# A Temporal Type-2 Fuzzy System for Time-dependent Explainable Artificial Intelligence

Mehrin Kiani, Javier Andreu-Perez*, *Senior Member, IEEE*, and Hani Hagras, *Fellow, IEEE*

***Abstract*—Explainable Artificial Intelligence (XAI) is a paradigm that delivers transparent models and decisions, which are easy to understand, analyze, and augment by a non-technical audience. Fuzzy Logic Systems (FLS) based XAI can provide an explainable framework, while also modeling uncertainties present in real-world environments, which renders it suitable for applications where explainability is a requirement. However, most real-life processes are not characterized by high levels of uncertainties alone; they are inherently time-dependent as well, i.e., the processes change with time. To account for the temporal component associated with a process, in this work, we present novel *Temporal Type-2 FLS Based Approach* for time-dependent XAI (TXAI) systems, which can account for the likelihood of a measurement's occurrence in the time domain using (the measurement's) frequency of occurrence. In Temporal Type-2 Fuzzy Sets (TT2FSs), a four-dimensional (4D) time-dependent membership function is developed where relations are used to construct the inter-relations between the elements of the universe of discourse and its frequency of occurrence. The proposed TXAI system with TT2FSs is exemplified with a step-by-step numerical example and an empirical study using a real-life intelligent environments dataset to solve a time-dependent classification problem (predict whether or not a room is occupied depending on the sensors readings at a particular time of day). The TXAI system performance is also compared with other state-of-the-art classification methods with varying levels of explainability. The TXAI system manifested better classification prowess, with 10-fold test datasets, with a mean recall of 95.40% than a standard XAI system (based on non-temporal general type-2 (GT2) fuzzy sets) that had a mean recall of 87.04%. TXAI also performed significantly better than most non-explainable AI systems between 3.95%, to 19.04% improvement gain in mean recall. Temporal convolution network (TCN) was marginally better than TXAI (by 1.98% mean recall improvement) although with a major computational complexity. In addition, TXAI can also outline the most likely time-dependent trajectories using the frequency of occurrence values embedded in the TXAI model; viz. given a rule at a determined time interval, what will be the next most likely rule at a subsequent time interval. In this regard, the proposed TXAI system can have profound implications for delineating the evolution of real-life time-dependent processes, such as behavioural or biological processes.**

## I. Introduction

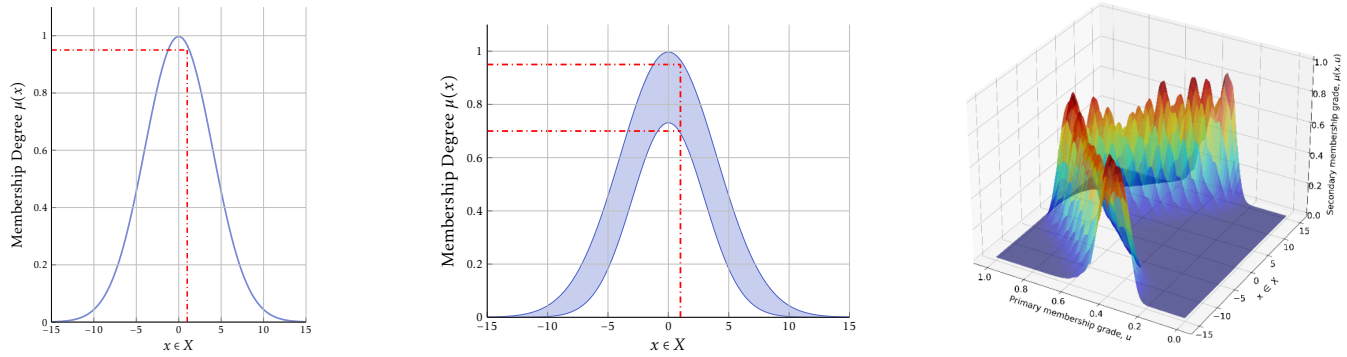
Over the last few decades, the widespread application of artificial intelligence (AI) systems have enhanced many aspects of everyday life from risk management [1], sky shepherding of sheep [2], medical image segmentation [3], recognition of expertise level [4], mobile applications [5] to Covid-19 detection based on cough samples [6]. Although opaque AI systems offer remarkable prediction accuracy, they are limited by a lack of explanation behind their predictions. A lack of explanation renders the AI systems untrustworthy, and particularly inapplicable where users want to understand the decision process of the AI system. To this end, there is a growing need for transparent, human-understandable AI systems called explainable AI (XAI) systems [7]. Several approaches taken towards the development of XAI systems include: 1) Intrinsic: a method in which model inference structure is fully transparent such as short decision trees or sparse linear models, and 2) Post-hoc: a model-agnostic meta-model is used to decipher the inference rationale of a black-box model permutation feature importance can be computed for decision trees. Within post-hoc methods attempts to unravel a black-box model into a surrogate intrinsic model have also been undertaken. A particular category of these are the anchor-based models.

Although anchor-based approach provides a step towards implementing human-understandable explanations [8], explanatory patterns rest on hard thresholds and are constrained by Boolean logic. However, real-life processes are characterised with uncertainty and therefore hard thresholds based models are not particular well-suited to model them (real-life processes). In this regard, another approach to implement XAI systems is fuzzy logic systems (FLS) [7, 9]. The FLS based XAI systems are well-suited for explainable modelling of real-life processes because of FLS capability to handle uncertainty in the input data, and subsequently improve the process model and performance. In addition, the use of conceptual labels (CoLs) that model uncertainty and axioms of FLS based XAI systems pave way for human-understandable models for describing complex, real-life processes.

The FLS based XAI systems handle uncertainty in the input data using fuzzy sets that convert crisp numbers (viz. uncertain observations) to CoLs characterised with membership values [9, 10]. The fuzzy sets are defined by membership functions (MFs) and represent a given CoL. The membership value is usually in the range [0,1]

* Corresponding author: javier.andreu@essex.ac.uk

M. Kiani, J. Andreu-Perez and H. Hagras are with the School of Computer and Electronic Engineering, University of Essex, Colchester, CO4 3SQ, United Kingdom.




(a) Type-1 (T1) Fuzzy Set (b) Interval Type-2 (IT2) Fuzzy Set (c) General Type-2 (GT2) Fuzzy Set

Fig. 1: The three types of fuzzy sets: (a) Type-1 (T1) fuzzy sets where each crisp measurement, $x \in X$, gets assigned a membership degree, $\mu_{T1}(x) \subseteq [0,1]$, but there is no ambiguity in the membership degree, for example as shown by the red dashed line: $\mu_{T1}(x=1) = 0.95$. (b) Interval type-2 (IT2) fuzzy sets have lower and upper membership degrees assigned to each crisp measurement for example $\mu_{IT2}(x=1) = [0.7, 0.95]$. (c) General type-2 (GT2) fuzzy sets have T1 fuzzy sets as membership degree for a crisp measurement i.e. $\mu_{T2}(x=1) = \{u, \mu_{T1}(u) | \forall u \in [0,1], \forall \mu_{T1} \in [0,1]\}$ where $u$ is called the primary membership degree and $\mu$ is called the secondary membership degree.

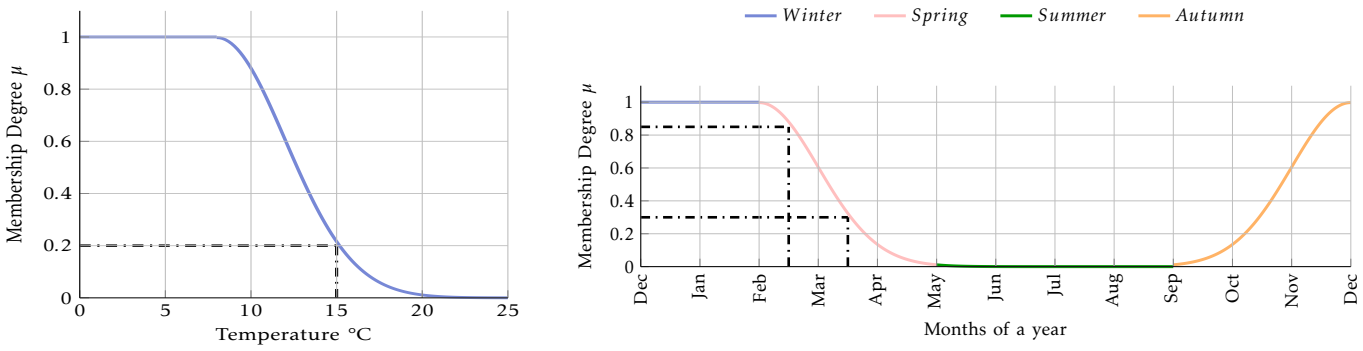
and is a soft measure of the degree of association the associated fuzzy set has for a given crisp measurement to belong to the CoL represented by the fuzzy set [10]. For example, an XAI system modelling the heights of people in a community using type-1 fuzzy sets may represent height using CoLs of *Tall*, *Medium*, and *Short*. The MF associated with each CoL's MF will assign a crisp number for the height of a person with a membership grade; for example, a height of 6ft may get assigned membership grades of $0.8, 0.5, 0.1$ to represent CoLs of *Tall*, *Medium*, and *Short* respectively.

In general, fuzzy sets can model uncertainty in the feature domain at different levels: Type 1 (T1), interval type-2 (IT2), and general type-2 (GT2) fuzzy sets; illustrated in Fig. 1. Despite the variability in the extent for uncertainty modelling amongst the types of fuzzy sets, all fuzzy sets are modelling uncertainty from a single time snapshot of the feature domain. More specifically, fuzzy sets do not integrate associated temporal information in their membership grade calculation. This is a critical limitation of the fuzzy sets since most real-life systems are time-variant, i.e., their behaviour changes with time. To model time-dependent real-life systems more effectively, in this work, **we present the theory of a new *Temporal Type-2 Fuzzy Set (TT2FS)* based approach for time-dependent XAI (TXAI)**.

The prowess of TXAI system for incorporating time information for modelling time-variant processes is of paramount significance since the insights provided by a TXAI system can shed light on both spatial (feature domain) and temporal behaviour of the time-dependent process. More specifically, the TXAI is able to inform not only about the relation between input features but can also describe the impact of time on the evolution of the inter-relation of the features. As an example, let's consider a standard XAI system composed of a T1 fuzzy set for modelling thermal sensation 'Cold' in the domain of values of temperature T °C as shown in Fig. 2 (a), and a T1 fuzzy set for the time of occurrence of concept 'Cold' during the months of a year as shown in Fig. 2 (b). The notion is that the perception of 'cold' is mostly associated with the months of winter than in the months of spring. Hence, using the time information associated with a fuzzy concept (such as Cold in this case), a temperature can belong to the concept (Cold) differently according to a particular point in time (e.g., months of a year).

Crediting a fuzzy membership with its associated time information is particularly advantageous for the modelling of time-dependent noise-prone processes. Moreover, for dynamic processes, the ability to delineate its' (dynamic process) trajectories across time would inform the evolution of the temporal dynamics of the process. To this end, our proposed TXAI system has been designed to integrate temporal information as well as able to outline the trajectories of a time-dependent process. To demonstrate the efficacy of TXAI system for time-dependent process modelling, in this work, an occupancy dataset is used [11]. Using the values of temperature, light and carbon dioxide ($CO_2$), and the time the aforementioned measurements are taken, the TXAI system is used to make a prediction of whether or not the room is occupied.

The rest of the paper is organised as follows: in Section II related works are outlined, Section III presents the TXAI system definition and operations, Section IV outlines the TXAI inference system with a numerical step-by-step example as well as the evolution of a TXAI model using temporal trajectories. An empirical study using TXAI system, as well as state-of-the-art systems (with varying levels of explainability) for performance comparison, on the aforementioned occupancy dataset [11] is presented in Section V, with conclusion and future research in Section VI.




(a) 'Cold' membership function in temperature domain.

(b) 'Cold' membership function in time domain.

Fig. 2: An illustrative type-1 (T1) membership function (MF) for the fuzzy concept of 'Cold' in the (a) universe of temperature in °C and (b) in the universe of time: months of a year. In this case, the membership degree for experiencing 'Cold' at 15 °C is $\mu_{Cold_{temp}}(15°C) = 0.2$. Likewise considering the prevalence of particular linguistic variable 'Cold', viz. the *likelihood of observing* 'Cold' can be different in February $\mu_{Cold_{time}}(February) = 0.85$ than March $\mu_{Cold_{time}}(March) = 0.3$. In this regard, the additional information of time can credit the primary membership in feature-domain through a fuzzy relation.

## II. Related Works

Fuzzy sets have enabled explainable models of complex real-life processes which prove too ill-defined for closed form mathematical analysis. In this regard, although uncertainty in complex processes could be handled by fuzzy sets, the time-variant characteristics of complex processes have not been integrated into the modelling by standard XAI systems based on state-of-the-art fuzzy sets.

There have been few notable attempts in the literature to model time in the MFs. The work by Garibaldi *et al.* [12] on *non-stationary* fuzzy sets proposed that variation within a MF can be incorporated by perturbing the parameters of the MF. Their work aims to develop non-deterministic fuzzy reason as a way to model the variability in fuzzy decision making to mimic the variability in expert opinions. The ability of *non-stationary* fuzzy sets to integrate differing experts' opinions is a significant contribution since it allows for a more comprehensive model that takes into account all experts' opinions. However, their work does not incorporate the variation within a fuzzy concept with respect to time, which is the aim of the present work, to represent the time-variant transformation of a same fuzzy linguistic variable.

Similarly, the work by Kostikova *et al.* [13] propose *dynamic* fuzzy sets by extending the classical fuzzy set to include a time dimension for representing MF at different time points. They propose four different types of dynamic MFs depending on how many parameters are changed in the definition of the dynamic MF. They simulated their dynamic MFs by using differing expert assessments on multilevel fuzzy description of a complex system. However, the dynamic MF is essentially a set of functions determined at different time points with no bearing on the temporal variation in the fuzzy concept.

In another work by Maeda *et al.* [14], they propose *dynamic fuzzy reason* to deal with the notion of *time delay* between premise and consequent. An example of where a time delay between premise and consequent assumes critical importance is: 'If it starts snowing, the traffic on road will increase about 30 minutes later'. They propose the use of fuzzy relations between a fuzzy concept and its fuzzy time interval to assign a credit degree to the concept. The temporal fuzzy reasoning provides a framework for modelling delay in fuzzy reasoning and the temporal dynamics of a fuzzy concept. In this work, we have built on the work of Maeda *et al.* [14] to credit the membership grade of a concept based on time.

To the best of the authors' knowledge, there is no work in the literature on fuzzy sets that delineates the incorporation of time-based variation in a fuzzy concept to compute the membership grade for the crisp values of the fuzzy concept. In addition, no previous work has aimed at delineating the trajectories of a time-variant process with respect to time. To this end, in this work, we propose TXAI systems that can integrate information from both the feature domain and time domain. More details on the proposed TXAI are outlined in Section III.

## III. Time-dependent Explainable Artificial Intelligence (TXAI) Systems

In this section, we present the TXAI system based on TT2FS (temporal type-2 fuzzy sets) that incorporate information from not only the uncertainty in the input domain of the fuzzy linguistic term, but also from its time of occurrence. In particular, the information from the time of occurrence is integrated into the membership grade of the TT2FS using fuzzy relations such that it (the membership grade of the TT2FS) varies with respect to time (time-dependent).

In the next section, we present the most common fuzzy relations and outline how they can be used for implementing TT2FS.

TABLE I: Fuzzy relations between the universe of concept X and time domain T.

| Name | Definition of the relation |
|---|---|
| Godel | $R_G(t,x) = \begin{cases} 1 & \text{if } \mu_{T_A}(t) \le \mu_A(x) \\ \mu_A(x) & \text{if } \mu_{T_A}(t) > \mu_A(x) \end{cases}$ |
| Lukasiewicz | $R_L(t,x) = 1 \wedge (1 - \mu_{T_A}(t) + \mu_A(x))$ |
| Gaines-Rescher | $R_{GR}(t,x) = \begin{cases} 1 & \text{if } \mu_{T_A}(t) \le \mu_A(x) \\ 0 & \text{if } \mu_{T_A}(t) > \mu_A(x) \end{cases}$ |
| Mamdani | $R_M(t,x) = \mu_{T_A}(t) \wedge \mu_A(x)$ |

### *A. Fuzzy relations between fuzzy linguistic variables and time related measures*

In this work, fuzzy relations are used to interrelate the information with respect to the degree of truth of a determined linguistic term or CoL, $A$, within the domain X, and time, T, to form TT2FSs such that the likelihood of occurrence of A in $x \in X$, i.e. the primary membership grade $\mu_A(x)$, is credited by a measure that is dependent on time such as frequency. The application of fuzzy relation, for constructing TT2FSs, is motivated by the work on *dynamic fuzzy reasoning models* in [14]. They outline fuzzy relations that can be used to model time dependencies, as noted in Table I.

Before reviewing the different relations that can be applied to construct a TT2FS, the conditions that need to be fulfilled by the associated temporal MF (TMF) are listed below:

(i) The TMF should be continuous.
(ii) The TMF should be convex.
(iii) The range of the TMF $\subseteq [0,1]$.
(iv) The TMF should reflect in the value of membership grade the intrinsic magnitudes of membership grade in feature domain and in frequency of occurrence domain, i.e., they should be directly proportional. For example, if $\mu_A(x)$ is high and the time representation is also high then the result from the relation between them should also be high and vice versa.

An illustrative comparison of the TT2FSs formed for the CoL 'Cold' of feature thermal concept using the fuzzy relations listed in Table I is shown in Fig. 3. The fuzzy relations are applied on hypothetical primary membership function of 'Cold' in feature domain (temperature) and time domain (months of a year). As can be seen in Fig. 3, the different fuzzy relations are encapsulating distinct inter-dependencies between time and feature domain. All relations meet the criteria (i) - (iii) listed above however, only the Mamdani relation meets the criterion (iv) as well since it gives credit to $\mu_{Cold}$ based on the variable frequency of occurrence of 'Cold' as observed in different months of the year. Hence, in this work, the Mamdani relation is used to construct the TT2FSs.

### *B. Conditional relative frequency distribution of a fuzzy linguistic term*

In our TT2FS we employ a measure of conditional relative frequency between time and the occurrence of a linguistic term. We denote as $A$ an instance of a linguistic term from a set of conceptual labels (also called words of the universe of discourse), $CoLs := [CoL_1, CoL_2, ..., CoL_J]$ of a specific linguistic variable or input.

**Definition III.1** (Discrete conditional relative frequency with respect to time)**.** *The discretized conditional relative frequency is defined as the likelihood of observing a linguistic term A based on its membership grade, across time. This is denoted as $g_A(t_n, \mu_A(x))$ with time t discretised over N time points ($t_n$) such as $t_n \in [t_1, ..., t_N]$, and is given by:*

$$g_A(t_n, \mu_A(x)) = \frac{\sum\limits_{x \in X, t_n} \delta_{nj}}{\max\limits_{[t_1,...,t_N]} \left( \sum\limits_{x \in X, t_n} \delta_{nj} \right)} \tag{1}$$

*$\delta_{nj}$ is a Kronecker delta function [15] (e.g. $\delta_{ab} = 0$ if $a \neq b$, $\delta_{ab} = 1$ if a=b) that takes the value of 1 when the following condition applies, $\exists\ argmax_j(\mu_{CoL_j}(x^{t_n})) : Col_j = A,\ \forall j \in [1,...,J]$, and 0 otherwise. Note $x^{t_n}$ is a realisation of x at time $t_n$.*

The numerator in (1) finds the count of occurrences of a given $A$ for a determined time point $t_n$ across all data instances, whereas the denominator is finding the maximum value of the count of occurrences of $A$ across all $N$ time points and all data instances. The resultant discrete conditional relative frequency $g_A(t_n, \mu_A(x))$ is interpolated to form a conditional distribution $f_A(t, \mu_A(x))$. For the sake of notational simplicity, we denote the later distribution as $f_A$ and the discrete conditional relative frequency as $g_A$ from here onwards.

Let us assume that the linguistic variable is thermal sensation defined on the input domain ($x \in X$) of temperature in °C and the associated CoLs be: [Cold, Comfortable, Hot]. For a given crisp input of temperature such as 15°C, the associated primary membership grade for all three CoLs of Cold, Comfortable, and Hot be $\mu_{Cold}(15\,°C) = [0.4]$, $\mu_{comf.}(15°C) = [0.3]$, $\mu_{hot}(15°C) = [0]$ respectively. In this illustrative case, the temperature of 15 °C has a maximum membership grade, amongst all CoLs, for *Cold* and hence 15 °C is assigned with the CoL of *Cold*. Referring back to (1), for computing the conditional relative frequency for *Cold* the numerator is going to sum all the data instances where the crisp inputs are assigned with *Cold* for a given time point $t_n$ such as a particular month of a year. The denominator finds the mode of occurrence of *Cold* across all months. The result of the division will scale the $g_{Cold}$ values to [0,1].

An illustration for calculating the $g_{Cold}$ values using (1), with a total of 12 time points as the months of a year is shown in Fig. 4 (b) with continuous values of $f_{Cold}$, found using interpolation of $g_{Cold}$, plotted in Fig. 4 (c). Please note the associated time intervals, (as listed in the illustration in Fig. 4 are seasons in a year such as Winter, Spring, Summer, and Autumn), are for easing the computational complexity of the four-dimensional (4D) TT2FSs as will be explained later in section III-D by taking time interval based slice of the TT2FS.

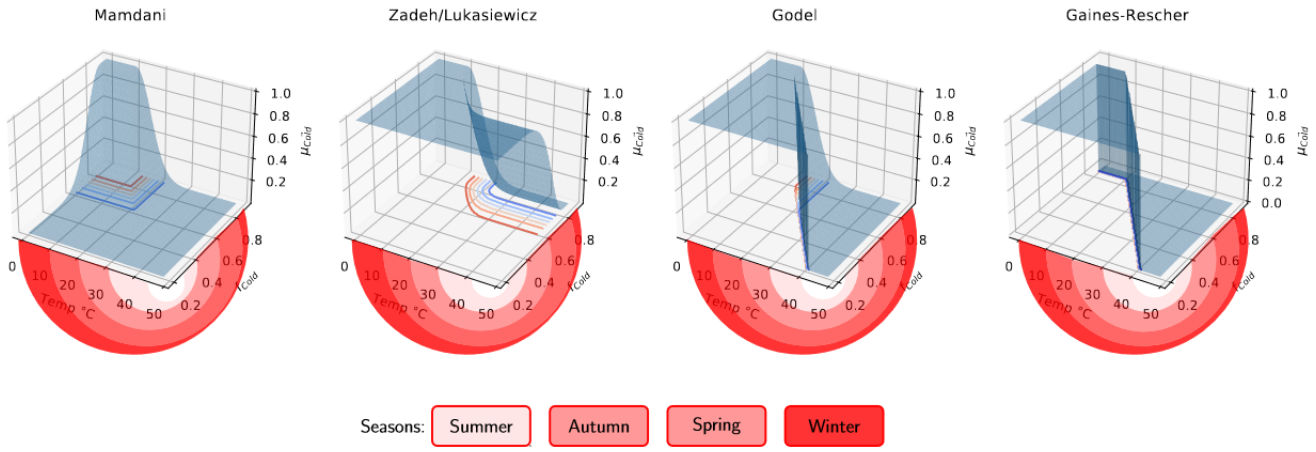



Fig. 3: A comparison of TT2FSs for the conceptual label (CoL) 'Cold' for feature thermal concept constructed with the most commonly used fuzzy relations namely Mamdani, Zadeh/Lukasiewicz, Godel, and Gaines-Rescher, see Table I for their respective definitions. In these illustrative plots, the feature domain i.e. temperature in °C is plotted on the x-axis, with conditional distribution, $f_{Cold}$ on y-axis, and the time is plotted on the axis connecting the x- and y- axis, i.e. the arc axis, with the 4 time intervals representing the typical seasons in a year. The z-axis has the values of temporal membership function (TMF), $\mu_{\vec{Cold}}(x,t,f_{Cold})$.

### C. Temporal Type-2 Fuzzy Sets (TT2FS)

In this section, a formal definition of temporal type-2 fuzzy sets (TT2FS) is presented. TT2FS are 4D as they incorporate information from the input domain ($X$), time domain ($T$), frequency of occurrence domain ($F$) and are characterised by a temporal membership function (TMF).

The computation of TMF, hereby termed as *temporal fuzzification*, involves two stages: 1) fuzzification of crisp input values of $A$ from feature domain $X$ to form T1 $\mu_A(x)$, as undertaken in standard T1 fuzzy sets; and 2) computation of the conditional distribution of $A$, $f_A$. The temporal fuzzification is illustrated in Fig. 4 (a) and defined next.

**Definition III.2** (Temporal membership function)**.** *The temporal membership function (TMF) can be defined as*

$$\mu_{\vec{A}}(x,t,f_A) = \mu_A(x) \otimes f_A \tag{2}$$

*where $\otimes$ is a relation operator, $\mu_A(x)$ is the primary membership of $A$ in feature domain credited by the conditional distribution of $A$, denoted $f_A$, using the Mamdani relation (outlined earlier in Sec III-A).*

**Theorem III.1.** *The TMF of A, constructed using Mamdani relation (2), $\mu_{\vec{A}}(x,t,f_A)$ is $\subseteq [0,1]$.*

*Proof.* The range of $\mu_{\vec{A}}(x,t,f_A)$ follows directly from the range of primary MF of A: $\mu_A(x) \subseteq [0,1]$, and the conditional distribution of A: $f_A \subseteq [0,1]$. Hence, by crediting $\mu_A(x)$ with $f_A$ using Mamdani relation (taking the min or product), it follows that the range of $\mu_{\vec{A}}(x,t,f_A) \subseteq [0,1]$. ■

**Proposition III.1.1.** *If the primary membership of TMF is normal and the conditional distribution $f$ is normal, according to (1), then the resultant TMF membership after applying the Mamdani relation yields a normal temporal membership function, therefore we can imply that*

$$\sup_{x \in X} \mu_{\vec{A}}(x,t,f_A) = 1 \tag{3}$$

*Proof.* Given a $f_A \subseteq [0,1]$ and a $\mu_A(x) \subseteq [0,1]$ both with $sup = 1$, $\forall x \in X$ by deduction, $\exists x : f_A \times \mu_A(x) \vee min(f_A, \mu_A(x)) = 1$ ■

Next, we define the TT2FS which are characterised by a TMF.

**Definition III.3** (Temporal Type-2 Fuzzy Sets (TT2FS))**.** *A TT2FS $\vec{A}$ of the universe of discourse $X \times T \times F$ is characterised by a credited TMF $\mu_{\vec{A}}(x,t,f_A) : X \times T \times F \to [0,1]$ where $X$ is the feature domain of $A$ characterised by a T1 MF $\mu_A(x)$, $T$ is the time domain of $A$, $F$ is the frequency of occurrence domain of $A$ characterised by conditional frequency distribution with respect to time $f_A$. In mathematical set notation, $\vec{A}$ can be written as (4):*

$$\begin{aligned} \vec{A} = \{(x,t,f_A,\mu_{\vec{A}}(x,t,f_A)) \mid \\ \forall x \in X, \forall t \in T, \forall \mu_A(x) \subseteq [0,1], \\ \forall f_A \in F \subseteq [0,1]\} \end{aligned} \tag{4}$$

*where $\mu_{\vec{A}}(x,t,f_A) \subseteq [0,1]$. Please note the conditional distribution, $f_A$, is a continuous distribution interpolated from discrete conditional relative frequency, $g_A$, and is defined mathematically earlier in (1). $\vec{A}$ can also be expressed as:*

$$\vec{A} = \int_{x \in X} \int_{t \in T} \int_{f_A \in F} \mu_{\vec{A}}(x,t,f_A)/f_A/t/x \tag{5}$$

where $\int\int\int$ denotes the aggregation over all admissible values of $x$, $t$, and $f_A$. The associated TMF,

$\mu_{\vec{A}}(x,t,f_A)) \subseteq [0,1]$, scales the $\mu_A(x)$ based on its conditional distribution $f_A$ as defined in (2).

### D. Operations on TT2FSs

In this section, the common operations for TT2FSs such as the union and intersection, as well as defuzzification are outlined. TT2FSs, on account of being 4D, are more computationally intense than GT2 fuzzy sets, which are three-dimensional (3D). A popular approach for minimising the computational demand of 3D GT2 fuzzy sets is to use z-slice based framework [16]. Motivated from the effectiveness of z-slice based framework for simplifying the computations for GT2 fuzzy sets, in this work, the approach of taking time interval slice followed by z-slice (TS-ZS) is taken for performing operations on TT2FSs. The TS-ZS approach is explained in more detail as follows:

(i) TS: Time interval based slice to convert 4D TT2FSs into 3D. The 3D time interval based TT2FS is similar to 3D GT2 fuzzy set, with both sharing the feature domain on $x$–axis. On $y$–axis is the frequency of occurrence domain, for that time interval, for time interval based TT2FS, while for GT2 fuzzy sets, primary membership grade is on $y$–axis. And on $z$–axis is the temporal membership grade for time interval based TT2FS while for GT2 fuzzy set secondary membership grade is on $z$–axis.

(ii) ZS: z-Slice based approach for the time interval based 3D TT2FS as utilised for GT2 fuzzy sets. The z-slices at specific z-levels render a given 3D fuzzy set to an equivalent IT2 fuzzy set with lower and upper primary membership grades. For the case of TS-ZS based TT2FSs, the primary membership grades are the conditional distribution values for that time interval at a given z-level.

In the following sections, a formal definition for the operations on TT2FSs is given with $\vec{A}$ and $\vec{B}$ denoting two TT2FSs characterised by TMFs $\mu_{\vec{A}}(x,t,f_A)$ and $\mu_{\vec{B}}(x,t,f_B)$ respectively as outlined in (6):

$$\vec{A} = \int_{x\in X}\int_{t\in T}\int_{f\in F} \mu_{\vec{A}}(x,t,f_A)/f_A/t/x$$
$$\vec{B} = \int_{x\in X}\int_{t\in T}\int_{f\in F} \mu_{\vec{B}}(x,t,f_B)/f_B/t/x \quad (6)$$

where $X$ is the feature domain, $T$ is the time domain, and $F$ is the frequency of the occurrence domain.

#### 1) Union and Intersection Operations

A general procedure for undertaking the union and intersection operations on the 4D TMFs is outlined in Algorithm 1. The union of two TT2FSs $\vec{A}$ and $\vec{B}$ is a TT2FS defined as $\vec{A}\cup\vec{B}$ in (7):

$$\vec{A}\cup\vec{B} = \int_{x\in X}\int_{t\in T}\int_{f\in F} \mu_{\vec{A}\cup\vec{B}}(x,t,f)/f/t/x \quad (7)$$

**Algorithm 1:** Union and Intersection Operations on TT2FSs

**Result:** Resultant Temporal Membership Function (TMF) $\mu_{\vec{A}\oslash\vec{B}}(x,t,f_{A\oslash B})$ where $\oslash$ denotes the operation of either union or intersection.

Let concepts A and B on feature domain X (input to the algorithm) have TMFs denoted by $\mu_{\vec{A}}(x,t,f_A(t,\mu_A(x)))$ and $\mu_{\vec{B}}(x,t,f_B(t,\mu_B(x)))$ respectively with time intervals $\Delta t_q \in [\Delta t_1,...,\Delta t_Q]$ and zslices discretised at $z_i \in [z_1,z_2,...,z_I]$;

For each time interval $\Delta t_q$ the operation (union or intersection) on 3D time interval based TMF is computed independently by first taking the z-slices at $z_i \in [z_1,z_2,...,z_I]$ which renders the 3D time interval based TMF into interval type 2 (IT2) TMFs;

For each IT2 TMF, the operation is done as shown below in eq. (Alg 1.1);

**for** $x \in X$ **do**
  **for** $z_i < z_I$ **do**

$$\mu_{\vec{A}\oslash\vec{B},\Delta t_q}(x,f_{\Delta t_q}) = \sum_x \sum_{f_{\Delta t_q}\in[\odot(l_A,l_B),\odot(u_A,u_B)]} z_i/f_{\Delta t_q} \quad \text{(Alg 1.1)}$$

  **end**
**end**

where the summation signs in eq. (Alg 1.1) denotes the aggregation in set theoretic operation, $l$ and $u$ are the lower and upper conditional distribution values respectively of set $\vec{A}$ and $\vec{B}$ on z-slice $z_i$ and time interval $\Delta t_q$. For union operation, in eq. (Alg 1.1), the $\odot$ denotes *max* and for intersection operation $\odot$ denotes *min*.

where $\mu_{\vec{A}\cup\vec{B}}$ can be calculated by discretising the $T$ domain, and taking z-slices on $\mu_{\vec{A}\cup\vec{B},\Delta t_q}(x,t,f)$ values as outlined in (Alg 1.1) of Algorithm 1. In particular, for union operation, at time interval $\Delta t_q$ (Alg 1.1) takes the form of (8) when using the max t-conorm:

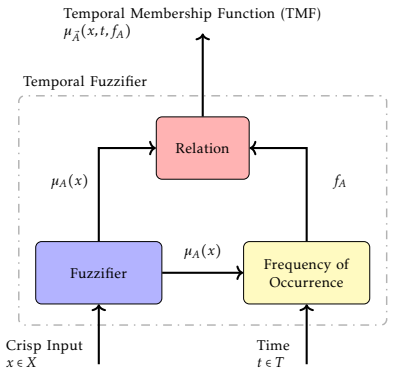
$$\mu_{\vec{A}\cup\vec{B},\Delta t_q}(x,f_{\Delta t_q}) = \sum_x \sum_{f_{\Delta t_q}\in[max(l_A,l_B),max(u_A,u_B)]} z_i/f_{\Delta t_q} \quad (8)$$

Likewise, the intersection of TT2FSs can be written as shown in (9)

$$\vec{A}\cap\vec{B} = \int_{x\in X}\int_{t\in T}\int_{f\in F} \mu_{\vec{A}\cap\vec{B}}(x,t,f)/f/t/x \quad (9)$$

where $\mu_{\vec{A}\cap\vec{B}}$ can be calculated by discretising the $T$ domain, and taking z-slices on $\mu_{\vec{A}\cap\vec{B},\Delta t_q}(x,t,f)$ values as outlined in (Alg 1.1) of Algorithm 1. In particular, for intersection operation, at time interval $\Delta t_q$ (Alg 1.1) takes the form of (10) when using the min t-norm. However, please note either product or min can be applied.

$$\mu_{\vec{A}\cap\vec{B},\Delta t_q}(x,f_{\Delta t_q}) = \sum_x \sum_{f_{\Delta t_q}\in[min(l_A,l_B),min(u_A,u_B)]} z_i/f_{\Delta t_q} \quad (10)$$

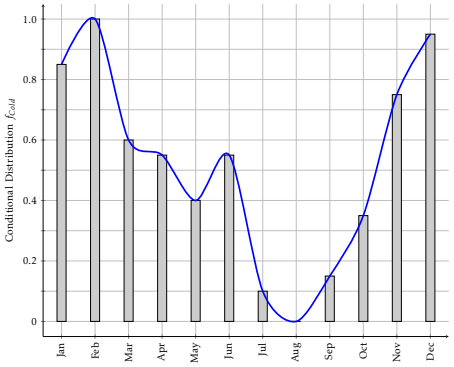



(a) Temporal Fuzzifier.

| $\Delta t_q$ | $t_n$ | $\text{Count}_{Cold}$ | $g_{Cold} \subseteq [0,1]$ |
|---|---|---|---|
| Winter | Jan | 17 | 17/20 = 0.85 |
| | Feb | 20* | 20/20 = 1.00 |
| Spring | Mar | 12 | 12/20 = 0.60 |
| | Apr | 11 | 11/20 =0.55 |
| | May | 8 | 8/20 =0.40 |
| Summer | Jun | 11 | 11/20 =0.55 |
| | Jul | 2 | 2/20=0.10 |
| | Aug | 0 | 0/20=0 |
| Autumn | Sep | 3 | 3/20=0.15 |
| | Oct | 7 | 7/20= 0.35 |
| | Nov | 15 | 15/20=0.75 |
| Winter | Dec | 19 | 19/20=0.95 |

* denotes the mode of observing *Cold*.

(b) Conditional relative frequency $g_{Cold}$




(c) Conditional distribution $f_{Cold}$

Fig. 4: (a) A schematic of temporal fuzzification for constructing temporal membership function (TMF). First, crisp values of input data from feature domain (i.e. $x \in X$) for a feature A are used to find primary membership function (MF) $\mu_A(x)$. The values of $\mu_A(x)$ associated with time $t \in T$ are then transformed into a conditional distribution, $f_A$, for each conceptual label (CoL) associated with A using discrete conditional relative frequency, $g_A$ as outlined in (1). The TMF for A, i.e. $\mu_{\vec{A}}(x,t,f_A)$, is computed by applying a fuzzy relation (such as Mamdani relation) on $\mu_A(x)$ and $f_A(t,\mu_A(x))$. (b) A hypothetical calculation for conditional relative frequency of conceptual label *Cold* where A denotes the thermal sensation, i.e. $g_{Cold}$ with respect to $N = 12$ discrete time points ($t_n$) i.e. the months and $Q = 4$ time intervals ($\Delta t_q$) representing the seasons in a year. The column $\text{Count}_{Cold}$ denotes the total number of times *Cold* was observed in the corresponding months i.e. the numerator in (1). The mode for $\text{Count}_{Cold}$ is 20, and is observed in February, which becomes the denominator of (1). (c) A bar plot of $g_{Cold}$ for all individual discrete time points (months) with an interpolated continuous $f_{Cold}$ superimposed in blue coloured solid line.

2) Defuzzification

In general, defuzzification converts a fuzzy set to an equivalent crisp number, and can be thought of as the inverse of fuzzification. For T1 fuzzy sets, defuzzification usually involves computing the centroid of the T1 fuzzy set [17] to compute a representative crisp number, as shown in (11).

$$x^* = \frac{\sum_{b=1}^{B} x_b \mu(x_b)}{\sum_{b=1}^{B} \mu(x_b)} \tag{11}$$

where $x^*$ is the centroid of the T1 MF defined on the domain $x \in X$. Here, the summation sign is used as in typical mathematical equations, i.e., for the case of the numerator, it is summing the product of $x$ values and their corresponding membership values whereas for the denominator it is summing the membership values corresponding to all $x_b$ values $\forall b \in [1,...,B]$.

For a 3D GT2 fuzzy set, defuzzification usually involves three steps, outlined as follows:

(i) Transforming a 3D GT2 fuzzy set to IT2 fuzzy sets by slicing the GT2 fuzzy set at given z-levels such as $z_i \in [z_1,...,z_I]$.

(ii) Type reducing the z-level based IT2 fuzzy sets results in two T1 fuzzy sets using Karnik Mendel (KM) method [18]. The type-reduced T1 fuzzy sets are composed of the left and right centroids of the IT2 fuzzy sets. More specifically, the KM method requires iterative process to compute left and right centroids resulting in two T1 fuzzy sets: $[y_{l_{z_1}}, y_{l_{z_2}},..., y_{l_{z_I}}]$ and $[y_{r_{z_1}}, y_{r_{z_2}},..., y_{r_{z_I}}]$ where $y_{l_{z_1}}$ is the left centroid at z-level 1 and $y_{r_{z_1}}$ is the right centroid at z-level 1 and so on.

(iii) Defuzzifcation of the type reduced T1 fuzzy sets, using centroid average, to find equivalent $y_l$ and $y_r$.

$$y_l = \frac{(z_1 * y_{l_{z_1}}) + (z_2 * y_{l_{z_2}}) + ... + (z_I * y_{l_{z_I}})}{z_1 + z_2 + ... + z_I} \tag{12}$$

$$y_r = \frac{(z_1 * y_{r_{z_1}}) + (z_2 * y_{r_{z_2}}) + ... + (z_I * y_{r_{z_I}})}{z_1 + z_2 + ... + z_I} \tag{13}$$

(iv) The final type-reduced crisp value is found using the Nie-Tan method [19] on $y_l$ and $y_r$.

In this work, the defuzzification of 4D TT2FS also involves TS-ZS approach (explained earlier in section III-D), i.e., taking the time interval based slice followed by z-slices. The time interval based TMF is 3D, and for each of the time interval ($\Delta t_q$) based TMF, z-slices at particular $z_i$ levels renders them as IT2 fuzzy sets. The KM procedure [18] can be applied on IT2 fuzzy sets, at each z-level, to compute T1 fuzzy sets composed of $[y_{l_{z_i,\Delta t_q}}, y_{r_{z_i,\Delta t_q}}]$ as outlined in (Alg 2.1). Using the centroid defuzzifier, the T1 fuzzy sets are defuzzified to give one equivalent $y_l$ and $y_r$, for that time interval, as outlined in (Alg 2.2) and (Alg 2.3). The Nie-Tan method [19] is then applied to compute one crisp value for that time interval. The defuzzification of TT2FSs, for a given time interval, is summarised in Algorithm 2. The procedure outlined in Algorithm 2 can be repeated for each time interval, i.e. $\Delta t_q$ where $q \in [1,...,Q]$, to obtain a crisp value for all time intervals.

**Algorithm 2:** Defuzzificaion of TT2FSs for a given time interval $\Delta t_q$

**Result:** Crisp value for a given time interval, denoted by $crisp_{\Delta t_q}$, where $\Delta t_q$ is the $qth$ time interval.

Let feature A on feature domain X have temporal membership function (TMF) denoted by $\mu_{\tilde{A}}(x,t,f_A(t,\mu_A(x)))$ with time intervals $\Delta t_q \in [\Delta t_1,...,\Delta t_Q]$ and z-slices discretised at $z_i \in [z_1,z_2,...,z_I]$;

For each 3D time interval based TMF, the defuzzification can be done independently, by first taking the z-slices at $z_i \in [z_1,z_2,...,z_I]$ which renders the 3D time interval based TMF into interval type 2 (IT2) MFs;

The left and right centroid for each IT2 TMF at z-location $z_i$, denoted by $C_{z_i,\Delta t_q}$, can be computed using Karnik-Mendel (KM) method [18] to give $[y_l,y_r]$ at that z-slice $z_i$ and time interval $\Delta t_q$ as outlined in eq. (Alg 2.1);

**for** $z_i \le z_I$ **do**

$$C_{z_i,\Delta t_q} = [y_{l_{z_i,\Delta t_q}},\ y_{r_{z_i,\Delta t_q}}] \qquad \text{(Alg 2.1)}$$

**end**

Defuzzifcation of the type reduced T1 fuzzy sets, using centroid average, to find equivalent $y_{l_{\Delta t_q}}$ and $y_{r_{\Delta t_q}}$;

$$y_{l_{\Delta t_q}} = \frac{(z_1 * y_{l_{z_1,\Delta t_q}}) + (z_2 * y_{l_{z_2,\Delta t_q}}) + ... + (z_I * y_{l_{z_I,\Delta t_q}})}{z_1 + z_2 + ... + z_I} \qquad \text{(Alg 2.2)}$$

$$y_{r_{\Delta t_q}} = \frac{(z_1 * y_{r_{z_1,\Delta t_q}}) + (z_2 * y_{r_{z_2,\Delta t_q}}) + ... + (z_I * y_{r_{z_I,\Delta t_q}})}{z_1 + z_2 + ... + z_I} \qquad \text{(Alg 2.3)}$$

A crisp value, $crisp_{\Delta t_q}$, can now be computed by applying Nie-Tan method [19] on $y_{l_{\Delta t_q}}$ and $y_{r_{\Delta t_q}}$.

## IV. TXAI Inference System (TXAI-IS)

In this section, the TXAI inference system (TXAI-IS) for classification problems is outlined. A general flowchart for the TXAI-IS is outlined in Fig. 5. The temporal fuzzifier constructs the 4D TT2FSs as outlined in Fig. 4 (a). To analyse a given dynamic process with respect to time, the TXAI-IS works for each time interval $\Delta t_q$ where $\Delta t_q \in [\Delta t_1,...,\Delta t_Q]$ independently. To this end, the 4D TT2FSs are first sliced based on the $\Delta t_q$, and inference is made on time sliced 3D TT2FSs using the temporal rules for the same $\Delta t_q$. Each time interval would entail a unique temporal rule base. The temporal rules can either be furnished by experts in the field or can be learnt from the input data using evolutionary algorithms such as genetic algorithm (GA) [20].

In addition, the assumptions of the proposed TXAI system with TT2FSs include: 1) the input features and output are observable, 2) a relation between input features and output exists, and 3) the relation between input features and output varies with time.

In the next subsections, the classification TXAI-IS is outlined in detail as the empirical study on which TXAI system is exemplified also undertakes a classification problem, i.e., occupancy dataset [11] is analysed to determine whether or not a room is occupied.

### A. Classification

For the classification problem, the TXAI-IS will predict one class or label for a given data instance for each time interval. The overall TXAI-IS for classification undertakes the following steps:

(i) Compute the membership degree for the time interval based 3D TT2FSs.

- The time interval based 3D TT2FSs are transformed into IT2 fuzzy sets by taking slices at predefined z-levels. The degree of membership at each z-level, such as $z_i \in [z_1,...,z_I]$ where $I$ is the total number of z slices, for a given 3D TT2FS A is given as follows [16]:

$$\tilde{A} = \{(x,u,z) | \forall x \in X, \ \forall u \in [\underline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{A}}(x)] \subseteq [0,1]\} \qquad (14)$$

where $\mu_{\tilde{A}}$ is the membership degree of the IT2 fuzzy set $\tilde{A}$ at the predefined $z$ level.

(ii) Compute the firing strength for each rule, at each z-level.

- The $\overline{upper}$ and $\underline{lower}$ firing strength of a given rule $p$, $\overline{w}_p$ and $\underline{w}_p$ respectively, is the degree of match between the rule $p$ and the data instance $x$. It is computed as:

$$\overline{w}_p(x^k) = \prod_{k=1}^{a} \overline{\mu}_{\tilde{A}}(x^k)$$
$$\underline{w}_p(x^k) = \prod_{k=1}^{a} \underline{\mu}_{\tilde{A}}(x^k) \qquad (15)$$

where $p$ is the rule number, $a$ is the total number of antecedents in the rule $p$ and $x^k$ is an input ($k$) of the actual data instance to be classified.

(iii) Compute the rule weight (RW) for each rule, at each z-level.

- The RW is a measure of a given rule's dominance and is computed as shown in (16).

$$\overline{RW}_p = \overline{c}_p \times \overline{s}_p$$
$$\underline{RW}_p = \underline{c}_p \times \underline{s}_p \qquad (16)$$

where $c$ is the confidence of the rule $p$ and $s$ is the support of the $p$th rule.

- The confidence of a rule is a measure of the likelihood to correctly classify a given data instance. It is calculated as shown in eq. (17)

$$\overline{c}_p(Ants_p \Rightarrow Cons_p) = \frac{\sum_{x\in(Ants_p\Rightarrow Cons_p)} \overline{w}_p(x)}{\sum_{p=1,x\in(Ants_p)}^{P} \overline{w}_p(x)} \quad (17)$$

$$\underline{c}_p(Ants_p \Rightarrow Cons_p) = \frac{\sum_{x\in(Ants_p\Rightarrow Cons_p)} \underline{w}_p(x)}{\sum_{p=1,x\in(Ants_p)}^{P} \underline{w}_p(x)}$$

where $Ants_p$ and $Cons_p$ are the antecedents and consequent respectively of the rule $p$. The numerator sums the firing strength of all the data instances that have the same antecedents and consequent as the rule $p$. Whereas the denominator sums the firing strength of all the data instances that have the same antecedents as the rule $p$ irrespective of the consequent- for all the rules $[1,...,P]$, where $P$ is the total number of rules.

- The support of a rule is calculated as shown in (18)

$$\overline{s}_p(Ants_p \Rightarrow Cons_p) = \frac{\sum_{x\in(Ants_p\Rightarrow Cons_p)} \overline{w}_p(x)}{P} \quad (18)$$

$$\underline{s}_p(Ants_p \Rightarrow Cons_p) = \frac{\sum_{x\in(Ants_p\Rightarrow Cons_p)} \underline{w}_p(x)}{P}$$

with $P$ as the total number of rules.

(iv) Compute the association degree of each rule, with a given data instance, for each z-level.

- The association degree of a rule $p$ with a given data instance $x$ is computed as shown in (19):

$$\overline{h}_p = \overline{w}_p(x) \times \overline{RW}_p \quad (19)$$
$$\underline{h}_p = \underline{w}_p(x) \times \underline{RW}_p$$

(v) Predict the label.

- Find a value of the association degree, $h$, for each rule by using Nie-Tan [19] method on the $\underline{h}$ and $\overline{h}$ which are found using (Alg 2.2) and (Alg 2.3).
- The rule with the highest association degree, $h$, predicts the label for the given data instance.

(vi) The steps outlined above (i)-(v) are repeated for each time interval to predict a label for all time intervals.

### B. Numerical Step-wise Example

In this section, a binary classification problem using TXAI-IS is exemplified using a hypothetical dataset with two input features, *Feature1* and *Feature2*, and one output. Let time intervals be defined over a day such as Morning, Daytime, and Evening with three CoLs associated with the inputs (*Feature1* and *Feature2*) be: [Low, Medium, High] and output labels be *Output1* and *Output2*.

First, TT2FSs for both inputs (*Feature1* and *Feature2*) are constructed using temporal fuzzifier, as outlined in Fig. 4. Also, for each time interval, the rules will be different but the overall process to determine the output label is same. In the following steps, we exemplify how the output label is predicted for one time interval, in this example, Morning.

Let the rules (R) outlining the relation between input features and output for Morning be as listed in (20). The corresponding lower and upper rule weights (RW) at each z-level are as listed in Table III. In the following steps i)- iv) we show how a corresponding label for *Output* is predicted using TXAI-IS for input values of *Feature1* = 19.7 and *Feature2* be = 4.3. In this example, the z-level is discretised at $z_{0.2}$, $z_{0.4}$, $z_{0.6}$, $z_{0.8}$, and $z_{1.0}$.

$R_1$ : IF *Feature1* is *Low* and *Feature2* is *Medium* THEN *Output* is *Output2*
$R_2$ : IF *Feature1* is *Medium* and *Feature2* is *Medium* THEN *Output* is *Output1*
$R_3$ : IF *Feature1* is *High* and *Feature2* is *High* THEN *Output* is *Output1* (20)

(i) The degree of membership for each CoL of the inputs *Feature1* and *Feature2* is determined from the time interval (Morning) based 3D TMF. The membership degree is the value of the conditional distribution at a given input value and corresponding z-level as outlined in (14). Let the corresponding membership degrees for each CoL of the inputs *Feature1* and *Feature2* be as noted in Table II.

TABLE II: The hypothetical lower (L) and upper (U) degree of membership values of the conceptual labels (CoLs) of *Feature1* and *Feature2* for the time interval Morning for five z levels: $z_{0.2}$, $z_{0.4}$, $z_{0.6}$, $z_{0.8}$, and $z_{1.0}$ with input value of *Feature1* = 19.7, and *Feature2* = 4.3.

| CoLs | CoLs | | $z_{0.2}$ | $z_{0.4}$ | $z_{0.6}$ | $z_{0.8}$ | $z_{1.0}$ |
|---|---|---|---|---|---|---|---|
| *Feature1* | Low | L | 0.50 | 0.52 | 0.54 | 0.52 | 0.51 |
| | | U | 0.61 | 0.63 | 0.64 | 0.61 | 0.60 |
| | Med. | L | 0.63 | 0.63 | 0.65 | 0.63 | 0.61 |
| | | U | 0.77 | 0.78 | 0.78 | 0.77 | 0.75 |
| | High | L | 0.65 | 0.64 | 0.64 | 0.63 | 0.63 |
| | | U | 0.69 | 0.69 | 0.68 | 0.68 | 0.67 |
| *Feature2* | Low | L | 0.31 | 0.31 | 0.31 | 0.31 | 0.31 |
| | | U | 0.32 | 0.32 | 0.32 | 0.32 | 0.32 |
| | Med. | L | 0.50 | 0.55 | 0.55 | 0.54 | 0.53 |
| | | U | 0.58 | 0.59 | 0.59 | 0.58 | 0.57 |
| | High | L | 0.40 | 0.40 | 0.40 | 0.42 | 0.44 |
| | | U | 0.43 | 0.43 | 0.46 | 0.46 | 0.49 |

(ii) The firing strength of each rule listed in (20) are found, using the membership degree in Table II, as outlined in (15) and listed in Table III. As an example, for $R_1$ the lower firing strength at $z = 0.6$, $\underline{w}_{1_{z=0.6}}$, can be calculated as follows:

$$\underline{w}_{1_{z=0.6}}(x = [19.7, 4.3]) = \prod_{k=1}^{2} \underline{\mu}(x^k)$$
$$= 0.54 * 0.55 = 0.297 \quad (21)$$

(iii) The association degree of each rule with the input data instance is determined, using the firing

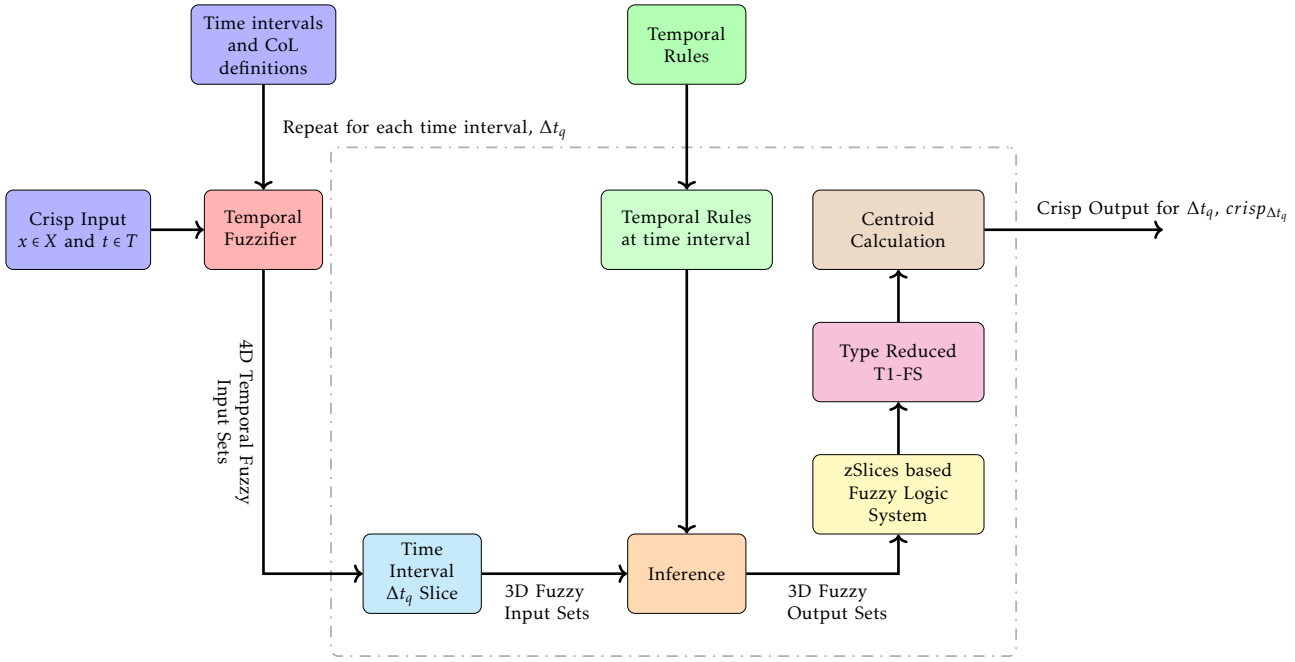



Fig. 5: A general schematic representation delineating the interlinks between salient components of a time-dependent explainable artificial intelligence (TXAI) inference system (TXAI-IS).

TABLE III: The lower and upper firing strengths, $\underline{w}$ and $\overline{w}$ respectively, for the hypothetical rules listed in (20) for time interval Morning. The rule weights (RW) at each z-level are also listed.

| Rule | Firing Strength, $w$ | z-level | | | | | Consequent | Rule Weight RW | z-level | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $z_{0.2}$ | $z_{0.4}$ | $z_{0.6}$ | $z_{0.8}$ | $z_{1.0}$ | | | $z_{0.2}$ | $z_{0.4}$ | $z_{0.6}$ | $z_{0.8}$ | $z_{1.0}$ |
| $R_1$ | Lower | 0.25 | 0.286 | 0.297 | 0.281 | 0.27 | *Output2* | Lower | 0.31 | 0.30 | 0.30 | 0.29 | 0.27 |
| | Upper | 0.354 | 0.372 | 0.378 | 0.354 | 0.342 | | Upper | 0.35 | 0.34 | 0.34 | 0.31 | 0.30 |
| $R_2$ | Lower | 0.315 | 0.347 | 0.358 | 0.34 | 0.323 | *Output1* | Lower | 0.69 | 0.69 | 0.68 | 0.66 | 0.66 |
| | Upper | 0.447 | 0.46 | 0.46 | 0.447 | 0.427 | | Upper | 0.73 | 0.73 | 0.72 | 0.72 | 0.72 |
| $R_3$ | Lower | 0.26 | 0.256 | 0.256 | 0.265 | 0.277 | *Output1* | Lower | 0.22 | 0.21 | 0.21 | 0.21 | 0.21 |
| | Upper | 0.297 | 0.297 | 0.313 | 0.313 | 0.328 | | Upper | 0.24 | 0.22 | 0.22 | 0.22 | 0.22 |

strength in Table III, as outlined in (19). The upper and lower values of the association degree for the five z-levels are as listed in Table IV. As an example, for $R_2$ the upper association degree at $z = 0.2$, $\overline{h}_{2_{z=0.2}}$, can be calculated as follows:

$$\begin{aligned}\overline{h}_{2_{z=0.2}} &= \overline{w}_{2_{z=0.2}}(x) \times \overline{RW}_{2_{z=0.2}} \\ &= 0.447 * 0.73 = 0.326\end{aligned} \tag{22}$$

(iv) The consequent of the rule with the highest association degree with the input data instance becomes the predicted label for a given time interval. The crisp value for the association degree of each rule is found using (Alg 2.2) and (Alg 2.3). As an example, the crisp value of association degree for $R_3$ is found as follows:

$$\begin{aligned}h_{3_l} &= \frac{0.2 * (\underline{h}_{3_{0.2}}) + ... + 1.0 * (\underline{h}_{3_{1.0}})}{0.2 + 0.4 + 0.6 + 0.8 + 1.0} \\ &= \frac{0.2 * 0.057 + 0.4 * 0.054 + ... + 1 * 0.058}{3} = 0.056 \\ h_{3_u} &= \frac{0.2 * (\overline{h}_{3_{0.2}}) + ... + 1.0 * (\overline{h}_{3_{1.0}})}{0.2 + 0.4 + 0.6 + 0.8 + 1.0} \\ &= \frac{0.2 * 0.071 + 0.4 * 0.065 + ... + 1 * 0.072}{3} = 0.0696 \\ h_{3_{crisp}} &= \frac{0.056 + 0.0696}{2} = \frac{0.1256}{2} = 0.063\end{aligned} \tag{23}$$

In this illustrative example, $R_2$ has the highest association degree (tabulated in Table IV) hence the predicted output for the input data instance (*Feature1* = 19.7 and *Feature2* be = 4.3) for time interval Morning is the consequent of $R_2$, i.e., *Output1*.

The same process can be repeated for each time interval with their respective rules to predict a label for the output. Hence, in this numerical example, there will be three output labels for a total of three time intervals.

TABLE IV: The lower (L) and upper (U) association degrees, $h$, for each of the three rules (R) listed in (20) with input data instance: Feature 1= 19.7, Feature 2 = 4.3. The association degrees' crisp value, for each of the rules $R_1$-$R_3$, denoted $h_{crisp}$ is also listed.

| **R** | **h** | $z_{0.2}$ | $z_{0.4}$ | $z_{0.6}$ | $z_{0.8}$ | $z_{1.0}$ | $h_{crisp}$ |
|---|---|---|---|---|---|---|---|
| $R_1$ | $L$ | 0.077 | 0.086 | 0.089 | 0.081 | 0.073 | 0.097 |
| | $U$ | 0.124 | 0.126 | 0.128 | 0.11 | 0.103 | |
| $R_2$ | $L$ | 0.217 | 0.239 | 0.243 | 0.225 | 0.213 | 0.274 |
| | $U$ | 0.326 | 0.336 | 0.331 | 0.322 | 0.308 | |
| $R_3$ | $L$ | 0.057 | 0.054 | 0.054 | 0.056 | 0.058 | 0.063 |
| | $U$ | 0.071 | 0.065 | 0.069 | 0.069 | 0.072 | |

### C. Estimating Temporal Trajectories from TXAI Models

The temporal trajectories of a dynamic system can be outlined by the TXAI system by making use of the conditional distribution integrated into the TXAI system. The trajectories of a TXAI model is motivated by the work of Filev et al. [21] that embodies fuzzy transition events defined by joint possibility encompassing the current and future prototypical rules. More specifically, the TXAI system can delineate a rule transition matrix (RTM) which will entail the joint possibility of the rules in present ($\Delta t$) and future ($\Delta t^+$) time intervals. In mathematical terms, for a total of U rules in time interval $\Delta t$, and a total of V rules in time interval $\Delta t^+$, the RTM can be written as follows [21]:

$$RTM(\Delta t, \Delta t^+) = \begin{bmatrix} \pi_{11} & \dots & \pi_{1N} \\ \vdots & \ddots & \vdots \\ \pi_{M1} & \dots & \pi_{UV} \end{bmatrix} \tag{24}$$

where $\pi_{cd}$ is the rule transition possibility (RTP) for the $c^{th}$ rule, $r_c$, in time interval $\Delta t$ and the $d^{th}$ rule, $r_d$, in time interval $\Delta t^+$ as given by (25).

$$\pi_{cd} = \eta_{cd} \times \frac{S_{cd}}{S_{\Delta t^+}} \tag{25}$$

where $\eta_{cd}$ is the joint possibility for the two rules to be prototypical in their respective time intervals, and the ratio $\frac{S_{cd}}{S_{\Delta t^+}}$ entails the number of times $r_c$ and $r_d$ are observed in their respective time intervals with respect to all V rules in $\Delta t^+$. The following equations, (26) - (28), outline how $\eta_{cd}$ and the ratio $\frac{S_{cd}}{S_{\Delta t^+}}$ are computed.

$$\eta_{cd}(r_{c,\Delta t}, r_{d,\Delta t^+}) = \gamma_c(r_c, \Delta t) \times \gamma_d(r_d, \Delta t^+) \tag{26}$$

Where $\gamma$ is computed by applying the t-norm operator (product or minimum type) to the conditional distribution values of the antecedents of a given rule ($r$) in a given time interval ($\Delta t$ or $\Delta t^+$); mathematically expressed as shown in equation (27) for rule ($r_c$) in time interval ($\Delta t$). The computation of the conditional distribution, $f$, is previously outlined in Section III-C (in particular see (1)).

$$\gamma_c(r_{c,\Delta t}) = f_c(Ant_{1,r_c}, \Delta t) \times f_c(Ant_{2,r_c}, \Delta t) \times \dots \times f_c(Ant_{a,r_c}, \Delta t) \tag{27}$$

where $a$ is the total number of antecedents (Ant) of rule $r_c$. The elements for computing the ratio $\frac{S_{cd}}{S_{\Delta t^+}}$ are outlined in (28):

$$S_{cd} = \sum r_{c,\Delta t} r_{d,\Delta t^+}$$
$$S_{\Delta t^+} = \sum_{d=1}^{V} r_{d,\Delta t^+} \tag{28}$$

where the numerator, $S_{cd}$, represents the sigma count of the number of times $r_c$ and $r_d$ are observed in their respective time intervals, and the denominator, $S_{\Delta t^+}$, denotes the sigma count of observing all V rules in $\Delta t^+$.

## V. Case Study: Time-dependant Occupancy Dataset

In this section, a temporal occupancy dataset [11] is used to exemplify the proposed TXAI system modelling. The occupancy dataset entails measurements of a room along with the time of when the measurement is recorded. In particular, it includes measurements of the room temperature, light, $CO_2$, and a binary label of whether or not the room is occupied. There are 8,143 data instances in the dataset taken over a period of a few weeks.

In this work, the dataset [11] is used for classification problem where TXAI system predicts whether or not the room is occupied based on the room measurements. The inputs of temperature, light, and $CO_2$ are used to predict whether or not the room is occupied. Three CoLs of Low, Medium, and High are associated with inputs of temperature, light, and $CO_2$. The primary MF of the CoLs for all inputs are empirically found. The time is discretised at each hour of the day hence a total of $N = 24$ time points with a total of three time intervals defined at Morning, Daytime, and Evening, as also summarised in Table V. The z-slices are obtained on locations [0.2, 0.4, 0.6, 0.8, 1.0]. All aforementioned parameters values are selected so as to reflect the inherent dynamics of the system (such as discretising time at each hour) and to obtain a good enough TXAI model without adding too much computational complexity, for example, the more the z-slices the more accurate the TXAI model would be but at a greater computational cost($p = Q * z_I$ but independent of the data size in each $\Delta_{t_q}$).

The conditional distribution for each CoL of every input is computed on the entire dataset. Once the conditional distributions are computed, the learning procedure focuses on the data belonging to each interval. A 10-repeated nested cross-validation procedure is adopted. The dataset is split into a disjoint stratified train, validation and the test set to ensure a random selection of the datasets (train, validation, and test) is not creating any bias in the results. Each repetition, 20% of the dataset is held out as a test set, and the remaining is used to build the train and validation sets.

TABLE V: The classification problem is exemplified using the proposed Time-dependent eXplainable Artificial Intelligence (TXAI) system with occupancy dataset [11]. The output for the classification problem predicts the label of whether the room is occupied or not. The time points ($t_n$) for calculating the frequency of occurrence are 24 on account of the number of hours in a given day with a total of three corresponding time intervals ($\Delta t_q$) of Morning (time < 11 am), Daytime (11 am < time < 7pm), and Evening (time > 7pm).

| Problem | Input/Output | Feature/Label | Conceptual Labels (CoLs) | N | Time Intervals, $\Delta t_q$ |
|---|---|---|---|---|---|
| Classification | Input | Temperature | Low, Medium, High | 24 | Morning, Daytime, Evening |
| | | Light | Low, Medium, High | 24 | Morning, Daytime, Evening |
| | | $CO_2$ | Low, Medium, High | 24 | Morning, Daytime, Evening |
| | Output | Occupied<br>Not Occupied | - | - | Morning, Daytime, Evening |

Train and validation sets are determined in an inner 10-fold procedure, where a fold is used for validation and the rest for training to determine the rule weights. Balanced accuracy and other performance metrics are computed over each validation and test set.

A rule-base is formed for each time interval. The rules are learned using GA [22] such that they (rules) attain optimally balanced accuracy on the validation datasets. The GA parameters specification includes the number of generations, set at 20, with each generation having a population of 50. Moreover, the GA is leveraged to find the rules that are prototypical for each time interval. The number of antecedents in each rule can be at most 3 but not more to underpin explainability and hamper model complexity therefore precluding over-fitting. For the same reason, the maximum number of rules in each candidate rule-base for each time interval was limited to 30, although further pruned when its weight (eq. (16)) does not surpass a tolerance threshold of 0.001.

In order to compare the performance of the proposed TXAI system, numerous state-of-the-art classifiers which can both analyse time-series data and/or are explainable have been used. More specifically, for comparison with temporal analysis Long Short-Term Memory (LSTM) [23] and Hidden Markov Models (HMM) [24] are used, for comparison with explainable models the standard GT2 based XAI system is used, and for partial explainablility Decision Trees (DT) [25] is used. In addition, a comparison is also made with a temporal convolutional network (TCN) [26] for comparison with deep learning methods [27]. Parametrization and configuration was set to default mode of their respective libraries (Sklearn and Keras). For methods with no modelling with respect to a time component, time is given as an extra input feature. Moreover, the train, validation, and test dataset splits are similar across all methods and for GT2 based XAI in particular, the location of z-slices, and the GA parameters for rule learning are also identical to those of TXAI system.

### A. Results

For the classification problem undertaken, using the occupancy dataset, the proposed TXAI system and numerous state-of-the-art classification methods predict whether or not the room is occupied. The mean (and standard deviation) f-score obtained using TXAI system on the 10 test datasets is 95.30% which is the highest score on the test dataset across all classifiers except TCN. The other classification metrics investigated in this work are balanced accuracy, recall, and precision. A bar plot for the aforementioned classification metrics for both the proposed TXAI and the state-of-the-art AI methods (TCN, LSTM, DT, HMM, GT2 based XAI) on 10 times repeated 10-fold validation and test datasets is shown in Fig. 6 (a) and (b) respectively. In addition, a convergence graph that outlines how the GA optimisation converges with respect to balanced accuracy for both TXAI and GT2 based XAI systems is also shown in Fig. 6 (c).

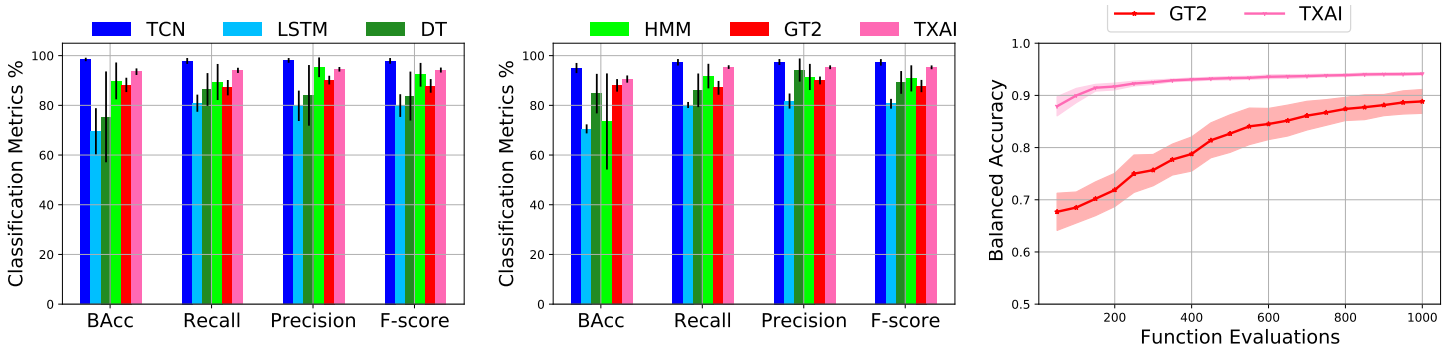
The rules outlined by TXAI and GT2 based XAI systems which are prototypical for whether or not the room is occupied are listed in Table VI. For the TXAI system, please note that the rules are found separately for each time interval (Morning, Daytime, and Evening) whereas, for GT2 based XAI system, the time intervals are one of the antecedents of the rules. In general, for both TXAI and GT2 based XAI systems, the rules outline that when the room measurements have higher values, the room is more likely to be occupied, and when the room measurements are on the lower end, the room is more likely to be not occupied.

For the TXAI system, the temporal trajectories of a time-variant system can also be investigated using the rule transition matrices (RTMs), previously outlined in section IV-C. The individual RTMs transitioning from one time interval ($\Delta t$) to another i.e., from Morning to Daytime, from Daytime to Evening, and from Evening to Morning, represent the joint possibilities of observing a given rule in $\Delta t^+$ with respect to the rules in $\Delta t$. The rules corresponding to the highest RTPs (rule transitioning possibilities) are also joined with lines in the column *RT* (rule transitions) in Table VI and illustrated in a schematic in Fig. 7.

### B. Discussion

In this work, the proposed TXAI system is used to model an occupancy dataset [11] for the classification problem of whether or not the room is occupied. For comparison purposes, several state-of-the-art explain-

Fig. 6: A comparison of the classification prowess of the proposed time-dependent eXplainable artificial intelligence (TXAI) system with numerous state-of-the-art classification systems: temporal convolutional networks (TCN), Long Short-Term Memory (LSTM), Decision Trees (DT), Hidden Markov Models (HMM), and the standard general type-2 (GT2) based XAI system for the classification problem using an occupancy dataset [11]. a) and b) show the classification metrics of the aforementioned systems on 10 times the 10-fold stratified validation and test dataset respectively. The classification metrics are balanced accuracy (BAcc), recall, precision, and f-score. c) A comparison of the convergence of TXAI with GT2 based XAI system using balanced accuracy for a total of 20 generations with a population of 50 each resulting in a total of 1000 function evaluations.




a) Validation datasets b) Test datasets c) Convergence graph

able (GT2 based XAI system), partially explainable (DT), and non-explainable methods that can analyse temporal information (LSTM and HMM) as well as TCN are also applied to the aforementioned classification problem. As can be noted from the Fig. 6 (a) and (b), TXAI offers greater classification performance than all classifiers (for e.g. for mean fscore TXAI performs better than LSTM by 18.19%, DT by 6.81%, HMM by 4.90% , GT2 based XAI system by 8.58% on test datasets) except TCN (for mean fscore TCN performs better than TXAI by 4.67% on test datasets). However, the TCN classification mechanism is not explainable hence unable to shed light on the prediction of the room occupancy based on input features of Temperature, Light, $CO_2$, and Time.

With respect to the comparison with the GT2 based XAI system, the only explainable system apart from the proposed TXAI system, a convergence graph plotted in Fig. 6 (c) also highlights that TXAI system converges (~500 function evaluations) twice as faster than standard GT2 based XAI system (~1000 function evaluations) whilst also yielding higher classification metrics (Fig. 6 (a) and (b)). Moreover, the rules outlined by the explainable systems, TXAI and XAI systems, are listed in Table VI, and both systems are in agreement that when the room measurements (Temperature, Light, and $CO_2$) have higher values, then the room is likely to be occupied, and when the room measurements are lower, then the room is likely to be not occupied. However, the rules for TXAI also offer greater insight into how the room measurements are interlinked with respect to predicting room occupancy. For example, for the time interval Morning, rule no 5 (see Table VI) outlines that if both inputs of Temperature and Light have high values then the room is likely to be occupied. In this regard, rules across time intervals shed light on the intertwined CoLs of the inputs prototypical for decoding the room occupancy.

Furthermore, the TXAI systems are also able to shed light on the temporal trajectories of the system being modelled using RTMs, previously outlined in Section IV-C, and illustrated in Fig. 7. The RTPs (rule transition possibilities), which are the elements of the RTMs, represent the joint likelihood of observing a rule in one time interval (rows) and then observing another rule in the next time interval (columns). For example, in the RTM transitioning from Morning to Daytime, the rules with the highest RTP are rule number 12 (for time interval Morning) and rule number 5 (for the next time interval Daytime). For the particular case of the occupancy datasets, the RTMs and the corresponding RTPs outline the trajectory across time as the TXAI model transitions from one time interval to another. In this case, an analysis of the occupancy dataset can be leveraged for the efficient energy management of smart homes using the predictive power of the RTMs [28].

Indeed, the motivation for developing the TXAI systems is to be able to analyse time-dependent real processes across time. In this regard, conditional distribution integrated within the TXAI model can be used to obtain the RTMs. The RTMs entail the likelihood of observing the transition of a real-life process from one time point to another. The proposed TXAI system can shed light not only on which rules are prototypical for each of the time intervals but also on the likelihood of observing the rules across the different time points.

## VI. Conclusion

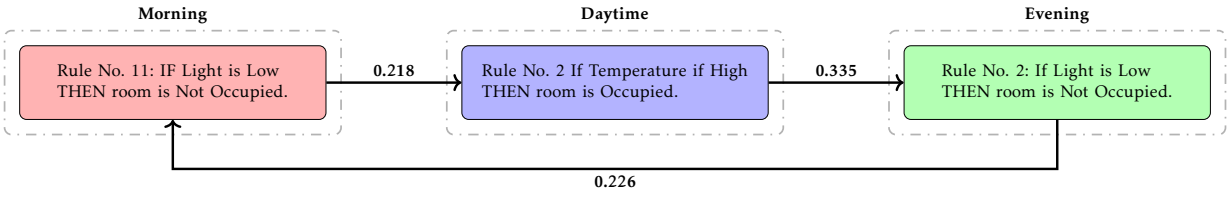
The ability of an explainable system to model a real-life process in terms of its characteristic features is of paramount significance to inform about the nature of the process. In this regard, XAI systems have proved pivotal for increasing our understanding of numerous complex real-life processes. However, non time-dependent XAI systems are not able to analyse real-life processes

TABLE VI: The prototypical rules were obtained by the proposed time-dependent explainable artificial intelligence (TXAI) system for the binary classification problem (room occupied or not) using the occupancy dataset. In the column RT (Rule Transition), the rules with the highest rule transition possibility (RTP) for transitioning from one time interval to another are marked with connecting lines: red line connects the rules with the highest RTP for going from Morning to Daytime, blue line connects the rules with the highest RTP for going from Daytime to Evening, and green lines connects the rules with the highest RTP for going from Evening to Morning of the next day. The numerical values of the corresponding RTPs are also listed. The rules obtained using the standard general type-2 (GT2) explainable artificial intelligence (XAI) system with time as another input are also outlined at the end of the table for comparison purposes.

| Method | Time | Rule No. | Rule | Rule Weight | Rule Transition (RT) |
|---|---|---|---|---|---|
| Time-dependent eXplainable Artificial Intelligence (TXAI) | Morning | 1 | IF Light is High THEN room is Occupied | 0.346 | |
| | | 2 | IF Temperature is High THEN room is Occupied | 0.079 | |
| | | 3 | IF $CO_2$ is Medium THEN room is Occupied | 0.050 | |
| | | 4 | IF $CO_2$ is High THEN room is Occupied | 0.046 | |
| | | 5 | IF Temperature is High AND Light is High THEN room is Occupied | 0.018 | |
| | | 6 | IF Light is High AND $CO_2$ is Medium THEN room is Occupied | 0.014 | |
| | | 7 | IF Light is High AND $CO_2$ is High THEN room is Occupied | 0.012 | |
| | | 8 | IF Temperature is Medium THEN room is Occupied | 0.012 | |
| | | 9 | IF Temperature is High AND $CO_2$ is High THEN room is Occupied | 0.011 | |
| | | 10 | IF Temperature is Medium AND $CO_2$ is Medium THEN room is Occupied | 0.007 | |
| | | 11 | IF Light is Low THEN room is Not Occupied | 1.000 | |
| | | 12 | IF Temperature is Low THEN room is Not Occupied | 0.073 | |
| | Daytime | 1 | IF Light is High THEN room is Occupied | 0.473 | |
| | | 2 | IF Temperature is High THEN room is Occupied | 0.277 | |
| | | 3 | IF $CO_2$ is High THEN room is Occupied | 0.110 | |
| | | 4 | IF Temperature is Medium AND Light is High THEN room is Occupied | 0.017 | |
| | | 5 | IF Temperature is High AND Light is High AND $CO_2$ is High THEN room is Occupied | 0.015 | |
| | | 6 | IF Light is Low THEN room is Not Occupied | 1.000 | |
| | | 7 | IF $CO_2$ is Low THEN room is Not Occupied | 0.50 | |
| | | 8 | IF Light is Medium THEN room is Not Occupied | 0.147 | |
| | | 9 | IF Temperature is High AND Light is Low THEN room is Not Occupied | 0.011 | |
| | Evening | 1 | IF Light is High THEN room is Occupied | 0.005 | |
| | | 2 | IF Light is Low THEN room is Not Occupied | 1.000 | |
| | | 3 | IF Light is Low AND $CO_2$ is Low THEN room is Not Occupied | 0.108 | |
| | | 4 | IF Temperature is High AND Light is Low THEN room is Not Occupied | 0.041 | |
| eXplainable Artificial Intelligence (XAI) | | 1 | IF Light is High AND Time is Daytime THEN room is Occupied | 0.580 | |
| | | 2 | IF Light is High AND Time is Morning THEN room is Occupied | 0.425 | |
| | | 3 | IF Temperature is High AND Time is Daytime THEN room is Occupied | 0.419 | |
| | | 4 | IF Light is Low AND Time is Morning THEN room is Not Occupied | 1.000 | |
| | | 5 | IF $CO_2$ is Medium AND Time is Evening THEN room is Not Occupied | 0.789 | |



Fig. 7: A schematic presenting the evolution of the occupancy system based on the rules with the highest rule transition possibilities (RTPs) found by the proposed time-dependent explainable artificial intelligence (TXAI). The rules in two consecutive time intervals with the highest RTPs are linked together to show how the occupancy system is evolving from one time interval to another. A complete list of all the rules delineated by TXAI for the occupancy dataset is enumerated in Table VI.



across time. This is a critical limitation of standard XAI systems for modelling time-variant real-life processes, especially where time is a defining parameter for the model, i.e., the real-life process behaves differently across time (for example, functional brain development [29, 30]). To this end, in this work, we propose a new time-dependent XAI system, called TXAI, characterised with time-conditioned distribution for analysing a time-variant real-life process across time.

The proposed TXAI system can delineate the trajectories of a dynamic, real-life process across time. In addition, a comparison with state-of-the-art AI systems, with varying levels of explainability, manifested that the proposed TXAI performed better than most of the compared AI systems (for e.g. for mean fscore TXAI performs better than LSTM by 18.19%, DT by 6.81%, HMM by 4.90%, GT2 based XAI system by 8.58% on test datasets) except TCN which is a much more complex, and a black-box method. XAI systems based on standard FLS (e.g., T1, IT2 or GT2) are unable to integrate information relative to the time dimension. More specifically, TXAI system credit the membership value of a fuzzy concept given the fuzzy concept is likely to occur at the time of observation of fuzzy concept using conditional distribution. The conditional distribution is then utilised to investigate the evolution of the process across different time intervals. In this way, the TXAI is able to predict the likelihood of observing prototypical rules of the process across different time intervals. For future works, the proposed TXAI system can have profound implications to contribute to our understanding of temporal real-life processes, for instance human-centred systems and life sciences. Further, for these future life science studies, we would also endeavour that TXAI entails all ethical concerns accounted for a more fair, and complete TXAI analysis.

**Acknowledgement**: We would like to gratefully acknowledge *Oracle for Research*, and specially Richard Pitts, and Mike Reilly for their technical assistance and support in the computational resources in Oracle Cloud Resource for this research.

## References

[1] M. S. A. Lee and L. Floridi and A. Denev, *Innovating with confidence: embedding AI governance and fairness in a financial services risk management framework*. Springer, 2021, ch. 9, pp. 353–371.

[2] K. J. Yaxley, K. F. Joiner, and H. Abbass, "Drone approach parameters leading to lower stress sheep flocking and movement: sky shepherding," *Scientific reports*, vol. 11, pp. 1–9, 2021.

[3] R. Li, D. Auer, C. Wagner, and X. Chen, "A generic ensemble based deep convolutional neural network for semi-supervised medical image segmentation," *In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pp. 1168–1172, 2020.

[4] M. Kiani, J. Andreu-Perez, H. Hagras, E. I. Papageorgiou, M. Prasad, and C. T. Lin, "Effective Brain Connectivity for fNIRS with Fuzzy Cognitive Maps in Neuroergonomics," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, pp. 50–63, 2022.

[5] H. Lu, Y. Li, M. Chen, H. Kim, and S. Serikawa, "Brain intelligence: go beyond artificial intelligence," *Mobile Networks and Applications*, vol. 23, pp. 368–375, 2018.

[6] J. Andreu-Perez, H. Pérez-Espinosa, E. Timonet, M. Kiani, and et al., "A generic deep learning based cough analysis system from clinically validated samples for point-of-need covid-19 test and severity levels," *IEEE Transactions on Services Computing*, 2021.

[7] H. Hagras, "Toward Human-Understandable, Explainable AI," *Computer*, vol. 51, no. 9, pp. 28 – 36, 2018.

[8] G. Alicioglu and B. Sun, "A survey of visual analytics for explainable artificial intelligence methods," *Computers & Graphics*, 2021.

[9] L. A. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, no. 1, pp. 28 – 44, 1975.

[10] ——, "Fuzzy Sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[11] L. M. Candanedo and V. Feldheim, "Accurate occupancy detection of an office room from light, temperature, humidity and CO2 measurements using statistical learning models," *Energy and Buildings*, vol. 112, pp. 28–39, 2016.

[12] J. M. Garibaldi, M. Jaroszewski, and S. Musikasuwan, "Nonstationary Fuzzy Sets," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 4, pp. 1072 – 1086, 2008.

[13] A. V. Kostikova, P. V. Tereliansky, A. V. Shuvaev, V. N. Parakhina, and P. N. Timoshenko4, "Expert Fuzzy Modeling of dynamic properties of complex systems," *ARPN Journal of Engineering and Applied Sciences*, vol. 11, no. 17, pp. 10 222–10 230, 2016.

[14] H. Maeda, S. Asaoka, and S. Murakami, "Dynamical fuzzy reasoning and its application to system modeling," *Fuzzy Sets and Systems*, vol. 80, no. 1, pp. 101–109, 1996.

[15] D. Kozen and M. Timme, "" Indefinite summation and the Kronecker delta," *https://hdl.handle.net/1813/8352*, 2007.

[16] C. Wagner and H. Hagras, "Toward general type-2 fuzzy logic systems based on zSlices," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 4, pp. 637–660, 2010.

[17] Q. Liang and J. Mendel, "Interval type-2 fuzzy logic systems: theory and design," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 5, pp. 535–550, 2000.

[18] C. Chen, D. Wu, J. M. Garibaldi, R. I. John, J. Twycross, and J. M. Mendel, "A Comprehensive Study of the Efficiency of Type-Reduction Algorithms," *IEEE Transactions on Fuzzy Systems*, vol. 29, pp. 1556–1566, 2021.

[19] M. Nie and W. W. Tan, "Towards an efficient type-reduction method for interval type-2 fuzzy logic systems," *2008 IEEE International Conference on Fuzzy Systems (IEEE World Congress on Computational Intelligence)*, pp. 1425–1432, 2008.

[20] S. Mirjalili, "Genetic algorithm," *In Evolutionary algorithms and neural networks*, pp. 43–55, 2019.

[21] D. P. Filev and I. Kolmanovsky, "Generalized markov models for real-time modeling of continuous systems," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 4, pp. 983–998, 2013.

[22] F. Herrera, "Genetic fuzzy systems: taxonomy, current research trends and prospects," *Evolutionary Intelligence*, vol. 1, no. 1, pp. 27–46, 2008.

[23] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.

[24] S. Fine, Y. Singer, and N. Tishby, "The hierarchical hidden Markov model: Analysis and applications," *Machine learning*, vol. 32, pp. 41–62, 1998.

[25] Y. Ben-Haim and E. Tom-Tov, "A Streaming Parallel Decision Tree Algorithm. Journal of Machine Learning Research," *Journal of Machine Learning Research*, vol. 11, 2010.

[26] P. Remy, "Temporal convolutional networks for keras," https://github.com/philipperemy/keras-tcn, 2020.

[27] I. Goodfellow, Y. Bengio, and A. Courville, "Deep learning," *MIT Press*, 2016.

[28] H. R. Rocha, I. H. Honorato, R. Fiorotti, W. C. Celeste, L. J. Silvestre, and J. A. Silva, "An artificial intelligence based scheduling algorithm for demand-side energy management in smart homes," *Applied Energy*, vol. 282, p. 116145, 2021.

[29] M. Kiani, J. Andreu-Perez, H. Hagras, S. Rigato, and M. L. Filippetti, "Towards Understanding Human Functional Brain Development with Explainable Artificial Intelligence: Challenges and Perspectives," *IEEE Computational Intelligence Magazine*, vol. 17, pp. 16–33, 2022.

[30] J. Andreu-Perez, L. L. Emberson, M. Kiani, M. L. Filippetti, H. Hagras, and S. Rigato, "Explainable Artificial Intelligence Based Analysis for Interpreting Infant fNIRS Data in Developmental Cognitive Neuroscience," *Communications Biology*, vol. 4, p. 1077, 2021.