SELECTING FEATURES BY UTILIZING INTUITIONISTIC FUZZY ENTROPY METHOD

Kiran Pandey[1], Arunodaya Raj Mishra[2], Pratibha Rani[3*], Jabil Ali[4] and Ripon Chakrabortty[5]

[1] Department of Computer Science and Engineering, Technocrats Institute of Technology, Bhopal, Madhya Pradesh, India;
[2] Department of Mathematics, Government College Raigaon, Satna, Madhya Pradesh, India;
[3] Department of Engineering Mathematics, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India;
[4] School of Engineering and Technology, Department of Computer Science and Engineering, Sharda University, Greater Noida, Uttar Pradesh, India;
[5] Capability Systems Centre, School of Engineering and Information Technology, UNSW Canberra, Australia

Original scientific paper

Abstract: *Feature selection is the most significant pre-processing activity, which intends to reduce the data dimensionality for enhancing the machine learning process. The evaluation of feature selection must consider classification, performance, efficiency, stability, and many factors. Nowadays, uncertainty is commonly occurred in the feature selection process due to time limitations, imprecise information, and the subjectivity of human minds. Moreover, the theory of intuitionistic fuzzy set has been proven as an extremely valuable tool to tackle the uncertainty and ambiguity that arises in many practical situations. Thus, this study introduces a novel feature selection framework using intuitionistic fuzzy entropy. In this regard, new entropy for IFS is proposed first and then compared with some of the previously developed entropy measures. As entropy is a measure of uncertainty present in data (features), features with higher entropy values are filtered out, and the remaining features having lower entropy values have been used to classify the data. To verify the effectiveness of the proposed entropy-based feature selection, some experiments are done with ten standard benchmark datasets by employing a support vector machine, K-nearest neighbor, and Naïve Bias classifiers. The outcomes of the study validate that the proposed entropy-based filter feature selection is more feasible and impressive than existing filter-based feature selection methods.*

Key words: *Intuitionistic fuzzy set, entropy, feature selection, classifier accuracy.*

* Corresponding author.
    E-mail addresses: kiran.mishra.cse07@iitbhu.ac.in (K. Pandey), arunodaya87@outlook.com (A. R. Mishra), pratibha138@gmail.com (P. Rani), jabirali.cse@gmail.com (J. Ali), r.chakrabortty@adfa.edu.au (R. Chakrabortty).

1. Introduction

Overfitting and the curse of dimensionality are the two biggest problems in machine learning. "Feature selection (FS)" assists to evade both of these concerns by decreasing the number of features in the model and optimizing the framework's performance. In addition, FS gives an additional benefit of model interpretation. In case of minor features, the output model becomes easier to understand and more reliable for a human to expect future forecasts made by the model. Existing studies have classified the FS method into three categories: filter, wrapper, and embedded techniques (Blum and Langley, 1997). The filter method chooses features independent of any classification algorithm (Parlak and Uysal, 2021). This method has broadly been used for FS of text data due to its simplicity and not being prone to overfitting (Revanasiddappa and Harish, 2018). The wrapper method selects features based on the learning algorithm (Kohavi and John, 1997). It is more computationally expensive and takes more time than the filter method, which results from calling the learning algorithm for each feature set considered (Blum and Langley, 1997), and makes it unreasonable. On the other hand, the embedded method works as a part of the classification algorithm and ranks the feature during the learning stage. Compared to wrapper and embedded techniques, the filter-based techniques operate independently of any learning algorithm-undesirable features are filtered out of the data before induction commences and thus are more adaptable and less computationally intensive for diverse datasets and classifiers (Parlak and Uysal, 2021).

Uncertainty is an inherent feature of information. In several scientific and industrial applications, we make decisions in an environment with diverse kinds of uncertainty. The "fuzzy set theory (FST)" invented by Zadeh (1965) has successfully been employed in varied areas and demonstrated its powerful ability to treat with the vague and uncertain information. In the literature, several doctrines and principles have been studied on FST (Precup et al., 2020; Kumar et al., 2022; Bairagi (2022). Further, Atanassov (1986) extended the FST to "intuitionistic fuzzy set (IFS)", which deals with the uncertain and ambiguous information more accurately. In IFSs, each object is defined with the degrees of membership and non-membership. The theory of IFS is one of the powerful and suitable tools to cope with the vagueness presented in numerous realistic decision-making applications. Research works on IFS theory and its applications in different settings are developing speedily, and several significant outcomes have been obtained (Kushwaha et al., 2020; Tripathi et al., 2022a, b; Hezam et al., 2022).

The notion of entropy provides a measure of information gained by comparing dissimilar attributes and a measure about the randomness of data. In the FS process, entropy can compute the amount of uncertainty and the quality of data information and improve the model's accuracy (Zhao et al., 2020; High et al., 2021). The Shannon entropy is usually considered the standard and most natural way to measure the expected value of the information. To deal with uncertain and vague information, the classic fuzzy entropy measure based on Shannon entropy has been widely applied for FS process (Tran et al., 2021). In the past, Lee et al. (2001) employed the fuzzy entropy to assess the information of pattern distribution by dividing the data into non-overlapping decision support systems. Further, Luukka (2011) used a feature selection approach using fuzzy entropy. As an extended version of FST, the notion of IFS (Atanassov, 1986) gives the details of both degrees of membership and non-membership along with the hesitancy margin of its elements. Thus, it has a strong capability to describe the vagueness of the data in comparison with the FST. Very few

authors have studied the FS methods based on intuitionistic fuzzy entropy (Revanasiddappa and Harish, 2018; Tiwari et al., 2019; Singh et al., 2019).

Based on the earlier studies, we identify the following research challenges: (i) Entropy, as an information measure, has been gained much attention in many fields such as deep learning, data science, machine learning, image segmentation, texture analysis etc. In the recent past, the notion of IFSs is proven as more flexible and superior way to model the vagueness and imprecision of complex real-life applications. However, several authors have proposed different intuitionistic fuzzy entropy measures, but these measures have some counter intuitive cases. (ii) In the literature, there is no study regarding the use of intuitionistic fuzzy entropy for feature selection.

Thus, the key contributions of the article are given by

- New entropy for intuitionistic fuzzy set is developed and compared with the extant entropies under IFSs.
- Proposed entropy has been utilized to choose the features which contain lower uncertainty (higher information) in the data. This work is based on "support vector machine (SVM)", "K-nearest neighbor (KNN)" and "Naïve Bias from the classification domain" for investigating the quality of solutions obtained by means of initial preprocessing and the information indicated by the intuitionistic fuzzy entropy.

The rest parts of the sections are summarized as Section 2 confers the related works of FS process. Section 3 offers the proposed entropy for IFSs and shows its validity by comparing extant entropies. Section 4 introduces a novel FS framework using developed IF-entropy. Section 5 explains the experimental work of the proposed algorithm. Section 6 presents the conclusion with the needful potential research direction.

2. Literature review

FS is defined as "one of the well-known dimensionality reduction methods, which can select a small subset of significant and non-redundant features from the original information systems" (Omuya et al., 2021). The main aim of FS is to remove redundant and/or irrelevant features, improve the learning algorithm's performance, reduce the cost of computation, and offer more explicit and concise description of data (Murugesan et al., 2021). Several notions of FS have been presented in machine learning, data mining, bioinformatics, text categorization, signal processing and others (Kim and Zzang, 2019; Álvarez et al., 2019, Ruan et al., 2021, Pintas et al., 2021). In general, the FS techniques are partitioned into a filter, wrapper and embedded methods (Tang et al., 2016; Rehman et al., 2017). A filter method performs statistical analysis to recognize the more relevant features from the less significant ones. Consequently, it is free from the classification algorithm (Revanasiddappa and Harish, 2018). The wrapper method uses the particle swarm optimization algorithm (Ji et al., 2020), the genetic algorithm (Rostami et al., 2021) or other searching algorithms to find an optimum feature set (Chen et al.,2021), while the embedded method selects the features during the model training. As compared to the wrapper and embedded techniques, the filter-based techniques are more flexible, speedy and have less computational complexity.

Due to the advantages of filter-based methods over other methods, numerous research efforts have been made in the literature. Existing studies have presented diverse filter-based methods, which are document frequency (Kim and Zzang, 2019),

term frequency-inverse document frequency (Thakkar and Chaudhari, 2020), mutual information (Gao and Wu, 2020), entropy-based FS (Zhang et al., 2021a) and others (Bahassine et al., 2020; Omuya et al., 2021). Amongst these FS methods, the entropy-based FS technique calculates the amount of uncertainty and the eminence of information. Several authors have focused their attention on entropy-based FS techniques. For illustration, Sun et al. (2012) adopted the concept of Shannon entropy to develop a universal FS method to select relevant and significant features. Inspired by the idea of entropy, Jaganathan and Kuppuchamy (2013) discussed an innovative FS method and its application in medical database classification. Zhang et al. (2016) formulated a hybrid FS model by combining entropy and fuzzy rough set for managing the mixed dataset. Inspired by Luukka (2011), Lohrmann et al. (2018) offered a fuzzy entropy and degree of similarity-based FS technique to distinguish the appropriate features. In a study, Sun et al. (2019) used the notions of Lebesgue and entropy measures to introduce a hybrid FS model for mixed and incomplete neighbourhood decision systems. Recently, Qu et al. (2020) studied an innovative FS model based on non-unique decision differential entropy for handling the nominal data. Based on correlation and relative entropy, Aremu et al. (2020) pioneered a novel feature engineering method for raw asset data. Later, Sun et al. (2021) recommended a hybrid fuzzy neighborhood entropy-based FS model for heterogeneous datasets. Motivated by the fuzzy-neighborhood relative decision entropy, Zhang et al. (2021b) designed an integrated feature ranking method and verified its validity through a data experiment and a numerical example. As an extended version of the fuzzy set, IFSs (Atanassov, 1986) are more reliable and efficient to tackle the uncertainty in real applications. Feature selection models based on intuitionistic fuzzy entropy can provide more suitability to extract the features in data mining, machine learning and text categorization. However, very few studies have been presented regarding the intuitionistic fuzzy entropy-based FS methods for classifiers in the literature (Singh et al., 2019; Tiwari et al., 2019).

3. Proposed intuitionistic fuzzy entropy

The current section firstly presents basic definitions. Then, new entropy for IFS is proposed with its enviable characteristics.

3.1. Preliminaries

This section concentrates on the demonstration of the decision information based on IFSs. In the following step, this study focuses on the aggregation based on the IF-CoCoSo method.

In the FSs doctrine, the MF of an element is represented based on the interval number of [0, 1], whereas the NF essentially is complement. Though, in concern, this hypothesis does not meet with human opinions. Hence, Atanassov (1986) defined the IFSs as follows:

Definition 1. In the following, Zadeh (1965) presents the mathematical definition of fuzzy set $\tilde{F}$ on a fixed universal set $Y = \{y_1, y_2, ..., y_n\}$:

$$\tilde{F} = \left\{ \left\langle y_i, \mu_{\tilde{F}}(y_i) \right\rangle : y_i \in Y \right\}, \tag{1}$$

wherein $\mu_{\tilde{F}}(y_i) \left(0 \le \mu_{\tilde{F}}(y_i) \le 1 \right)$ denotes the membership grade of y_i to $\tilde{F}$ in Y.

Definition 2. In 1986, Atanassov (1986) defined the mathematical form of an IFS F on Y as

$$F = \left\{ \left\langle y_i, F\left(\mu_F(y_i), \nu_F(y_i)\right)\right\rangle \middle| y_i \in Y \right\}, \tag{2}$$

where $\mu_F : Y \to [0,1]$ and $\nu_F : Y \to [0,1]$ represent the grades of membership and non-membership of $y_i \in Y$ to F, respectively, satisfying the constraint $0 \le \mu_F(y_i) + \nu_F(y_i) \le 1$. For each $y_i \in Y$, the indeterminacy degree is defined by $\pi_F(y_i) = 1 - \mu_F(y_i) - \nu_F(y_i)$. For convenience, $F(\mu_F(y_i), \nu_F(y_i))$ is defined as an "intuitionistic fuzzy number (IFN)" and denoted by $\theta = F(\mu_\theta, \nu_\theta)$, where $\mu_\theta, \nu_\theta \in [0,1]$ and $0 \le \mu_\theta + \nu_\theta \le 1$.

Definition 3 (Atanassov, 1986). For any two intuitionistic fuzzy numbers $\theta_1 = F\left(\mu_{\theta_1}, \nu_{\theta_1}\right)$ and $\theta_2 = F\left(\mu_{\theta_2}, \nu_{\theta_2}\right)$, the basic operational laws are presented as

$$\theta_1 \oplus \theta_2 = F\left(\mu_1 + \mu_2 - \mu_1\mu_2, \nu_1\nu_2\right); \tag{3}$$

$$\theta_1 \otimes \theta_2 = F\left(\mu_1\mu_2, \nu_1 + \nu_2 - \nu_1\nu_2\right); \tag{4}$$

$$\theta_1^\lambda = F\left((\mu_1)^\lambda, 1 - (1-\nu_1)^\lambda\right); \tag{5}$$

$$\lambda\theta_1 = F\left(1 - (1-\mu_1)^\lambda, (\nu_1)^\lambda\right). \tag{6}$$

Definition 4 (Szmidt and Kacprzyk, 2001). Let $F, G \in IFSs(Y)$. A mapping $e : IFS(Y) \to [0,1]$ is called entropy for IFS if

(a1). $0 \le e(F) \le 1$,

(a2). $e(F) = 0$ iff F is a crisp set,

(a3). $e(F) = 1$ iff $\pi_F(y_i) = 1, \ \forall \ y_i \in Y$,

(a4). $e(F) = e(F^c)$,

(a5). For each $y_i \in Y$, $e(F) \ge e(G)$ iff the hesitancy index of the elements of G is less than the hesitancy index of the elements of F, i.e., $\pi_G \le \pi_F$.

3.2. Entropy for IFSs

Szmidt and Kacprzyk (2001) firstly studied the definition of intuitionistic fuzzy (IF)-entropy measure. Further, several IF-entropy measures have been put forward with lots of applications (Bustince and Burillo, 1996; Vlachos and Sergiadis, 2007; Wei et al., 2012; Mishra 2016; Rani and Jain, 2017; Ansari et al., 2018; Mishra and Rani, 2019; Mishra et al., 2017a,b, 2019a,b, 2020a,b; Rahimi et al., 2021; Rani et al., 2021; Thao and Chou, 2022). To evade the drawback of some existing entropies, this section develops novel entropy for IFS and then apply it to develop a new feature selection method in the next section.

For this, consider $F \in IFS(Y)$, then the entropy measure for IFS is given by

$$e(F) = 1 - \frac{1}{n}\sum_{i=1}^{n}\left[1 - \frac{1}{2}\left(1 - \mu_F(y_i) - \nu_F(y_i)\right) + \frac{4}{\pi}\tan^{-1}\left(1 - \mu_F(y_i) - \nu_F(y_i)\right)^2\right]. \tag{7}$$

Theorem 1. A mapping $e(F)$ given by Eq. (7), is a valid IF-entropy.

Proof: In order to verify the theorem, the mapping $e(F)$ must hold the postulates (a1)-(a5) of Definition 4.

(a1). In IFS, we know that $0 \leq \mu_F(y_i), \nu_F(y_i) \leq 1$ and $0 \leq \mu_F(y_i) + \nu_F(y_i) \leq 1$. Thus, $0 \leq e(F) \leq 1$.

(a2). For a crisp set F, we know that $\mu_F(y_i) = 1$, $\nu_F(y_i) = 0$ or $\mu_F(y_i) = 0$, $\nu_F(y_i) = 1$. Then, from Eq. (7), we have

$$e(F) = 1 - \frac{1}{n} \sum_{i=1}^{n} \left[1 - \frac{1}{2}(1-1-0) + \frac{4}{\pi} \tan^{-1}(1-1-0)^2 \right] = 0.$$

Conversely, assume that $e(F) = 0$. This implies that

$$1 - \frac{1}{n} \sum_{i=1}^{n} \left[1 - \frac{1}{2}(1 - \mu_F(y_i) - \nu_F(y_i)) + \frac{4}{\pi} \tan^{-1}(1 - \mu_F(y_i) - \nu_F(y_i))^2 \right] = 0,$$

$$\sum_{i=1}^{n} \left[1 - \frac{1}{2}(1 - \mu_F(y_i) - \nu_F(y_i)) + \frac{4}{\pi} \tan^{-1}(1 - \mu_F(y_i) - \nu_F(y_i))^2 \right] = n,$$

$$\sum_{i=1}^{n} \left[1 - (\mu_F(y_i) + \nu_F(y_i)) + \frac{4}{\pi} \tan^{-1}(1 - \mu_F(y_i) - \nu_F(y_i))^2 \right] = 0,$$

$$\sum_{i=1}^{n} \left[(\mu_F(y_i) + \nu_F(y_i)) - \frac{4}{\pi} \tan^{-1}(1 - \mu_F(y_i) - \nu_F(y_i))^2 \right] = n. \tag{8}$$

For each $y_i \in Y$, the left-hand side of Eq. (8) cannot exceed 1. And, Eq. (1) is satisfied only when for all $y_i \in Y$, the left-hand side should be 1.

$$(\mu_F(y_i) + \nu_F(y_i)) - \frac{4}{\pi} \tan^{-1}(1 - \mu_F(y_i) - \nu_F(y_i))^2 = 1.$$

Since $0 \leq \mu_F(y_i) + \nu_F(y_i) \leq 1$, therefore, Eq. (7) can only be satisfied when $\tan^{-1}\left[(1 - \mu_F(y_i) - \nu_F(y_i))^2\right] = 0$ and this can only be zero when $\mu_F(y_i) + \nu_F(y_i) = 1$.

This is only possible when $\mu_F(y_i) = 1$, $\nu_F(y_i) = 0$ or $\mu_F(y_i) = 0$, $\nu_F(y_i) = 1$. This implies F is a crisp set.

(a3). If $e(F) = 1$,

$$\Leftrightarrow 1 - \frac{1}{n} \sum_{i=1}^{n} \left[1 - \frac{1}{2}(1 - \mu_F(y_i) - \nu_F(y_i)) + \frac{4}{\pi} \tan^{-1}(1 - \mu_F(y_i) - \nu_F(y_i))^2 \right] = 1,$$

$$\Leftrightarrow \sum_{i=1}^{n} \left[1 - \frac{1}{2}(1 - \mu_F(y_i) - \nu_F(y_i)) + \frac{4}{\pi} \tan^{-1}(1 - \mu_F(y_i) - \nu_F(y_i))^2 \right] = 0,$$

$$\Leftrightarrow \sum_{i=1}^{n} \left[(\mu_F(y_i) + \nu_F(y_i)) - \frac{4}{\pi} \tan^{-1}(1 - \mu_F(y_i) - \nu_F(y_i))^2 \right] = -n. \tag{9}$$

Since $\dfrac{4}{\pi}\tan^{-1}\left(1-\mu_F(y_i)-\nu_F(y_i)\right)^2 \leq 1$, therefore, for each $y_i \in Y$, the left-hand side of Eq. (9) has minimum value -1. Hence, Eq. (4) is satisfied only when the left-hand side equals to -1 for all $y_i \in Y$,

$$\left(\mu_F(y_i)+\nu_F(y_i)\right)-\frac{4}{\pi}\tan^{-1}\left(1-\mu_F(y_i)-\nu_F(y_i)\right)^2 = -1. \tag{10}$$

Eq. (10) is obtained when $\dfrac{4}{\pi}\tan^{-1}\left(1-\mu_F(y_i)-\nu_F(y_i)\right)^2 = 1$ and it is true for $\mu_F(y_i)+\nu_F(y_i)=0$. This is possible only when $\mu_F(y_i)=0$ and $\nu_F(y_i)=0$, i.e., $\pi_F(y_i)=1, \ \forall \ y_i \in Y$.

(a4). It is evident from Eq. (3) that $e(F)=e(F^c)$.

(a5). For $F,G \in IFSs(Y)$, if $e(F) \geq e(G)$,

$$\Leftrightarrow 1-\frac{1}{n}\sum_{i=1}^{n}\left[1-\frac{1}{2}\left(1-\mu_F(y_i)-\nu_F(y_i)\right)+\frac{4}{\pi}\tan^{-1}\left(1-\mu_F(y_i)-\nu_F(y_i)\right)^2\right]$$

$$\geq 1-\frac{1}{n}\sum_{i=1}^{n}\left[1-\frac{1}{2}\left(1-\mu_G(y_i)-\nu_G(y_i)\right)+\frac{4}{\pi}\tan^{-1}\left(1-\mu_G(y_i)-\nu_G(y_i)\right)^2\right];$$

$$\Leftrightarrow \sum_{i=1}^{n}\left[\frac{1}{2}\left(1-\mu_F(y_i)-\nu_F(y_i)\right)+\frac{4}{\pi}\tan^{-1}\left(1-\mu_F(y_i)-\nu_F(y_i)\right)^2\right]$$

$$\leq \sum_{i=1}^{n}\left[\frac{1}{2}\left(1-\mu_G(y_i)-\nu_G(y_i)\right)+\frac{4}{\pi}\tan^{-1}\left(1-\mu_G(y_i)-\nu_G(y_i)\right)^2\right];$$

$$\Leftrightarrow \left(1-\mu_F(y_i)-\nu_F(y_i)\right)+\frac{4}{\pi}\tan^{-1}\left(1-\mu_F(y_i)-\nu_F(y_i)\right)^2$$

$$\leq \left(1-\mu_G(y_i)-\nu_G(y_i)\right)+\frac{4}{\pi}\tan^{-1}\left(1-\mu_G(y_i)-\nu_G(y_i)\right)^2;$$

$$\Leftrightarrow \left(1-\mu_F(y_i)-\nu_F(y_i)\right) \leq \left(1-\mu_G(y_i)-\nu_G(y_i)\right);$$

$$\Leftrightarrow \pi_F(y_i) \leq \pi_G(y_i); \ \ \forall \ y_i \in Y.$$

This completes the proof.

3.3. Performance of the proposed entropy for IFS with extant entropies

In the following, we first recall some of the extant intuitionistic fuzzy entropy measures to show the usefulness and utility of proposed entropy for IFSs:

$$e_{bb}(F) = \frac{1}{n}\sum_{i=1}^{n}\left(1-\mu_F(y_i)-\nu_F(y_i)\right), \tag{11}$$

$$e_{sk}(F) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\left(1-\nu_F(y_i)\right)\wedge\left(1-\mu_F(y_i)\right)}{\left(1-\nu_F(y_i)\right)\vee\left(1-\mu_F(y_i)\right)}\right), \tag{12}$$

$$e_{zj}(F) = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\mu_F(y_i) \wedge \nu_F(y_i)}{\mu_F(y_i) \vee \nu_F(y_i)} \right), \tag{13}$$

$$e_{hc}^2(F) = \frac{1}{n} \sum_{i=1}^{n} \left(1 - \mu_F^2(y_i) - \nu_F^2(y_i) - \pi_F^2(y_i)\right), \tag{14}$$

$$e_s(F) = $$
$$-\frac{1}{n} \sum_{i=1}^{n} \left(\mu_F(y_i) \ln \mu_F(y_i) + \nu_F(y_i) \ln \nu_F(y_i) + \pi_F(y_i) \ln \pi_F(y_i) \right), \tag{15}$$

$$e_{vs}(F) = -\frac{1}{n \ln 2} \sum_{i=1}^{n} \left(\mu_F(y_i) \ln \mu_F(y_i) + \nu_F(y_i) \ln \nu_F(y_i) \right.$$
$$+ \left(\mu_F(y_i) + \nu_F(y_i) \right) \ln \left(\mu_F(y_i) + \nu_F(y_i) \right) - \pi_F(y_i) \ln 2, \tag{16}$$

$$e_w(F) = \frac{1}{n} \sum_{i=1}^{n} \left[\left\{ \sqrt{2} \cos \left(\frac{\mu_F(y_i) - \nu_F(y_i)}{4} \right) \pi - 1 \right\} \times \frac{1}{\sqrt{2} - 1} \right], \tag{17}$$

$$e_m(F) = \frac{1}{n\sqrt{e}\left(\sqrt{e}-1\right)} \sum_{i=1}^{n} \left[e - \left(\frac{\mu_F(y_i) + 1 - \nu_F(y_i)}{2} \right) e^{\left(\frac{\mu_F(y_i) + 1 - \nu_F(y_i)}{2} \right)} \right.$$

$$\left. - \left(\frac{\nu_F(y_i) + 1 - \mu_F(y_i)}{2} \right) e^{\left(\frac{\nu_F(y_i) + 1 - \mu_F(y_i)}{2} \right)} \right], \tag{18}$$

$$e_a(F) = \frac{1}{2^{1-\alpha} - 1} \left[\exp \left\{ (\alpha - 1) \sum_{i=1}^{n} \left(\mu_F(y_i) \ln \left(\frac{\mu_F(y_i)}{\mu_F(y_i) + \nu_F(y_i)} \right) \right. \right. \right.$$

$$\left. \left. \left. + \nu_F(y_i) \ln \left(\frac{\nu_F(y_i)}{\mu_F(y_i) + \nu_F(y_i)} \right) - \pi_F(y_i) \ln 2 \right\} - 1 \right],$$

(19)

$$e_b(F) = \frac{-1}{n\left(\exp(2) - 1\right)} \sum_{i=1}^{n} \left[\left(\mu_F(y_i) - \nu_F(y_i) \right) \left\{ \exp \left(\frac{2 \mu_F(y_i)}{\mu_F(y_i) + \nu_F(y_i)} \right) \right. \right.$$

$$\left. \left. - \exp \left(\frac{2 \nu_F(y_i)}{\mu_F(y_i) + \nu_F(y_i)} \right) \right\} - \exp(2) + 1 \right]. \tag{20}$$

Example 1 (Zhang and Jiang, 2008; Ansari et al., 2018). Let $F \in IFS(Y)$. For any positive real number n, we have

$$F^n = \left\{ \left\langle y_i, \left[\mu_F \left(y_i \right) \right]^n, 1 - \left[1 - \nu_F \left(y_i \right) \right]^n \right\rangle : y_i \in Y \right\}. \tag{21}$$

We assume $F \in IFS(Y)$ given by

$$F = \left\{ (6, 0.1, 0.8), (7, 0.3, 0.5), (8, 0.5, 0.4), (9, 0.9, 0), (10, 1.0, 0) \right\} \tag{22}$$

By seeing the categorization of linguistic terms, Mishra and Rani (2019) observed F as "LARGE" on Y. Using Eq. (21), we obtain

$F^{1/2}$ considered as "More or less LARGE"

F^2 considered as "Very LARGE"

F^3 considered as "Quite very LARGE"

F^4 considered as "Very very LARGE".

Now, we compare the outcomes achieved by developed and extant IF-entropies. The entropy measures follow the pattern of Eq. (23).

$$e\left(F^{1/2}\right) \geq e\left(F\right) \geq e\left(F^2\right) \geq e\left(F^3\right) \geq e\left(F^4\right). \tag{23}$$

Table 1. Results by different entropies

Entropies	$F^{1/2}$	F	F^2	F^3	F^4
$e_{bb}(F)$	**0.4090**	**0.5000**	**0.4900**	**0.4670**	**0.4670**
$e_{sk}(F)$	**0.3450**	**0.3740**	**0.1970**	**0.1310**	**0.1090**
$e_{hy}^2(F)$	**0.3420**	**0.3440**	**0.2610**	**0.1990**	**0.1610**
$e_s(F)$	0.4330	0.4310	0.3270	0.2530	0.2080
$e_{vs}(F)$	0.5518	0.5217	03491	0.2354	0.1417
$e_{zj}(F)$	**0.2851**	**0.3050**	**0.1042**	**0.0383**	**0.0161**
$e_w(F)$	0.4545	0.4377	0.3029	0.2159	0.1709
$e_m(F)$	0.5522	0.5333	0.3758	0.2719	0.2149
$e_a(F)$	0.4659	0.4356	0.2737	0.1775	0.1032
$e_b(F)$	0.5416	0.5072	0.3255	0.2113	0.1523
$e(F)$	0.6579	0.6257	0.4287	0.3139	0.2620

According to Table 1, the results obtained by $e_{bb}(F)$, $e_{sk}(F)$, $e_{hy}^2(F)$ and $e_{zj}(F)$ have some unreasonable cases because these measures do not hold the pattern Eq. (5). However, the measures $e_s(F)$, $e_{vs}(F)$, $e_w(F)$, $e_m(F)$, $e_a(F)$, $e_b(F)$ and $e(F)$ perform well on given examples. From Example 1, we accomplish that the performance of entropy $e(F)$, given in Eq. (7), is better than various extant entropies.

4. Calculation of proposed intuitionistic fuzzy entropy for the dataset

The current section contains an algorithm for the calculation of IF-entropy of the attributes for the dataset. For calculating membership and non-membership

functions, we have used the Bell-shaped function implemented by Rangasamy (2021) and mentioned in Eq. (24) and Eq. (25). Here, $\mu_P(j)$, $\nu_P(j)$ and $\pi_P(j)$ represent the grades of membership, non-membership and indeterminacy for jth feature of feature set P, respectively.

$$\mu_P(j) = 1 - \pi_P(j) - \left(\frac{1}{1 + \left| \frac{j-c}{a} \right|^{2b}} \right), \tag{24}$$

$$\nu_F(j) = \frac{1}{1 + \left| \frac{j-c}{a} \right|^{2b}}, \tag{25}$$

where three parameters a, b, c and usually the parameter 'b' is positive. The parameter 'c' places at the center of the curve and 'b' control the slopes at the crossover points and $0 \le \pi_P(j) \le 1$.

Algorithm 1: Algorithm for calculation of intuitionistic fuzzy entropy for different attributes of the dataset

---

Input: P, n, m (P is the set of attributes, n and m are the numbers of attributes and the observations in the dataset, respectively)

Output: e(F),the set of entropies for P

Set E=0, e(F)={ϕ}

For each i=1 to n

    For j= 1 to m

        Normalize the data;

        Calculate $\mu_P(j), \nu_P(j)$ and $\pi_P(j)$(membership, non-membership and indeterminacy;

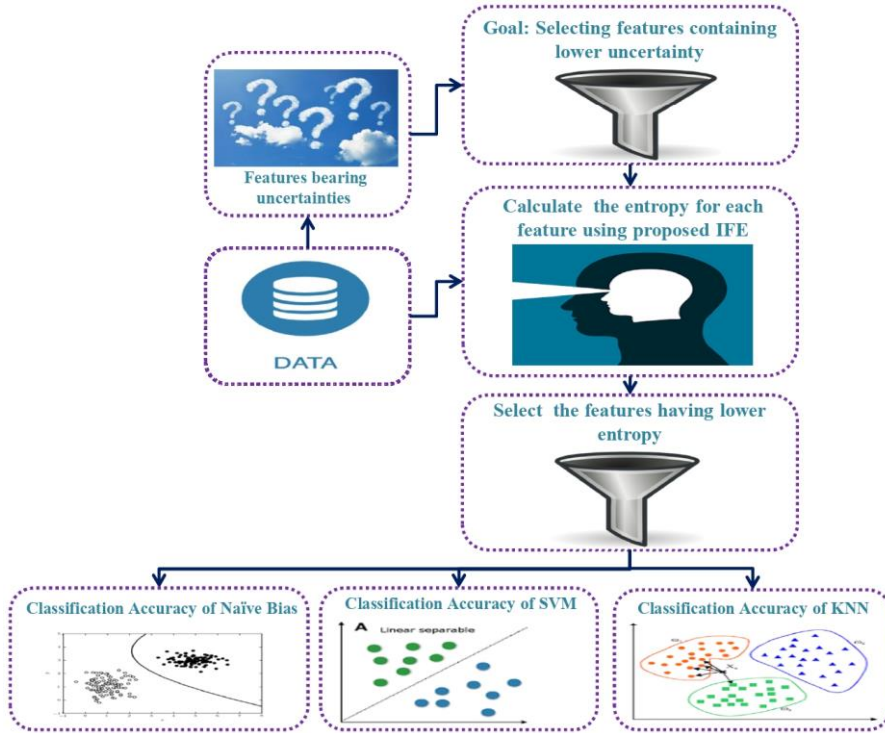   *Calculate e(F) (Use Eq. (7) for the calculation of entropy);*

        E=E+E$_j$;

    End For

    Set e(F)=e(F)⬚ {E};

End For

Return e(F)

Selecting features by utilizing intuitionistic fuzzy entropy method

Figure 1. Flowchart of the proposed methodology

Figure 1 presents the utilization of proposed IFE for feature selection. We have taken datasets and the features of these datasets bear uncertainty. We need to filter out the higher uncertainty features before performing the classification task. As entropy is utilized as a measure of uncertainty (Zhang et al., 2021a), we have calculated the entropy of features using proposed intuitionistic fuzzy entropy and then filter out the features with higher entropies and provide the dataset (with filtered attributes) to the SVM, KNN and Naïve Bias classifiers. Then, we have measured classification accuracy with raw data as well as with filtered data. In most cases, we are getting better accuracy with a reduced feature set (as discussed below in Section 5).

5. Experimentation

For evaluating the proposed approach, we have used 10 benchmark datasets. All datasets except cancer patients are obtained from the "UCI (University of California, Irvine) machine learning repository" (Dua and Graff, 2019) for our proposed algorithm's qualitative and quantitative analysis. The indicators within each dataset have a diversity of characteristics (i.e., several binary/discrete and several continuous). Table 2 contains the details of benchmark datasets used for experimentation.

Table 2. Benchmark data for experimental work

S. No.	Dataset	Number of Observations	Number of Features	Number of Classes	References
1.	Cancer Coimbra	116	10	2	Dua and Graff (2019)
2	Breast tissue	101	10	6	Dua and Graff (2019)
3	Cancer patient	1000	24	3	www.kaggle.com, 2021
4	Seeds	211	8	3	Dua and Graff (2019)
5	Sonar	208	61	2	Dua and Graff (2019)
6	Heart	303	14	2	Dua and Graff (2019)
7	Pima diabetes	768	9	2	Dua and Graff (2019)
8	Wine	1599	12	6	Dua and Graff (2019)
9	Internet firewall	65532	12	4	Dua and Graff (2019)
10	Glass	214	10	7	Dua and Graff (2019)

Table 3. Lowly ranked features using various filter feature selection algorithms

Data Sets	Rank	IFE-FS	FE-FS	ReliefF	Lasso	Laplacian	mRMR	Mutinffs
Cancer Coimbra	R1	p5	p8	p8	p9	p2	p8	p8
	R2	p7	p7	p1	p5	p9	p6	p3
	R3	p8	p4	p9	p7	p5	p2	p1
Breast Tissue	R1	p6	p3	p4	p6	p7	p7	p7
	R2	p3	p2	p5	p2	p6	p9	p6
	R3	p2	p6	p9	p5	p3	p4	p5
Cancer Patient	R1	p1	p2	p17	p24	p9	p7	p17
	R2	p5	p9	p22	p12	p13	p22	p18
	R3	p14	p11	p19	p23	p2	p13	p22
Seeds	R1	p3	p3	p6	p8	p5	p7	p4
	R2	p5	p5	p8	p4	p8	p5	p5
	R3	p6	p7	p5	p6	p2	p8	p6
Sonar	R1	p13	p47	p12	p13	p1	p5	p46
	R2	p12	p39	p34	p32	p14	p50	p51
	R3	p14	p41	p44	p51	p15	p16	p36
Heart	R1	p5	p3	p13	p1	p10	p13	p14
	R2	p8	p11	p12	p11	p7	p4	p13
	R3	p4	p12	p2	p13	p14	p5	p12
Pima Diabetes	R1	p1	p1	p5	p7	p9	p4	p9
	R2	p7	p7	p9	p5	p5	p6	p8
	R3	p9	p5	p7	p3	p2	p8	p7
Wine	R1	p8	p8	p7	p10	p10	p4	p11
	R2	p10	p10	p6	p5	p9	p9	p10
	R3	p4	p3	p5	p7	p7	p3	p9
Internet Firewall	R1	p1	p9	p5	p9	p3	p9	p1
	R2	p3	p11	p11	p11	p11	p10	p12

Data Sets	Rank	IFE-FS	FE-FS	ReliefF	Lasso	Laplacian	mRMR	Mutinffs
	R3	p2	p3	p12	p10	p7	p8	p6
	R1	p5	p1	p1	p1	p1	p1	p7
Glass	R2	p2	p8	p4	p4	p6	p9	p4
	R3	p7	p4	p2	p3	p8	p2	p8

Table 3 presents the features obtained by various filter-based methods and our proposed IFE-FS method for benchmark datasets for filtering out. Here, the rank represents order of features obtained by feature selection methods for filtering out, i.e., if the threshold number is 2 then features at rank 1 and rank 2 will be filtered out. If the threshold number is 3, then 3 features (at rank 1, rank 2 and rank 3) will be filtered out. For each dataset, we have assigned numbers to features to increase their presence in datasets. The threshold number is decided experimentally. We have tried different features and selected the number of features giving the highest accuracy. For lack of space, only the first three ranks are shown in Table. As we can see in the first row of the table (Cancer Coimbra dataset) that our proposed IFE based feature selection (IFE-FS) method obtains p5 at first rank while FE-FS (fuzzy entropy-based feature selection), ReliefF, mRMR and Mutinf obtains the same feature p8 at first rank. Lasso method obtains p9 and the Laplacian method obtains p2. For each data set three rows are mentioning first three features for filtering out.

To prove the performance of introduced technique, we have provided selected features (after removing/filtering out higher entropy features) based on our proposed IFE-FS method and six other filter methods, i.e., fuzzy entropy base feature selection(FE-FS) (Luukka, 2011), ReliefF (Eiras-Franco et al., 2021), Lasso (Coelho et al., 2020), Laplacian (He et al., 2005), mRMR (Jo et al., 2019) and Mutinf (Estevez et al., 2009) to three types of classifiers, i.e., SVM and KNN and Naïve Bias classifiers. For all experimental work, we have used MATLAB 2020. For classification, we have used Classification Learner App available in MATLAB 2020. We have used Linear SVM and Fine KNN and Naïve Bias Classifiers. Table 4(a)-4(c) summarize the accuracy of SVM, KNN and Naïve Bias classifiers with the benchmark datasets using raw data as well as with selected features. The brief introductions of these classifiers are given below:

Naïve Bias: "Naïve Bayesian classifier is a statistical supervised machine learning algorithm that predicts class membership probabilities. NB achieves high accuracy and speed when applied to a large dataset (Shah and Jivani, 2013), but it also works very well in small datasets (Delizo et al, 2020). The Naïve Bayes algorithm does not depend on the presence of other parameters, and that is why it is called naïve. Naïve Bias is a greedy classifier and has interpretability (Zaidi et al., 2013)".

Support Vector Machine: "The support vector machine (SVM) is preferred by data scientists because it can achieve good generalization performance without the need for former knowledge or experience (Chapelle et al., 2018). The SVM algorithm makes use of a hyperplane that separates the instances, putting the same classes in the same division while maximizing each group's distance in the same division while maximizing each group's distance from the dividing hyperplane".

k-Nearest Neighbors: "k-Nearest Neighbors (k-NN)" is one of the simplest and oldest supervised machine learning algorithms used in classification; it classifies a given instance via the majority of the classes among its k-nearest neighbors found in the dataset (Sun and Huang, 2010). This algorithm relies on the distance metric used to determine the nearest neighbors of the given instance. The two primary benefits of the k-Nearest Neighbor algorithm are efficiency and flexibility".

It is evident from Table 4(a) that the classification accuracy of the SVM classifier with selected features obtained by our proposed method can be enhanced or kept by in all of the cases (for all benchmark datasets) compared with the raw data. For 60 % of benchmark datasets, SVM has achieved highest or similar to highest accuracy with selected features using our proposed method except (40% cases) Cancer Coimbra, Sonar, Heart and Glass datasets, e.g., for Cancer Coimbra, the highest accuracy is 76.7% (Lasso method). In comparison, in our case, it is 72.4%. For the Heart dataset highest accuracy is 83.2% (Laplacian Method), while in our case it is 82.8%. For Sonar dataset, the highest accuracy is 78.4% (mRMR method) while it is 76%. For the Glass dataset highest accuracy is 64.5% (mRMR method), while in our case, it is 64%. Moreover, the average accuracy obtained by the SVM classifier is highest, i.e., 78.83%using selected features by our proposed method.

Table 4: Comparative results showing performance of SVM and KNN and Naïve Bias Classifiers

4(a) Performance of SVM with raw data and with selected features obtained by various filter methods

Accuracy with SVM classifier									
S No.	Data Set	Raw data	IFE-FS	FE-FS	ReliefF	Lasso	Laplacian	mRMR	Mutinf
1	Cancer Coimbra	69.00	72.4	69.8	71.6	76.7	75	67.2	63.8
2	Breast Tissue	65.3	70.3	66.3	69.3	63.4	69.3	66.3	67.3
3	Cancer patient	100	100	98.4	100	99.9	98.1	100	100
4	Seeds	92.4	94.8	88.1	94.3	89.5	89.5	94.8	85.7
5	Sonar	74	76	75.5	76.4	76.9	74	78.4	75.5
6	Heart	77.2	82.8	76.9	79.2	83.2	83.2	81.2	79.2
7	Pima diabetes	77.2	77.5	77.5	77.1	76.8	77.1	76.8	77.2
8	Wine quality red	58.1	58.1	58.1	57.1	56.7	56.9	57.6	55
9	Internet firewall	91.1	92.4	85.9	85.9	80.2	86	86	92.2
10	Glass	63.1	64	61.7	61.7	61.7	63.1	64.5	61.7
	Average	76.74	78.83	75.82	77.26	76.50	77.22	77.28	75.76

4(b) Performance of KNN classifier with raw data and with selected features obtained by various filter methods

Accuracy with KNN classifier									
S No.	Data Set	Raw data	IFE-FS	FE-FS	ReliefF	Lasso	Laplacian	mRMR	Mutinf
1	Cancer Coimbra	64.7	69	64.7	64.7	66.4	65.5	64.7	56.9
2	Breast tissue	69.3	71.3	62.4	69.3	65.3	67.3	67.3	67.3
3	Cancer patient	100	100	100	100	100	100	100	100
4	Seeds	91.4	92.4	90	90.5	88.1	88.1	93.3	83.3
5	Sonar	85.6	85.6	84.6	85.1	86.5	84.4	88.9	85.6
6	Heart	76.6	80.9	69.3	64	74.9	73.3	71	67.7
7	Pima diabetes	70.3	71.6	71.6	68	69.7	68	71.6	68.5
8	Wine quality red	62.2	62.8	62.8	62	62.5	62.5	61.9	62.3
9	Internet firewall	72.5	84.8	72.4	78.9	68.2	78.9	68.3	72.6
10	Glass	71.5	72.9	65.4	66.8	66.8	66.8	72.9	66.8
	Average	76.41	79.13	74.32	74.93	74.84	75.48	75.99	73.10

4(c) Performance of Naïve Bias classifier with raw data and with selected features obtained by various filter methods

| S No.	Data Set	Raw data	IFE-FS	FE-FS	ReliefF	Lasso	Laplacian	mRMR	Mutinf
				Accuracy with Naïve Bias classifier					
1	Cancer Coimbra	62.9	67.2	62.1	62.1	68.1	60.3	60.3	58.6
2	Breast tissue	61.4	68.3	63.4	58.4	60.4	61.4	61.4	61.4
3	Cancer patient	89.1	90.8	89	89.1	85.5	90	90.1	87.1
4	Seeds	90	92.9	89	92.4	89	88.1	92.4	92.4
5	Sonar	67.3	68.3	64.4	66.8	68.3	66.8	64.9	68.8
6	Heart	82.5	82.5	77.2	75.2	78.5	74.6	75.9	72.9
7	Pima diabetes	75.7	77.2	76.4	75.4	75.7	69.8	74.7	73.3
8	Wine quality red	55	56	46.5	45.7	48.8	54.4	55.2	47.8
9	Internet firewall	78.6	81.2	69.7	78.8	70.4	68.3	78	75.3
10.	Glass	61.2	65.9	58.9	58.9	58.9	58.9	58.9	53.7
	Average	72.37	75.03	69.66	70.28	70.36	69.26	71.18	69.13

It is evident from Table 4(b) that the classification accuracy of the KNN classifier with selected features obtained by our proposed method can be enhanced or kept in all of the cases (for all benchmark datasets) compared with the raw data. For 80% of benchmark datasets, KNN has achieved the highest or similar to highest accuracy with selected features using our proposed method except for (20% benchmark datasets) Seeds and Sonar datasets, e.g., for seeds dataset, the highest accuracy is 93.3% (mRMR method). In comparison, in our case, it is 92.4%. For the Sonar dataset highest accuracy is 83.2% (Laplacian Method), while in our case, it is 82.8%. For the Sonar dataset highest accuracy is 88.9% (mRMR method), while in our case, it is 85.6%. However, the average accuracy obtained by the KNN classifier is highest, i.e., 79.13% using selected features by our proposed method.

It is evident from Table 4(c) that the classification accuracy of the Naïve Bias classifier with selected features obtained by our proposed method can be enhanced or kept in all of the cases (for all benchmark datasets) when compared with the raw data. For 80% of benchmark datasets, KNN has achieved highest or similar to highest accuracy with selected features using our proposed method except for (20% benchmark datasets) Cancer Coimbra and Sonar datasets, e.g., for Cancer Coimbra dataset, the highest accuracy is 68.1% (Lasso method). In comparison, in our case, it is 67.2%. For the Sonar dataset highest accuracy is 68.8% (Mutinf Method), while in our case, it is 68.3%. However, the average accuracy obtained by the Naïve Bias classifier is highest, i.e., 75.03% using selected features by our proposed method.
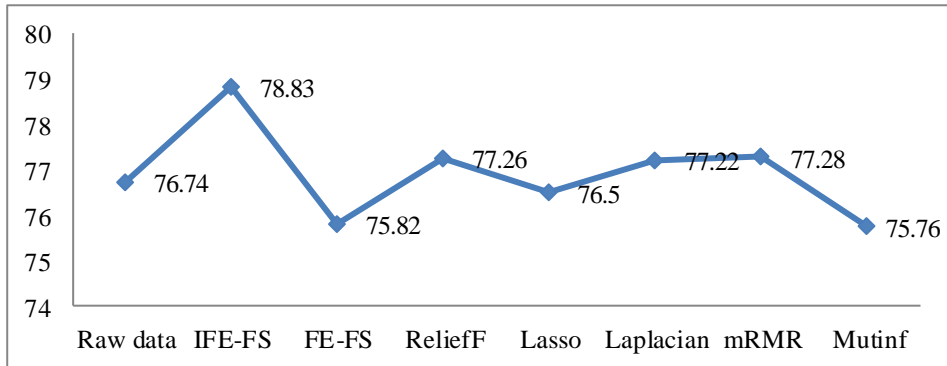
Figure 2(a): Average classification accuracy with SVM
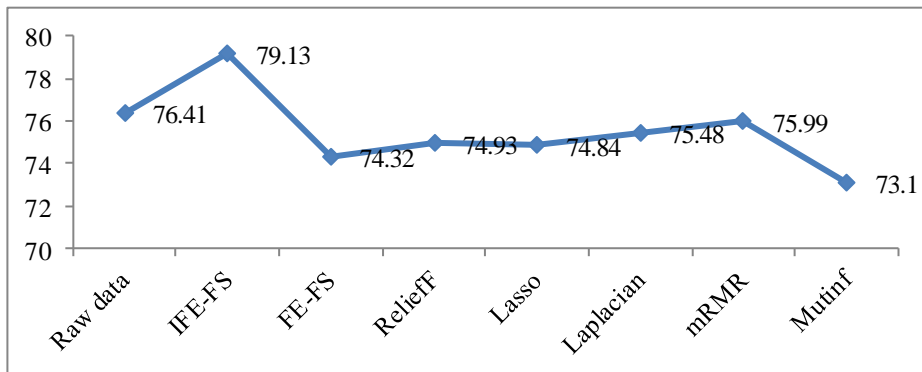
Figure 2(b). Average classification accuracy with KNN
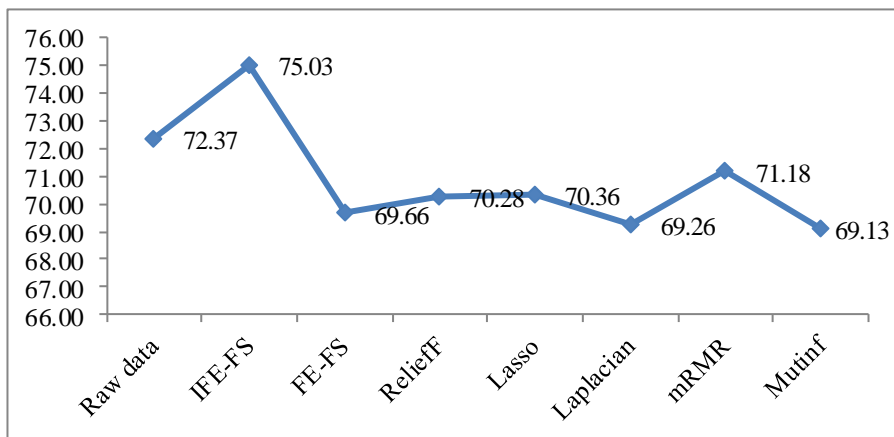
Figure 2(c). Average classification accuracy with Naïve Bias

Figure 2. Average classification accuracy of the developed technique with different classifiers

Figures 2(a), 2(b) and 2(c) present the average accuracy obtained by SVM, KNN and Naïve Bias classifiers, respectively. It is evident that the highest average accuracy has been obtained by utilizing the proposed IF-entropy for feature selection in comparison to different filter-based feature selection methods.

6. Conclusions

This study proposes a new intuitionistic fuzzy entropy-based algorithm for feature selection before classification tasks in an information system. For this purpose, new intuitionistic fuzzy entropy has been developed to measure feature entropy (uncertainty) as parameters for feature selection. The experimental results are compared with the existing filter-based technique and prove that the proposed technique perfectly fits the hidden information. Afterwards, the developed model was executed on ten real benchmark datasets. Classification accuracies for diverse classifiers, i.e., SVM, KNN and Naïve Bias, were evaluated. Based on experimental results, it has been acquired that the average accuracy was 78.83% with the SVM classifier using selected feature obtained by our proposed IFE-FS, which is the highest among all the accuracies which have been obtained using selected features obtained by other filter methods i.e., with fuzzy entropy based feature selection (75.82%), ReliefF (77.26%), Lasso (76.50%), Laplacian (77.22%), mRMR (77.28%) and Mutinf (75.76%), while with raw data it was 76.77%. In addition, we have noticed that the average performance of the SVM classifier with features obtained by the proposed IFE-FS model obtains an increment of 2.09%, which is the highest increment (with Relief (0.52%), with Laplacian (0.48%), with mRMR (0.54%) compared to the raw data, while with the fuzzy entropy-based feature selection algorithm, with Lasso feature selection method and with mutinfit obtains decrement of 0.24% and 0.92% and 0.98% respectively. The average accuracy was 79.41% with the KNN classifier using selected feature obtained by proposed IFE-FS which is highest among all the accuracies which have been obtained using selected features obtained by other filter methods, i.e., with fuzzy entropy-based feature selection (74.32%), ReliefF (74.93%), Lasso (74.83%), Laplacian (75.48%), mRMR (75.99%) and Mutinf (73.10%), while with raw data it was 76.41%. In addition, we have noticed that the average performance of KNN classifier with selected features obtained by the proposed IFE-FS method obtains an increment of 2.72%, while with selected features obtained by other filter methods, i.e., Fuzzy entropy-based method, Relief, Lasso, Laplacian, mRMR, and Mutinf it is getting decrement of 2.09%, 1.48%, 1.57%, 0.93%, 0.42% and 3.31% respectively, compared to the raw data. A similar observation has been obtained with the Naïve Bias classifier also. It has been found that the reduced datasets perform quite well than unreduced datasets in terms of resulting classification accuracy. While comparing FE-FS and some other filter filter-based feature selection methods, we have obtained different features at the same rank. While doing feature selection by filtering out features with higher entropies and providing remaining features to the classifiers, we have acquired better accuracy in most cases with our proposed method than other filter-based feature selection methods. Thus, we can conclude that our technique is more efficient in dealing with uncertainties of features and is perfectly fit to locate the hidden information while doing FS.

The certain limitations of the developed framework are important to be aware of. A practical difficulty is that the experts must be trained with the preference style to properly use the flexibility and potential of intuitionistic fuzzy numbers. In addition, this work has limitation in dealing with more uncertain decision-making problems

because of the constraint condition of intuitionistic fuzzy set. Future research studies will try to handle the limitations of this work. Moreover, it would be exciting to use the presented model for feature selection with high-dimensional regression in the future. Also, we can generalize the introduced technique under different uncertain contexts such as spherical set, Pythagorean fuzzy set, neutrosophic set, q-rung orthopair fuzzy rough set and linguistic generalized orthopair fuzzy set.

References

Álvarez, J.D., Matias-Guiu, J.A., Cabrera-Martín, M.N., Risco-Martín, J.L., & Ayala, J.L. (2019). An application of machine learning with feature selection to improve diagnosis and classification of neurodegenerative disorders. BMC Bioinformatics, 20, 01-12.

Ansari, M.D., Mishra, A.R., & Ansari, F.T. (2018). New divergence and entropy measures for intuitionistic fuzzy sets on edge detection, International Journal of Fuzzy Systems, 20(2), 474-487.

Aremu, O.O., Cody, R.A., Hyland-Wood, D., & McAree, P.R. (2020). A relative entropy based feature selection framework for asset data in predictive maintenance. Computers & Industrial Engineering, 145, 106536, https://doi.org/10.1016/j.cie.2020.106536.

Atanassov, K.T. (1986). Intuitionistic fuzzy sets. Fuzzy Sets and Systems, 20(1), 87-96.

Bahassine, S., Madani, A., Al-Sarem, M., & Kissi, M. (2020). Feature selection using an improved Chi-square for Arabic text classification. Journal of King Saud University-Computer and Information Sciences, 32(2), 225-231.

Bairagi, B. (2022). A new framework for green selection of material handling equipment under fuzzy environment. Decision Making: Applications in Management and Engineering. https://doi.org/10.31181/dmame0313052021b.

Blum, A., & Langley, P. (1997). Selection of relevant features and examples in machine learning. Artificial Intelligence, 97, 245–271.

Bustince, H., & Burillo, P. (1996). Vague sets are intuitionistic fuzzy sets. Fuzzy Sets and Systems, 79, 403-405

Chapelle, O., Haffner, P., & Vapnik, V.N. (2018). Support Vector Machines for Histogram-Based Image Classification. IEEE Transactions on Neural Networks, 10, 1055–1064.

Chen J., Yuan S.H., Lv D., & Yang, X. (2021). A novel self-learning Feature selection approach based on feature attributions. Expert Systems with Applications, https://doi.org/10.1016/j.eswa.2021.115219.

Coelho, F., Costa, M., Verleysen, M., & Braga, A.P. (2020). LASSO multi-objective learning algorithm forfeature selection. Soft Computing 24, 13209–13217.

Delizo, J.P.D., Abisado, M.B., & Trinos, M.I.D. (2020) Philippine twitter sentiments during covid-19 pandemic using multinomial naïve-bayes. International Journal of Advanced Trends in Computer Science and Engineering, 9, 408–412.

Dua, D., & Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

Eiras-Franco, C., Guijarro-Berdiñas, B., Alonso-Betanzos, A., & Bahamonde, A. (2021). Scalable feature selection using ReliefF aided by locality-sensitive hashing. International Journal of Intelligent Systems, 36, 6161-6179.

Estevez, P., Tesmer, M., Perez, C., & Zurada, J.M. (2009). Normalized Mutual Information Feature Selection. IEEE Transactions on Neural Networks, 20(2), 189-201.

Gao, L., & Wu, W. (2020). Relevance assignation feature selection method based on mutual information for machine learning. Knowledge-Based Systems, 209, 106439, doi:10.1016/j.knosys.2020.106439.

He, X.F., Cai, D., & Niyogi, P. (2005). Laplacian score for feature selection. Proceedings of the 18th International Conference on Neural Information Processing Systems, 5-8 December 2005, 507-514.

Hezam, I.M., Mishra, A.R., Rani, P., Cavallaro, F., Saha, A., Ali, J., Strielkowski, W., & Štreimikienė, D. (2022). A hybrid intuitionistic fuzzy-MEREC-RS-DNMA method for assessing the alternative fuel vehicles with sustainability perspectives. Sustainability 14, 1-32.

High, R., Eyres, G. T., Bremer, P., & Kebede, B. (2021). Characterization of blue cheese volatiles using finger printing, self-organizing maps, and entropy-based feature selection. Food Chemistry 347, 128955, https://doi.org/10.1016/j.foodchem.2020.128955.

Jaganathan, P., & Kuppuchamy, R. (2013). A threshold fuzzy entropy based feature selection for medical database classification. Computer in Biology and Medicine, 43(12), 2222–2229.

Ji, B., Lu, X., Sun, G., Zhang, W., Li, J., & Xiao, Y. (2020). Bio-Inspired Feature Selection: An Improved Binary Particle Swarm Optimization Approach. IEEE Access, 8, 85989-86002.

Jo, I., Lee, S., & Oh, S. (2019). Improved Measures of Redundancy and Relevance for mRMR Feature Selection. Computers, 8, 01-15.

Kim, K., & Zzang, S.Y. (2019). Trigonometric comparison measure: A feature selection method for text categorization. Data & Knowledge Engineering, 119, 1-21.

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. Artificial Intelligence, 7, 273–323.

Kumar, S., Maity, S. R., & Patnaik, L. (2022). Optimization of Wear Parameters for Duplex-TiAlN Coated MDC-K Tool Steel Using Fuzzy MCDM Techniques. Operational Research in Engineering Sciences: Theory and Applications. https://doi.org/10.31181/110722105k.

Kushwaha, D. K., Panchal, D., & Sachdeva, A. (2020). Risk analysis of cutting system under intuitionistic fuzzy environment. Reports in Mechanical Engineering, 1(1), 162-173.

Lee H. M., Chen, C. M., Chen, J. M., & Jou, Y. L. (2001). An efficient fuzzy classifier with feature selection based on fuzzy entropy. IEEE Transactions on Systems, Man and Cybernetics-Part B: Cybernetics, 31(3), 425-432.

Lohrmann, C., Luukka, P., Jablonska-Sabuka, M., & Kauranne, T. (2018). A combination of fuzzy similarity measures and fuzzy entropy measures for supervised feature selection. Expert Systems with Applications, 110, 216-236.

Luukka, P. (2011). Feature selection using fuzzy entropy measures with similarity classifier Expert Systems with Applications, 38, 4600–4607.

Mishra, A.R. (2016). Intuitionistic fuzzy information with application in rating of township development. Iranian Journal of Fuzzy Systems, 13, 49–70

Mishra, A.R., Jain, D., & Hooda, D.S., (2017a). Exponential intuitionistic fuzzy information measure with assessment of service quality. International Journal of Fuzzy Systems, 19(3), 788-798.

Mishra, A.R., Kumari, R., & Sharma, D.K. (2019a). Intuitionistic fuzzy divergence measure-based multi-criteria decision-making method. Neural Computing and Applications 31, 2279-2294.

Mishra, A.R., & Rani, P. (2019). Shapley divergence measures with VIKOR method for multi-attribute decision making problems. Neural Computing and Applications 31(2), 1299-1316.

Mishra, A.R., Rani, P., & Jain, D. (2017b) Information measures based TOPSIS method for multicriteria decision making problem in intuitionistic fuzzy environment. Iranian Journal of Fuzzy Systems 14(6), 41-63.

Mishra, A.R., Singh, R.K., & Motwani, D. (2019b). Multi-criteria assessment of cellular mobile telephone service providers using intuitionistic fuzzy WASPAS method with similarity measures. Granular Computing, 4, 511-529,

Mishra, A.R., Singh, R.K., & Motwani, D. (2020b). Intuitionistic fuzzy divergence measure-based ELECTRE method for performance of cellular mobile telephone service providers. Neural Computing and Applications, 32, 3901-3921.

Mishra, A.R., Sisodia, G., Pardasani, K.R., & Sharma, K. (2020a). Multicriteria IT personnel selection on intuitionistic fuzzy information measures and ARAS methodology. Iranian Journal of Fuzzy Systems, 17(4), 55-68.

Murugesan, S., Bhuvaneswaran, R.S., Nehemiah, H.K., Sankari, S.K., & Jane, Y.N. (2021) Feature selection and classification of clinical datasets using bioinspired algorithms and super learner. Computational and Mathematical Methods in Medicine, Article ID 6662420, https://doi.org/10.1155/2021/6662420.

Omuya, E.O., Okeyo, G.O., & Kimwele, M.W. (2021). Feature selection for classification using principal component analysis and information gain. Expert Systems with Applications, 174, 114765, https://doi.org/10.1016/j.eswa.2021.114765.

Parlak, B., & Uysal, A.K. (2021). A novel filter feature selection method for text classification: Extensive Feature Selector. Journal of Information Science. doi:10.1177/0165551521991037.

Pintas, J.T., Fernandes, L.A.F., & Garcia, A.C.B. (2021). Feature selection methods for text classification: a systematic literature review. Artificial Intelligence Review, 54, 6149–6200.

Precup, R.-E., Preitl , S., Petriu, E., Bojan-Dragos , C.-A., Szedlak-Stinean, A.-I., Roman, R.-C., & Hedrea, E.-L. (2020). Model-Based Fuzzy Control Results for Networked Control Systems. Reports in Mechanical Engineering, 1(1), 10-25.

Qu, Y., Li, R., Deng, A., Shang, C., & Shen, Q. (2020). Non-unique decision differential entropy-based feature selection. Neurocomputing, 393, 187-193.

Rahimi, M., Kumar, P., Moomivand, B., & Yari, G. (2021). An intuitionistic fuzzy entropy approach for supplier selection. Complex &Intelligent Systems, 7, 1869–1876.

Rangasamy, P. (2021). Intuitionistic Fuzzy (IF) Logic Toolbox (https://www.mathworks.com/matlabcentral/fileexchange/68198-intuitionistic-fuzzy-if-logic-toolbox), MATLAB Central File Exchange. Retrieved JANUARY 30, 2021.

Rani, P., & Jain, D. (2017). Intuitionistic fuzzy PROMETHEE technique for multi-criteria decision making problems based on entropy measure In: Singh, M., Gupta, P., Tyagi, V., Sharma, A., Ören, T., Grosky, W. (eds) Advances in Computing and Data Sciences. ICACDS 2016. Communications in Computer and Information Science, 721, 290–301.

Rani, P., Mishra, A. R., Ansari, M. D., & Ali, J. (2021). Assessment of performance of telecom service providers using intuitionistic fuzzy grey relational analysis framework (IF-GRA). Soft Computing, 25, 1983-1993.

Rehman, A., Javed, K., & Babri, H. A. (2017). Feature selection based on a normalized difference measure for text classification. Information Processing & Management, 53(2), 473-489.

Revanasiddappa, M.B., & Harish, B.S. (2018). A new feature selection method based on intuitionistic fuzzy entropy to categorize text documents. International Journal of Interactive Multimedia and Artificial Intelligence, 5(3), 106-117.

Rostami, M., Berahmand, K., & Forouzandeh, S. (2021). A novel community detection based genetic algorithm for feature selection. Journal of Big Data, 8, 2, https://doi.org/10.1186/s40537-020-00398-3

Ruan, F., Hou, L., Zhang, T., & Li, H. (2021). A novel hybrid filter/wrapper method for feature selection in archaeological ceramics classification by laser-induced break down spectroscopy. Analyst, 146, 1023-1031, DOI: 10.1039/d0an02045a.

Shah, C., & Jivani, (2013). A Comparison of Data Mining Classification Algorithms for Breast Cancer Prediction. 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), pp. 1-4, https://doi.org/10.1109/ICCCNT.2013.6726477.

Singh, S., Shreevastava, S., Som, T., & Jain, P. (2019). Intuitionistic Fuzzy Quantifier and Its Application in Feature Selection. International Journal of Fuzzy Systems, 21, 441-453.

Sun, L., Wang, L., Ding, W., Qian, Y., & Xu, J. (2021). Feature selection using fuzzy neighborhood entropy-based uncertainty measures for fuzzy neighborhood multigranulation rough sets. IEEE Transactions on Fuzzy Systems, 29(1), 19-33.

Sun, S., & Huang, R. (2010). An Adaptive k-Nearest Neighbor Algorithm. 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery, pp. 91-94, https://doi.org/10.1109/FSKD.2010.5569740.

Sun, X., Liu, Y.H., Li, J.Q., Zhu, J.Q., Liu, X., & Chen, H. (2012). Using cooperative game theory to optimize the feature selection problem. Neurocomputing, 97(15), 86-93.

Sun, L., Wang, L., Xu, J., & Zhang, S. (2019). A Neighborhood Rough Sets-Based Attribute Reduction Method Using Lebesgue and Entropy Measures. Entropy. 21(2), 138. https://doi.org/10.3390/e21020138

Szmidt, E., & Kacprzyk, J. (2001). Entropy for Intuitionistic Fuzzy Sets. Fuzzy Sets and Systems, 118, 467-477.

Tang, B., Kay, S., & He, H. (2016).Toward optimal feature selection in naive Bayes for text categorization. IEEE Transactions on Knowledge and Data Engineering, 28(9), 2508-2521.

Thakkar, A., & Chaudhari, K. (2020). Predicting stock trend using an integrated term frequency–inverse document frequency-based feature weight matrix with neural networks. Applied Soft Computing, 96, 106684, https://doi.org/10.1016/j.asoc.2020.106684.

Thao, N. X., & Chou, S. Y. (2022). Novel similarity measures, entropy of intuitionistic fuzzy sets and their application in software quality evaluation. Soft Computing 26, 2009–2020.

Tiwari, A. K., Shreevastava, S., Subbiah, K., & Som, T. (2019). An intuitionistic fuzzy-rough set model and its application to feature selection. Journal of Intelligent & Fuzzy Systems, 36, 4969-4979.

Tran, M.-Q., Elsisi, M., & Liu, M.-K. (2021). Effective feature selection with fuzzy entropy and similarity classifier for chatter vibration diagnosis. Measurement, 184, 109962, https://doi.org/10.1016/j.measurement.2021.109962.

Tripathi, D., Nigam, S. K., Mishra, A. R., & Shah, A. R. (2022a). A Novel Intuitionistic Fuzzy Distance Measure-SWARA-COPRAS Method for Multi-Criteria Food Waste Treatment Technology Selection. Operational Research in Engineering Sciences: Theory and Applications. https://doi.org/10.31181/oresta111022106t

Tripathi, D.K., Nigam, S.K., Rani, P., & Shah, A.R. (2022b). New intuitionistic fuzzy parametric divergence measures and score function-based CoCoSo method for decision-making problems. Decision Making: Applications in Management and Engineering. https://doi.org/10.31181/dmame0318102022t

Vlachos, I. K., & Sergiadis, G.D. (2007). Intuitionistic fuzzy information-applications to pattern recognition, Pattern Recognition Letters, 28, 197-206.

Wei, C.P. Gao, Z.H., & Guo, T.T. (2012). An intuitionistic fuzzy entropy measure based on the trigonometric function. Control and Decision, 27, 571-574

Zadeh, L.A. (1965), Fuzzy sets. Information & Control, 8, 338–353.

Zaidi, N., Cerquides, J., Carman, M., & Webb, G. (2013). Alleviating Naive Bayes Attribute Independence Assumption by Attribute Weighting. The Journal of Machine Learning Research, 14, 1947-1988.

Zhang, X., Mei, C., Chen, D., & Li, J. (2016.) Feature selection in mixed data: A method using a novel fuzzy rough set-based information entropy, Pattern Recognition 56, 1-15, https://doi.org/10.1016/j.patcog.2016.02.013

Zhang, Q. S., & Jiang, S.Y. (2008). A note on information entropy measures for vague sets and its applications. Information Sciences, 178, 4184-4191.

Zhang, X., Fan, Y., & Yang, J. (2021a). Feature selection based on fuzzy-neighborhood relative decision entropy. Pattern Recognition Letters, 146, 100-107, https://doi.org/10.1016/j.patrec.2021.03.001.

Zhang, Y., Huang, F., Deng X., & Jiang, W. (2021b). A New Total Uncertainty Measure from A Perspective of Maximum Entropy Requirement. Entropy, https://doi.org/10.3390/e23081061.

Zhao, J., Liang, J.-M., Dong, Z.-N., Tang, D.-Y., & Liu, Z. (2020). Accelerating information entropy-based feature selection using rough set theory with classified nested equivalence classes. Pattern Recognition, 107, 107517, https://doi.org/10.1016/j.patcog.2020.107517.