Ecological Informatics 77 (2023) 102206

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

CO_2 emission based GDP prediction using intuitionistic fuzzy transfer learning

Sandeep Kumar [a], Amit K. Shukla [b,c,1,*], Pranab K. Muhuri [a], Q.M. Danish Lohani [a]

[a] Department of Computer Science, South Asian University, Maidan Garhi, New Delhi 110068, India
[b] School of Technology and Innovations, University of Vaasa, Wolffintie 34, FI-65200 Vaasa, Finland
[c] Faculty of Information Technology, University of Jyväskylä, Box 35 (Agora), Jyväskylä 40014, Finland

ARTICLE INFO

Keywords:
Atanassov intuitionistic fuzzy sets
Hausdorff distance
Yager's generating function
GDP prediction
Fuzzy sets

ABSTRACT

The industrialization has been the primary cause of the economic boom in almost all countries. However, this happened at the cost of the environment, as industrialization also caused carbon emissions to increase exponentially. According to the established literature, Gross Domestic Product (GDP) is related to carbon emissions (CO_2) which could be optimally employed to precisely estimate a country's GDP. However, the scarcity of data is a significant bottleneck that could be handled using transfer learning (TL) which uses previously learned information to resolve new tasks, more specifically, related tasks. Notably, TL is highly vulnerable to performance degradation due to the deficiency of suitable information and hesitancy in decision-making. Therefore, this paper proposes 'Intuitionistic Fuzzy Transfer Learning (IFTL)', which is trained to use CO_2 emission data of developed nations and is tested for its prediction of GDP in a developing nation. IFTL exploits the concepts of intuitionistic fuzzy sets (IFSs) and a newly introduced function called the modified Hausdorff distance function. The proposed IFTL is investigated to demonstrate its actual capabilities for TL in modeling hesitancy. To further emphasize the role of hesitancy modelled with IFSs, we propose an ordinary fuzzy set (FS) based transfer learning. The prediction accuracy of the IFTL is further compared with widely used machine learning approaches, extreme learning machines, support vector regression, and generalized regression neural networks. It is observed that IFTL capably ensured significant improvements in the prediction accuracy over other existing approaches whenever training and testing data have huge data distribution differences. Moreover, the proposed IFTL is deterministic in nature and presents a novel way for mathematically computing the intuitionistic hesitation degree.

1. Introduction

Rich and developed countries possess higher per capita Gross Domestic Product (GDP) but at the cost of enormous energy consumption. Globally 46% of carbon emission is produced by energy-based industries (Bokde et al., 2021). Historically, industrialization and urbanization propelled the per capita GDP of a nation. Energy was produced by burning fossil fuels which ultimately caused an exponential increase in carbon emissions and degradation of the environment (Depren et al., 2022). The construction industry played a crucial role in urbanization and is also one of the biggest emitters of CO_2 (carbon dioxide) (Guo et al., 2012). Yao et al. (2015) have proved that economic growth caused an increase in CO_2 emissions for G20 countries. Similarly, Govindaraju and Tang, (2013b) demonstrated that industrialization represents the economic growth of a nation which is proportional to its CO_2 emission. Adedoyin et al., (2021) showed that both real GDP and use of non-renewable energy sources increased CO_2 emissions in 32 countries in Sub-Saharan Africa. Further, Gyamfi et al., (2021) have proved that a 1% increase in GDP caused a 0.400% increase in pollution for emerging industrialized seven (E7) economies. Chaabouni and Saidi, (2017) studied 51 countries and found that a 1% increase in CO_2 increased the economy by 0.011%, while 1% economic growth increased CO_2 emissions by 0.263%. Acheampong, (2018b) further proved that CO_2 emission positively impacts economic development by analyzing historical

* Corresponding author.
E-mail addresses: 2431sandeep@gmail.com (S. Kumar), amit.shukla@uwasa.fi (A.K. Shukla), pranabmuhuri@cs.sau.ac.in (P.K. Muhuri), danishlohani@cs.sau.ac.in (Q.M.D. Lohani).
[1] Corresponding author is currently at the School of Technology and Innovations, University of Vaasa, Finland.

https://doi.org/10.1016/j.ecoinf.2023.102206
Received 1 July 2022; Received in revised form 5 July 2023; Accepted 6 July 2023
Available online 11 July 2023

data from 116 countries. Globally, Stern, (2010b) showed that carbon-income elasticity is 1.509, indicating their proportional relationship.

This tightly coupled relationship between CO_2 emission and GDP motivated us to predict the GDP of a nation using its carbon emission. However, the available dataset of some countries is not sufficient enough to train a machine learning (ML) model. To overcome this limitation, this paper proposes a novel transfer learning methodology. Transfer learning (TL), a machine learning (ML) technique, utilizes previously learned information to resolve new tasks, more specifically, related tasks. In conventional ML techniques, every new task is learned from the scratch (Che et al., 2021), and the training and test data are drawn from the same distribution. However, if the data distribution is changed between the training and test data, the performance and reliability of the predicted output may degrade significantly (Lu et al., 2019). This may also lead to high computational costs and inappropriate outcomes (Pan and Yang, 2010). In such a scenario, TL is helpful as it transfers the learned knowledge from a related task to solve a new task in a related domain. TL contradicts semi-supervised ML algorithms (Van et al., 2020), as it can handle all those cases where both domains of training and testing data are of different distribution (Pedrycz et al., 1997). TL has been associated with many other terminologies, such as transfer of learning, domain adaptation (Pan and Yang, 2010), multi-task learning (MTL), meta-learning, etc. It has found its application in several domains, including intelligent environments (Shell and Coupland, 2015), failure prediction (Behbood et al., 2015), acoustic monitoring (Dufourq et al., 2022), detection and classification (Lu et al., 2021; Lumini and Nanni, 2019), etc.

A discrete numeric set of values often represents feature space matrices in TL methods. This is an essential factor to be considered in the cases of noisy data sets. Such noises, which are associated with information and decisions, can be modelled using fuzzy sets (FSs). The parameters in human cognition to handle uncertainty and imprecision during knowledge transfer are generally associated with epistemic uncertainties (Shell and Coupland, 2015). Various factors could influence these parameters, including uncertainties due to external noises, measurement uncertainty, non-agreement of the decision-makers, etc. Almost all the fuzzy TL (FTL) techniques are developed using concepts of ordinary FSs, each element of which is associated with a membership value. There are higher-order fuzzy sets such as type-2 fuzzy sets which considers membership function uncertainty for every element of traditional FSs (Debnath et al., 2018; Shukla and Muhuri, 2019). However, in a real-world applicative context, there are cases when besides membership value, there exists a need for non-membership value to accurately define its elements due to hesitancy. This hesitancy can be effectively handled by the generalization of FSs, called intuitionistic fuzzy sets (IFSs). It was proposed by Atanassov, (1986), where the hesitation margin was considered as hesitancy (Fan and Xiao, 2020; Ohta et al., 2020) that can enable better decision making. The sum of membership and non-membership values is unity in an FS; however, for an IFS, the hesitation margin is calculated by subtracting the sum of membership and non-membership values from unity. This property of IFSs have shown potential in wide range of applications such as adaptive sliding mode control method (Kutlu et al., 2020), time series analysis (Castillo et al., 2007), multi-criteria decision making (Shukla et al., 2022), hyperchaotic synchronization (Atan et al., 2020), etc.

To handle the hesitancy in a TL, this paper proposes a novel approach, which is termed 'Intuitionistic fuzzy transfer learning (IFTL)'. IFTL capably increases learning by optimally exploiting the hesitancy aspect of IFSs and a newly introduced similarity function called the modified Hausdorff distance function. IFTL models human behavior (of cautiously utilizing their previous experience) with a hesitancy degree using the Atanassov's IFSs. The novelty of the proposed IFTL is that the hesitancy degree is not randomly chosen or tuned, it is empirically computed using a novel mathematical formulation that captures the variance in the source and target domains. The input data is intuitionistically fuzzified using Yager's generating function (Ghosh et al., 2021; Kumar et al., 2016). The similarities between these intuitionistically fuzzified domains are optimally captured using modified Hausdorff distance, which adaptively refine the outcome of the learning model. This procedure effectively transfers knowledge across domains by using the differences in their data distribution. The proposed IFTL model is investigated to demonstrate its knowledge transfer abilities in a TL domain.

The performance of the IFTL is also compared with the FTL, where the ordinary FSs and Euclidean distance are utilized instead of IFSs and modified Hausdorff distance metric, respectively. Notably, there is no hesitancy degree in ordinary FSs; therefore, FTL (unlike IFTL) may not be able to model the human tendency to cautiously utilize their experience during the transferal of knowledge across hugely diverse domains. It is explicitly studied to signify the importance of hesitancy, without which FTL cannot enhance learning during the transferal of knowledge. Furthermore, the prediction accuracies of IFTL are compared with the widely used ML approaches such as extreme learning machine (ELM) (Specht, 1991), support vector regression (SVR) (Bi and Bennett, 2003; Vapnik, 2013), generalized regression neural network (GRNN) (http://www.ntu.edu.sg/home/egbhuang/elm_kernel.html, n.d.; Huang et al., 2006). The experimentations are performed where the model is trained for GDP prediction using only the CO_2 emission data of developed nations and is tested for its prediction for the developing nation. It is observed that IFTL capably ensured significant improvements in the prediction accuracy over other approaches.

The major contributions of the paper are mentioned below:

a) This paper proposes a novel domain adaptation approach that provides a mechanism to manage hesitancy in a TL problem. The approach is termed 'Intuitionistic fuzzy transfer learning (IFTL)', which can also enhance model learning during transfer learning.
b) A new distance function called 'the modified Hausdorff distance function' is also proposed in this paper, which is exploited by IFTL along with the hesitancy aspect of IFSs.
c) Further, to highlight the significance of hesitancy using IFTL, this paper also investigates and compares the problem with ordinary FS-based transfer learning (FTL).
d) The proposed IFTL approach is investigated to demonstrate its capability for TL, where it is trained for GDP prediction using CO_2 emission data of developed nations and then tested for its GDP prediction accuracies for a developing nation.
e) IFTL capably ensures significant improvements over ELM, SVR, GRNN, and FTL with better prediction accuracies across highly diverse domains.
f) IFTL is deterministic in nature and introduces an innovative way for calculating intuitionistic hesitation degree by utilizing the variance of the data differences among different domains.

The rest of the paper is organized as follows. Related research works in TL and domain adaptation are discussed in Section 2. In Section 3, essential mathematical preliminaries are explained. Section 4 explains the proposed IFTL approach and a novel procedure for selecting an intuitionistic hesitation degree. Section 5 discusses the experimental results of our approach. Finally, Section 6 concludes with a discussion of the work and highlights some future research scopes.

2. Related works

Rosenstein et al., (2005) empirically proved that, in case of extreme dissimilarity between two tasks, the technique's performance degrades even after knowledge transferal with the use of brute-force technique. Pedrycz et al. (1997)) proposed calibration of fuzzy sets via non-linear context adaptation. Pan et al., 2016 proposed Multi-Layer Transfer Learning (MLTL) which uses latent spaces to learn non-shared concepts. Firstly, specific latent feature spaces are produced by MLTL, which are then combined with common latent feature space to generate one latent feature space layer. Moreover, to learn the corresponding distributions on different layers, multiple layers are generated which learns with their pluralism, simultaneously. Salaken et al., (2019) employed deep learning in which a small sample of target domain is used as seed to transform source domain dataset. Sousa et al., (2020) proposed a transfer learning approach with data augmentation techniques for wildfire detection to overcome data limitations.

Xie et al., (2018) introduced the generalized hidden mapping transudative learning method to extend the utilization of the transfer learning method in various other classical as well as intelligent models. The authors demonstrated the applicability of the proposed approach in recognition of epileptic electroencephalogram (EEG) signals. A novel TL approach is introduced in Wang et al., (2018) to handle the problem of missing data. Here, TL may be utilized to enhance classification problems performance with incomplete training data by adopting the additive least squares support vector machine (LS-SVM). Chen et al., (2018b) proposed the domain space transfer Extreme Learning Machine (DST-ELM), which handles the problem of unsupervised domain adaptation, i.e., without the availability of target label data. Predictions are performed on real-world image and text datasets.

In several reported works, transfer learning approaches were utilized with several fuzzy techniques in various applications. Deng et al., (2014) proposed an inductive transfer learning approach based on Takagi-Sugeno-Kang Fuzzy logic System (TSK FLS) for email spam filtering with the help of neural networks (NNs). Transfer learning in Takagi-Sugeno fuzzy model was proposed by Zuo et al., (2016) for regression using particle swarm optimization (PSO) and differential evolution (DE). Domain adaptation was used in conjunction with FSs for various other applications, such as land use classification (Lu et al., 2021) and driver drowsiness estimation (Wu et al., 2016). FTL was utilized in intelligent environments and online fuzzy min max NNs (Seera and Lim, 2014). Also, unsupervised TL was used in texture image segmentation with fuzzy c-means (FCM) clustering by Qian et al., (2017).

The problem of domain adaptation has been discussed extensively in the research domain (Lu et al., 2015; Margolis, 2011; Pan and Yang, 2010) and can broadly be classified into five main classes: 1) All source domain samples are weighted to mimic the samples of the target domain using covariate shift methods with instance weighting, 2) Whenever confidence is reasonably high in the collected or balanced samples then knowledge transferral using instances are utilized (Saidi and Hammami, 2015), 3) Cluster-based learning methods, where samples in a single cluster will be having same labels, 4) Self-labeling methods, which integrates target domain's unlabeled samples during training, initializes their labels and then do label refinements iteratively, 5) Feature representation methods uses feature representation techniques which extract suitable features from data set.

Uncertainty handling using IFSs was demonstrated by Szmidt and Kacprzyk, (2001). In their research, it was shown that IFSs can suitably describe a problem by linguistic variables. Also, Szmidt and Kacprzyk, (2004) highlighted the requirement of IFSs in knowledge uncertainty, in the applications of human reasoning. Behbood et al., (2013a; 2015) introduced a novel domain adaptation technique for bank failure prediction called the fuzzy domain adaptation. Their technique has gone through a lot of refinement, as mentioned by Behbood and others in (Behbood et al., 2015), (Behbood et al., 2013a; Behbood et al., 2013b). The same algorithm was developed and thoroughly analyzed in Behbood et al., (2013a), which proved its superiority over its other evolutions. Table 1 depicts the comparative analyses of existing fuzzy transfer learning approaches.

3. Mathematical preliminaries

First, we have accumulated the frequently used symbols/notations in Table 2. This section further presents the mathematical preliminaries utilized for the proposed approach along with their explanations. To effectively handle the shortcomings associated with crisp sets, Zadeh presented the concept of FSs in 1965 (Zadeh, 1965). FSs have been used to model the various sources of imprecision and uncertainty in the data sets.

Definition 1. **- Fuzzy Sets:** A Fuzzy Set 'A' can be expressed as:

$$A(x) = \{(x, \mu_A(x)) \,|\, \forall x \in X\} \tag{1}$$

where, $\mu_A(x)$ is membership value attached to each x, and X is the universal space. Traditional FSs can also be stated as type-1 fuzzy sets (T1 FSs). In T1 FSs, membership grades for the elements are articulated as crisp values in a set. Membership grade is the measure with which an element belongs to a set.

Definition 2. **- Atanassov Intuitionistic Fuzzy Sets (IFSs) (Atanassov, 1986):** Let X be the universe of discourse containing a finite number of objects. Mathematically, an IFSs can be defined over X as follows:

Table 1
Comparative overview of proposed and existing approaches.

Methodology & Refs.	Soft computing techniques	Application	TL technique	Database
MFBRDA (Behbood et al., 2015)	T1 FSs, FNN, SVM	Bank failure detection	Feature based/ Domain adaptation	Banking financial data, 20 newsgroup, etc.
TrCbrBoost (Liu and Li, 2014)	Fuzzy case-based reasoning, TrAdaBoost	Land use classification	Domain adaptation	Middle part of the pearl river delta in china
TL in Takagi–Sugeno fuzzy models (Zuo et al., 2016)	Takagi–sugeno fuzzy model, PSO and DE	Regression	Transfer learning	Housing dataset, synthetic dataset, etc.
Online weighted adaptation regularization for regression (Wu et al., 2016)	FSs, Regression function	Driver drowsiness estimation	Domain adaptation	EEG signals
Transfer generalized hidden-mapping ridge regression (Deng et al., 2014)	FSs, TSK-FLS, NNs	Spam filtering, etc.	Inductive transfer learning	Email spam filtering text dataset, synthetic dataset etc.
Knowledge-leveraged transfer FCM (Qian et al., 2017)	FCM clustering	Texture image segmentation	Unsupervised transfer learning	Brodatz texture
Fuzzy transfer learning (Shell and Coupland, 2015)	Fuzzy logic	Intelligent environments	Transfer learning	Intel Berkeley dataset, de Montfort university robotics dataset, and robotics laboratory data, etc.
Multitask TSK (Jiang et al., 2015)	FCM, FSs	Multitask regression learning	Multitask learning	Glutamic acid fermentation process modeling, multivalued (MV) data modeling, synthetic dataset, etc.
Online fuzzy min–max neural (Seera and Lim, 2014)	Fuzzy min–max neural network		Transfer learning	20 Newsgroups, WiFi time, and Botswana
TSK-TL-FLS (Vapnik, 2013)	TSK-FLS, Regression and classification	Recognition of epileptic EEG Signals	Transductive transfer learning	Epileptic EEG data
Proposed IFTL	IFSs, ELM, SVM, GRNN	GDP prediction	Transfer Learning	CO_2 emission

Table 2
Symbols and Notations with their descriptions.

Symbols/ Notations	Descriptions
IFTL	Intuitionistic fuzzy transfer learning
IFS	Intuitionistic fuzzy set
ELM	Extreme learning machines
SVR	Support vector regression
GRNN	Generalized regression neural network
D_s	Source domain
D_t	Target domain
T_s	Learning task for a source domain D_s
T_t	Learning task for a target domain D_t
TUR	Target unaware regression
IFTL	Intuitionistic fuzzy transfer learning
PL	Target domain output predicted by TUR model
PL_{jF}	Predicted label for class *F* of j^{th} target instance
Dm	Mixture of the intuitionistically fuzzified instances of the source domain and the target domain.
IFSM	Intuitionistic fuzzy similarity matrix
γ	Refinement impact factor
Y	Source domain label
K	Nearest neighbor using minimum Hausdorff distance
N_j^{Dm}	Collection of K nearest neighbors for the j^{th} target, extracted from the set of Dm
MF	Refined target domain labels via IFTL

$$I = \{(x, \mu_I(x), \gamma_I(x)) \,|\, x \in X\} \tag{2}$$

where, $\mu_I : X \rightarrow [0,1]$ and $\gamma_I : X \rightarrow [0,1]$ with the property that $0 \leq \mu_I + \gamma_I \leq 1$, $\mu_I(x)$ denotes the membership degree of $x \in I$, non-membership degree of $x \in I$ is denoted by $\gamma_I(x)$. In fuzzy sets, the following assumption is observed:

$$\mu_I(x) + \gamma_I(x) = 1 \tag{3}$$

However, the hesitancy margin of each element in an intuitionistic fuzzy set ($x \in I$), is defined by:

$$\pi_I(x) = 1 - \mu_I(x) - \gamma_I(x) \tag{4}$$

where, $\pi_I(x)$ denotes the hesitancy margin for x. From Eq. (4), its deducible that $0 \leq \pi_I(x) \leq 1$.

Definition 3. - Intuitionistic fuzzy inference system (IFIS) (Castillo and Melin, 2003): An IFIS is pictorially represented in Fig. 1. Its output (y_f) may be represented as:

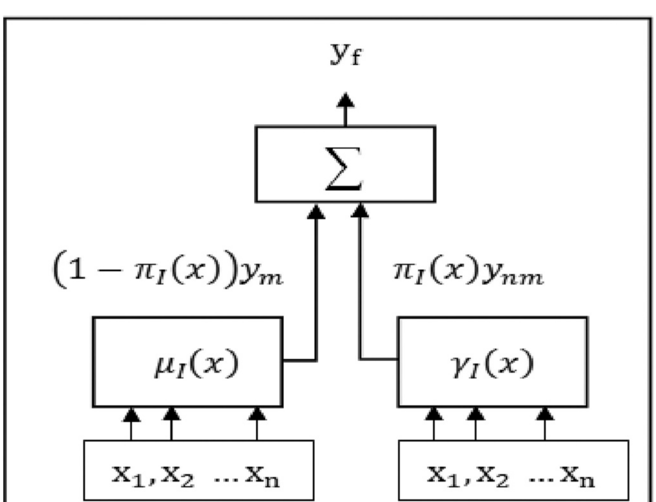


Fig. 1. Structure of a typical IFIS (Songwei et al., 2012).

$$y_f = (1 - \pi_I(x))y_m + \pi_I(x)y_{nm} \tag{5}$$

Here, y_m can be produced from the membership value $\mu_I(x)$ while y_{nm} is generated from the non-membership value $\gamma_I(x)$. If we put $\pi_I(x) = 0$, then $y_f = y_m$, which is the ordinary or T1 FS.

Definition 4. - Hausdorff distance: Let Y and Z be two IFSs defined in $X = \{x_1, x_2, \ldots, x_n\}$ and let $I_Y(x_i)$ and $I_Z(x_i)$ be definite with range [0,1]. Then IFSs Y and Z can be represented as:

$$I_Y(X_i) = [\mu_Y(x_i), \gamma_Y(x_i)] \tag{6}$$

$$I_Z(x_i) = [\mu_Z(x_i), \gamma_Z(x_i)], i = 1, 2, 3, \ldots, n. \tag{7}$$

Then, the Hausdorff distance between $I_Y(x_i)$ and $I_Z(x_i)$, can be written as:

$$H(I_Y(x_i), I_Z(x_i)) = \text{Max}\{(\mu_Y(x_i) - \mu_Z(x_i)), (\gamma_Y(x_i) - \gamma_Z(x_i))\} \tag{8}$$

Thus, the distance $d_H(Y, Z)$ between Y and Z may be represented as:

$$d_H(Y, Z) = {}^1/_n \sum_{i=1}^{n} H(I_Y(x_i), I_Z(x_i)) \tag{9}$$

Dissimilarity will be more between two instances if their Hausdorff distance is more significant. Experimental results proved that the Hausdorff distance function is precise in finding the linguistic similarity in IFSs (Hung and Yang, 2004).

Definition 5. - Domain: A domain is represented by $\{F, P(X)\}$, where $X = \{x_1, \ldots\ldots, x_n\}$, F denotes feature space and P(X) depicts every feature's marginal probability distribution, obtained when all the remaining features are considered constant.

Definition 6. - Task: A Task is defined by $T = \{Y, f()\}$, where $Y = \{y_1, \ldots, y_n\}$ depicts a label space. The predictive function is denoted by $f()$, which is trained using pairs (x_i, y_i) and the labels for new instances are predicted by this learned function $f()$.

Definition 7. - Transfer learning: If T_s is the learning task for a source domain D_s and T_t is the learning task for a target domain D_t, then the rationale of TL is to model a learning function $f_t()$ in D_t by utilizing D_s *and* T_s, when $D_s \neq D_t$ *or* $T_s \neq T_t$.

However, when both domains are precisely similar (i.e. $D_s = D_t$) and their corresponding learning tasks are also similar ($T_s = T_t$), the learning problem changes to a machine learning approach.

4. Proposed methodology

In this section, we will explain our proposed IFTL technique, which is implemented under the framework of label refinement. The complete architecture is depicted in Fig. 2, and the procedure is described in Algorithm 1.

First, target data labels are predicted by a regression model, which is trained using only the source domain data. This regression model is entirely unaware of the target domain's data distribution; hence, it is termed the target unaware regression (TUR) model. Secondly (under the IFTL), both the source and the target data are intuitionistically fuzzified using the IFSs. Then, K-nearest neighbors are chosen using a modified Hausdorff distance function. Finally, these K-nearest neighbors are optimally utilized to refine the labels predicted by TUR for effective learning during TL. Following subsections present detailed explanations of the components of Algorithm 1 (Fig. 2) in a step-by-step manner. Then, we have also discussed the special case of FTL and the computational complexity of IFTL in the following sub-sections.

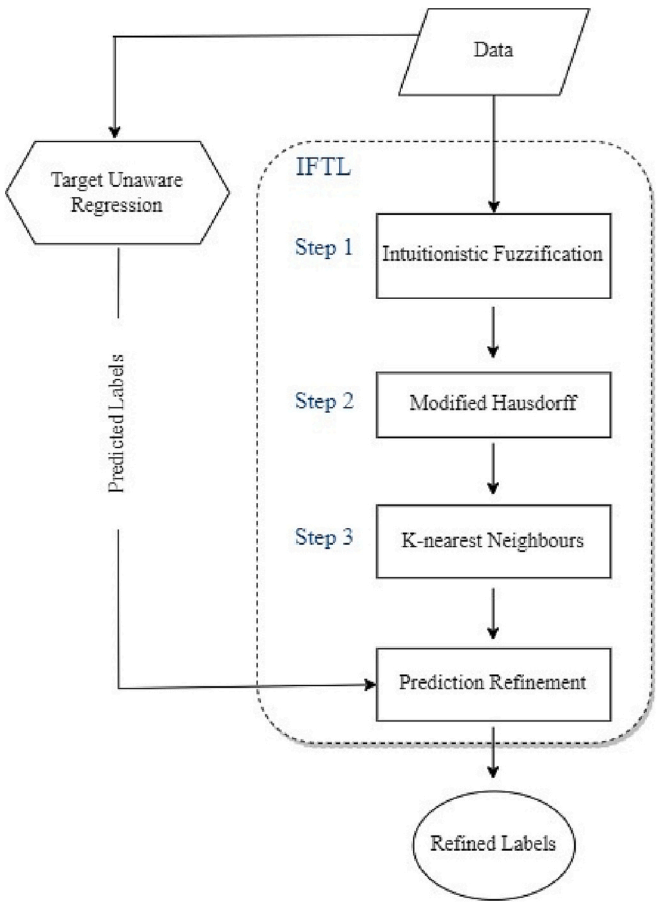


Fig. 2. Proposed IFTL approach under the label refinement framework.

Algorithm 1. Label refinement using IFTL.

Output*: Refined Labels*
Input*: Input Data*

1. Train the target unaware regression model using the source domain and obtain its predicted labels for the target domain.

2. IFTL:

{

Step 1: Perform intuitionistic fuzzification of both the source domain and target domain after hesitancy degree computation.

Step 2: Calculate the distance metric using the Modified Hausdorff distance function of all the data from Step 1.

Step 3: For each target data instance, do the following:

i. Extract K-nearest neighbors using the distance metric from Step 2.

ii. Use labels of these K-neighbors to refine the predicted target label.

}

1. Target unaware regression

Target unaware regression (TUR) is one of the core steps of label refinement, which uses the widely used ML regression approaches for label predictions, such as GRNN, SVR, and ELM. In this paper, the kernel version of ELM is implemented, which is deterministic in nature, as it does not use randomly initialized weights. These approaches are used to prove that the proposed IFTL enhances learning independently of the regression model chosen by exploiting the divergence in different domains.

2. Intuitionistic fuzzy transfer learning (IFTL)

The proposed IFTL refines target domain labels predicted by the TUR model. Features in the target and the source domain are exploited using IFTL algorithm. These features act as bridges between the source and the target domains. For the source domain hesitancy degree, we first find the variance (*Var*) of the differences between the source domain and the target domain features. Then we normalize it by its feature mean. Later, its absolute (*Abs*) value is chosen as the source domain hesitancy degree. Mathematically, for each feature F, the hesitancy degree is computed as follows:

$$Source\_hesitancy\_degree(F) =$$

$$Abs\left(\left(Var\left\{\frac{Source\_domain\_feature(F) - Target\_domain\_faeture(F)}{mean(Source\_domain\_faeture(F)}\right\}\right)\right.$$

$$Target\_hesitancy\_degree(F) = \frac{1}{Source\_hesitancy\_degree(F)}$$

This also gives $\alpha > 1$, which brings out the anomaly in Yager's generating function for its incorporation into IFTL. Hence, it efficiently captures the uncertainty that arises due to the variations in the data distribution across the source and the target domains. Thereby improving model prediction during transferral of the learned knowledge. As the knowledge is being transferred from the source domain to the target domain, the hesitancy degree of the target domain should be inverse to that of the source domain hesitancy degree. This helps in judiciously utilizing the already acquired leaning on a new domain. This hesitation degree is calculated for each feature of the source domain and the target domain, separately. Further, intuitionistic fuzzification of the source and target domain features are performed using Yager's generating function, as shown below in Step 1.

Step 1: *Intuitionistic fuzzification.*

A parametric intuitionistic fuzzy complement technique called 'Yager's generating function' efficiently finds the value of the hesitancy margin ($\pi_I(x)$) based on the experimental results (Ghosh et al., 2021; Kumar et al., 2016). It has the flexibility to properly tune the parameters of any application of real-world application by parameter optimization. Here, the hesitancy margin is controlled using membership value (Chaira, 2011). The Yager's function is defined as:

$$f(x) = x^{\alpha} \tag{10}$$

where α is the hesitancy degree. The Yager's generating function in conjunction with Atanassov intuitionistic fuzzy complement is denoted in the following way (Ghosh et al., 2021):

$$\gamma_Y(x_i) = (1 - x_i^{\alpha})^{\frac{1}{\alpha}} \tag{11}$$

where, $\alpha > 0$. For $x_i = 0$ and $1, \gamma_Y(0) = 1$ and $\gamma_Y(1) = 0$.

Now, IFSs can be expressed as:

$$I_Y(x_i) = \left\{ x_{i,}\mu_Y(x_i), (1 - \mu_Y(x_i)^{\alpha})^{\frac{1}{\alpha}} \right\} \tag{12}$$

The hesitancy margin ($\pi_Y(x)$) is given by:

$$\pi_Y(x) = 1 - \mu_Y(x) - (1 - \mu_Y(x_i)^{\alpha})^{\frac{1}{\alpha}} \tag{13}$$

where $0 \le \pi_Y(x) \le 1$. In IFSs, hesitancy margin should be within the interval [0,1] as stated in Eq. (13). However, if $\alpha > 1$, then $\pi_Y(x) \le 0$. This is against the principles of IFSs that lead to an anomaly in Yager's generating function.

Let us consider the following example: If membership value ($\mu_Y(x_i)$) of x_i in IFS Y is 0.2 and $\alpha = 2$, then the non-membership value is: $(1 - \mu_Y(x_i)^{\alpha})^{\frac{1}{\alpha}} = 0.979$ and $\pi_Y(x) = -0.179$, which results in $\pi_Y(x) < 0$. Here, the anomaly is: $\mu_Y(x) + (1 - \mu_Y(x_i)^{\alpha})^{\frac{1}{\alpha}} > 1$. This is a contradictory formulation in the context of FS theory, since it has a condition that maximum value obtained by adding $\mu_I(x)$ and $\gamma_I(x)$ is 1. Here, the hesitancy margin is <0 i.e. $\pi_Y(x) < 0$. However, this anomaly is handled by the introduced IFTL technique. The logic behind using it is that it may help in penalizing those instances in the target domain which have high cohesion but are of no use. Therefore, it helps in enhancing learning while transferring knowledge. Then, the intuitionistically fuzzified data points of the source and target domain are mixed with increasing the data size for better learning. We further propose a modified Hausdorff distance metric to compute the similarity between the target domain data and the mixed domain data. This distance matrix is termed an Intuitionistic Fuzzy Similarity Matrix (IFSM).

Step 2: *Proposed modified Hausdorff distance.*

Let Y and Z be two IFSs in $X = \{x_1, x_2, \ldots, x_n\}$ and $I_Y(x_1)$ and $I_Z(x_i)$ with range [0, 1] are defined as follows:

$$I_Y(x_i) = [x_{i,}\mu_Y(x_i), \gamma_Y(x_i)] \tag{14}$$

$$I_Z(x_i) = [x_{i,}\mu_Z(x_i), \gamma_Z(x_i)], i = 1, 2, 3\ldots, n. \tag{15}$$

$\pi_Y(x)$ is hesitation margin as defined below:

$$\pi_Y(x) = 1 - \mu_Y(x_i) - \gamma_Y(x_i) \tag{16}$$

Let $H_m(I_Y(x_i), I_Z(x_i))$ is modified Hausdorff distance between $I_Y(x_i)$ and $I_Z(x_i)$ and is defined as:

$$H_m(I_Y(x_i), I_Z(x_i)) = |\pi_Y - \pi_Z(x)| * [Max\{(\mu_Y(x_i) - \mu_Z(x_i)), (\gamma_Y(x_i) - \gamma_Z(x_i)\}] \tag{17}$$

The modification in Hausdorff distance function is the addition of Coefficient ($|\pi_Y - \pi_Z(x)|$) in Eq. (17). Here, Eq. (9) is disregarded as the similarity is being calculated between two data points, and not between two sets. In addition, the hesitation margin difference ($|\pi_Y(x) - \pi_Z(x)|$) is added in order to exploit the influence arising due to Yager generating function anomaly.

This proposed modified Hausdorff distance is then used for IFSM computation. Using IFSM, for the each of the j^{th} instance, a set N_j^{Dm} is selected in this step. This set N_j^{Dm} consist of K-nearest neighbors from the set Dm, and its elements are chosen using their minimum value in the IFSM as denoted by the following Eq. (18):

$$N_j^{Dm} = \left\{ n_{1:K} = argmin\ \{IFSM(j, *)\} \right\} \tag{18}$$

Step 3: *Label refinement.*

From the above set of mixed features, k-nearest neighbors are extracted for efficient label refinement in the proposed approach. We have compiled the label refinement steps in the form of a flowchart in Fig. 3. The input and output terminologies for the IFTL label refinement are summarized in Table 3. Various stages of this flowchart are explained as follows:

Step A: Choose an instance of the target domain iteratively and for each instance, follow Steps B to E.

Step B: For the instance chosen in Step A, select the class to which it belongs and proceed to Step D.

Step C: For all instances in the target instance, calculate IFSM (Intuitionistic fuzzy similarity matrix) with respect to each datapoint of Dm.

Step D: Refine the predicted labels using Eq. (19):

$$MF_{JF} = \gamma \left(\sum_{x_0 \in N_j^{Dm}} IFSM(x_j, x_0) * \delta P_{jF} \right) + (1 - \gamma) PL_{jF} \tag{19}$$

If $x_0 \in D_s$, then,

$$\delta P_{jF} = PL_{jF} - \mu_{Y_F}(x_0) \tag{20}$$

else if $x_0 \in D_t$, then

$$\delta P_{jF} = 0\ (\text{as } PL_{jF} - PL_{jF}) \tag{21}$$

Here, MF_{JF} denotes the IFTL-refined F^{th} class label of the j^{th} instance of the target domain. The refinement of the predicted labels is done via $\sum_{x_0 \in N_j^{Dm}} IFS(x_j, x_0) * \delta P_{jF}$ in the Eq. (19).

It denotes that for all $x_0 \in N_j^{Dm}$, multiply δP_{jF} with the IFSM values of the x_j $(x_j \in D_t)$ and $x_0 (x_0 \in D_{Dm})$, and then take their summation. The impact of refinement in Eq. (19) is controlled by γ (gamma). The value of γ if chosen too high, then the importance of good prediction by the TUR model is not optimally utilized, and if γ is too small, then IFTL will give results nearly equivalent to that of the prediction of ML approaches.

Therefore, it is kept constant at 10% for every experiment of IFTL. This value of γ (a hyper-parameter) is experimentally chosen and then kept constant at 10% for all experiments of IFTL. When $\gamma = 0$, then $MF_{JF} = PL_{jF}$, which represents the prediction done by TUR. If Eq. (19), results are in a negative value, then we discard it, and no update is done. N_j^{Dm} is a set of those instances from the mixture set (Dm), that are the K-nearest neighbors of the j^{th} instance of the target domain. As nearest neighbors are taken from the mixture of all the intuitionistically fuzzified instances of the target and the source domain, i.e., Dm, δP_{jF} is calculated by two ways: First, if x_0 in set N_j^{Dm} belongs to the source domain (D_s), whose labels for F^{th} class, $\mu_{Y_F}(x_0)$ are available, then δP_{jF} is

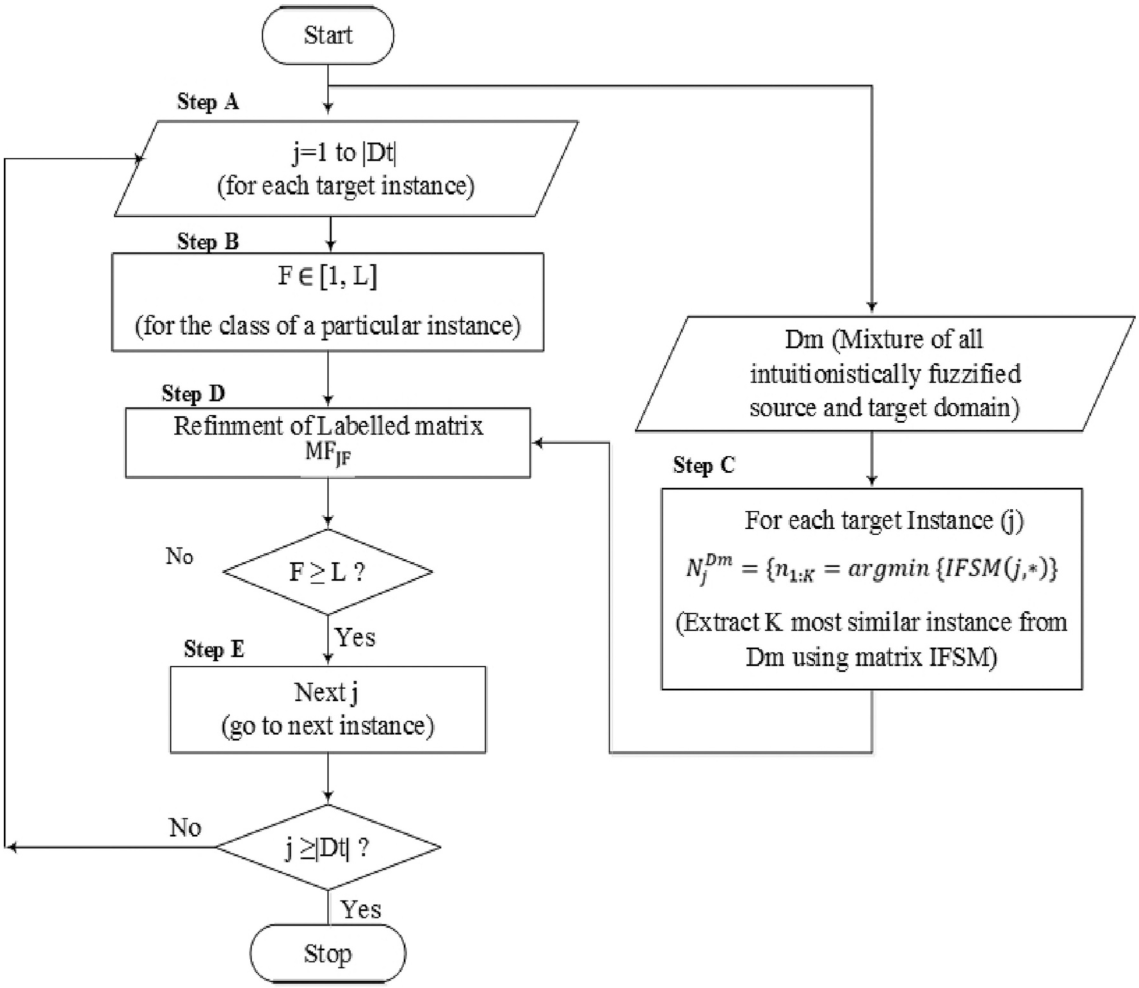


Fig. 3. Algorithmic flow of IFTL label refinement.

Table 3
Input and Output terminologies.

Input	Output
D_s, D_t, PL, PL_{jF}, Dm, IFSM, γ, Y, K, N_j^{Dm}	MF

derived using Eq. (20). Here, $\mu_{Y_F}(x_0)$ depicts F[th] class of $x_0 \in D_s$. Second, if x_0 belongs to the target domain, then δP_{jF} in Eq. (21) is equal to zero.

Step E: Repeat Step A till the labels of all the target instances are not completed; else, stop.

4.1. Special case: FTL (when hesitancy margin = 0)

When the hesitancy margin is not considered (i.e., taken as zero), IFTL reduces to FTL, which considers ordinary FSs for fuzzifying the dataset and employs Euclidean distance instead of modified Hausdorff distance for determining the K-nearest neighbors. The FTL is also studied to show the necessity of hesitancy margin in handling human cognitive behavior, which is efficiently achieved using IFTL. Thus, FTL would not be effective during TL while using previously acquired experience to solve a new and related task, which our proposed IFTL can efficiently accomplish.

4.2. Computational complexity

While analyzing the computational complexity of IFTL, it is observed that Step A to Step E in Fig. 3 is repeated $|D_t|$ number of times. All these steps have an individual computational complexity of O (1). Step B is ignored, as number of classes are not always too large (an assumption that they are <100). Step C in Fig. 3 has a runtime equal to that of the k-nearest neighbors, i.e., $O(|\mathrm{Dm}|d + |\mathrm{Dm}|K)$. Here, d is the dimension of the dataset. Modified Hausdorff distance metric uses intuitionistically fuzzified values in terms of membership, non-membership, and hesitancy margin of each feature of a data point. However, it does not affect the overall computational complexity. Due to the three different components in the modified Hausdorff distance, the computational complexity is $(3^*d) = O(d)$, which is equivalent to the complexity of the Euclidean distance.

So, the total computational cost involving Step A to Step E is $O(|D_t|^*(|\mathrm{Dm}|d + |\mathrm{Dm}|K))$. Also, the intuitionistic fuzzification of all the source and the target data has a computational complexity of $O(|D_t| + |D_s|)$ since each data point can be intuitionistically fuzzified cost of $\Theta(1)$, which is equivalent to $O(|\mathrm{Dm}|)$ as $|\mathrm{Dm}| = (|D_t| + |D_s|)$.

Therefore, the computational complexity of IFTL is computed by adding the computational cost of Step A to Step E (i.e., $O(|D_t|^*(|\mathrm{Dm}|d + |\mathrm{Dm}|K))$) and the computational cost of intuitionistic fuzzification (i.e. $O(|\mathrm{Dm}|)$). Hence, the computational complexity of IFTL is $O(|\mathrm{Dm}| + (|D_t|^*(|\mathrm{Dm}|d + |\mathrm{Dm}|K)))$.

5. Experiments and results

This section presents the experimental setup considered for evaluating the proposed IFTL. We first discuss the dataset description for GDP prediction from carbon emission data. For the experiments, we start by using widely used ELM for the label prediction in the TUR. Later, IFTL and FTL are implemented using GRNN and SVR as the TUR model. It highlights that whenever training and testing data (including labels) have huge distribution differences, then IFTL outperformed FTL by being the most accurate due to the crucial role of hesitancy.

Further discussions analyze IFTL and FTL for their abilities to improve learning by restricting overconfidence (controlling hesitancy) during TL by considering GRNN and SVR as the TUR. Overconfidence in ELM, GRNN, or SVR arises when they make predictions on a dataset (target domain) that has a huge data distribution difference from the source domain data on which they are trained. In this scenario, they just use their training experience to make predictions during testing without considering the distribution divergence in the testing dataset from the training dataset. We conclude this section with the execution time analysis of the approaches.

a) *GDP prediction using only CO_2 emission data*[2]

For evaluating the performance of machine learning algorithms, well-known benchmark datasets are utilized, which are partitioned (by KFold or StratifiedKFold, etc.) in test data (target domain) and training data (source domain). Here both these domains are drawn from the same domain distribution; hence both of them are equal in their distribution i.e. $D_s = D_t$ and $T_s = T_t$. However, IFTL is hypothesized for transfer learning, and to evaluate its performance precisely, the source domain and target domain must be highly diverse but related i.e. $D_s \neq D_t$ or $T_s \neq T_t$ or both.

So, this paper takes up a very innovative real-life application in which the IFTL approach is analyzed to predict the GDP of a developing county (target domain) using only its carbon dioxide (CO_2) gas emissions data. However, the training is performed on developed countries (source domain). For that, the European Union's data comprising GDP and its CO_2 emission are used as the source domain. For the testing phase, target domain data is collected from India. European Union (EU) comprises 28 countries in Europe, which are highly developed in terms of infrastructure and economy. As compared to any EU country, India is a developing Asian country, with twice the total population that of the European Union. Also, the EU and India have substantial geographical, educational, cultural, political, and environmental differences. Hence, the data distribution of these two domains (CO_2 emission data and per capita GDP) are highly different from each other i.e. $D_s \neq D_t$ and $T_s \neq T_t$. Thus, it is a perfect case for a TL problem. All three regression models (GRNN, SVR, and ELM) are trained on source domain (EU) data and tested on target domain (India) data. The experimental setup encompasses input and output parameters, as shown in Table 4 and Table 5, respectively. These input and output parameters of source and target domains are collected from the World Bank database of the EU and India (http://data.worldbank.org/).

Table 4
Input Parameters.

Feature	Input Parameter
1	CO_2 emissions due to gaseous fuel consumption (% of total)
2	CO_2 emissions due to liquid fuel consumption (% of total)
3	CO_2 emissions due to solid fuel consumption (% of total)
4	CO_2 emissions (in metric tons per capita)

Table 5
Output Parameters.

Feature	Output Parameters
Output	GDP per capita (at current US$)

The GDP of a country is directly proportional to its carbon dioxide emissions, as empirically demonstrated in numerous studies (Acheampong, 2018a; Govindaraju and Tang, 2013a; Stern, 2010a), (Chen et al., 2018a; Saidi and Hammami, 2015). So, to efficiently reflect this relationship during fuzzification, each feature of a domain is fuzzified by dividing it by its maximum value. Thus, a higher membership value ensures a higher per capita GDP value and vice versa. Therefore, for intuitionistic fuzzification, the source domain data (EU) and target domain data (IND) features (Table 4) are divided with their respective maximum values to compute the corresponding membership values. The non-membership values are calculated using Eq. (11) of Yager's generating function, in which α (the hesitancy degree) is calculated by exploiting the data distribution differences of source (EU) and the target domain (IND) as explained in Subsection C of Section 4. Then, K-nearest neighbors (Step C of Fig. 3) are chosen for each of the target (IND) data points using minimum distance values calculated by Eq. (17) from the mixture of intuitionistically fuzzified source (EU) and the target data (IND). In the end, these K-nearest neighbors are optimally utilized using Eq. (19) to refine the labels predicted (Step D in Fig. 3) by the TUR model (best of the ELM, GRNN, and SVR, i.e., ELM) for effective TL. During FTL, ordinary fuzzification is performed instead of intuitionistic fuzzification, and Euclidean distance is calculated using membership and non-membership values. The rest of the procedure is similar to that of IFTL. For both IFTL and FTL, the refinement impact factor 'gamma' is kept constant at 10% ($\gamma = 0.1$), which governs the impact of refinement in IFTL.

From the experimental result (as shown in Fig. 4 (a)), it is found that the RMSE for SVR, GRNN, and ELM are 1989.4, 610.3343, and 362.2708, respectively. The comparison between ELM and IFTL with actual GDP values is shown in Fig. 5. Here, all these regression models (GRNN, SVR, and ELM) are first trained using only EU data, and then these trained models are used to predict the GDP. There is an overconfidence in the predictions of these models as they simply use their experiences acquired from training without considering that both India and the EU have considerable differences in their data distribution. This overconfidence of these regression techniques is restricted by IFTL, which optimally exploits these data distribution differences to its advantage by using hesitation degree for modeling by exerting caution in utilizing their experience in a new environment.

In Fig. 5, it can also be observed that IFTL attempts to reduce the overconfidence of ELM predictions by refining it and pulling it toward the actual GDP values, which FTL also could not perform as it doesn't have a hesitancy degree. FTL and IFTL nearly overlap until the 48th point on the horizontal x-axis (the year 2007). However, as the difference between the predicted values (ELM prediction) and the actual GDP values grows after the 48th point, IFTL outperforms FTL by pulling the labels more effectively toward actual values. This behavior is consistent till the year 2013.

For IFTL, RMSE of 279.1 can be seen in Fig. 4 (a) when the TUR uses ELM (also denoted as ELM_IF). However, for FTL (or ELM_F), the RMSE is 366.83, which is even worse than the RMSE of 362.2708 as predicted by ELM. It behaves poorly after the 48th point on the horizontal x-axis of Fig. 5 once the difference between the actual GDP values and the predicted GDP values grows significantly. Hence, IFTL (ELM_IF) performed exceptionally well and secured at least 22.9% improvement in the prediction accuracy over ELM and FTL (ELM_F). Furthermore, it also validated its abilities to improve learning during TL, in which USA data is used as the source domain and India as the target domain.

[2] Benchmark datasets are investigated for ML methodologies where training and testing data are drawn from same distribution. The proposed IFTL is a generic approach based on the principle of transfer learning which may primarily work in any real-world application with training and testing domains having different data distributions.

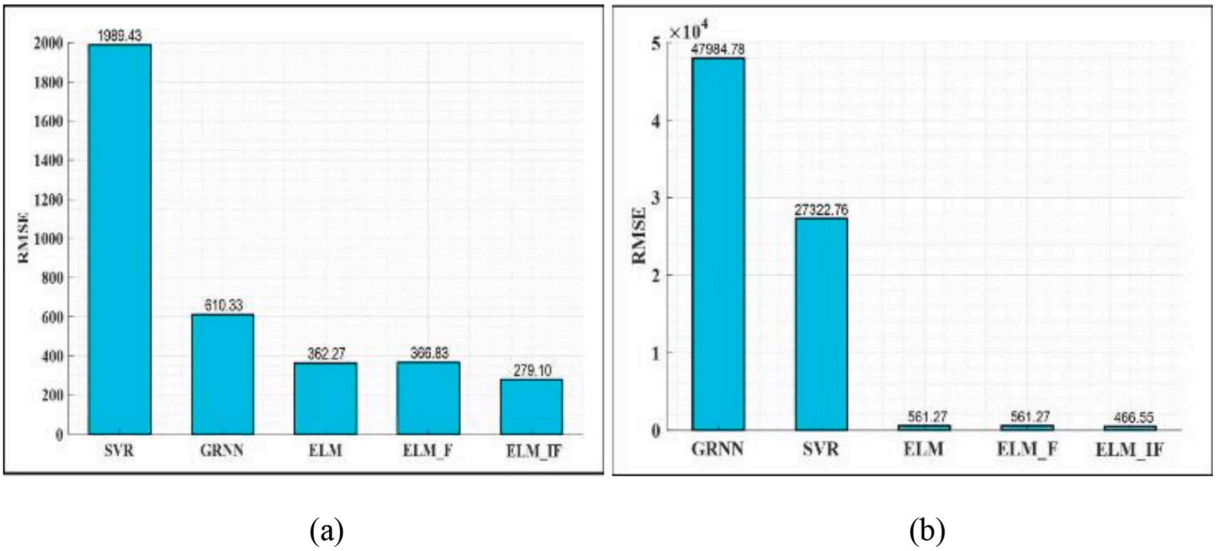


Fig. 4. IFTL, source domain: (a) EU, (b) USA.

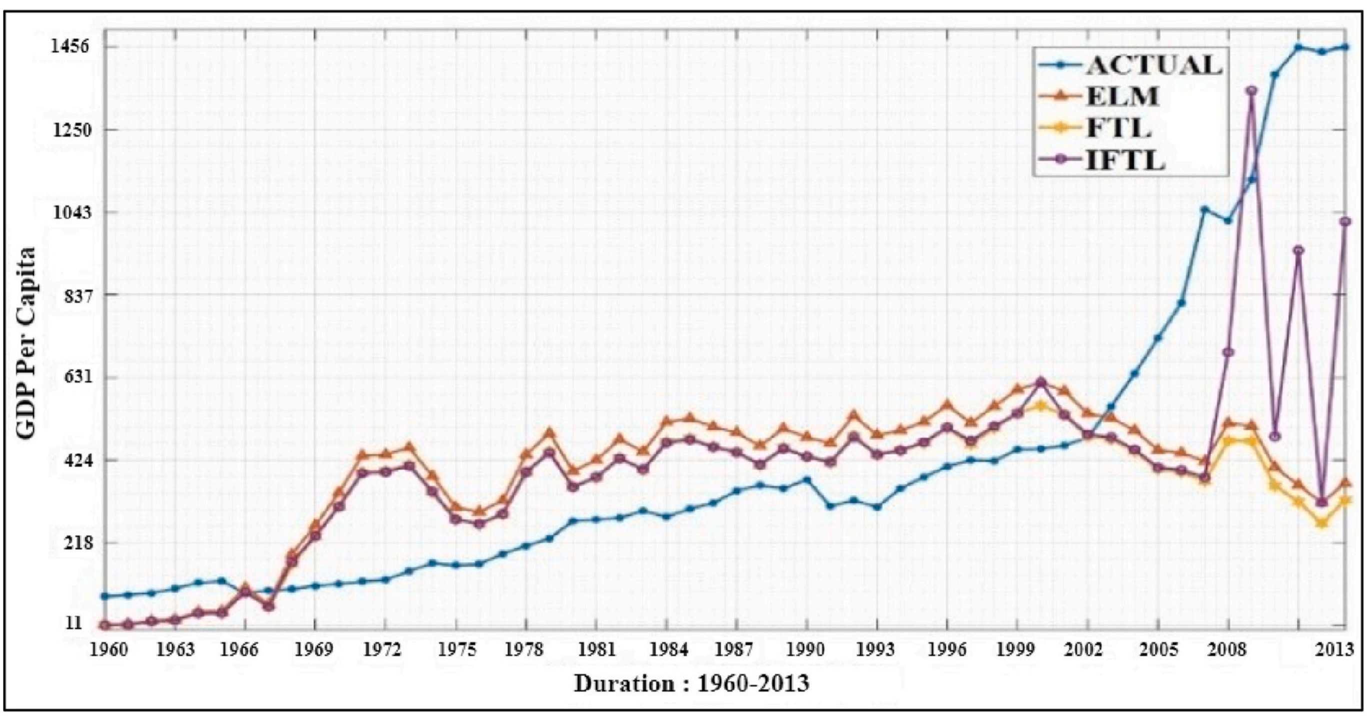
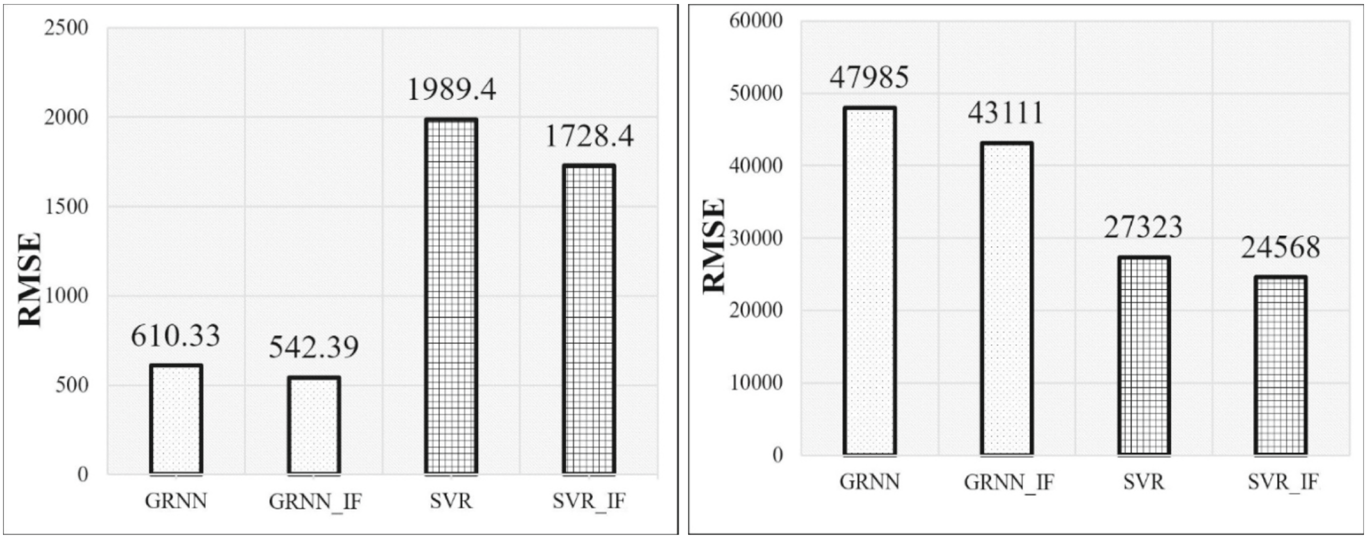


Fig. 5. Performance analysis of ELM, FTL and IFTL (source domain: EU).

As shown in Fig. 4 (b), IFTL (ELM_IF) predicted more accurately with an RMSE value of 466.5, whereas the best regression model of ELM returned RMSE of 561.27. Again, FTL (ELM_F) could not improve the results of ELM. It predicted similar RMSE values. Thus, IFTL outperforms the standalone regression model and ordinary FSs-based FTL by suitably restricting overconfidence, thereby enhancing learning during TL and providing more accurate predictions.

b) *IFTL Refinement over GRNN and SVR*

In previous results, ELM is chosen as TUR, and IFTL is implemented on the ELM predictions to improve its knowledge transfer. Here, GRNN and SVR are chosen one by one as a TUR model in the process described in Fig. 2. Then, IFTL with the same parameter settings (as was in the case of ELM in the previous subsection, where $\gamma = 0.1$) is applied over their prediction. The improvement of IFTL over these approaches are depicted in Fig. 6, where 'GRNN_IF' denotes the error in IFTL prediction when IFTL is implemented, choosing GRNN as TUR. Similarly, 'SVR_IF' denotes the IFTL error when IFTL is implemented, choosing SVR as TUR.



(a) (b)

Fig. 6. IFTL refinement, source domain: (a) EU, (b) USA.

Table 6
IFTL refinement effectiveness.

	RMSE	IFTL_error	Improvement (%)
Source Domain: EU and Target Domain: India			
GRNN	610.33	542.39	11.13
SVM	1989.4	1728.4	13.12
ELM	362.27	279.1	22.96
Source Domain: USA And Target Domain: India			
GRNN	47,985	43,111	10.16
SVM	27,323	24,568	10.08
ELM	561.27	466.55	16.88

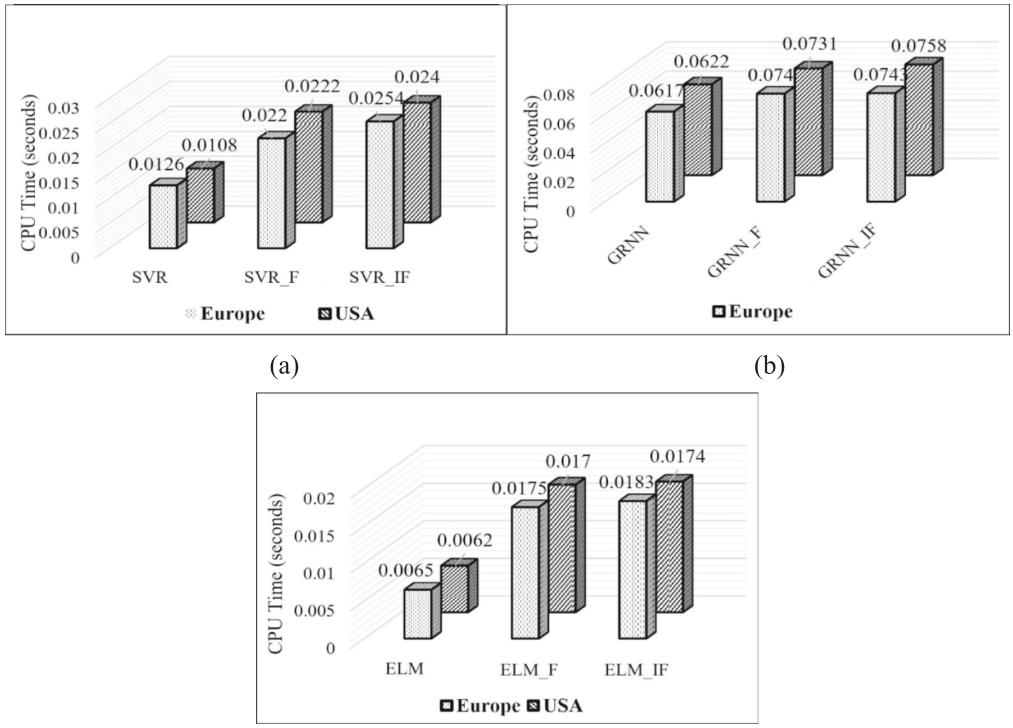
Fig. 6 (a) depicts the experimental results when models are trained using EU data as the source domain and India data as the target domain. It can be seen that GRNN prediction error (RMSE) is 610.33, which is significantly improved when IFTL is applied, which reduces the prediction error, i.e., GRNN_ IF error (RMSE) is 542.39. Similarly, the SVR error (RMSE) is 1989.4, while the SVR_IF error is 1728.4. Fig. 6 (b) demonstrates the results of the model trained using USA data (source domain) and tested on INDIA data (target domain). Here, the GRNN prediction error is 47,985, which is greater than the GRNN_IF error of 43,111. Similarly, the SVR error is 27,323, which is poorer than the SVR_IF error of 24,568. To conclude, the above results indicate the improved transfer of knowledge through IFTL over GRNN and SVR.

Hence, it is validated that IFTL is a refinement approach that can be applied to the predictions done by any other existing techniques. Table 6 summarizes these results along with ELM predictions. Third column of this table depicts the percentage (%) reduction in RMSE, which ranges from 10% to 23%. The proposed IFTL is used as a refinement over the results of GRNN, SVR, and ELM trained in different source domains while keeping γ (Gamma: the Refinement Impact Factor) fixed at 0.1. IFTL carefully considers the vast differences in the data distribution of the source domain and the target domain by which it ensures better TL in GRNN, SVR, or ELM.

c) *Execution time analysis*

The execution time of IFTL and FTL are analyzed with respect to their TUR model. In Fig. 7 (a), execution time of SVR is shown for two different experiments. The CPU time of 0.0126 *s* denotes the execution time of SVR, which is trained using EU as the source domain and INDIA as the target domain. The execution time of 0.0108 *s* is observed for SVR when the source domain is changed to USA. The SVR_F and SVR_IF denote the FTL and IFTL execution time, respectively, in which SVR is used as TUR model. SVR_F executes in 0.22 and 0.222 *s*, when the source domain is either EU or USA, respectively. Similarly, SVR_IF finished with an execution time of 0.0254 *s* when the source domain is EU and 0.024 *s* when USA is the source domain.

In the same way, execution times (including training and testing) of FTL and IFTL are computed by considering GRNN and ELM as TUR models. The respective results are compiled in the Fig. 7 (b) and Fig. 7 (c), respectively. It can be seen from the Figs. 7 (a - c), that IFTL is not very time intensive as its extra computational time is almost equivalent to that of SVR, i.e., it takes only about 120 CPU time (seconds) extra, in addition to that of the TUR in all these cases. Moreover, FTL and IFTL execution times are comparatively equal. These experiments have been implemented in MATLAB on a workstation having Xeon® processor with 64 GB RAM on Windows 7 OS.



(a) (b)

(c)

Fig. 7. Execution time using (a) SVR as TUR, (b) GRNN as TUR, (c) ELM as TUR.

Table 7
Comparative Overview of Some Recent Works Utlizing Ecological Attribute.
(Highlighting computational tools/aproaches used, applications considered, and ecologocal implications)

Ref.	Ecological Attribute	Computational Tools/ Approaches	Application	Ecological Implications
(Khatua et al., 2020)	Water temperature and pollution, Global warming, and Fish harvesting	Fuzzy Rule based inference model	Production of Hilsa fishes	• High temperatures and increasing water pollution affect the production of Hilsa fishes significantly. • A novel fuzzy mathematical model is utilized to form a relationship between global warming and Hilsa production
(Lu et al., 2021)	Sound of marine mammals	AlexNet with TL	Detection and classification of marine mammal sounds	• The proposed hybrid ML model detects and classifies the sounds of different marine mammal species which inhabit overlapping living areas. • It helps in monitoring marine mammal populations and adds to the conservation efforts.
(Guo et al., 2012)	CO_2 emissions induced by fossil fuel combustion	Input-output analysis	Investigation of embodied CO_2 emissions	• The study examines the concrete emissions inventory of Beijing's economy in 2007 and analyses the CO_2 emissions resulting from fossil fuel combustion. • This provides valuable information for understanding the environmental impact of energy consumption and helps in formulating effective policies to reduce CO_2 emissions.
(Debnath et al., 2018)	Air pollution	IT2 FSs, weighted interval type-2 fuzzy reasoning	Evaluation of air quality status, assessment of Air Quality Index (AQI)	• Proposed work assesses the air quality where concentration and toxicity levels are modelled using IT2 FSs. • Helps to identify contributing components for the resulting air quality, which further may enhance decision-making processes related to environmental policies and health-related issues.
(Wang et al., 2022)	Reproductive performance of sows	Acoustic data, deep convolutional neural network (CNN) algorithms	sow oestrus sound monitoring	• This work utilizes acoustic data and deep CNN to identify oestrus and non-oestrus sounds of sows. • Also, provide ways for monitoring and early warning systems on pig farms. • Better reproduction capability in sows may upscale the economic efficiency and production level of pigs. • A sustainable perspective for the pig farming operations.
Proposed IFTL	Carbon emission of developed and developing countries	IFS and TL	GDP Prediction	• The ecological implications of the proposed work are directly associated with the increasing CO_2 emissions resulting from industrialization. • Though it has negative effects on the environment, the proposed IFTL utilize its attribute to predict GDP of economies with unavailable information.

6. Discussion

Literature is filled with numerous prediction and forecasting studies utilizing uncertainty-based approaches (Pearson and Clark-Wolf, 2023; Tan and Zhao, 2023). Both traditional and non-traditional FSs have been effectively used for prediction studies in various ecological applications (Debnath et al., 2018; Khatua et al., 2020). Khatua et al., (2020) studied the impact of the environmental conditions on the production of Hilsa fishes using a traditional fuzzy logic-based inference system. Debnath et al., (2018) utilized an interval type-2 fuzzy set based inference system for assessing air quality. In addition, significant improvements in the GDP forecasting accuracy were observed in several fuzzy based approaches, e.g., (Jha et al., 2022; Mirbagheri, 2010), etc. Traditionally, GDP prediction or forecasting has depended on several economic indicators and econometric models. However, the integration of environmental factors such as carbon emission has brought added significance to the problem of GDP prediction as it has helped to address the limitation of unavailability of those economic indicators in war-torn and closed countries (Tang, 2013b). The literature has constantly provided scientific evidence of the proportional relationship between carbon emission and the GDP growth of a nation (Adedoyin et al., 2021; Tang, 2013b). Similarly, many such studies with tremendous ecological implications have been carried out in recent times by utilizing ecological attributes and indicators. In Table 7, we have summarized some recent studies, including the proposed IFTL, which have utilized ecological attributes/indicators, and have demonstrated their individual ecological implications.

The ecological implications of the proposed work are directly associated with the increasing CO_2 emissions from industrialization. Though it has adverse effects on the environment, the proposed IFTL utilizes its attribute to predict the GDP of economies with unavailable information. Specifically, this work aims to predict a country's GDP using only its carbon emission data. To precisely predict a nation's GDP, this paper explored the non-linear relationship among ecological indicators such as carbon emissions from gaseous fuel, liquid fuel, and solid fuel. The differential usage of these fuels is also related to the economic condition of a nation. Solid fuel is more predominant in poorer countries. When a country economically progresses, it moves toward gaseous fuel (cleaner fuel) from liquid or solid fuels. This work has strongly established the need for computational intelligence based tools and approaches (e.g., IFS and TL) to utilize ecological indicators such as carbon emission to efficiently predict the economic aspect (GDP) of a country. By studying this correlation between CO_2 and GDP, we have also presented valuable insights into the environmental consequences of commercialization and industrialization.

7. Conclusion and future works

This study has found its motivation from the increased level of carbon emissions due to industrialization and how economic indicators are associated with it. The ecological implication is clearly the environmental degradation due to these carbon emissions, though it is strongly associated with economic development. The available carbon emission and GDP data of a country are not sufficient enough to build a robust predictive machine learning (ML) model to effectively learn the nonlinearity in the data from different domains. Therefore, to precisely estimate the per-capita GDP of a nations using only its CO_2 emission data, this paper proposed a unique deterministic algorithm called 'Intuitionistic fuzzy transfer learning (IFTL)' for enhanced learning during TL. The anomaly observed regarding the hesitancy margin of Yager's generating function is optimally incorporated in modified Hausdorff distance to avoid overconfidence for better TL. We also proposed a novel way for calculating the intuitionistic hesitation degree by utilizing the variance in the data differences among different domains. A novel fuzzy transfer learning (FTL) approach is also introduced, which unlike IFTL, uses only FSs and Euclidean distance. FTL is incorporated to emphasize the critical role of hesitancy in utilizing previously learned skills to solve a new and related task. IFTL includes a hesitation degree for modeling the human tendency to exert caution while using their experience in a new environment. FTL does not contain any hesitancy degree to model this human behavior.

IFTL is examined for its efficacy compared to FTL and three ML approaches, viz., SVR, GRNN, and ELM, on the real-world problems of GDP prediction only from carbon emission data. These ML approaches only use their training experience to make predictions on the testing dataset (i.e., target domain dataset) without considering its distribution differences with the training dataset (i.e., source domain dataset). At the same time, the proposed IFTL refines their prediction by optimally incorporating these data distribution differences for efficient TL from one domain to another, along with human behavior of using caution. For GDP prediction, IFTL performance has been tested where it is trained on the data of developed nations and tested over the data of a developing nation. When ELM is considered as Target unaware regression (TUR), IFTL performed exceptionally well and secured at least 22.9% improvement in the prediction accuracy over traditional ELM and FTL. The overall IFTL refinement effectiveness can be seen by the percentage (%) reduction in RMSE, which ranges from 10% to 23% when GRNN, SVR, and ELM are considered as TUR.

We have found that the proposed approach efficiently captured the uncertainty produced by extreme variations in data distribution or the predictive tasks across the source and target data. It suitably restricted overconfidence for better learning while transferring the learned knowledge. IFTL outperformed FTL, where the training and testing domains (including labels) have huge data distribution differences. This validated the effectiveness of hesitancy in restricting overconfidence during TL. IFTL is also thoroughly analyzed for its asymptotic computational complexity and execution time. The limitation of IFTL is that it adds additional computation by calculating membership, nonmembership, hesitancy margin of each feature, and the distance metric for refining the output labels predicted by other ML approaches.

This work has successfully demonstrated the utility of ecological attributes in addressing the limitation of GDP prediction approaches due to the unavailability of economic indicators for some countries. In addition, the implication of ecological attributes may find their due utility in many such fields beyond the scope of this study, where the ecological considerations may play a significant role. Therefore, future works may focus on exploring the proposed technique in several other real-world problems, such as economics, finance, human development index, text analytics, bank failure predictions, etc. IFTL shall be analyzed by considering different distance metrics for k-nearest neighbors extraction. Moreover, the proposed IFTL approach shall be used to build TL models for estimating the GDP of war-torn countries or inaccessible countries since their reliable macroeconomic data is not entirely satisfactory in calculating their GDP. This will undoubtedly help in the efficient socio-economic analysis of the people residing in these countries.

Declaration of Competing Interest

None.

Data availability

The data information is mentioned in the manuscript. If not clear, datafile may be requested from any of the authors.

Acknowldegement

The first author gratefully acknowledges the financial assistance received from the Department of Science and Technology, Government of India, in the form of INSPIRE research fellowship. The authors are also thankful to Mr. Kamakhya Prasad Seal for proof reading the manuscript.

References

Acheampong, A.O., 2018a. Economic growth, CO2 emissions and energy consumption: what causes what and where? Energy Econ. 74, 677–692.

Acheampong, A.O., 2018b. Economic growth, CO2 emissions and energy consumption: what causes what and where? Energy Econ. 74, 677–692.

Adedoyin, F.F., Ozturk, I., Agboola, M.O., Agboola, P.O., Bekun, F.V., 2021. The implications of renewable and non-renewable energy generating in sub-Saharan Africa: the role of economic policy uncertainties. Energy Policy 150, 112115.

Atan, Ö., Kutlu, F., Castillo, O., 2020. Intuitionistic fuzzy sliding controller for uncertain hyperchaotic synchronization. Int. J. Fuzzy Syst. 22 (5), 1430–1443.

Atanassov, K.T., 1986. Intuitionistic fuzzy sets. Fuzzy Sets Syst. 20 (1), 87–96.

Behbood, V., Lu, J., Zhang, G., 2013a. Text categorization by fuzzy domain adaptation. In: Fuzzy System (FUZZ), 2013 IEEE Int. Conf on. IEEE.

Behbood, V., Lu, J., Zhang, G., 2013b. Fuzzy bridged refinement domain adaptation: long-term bank failure prediction. Int. J. Comput. Intell. Appl. 12 (01), 1350003.

Behbood, V., Lu, J., Zhang, G., Pedrycz, W., 2015. Multi-Step Fuzzy Bridged Refinement Domain Adaptation Algorithm and Its Application to Bank Failure Prediction.

Bi, J., Bennett, K.P., 2003. A geometric approach to support vector regression. Neurocomputing 55 (1), 79–108.

Bokde, N.D., Tranberg, B., Andresen, G.B., 2021. Short-term CO2 emissions forecasting based on decomposition approaches and its impact on electricity market scheduling. Appl. Energy 281, 116061.

Castillo, O., Melin, P., 2003. A new method for fuzzy inference in intuitionistic fuzzy systems. In: Fuzzy Information Processing Society, 2003. NAFIPS 2003. 22nd Int. Conf of the North American. IEEE.

Castillo, O., Alanis, A., Garcia, M., Arias, H., 2007. An intuitionistic fuzzy system for time series analysis in plant monitoring and diagnosis. Appl. Soft Comput. 7 (4), 1227–1233.

Chaabouni, S., Saidi, K., 2017. The dynamic links between carbon dioxide (CO2) emissions, health spending and GDP growth: a case study for 51 countries. Environ. Res. 158, 137–144.

Chaira, T., 2011. A novel intuitionistic fuzzy C means clustering algorithm and its application to medical images. Appl. Soft Comput. 11 (2), 1711–1717.

Che, X., Zuo, H., Lu, J., Chen, D., 2021. Fuzzy multi-output transfer learning for regression. IEEE Trans. Fuzzy Syst. 30 (7), 2438–2451.

Chen, J., Wang, P., Cui, L., Huang, S., Song, M., 2018a. Decomposition and decoupling analysis of CO2 emissions in OECD. Appl. Energy 231, 937–950.

Chen, Y., Song, S., Li, S., Yang, L., Wu, C., 2018b. Domain space transfer extreme learning machine for domain adaptation. IEEE Trans. Cybernetics 49 (5), 1909–1922.

Debnath, J., Majumder, D., Biswas, A., 2018. Air quality assessment using weighted interval type-2 fuzzy inference system. Ecol. Inform. 46, 133–146.

Deng, Z., Choi, Kup-Sze, Jiang, Yizhang, Wang, Shitong, 2014. Generalized hidden-mapping ridge regression, knowledge-leveraged inductive transfer learning for neural networks, fuzzy systems and kernel methods. IEEE Trans. Cybernetics 44 (12), 2585–2599.

Depren, S.K., Kartal, M.T., Çelikdemir, N.Ç., Depren, Ö., 2022. Energy consumption and environmental degradation nexus: a systematic review and meta-analysis of fossil fuel and renewable energy consumption. Ecol. Inform. 70, 101747.

Dufourq, E., Batist, C., Foquet, R., Durbach, I., 2022. Passive acoustic monitoring of animal populations with transfer learning. Ecol. Inform. 101688.

Fan, Y., Xiao, F., 2020. TDIFS: two dimensional intuitionistic fuzzy sets. Eng. Appl. Artif. Intell. 95, 103882.

Ghosh, S.K., Mitra, A., Ghosh, A., 2021. A novel intuitionistic fuzzy soft set entrenched mammogram segmentation under multigranulation approximation for breast cancer detection in early stages. Expert Syst. Appl. 169, 114329.

Govindaraju, V.G.R.C., Tang, C.F., 2013a. The dynamic links between CO_2 emissions, economic growth and coal consumption in China and India. Appl. Energy 104, 310–318.

Govindaraju, V.G.R.C., Tang, C.F., 2013b. The dynamic links between CO2 emissions, economic growth and coal consumption in China and India. Appl. Energy 104, 310–318.

Guo, S., Shao, L., Chen, H., Li, Z., Liu, J.B., Xu, F.X., Li, J.S., et al., 2012. Inventory and input–output analysis of CO2 emissions by fossil fuel consumption in Beijing 2007. Ecol. Inform. 12, 93–100.

Gyamfi, B.A., Adedoyin, F.F., Bein, M.A., Bekun, F.V., Agozie, D.Q., 2021. The anthropogenic consequences of energy consumption in E7 economies: juxtaposing roles of renewable, coal, nuclear, oil and gas energy: evidence from panel quantile method. J. Clean. Prod. 295, 126373.

http://data.worldbank.org/.

http://www.ntu.edu.sg/home/egbhuang/elm_kernel.html.

Huang, G.B., Zhu, Q.Y., Siew, C.K., 2006. Extreme learning machine: theory and applications. Neurocomputing 70 (1), 489–501.

Hung, W.L., Yang, M.S., 2004. Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance. Pattern Recogn. Lett. 25 (14), 1603–1611.

Jha, V., et al., 2022. An improved monarch butterfly optimization based multivariate fuzzy time series approach for forecasting GDP of India. Evol. Intel. 1–15.

Jiang, Y., Chung, F.-L., Ishibuchi, H., Deng, Z., Wang, S., 2015. Multitask TSK fuzzy system modeling by mining intertask common hidden structure. IEEE Trans. Cybernetics 45 (3), 534–547.

Khatua, A., Jana, S., Kar, T.K., 2020. A fuzzy rule-based model to assess the effects of global warming, pollution and harvesting on the production of Hilsa fishes. Ecol. Inform. 57, 101070.

Kumar, S., Shukla, A.K., Muhuri, P.K., Lohani, Q.D., 2016, July. Atanassov intuitionistic fuzzy domain adaptation to contain negative transfer learning. In: Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference on. IEEE, pp. 2295–2301.

Kutlu, F., Atan, Ö., Silahtar, O., 2020. Intuitionistic fuzzy adaptive sliding mode control of nonlinear systems. Soft. Comput. 24 (1), 53–64.

Liu, Y., Li, Xia, 2014. Domain adaptation for land use classification: a spatio-temporal knowledge reusing method. ISPRS J. Photogramm. Remote Sens. 98, 133–144.

Lu, J., Behbood, V., Hao, P., Zuo, H., Xue, S., Zhang, G., 2015. Transfer learning using computational intelligence: A survey. Knowl.-Based Syst. 80, 14–23.

Lu, J., Zuo, H., Zhang, G., 2019. Fuzzy multiple-source transfer learning. IEEE Trans. Fuzzy Syst. 28 (12), 3418–3431.

Lu, T., Han, B., Yu, F., 2021. Detection and classification of marine mammal sounds using AlexNet with transfer learning. Ecol. Inform. 62, 101277.

Lumini, A., Nanni, L., 2019. Deep learning and transfer learning features for plankton classification. Ecol. Inform. 51, 33–43.

Margolis, A., 2011. A Literature Review of Domain Adaptation with Unlabeled Data. University of Washington, Rapport Technique.

Mirbagheri, M., 2010. Fuzzy-logic and neural network fuzzy forecasting of Iran GDP growth. Afr. J. Bus. Manag. 4 (6), 925.

Ohta, R., Salomon, V., Silva, M.B., 2020. Classical, fuzzy, hesitant fuzzy and intuitionistic fuzzy analytic hierarchy processes applied to industrial maintenance management. J. Intell. Fuzzy Syst. 38 (1), 601–608.

Pan, J., Yang, Q., 2010. A survey on transfer learning. In: Knowledge and Data Engineering, IEEE Transactions on 22.10, pp. 1345–1359.

Pan, J., Hu, X., Li, P., Li, H., He, W., Zhang, Y., Lin, Y., 2016. Domain adaptation via multi-layer transfer learning. Neurocomputing 190, 10–24.

Pearson, D.E., Clark-Wolf, T.J., 2023. Predicting ecological outcomes using fuzzy interaction webs. Ecology 104 (7), e4072.

Pedrycz, W., Gudwin, R.R., Gomide, F.A.C., 1997. Nonlinear context adaptation in the calibration of fuzzy sets. FSS 88 (1), 91–97.

Qian, P., Zhao, K., Jiang, Y., Su, K.-H., Deng, Z., Wang, S., Muzic Jr., R.F., 2017. Knowledge-leveraged transfer fuzzy c-means for texture image segmentation with self-adaptive cluster prototype matching. Knowl.-Based Syst. 130, 33–50.

Rosenstein, M.T., Marx, Z., Kaelbling, L.P., Dietterich, T.G., 2005. To transfer or not to transfer. In: *NIPS* 2005 Workshop on Inductive Transfer: 10 Years Later, Vol. 2.

Saidi, K., Hammami, S., 2015. The impact of energy consumption and CO 2 emissions on economic growth: fresh evidence from dynamic simultaneous-equations models. Sustain. Cities Soc. 14, 178–186.

Salaken, S.M., Khosravi, A., Nguyen, T., Nahavandi, S., 2019. Seeded transfer learning for regression problems with deep learning. Expert Syst. Appl. 115, 565–577.

Seera, M., Lim, C.P., 2014. Transfer learning using the online Fuzzy Min–Max neural network. Neural Comput. & Applic. 25 (2), 469–480.

Shell, J., Coupland, S., 2015. Fuzzy transfer learning: methodology and application. Inf. Sci. 293, 59–79.

Shukla, A.K., Muhuri, P.K., 2019. Big-data clustering with interval type-2 fuzzy uncertainty modeling in gene expression datasets. Eng. Appl. Artif. Intell. 77, 268–282.

Shukla, A.K., Prakash, V., Nath, R., Muhuri, P.K., 2022. Type-2 intuitionistic fuzzy TODIM for intelligent decision-making under uncertainty and hesitancy. Soft. Comput. 1–18.

Songwei, L., Xiaoguang, Z., Renhou, Z., 2012. Takagi-Sugeno Intuitonsitic fuzzy inference systems based on game theory. In: (CSSS), 2012 International Conference on. IEEE.

Sousa, M.J., Moutinho, A., Almeida, M., 2020. Wildfire detection using transfer learning on augmented datasets. Expert Syst. Appl. 142, 112975.

Specht, D.F., 1991. A general regression neural network. IEEE Trans. Neural Netw. 2 (6), 568–576.

Stern, D.I., 2010a. Between estimates of the emissions-income elasticity. Ecol. Econ. 69 (11), 2173–2182.

Stern, D.I., 2010b. Between estimates of the emissions-income elasticity. Ecol. Econ. 69 (11), 2173–2182.

Szmidt, E., Kacprzyk, J., 2001. Intuitionistic Fuzzy Sets in some Medical Applications. Berlin Heidelberg, Computational Intelligence. Theory and Applications. Springer, pp. 148–151.

Szmidt, E., Kacprzyk, J., 2004. Medical diagnostic reasoning using a similarity measure for intuitionistic fuzzy sets. Note IFS 10 (4), 61–69.

Tan, T., Zhao, Tao, 2023. A data-driven fuzzy system for the automatic determination of fuzzy set type based on fuzziness. Inf. Sci. 642, 119173.

Van, E., Jesper, E., Hoos, H.H., 2020. A survey on semi-supervised learning. Mach. Learn. 109 (2), 373–440.

Vapnik, V., 2013. The Nature of Statistical Learning Theory. Springer Science & Business Media.

Wang, G., Lu, J., Choi, K.-S., Zhang, G., 2018. A transfer-based additive LS-SVM classifier for handling missing data. IEEE Trans. Cybernetics 50 (2), 739–752.

Wang, Y., et al., 2022. A lightweight CNN-based model for early warning in sow oestrus sound monitoring. Ecol. Inform. 72, 101863.
Wu, D., Lawhern, Vernon, J., Gordon, Stephen, Lance, Brent, J., Lin, Chin-Teng, 2016. Driver drowsiness estimation from EEG signals using online weighted adaptation regularization for regression (OwARR). IEEE Trans. Fuzzy Syst. 25 (6), 1522–1535.
Xie, L., Deng, Z., Xu, P., Choi, K.-S., Wang, S., 2018. Generalized hidden-mapping transductive transfer learning for recognition of epileptic electroencephalogram signals. IEEE Trans. Cybernetics 49 (6), 2200–2214.
Yao, C., Feng, K., Hubacek, K., 2015. Driving forces of CO2 emissions in the G20 countries: an index decomposition analysis from 1971 to 2010. Ecol. Inform. 26 (1), 93–100.
Zadeh, L.A., 1965. Fuzzy sets. Inf. Control. 8 (3), 338–353.
Zuo, H., Zhang, G., Pedrycz, W., Behbood, V., Lu, J., 2016. Fuzzy regression transfer learning in takagi-sugeno fuzzy models. IEEE Trans. Fuzzy Syst. 25 (6), 1795–1807.