International Journal of Engineering Research and Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 22, Issue 6 (June 2026), PP 101-109

A Jensen-Shannon fuzzy divergence measure and its application in Pattern Recognition

Tamanna Sethi
Panipat Institute of Engineering and Technology
Department of Applied Sciences and Himanities
Samalkha, Panipat,Haryana, India

Abstract
The requirement of appropriate divergence measures emerge as they play a vital part in segregation of two likelihood disseminations. The display communication is committed to the presentation of one such divergence measure utilizing Jensen inequality and Shannon entropy and its approval. On basis of the proposed divergence measure, a new dissimilarity measure is presented. Other than building up approval, a few of its major properties are moreover presented. Advance, based on proposed dissimilarity measure ,a new multiple attribute decision making method is presented and is altogether clarified with the assistance of an outlined case. In last paper is briefed with an application of the proposed dissimilarity measure in pattern recognition.

***Keywords**: fuzzy set, divergence measure, Jensen measure, pattern recognition*

Date of Submission: 05-06-2026 Date of acceptance: 16-06-2026

I. Introduction

A prominent role plays by divergence measure in measuring the distinction within two probability distributions. This measure has its own superiorities. The divergence value changes with change in parameters.The parameters do a favorablepart in explaining the divergence measure. In such wise, how the divergence is affected by parameters is an immersing content to learn. The idea to study the distance measure within two probability distributions was firstly introduced by Mahalanobis (1936). Based on Shannon entropy, the notion of divergence measure proposed by Kullback and Leibler (1951). Thereafter, Rao (1982)proposed variousdivergence measures from their perspective. To compare the effectiveness of two events the point of divergence measure is also determined. That's why, from the above argument, consequence of divergence measures could be simply determined. Before the initiation of theory of fuzzy set given by Zadeh (1965), to measure the uncertainty, the tool was probability theory.To enhance the researchers Zadeh (1965) introduced the fuzzy set theory with a new tool of estimating the uncertainty or fuzziness. In Zadeh theory, in universal set every element is allotted a membership degree whose value lies between 0 and 1. Associatedto fuzzy set theorythere is a prominent concept in Fuzzy entropy which is helpful to calculate the amount of fuzziness created due to vagueness. Here we check either an element belongs to given set or does not belong to set. Afterwards, the theory of Fuzzy entropy axiomatized by De Luca and Termini (1962). In numerous real-life problems serving as clustering, image segmentation, medical diagnosis, decision making, pattern recognition, bacteria detection, image segmentation, and financial diagnosis etc. fuzzy sets used.

The constitution on solicitation and advancement of divergence information measures between fuzzy sets has enlargedmeasurably in past few years; still, an augmentation that the superior divergence measures could be construct which will helpful in applications of multiplicity areas. By taking into account the all-inclusive studies and importance of divergence measures could be prominently judged. On the basis of circumstances, different divergence measures are required. Therefore, just complete the requirement yet efficiency that divergence measures are not requiredneed is to research new divergence measures. By using this idea, Lin(1991) proposed Jensen–Shannon divergence, a new similarity measure proposedwhich is based on Jensen–Shannon fuzzy measure. Various properties are also examined of this proposed divergence. After that, influence of generalized fuzzy divergence which is proposed measure has been exhibited by its applications applying on pattern recognition and compares the result with exiting measures.

II. Jensen–Shannon Divergence Measure

We study the hypothetical concept of information theory in this part and after that instigate the Jensen–Shannon divergence, we explain the properties in probabilistic manner.

Let us consider a set $\Delta_t = \{L = (l_1, l_2, \dots, l_t): l_i \geq 0,; \sum_{i=1}^{t} l_i = 1\}, t \geq 2$ of complete probability distributions. For the probability distribution L $=(l_1, l_2, \dots, l_t) \in \Delta_t$, the famous Shannon's entropy is explained as:

$H(L) = -\sum_{i=1}^{t} l_i \log(l_i)$ (1)

Cover and Thomas (1991) stated that the cross entropy depend on observed probability distribution $Q = (q_1, q_2 \dots q_t)$ as related to true probability distribution L $=(l_1, l_2, \dots, l_t)$ as follows:

$$H(L:Q) = \sum l_i \log\left(\frac{1}{q_i}\right) = -\sum l_i \log(q_i) \quad (2)$$

Corresponding to (1), for any $L, Q \in \Delta_t$, divergence measure explained by Kullback and Leibler (1951)as

$KL(L,Q) = \sum_{i=1}^{t} l_i \log\left(\frac{l_i}{q_i}\right) = H(L,Q) - H(L)$ (3)

Where l_i and q_i are the probability of occurrence for probability distributions, respectively. It is essential that L need to be absolutely continuous with respect to$Q$ i.e.l = 0 when$q$= 0. To get over this limitation, another divergence discussed by Lin (1991) for the probability distribution L and Qas

$P(L,Q) = \sum_{i=1}^{t} l log\left(\frac{l}{\frac{l+q}{2}}\right) = KL(L, \frac{L+Q}{2})$ (4)

Based on Ia symmetric divergence measure can be explained by

$$I(L,Q) = \frac{P(L,Q) + P(Q,L)}{2} = H\left(\frac{L+Q}{2}\right) - \left(\frac{H(L)+H(Q)}{2}\right)$$

Now, we have presumed that the probability distributions L and $Q$with the like weights 1/2. For the probability distributions L and Q,$\alpha_1, \alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$ is random weights. Jensen–Shannon divergence is direct generalization of I:

$JS(L,Q) = H(\alpha_1 L + \alpha_2 Q) - (\alpha_1 H(L) + \alpha_2 H(Q))$ (5)

Jensen Inequality$JS(L, Q)$ is non negative and when $L = Q$ then it will be zero. We can allocate weights according to their significance to probability distributions. It is major benefits of Jensen-Shannon inequality. in decision-making problem this property assists. As a resulton the requirement of symmetric divergence measuresbased on Shannon entropy and Jensen inequality, introduced a new idea named as Jensen–Shannon divergence defined as.

Take two probability distributions E, $G \in \Delta_t$withα_1 andα_2 are their respective weights fulfilling the condition$\alpha_1 + \alpha_2 = 1$, new divergence measure introduced

$SHN_\alpha(E,G) = -\sum_{i=1}^{t}(-(e_i \log(e_i) + g_i \log(g_i)) + (\alpha_1 e_i + \alpha_2 g_i)\log(\alpha_1 e_i + \alpha_2 g_i) + (\alpha_1 g_i + \alpha_2 e_i)\log(\alpha_1 g_i + \alpha_2 e_i))$(6)

Propertiesof $SHN_\alpha(E,G)$

Few main properties of above divergence measure are

(1) $SHN_\alpha(E,G) \geq 0$ under condition $SHN_\alpha(E,G) = 0$ when $E = G$

(2) $SHN_\alpha(E,G)$is convex function of E and G

(3) When E and G are degenerate distributions then $SHN_\alpha(E,G)$ acquire its maximum value

Proof: To prove $SHN_\alpha(E,G) \geq 0$.
As concave function in (1), for any E, $G \in \Delta_t$ Jensen inequality implies

$\alpha_1 H(E) + \alpha_2 H(G) \leq H(\alpha_1 E + \alpha_2 G)$ (7)

which gives

$$-\alpha_1 H(E) - \alpha_2 H(G) \geq -H(\alpha_1 E + \alpha_2 G); \tag{8}$$

From (8), we get

$$-\alpha_1\left(-\sum_{i-1}^{t} e_i \log(e_i)\right) - \alpha_2\left(-\sum_{i=1}^{t} g_i \log(g_i)\right) \geq -\left(-\sum_{i=1}^{t}(\alpha_1 e_i + \alpha_2 g_i) log(\alpha_1 e_i + \alpha_2 g_i)\right) \tag{9}$$

Similarly

$$-\alpha_2\left(-\sum_{i-1}^{t} e_i \log(e_i)\right) - \alpha_1\left(-\sum_{i=1}^{t} g_i \log(g_i)\right) \geq -\left(-\sum_{i=1}^{t}(\alpha_1 e_i + \alpha_2 g_i) log(\alpha_1 e_i + \alpha_2 g_i)\right) \tag{10}$$

Adding (9) and (10), we get

$$\begin{aligned} SHN_\alpha(E,G) = -\sum_{i=1}^{t}\Big(&-(e_i \log(e_i) + g_i \log(g_i)) + (\alpha_1 e_i + \alpha_2 g_i) \log(\alpha_1 e_i + \alpha_2 g_i) \\ &+ (\alpha_1 g_i + \alpha_2 e_i) \log(\alpha_1 g_i + \alpha_2 e_i)\Big) \geq 0 \end{aligned} \tag{11}$$

This explained the non-negativity of SHN_α
From (6) we can directly prove properties 2 and 3

Next, weestablished a new divergence measure as generalize the idea from probabilistic to fuzzy set theory. This measure is related to Jensen–Shannondivergence measure.

III. Jensen- Shannon Fuzzy Divergence Measure

A concise discussion on definitions related to FSs and proposed Jensen-Shannon Fuzzy Divergence measure is given in this segment

Definition 3.1 For fuzzy set in χ (fixed set), construct

$$A = \{(t_p, \mu_A(t_p)) | t_p \in \chi\}, \tag{12}$$

in which$\mu_A : \chi \rightarrow [0, 1]$ is a membership function in χ and$\mu_A(t_p) \in [0, 1]$. Then A is known as fuzzy set.

Definition 3.2 Consider a universe of discourse= $\{k_1, k_2, \dots, k_t\}$. Explain fuzzy divergence measure $SHN(S,T)$for $S,T \in FS(X)$as

$$\begin{aligned} SHN(S,T) = -\sum_{i=1}^{t}\Big(&(\alpha_1\mu_S(k_i) + \alpha_2\mu_T(k_i))(log(\alpha_1\mu_S(k_i) + \alpha_2\mu_T(k_i))) + (\alpha_1\mu_T(k_i) + \\ &\alpha_2\mu_S(k_i))(log(\alpha_1\mu_T(k_i) + \alpha_2\mu_S(k_i))) + (\alpha_1\nu_S(k_i)) + \alpha_2(\nu_T(k_i)))(log(\alpha_1\nu_S(k_i)) + \alpha_2(\nu_T(k_i)))) + \\ &(\alpha_1\nu_T(k_i)) + \alpha_2(\nu_S(k_i)))(log(\alpha_1\nu_T(k_i)) + \alpha_2\nu_S(k_i)))) - (\mu_S(\mathrm{k_i})\log(\mu_S(k_i)) + (\mu_T(\mathrm{k_i})\log(\mu_T(k_i)) - \\ &(\nu_S(\mathrm{k_i}))\log(\nu_S(k_i))) + (\nu_T(\mathrm{k_i}))\log(\nu_T(k_i)))\Big) \end{aligned} \tag{13}$$

3.3 Justification

At that moment, a main logical question leaps to mind about the proposed measure validation.

Proof: Firstly we will prove the non-negativity condition for proposed measure.
As a concave function (1), by the use of Jensen inequality, we get

$$-\alpha_1(-\sum_{i=1}^{t}(\mu_S(k_i)\log(\mu_S(k_i)) + (\nu_S(k_i))\log(\nu_S(k_i)))) - \alpha_2(-\sum_{i=1}^{n}(\mu_T(k_i)\log(\mu_T(k_i)) + (\nu_T(k_i))\log(\nu_T(k_i)))) \geq -\left(-\sum_{i=1}^{n}((\alpha_1(\mu_S(k_i) + (\nu_S(k_i))) + \alpha_2\left(\mu_T(k_i) + (\nu_T(k_i))\right)(\log(\alpha_1(\mu_S(k_i) + (\nu_S(k_i))) + \alpha_2(\mu_T(k_i) + (\nu_T(k_i))))))\right) \quad (14)$$

Similarly

$$-\alpha_2(-\sum_{i=1}^{t}(\mu_S(k_i)\log(\mu_S(k_i)) + (\nu_S(k_i))\log(\nu_S(k_i)))) - \alpha_1(-\sum_{i=1}^{n}(\mu_T(k_i)\log(\mu_T(k_i)) + (\nu_T(k_i))\log(\nu_T(k_i)))) \geq -\left(-\sum_{i=1}^{n}((\alpha_2(\mu_S(k_i) + (\nu_S(k_i))) + \alpha_1\left(\mu_T(k_i) + (\nu_T(k_i))\right)(\log(\alpha_2(\mu_S(k_i) + (\nu_S(k_i))) + \alpha_1(\mu_T(k_i) + (\nu_T(k_i))))))\right) \quad (15)$$

Adding (14) and (15), we get $SHN(S,T) \geq o$

In coming segment we study various main properties of $SHN(S,T)$

IV. Properties of proposed dissimilarity measure

Theorem 4.1 For $S,T,R \in IFS(X)$,

(1) $SHN(S,T) = SHN(T,S)$
(2) $SHN(S,T) = 0\ ifandonlyif S = T$
(3) $SHN(S,S^c) = 0\ ifandonlyif \mu_S(k_i) = \nu_S(k_i) forallk_i \in X$
(4)$SHN(S,S\cup T) = SHN(S\cap T,T)$= SHN(S, T)
(5) $SHN(S\cup T,S\cap T) = SHN(S\cap T,S\cup T) = SHN(S,T)\backslash$
(6) $SHN(S,SUT)\ +\ SHN(S,S\cap T) = 2SHN(S,T)$
(7) $SHN(T,S\cup T) + SHN(T,S\cap T) = 2SHN(T,S)$
(8) $SHN(S\cup T,R) \leq SHN(S,R) + SHN(T,R)$
(9) $SHN(S\cap T,R) \leq SHN(S,R) + SHN(T,R)$
(10)
$SHN(S\cup T,R) + SHN(S\cap T,R) = SHN(S,R) + SHN(T,R)$
(11) $SHN(S,T) = SHN(S^C,T^C)$;
(12) $SHN(S,T^C) = SHN(S^C,T)$
(13) $SHN(S,T) + SHN(S^c,T) = SHN(S^c,T^c) + SHN(S,T^c)$
$whereS^c$is complement of S and T^c is complement of T.

Proof: Above properties proved in supplementary section

In upcoming part, proposed divergence measure applications applied for pattern recognition with some illustrations.

V. Pattern Recognition Application

Here, our main purpose is that from a set of patterns we haceto select a particular pattern by comparing their features. Procedure is explained as follows:
Consider a set of n patterns i.e. $\psi_1,\psi_2,\dots,\psi_n$ with classifications$\phi_1,\phi_2,\dots,\phi_n$. Suppose the patterns are defined by fuzzy sets
$\psi_i = \{s_i,\mu_{\psi_i}(s_i),\nu_{\psi_i}(s_i): s_i \in \psi\}, i = 1,2,\dots,n$ (16)

where $\psi = (s_1, s_2, \ldots, s_n)$ is universe of discourse. Let there is a known pattern defined by fuzzy set as

$P = \{\{s_j, \mu_F(s_j), \nu_F(s_j): s_j \in \psi\}$ (17)

For matching P to one of $\phi_1, \phi_2, \ldots, \phi_n$, use principle of minimum divergence explained by Shore and Grey [32] given as

$z^* = arg \min_z \{\psi_z, P\}$ (18)

Using the proposed measure, identify the one pattern which is resembles with the given pattern. Consider an example

Example: Let ψ_1, ψ_2, ψ_3 be three patterns with classifications ϕ_1, ϕ_2, ϕ_3 represented by fuzzy sets as below

$$\phi_1 = \{(s_1, 0.2,0.5), (s_2, 0.5,0.4), (s_3, 0.2,0.4)\}$$
$$\phi_2 = \{(s_1, 0.4,0.3), (s_2, 0.6,0.1), (s_3, 0.5,0.2)\} \quad (19)$$
$$\phi_3 = \{(s_1, 0.1,0.4), (s_2, 0.3,0.5), (s_3, 0.7,0.1)\}$$

Suppose P is a particular pattern defined by

$P = \{(s_1, 0.4,0.5), (s_2, 0.3,0.2), (s_3, 0.6,0.3)$ (20)

Our motive is to classify P to one of the class ϕ_1, ϕ_2, ϕ_3. The proposed measure values calculated are showed in Table *I*

Table *I*: valued calculated from proposed measure

$SHN(\psi_i, P)$	ϕ_1	ϕ_2	ϕ_3
$\alpha_1 = 0.1, \alpha_2 = 0.9$	0.0327	0.0170	0.0363
$\alpha_1 = 0.2, \alpha_2 = 0.8$	0.0576	0.0301	0.0635
$\alpha_1 = 0.3, \alpha_2 = 0.7$	0.0752	0.0394	0.0826
$\alpha_1 = 0.4, \alpha_2 = 0.6$	0.0856	0.0458	0.0938
$\alpha_1 = 0.5, \alpha_2 = 0.5$	0.0890	0.0468	0.0975

In this table it is clearly visible that pattern P is classified with ϕ_2 for different - different values of α_1 and α_2. And clearly we can say that ψ_2 is the perfect option.

VI. Comparitive Study

Now compare results of proposed measure (13) with measures that are existing. Here we are considered Li(2004), Zhang and Jiang (2008), Rajesh Joshi (2018).

The IF fizzy divergence measure between any two Intuitionstic fuzzy sets proposed by Li(2004) is given by

$d_L(S,T) = \sum_{i=1}^{n} \left(\frac{|\mu_S(r_i) - \mu_T(r_i)| + |\nu_S(r_i) - \nu_T(r_i)|}{2} \right)$ (21)

By Zhang and Jiang (2008), intuitionist fuzzy measure is given by

$$D(S,T) = \sum_{i=1}^{n} \left[\frac{\mu_S(r_i) + 1 - \nu_S(r_i)}{2} \log_2 \left(\frac{[\mu_S(r_i) + 1 - \nu_S(r_i)]/2}{[(\mu_S(r_i) + 1 - \nu_S(r_i)) + (1 + \mu_T(r_i) - \nu_T(r_i))]/4} \right) + \frac{-\mu_S(r_i) + 1 + \nu_S(r_i)}{2} \log_2 \left(\frac{[-\mu_S(r_i) + 1 + \nu_S(r_i)]/2}{[(-\mu_S(r_i) + 1 + \nu_S(r_i)) + (1 - \mu_T(r_i) + \nu_T(r_i))]/4} \right) \right]$$

(22)

The intuitionistic fuzzy measure by Rajesh Joshi (2018) is given by

$$d(S,T) = -\sum_{i=1}^{n}\begin{pmatrix} (\eta_1\mu_S(k_i)+\eta_2\mu_T(k_i)(log(\eta_1\mu_S(k_i)+\eta_2\mu_T(k_i)) \\ +(\eta_1\mu_T(k_i)+\eta_2\mu_S(k_i)(log(\eta_1\mu_T(k_i)+\eta_2\mu_S(k_i)) \\ +(\eta_1\nu_S(k_i)+\eta_2\nu_T(k_i)(log(\eta_1\nu_S(k_i)+\eta_2\nu_T(k_i)) \\ +(\eta_1\nu_T(k_i)+\eta_2\nu_S(k_i)(log(\eta_1\nu_T(k_i)+\eta_2\nu_S(k_i)) \\ +(\eta_1\pi_S(k_i)+\eta_2\pi_T(k_i)(log(\eta_1\pi_S(k_i)+\eta_2\pi_T(k_i)) \\ +(\eta_1\pi_T(k_i)+\eta_2\pi_S(k_i)(log(\eta_1\pi_T(k_i)+\eta_2\pi_S(k_i)) \\ -\left(\mu_S(k_i)\log\left(\mu_S(k_i)\right)+\nu_S(k_i)\log\left(\nu_S(k_i)\right)+\pi_S(k_i)\log\left(\pi_S(k_i)\right)\right) \\ -\left(\mu_T(k_i)\log\left(\mu_T(k_i)\right)+\nu_T(k_i)\log\left(\nu_T(k_i)\right)+\pi_T(k_i)\log\left(\pi_T(k_i)\right)\right) \end{pmatrix}$$

(23)

Table *II* shows the calculated values of divergence measure (13), (21),(22) and (23) as follows:

Divergence measure	ϕ_1	ϕ_2	ϕ_3
$d_L(\phi_i,P)$	0.5500	0.4000	0.5000
$d_{HY}(\phi_i,P)$	0.2667	0.2000	0.2667
$d_{RJ}(\phi_i,P)$	0.4933	0.2378	0.4535
$SHN(\phi_i,P)$	0.0327	0.0170	0.0363

From above table, it could be noticed from all divergence measure the pattern Psupported withϕ_2. Hence the performance of our proposed measure could be considered good.

VII. Conclusion

Here a new divergence measure developed based on Jensen Shannon information measure and some properties are discussed and showed the effectiveness of proposed measure with the help of example and with comparing the results with previous measures on Pattern Recognition. We will do work on theirmore generalizations further.

Supplement Section: Proofs of Properties

Proof: For verifyingthe properties, we fractionate χ into χ_1 and χ_2 such that

$\chi_1 = \{e_i \in \chi: S(e_i) \subseteq T(e_i)\}$and $\chi_2 \in \{\, b_i \in \chi: S(e_i) \supseteq T(e_i)\}$ (APP-1)

From (APP-1), for all $e_i \in \chi_1$
$\mu_S(e_i) \le \mu_T(e_i)$ and $\nu_S(e_i) \ge \nu_T(e_i)$ (APP-2)
for all $e_i \in \chi_2$
$\mu_S(e_i) \ge \mu_T(e_i)$ and $\nu_S(e_i) \le \nu_T(e_i)$ (APP-3)

1. Statement (13) directly followed the proof of 1,2 and 3
2. Proof (Property 4): From (13), we get

$$\begin{aligned} SHN(S,S\cup T) = -\sum_{i=1}^{t}\Big((\alpha_1\mu_S(k_i)+\alpha_2\mu_{S\cup T}(k_i))(log(\alpha_1\mu_S(k_i)+\alpha_2\mu_{SUT}(k_i)))+(\alpha_1\mu_{S\cup T}(k_i) \\ +\alpha_2\mu_S(k_i))(log(\alpha_1\mu_{S\cup T}(k_i)+\alpha_2\mu_S(k_i)))+(\alpha_1\nu_S(k_i)) \\ +\alpha_2(\nu_{S\cup T}(k_i)))(log(\alpha_1\mu_S(k_i))+\alpha_2(\mu_{S\cup T}(k_i))))+(\alpha_1\nu_{S\cup T}(k_i)) \\ +\alpha_2\nu_S(k_i)))(log(\alpha_1\nu_{S\cup T}(k_i))+\alpha_2\nu_S(k_i))))-(\mu_S(\mathrm{k_i})\log(\mu_S(k_i)) \\ +(\mu_{S\cup T}(\mathrm{k_i})\log(\mu_{S\cup T}(k_i)))-((\nu_S(\mathrm{k_i}))\log(\nu_S(k_i)))+((\nu_{S\cup T}(\mathrm{k_i}))\log(\nu_{S\cup T}(k_i))))\Big) \end{aligned}$$

(APP-4)

$$= \quad -\sum_{\chi_1}\Big((\alpha_1\mu_S(k_i)+\alpha_2\mu_T(k_i))(log(\alpha_1\mu_S(k_i)+\alpha_2\mu_T(k_i)))+(\alpha_1\mu_T(k_i)$$
$$+\alpha_2\mu_S(k_i))(log(\alpha_1\mu_T(k_i)+\alpha_2\mu_S(k_i)))+(\alpha_1\nu_S(k_i))$$
$$+\alpha_2(\nu_T(k_i)))(log(\alpha_1\nu_S(k_i))+\alpha_2(\nu_T(k_i))))+(\alpha_1\nu_T(k_i))$$
$$+\alpha_2\nu_S(k_i)))(log(\alpha_1\nu_T(k_i))+\alpha_2\nu_S(k_i))))-(\mu_S(\mathrm{k_i})\log(\mu_S(k_i))+(\mu_T(\mathrm{k_i})\log(\mu_T(k_i)))$$
$$-\Big((\nu_S(\mathrm{k_i}))\log(\nu_S(k_i))\Big)+\Big((\nu_T(\mathrm{k_i}))\log(\nu_T(k_i))\Big))\Big)$$

$$-\sum_{\chi_2}\big((\alpha_1\mu_S(k_i)+\alpha_2\mu_S(k_i))(log(\alpha_1\mu_S(k_i)+\alpha_2\mu_S(k_i)))+(\alpha_1\mu_S(k_i)$$
$$+\alpha_2\mu_S(k_i))(log(\alpha_1\mu_S(k_i)+\alpha_2\mu_S(k_i)))+(\alpha_1\nu_S(k_i))$$
$$+\alpha_2(\nu_S(k_i)))(log(\alpha_1\nu_S(k_i))+\alpha_2(\nu_S(k_i))))+(\alpha_1\nu_S(k_i))$$
$$+\alpha_2\nu_S(k_i)))(log(\alpha_1\nu_S(k_i))+\alpha_2\nu_S(k_i))))-(\mu_S(\mathrm{k_i})\log(\mu_S(k_i))+(\mu_S(\mathrm{k_i})\log(\mu_S(k_i)))$$
$$-((\nu_S(\mathrm{k_i}))\log(\nu_S(k_i)))+((\nu_S(\mathrm{k_i}))\log(\nu_S(k_i)))))\big)$$

(APP-5)

$$= -\sum_{i=1}^{t}\big((\alpha_1\mu_S(k_i)+\alpha_2\mu_T(k_i))(log(\alpha_1\mu_S(k_i)+\alpha_2\mu_T(k_i)))+(\alpha_1\mu_T(k_i)$$
$$+\alpha_2\mu_S(k_i))(log(\alpha_1\mu_T(k_i)+\alpha_2\mu_S(k_i)))+(\alpha_1\nu_S(k_i))$$
$$+\alpha_2(\nu_T(k_i)))(log(\alpha_1\nu_S(k_i))+\alpha_2(\nu_T(k_i))))+(\alpha_1\nu_T(k_i))$$
$$+\alpha_2\nu_S(k_i)))(log(\alpha_1\nu_T(k_i))+\alpha_2\nu_S(k_i))))-(\mu_S(\mathrm{k_i})\log(\mu_S(k_i))+(\mu_T(\mathrm{k_i})\log(\mu_T(k_i)))$$
$$-((\nu_S(\mathrm{k_i}))\log(\nu_S(k_i)))+((\nu_T(\mathrm{k_i}))\log(\nu_T(k_i)))))\big)$$

$= SHN(S,T)$ (APP-6)

Similarly we can prove $SHN(S\cap T,T)=SHN(S,T)$. Property 4 proved.

3. Proof (Property 5): In the same manner of Property 4 we can prove the Property 5.
4. Proof (Property 6): In Property 4, $SHN(S,S\cup T)=SHN(S,T)$
   In Property 5, $SHN(S,S\cap T)= SHN(S,T)$
   So by adding these $SHN(S,S\cup T)+SHN(S,S\cap T)= 2SHN(S,T)$
5. Proof (Property 7): It can be prove by same procedure as property 6.
6. Proof (Property 8): Let us take

$$SHN(S,R)+SHN(T,R)-SHN(S\cup T,R)=$$

$$-\sum_{i=1}^{t}\big((\alpha_1\mu_S(k_i)+\alpha_2\mu_R(k_i))(log(\alpha_1\mu_S(k_i)+\alpha_2\mu_R(k_i)))+(\alpha_1\mu_R(k_i)$$
$$+\alpha_2\mu_S(k_i))(log(\alpha_1\mu_R(k_i)+\alpha_2\mu_S(k_i)))+(\alpha_1\nu_S(k_i))$$
$$+\alpha_2(\nu_R(k_i)))(log(\alpha_1\nu_S(k_i))+\alpha_2(\nu_R(k_i))))+(\alpha_1\nu_R(k_i))$$
$$+\alpha_2\nu_S(k_i)))(log(\alpha_1\nu_R(k_i))+\alpha_2\nu_S(k_i))))-(\mu_S(\mathrm{k_i})\log(\mu_S(k_i))+(\mu_R(\mathrm{k_i})\log(\mu_R(k_i)))$$
$$-((\nu_S(\mathrm{k_i}))\log(\nu_S(k_i)))+((\nu_R(\mathrm{k_i}))\log(\nu_R(k_i)))))\big)$$

$$-\sum_{i=1}^{t}\big((\alpha_1\mu_T(k_i)+\alpha_2\mu_R(k_i))(log(\alpha_1\mu_T(k_i)+\alpha_2\mu_R(k_i)))+(\alpha_1\mu_R(k_i)$$
$$+\alpha_2\mu_T(k_i))(log(\alpha_1\mu_R(k_i)+\alpha_2\mu_T(k_i)))+(\alpha_1\nu_T(k_i))$$
$$+\alpha_2(\nu_R(k_i)))(log(\alpha_1\nu_T(k_i))+\alpha_2(\nu_R(k_i))))+(\alpha_1\nu_R(k_i))$$
$$+\alpha_2\nu_T(k_i)))(log(\alpha_1\nu_R(k_i))+\alpha_2\nu_T(k_i))))-(\mu_T(\mathrm{k_i})\log(\mu_T(k_i))+(\mu_R(\mathrm{k_i})\log(\mu_R(k_i)))$$
$$-((\nu_T(\mathrm{k_i}))\log(\nu_T(k_i)))+((\nu_R(\mathrm{k_i}))\log(\nu_R(k_i)))))\big)$$

$$-\sum_{i=1}^{t}\Big((\alpha_1\mu_T(k_i)+\alpha_2\mu_R(k_i))(log(\alpha_1\mu_T(k_i)+\alpha_2\mu_R(k_i)))+(\alpha_1\mu_R(k_i)$$
$$+\alpha_2\mu_T(k_i))(log(\alpha_1\mu_R(k_i)+\alpha_2\mu_T(k_i)))+(\alpha_1\nu_T(k_i))$$
$$+\alpha_2(\nu_R(k_i)))(log(\alpha_1\nu_T(k_i))+\alpha_2(\nu_R(k_i))))+(\alpha_1\nu_R(k_i))$$
$$+\alpha_2\nu_T(k_i)))(log(\alpha_1\nu_R(k_i))+\alpha_2\nu_T(k_i))))-(\mu_T(\mathrm{k_i})\log(\mu_T(k_i))+(\mu_R(\mathrm{k_i})\log(\mu_R(k_i)))$$
$$-\Big((\nu_T(\mathrm{k_i}))\log(\nu_T(k_i))\Big)+\Big((\nu_R(\mathrm{k_i}))\log(\nu_R(k_i))\Big))\Big)$$

$$+\sum_{i=1}^{t}\big((\alpha_1\mu_S(k_i)+\alpha_2\mu_{S\cup T}(k_i))(log(\alpha_1\mu_S(k_i)+\alpha_2\mu_{SUT}(k_i)))+(\alpha_1\mu_{S\cup T}(k_i) + \alpha_2\mu_S(k_i))(log(\alpha_1\mu_{S\cup T}(k_i)+\alpha_2\mu_S(k_i)))+(\alpha_1\nu_S(k_i)) + \alpha_2(\nu_{S\cup T}(k_i)))(log(\alpha_1\nu_S(k_i))+\alpha_2(\nu_{S\cup T}(k_i))))+(\alpha_1\nu_{S\cup T}(k_i)) + \alpha_2\nu_S(k_i)))(log(\alpha_1\nu_{S\cup T}(k_i))+\alpha_2\nu_S(k_i))))-(\mu_S(\mathrm{k_i})\log(\mu_S(k_i)) + (\mu_{S\cup T}(\mathrm{k_i})\log(\mu_{S\cup T}(k_i)))-((\nu_S(\mathrm{k_i}))\log(\nu_S(k_i)))+((\nu_{S\cup T}(\mathrm{k_i}))\log(\nu_{S\cup T}(k_i))))\big)$$

(APP-7)

And then

=

$$-\sum_{i=1}^{t}\big((\alpha_1\mu_S(k_i)+\alpha_2\mu_R(k_i))(log(\alpha_1\mu_S(k_i)+\alpha_2\mu_R(k_i)))+(\alpha_1\mu_R(k_i) + \alpha_2\mu_S(k_i))(log(\alpha_1\mu_R(k_i)+\alpha_2\mu_S(k_i)))+(\alpha_1\nu_S(k_i)) + \alpha_2(\nu_R(k_i)))(log(\alpha_1\nu_S(k_i))+\alpha_2(\nu_R(k_i))))+(\alpha_1\nu_R(k_i)) + \alpha_2\nu_S(k_i)))(log(\alpha_1\nu_R(k_i))+\alpha_2\nu_S(k_i))))-(\mu_S(\mathrm{k_i})\log(\mu_S(k_i))+(\mu_R(\mathrm{k_i})\log(\mu_R(k_i))) - ((\nu_S(\mathrm{k_i}))\log(\nu_S(k_i)))+((\nu_R(\mathrm{k_i}))\log(\nu_R(k_i))))\big)$$

$$-\sum_{I=1}^{n}\Big((\alpha_1\mu_T(k_i)+\alpha_2\mu_R(k_i))(log(\alpha_1\mu_T(k_i)+\alpha_2\mu_R(k_i)))+(\alpha_1\mu_R(k_i) + \alpha_2\mu_T(k_i))(log(\alpha_1\mu_R(k_i)+\alpha_2\mu_T(k_i)))+(\alpha_1\nu_T(k_i)) + \alpha_2(\nu_R(k_i)))(log(\alpha_1\nu_T(k_i))+\alpha_2(\nu_R(k_i))))+(\alpha_1\nu_R(k_i)) + \alpha_2\nu_T(k_i)))(log(\alpha_1\nu_R(k_i))+\alpha_2\nu_T(k_i))))-(\mu_T(\mathrm{k_i})\log(\mu_T(k_i))+(\mu_R(\mathrm{k_i})\log(\mu_R(k_i))) - \Big(\big(\nu_T(\mathrm{k_i})\big)\log\big(\nu_T(k_i)\big)\Big)+\Big(\big(\nu_R(\mathrm{k_i})\big)\log\big(\nu_R(k_i)\big)\Big))\Big)$$

$$+\sum_{X_1}\Big((\alpha_1\mu_S(k_i)+\alpha_2\mu_T(k_i))(log(\alpha_1\mu_S(k_i)+\alpha_2\mu_T(k_i)))+(\alpha_1\mu_T(k_i) + \alpha_2\mu_S(k_i))(log(\alpha_1\mu_T(k_i)+\alpha_2\mu_S(k_i)))+(\alpha_1\nu_S(k_i)) + \alpha_2(\nu_T(k_i)))(log(\alpha_1\nu_S(k_i))+\alpha_2(\nu_T(k_i))))+(\alpha_1\nu_T(k_i)) + \alpha_2\nu_S(k_i)))(log(\alpha_1\nu_T(k_i))+\alpha_2\nu_S(k_i))))-(\mu_S(\mathrm{k_i})\log(\mu_S(k_i))+(\mu_T(\mathrm{k_i})\log(\mu_T(k_i))) - \Big(\big(\nu_S(\mathrm{k_i})\big)\log\big(\nu_S(k_i)\big)\Big)+\Big(\big(\nu_T(\mathrm{k_i})\big)\log\big(\nu_T(k_i)\big)\Big))\Big)$$

$$+\sum_{X_2}\big((\alpha_1\mu_S(k_i)+\alpha_2\mu_S(k_i))(log(\alpha_1\mu_S(k_i)+\alpha_2\mu_S(k_i)))+(\alpha_1\mu_S(k_i) + \alpha_2\mu_S(k_i))(log(\alpha_1\mu_S(k_i)+\alpha_2\mu_S(k_i)))+(\alpha_1\nu_S(k_i)) + \alpha_2(\nu_S(k_i)))(log(\alpha_1\nu_S(k_i))+\alpha_2(\nu_S(k_i))))+(\alpha_1\nu_S(k_i)) + \alpha_2\nu_S(k_i)))(log(\alpha_1\nu_S(k_i))+\alpha_2\nu_S(k_i))))-(\mu_S(\mathrm{k_i})\log(\mu_S(k_i))+(\mu_S(\mathrm{k_i})\log(\mu_S(k_i))) - ((\nu_S(\mathrm{k_i}))\log(\nu_S(k_i)))+((\nu_S(\mathrm{k_i}))\log(\nu_S(k_i))))\big)$$

≥ 0 (APP-8)

Hence $SHN(S, S\cup T) \leq SHN(S,R) + SHN(T,R)$

7. Proof (Property 9): Property 9 can prove same as above.
8. Proof (Property 10): with the help of property 4 we can easily prove property 10
9. Properties 11 and 12 are easily proved by definition and Property 13 we can easily fine by adding property 11 and property 12.

Conflict of Interest: There is no conflict of interest as per the declaration of author.

Data Availability

The data that support the findings of this study are already included in the manuscript. There is no additional data to share.

References

[1]. Atanassov KT. Intuitionistic fuzzy sets. *Fuzzy Sets Syst.* 1986;20:87–96.
[2]. Atanassov KT. Intuitionistic Fuzzy Sets. Heidelberg, Germany: Physica; 1999.
[3]. Burillo P, Bustince H. Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets. *Fuzzy Sets Syst.*2001;118:305–316.
[4]. Cover, T.M., Thomas,J.A.: Elements of Information Theory. Wiley-Interscience, New York(1991)

[5]. A. De Luca and S. Termini(1972), A definition of a non- probabilistic entropy in thesetting of fuzzy sets theory, Information and Control 20 , 301-312.
[6]. Hung WL, Yang MS. On the *J*-divergence of intuitionistic fuzzy sets and its application to pattern recognition. *Inform Sci.* 2008; 178(6):1641–1650.
[7]. Joshi R, Kumar S. A new approach in multiple attribute decision making using *R*-norm entropy andHamming distance measure. *Int*
[8]. *J Inform Manag Sci.* 2016;27(3):253–268.
[9]. Joshi R, Kumar S. An (*R, S*)-norm fuzzy information measure with its applications in multiple-attribute Decision making,*Comput*
[10]. *Appl Math.* 2018; 37(3):2943–2964.
[11]. Joshi R, Kumar S. A new exponential fuzzy entropy of order-(α, β) and its application in multiple attribute decision making.
[12]. *Commun Math Stat.* 2017;5(2):213–229.
[13]. Joshi R, Kumar S. Application of interval-valued intuitionistic fuzzy *R*-norm entropy in multiple attribute Decision making. *Int J Inform Manag Sci.* 2017;28(3):233–251.
[14]. Li DF. Some measures of dissimilarity in intuitionistic fuzzy structures. *J Comput Syst Sci.* 2004;68(1):115– 122.
[15]. Mahalanobis (1936). "On the generalised distance in statistics" (PDF). *Proceedings of the National Institute of Sciences of India*. **2** (1): 49–55. Retrieved 2016-09-27.
[16]. Rao, C.R.: Diversity and dissimilarity coefficients: a unified approach. Theor. Popul. Biol.21,24-43(1982)
[17]. S. Kullback and R.A. Leibler(1951), On information and su_ciency, Annals of MathematicalStatistics22 , 79-86.
[18]. S. Kullback(1959), Information theory and statistics, Wiley, New York
[19]. S. Kullback(1968), Information Theory and Statistics, Courier Corporation
[20]. Shang X-G, Jiang W-S. A note on fuzzy information measures. *Pattern Recognit Lett.* 1997;18:425–432.
[21]. T.M. Cover and J.A. Thomas(2012), Elements of Information Theory, John Wiley andSons.
[22]. Vlachos IK, Sergiadis GD. Intuitionistic fuzzy information—applications to pattern recognition. *Pattern RecognitLett.*2007;28(2):197–206.
[23]. Wang WQ,XinXL.Distance measures between intuitionistic fuzzy sets. *Pattern Recognit Lett.* 2005;26:2063–2069.
[24]. Zhang Q, Jiang S. A note on information entropy measure for vague sets. *Inform Sci.* 2008; 178:4184–4191.
[25]. Joshi R, Kumar S. Parametric (*R, S*)-norm entropy on intuitionistic fuzzy sets with a new approach in Multiple attribute decision making. *Fuzzy Inform Eng.* 2017; 9:181–203.
[26]. Joshi R, Kumar S. An intuitionistic fuzzy (δ, γ)-norm entropy with its application in supplier selection problem, *Comput Appl Math.*, 2018; https://doi.org/10.1007/s40314-018-0656-9.
[27]. Joshi R, Kumar S. An intuitionistic fuzzy information measure of order (α, β) with a new approach in Supplier selection problems using an extended VIKOR method, *Journal of Applied Mathematics and Computing*, 2018; https://doi.org/10.1007/s12190-018-
[28]. 1202-z.
[29]. Joshi R, Kumar S. A new parametric intuitionistic fuzzy entropy and its applications in multiple attribute Decisionmaking, *Int. J.*
[30]. *Appl. Comput. Math* (2018) 4:52–74.
[31]. Verma R, Sharma BD. Intuitionistic fuzzy Jensen Renyi divergence: application to multiple attribute decision making. *Informatica.* 2013;37(4):399–409.
[32]. Verma R, Sharma BD. On generalized intuitionistic fuzzy divergence (relative information) and theirproperties. *JUncertain Syst.* 2012;6(4):308–320.
[33]. Zhang SF, Liu SY. GRA-based intuitionistic multi criteria decision making method for personal selection. *Experts Syst Appl.* 2011;38(9):11401–11405.
[34]. Joshi R, Kumar S., Gupta D., Kaur H. A Jensen-α-norm dissimilarity measure for intuitionistic fuzzy sets and its applications in
[35]. multiple attribute decision making, *Int. J. Fuzzy Syst.* (2018) 20(4):1188–1202.
[36]. Joshi R, Kumar S. An exponential Jensen fuzzy divergence measure with applications in multiple attribute decision making, *Mathematical problems in engineering*, 2018; Volume 2018, Article ID 4342098, 9 pages.(https://doi.org/10.1155/2018/4342098)
[37]. Shannon CE. The mathematical theory of communication. *Bell Syst Tech J.* 1948;27:379–423, 623–656.
[38]. Shore JE, Gray RM.Minimization cross-entropy pattern classification and cluster analysis. *IEEE Trans Pattern Anal Mach Intell.* 1982;4(1):11–17.
[39]. Lin J. Divergence measure based on Shannon entropy. *IEEE Trans Inform Theory.* 1991;37(1):145–151.
[40]. Joshi R, Kumar S. An (*R', S'*)-norm fuzzy relative information measure and its applications in strategic Decision making. *Comput*
[41]. *Appl Math.* 2018; https://doi.org/10.1007/s40314-018-0582-x.
[42]. Zadeh LA. Fuzzy sets. *Inform Comput.* 1965;8:338–353.