Int. J. Biom., vol. 8, 2016.

[5] J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," Neural processing letters, vol. 9, pp. 293–300, June 1999.

[6] L. E. Peterson, "K-nearest neighbor," Scholarpedia, vol. 4, p. 1883, 2009.

[7] S. Mitra and M. Savvides, "Gaussian mixture models based on the frequency spectra for human identification and illumination classification," in IEEE Workshop on Automatic Identification Advanced Technologies, 17-18 October 2005, pp. 245–250.

[8] K. Thiel and M. Berthold, "The knime text processing feature: An introduction," KNIME, Tech. Rep., 2012.

[9] K. Reynolds, A. Kontostathis, and L. Edwards, "Using machine learning to detect cyberbullying," in International Conference on Machine Learning and Applications, 18-21 December 2011, pp. 241–244.

[10] R. Zhao, A. Zhou, and K. Mao, "Automatic detection of cyberbullying on social networks based on bullying features," in International Conference on Distributed Computing and Networking, 4-7 January 2016.

[11] M. Corney, O. de Vel, A. Anderson, and G. Mohay, "Gender-preferential text mining of e-mail discourse," in Computer Security Applications Conference, 9-13 December 2002, pp. 282–289.

[12] A. Mukherjee and B. Liu, "Improving gender classification of blog authors," in Conference on Empirical Methods in natural Language Processing, 9-11 October 2010, pp. 207–217.

[13] J. P. Carvalho, F. Batista, and L. Coheur, "A critical survey on the use of fuzzy sets in speech and natural language processing," in IEEE International Conference on Fuzzy Systems, Brisbane, QLD, Australia, 10-15 June 2012.

[14] F. Batista and J. P. Carvalho, "Text based classification of companies in crunchbase," in IEEE International Conference on Fuzzy Systems, Istanbul, Turkey, 2-5 August 2015.

[15] D. Chandran, K. A. Crockett, D. Mclean, and A. Crispin, "An automatic corpus based method for a building multiple fuzzy word dataset," in IEEE International Conference on Fuzzy Systems, Istanbul, Turkey, 2-5 August 2015.

[16] M. Vicente, F. Batista, and J. P. Carvalho, "Twitter gender classification using user unstructured information," in IEEE International Conference on Fuzzy Systems, Istanbul, Turkey, 2-5 August 2015.

[17] H. Liu and M. Cocea, "Fuzzy rule based systems for interpretable sentiment analysis," in International Conference on Advanced Computational Intelligence, Doha, Qatar, 4-6 February 2017, pp. 129–136.

[18] C. Jefferson, H. Liu, and M. Cocea, "Fuzzy approach for sentiment analysis," in IEEE International Conference on Fuzzy Systems, Naples, Italy, 9-12 July 2017.

[19] E. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," Int. J. Human-Computer Stud, vol. 51, pp. 135–147, January 1999.

[20] S.-M. Chen, "A fuzzy reasoning approach for rule-based systems based on fuzzy logics," IEEE Trans. Syst., Man, Cybern. B, vol. 26, pp. 769–778, October 1996.

[21] F. Bergadano and V. Cutello, "Learning membership functions," in European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Granada, Spain, 8-10 November 1993, pp. 25–32.

[22] H. Liu and M. Cocea, "Semi-random partitioning of data into training and test sets in granular computing context," Granul. Comput, vol. 2, pp. 357–386, December 2017.