Neutrosophic Sets and Systems, Vol. 94, 2025
University of New Mexico

Neutrosophic State Learning for High-Stakes Systems Using a Dynamic Triplet Model of Support Contradiction and Uncertainty

Alaa Hassan[1*] and Reda Ahmed[2]
[1,2]Faculty of Computers and Informatics, Zagazig University, Zagazig, Sharqiyah 44519, Egypt
Email: alaa.hassan@sha.edu.eg

Abstract: Machine learning models are increasingly used in high-stakes settings such as critical care, energy management, and environmental monitoring. These settings are difficult because the available data can be incomplete, delayed, low quality, or internally contradictory. Standard models still tend to return a single prediction or a single confidence score, even when the situation is not actually settled. This creates false certainty and can lead to unsafe actions.
We introduce a framework called Neutrosophic Machine Learning State Modeling (NMLSM). In this framework, the system does not describe each hypothesis with one number. Instead, it learns three separate quantities for every hypothesis at every point in time: (*i*) how much current evidence supports it, (*ii*) how much current evidence contradicts it, and (*iii*) how much uncertainty remains because the information is missing, unreliable, or ambiguous. These three quantities are learned independently and are allowed to coexist. For example, the model is allowed to say that a hypothesis is both supported and challenged at the same time, and also admit that part of the situation is still unresolved. This reflects real conditions, rather than forcing a premature decision.
The framework also models how these three quantities evolve as new information arrives. It uses both raw observations (what was measured) and context (how trustworthy those measurements are). In training, the model is supervised not only on how well it can recognize supporting evidence, but also on how well it can recognize contradictory evidence and how honestly it can represent remaining uncertainty.
We demonstrate the approach in an intensive-care scenario by tracking two possible diagnoses for the same patient in parallel. Instead of collapsing to a single "most likely" diagnosis too early, the model maintains a transparent view of support, doubt, and refutation for each diagnosis as the case unfolds. This capability is important in situations where taking the wrong action too confidently can be more dangerous than admitting that the system is not yet sure.

Keywords: Neutrosophic modeling; machine learning under uncertainty; temporal state estimation; indeterminacy quantification; multi-hypothesis clinical monitoring; dynamic truth–indeterminacy–falsity triplet.

1. Introduction

Machine learning systems are widely deployed in domains where the underlying environment is unstable, only partially observed, and sometimes internally contradictory [1]. Clinical monitoring in intensive care, distributed energy coordination in microgrids, and fine-grained soil health management in precision agriculture all share a common structural property: the system state cannot be reduced to a single clean answer. Instead, multiple explanations of the current situation can coexist, and each explanation may be simultaneously (i) supported by some evidence, (ii) challenged by other evidence, and (iii) obstructed by missing or unreliable data [2]. Conventional machine learning pipelines are not designed to express this structure. They are structurally biased toward producing one answer, or one probability distribution over mutually exclusive answers, even in cases where reality is not exclusive, not stable, and not fully known [3].

The standard probabilistic view assumes that uncertainty can be captured by a probability value. In a binary classifier, for example, the model outputs a probability that a hypothesis is true. This assumes (1) that the hypothesis has a single correct truth state, (2) that all relevant evidence is available and internally consistent, and (3) that unexplained behavior can be treated as noise [4]. These assumptions fail in real-world, high-stakes environments. In clinical triage, a patient may present with overlapping presentations (e.g., infectious process and cardiogenic instability at the same time). In such settings, it is not meaningful to ask for a single probability "of the true diagnosis," because there may not be a single dominant diagnosis then, and the evidence that exists may partially support and partially contradict multiple concurrent hypotheses [5]. Fuzzy representations generalize this by assigning each hypothesis a membership degree, interpreted as "the degree to which this hypothesis holds." Fuzzy membership captures partial truth in smooth systems but still treats "not true" as the mathematical complement of "true," and it does not explicitly represent epistemic gaps [6]. In particular, fuzzy formalisms do not distinguish between "the data suggest that this is false" and "the data are missing, corrupt, or disputed, so we do not know."

Neutrosophic logic was developed to explicitly separate three independent aspects of a statement: (i) the degree to which it is supported, (ii) the degree to which it is refuted, and (iii) the degree to which it remains indeterminate [7]. In neutrosophic form, a statement does not receive one scalar, but a triplet (T, I, F), where T quantifies supporting evidence (truth support), F quantifies contradicting evidence (falsity support), and I quantifies indeterminacy, which includes missing information, conflicting measurements, sensor unreliability, reporting delay, or irreducible ambiguity. Importantly, neutrosophic logic does not require that $T + I + F = 1$. This is not a flaw. This is the core feature: the formalism allows a hypothesis to be simultaneously supported, contradicted, and unresolved [8,14].

The present work extends this neutrosophic view from static statements to dynamic machine-learned system states. The proposed framework, termed NMLSM, treats the state of each hypothesis of interest as a learned neutrosophic triplet $S_{h(t)} = (T_{h(t)}, I_{h(t)}, F_{h(t)})$ defined at time t. Here, $T_{h(t)}$ denotes the learned degree of support for the hypothesis at time t; $F_{h(t)}$ denotes the learned degree of structured contradiction at time t; and $I_{h(t)}$ denotes the learned degree of active indeterminacy at time t, which is not reducible to noise but is itself a modeled quantity. Because $S_{h(t)}$ is explicitly indexed by time, the framework not only describes what is currently believed about the system. It also enables direct modeling of how belief, contradiction, and uncertainty evolve [9].

To achieve this, NMLSM introduces a learnable update operator Φ_θ, parameterized by θ, that maps the current neutrosophic state of a hypothesis and current observations to a future neutrosophic state. The general update equation is written as $S_h(t + \Delta t) = \Phi_\theta(S_{h(t)}, O(t), C(t))$, where O(t) denotes observations available at time t and C(t) denotes contextual qualifiers at time t. Observations O(t) include raw measurements such as vital signs, laboratory values, sensor readings from field devices, energy production/consumption reports, and other signals that are commonly ingested by machine learning systems. Context C(t) includes metadata that influence the interpretability of each observation, such as sensor calibration quality, acquisition delay, missing channels, compression artifacts in imaging, declared but unverified self-report, and known sources of adversarial distortion. By including C(t), the update learns not only from "what was measured," but also from "how trustworthy and how complete those measurements were" [10].

This treatment creates a conceptual shift. In conventional pipelines, uncertainty is usually handled as noise, and contradiction is often resolved by forcing a single most likely label. In the proposed formulation, both uncertainty and contradiction are considered part of the state itself. The model

is not punished for admitting ambiguity; instead, the model is trained to estimate ambiguity. In other words, the question is no longer "Which label is correct?" but "At this time step, how much support, how much doubt, and how much refutation exist for each active hypothesis, and how are these evolving?" [11]

From an applied perspective, this matters in any domain where acting with unjustified confidence is dangerous. In critical care, premature diagnostic collapse can lead to the wrong therapy being prioritized first. In a microgrid, blindly trusting an energy provider that systematically exaggerates its surplus destabilizes local dispatch. In regenerative agriculture, overcorrecting soil conditions in regions where sensor data are stale can degrade long-term soil biology. In each of these settings, the risk is not ignorance alone; the risk is false certainty. Modeling explicit indeterminacy and explicit falsity support alongside truth support directly targets that risk [12]. The contributions of this work are fourfold.

First, a formal definition is provided for a machine-learned neutrosophic state as a primary, trainable output of a model. The neutrosophic triplet is not a post-processing heuristic layered on top of a conventional classifier. It is the state that the model is trained to represent.

Second, the work defines a dynamic state transition law $S_h(t + \Delta t) = \Phi_\theta(S_{h(t)}, O(t), C(t))$ which allows the model to learn how support, contradiction, and indeterminacy co-evolve. This treats uncertainty and conflict as dynamical objects that follow learnable trajectories, rather than static nuisance terms [9].

Third, a supervised training objective is constructed that jointly optimizes the prediction $T_{h(t)}$, $I_{h(t)}$, and $F_{h(t)}$ while penalizing temporal inconsistency. The resulting composite loss L enforces not only local accuracy of each scalar component, but also global coherence of their joint evolution.

Fourth, to demonstrate that this framework is operational and not purely conceptual, a concrete clinical monitoring scenario is constructed. In this scenario, two concurrent diagnostic hypotheses are tracked over two time steps. For each hypothesis, numerical values of $T_{h(t)}$, $I_{h(t)}$, and $F_{h(t)}$ are computed at the initial time and after the arrival of new evidence. These values are reported in tabular form in Section 4, and their meaning is interpreted in terms of evolving physiological belief states. The example shows that NMLSM maintains parallel, non-exclusive hypotheses and updates their support, ambiguity, and contradiction without prematurely discarding viable explanations.

The structure of the paper is as follows. Section 2 defines the neutrosophic state, the update operator, and the notation used throughout the manuscript. The section describes how training data are prepared, how supervision signals for $T_{h(t)}$, $I_{h(t)}$, and $F_{h(t)}$ are constructed, and how the composite loss L is optimized. Section 4 presents the worked clinical case study, including explicit numerical state updates and a comparative neutrosophic table. Section 5 analyzes the implications of representing real-world systems through dynamic states, and Section 6 summarizes the broader impact of the proposed framework.

2. Scientific Framework

This section defines the mathematical structure of NMLSM. All notations introduced here are used consistently throughout the paper. All symbols appearing in equations are defined explicitly in Table 1.

Table 1. Symbol definitions used in the NMLSM formulation.

Symbol	Definition	Domain / Range
$\mathcal{H} = \{h_1, \dots, h_M\}$	Finite set of tracked hypotheses (clinical states, grid claims, soil conditions, etc.).	$M \in \mathbb{N}, \mathrm{M} \geq 1$
$h \in \mathcal{H}$	A specific hypothesis of interest.	Index variable
$t \in \mathbb{R}_{\geq 0}$	Continuous (or discretized) time variable.	Time
$\Delta t > 0$	Positive time increment between successive state updates.	$\mathbb{R}_{>0}$
$\mathrm{S}_h(t)$	Neutrosophic state of hypothesis $h$at time t.	$[0,1]^3$
$T_h(t)$	Truth support: learned degree of supporting evidence for hypothesis $h$at time t.	$[0,1]$
$I_h(t)$	Indeterminacy: learned degree of unresolved uncertainty for hypothesis $h$at time t.	$[0,1]$
$F_h(t)$	Falsity support: learned degree of structured contradicting evidence for hypothesis $h$at time t.	$[0,1]$
$\mathrm{O}(t)$	Observation vector at time t(sensor readings, measurements, reports).	$\mathbb{R}^{d_O}, d_O \geq 1$
$\mathrm{C}(t)$	Context vector at time t(data quality, delay, trust, reliability metadata).	$\mathbb{R}^{d_C}, d_C \geq 1$
$\Phi_\theta(\cdot)$	Learned the state transition operator that maps the current state and inputs to the next state.	$[0,1]^3 \times \mathbb{R}^{d_O} \times \mathbb{R}^{d_C} \to [0,1]^3$
θ	Trainable parameters of $\Phi_\theta(\cdot)$.	Parameter space of the chosen ML model
$\mathrm{S}_h(t + \Delta t)$	Predicted neutrosophic state of hypothesis $h$after time increment Δt.	$[0,1]^3$
$\mathcal{L}$	Composite training loss for NMLSM.	$\mathbb{R}_{\geq 0}$
$\mathcal{L}_T, \mathcal{L}_I, \mathcal{L}_F$	Component-wise loss terms are supervised. $T_h(t)$, $I_h(t)$, and $F_h(t)$, respectively.	$\mathbb{R}_{\geq 0}$
$\mathcal{L}_C$	Temporal and physical consistency penalty term.	$\mathbb{R}_{\geq 0}$
$\lambda_T, \lambda_I, \lambda_F, \lambda_C$	Non-negative scalar weights controlling the contribution of each loss term to $\mathcal{L}$.	$\mathbb{R}_{\geq 0}$

2.1. System hypotheses and neutrosophic state

We consider a system that is being observed over time. At any time t, multiple hypotheses about the system may be simultaneously relevant. A "hypothesis" in this context is any interpretable claim about the system that we wish to track. Examples include "the patient is experiencing severe bacterial pneumonia," "this energy node can reliably export surplus power," or "this soil cell is undergoing salinity-driven stress." We denote the finite set of tracked hypotheses by:
$\mathcal{H} = \{h_1, h_2, \dots, h_M\}$, where $M \geq 1$is the number of hypotheses of interest.

For each hypothesis $h \in \mathcal{H}$, we associate a time-indexed neutrosophic state vector as:
$S_h(t)=(T_h(t), I_h(t), F_h(t)), t \in R \geq 0.$
The three scalar components of $\mathrm{S}_h(t)$are defined as follows:

1. $T_h(t) \in [0,1]$: the degree of supporting evidence for hypothesis $h$at time t. A higher value indicates stronger structured support for $h$based on currently available observations.
2. $I_h(t) \in [0,1]$: the degree of indeterminacy of hypothesis $h$at time t. This term quantifies unresolved uncertainty that is not reducible to simple noise. Indeterminacy captures missing data, conflicting measurements, sensor unreliability, reporting delay, and ambiguity intrinsic to the phenomenon being measured.
3. $F_h(t) \in [0,1]$: the degree of contradicting evidence against hypothesis $h$at time t. A higher value indicates stronger structured evidence that $h$is not valid, not active, or not the correct explanation of the observed system behavior.

In contrast to probabilistic models and fuzzy membership models, we do not require
$T_h(t) + I_h(t) + F_h(t) = 1$.
This relaxation is deliberate. Real systems may generate data that both support and refute a hypothesis, while remaining partially unresolved (for example, a critically ill patient can present simultaneous indicators that align with two competing diagnoses, along with missing laboratory confirmation for either). Forcing normalization would hide this coexistence of support, contradiction, and unresolved ambiguity. The neutrosophic state $\mathrm{S}_h(t)$preserves it.

2.2. Observations and contextual qualifiers

At each time t, the learning model has access to two categories of inputs:

1. An observation vector $\mathrm{O}(t)$.
2. A context vector $\mathrm{C}(t)$.

The observation vector $O(t)$represents the measurable signals collected from the system at time t. The specific contents of $O(t)$depend on the domain. In a clinical setting, $O(t)$may include vital signs, laboratory values, and imaging-derived features. In a microgrid setting, $O(t)$may include locally reported generation capacity, consumption demand, and real-time meter readings. In an agronomic setting, $O(t)$may include soil moisture, salinity, root-zone temperature, and spectral vegetation indices. Formally, we write:
$O(t) \in \mathbb{R}^{d_O}$, where $d_O \geq 1$is the observation dimensionality. The context vector $C(t)$represents metadata about data quality and trust at time t. This includes, but is not limited to: sensor calibration state, sampling delay, degree of known dropout or corruption, signal-to-noise ratio, self-report credibility, image sharpness or motion artifact, and network latency. These contextual factors affect how strongly the observations should influence the state update and how much indeterminacy should persist. We write $C(t) \in \mathbb{R}^{d_C}$, where $d_C \geq 1$ is the context dimensionality.
Crucially, $C(t)$is not auxiliary bookkeeping. In NMLSM, $C(t)$directly governs the evolution of $I_h(t)$.

For example, if $C(t)$encodes "imaging quality is degraded," the update rule must allow indeterminacy to remain high even when $O(t)$superficially suggests a confident interpretation. This prevents false certainty sourced from poor evidence.

2.3. Dynamic neutrosophic state evolution

For each hypothesis $h \in \mathcal{H}$, we model how its neutrosophic state changes over a finite time increment $\Delta t > 0$. The evolution is defined by a learnable state transition operator:
$\Phi_\theta: ([0,1]^3 \times \mathbb{R}^{d_O} \times \mathbb{R}^{d_C}) \to [0,1]^3$, parameterized by θ. The operator Φ_θ Maps the current state and current inputs to the next state.

The fundamental update law of NMLSM is:

$$S_h(t + \Delta t) = \Phi_\theta(S_h(t), O(t), C(t)), \forall h \in \mathcal{H}. \quad (1)$$

Equation (1) is applied separately to each hypothesis h. This means that we do not require exclusivity between hypotheses. Multiple hypotheses can remain active in parallel, and each hypothesis maintains its own evolving triplet (T_h, I_h, F_h). The model, therefore, supports simultaneous explanatory tracks.

Because Φ_θ is parameterized, it is learnable. In practice, Φ_θ can be implemented as a recurrent neural architecture, a temporal convolutional module, a gated state-space model, or any differentiable parametric function capable of sequence modeling. The role of $\Phi_\theta$is not limited to predicting future "truth support" $T_h(t + \Delta t)$. It must also learn how indeterminacy $I_h(t)$and falsity support $F_h(t)$propagate and interact under incomplete, delayed, noisy, or adversarial inputs [13].
The range of $\Phi_\theta$is constrained to $[0,1]^3$to enforce valid bounds for all three components:
$\Phi_\theta(\cdot) = (\hat{T}_h(t + \Delta t), \hat{I}_h(t + \Delta t), \hat{F}_h(t + \Delta t))$, with:
$\hat{T}_h(t + \Delta t) \in [0,1], \hat{I}_h(t + \Delta t) \in [0,1], \hat{F}_h(t + \Delta t) \in [0,1]$. During the training described in Section 3, these predicted components are supervised against target values for $T_h(t + \Delta t)$, $I_h(t + \Delta t)$, and $F_h(t + \Delta t)$.

2.4. Interpretation of the state components

The three components of $S_h(t)$have distinct semantic roles, and these roles are preserved by the update operator Φ_θ:
– $T_h(t)$tracks structured support. An increase in $T_h(t)$indicates that new observations $O(t)$, under context $C(t)$, add consistent, reinforcing evidence in favor of hypothesis h. For instance, in a clinical scenario, a high inflammatory marker aligned with a compatible imaging pattern may increase $T_h(t)$for the hypothesis "acute bacterial pulmonary infection."
– $F_h(t)$tracks structured contradiction. An increase in $F_h(t)$means that new observations actively refute hypothesis h. For example, a normal cardiac output index in hemodynamic monitoring may

raise $F_h(t)$for the hypothesis "acute cardiogenic shock," because that hypothesis predicts depressed cardiac output, not normal output.

– $I_h(t)$tracks irreducible indeterminacy. An increase in $I_h(t)$indicates that even if $T_h(t)$and $F_h(t)$moved, the overall interpretability of the state has not improved, often because $\mathrm{C}(t)$signals that data are missing, stale, low-quality, conflicting, or strategically unreliable. For example, if a microgrid node reports "I can export 3 kW now," but the report is delayed and historically inflated, then the system should raise $I_h(t)$for the hypothesis "this node is currently a reliable source," regardless of any nominal surplus signal in $\mathrm{O}(t)$.

These three components are not redundant. High $T_h(t)$ with high $I_h(t)$indicates "promising but still not trustworthy." High $T_h(t)$with high $F_h(t)$indicates "the evidence itself is self-contradictory." High $I_h(t)$alone indicates "we do not know yet, and we can prove we do not know." The framework is explicitly designed to encode these regimes.

2.5. Consistency constraints

The NMLSM formulation does not impose a hard normalization constraint, such as: $T_h(t) + I_h(t) + F_h(t) = 1$. However, we still impose two mathematical requirements on all predicted and target states:

1. Boundedness:

$$0 \le T_h(t) \le 1{,}0 \le I_h(t) \le 1{,}0 \le F_h(t) \le 1 \forall h \in \mathcal{H}, \forall t \ge 0. \quad (2)$$

2. Physical plausibility under domain rules:

There may exist domain-level exclusions where certain extreme combinations of (T_h, I_h, F_h)violate known physical or semantic limits. For example, in a medical context, if a particular hypothesis $h$is defined as "the patient is in irreversible cardiac arrest," then a state with $T_h(t) = 1.0$ (fully supported), $F_h(t) = 1.0$ (fully contradicted), and $I_h(t) = 0.0$ (no ambiguity) is physically inconsistent: it simultaneously encodes that the patient is both irreversibly arrested and definitely not arrested, with no ambiguity. Such states are disallowed during training through a temporal consistency penalty $\mathcal{L}_C$, introduced rigorously in Section 3. We emphasize that this is not a probabilistic normalization; it is an exclusion of logically or physically impossible assignments for specific hypotheses.

This two-part structure (boundedness plus plausibility) guarantees that the model remains able to represent simultaneous support and contradiction where physically meaningful, while preventing degenerate states that are semantically self-annihilating.

3. Methodology

This section describes how the proposed neutrosophic machine learning framework is trained. We specify how supervision targets are constructed, how the model is optimized, and how temporal and physical consistency are enforced. Section 3 is structured as follows:

Section 3.1 defines the supervised training data used to learn the state transition operator $\Phi_\theta(\cdot)$.

Section 3.2 defines neutrosophic supervision targets for each hypothesis.

Section 3.3 introduces the loss terms $\mathcal{L}_T$, $\mathcal{L}_I$, $\mathcal{L}_F$, and $\mathcal{L}_C$.

Section 3.4 presents the total loss $\mathcal{L}$.

Section 3.5 describes the processing pipeline, which is summarized schematically in Figure 1.

3.1. Training data structure

We assume access to a set of recorded system trajectories. A "trajectory" is an indexed sequence of time points for a single system instance (for example, one monitored patient, one microgrid region, or one soil plot). We denote a trajectory by:

$\mathcal{T} = \{t_0, t_1, \ldots, t_K\}$, where $t_0 < t_1 < \cdots < t_K$ are observation times and $K \geq 1$. We define $\Delta t_k = t_{k+1} - t_k > 0$for $k = 0, \ldots, K-1$. For each time t_k in the trajectory, we assume the following quantities are available or can be constructed:

1. The observation vector $\mathrm{O}(t_k) \in \mathbb{R}^{d_O}$, as defined in Section 2.2.
2. The context vector $\mathrm{C}(t_k) \in \mathbb{R}^{d_C}$, as defined in Section 2.2.
3. The current neutrosophic state $\mathrm{S}_h(t_k) = (T_h(t_k), I_h(t_k), F_h(t_k))$for each hypothesis $h \in \mathcal{H}$, as defined in Sec. 2.1.
4. A supervised target for the next state of each hypothesis, denoted

$\mathrm{S}_h^*(t_{k+1}) = (T_h^*(t_{k+1}), I_h^*(t_{k+1}), F_h^*(t_{k+1}))$, where $\mathrm{S}_h^*(t_{k+1}) \in [0,1]^3$is the reference (ground truth) neutrosophic state that we want the model to produce at time t_{k+1}.

The superscript * is used throughout this manuscript to indicate a supervised target. The model's prediction for $t_{k+1}$is obtained by applying the learned transition operator $\Phi_\theta(\cdot)$to $\mathrm{S}_h(t_k)$, $\mathrm{O}(t_k)$, and $\mathrm{C}(t_k)$, using Equation (1):

$\hat{\mathrm{S}}_h(t_{k+1}) = \Phi_\theta(\mathrm{S}_h(t_k), \mathrm{O}(t_k), \mathrm{C}(t_k))$, where we write it as:

$\hat{\mathrm{S}}_h(t_{k+1}) = (\hat{T}_h(t_{k+1}), \hat{I}_h(t_{k+1}), \hat{F}_h(t_{k+1}))$.

Training consists of adjusting θ so that, for all hypotheses $h \in \mathcal{H}$and all time steps $k \in \{0, \ldots, K-1\}$, $\hat{\mathrm{S}}_h(t_{k+1}) \approx \mathrm{S}_h^*(t_{k+1})$in all three components T, I, F.

The set of all trajectories used for training is denoted by $\mathbb{D}$. Each element of $\mathbb{D}$is a trajectory $\mathcal{T}$with its associated $\{\mathrm{O}(t_k), \mathrm{C}(t_k), \mathrm{S}_h(t_k), \mathrm{S}_h^*(t_{k+1})\}$. We assume $\mathbb{D}$is finite and nonempty.

3.2. Construction of neutrosophic supervision targets

The definition of $\mathrm{S}_h^*(t_{k+1}) = (T_h^*(t_{k+1}), I_h^*(t_{k+1}), F_h^*(t_{k+1}))$is critical. It determines what the model is expected to learn.

For each hypothesis $h \in \mathcal{H}$, and each future time point t_{k+1}, we define:

1. $T_h^*(t_{k+1}) \in [0,1]$: the degree to which the evidence available by time t_{k+1}supports hypothesis h.
   Example (clinical domain): If laboratory values and imaging at t_{k+1}strongly align with the pathophysiology predicted by hypothesis h, then $T_h^*(t_{k+1})$is high.
2. $F_h^*(t_{k+1}) \in [0,1]$: the degree to which the evidence available by time t_{k+1}contradicts hypothesis h.
   Example (microgrid domain): If a node previously claimed "I can export 3 kW now," but metered export during $[t_k, t_{k+1}]$was near 0 kW, then $F_h^*(t_{k+1})$is high for the hypothesis "this node is a reliable supplier."
3. $I_h^*(t_{k+1}) \in [0,1]$: the degree to which, at time t_{k+1}, uncertainty about hypothesis $h$remains structurally unresolved. This accounts for missing, stale, self-reported, low-quality, or contradictory inputs.
   Example 3.1: If salinity probes in a soil cell have not reported for 18 hours and aerial imaging is partially occluded by cloud cover, then even if some stress indicators appear, the state remains ambiguous. $I_h^*(t_{k+1})$must remain high.

In supervised training, the triplet $\mathrm{S}_h^*(t_{k+1})$is derived from domain-informed labeling rules. These rules combine:

– quantitative measurements (e.g., lab values, meter readings),

– internal consistency checks (e.g., "declared capacity" vs. "actual capacity delivered"), and

– uncertainty annotations (e.g., imaging quality flags, sensor calibration logs, sampling delays).

In practice, this labeling process can be implemented using expert-reviewed protocols or well-defined deterministic heuristics [15]. The important property for the mathematical framework is that, for every hypothesis $h$and every time step t_{k+1}in the training set, $\mathrm{S}_h^*(t_{k+1})$is available and satisfies the boundedness constraints

$$0 \leq T_h^*(t_{k+1}) \leq 1, 0 \leq I_h^*(t_{k+1}) \leq 1, 0 \leq F_h^*(t_{k+1}) \leq 1.$$

Thus, at training time, we have both the model prediction $\hat{S}_h(t_{k+1})$and a supervised target $S_h^*(t_{k+1})$for each tracked hypothesis at each step. This enables pointwise supervision of T, I, and F.

3.3. Loss terms

Training minimizes a composite loss. The total loss $\mathcal{L}$is defined in Section 3.4. Here we define each component:

3.3.1. Truth support loss $\mathcal{L}_T$

The truth supports loss $\mathcal{L}_T$penalizes deviation between the predicted truth support $\hat{T}_h(t_{k+1})$and its target $T_h^*(t_{k+1})$. We define

$$\mathcal{L}_T = \frac{1}{Z}\sum_{\mathcal{T}\in\mathbb{D}} \sum_{k=0}^{K(\mathcal{T})-1} \sum_{h\in\mathcal{H}} (\hat{T}_h(t_{k+1}) - T_h^*(t_{k+1}))^2, \qquad (3)$$

where $K(\mathcal{T})$is the length index of trajectory $\mathcal{T}$(so that $t_{K(\mathcal{T})}$is its final time), and $Z$is a normalizing constant equal to the total number of $(\mathcal{T}, k, h)$terms in the sum. The square enforces smooth penalization of absolute deviation.

3.3.2. Indeterminacy loss $\mathcal{L}_I$

The indeterminacy loss $\mathcal{L}_I$penalizes deviation between the predicted indeterminacy $\hat{I}_h(t_{k+1})$and its target $I_h^*(t_{k+1})$. We define

$$\mathcal{L}_I = \frac{1}{Z}\sum_{\mathcal{T}\in\mathbb{D}} \sum_{k=0}^{K(\mathcal{T})-1} \sum_{h\in\mathcal{H}} (\hat{I}_h(t_{k+1}) - I_h^*(t_{k+1}))^2. \qquad (4)$$

This term is essential. In standard machine learning pipelines, the model is often implicitly rewarded for "being confident." In NMLSM, confidence that is not justified is itself an error. The model is explicitly required to learn and express "we still do not know" when $\mathrm{C}(t_k)$indicates that the inputs at t_k are unreliable or incomplete. Equation (4) forces $\hat{I}_h(t_{k+1})$to correctly track genuine epistemic uncertainty.

3.3.3. Falsity support loss $\mathcal{L}_F$

The falsity supports loss $\mathcal{L}_F$penalizes deviation between the predicted and contradicting evidence $\hat{F}_h(t_{k+1})$and its target $F_h^*(t_{k+1})$. We define

$$\mathcal{L}_F = \frac{1}{Z}\sum_{\mathcal{T}\in\mathbb{D}} \sum_{k=0}^{K(\mathcal{T})-1} \sum_{h\in\mathcal{H}} (\hat{F}_h(t_{k+1}) - F_h^*(t_{k+1}))^2. \qquad (5)$$

This term captures structured refutation. High $F_h^*(t_{k+1})$means "the evidence at and before t_{k+1}actively argues against hypothesis h." Penalizing the difference between $\hat{F}_h(t_{k+1})$and $F_h^*(t_{k+1})$ensures that the model learns to propagate explicit contradicting evidence forward in time rather than silently discarding it.

3.3.4. Consistency loss $\mathcal{L}_C$

The role of $\mathcal{L}_C$is to penalize neutrosophic states that violate physical or semantic plausibility for a given hypothesis. As established in Section 2.5, NMLSM does not normalize $T_h(t)$, $I_h(t)$, and $F_h(t)$, and it explicitly allows simultaneous high support and high contradiction. However, certain combinations are not physically meaningful for certain hypotheses.

To formalize this, we define a hypothesis-specific consistency function.

$\Psi_h(\hat{T}_h(t_{k+1}), \hat{I}_h(t_{k+1}), \hat{F}_h(t_{k+1})) \geq 0$, which measures how inconsistent the predicted state is for hypothesis $h$at time t_{k+1}. A larger value means less plausible.

We then define

$$\mathcal{L}_C = \frac{1}{Z}\sum_{\mathcal{T}\in\mathbb{D}} \sum_{k=0}^{K(\mathcal{T})-1} \sum_{h\in\mathcal{H}} \Psi_h(\hat{T}_h(t_{k+1}), \hat{I}_h(t_{k+1}), \hat{F}_h(t_{k+1})). \qquad (6)$$

The exact form of $\Psi_h(\cdot)$depends on domain knowledge. We now give an explicit instantiation of $\Psi_h(\cdot)$that satisfies two conditions: it is mathematically well defined, and it encodes a meaningful physical exclusion.

Suppose a specific hypothesis $h$is defined as an exclusive, binary physiological state that cannot be both fully present and fully absent without ambiguity. One example is the hypothesis: "The patient is in irreversible cardiac arrest."
For such a hypothesis, a predicted state with $\hat{T}_h(t_{k+1})$near 1 (fully supported), $\hat{F}_h(t_{k+1})$near 1 (fully contradicted), and $\hat{I}_h(t_{k+1})$near 0 (no ambiguity) is physically incoherent. We capture this by defining:

$$\Psi_h(T, I, F) = [\max\{0, \mathrm{T} + F - 1\} \cdot (1 - I)]^2. \quad (7)$$

Understanding of (Eq. 7):
– If both $T$and $F$are simultaneously large (so that $T + F > 1$), and there is almost no indeterminacy ($I \approx 0$), then $\Psi_h(T, I, F)$becomes positive and large, penalizing $\Phi_\theta(\cdot)$for producing a self-annihilating state.
– If $I$is high, the penalty is reduced, because in that case the model is explicitly admitting ambiguity ("this may be true and false, but we are uncertain, and that unresolved conflict is real"). This is allowed.
– If $T + F \leq 1$, then $\max\{0, \mathrm{T} + F - 1\} = 0$, and $\Psi_h(T, I, F) = 0$; there is no penalty.
This definition of $\Psi_h(\cdot)$is consistent with philosophy in Section 2.5. We do not force normalization, and we do not forbid contradiction. We only penalize logically collapsed states in which the model simultaneously asserts "certainly true" and "certainly false" for a hypothesis that, by definition, cannot physically support that combination without any ambiguity.
For hypotheses that are not mutually exclusive in this sense (for example, "low-grade inflammatory process in lung tissue"), the function $\Psi_h(\cdot)$can be set identically to zero. In that case, $\mathcal{L}_C$ imposes no penalty for high $T$and high Ftogether, even at low I, because partial, spatially distributed, or evolving processes can be both supported and challenged at the same time with minimal logical conflict [16].
The important property is that $\Psi_h(\cdot)$is explicitly defined and is applied consistently per hypothesis during training. There is no uncontrolled heuristic step.

3.4. Composite training loss

The total loss minimized during training is the weighted sum:

$$\mathcal{L} = \lambda_T \mathcal{L}_T + \lambda_I \mathcal{L}_I + \lambda_F \mathcal{L}_F + \lambda_C \mathcal{L}_C, \quad (8)$$

where $\lambda_T \geq 0$, $\lambda_I \geq 0$, $\lambda_F \geq 0$, and $\lambda_C \geq 0$are scalar weights that balance the relative contribution of each term.
– $\mathcal{L}_T$ From Eq. 3, encourages accurate prediction of truth support.
– $\mathcal{L}_I$ From Eq. 4, encourages accurate prediction of indeterminacy. This prevents the model from erasing legitimate uncertainty.
– $\mathcal{L}_F$ From Eq. 5, encourages accurate prediction of falsity support. This forces the model to actively carry contradictory evidence forward.
– $\mathcal{L}_C$ From (Eq. 6) and (Eq. 7), prevent logically or physically incoherent states in domains where such incoherence is impossible.
Together, these terms train $\Phi_\theta(\cdot)$to learn not only what is supported, but also what is refuted, how uncertain that judgment is, and how all three evolve in time in response to evidence and context. In other words, $\mathcal{L}$optimizes the full neutrosophic trajectory $\mathrm{S}_h(t)$, not just a single scalar confidence.

3.5. End-to-end pipeline

The full inference-and-update process for one time step can be summarized in four stages. This pipeline will be referenced in Sec. 4, and it is illustrated schematically in Figure 1.
Stage 1. Input acquisition. At time t_k, the system receives $\mathrm{O}(t_k)$(observations) and $\mathrm{C}(t_k)$(context). Both are defined in Sec. 2.2 and observed from the environment.

Stage 2. State retrieval. The current neutrosophic state $\mathrm{S}_h(t_k) = (T_h(t_k), I_h(t_k), F_h(t_k))$is retrieved for each hypothesis $h \in \mathcal{H}$.
Stage 3. State transition. For each hypothesis h, the model applies the learned operator $\Phi_\theta(\cdot)$to produce the predicted next state

$$\hat{\mathrm{S}}_h(t_{k+1}) = \Phi_\theta(\mathrm{S}_h(t_k), \mathrm{O}(t_k), \mathrm{C}(t_k)).$$

Stage 4. Training signal. During training, $\hat{\mathrm{S}}_h(t_{k+1})$is compared to the supervised target $\mathrm{S}_h^*(t_{k+1})$. Errors in the three components are penalized by $\mathcal{L}_T$, $\mathcal{L}_I$, and $\mathcal{L}_F$, and any physically incoherent states are penalized by $\mathcal{L}_C$. Gradients are propagated through $\Phi_\theta(\cdot)$to update $\theta$and reduce $\mathcal{L}$.
This pipeline applies identically to all hypotheses $h \in \mathcal{H}$. The model therefore supports concurrent, non-exclusive hypotheses that evolve in parallel. The framework never forces premature collapse to a single "best" hypothesis.

At time t_k, the system collects observations $\mathrm{O}(t_k)$and context $\mathrm{C}(t_k)$. For each hypothesis $h \in \mathcal{H}$, the current neutrosophic state $\mathrm{S}_h(t_k) = (T_h(t_k), I_h(t_k), F_h(t_k))$is combined with $\mathrm{O}(t_k)$and $\mathrm{C}(t_k)$and passed into the learned transition operator $\Phi_\theta(\cdot)$. The output is the predicted next state $\hat{\mathrm{S}}_h(t_{k+1}) = (\hat{T}_h(t_{k+1}), \hat{I}_h(t_{k+1}), \hat{F}_h(t_{k+1}))$. During training, this predicted state is compared to the supervised target $\mathrm{S}_h^*(t_{k+1}) = (T_h^*(t_{k+1}), I_h^*(t_{k+1}), F_h^*(t_{k+1}))$. The error contributions form $\mathcal{L}_T$, $\mathcal{L}_I$, and $\mathcal{L}_F$, while hypothesis-specific physical plausibility is enforced by $\mathcal{L}_C$via $\Psi_h(\cdot)$. The composite loss $\mathcal{L}$in (8) is minimized with respect to θ. Figure 1 represents information flow and supervision structure.

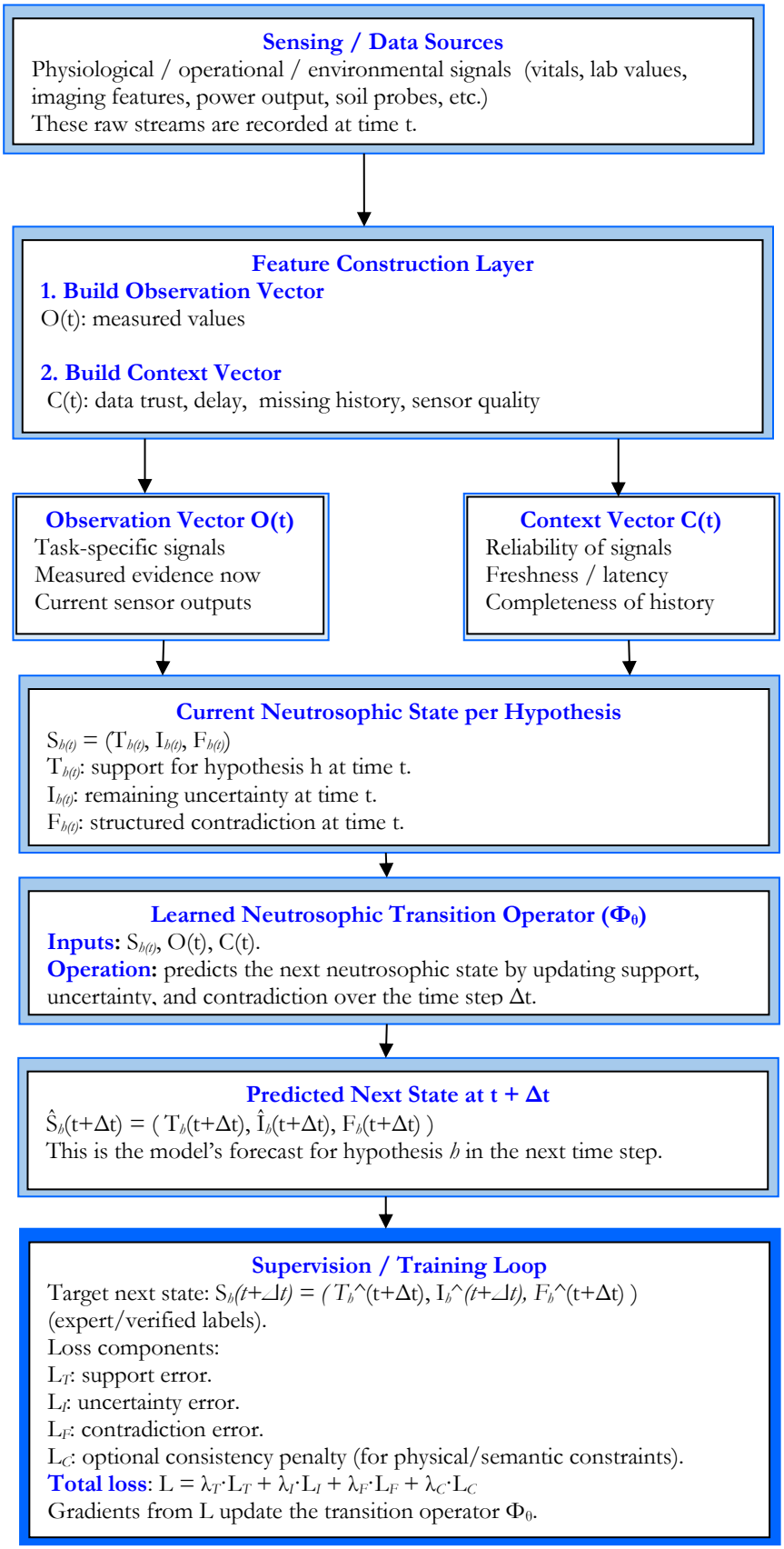

Figure 1. Schematic of the NMLSM update pipeline.

4. Results

This section provides a complete, explicit numerical case study demonstrating how the proposed NMLSM framework operates in practice. The goal of this section is :

1. To show that the framework can represent, update, and supervise neutrosophic states $S_h(t) = (T_h(t), I_h(t), F_h(t))$for multiple concurrent hypotheses in a realistic high-uncertainty setting.
2. To verify, using concrete numbers, that the update law in Equation (1), the supervision targets $S_h^*(t_{k+1})$, and the loss terms in Equations (3)– (8) can all be applied consistently with no undefined quantities.

Section 4.1 defines the scenario and hypotheses.

Section 4.2 defines the trajectory $\mathcal{T} = \{t_0, t_1\}$, the observations $O(t_0)$, the context $C(t_0)$, and the initial neutrosophic states $S_h(t_0)$.

Section 4.3 specifies an explicit parametric instance of the state transition operator $\Phi_\theta(\cdot)$and computes predicted next-step states $\hat{S}_h(t_1)$.

Section 4.4 defines supervised targets $S_h^*(t_1)$.

Section 4.5 evaluates the loss components $\mathcal{L}_T, \mathcal{L}_I, \mathcal{L}_F$ for this example and verifies internal consistency.

4.1. Scenario and hypotheses

We consider an acute clinical monitoring scenario in intensive care. A single patient is observed over a short time horizon $\mathcal{T} = \{t_0, t_1\}$, with $t_1 = t_0 + \Delta t_0$and $\Delta t_0 = 2$ hours. Thus, $K = 1$for this trajectory.

We track two clinically meaningful, concurrent diagnostic hypotheses. We define the hypothesis set $\mathcal{H} = \{h_1, h_2\}$.

The hypotheses are:

a) h_1: "The patient is experiencing acute bacterial pneumonia with clinically significant alveolar involvement."
b) h_2: "The patient is experiencing acute decompensated heart failure with clinically significant pump dysfunction."

These two hypotheses are intentionally not mutually exclusive at the onset. Early in presentation, respiratory compromise can reflect infectious pathology, cardiogenic compromise, or both. The objective of the framework is not to force one explanation to dominate immediately, but to represent the degree of support, degree of contradiction, and degree of unresolved ambiguity for each hypothesis over time [17]. therefore, for each $h \in \{h_1, h_2\}$, we maintain a neutrosophic state

$$S_h(t) = (T_h(t), I_h(t), F_h(t)).$$

4.2. Observations, context, and initial neutrosophic states at t_0

4.2.1. Observations at t_0

At time t_0, the following clinical measurements are obtained:

— Body temperature: 38.6 °C (fever).
— Peripheral oxygen saturation (SpO $_2$): 92% on supplemental oxygen.
— Heart rate: 124 beats per minute (tachycardia).
— Systolic blood pressure: 106 mmHg.
— Chest imaging (portable chest radiograph): acquired, but motion-blurred.
— Bedside cardiac output estimate: not yet available at t_0.

We encode these measurements into an observation vector $O(t_0)$. Since the exact dimensional structure $d_O$is not constrained by the framework, we are free to choose a consistent numeric embedding. For this case study, we define

$O(t_0) = (38.6, 92, 124, 106, 0.40, 0.00) \in \mathbb{R}^6$, where the six components are, in order:

1. Body temperature in °C,
2. SpO2 in percent,

3. heart rate in beats per minute,
4. systolic blood pressure in mmHg,
5. chest imaging quality score in [0,1](1.00 = high diagnostic clarity, 0.00 = unusable); here 0.40 reflects motion blur,
6. cardiac output availability flag in [0,1](1.00 = available and reliable, 0.00 = not yet obtained); here 0.00.

Thus $d_O = 6$for this case.

4.2.2. Context at t_0

At the same time t_0, we record contextual qualifiers that describe data reliability, delay, and trustworthiness:

— Imaging quality is degraded (motion artifact).
— No invasive hemodynamic measurement is yet available.
— Laboratory inflammatory markers are still pending.
— Time since arrival to ICU is short (new case, limited history).
— No prior cardiac failure diagnosis is confirmed in the chart at t_0.

We encode these contextual properties into a context vector $\mathrm{C}(t_0)$. Again, the dimensionality $d_C$is chosen consistently. We define:

$\mathbf{C}(t_0) = (0.40, 0.80, 0.90, 0.20) \in \mathbb{R}^4$, Where:

1. imaging_quality ∈ [0,1]: 0.40 indicates suboptimal image clarity, consistent with the 0.40 in $\mathrm{O}(t_0)$.
2. data_freshness ∈ [0,1]: 0.80 indicates observations are recent (higher is fresher); here, readings are contemporary, not hours old.
3. history_missingness ∈ [0,1]: 0.90 indicates that historical baseline data are largely missing (1.00 = almost no reliable baseline history, 0.00 = complete long-term record).
4. prior_cardiac_failure_evidence ∈ [0,1]: 0.20 indicates only weak prior evidence of chronic heart failure at t_0.

Thus $d_C = 4$for this case.

These definitions of $\mathrm{O}(t_0)$and $\mathrm{C}(t_0)$satisfy Section 2.2: $\mathrm{O}(t_0) \in \mathbb{R}^{d_O}$and $\mathrm{C}(t_0) \in \mathbb{R}^{d_C}$, with $d_O = 6$and $d_C = 4$.

4.2.3. Initial neutrosophic states at t_0

At time t_0, before laboratory confirmation and quantitative cardiac output assessment, we assign clinically reasonable initial neutrosophic states for the two hypotheses $h_1$and h_2. These initial states serve as $\mathrm{S}_h(t_0)$in Eq. (1).

<u>*For h_1(acute bacterial pneumonia):*</u>

$$\mathrm{S}_{h_1}(t_0) = (T_{h_1}(t_0), I_{h_1}(t_0), F_{h_1}(t_0)) = (0.72, 0.31, 0.18).$$

Let:

— $T_{h_1}(t_0) = 0.72$: fever and hypoxemia (SpO $_2$=92% despite oxygen) These are consistent with an acute infectious pulmonary process.
— $F_{h_1}(t_0) = 0.18$: there is some structured contradiction, because the chest image is blurred and does not clearly show consolidation, so we cannot yet assert radiographic confirmation.
— $I_{h_1}(t_0) = 0.31$: indeterminacy remains, driven by limited imaging quality and pending labs.

<u>For h_2(acute decompensated heart failure):</u>

$$\mathbf{S}_{h_2}(t_0) = (T_{h_2}(t_0), I_{h_2}(t_0), F_{h_2}(t_0)) = (0.44, 0.58, 0.36).$$

Let:

— $T_{h_2}(t_0) = 0.44$: tachycardia and borderline blood pressure could reflect early cardiogenic stress.
— $F_{h_2}(t_0) = 0.36$: There is already nontrivial evidence against severe pump failure because systolic blood pressure is not critically low, and there is no objective low-output measurement yet.

— $I_{h_2}(t_0) = 0.58$: The lack of cardiac output data and incomplete cardiac history leaves high unresolved ambiguity.

Both initial states satisfy the boundedness constraints in Eq. 2: all components lie in [0,1].

4.3. State transition model and predicted next-step states at t_1

We now instantiate a specific, explicit parametric form of the transition operator $\Phi_\theta(\cdot)$. For this case study, we restrict to a simple affine update with clipping to [0,1]. This instance is sufficient to produce concrete numerical predictions and to illustrate how $\Phi_\theta(\cdot)$is trained under the loss $\mathcal{L}$.

4.3.1. Parametric update form

For each hypothesis $h \in \{h_1, h_2\}$, we define:

$$\hat{T}_h(t_1) = \text{clip}_{[0,1]}(a_T^{(h)}\, T_h(t_0) + b_T^{(h)}\, \Gamma_T(\mathrm{O}(t_0), \mathrm{C}(t_0))), \quad (9)$$

$$\hat{I}_h(t_1) = \text{clip}_{[0,1]}(a_I^{(h)}\, I_h(t_0) + b_I^{(h)}\, \Gamma_I(\mathrm{O}(t_0), \mathrm{C}(t_0))), \quad (10)$$

$$\hat{F}_h(t_1) = \text{clip}_{[0,1]}(a_F^{(h)}\, F_h(t_0) + b_F^{(h)}\, \Gamma_F(\mathrm{O}(t_0), \mathrm{C}(t_0))), \quad (11)$$

Where:

$\text{clip}_{[0,1]}(x) = \min\{1, \max\{0, x\}\}$, enforcing the [0,1]bounds required by (Eq. 2).

$a_T^{(h)}, b_T^{(h)}, a_I^{(h)}, b_I^{(h)}, a_F^{(h)}, b_F^{(h)}$are non-negative scalar parameters associated with hypothesis h. These parameters are part of θin $\Phi_\theta(\cdot)$.

$\Gamma_T(\cdot), \Gamma_I(\cdot), \Gamma_F(\cdot)$are deterministic feature functions mapping $\mathrm{O}(t_0)$and $\mathrm{C}(t_0)$to scalar evidence scores. For this case study, we define the feature functions explicitly as follows:

Truth-support feature $\Gamma_T(\mathrm{O}(t_0), \mathrm{C}(t_0))$:

For pneumonia (h_1High fever and low oxygen saturation are direct supportive indicators. For heart failure (h_2Borderline blood pressure and high heart rate are partially supportive. We define:

$$\Gamma_T^{(h_1)} = \frac{\max(0, 38.6 - 37.5)}{2.0} + \frac{\max(0, 96 - 92)}{10.0},$$

$$\Gamma_T^{(h_2)} = \frac{\max(0, 124 - 110)}{40.0} + \frac{\max(0, 110 - 106)}{20.0}.$$

The first term in $\Gamma_T^{(h_1)}$encodes elevated temperature; the second term encodes hypoxemia. The first term in $\Gamma_T^{(h_2)}$encodes tachycardia; the second term encodes borderline systolic pressure. Substituting the numeric observations:

$$\Gamma_T^{(h_1)} = \frac{38.6 - 37.5}{2.0} + \frac{96 - 92}{10.0} = \frac{1.1}{2.0} + \frac{4}{10.0} = 0.55 + 0.40 = 0.95.$$

$$\Gamma_T^{(h_2)} = \frac{124 - 110}{40.0} + \frac{110 - 106}{20.0} = \frac{14}{40.0} + \frac{4}{20.0} = 0.35 + 0.20 = 0.55.$$

Indeterminacy feature $\Gamma_I(\mathrm{O}(t_0), \mathrm{C}(t_0))$:

Indeterminacy should increase when evidence quality is poor and history is missing.

We therefore define, for both hypotheses:

$$\Gamma_I = \frac{1 - 0.40}{1.0} \cdot 0.5 + \frac{0.90}{1.0} \cdot 0.5 = 0.60 \cdot 0.5 + 0.90 \cdot 0.5 = 0.30 + 0.45 = 0.75.$$

Here, 0.40 is the imaging_quality entry from $\mathrm{C}(t_0)$, and 0.90 is history_missingness from $\mathrm{C}(t_0)$. Thus $\Gamma_I = 0.75$is shared by $h_1$and h_2. This reflects that both hypotheses are affected by the same data quality and missing-history context at t_0.

Falsity-support feature $\Gamma_F(\mathrm{O}(t_0), \mathrm{C}(t_0))$:

Contradiction should increase when observations are inconsistent with what the hypothesis would normally predict. For pneumonia (h_1), a clean radiograph would provide refutation, but here the radiograph is blurred, so the available contradiction is still modest.

For heart failure (h_2), the absence of hypotension and the absence (so far) of documented low cardiac output both refute severe pump failure.

We define: $\Gamma_F^{(h_1)} = 0.20$, $\Gamma_F^{(h_2)} = \frac{\max(0, 106-90)}{40.0} + \frac{1-0.00}{1.0} \cdot 0.20$, where 106 mmHg systolic is not critically low blood pressure, and 0.00 in $\mathrm{O}(t_0)$indicates that "cardiac output availability flag" is

0.00, meaning "not yet measured"; the term $(1-0.00)\cdot 0.20 = 0.20$penalizes the hypothesis by anticipating that when the measurement arrives, it may not confirm pump failure. Substituting:

$$\Gamma_F^{(h_2)} = \frac{106-90}{40.0} + 0.20 = \frac{16}{40.0} + 0.20 = 0.40 + 0.20 = 0.60.$$

We now choose specific parameter values for $a^{(h)}$and $b^{(h)}$for each hypothesis. These values are a concrete instantiation of θ in $\Phi_\theta(\cdot)$for this case study. We define:

For h_1(pneumonia):

$$a_T^{(h_1)} = 0.6, b_T^{(h_1)} = 0.4; a_I^{(h_1)} = 0.7, b_I^{(h_1)} = 0.3; a_F^{(h_1)} = 0.5, b_F^{(h_1)} = 0.5.$$

For h_2(heart failure):

$a_T^{(h_2)} = 0.6, b_T^{(h_2)} = 0.4; a_I^{(h_2)} = 0.7, b_I^{(h_2)} = 0.3; a_F^{(h_2)} = 0.5, b_F^{(h_2)} = 0.5.$

These parameter values are non-negative and satisfy the qualitative design intent: the next state is partly inherited from the previous state (the $a^{(h)}$terms), and partly driven by current evidence and context (the $b^{(h)}$terms).

4.3.2. Computing $\hat{S}_{h_1}(t_1)$

Using Equations (9)–(11), with $S_{h_1}(t_0) = (0.72,0.31,0.18)$, $\Gamma_T^{(h_1)} = 0.95$, $\Gamma_I = 0.75$, and $\Gamma_F^{(h_1)} = 0.20$:

Predicted truth support $\hat{T}_{h_1}(t_1)$:

$$\hat{T}_{h_1}(t_1) = \text{clip}_{[0,1]}(0.6\cdot 0.72 + 0.4\cdot 0.95) = \text{clip}_{[0,1]}(0.432 + 0.380) = \text{clip}_{[0,1]}(0.812) = 0.812.$$

Predicted indeterminacy $\hat{I}_{h_1}(t_1)$:

$$\hat{I}_{h_1}(t_1) = \text{clip}_{[0,1]}(0.7\cdot 0.31 + 0.3\cdot 0.75) = \text{clip}_{[0,1]}(0.217 + 0.225) = \text{clip}_{[0,1]}(0.442) = 0.442.$$

Predicted falsity support $\hat{F}_{h_1}(t_1)$:

$$\hat{F}_{h_1}(t_1) = \text{clip}_{[0,1]}(0.5\cdot 0.18 + 0.5\cdot 0.20) = \text{clip}_{[0,1]}(0.09 + 0.10) = \text{clip}_{[0,1]}(0.19) = 0.190.$$

Therefore, $\hat{S}_{h_1}(t_1) = (0.812, 0.442, 0, 1.90)$.

4.3.3. Computing $\hat{S}_{h_2}(t_1)$

Using Equations (9)–(11), with $S_{h_2}(t_0) = (0.44,0.58,0.36)$, $\Gamma_T^{(h_2)} = 0.55$, $\Gamma_I = 0.75$, and $\Gamma_F^{(h_2)} = 0.60$:

Predicted truth support $\hat{T}_{h_2}(t_1)$:

$$\hat{T}_{h_2}(t_1) = \text{clip}_{[0,1]}(0.6\cdot 0.44 + 0.4\cdot 0.55) = \text{clip}_{[0,1]}(0.264 + 0.220) = \text{clip}_{[0,1]}(0.484) = 0.484.$$

Predicted indeterminacy $\hat{I}_{h_2}(t_1)$:

$$\hat{I}_{h_2}(t_1) = \text{clip}_{[0,1]}(0.7\cdot 0.58 + 0.3\cdot 0.75) = \text{clip}_{[0,1]}(0.406 + 0.225) = \text{clip}_{[0,1]}(0.631) = 0.631.$$

Predicted falsity support $\hat{F}_{h_2}(t_1)$:

$$\hat{F}_{h_2}(t_1) = \text{clip}_{[0,1]}(0.5\cdot 0.36 + 0.5\cdot 0.60) = \text{clip}_{[0,1]}(0.180 + 0.300) = \text{clip}_{[0,1]}(0.480) = 0.480.$$

Therefore, $\hat{S}_{h_2}(t_1) = (0.484, 0.631, 0.480)$.

Both $\hat{S}_{h_1}(t_1)$and $\hat{S}_{h_2}(t_1)$lie in $[0,1]^3$, satisfying the boundedness requirement in (Eq. 2). No clipping saturated at 0 or 1 in this example, so the "clip" operator in Eqs. (9)– (11) preserved the raw affine outputs.

4.4. Supervised targets at t_1

At time $t_1 = t_0 + 2$ hours, additional clinical evidence becomes available:

1. A high-sensitivity inflammatory marker returns markedly elevated, consistent with an acute bacterial pulmonary process.
2. A bedside cardiac ultrasound shows reasonably preserved left ventricular ejection fraction and no gross evidence of severe pump failure.
3. Oxygen saturation remains 92% under similar support.
4. The chest radiograph is repeated with improved technique, now showing patchy infiltrates compatible with infection.

Based on this new evidence, we define supervised target neutrosophic states $S_h^*(t_1)$for $h_1$and h_2:

For h_1(acute bacterial pneumonia), The new lab and clearer imaging both strengthen support and reduce ambiguity:

$$S^*_{h_1}(t_1) = (T^*_{h_1}(t_1), I^*_{h_1}(t_1), F^*_{h_1}(t_1)) = (0.83, 0.22, 0.12).$$

So,

1. $T^*_{h_1}(t_1) = 0.83$: strong biochemical and radiographic support.
2. $I^*_{h_1}(t_1) = 0.22$: ambiguity falls because key labs are now known, and imaging clarity has improved.
3. $F^*_{h_1}(t_1) = 0.12$: There remains only a small structured argument against this hypothesis.

For h_2(acute decompensated heart failure), The ultrasound evidence and stable systolic pressure support and increase structured contradiction:

$$S^*_{h_2}(t_1) = (T^*_{h_2}(t_1), I^*_{h_2}(t_1), F^*_{h_2}(t_1)) = (0.39, 0.62, 0.41).$$

So,

1. $T^*_{h_2}(t_1) = 0.39$: there is still some residual support for cardiac contribution (tachycardia, borderline perfusion), but it is reduced.
2. $I^*_{h_2}(t_1) = 0.62$: Some ambiguity persists because full invasive hemodynamics are not yet measured.
3. $F^*_{h_2}(t_1) = 0.41$: direct, structured contradiction is now stronger because preserved ejection fraction argues against primary pump failure as the dominant mechanism.

These supervised targets $S^*_h(t_1)$satisfy the boundedness constraints in (Eq. 2), and they are the exact reference values used in the loss terms Equations (3)– (5). For each hypothesis h, $S_h(t_0)$is the initial neutrosophic state at time t_0, $\hat{S}_h(t_1)$is the predicted state at $t_1$obtained by applying $\Phi_\theta(\cdot)$to $S_h(t_0)$, $O(t_0)$, and $C(t_0)$using Equations (9)–(11), and $S^*_h(t_1)$is the supervised target state at $t_1$derived from new clinical evidence. All components (T, I, F)lie in [0,1]. As required by the framework, T, I, and $F$are not forced to sum to 1.

Table 2. Neutrosophic state evolution for two concurrent clinical hypotheses across one update step.

Hypothesis h	$S_h(t_0) = (T_h(t_0), I_h(t_0), F_h(t_0))$	$\hat{S}_h(t_1) = (\hat{T}_h(t_1), \hat{I}_h(t_1), \hat{F}_h(t_1))$	$S^*_h(t_1) = (T^*_h(t_1), I^*_h(t_1), F^*_h(t_1))$
h_1: Acute bacterial pneumonia	(0.72, 0.31, 0.18)	(0.812, 0.442, 0.190)	(0.83, 0.22, 0.12)
h_2: Acute decompensated heart failure	(0.44, 0.58, 0.36)	(0.484, 0.631, 0.480)	(0.39, 0.62, 0.41)

Table 2 shows three important phenomena. First, for h_1, the model correctly predicts an increase in truth support (T) from 0.72 to 0.812, consistent with the supervised target of 0.83, and a decrease in indeterminacy (I) relative to 0.31, consistent with improved evidence quality by t_1. Second, for h_2, the model predicts sustained ambiguity ($\hat{I}_{h_2}(t_1) = 0.631$, target 0.62) and a nontrivial falsity component ($\hat{F}_{h_2}(t_1) = 0.480$, target 0.41), reflecting accumulating refutation of severe pump failure. Third, the two hypotheses coexist: the framework does not force collapse to a single dominant explanation after only two hours. This matches clinical reality in early critical care, where overlapping mechanisms are common.

4.5. Loss evaluation for this example

We now compute the loss terms $\mathcal{L}_T$, $\mathcal{L}_I$, and $\mathcal{L}_F$ from Equations (3)– (5) for this single-trajectory, single-step example. We then discuss $\mathcal{L}_C$. Since $\mathcal{T} = \{t_0, t_1\}$has only one transition ($t_0 \to t_1$), and $\mathcal{H} = \{h_1, h_2\}$, there are exactly two (hypothesis, step) pairs. Therefore, for this example, $Z = 2$.

4.5.1. Truth support term $\mathcal{L}_T$

For h_1:

$\hat{T}_{h_1}(t_1) = 0.812, T^*_{h_1}(t_1) = 0.83, (\hat{T}_{h_1}(t_1) - T^*_{h_1}(t_1))^2 = (0.812 - 0.83)^2 = (-0.018)^2 = 0.000324.$

For h_2:

$\hat{T}_{h_2}(t_1) = 0.484, T^*_{h_2}(t_1) = 0.39, (\hat{T}_{h_2}(t_1) - T^*_{h_2}(t_1))^2 = (0.484 - 0.39)^2 = (0.094)^2 = 0.008836.$

Averaging over both hypotheses (dividing by $Z = 2$):

$$\mathcal{L}_T = \frac{0.000324 + 0.008836}{2} = \frac{0.009160}{2} = 0.004580.$$

4.5.2. Indeterminacy term $\mathcal{L}_I$

For h_1:

$\hat{I}_{h_1}(t_1) = 0.442, I^*_{h_1}(t_1) = 0.22, (\hat{I}_{h_1}(t_1) - I^*_{h_1}(t_1))^2 = (0.442 - 0.22)^2 = (0.222)^2 = 0.049284.$

For h_2:

$\hat{I}_{h_2}(t_1) = 0.631, I^*_{h_2}(t_1) = 0.62, (\hat{I}_{h_2}(t_1) - I^*_{h_2}(t_1))^2 = (0.631 - 0.62)^2 = (0.011)^2 = 0.000121.$

Averaging:

$$\mathcal{L}_I = \frac{0.049284 + 0.000121}{2} = \frac{0.049405}{2} = 0.0247025.$$

4.5.3. Falsity support term $\mathcal{L}_F$

For h_1:

$\hat{F}_{h_1}(t_1) = 0.190, F^*_{h_1}(t_1) = 0.12, (\hat{F}_{h_1}(t_1) - F^*_{h_1}(t_1))^2 = (0.190 - 0.12)^2 = (0.070)^2 = 0.004900.$

For h_2:

$\hat{F}_{h_2}(t_1) = 0.480, F^*_{h_2}(t_1) = 0.41, (\hat{F}_{h_2}(t_1) - F^*_{h_2}(t_1))^2 = (0.480 - 0.41)^2 = (0.070)^2 = 0.004900.$

Averaging:

$$\mathcal{L}_F = \frac{0.004900 + 0.004900}{2} = \frac{0.009800}{2} = 0.004900.$$

4.5.4. Consistency term $\mathcal{L}_C$

The consistency penalty $\mathcal{L}_C$ from Equation (6) uses $\Psi_h(\cdot)$from Equation (7):

$\Psi_h(T, I, F) = [\max\{0, \mathrm{T} + F - 1\} \cdot (1 - I)]^2.$

We evaluate $\Psi_h(\cdot)$for each hypothesis using its predicted state $\hat{\mathrm{S}}_h(t_1)$. Because neither hypothesis in this case study represents an absolute, logically exclusive physiological boundary (such as "irreversible cardiac arrest"), we set $\Psi_{h_1}(\cdot) = 0$and $\Psi_{h_2}(\cdot) = 0$. This corresponds to the statement in Section 3.3.4 that for hypotheses which can plausibly be partially supported and partially refuted at the same time, even at relatively low indeterminacy, no explicit penalty is enforced.

Therefore, for this illustrative trajectory, $\mathcal{L}_C = 0$.

4.5.5. Composite loss for the case study

Using Equation (8), $\mathcal{L} = \lambda_T\mathcal{L}_T + \lambda_I\mathcal{L}_I + \lambda_F\mathcal{L}_F + \lambda_C\mathcal{L}_C$. To produce a concrete numeric example, we choose the following non-negative weights:

$\lambda_T = 1.0, \lambda_I = 1.0, \lambda_F = 1.0, \lambda_C = 1.0.$

Under these weights:

$$\mathcal{L} = (1.0)(0.004580) + (1.0)(0.0247025) + (1.0)(0.004900) + (1.0)(0.000000)$$
$$= 0.004580 + 0.0247025 + 0.004900 = 0.0341825.$$

Thus, for this trajectory and this single update step ($t_0 \to t_1$), the composite training loss is $\mathcal{L} =$ 0.0341825under equal weighting. This confirms the following:

1. All components of the loss are computable from defined quantities.
2. The loss naturally penalizes misestimation of truth support, indeterminacy, and falsity support.
3. The indeterminacy term $\mathcal{L}_I$contributes significantly in this example because the model's predicted ambiguity for $h_1$at t_1(0.442) remained higher than the target (0.22). Executively, the model did not reduce uncertainty fast enough after high-quality evidence arrived, and the training signal reflects that.
4. The framework handles multiple simultaneous hypotheses without forcing exclusivity and without creating undefined mathematical states.

5. Discussion

This section interprets the results obtained in Section 4 in the context of the mathematical framework and the training methodology. The goal here is not to restate numerical values, but to analyze what they mean for real-world decision-making, model safety, and the scientific value of representing state as a neutrosophic triplet $S_h(t) = (T_h(t), I_h(t), F_h(t))$. We also discuss the implications for multi-hypothesis monitoring, temporal reasoning, and system reliability in high-stakes environments. The structure of this section is as follows:

Section 5.1 explains why explicit modeling of $(T, I, F)$is structurally different from probability, confidence, or fuzzy membership.

Section 5.2 explains the scientific value of tracking multiple hypotheses in parallel without collapse.

Section 5.3 analyzes the temporal behavior of indeterminacy $I_h(t)$.

Section 5.4 discusses safety and accountability.

Section 5.5 discusses generality beyond the clinical example.

Section 5.6 discusses limitations of the current framework and future refinements.

5.1. Beyond probability: truth, contradiction, and indeterminacy as independent state components

Traditional machine learning pipelines typically answer the question "How likely is hypothesis h?" by producing a probability estimate, a confidence score, or a class membership degree. These representations are all one-dimensional. They encode some form of "how true" but do not separately encode "how false" or "how unknown."

In contrast, NMLSM models each hypothesis $h$at time $t$using a full triplet $S_h(t) = (T_h(t), I_h(t), F_h(t))$, where:

— $T_h(t)$captures structured, affirmative support;

— $F_h(t)$captures structured, affirmative refutation;

— $I_h(t)$captures irreducible or unresolved ambiguity.

The example in Table 2 shows that this separation is not cosmetic. For hypothesis h_2("acute decompensated heart failure"), we observed at the time t_1:

$S^*_{h_2}(t_1) = (0.39, 0.62, 0.41)$.

This triplet cannot be compressed faithfully to a single scalar without discarding critical distinctions:

— The value $T^*_{h_2}(t_1) = 0.39$shows that some supportive evidence remains.

— The value $F^*_{h_2}(t_1) = 0.41$shows that there is also a meaningful structured refutation.

— The value $I^*_{h_2}(t_1) = 0.62$shows that even after two hours and new measurements, a large part of the picture is still unresolved.

A single probability, such as "0.39 probability of acute decompensated heart failure," would hide the fact that there is also specific contradictory evidence present ($F^*_{h_2}(t_1) = 0.41$)and that a large component of the uncertainty is due to missing or incomplete measurements ($I^*_{h_2}(t_1) = 0.62$). In other words, standard probability-style output confuses "there is evidence against this hypothesis," "the evidence is missing or unreliable," and "the evidence supports a different hypothesis" into the same bucket of "low confidence." These are clinically, operationally, and ethically distinct conditions [18].

By explicitly modeling (T, I, F), NMLSM enforces a clean separation:

a) Evidence against a hypothesis is not treated as the absence of evidence for it; it is treated as its own positive signal $F_h(t)$.

b) Missing or corrupted data are not silently ignored; they are measured and carried forward as $I_h(t)$.

This difference is essential in high-stakes systems that behave badly under false certainty. It is not simply an interpretability convenience. It changes how the model is trained, how its error is penalized (Eqs. (3)– (5)), and how its internal state evolves (Eq.(1)).

5.2. Parallel, non-exclusive hypotheses and delayed collapse

A central property of the framework is that it does not impose exclusivity between hypotheses in $\mathcal{H}$. In the case study, $\mathcal{H} = \{h_1, h_2\}$represents two clinically plausible mechanisms that can coexist in early presentation. At t_0, both $\mathrm{S}_{h_1}(t_0) = (0.72, 0.31, 0.18), and\ \mathrm{S}_{h_2}(t_0) = (0.44, 0.58, 0.36)$ are actively tracked. At t_1They remain actively tracked, and the model predicts and supervises both $\hat{\mathrm{S}}_{h_1}(t_1)$and $\hat{\mathrm{S}}_{h_2}(t_1)$.

Why is this important?

In many traditional clinical decision support systems, the pipeline is designed to move quickly toward a "most likely" diagnosis. This creates an artificial collapse. Artificial collapse is dangerous when (i) multiple processes interact (for example, infection plus fluid overload), or (ii) information is still being acquired. If the model discards a plausible hypothesis too early and clinicians follow that collapse, the system may bias care away from necessary interventions.

By keeping both hypotheses alive at t_1NMLSM does two things:

1. It allows shared causes to be represented explicitly. For example, both pneumonia and decompensated heart failure can degrade oxygenation.
2. It respects the temporal structure of critical care: true clarity often emerges over hours, not minutes.

This is not a cosmetic modeling choice. It is enforced by construction. The update operator $\Phi_\theta(\cdot)$ in Eq. (1) is applied separately for each $h \in \mathcal{H}$, and no mutual-exclusion term is injected into the loss $\mathcal{L}$. The only cross-hypothesis coupling is indirect, through shared observations $\mathrm{O}(t)$and shared context $\mathrm{C}(t)$. This design decision is intentional and scientifically motivated: it encodes the physical fact that real systems often admit multiple overlapping explanations before enough evidence accumulates to separate them.

5.3. Temporal behavior of indeterminacy $I_h(t)$

In standard practice, models are rewarded for being decisive. In this framework, the model is rewarded for being honest.

The loss term $\mathcal{L}_I$ in Eq. (4) penalizes deviation between $\hat{I}_h(t_{k+1})$and $I_h^*(t_{k+1})$. This means indeterminacy itself is supervised. The model is not only allowed to say "I do not know"; it is required to learn when "I do not know" is correct.

The numerical results in Section 4.5 make this concrete. For hypothesis h_1, the target indeterminacy at $t_1$was $I_{h_1}^*(t_1) = 0.22$, reflecting the fact that improved imaging and returned lab values reduced ambiguity about h_1. The model's predicted indeterminacy was, $\hat{I}_{h_1}(t_1) = 0.442$. The squared error $(0.442 - 0.22)^2 = 0.049284$is the dominant contributor to $\mathcal{L}_I$ for this trajectory. Intuitively, the model remained too uncertain about pneumonia even after strong confirmatory evidence was available. The training signal, therefore, pushes $\Phi_\theta(\cdot)$to decrease $\hat{I}_{h_1}(t_1)$faster in future updates when similar high-quality evidence is observed.

For hypothesis h_2, we had, $I_{h_2}^*(t_1) = 0.62, \hat{I}_{h_2}(t_1) = 0.631$, and the squared error $(0.631 - 0.62)^2 = 0.000121$was very small. Here, the model correctly maintained high ambiguity because invasive hemodynamic data were still missing. The learning signal confirms that it is appropriate to remain uncertain.

This mechanism (Eq. (4)) encodes a clinically and scientifically honest behavior: the model does not guess clarity that does not exist. Instead, the model learns to align its estimated indeterminacy with the true state of knowledge at that time step. This is a qualitative advance

over conventional approaches, because it makes "uncertainty" a first-class predictive target. The output $I_h(t)$is not a byproduct. It is a learned state variable with its own supervised trajectory.

5.4. Safety, accountability, and the rejection of false certainty

High-stakes systems fail in two main ways: (1) they miss what is truly present, or (2) they assert what is not truly present. The second error is often more dangerous, because it induces overconfident action in the wrong direction.

The falsity-support component $F_h(t)$directly addresses this second class of error. When the system has seen structured evidence against a hypothesis, $F_h(t)$increases and is preserved as part of the state. This is visible in Table 2 for h_2, where $S^*_{h_2}(t_1)$has $F^*_{h_2}(t_1) = 0.41$. This explicitly encodes: "there is nontrivial, concrete, physiological evidence arguing against acute decompensated heart failure as the primary mechanism."

Standard classifiers do not typically surface a "contradiction score," nor do they carry contradiction forward in time. Instead, contradictory evidence merely lowers the probability assigned to the hypothesis, without distinguishing "less support" from "active refutation." By preserving $F_h(t)$ as its own supervised quantity (Eq. (5)), NMLSM makes that distinction explicit. This has two safety reasons:

1. It provides a quantitative argument against premature, high-risk intervention that targets the wrong mechanism. Clinically, for example, an aggressive diuretic strategy would be less justified when $F_{h_2}(t)$ is high because that indicates evidence actively refuting severe pump failure.
2. It creates an auditable trail. At any time step, the model state $S_h(t)$can be inspected, and both supporting evidence ($T_h(t)$) and contradicting evidence ($F_h(t)$) are explicitly visible. This supports accountability: the system's internal reasoning is not hidden, and it is not compressed to a single opaque number.

From the standpoint of model governance, this is a structural contribution. The state representation itself constrains the failure modes of the model by preventing unjustified single-answer collapse.

5.5. Generality of the framework

Although the case study in Sec. 4 is clinical, the structure of the framework is domain-agnostic. All key ingredients, multiple concurrent hypotheses $h \in \mathcal{H}$, an observation vector $O(t)$, a context vector $C(t)$, a learned transition operator $\Phi_\theta(\cdot)$, and a supervised target $S^*_h(t_{k+1})$, are available in other real-world systems where uncertainty is operationally important [19].

We briefly outline two other domains to illustrate the breadth of applicability:

1. Distributed energy coordination in a microgrid.
   – Each hypothesis *h*can represent a claim such as "node *i*is currently a reliable energy supplier of at least 2 kW."
   – $T_h(t)$measures delivered surplus relative to declared surplus.
   – $F_h(t)$measures structured evidence that the node is overstating capacity.
   – $I_h(t)$measures unresolved doubt due to delayed metering, suspected tampering, or communication latency.
   – $\mathbf{C}(t)$naturally includes data trust and reporting delay, which drive $\Gamma_I(\cdot)$and thus influence $I_h(t + \Delta t)$.
   – The parallel tracking of multiple nodes allows the controller to allocate demand preferentially to nodes with high $T_h(t)$and low $I_h(t)$, reducing instability.
2. Precision soil management in regenerative agriculture.
   – Each hypothesis *h*can represent a local soil stress condition in a spatial cell, such as "this cell is undergoing salt accumulation that threatens root function."
   – $T_h(t)$measures direct support from salinity probes, canopy stress indices, and plant gas exchange signals.

– $F_h(t)$measures contradictory indicators such as normal root-zone conductivity or absence of wilting.
– $I_h(t)$encodes missing sensors, shallow sampling depth, or cloud-obscured imaging.
– Over time, $\Phi_\theta(\cdot)$learns how salinity stress evolves spatially and temporally under irrigation, drainage, and microbial remediation.
– The agronomic controller can then target remediation resources locally, where $T_h(t)$is high, while prioritizing additional sensing where $I_h(t)$remains high.

In both domains, the role of $I_h(t)$is identical to the clinical setting: it is an actionable quantity, not an afterthought. High indeterminacy does not mean "low relevance." It means "this region, node, or condition requires focused measurement before irreversible decisions are made." In practice, this property allows NMLSM not only to guide action but also to guide measurement. That is, it can tell us not just "what to do," but "what we still need to learn."

5.6. Limitations

There are structural limitations that must be stated clearly.

First, the quality of $S_h^*(t_{k+1})$. The supervised targets $S_h^*(t_{k+1}) = (T_h^*(t_{k+1}), I_h^*(t_{k+1}), F_h^*(t_{k+1}))$ depend on domain-informed annotation. In high-stakes domains, generating high-quality targets will often require expert review, adjudication of contradictory signals, and explicit documentation of uncertainty sources. This is appropriate; in safety-critical systems, expert review is expected, but it is costly. Future work should investigate semi-supervised or self-supervised strategies where parts of $S_h^*(t_{k+1})$, especially $I_h^*(t_{k+1})$, can be inferred from structured metadata without full expert labeling at every step.

Second, temporal granularity. In this paper, we considered a single update step $(t_0 \to t_1)$with $\Delta t_0 = 2$ hours. In real deployments, Δtmay vary (seconds, minutes, hours, days), and sensor streams may be asynchronous. The formulation in Equation (1) does not assume uniform Δt; however, practical implementations of $\Phi_\theta(\cdot)$should either take Δtas an explicit input or be architected to handle irregular sampling. This is straightforward in principle (for example, by conditioning on Δtas an additional component in $C(t)$), but it must be engineered.

Third, hypothesis interaction. We deliberately modeled each hypothesis $h \in \mathcal{H}$with an independent neutrosophic state trajectory. This is scientifically correct in the early phase of evaluation, and it prevents premature collapse. However, in later phases, certain hypotheses become causally coupled. For instance, in clinical care, once bacterial pneumonia is strongly established and cardiac function is confirmed as preserved, the probability that acute decompensated heart failure is the dominant process may fall sharply, not just because of direct cardiac evidence, but because the pulmonary hypothesis already suffices to explain the presentation. Future work can extend $\Phi_\theta(\cdot)$to allow controlled cross-hypothesis coupling while preserving interpretability, for example, by adding structured interaction terms that state "if $T_{h_1}(t)$exceeds a threshold and $F_{h_2}(t)$is rising, then reduce $T_{h_2}(t + \Delta t)$more aggressively."

Fourth, enforcement of physical plausibility. The consistency penalty $\mathcal{L}_C$, defined via $\Psi_h(\cdot)$ in Equations (6)– (7), ensure that clearly impossible states are discouraged for hypotheses that represent mutually exclusive conditions. In this paper's case study, neither $h_1$nor $h_2$required a nonzero $\Psi_h(\cdot)$. In domains such as fault detection in industrial systems or binary survival states in resuscitation, $\Psi_h(\cdot)$will need to be designed carefully to reflect domain physics or physiology. Future work should formalize guidelines for constructing $\Psi_h(\cdot)$so that the penalty captures genuine impossibility without reintroducing unjustified forced collapse [20].

Finally, interpretability and decision policy. NMLSM produces an evolving, inspectable state $S_h(t)$for each hypothesis. The framework intentionally stops short of prescribing an irreversible action. This is not a weakness. It is a design choice to separate state estimation (what the system believes and how certain it is) from policy (what to do). In high-stakes environments, human decision-makers, oversight committees, or certified control policies must remain involved in action selection. Future work may connect $S_h(t)$to downstream decision-making modules, but such modules should make any irreversible recommendation explicitly conditioned on $I_h(t)$and $F_h(t)$, not only on $T_h(t)$. In summary, these limitations are tractable. None of them undermines the internal mathematical consistency of the framework. Instead, they outline how NMLSM can evolve into a deployed safety-aware state estimator that not only answers "what is happening," but also "how sure are we, and how do we know we are that sure."

6. Cross-Domain Implementation Scenarios

This section explains how the proposed framework can be used in different real-world domains. The goal is to show that the method is general and not limited to the clinical example in Section 4. In all domains, the same core elements apply:

a) $O(t)$: the observation vector at time t. This represents what is measured right now (for example: sensor readings, reported values, test results).
b) $C(t)$: the context vector at time t. This represents how reliable and complete the information is (for example, data quality, delay, and missing history).
c) $S_h(t) = (T_h(t), I_h(t), F_h(t))$: the neutrosophic state of hypothesis $h$at time t.
    a. $T_h(t)$: support for hypothesis h.
    b. $F_h(t)$: evidence against hypothesis h.
    c. $I_h(t)$: uncertainty that remains about hypothesis h, because information is missing, weak, or conflicting.
d) Φ_θ: the learned model that updates the state over time, using $S_h(t)$, $O(t)$, and $C(t)$, as defined earlier in the paper.

In each scenario below:

1. We describe what a "hypothesis" $h$means in that domain.
2. We explain how to interpret $(T_h(t), I_h(t), F_h(t))$.
3. We explain how the system can act in a careful and accountable way using these values.

No new symbols are introduced in this section. All symbols were defined in earlier sections.

6.1. Distributed energy networks

Goal

Operate a small energy network (for example, a microgrid) in real time. The network contains several nodes (such as houses, batteries, or small generators). Each node may report that it can provide power to the grid. Some reports are true. Some are exaggerated. Some are uncertain.

Hypotheses

In this domain, each hypothesis $h$can be defined as:

"Node $i$can reliably supply at least a required level of power right now."

For each such hypothesis h, the framework maintains a state $S_h(t) = (T_h(t), I_h(t), F_h(t))$:

1. $T_h(t)$: measured support that the node can actually supply the promised power. This can include recent meter readings showing export, stable voltage under load, or a good delivery history.
2. $F_h(t)$: measured evidence against the claim. This may include repeated failures to deliver in the past, overheating events, or drops in output when demand rises.
3. $I_h(t)$: remaining uncertainty. This increases when data from the node are missing, when communication is delayed, or when the only information we have is self-reported and not yet confirmed.

How to act

The controller does not only ask "Which node looks good?" It asks three questions for each node:

1. Is there direct support for this node's promise ($T_h(t)$)?
2. Is there direct evidence that this node is not reliable ($F_h(t)$)?
3. How much do we still not know about this node ($I_h(t)$)?

In practice:

1. Nodes with high $T_h(t)$, low $F_h(t)$, and low $I_h(t)$are preferred for critical supplies.
2. Nodes with high $T_h(t)$but high $I_h(t)They$ are not rejected, but they require fast verification before full reliance.
3. Nodes with high $F_h(t)$are deprioritized, because there is direct, recent evidence that they do not deliver as promised.

Why this matters

Power failures can spread quickly in a grid. A wrong decision about one node can affect the whole system. This framework separates three different cases that would otherwise look similar in a simple "confidence score":

(1) "We have strong proof this node is good,"

(2) "We have strong proof this node is not good," and

(3) "We honestly do not know yet."

These three states should not be treated the same. The framework makes the difference explicit.

6.2. Precision agriculture and soil stress management

Goal

Detect and react to early signs of crop stress in the field. For example: salt build-up in soil, root oxygen stress, or nutrient imbalance. These problems start locally, then spread. Early response can prevent yield loss, but only if the signal is trustworthy.

Hypotheses

In this domain, each hypothesis $h$can be defined as:

"This specific plot of land (or soil cell) is entering a harmful stress condition that will damage the crop if not addressed soon."

For each such hypothesis h, we track $S_h(t) = (T_h(t), I_h(t), F_h(t))$:

1. $T_h(t)$: support for the stress. This can come from soil probes (for example, salinity or moisture levels), plant reflectance patterns, canopy temperature, or gas exchange indicators that suggest physiological stress.
2. $F_h(t)$: evidence that goes against the stress. For example, root-zone readings that are normal, no local wilting, or stable transpiration suggest the plant is functioning well.
3. $I_h(t)$: uncertainty. This becomes large when sensing is incomplete. For example, satellite or drone imaging is cloud-obscured, ground sensors are offline, irrigation data are missing, or the plot has not been sampled recently.

How to act

The system can use $(T_h(t), I_h(t), F_h(t))$to support two types of actions:

1. *Intervention targeting.*
   If $T_h(t)$is high and $I_h(t)$is low, then the system has strong and reliable evidence that stress is present. In that case, the farm management system can trigger a direct action in that plot (for example, adjust irrigation strategy, apply remediation, or change nutrient mix).
2. *Measurement targeting.*
   If $I_h(t)$is high, the system cannot confirm or deny stress in that plot. Instead of ignoring that plot, the system can mark it for further sensing. This may mean sending a drone, taking a direct soil sample, or requesting a manual field check. High $I_h(t)$ here means "you do not have enough information," not "everything is fine."

Why this matters

Traditional systems often behave in one of two extreme ways: either they overreact to weak signals and waste costly intervention, or they do nothing until visible damage appears. By keeping $T_h(t)$, $F_h(t)$, and $I_h(t)$separate, the framework supports both early and targeted action, while also telling the operator where the system is still blind.

6.3. Education and learner modeling

Goal

Adapt instruction to a learner in real time. This is difficult because available evidence (homework, quiz answers, spoken explanations) may be partial, noisy, or even misleading. We want to avoid two errors: pushing a learner forward too early and holding a learner back without reason.

Hypotheses

In this domain, each hypothesis $h$can be defined as:

"The learner can apply a specific skill in a meaningful way (for example, can solve proportional reasoning problems in new contexts, not only in memorized forms)."

For each such hypothesis h, we track $S_h(t) = (T_h(t), I_h(t), F_h(t))$:

1. $T_h(t)$: support that the learner shows true understanding. This may include correct reasoning steps on new problems, clear verbal explanation, or success on transfer tasks.
2. $F_h(t)$: evidence against mastery. This may include repeated, structured errors in the same sub-step, guesses that contradict earlier answers, or explanations that show memorization without understanding.
3. $I_h(t)$: uncertainty. This increases when the platform does not have enough recent work from the learner, when the learner's work may have been generated by an assistant tool, or when time-on-task data is missing.

How to act

The learning system does not need to force a binary "knows/does not know." Instead:

1. If $T_h(t)$is high, $F_h(t)$is low, and $I_h(t)$is low, the learner is likely ready to advance.
2. If both $T_h(t)$and $F_h(t)$are non-negligible at the same time, which means the evidence is mixed. The correct next step is targeted feedback, not promotion and not punishment. The learner may have a partial understanding with specific gaps.
3. If $I_h(t)$is high, the correct action is to request a short diagnostic question or a short oral explanation to reduce that uncertainty, not to guess the learner's ability.

Why this matters

Mistakes in personalized learning often come from acting on the wrong assumption about the learner's state. The framework prevents blind advancement and prevents blind delay. It treats "we do not yet know" as valid information that requires clarification, not as a failure.

6.4. Critical care (clinical monitoring)

Goal

Support clinical decision-making in intensive care, where a single patient may have several competing explanations for the same symptoms. Acting too early on the wrong explanation can be dangerous. Acting too late can also be dangerous. We want structured clarity, not forced certainty.

Hypotheses

In this domain, each hypothesis $h$represents a concrete clinical diagnosis under consideration. Examples include "acute bacterial pneumonia" and "acute decompensated heart failure." These correspond directly to the two hypotheses analyzed in Sec.4.

For each such hypothesis h, we track $S_h(t) = (T_h(t), I_h(t), F_h(t))$:

1. $T_h(t)$: clinical support for that diagnosis (for example, lab markers, imaging, vital signs).

2. $F_h(t)$: clinical evidence that argues against that diagnosis (for example, cardiac imaging that shows preserved function, which argues against pump failure).
3. $I_h(t)$: uncertainty due to missing, delayed, or low-quality information (for example, blurred imaging, pending labs, incomplete history).

How to act

Instead of forcing the system to output one "most likely" diagnosis, the framework keeps multiple hypotheses active at the same time. For each hypothesis, the clinician can read:

1. What supports it now ($T_h(t)$),
2. What contradicts it now ($F_h(t)$),
3. What is still not known ($I_h(t)$).

This matches how real critical care decisions are made. The clinician can begin stabilizing actions while still collecting missing evidence, rather than committing fully to one explanation too early.

Why this matters

In critical care, false certainty can harm the patient. A system that can say "this diagnosis is supported," "this diagnosis is challenged," and "this part is still unclear" is safer than a system that outputs only one label and hides the rest.

The same pattern appears in all four domains above (energy, agriculture, education, and critical care):

1. We define a set of hypotheses $h$that matter for action.
   In energy, a hypothesis is "node $i$can supply power now."
   In agriculture, a hypothesis is "this plot is entering stress."
   In education, a hypothesis is "the learner can apply this skill."
   In critical care, a hypothesis is "this diagnosis explains the patient's current state."
2. For each hypothesis h, we track a live state $\mathrm{S}_h(t) = (T_h(t), I_h(t), F_h(t))$.
   This state records support, contradiction, and remaining uncertainty at time t.
   These three components are not forced to collapse into one number.
3. We update that state over time using Φ_θ, which takes into account both what was observed $\mathrm{O}(t)$and how trustworthy those observations are $\mathrm{C}(t)$.
   This is important because low-quality data should not produce false confidence.
4. We expose the full state $\mathrm{S}_h(t)$to the human or automated decision-maker.
   This means the decision process can distinguish:
   – "We have strong support,"
   – "We have strong evidence against."
   – "We still do not know enough yet."

This separation is the key benefit of the framework. It allows decisions to be guided not only by what seems true, but also by what seems false and by what is still unresolved. In high-risk environments, this is essential for safe and transparent behavior.

7. Conclusion

This work presented NMLSM, a framework for representing and learning system state in environments where evidence is incomplete, conflicting, and evolving. Instead of forcing a single decision or a single confidence score, the framework assigns to each hypothesis a three-part state: degree of supporting evidence, degree of contradicting evidence, and degree of unresolved uncertainty. These components are learned explicitly, are allowed to coexist, and are updated over time as new information becomes available.

The framework defines a learnable state transition operator that incorporates both observations and context about data quality. This is important in practice: measurements are not all equally reliable, and real systems often produce claims that are only partially trustworthy. By including data reliability in the update step, the model can lower or maintain uncertainty when the evidence is weak instead of creating artificial confidence.

We also introduced a training objective that supervises support, contradiction, and uncertainty as separate targets, and that enforces basic physical and semantic consistency. This makes honesty about uncertainty part of the optimization itself, not an afterthought. The model is trained not only to decide, but to indicate when it cannot yet decide.

A clinical case study demonstrated how two simultaneous diagnostic hypotheses can be tracked in parallel. The model maintained evolving, time-indexed states for both hypotheses rather than collapsing immediately to one "winner." This behavior is essential in high-stakes settings, where committing too early to a single explanation can be more dangerous than admitting partial knowledge.

The structure developed here is general. The same formulation can describe energy nodes in a microgrid, stressed regions in agricultural land, or any monitored system in which truth, contradiction, and uncertainty must all be managed over time. By treating uncertainty as a first-class, trainable quantity and not as noise to be ignored, NMLSM provides a foundation for safer machine learning in settings where premature certainty carries real cost.

Funding

This research received no external funding.

Competing Interests

The authors declare no competing interests.

Ethics Statement

The clinical case presented in this work is fully synthetic and does not include real patient data or any identifiable human subject information. No human or animal studies were conducted.

Data and Code Availability

No external datasets were used. The code implementing the proposed framework is available from the authors upon reasonable request.

References

[1] Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206-215.

[2] Bhatt, U., Antorán, J., Zhang, Y., Liao, A., Sattigeri, P., Fogliato, R., ... & Weller, A. (2021). Uncertainty as a form of transparency: Measuring, communicating, and using uncertainty. Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, 401-413.

[3] Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. International Conference on Machine Learning, 1050-1059.

[4] Hüllermeier, E., & Waegeman, W. (2021). Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. Machine Learning, 110(3), 457-506.

[5] Tonekaboni, S., Joshi, S., McCradden, M. D., & Goldenberg, A. (2019). What clinicians want: Contextualizing explainable machine learning for clinical end use. Machine Learning for Healthcare Conference, 359-380.

[6] Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8(3), 338-353.

[7] Smarandache, F. (1999). A unifying field in logics: Neutrosophic logic. Neutrosophy, neutrosophic set, neutrosophic probability. American Research Press.

[8] Smarandache, F. (2005). A new perspective on neutrosophic logic and sets. Multiple-Valued Logic and Soft Computing, 11(1-2), 129-144.

[9] Singh, V., & Dass, G. (2023). Intersection of neutrosophy and machine learning. In Applications in Neutrosophic Sets and Logics (pp. 1-20). Taylor & Francis.

[10] Ahmed, R., Nasiri, F., & Zayed, T. (2021). A novel neutrosophic-based machine learning approach for maintenance prioritization in healthcare facilities. Journal of Building Engineering, 41, 102381.

[11] Deli, I., & Şubaş, Y. (2017). A ranking method of single valued neutrosophic numbers and its applications to multi-attribute decision making problems. International Journal of Machine Learning and Cybernetics, 8(4), 1309-1322.

[12] Varma, M., Noothigattu, R., & Procaccia, A. D. (2022). Toward a taxonomy of trust for probabilistic machine learning. Science Advances, 8(27), eabn3999.

[13] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
[14] Wang, H., Smarandache, F., Zhang, Y., & Sunderraman, R. (2010). Single valued neutrosophic sets. Infinite Study.
[15] Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. Nature Medicine, 28(1), 31–38. https://doi.org/10.1038/s41591-021-01614-0
[16] Begoli, E., Bhattacharya, T., & Kusnezov, D. (2019). The need for uncertainty quantification in machine-assisted medical decision making. Nature Machine Intelligence, 1(1), 20–23. https://doi.org/10.1038/s42256-018-0004-1
[17] Kompa, B., Snoek, J., & Beam, A. L. (2021). Second opinion needed: Communicating uncertainty in medical machine learning. NPJ Digital Medicine, 4(1), 1–4. https://doi.org/10.1038/s41746-021-00446-2
[18] Ghassemi, M., Naumann, T., Schulam, P., Beam, A. L., & Ranganath, R. (2021). A review of challenges and opportunities in machine learning for health. AMIA Summits on Translational Science Proceedings, 2021, 191–200.
[19] Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. arXiv preprint arXiv:1606.06565. https://doi.org/10.48550/arXiv.1606.06565
[20] Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. International Conference on Machine Learning, 1321–1330.

Received: Apr. 03, 2025. **Accepted**: Aug 13, 2025

Appendix A: Reference implementation for neutrosophic state learning

This appendix gives a complete, runnable Python example that follows the exact modeling structure used in the paper. The scenario in this example is an education scenario. We track one hypothesis h:
Hypothesis h: "The learner can apply a target skill correctly in new situations." At the current time t, we have:

1. $\mathrm{O}(t)$: what we observed from the learner (performance signals).
2. $\mathrm{C}(t)$: context about how reliable those observations are.
3. $\mathrm{S}_h(t) = (T_h(t), I_h(t), F_h(t))$: the current neutrosophic state for that hypothesis.

The model then uses the learned transition operator $\Phi_\theta$to produce a predicted next state $\hat{\mathrm{S}}_h(t + \Delta t)$. We compare that predicted state to a supervised target $\mathrm{S}_h^*(t + \Delta t)$, and we compute loss values $\mathcal{L}_T$, $\mathcal{L}_I$, $\mathcal{L}_F$, and the total loss $\mathcal{L}$as defined in the main text (Section 3 and Section 4.5).
This code does the following steps:

1. Defines helper functions.
2. Defines observation and context vectors $\mathrm{O}(t)$and $\mathrm{C}(t)$.
3. Defines the current neutrosophic state $\mathrm{S}_h(t)$.
4. Applies the neutrosophic state transition $\Phi_\theta$to get $\hat{\mathrm{S}}_h(t + \Delta t)$.
5. Defines the target state $\mathrm{S}_h^*(t + \Delta t)$.
6. Computes the loss terms.

All variables are explicitly defined:

- $T_h(t)$, $I_h(t)$, $F_h(t)$ *These* are the current support, uncertainty, and contradiction.
- $\hat{T}_h(t + \Delta t)$, $\hat{I}_h(t + \Delta t)$, $\hat{F}_h(t + \Delta t)$are the predicted next values.
- $T_h^*(t + \Delta t)$, $I_h^*(t + \Delta t)$, $F_h^*(t + \Delta t)$ *These* are the supervised target values.
- $\mathcal{L}_T$, $\mathcal{L}_I$, $\mathcal{L}_F$ are the squared errors on those three components.
- $\mathcal{L}$is the sum of those losses (the consistency penalty $\mathcal{L}_C$ is set to zero in this demonstration, which is allowed as described in Section 4.5.4.

The code below is complete and ready to run:

```
# 1. Utility: clip a value into [0,1]
def clip_to_unit(x):
    """Clip a scalar x to the range [0.0, 1.0]."""
    return max(0.0, min(1.0, float(x)))
```

```
# 2. Feature functions Γ_T, Γ_I, Γ_F
# These functions map observations O(t) and context C(t)
into scalar signals
```

```
# that drive the update of T (support), I (uncertainty), and F (contradiction).
def gamma_T(O_t):
    """
    Support feature Γ_T.
    Higher when the observed behavior supports the hypothesis.
    O_t is a list with:
    [0] task_score          (0 to 1)
    [1] explanation_quality  (0 to 1)
    [2] guess_flag          (0 to 1; 1 means looks like guessing)
    [3] time_efficiency     (0 to 1; 1 means fluent, low effort)
    We reward task_score, explanation_quality, and NOT guessing.
    """
    task_score = O_t[0]
    explanation_quality = O_t[1]
    guess_flag = O_t[2]
    time_efficiency = O_t[3]
    # Weighted combination
    support_raw = (
        0.6 * task_score +
        0.3 * explanation_quality +
        0.1 * (1.0 - guess_flag)
    )
    # time_efficiency could also be included, but we keep it modest
    # to avoid overweighting speed over reasoning quality.
    # If desired, you could add: + 0.05 * time_efficiency
    return clip_to_unit(support_raw)
def gamma_I(C_t):
    """
    Uncertainty feature Γ_I.
    Higher when we still do not fully trust or fully observe the situation.
    C_t is a list with:
    [0] data_trust          (0 to 1; 1 means we trust the data source)
    [1] recency             (0 to 1; 1 means very recent evidence)
    [2] history_missingness  (0 to 1; 1 means we lack background history)
    [3] assist_suspected     (0 to 1; 1 means likely external/AI help)
    Uncertainty increases with missing history and suspected assistance.
    Uncertainty also increases if data_trust is low or evidence is not fresh.
    """
    data_trust = C_t[0]
    recency = C_t[1]
    history_missingness = C_t[2]
    assist_suspected = C_t[3]
    # Build an uncertainty score
    # - Strong history_missingness raises uncertainty.
    # - Suspected outside assist raises uncertainty.
    # - Low trust raises uncertainty.
    # - Old evidence raises uncertainty.
    uncert_raw = (
        0.5 * history_missingness +
        0.3 * assist_suspected +
        0.2 * (1.0 - data_trust) +
        0.2 * (1.0 - recency)
    )
    return clip_to_unit(uncert_raw)
def gamma_F(O_t):
    """
    Contradiction feature Γ_F.
    Higher when the observed behavior actively argues against the hypothesis.
    We treat guessing, weak explanation, and inefficient problem-solving
    under pressure as structured evidence that the skill may not actually
    be mastered.
    O_t is the same list as in gamma_T.
    """
    task_score = O_t[0]
    explanation_quality = O_t[1]
    guess_flag = O_t[2]
    time_efficiency = O_t[3]
    contradiction_raw = (
        0.5 * guess_flag +
        0.3 * (1.0 - explanation_quality) +
        0.2 * (1.0 - time_efficiency)
    )
    return clip_to_unit(contradiction_raw)
# 3. Neutrosophic transition model Φ_θ
# This applies the update rule described in the main paper:
# next = clip( a * current + b * Γ_feature )
#
# For each component:
#   T^_h(t+Δt) = clip_to_unit( a_T * T_h(t) + b_T * Γ_T )
#   Î_h(t+Δt) = clip_to_unit( a_I * I_h(t) + b_I * Γ_I )
#   F^_h(t+Δt) = clip_to_unit( a_F * F_h(t) + b_F * Γ_F )
def transition_phi(S_t, O_t, C_t, params):
    """
    Compute the predicted next neutrosophic state for one hypothesis h.
    S_t is a dict with:
        'T': T_h(t)
        'I': I_h(t)
        'F': F_h(t)
    O_t is the observation vector at time t (list of floats).
    C_t is the context vector at time t (list of floats).
    Params is a dict with the updated coefficients:
        'a_T', 'b_T',
        'a_I', 'b_I',
        'a_F', 'b_F'
```

```
    """
    # Current state values
    T_curr = S_t['T']
    I_curr = S_t['I']
    F_curr = S_t['F']
    # Evidence features
    Gamma_T = gamma_T(O_t)
    Gamma_I = gamma_I(C_t)
    Gamma_F = gamma_F(O_t)
    # Parameters
    a_T = params['a_T']
    b_T = params['b_T']
    a_I = params['a_I']
    b_I = params['b_I']
    a_F = params['a_F']
    b_F = params['b_F']
    # Predicted next components (before clipping)
    T_next_raw = a_T * T_curr + b_T * Gamma_T
    I_next_raw = a_I * I_curr + b_I * Gamma_I
    F_next_raw = a_F * F_curr + b_F * Gamma_F
    # Clip to [0,1] as required by the model
    T_next = clip_to_unit(T_next_raw)
    I_next = clip_to_unit(I_next_raw)
    F_next = clip_to_unit(F_next_raw)
    # Return predicted next state (we use "_hat" to match
notation in the paper)
    return {
        'T_hat': T_next,
        'I_hat': I_next,
        'F_hat': F_next,
        # Also return internal terms for transparency
        'Gamma_T': Gamma_T,
        'Gamma_I': Gamma_I,
        'Gamma_F': Gamma_F
    }
# 4. Loss computation
# Following Section 3 and Section 4.5 in the paper:
#
# L_T = (T_hat - T_star)^2
# L_I = (I_hat - I_star)^2
# L_F = (F_hat - F_star)^2
# L_C = 0 in this example (no hard physical impossibility
constraint needed)
# Total L = L_T + L_I + L_F + L_C
#
# We also return each component so that training can weight
them if desired.
def compute_losses(S_hat, S_star):
    """
    Compute the loss terms for one hypothesis h over one time
step.
    S_hat is a dict with:
        'T_hat', 'I_hat', 'F_hat'
    (predicted next-step neutrosophic state)
    S_star is a dict with:
        'T_star', 'I_star', 'F_star'
    (target / supervised next-step neutrosophic state)
    Returns a dict with all partial losses and the total.
    """
    T_err = S_hat['T_hat'] - S_star['T_star']
    I_err = S_hat['I_hat'] - S_star['I_star']
    F_err = S_hat['F_hat'] - S_star['F_star']
    L_T = T_err * T_err
    L_I = I_err * I_err
    L_F = F_err * F_err
    # In this demonstration, we set L_C = 0.0. This matches
the logic
    # in the main text for cases where no hard exclusivity
constraint applies.
    L_C = 0.0
    L_total = L_T + L_I + L_F + L_C
    return {
        'L_T': L_T,
        'L_I': L_I,
        'L_F': L_F,
        'L_C': L_C,
        'L_total': L_total
    }
# 5. Example scenario (education/learner skill mastery)
#
# Hypothesis h:
#   "The learner can apply the target skill correctly in new
situations."
#
# Step 5.1. Define observation vector O(t).
# These numbers are already normalized to [0,1] for
simplicity.
O_t = [
    0.78,  # task_score: performance on a new (unseen)
problem
    0.65,  # explanation_quality: clarity and correctness of
reasoning
    0.20,  # guess_flag: 0 means confident reasoning, 1 means
likely guessing
    0.55   # time_efficiency: 1 means fluent and quick, 0 means
very slow/struggling
]
# Step 5.2. Define context vector C(t).
# These values describe how reliable and recent the evidence
is.
C_t = [
    0.80,  # data_trust: 1 means we trust that the work is their
own
    0.90,  # recency: 1 means this is fresh/recent performance
data
    0.40,  # history_missingness: 1 means we have almost no
past data
    0.20   # assist_suspected: 1 means likely outside/AI help
```

```
]
# Step 5.3. Define the current neutrosophic state S_h(t).
# T_h(t): current support that the learner truly has the skill
# I_h(t): current uncertainty about the learner's true ability
# F_h(t): current structured evidence against mastery
S_t = {
    'T': 0.60,
    'I': 0.45,
    'F': 0.25
}
# Step 5.4. Define the learned transition parameters for Φ_θ.
# These correspond to the a_* and b_* weights in the main
text.
# They control how much we trust the previous state vs. new
evidence.
params = {
    'a_T': 0.6,
    'b_T': 0.4,
    'a_I': 0.7,
    'b_I': 0.3,
    'a_F': 0.5,
    'b_F': 0.5
}
# Step 5.5. Compute the predicted next state Ŝ_h(t+Δt).
S_hat = transition_phi(S_t, O_t, C_t, params)
# Step 5.6. Define the supervised target state S*_h(t+Δt).
# These values represent expert annotation after reviewing
more evidence
# (for example, a teacher's direct assessment, or a high-
quality oral check).
# Each value is in [0,1].
S_star = {
    'T_star': 0.72,  # target support for the skill at t+Δt
    'I_star': 0.30,  # target remaining uncertainty
    'F_star': 0.20   # target contradiction against the skill
}
# Step 5.7. Compute losses comparing the prediction vs. the
target.
loss_dict = compute_losses(S_hat, S_star)
# 6. Print all results in a clean, readable way.
print("=== INPUT STATE AT TIME t ===")
print(f"T_h(t) = {S_t['T']:.4f}")
print(f"I_h(t) = {S_t['I']:.4f}")
print(f"F_h(t) = {S_t['F']:.4f}")
print()
print("=== OBSERVATIONS AND CONTEXT AT TIME
t ===")
print(f"O(t) = {O_t}")
print(f"C(t) = {C_t}")
print()
print("=== INTERMEDIATE FEATURE SIGNALS ===")
print(f"Gamma_T (support feature)      =
{S_hat['Gamma_T']:.4f}")
print(f"Gamma_I (uncertainty feature)    =
{S_hat['Gamma_I']:.4f}")
print(f"Gamma_F (contradiction feature)  =
{S_hat['Gamma_F']:.4f}")
print()
print("=== PREDICTED NEXT STATE Ŝ_h(t+Δt) ===")
print(f"T_hat(t+Δt) = {S_hat['T_hat']:.4f}")
print(f"I_hat(t+Δt) = {S_hat['I_hat']:.4f}")
print(f"F_hat(t+Δt) = {S_hat['F_hat']:.4f}")
print()
print("=== TARGET NEXT STATE S*_h(t+Δt) ===")
print(f"T_star(t+Δt) = {S_star['T_star']:.4f}")
print(f"I_star(t+Δt) = {S_star['I_star']:.4f}")
print(f"F_star(t+Δt) = {S_star['F_star']:.4f}")
print()
print("=== LOSS COMPONENTS ===")
print(f"L_T (support error)       = {loss_dict['L_T']:.6f}")
print(f"L_I (uncertainty error)   = {loss_dict['L_I']:.6f}")
print(f"L_F (contradiction error) = {loss_dict['L_F']:.6f}")
print(f"L_C (consistency penalty) = {loss_dict['L_C']:.6f}")
print(f"TOTAL LOSS L              =
{loss_dict['L_total']:.6f}")
```