Trustworthy Neurosymbolic AI for Predictive Decision Support

Kamal Acharya, Ph.D.

University of Maryland, Baltimore County
Department of Information Systems

Department of Population Health Sciences
Weill Cornell Medicine

July 10, 2026

Research Journey

Artificial Intelligence

$\Downarrow$

Trustworthy AI

$\Downarrow$

Neurosymbolic AI

$\Downarrow$

Predictive Decision Support

$\Downarrow$

Healthcare **Transportation** Scientific Discovery

Research Background

Ph.D. in Information Systems at UMBC with 15+ publications in leading AI venues including IEEE Transactions on Artificial Intelligence, IJCAI, IEEE Big Data, IEEE ITSC, IEEE CAI, and AIAA.

Roadmap

1. **Motivation: Why Trustworthy AI?**
2. **Foundations of Neurosymbolic AI**
3. **Trustworthy Neurosymbolic AI**
   - Robustness
   - Uncertainty Quantification
   - Intervenability
4. **Application: Advanced Air Mobility**
5. **Research Vision**

Artificial Intelligence is Transforming Decision Making

- Artificial Intelligence (AI) has become a transformative technology across numerous application domains.
- AI systems increasingly assist or automate decision-making in
  - Healthcare
  - Transportation
  - Finance
  - Cybersecurity
  - Robotics
  - Scientific Discovery
- As AI becomes part of critical decision-making processes, ensuring its reliability and trustworthiness becomes increasingly important.

Why Can't We Fully Trust Today's AI?

Current Limitations

- Black-box decision making
- Limited interpretability
- Prediction uncertainty
- Distribution shift
- Limited human intervention
- Safety and robustness concerns

Data

$\Downarrow$

AI Model

$\Downarrow$

AI Decision

?

Can We Trust It?

From Symbolic AI to Neurosymbolic AI

Symbolic AI

- Logic, rules, and knowledge
- Transparent reasoning
- Human-readable decisions

Challenge: limited scalability and weak handling of noisy data.

Neural AI

- Data-driven learning
- Strong predictive performance
- High-dimensional pattern discovery

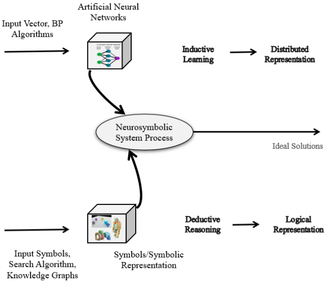
Challenge: black-box behavior and limited explicit reasoning.

Why Neurosymbolic AI?

Neurosymbolic AI combines learning and reasoning to support interpretable, robust, and trustworthy decision-making.

Foundations of Neurosymbolic AI

IEEE Transactions on Artificial Intelligence (2023)
Neurosymbolic Reinforcement Learning and Planning: A Survey



Research Outcome

This work established the conceptual foundation for my subsequent research on trustworthy neurosymbolic AI and predictive decision support.

Research Theme 1: Trustworthy Neurosymbolic AI

Arabian Journal for Science and Engineering (2025)
Neurosymbolic AI for Robustness, Uncertainty Quantification, and Intervenability

Research Objective

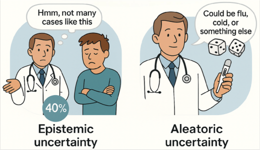
Develop trustworthy neurosymbolic AI
for high-stakes decision support.

Three Complementary Research Pillars

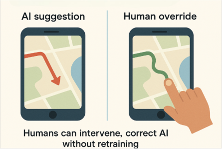
Robustness, Uncertainty Quantification and Intervenability

Three Pillars of Trustworthy AI

Stable Performance Under Varied Condition
Robustness



Measuring Confidence in Prediction
Uncertainty Quantification



Human Ability to Modify or Influence AI
Intervenability

Robustness: Reliable AI Beyond the Training Data

Robust AI should remain reliable under:

- Noisy or corrupted inputs
- Missing or incomplete information
- Distribution shifts
- Adversarial perturbations

Key Challenge

High predictive accuracy alone does not guarantee reliable performance in real-world deployment.

Training Data

$\Downarrow$

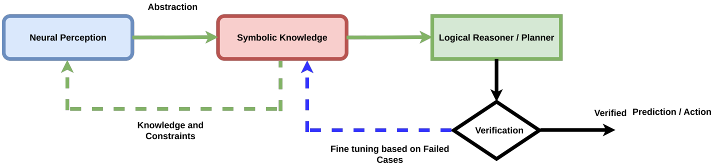
AI Model

$\Downarrow$

Real-World Conditions

Reliable?

Neurosymbolic Framework for Robustness



Key Idea

Symbolic knowledge provides explicit constraints and reasoning that complement neural learning, improving robustness under uncertain and changing conditions.

Research Insights on Robustness

- Neural models learn effectively from data but may fail under unexpected conditions.
- Symbolic knowledge introduces explicit constraints and domain rules.
- Neurosymbolic integration improves reliability while preserving predictive capability.

Takeaway

Robustness can be strengthened by augmenting neural learning with symbolic reasoning.

Uncertainty Quantification: Knowing When the Model Is Unsure

Uncertainty may arise from:

- Noisy or ambiguous data
- Limited or incomplete training data
- Model assumptions and limitations
- Stochastic training procedures

Key Challenge

A prediction is not sufficient unless the model can also communicate its confidence.

Input Data

$\Downarrow$

AI Model

$\Downarrow$

Prediction

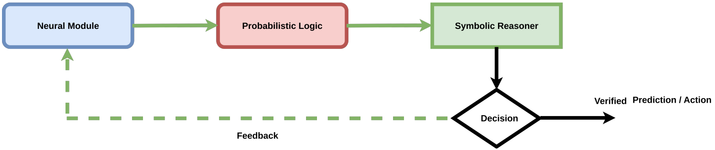
How certain is it?

Sources and Quantification of Uncertainty

Key Distinction

Aleatoric uncertainty arises from the data, while **epistemic uncertainty** arises from limited model knowledge.

Neurosymbolic Framework for Uncertainty Quantification



Key Idea

Neural models estimate predictive uncertainty, while symbolic knowledge provides constraints and reasoning to identify inconsistent or unreliable predictions.

Intervenability: Keeping Humans in the Loop

Intervenability enables humans to:

- Modify AI decisions when necessary
- Inject domain knowledge
- Correct reasoning errors
- Improve transparency and accountability

AI Prediction

$\Downarrow$

Human Intervention

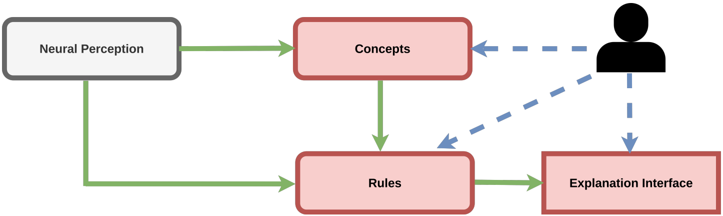
$\Downarrow$

Updated Decision

Key Challenge

AI should support human decision-making, not replace it.

Neurosymbolic Framework for Intervenability



Key Idea

Symbolic rules provide explicit intervention points, enabling humans to inspect, modify, and guide AI decisions without retraining the entire model.

Summary: Trustworthy Neurosymbolic AI

Robustness

$+$

Uncertainty Quantification

$+$

Intervenability

$\Longrightarrow$

Trustworthy Neurosymbolic AI

Research Theme 2: Neurosymbolic Predictive Decision Support

IWCMC 2025
Neurosymbolic Approach for Travel Demand Prediction:
Integrating Decision Tree Rules into Neural Networks

Research Question

Can symbolic knowledge improve predictive models
while preserving the learning capability of neural networks?

Decision Tree Rules

$$+$$

Neural Networks

$$\Longrightarrow$$

Interpretable Predictive Decision Support

Advanced Air Mobility: A Challenging AI Application

Advanced Air Mobility involves:

- Emerging electric aircraft and mobility services
- Limited historical demand data
- Complex spatial and socioeconomic interactions
- Uncertain adoption and travel behavior
- Safety-critical planning decisions

Limited Data

$\Downarrow$

Complex Demand

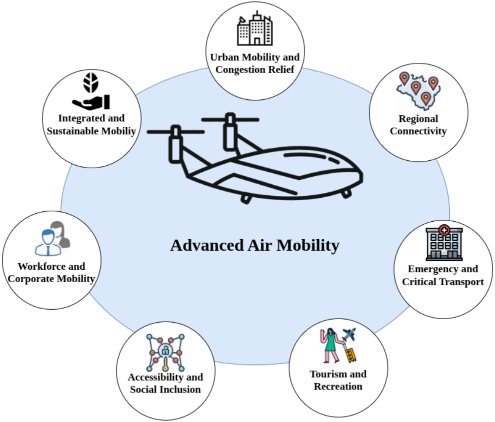
$\Downarrow$

Uncertain Decisions

Trustworthy AI Needed

Research Need

AAM requires predictive models that are accurate, interpretable, and reliable under uncertainty.

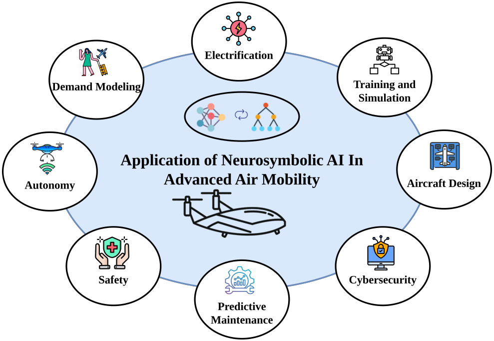


Decision-Support Requirements

Demand forecasting supports infrastructure planning, fleet allocation, route design, and service deployment.

Neurosymbolic AI for Advanced Air Mobility

IJCAI 2025 (Survey Track)
Neurosymbolic Artificial Intelligence for Advanced Air Mobility: A Survey



Research Insight

This survey highlights how integrating neural learning with symbolic reasoning can address the reliability, interpretability, and decision-support requirements of Advanced Air Mobility.

Predictive Decision Support: The Modeling Challenge

Neural Networks

- Strong nonlinear learning
- High predictive capability
- Handles complex feature interactions

Limitation: Predictions are difficult to interpret.

Decision Trees

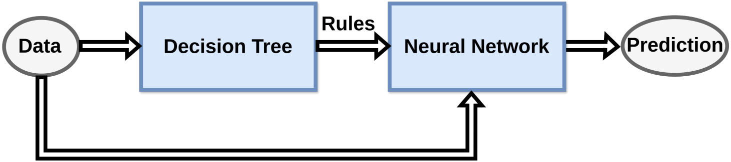
- Transparent decision structure
- Human-readable rules
- Explicit feature relationships

Limitation: Limited flexibility for complex patterns.

Research Goal

Integrate symbolic rules from decision trees with neural learning to support accurate and interpretable prediction.

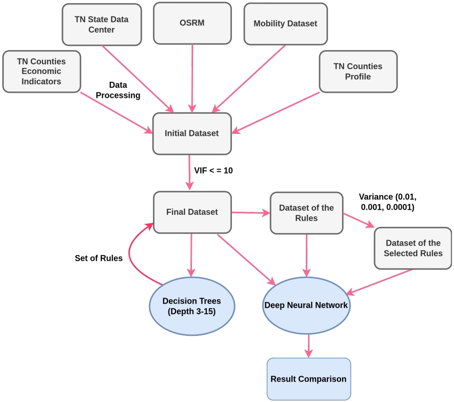
Neurosymbolic Predictive Framework



Key Contribution

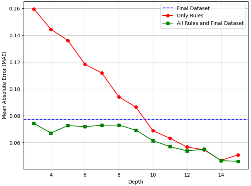
Decision-tree rules are encoded as symbolic features and integrated with the original data to train a neural predictive model.

Experimental Workflow

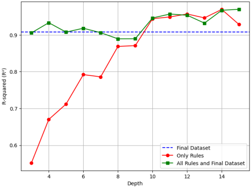


Evaluation Strategy

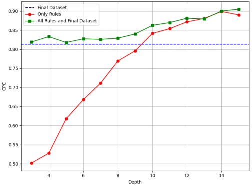
We compare the baseline dataset, symbolic rule representations, and combined neurosymbolic datasets across multiple decision-tree depths.



Mean Absolute Error



Coefficient of Determination



Common Part of Commuters

Key Observation

The combined neurosymbolic representation consistently provides the strongest predictive performance by combining symbolic rules with neural learning.

Key Findings

- Combining the original features with decision-tree rules produced the strongest predictive configurations.
- Symbolic rules preserved interpretable relationships among distance, points of interest, and travel demand.
- Rule selection reduced the symbolic feature space while retaining useful predictive information.
- The framework demonstrates how symbolic knowledge can complement neural learning in predictive decision support.

Takeaway

Neurosymbolic integration can improve predictive modeling while providing a more transparent representation of learned relationships.

Research Contributions

1. **Established a framework for trustworthy neurosymbolic AI**
   - Robustness
   - Uncertainty Quantification
   - Intervenability
2. **Developed a neurosymbolic predictive framework**
   - Integrated decision-tree rules with neural networks
   - Improved interpretability while maintaining predictive performance
3. **Demonstrated applicability to safety-critical decision support**
   - Advanced Air Mobility as a case study
   - Generalizable to other application domains

Limitations and Immediate Next Steps

- Rule selection currently depends on predefined variance thresholds.
- Static rules may not fully capture temporal or contextual changes.
- Future models should support adaptive rule extraction and uncertainty-aware reasoning.

Research Vision

Neural Learning

$+$

Symbolic Knowledge

$+$

Human Guidance

$\Longrightarrow$

Trustworthy Predictive Decision Support

Methodological Directions

- Adaptive knowledge integration
- Knowledge graphs
- Foundation models
- Human-in-the-loop learning

Application Directions

- Digital health
- Population health
- Transportation
- Scientific decision support

Emerging Research Direction: Neurosymbolic Foundation Models

IEEE Transactions on Artificial Intelligence (2024)
A Survey on Symbolic Knowledge Distillation of Large Language Models

Key Research Contributions

- Introduced the first comprehensive survey on Symbolic Knowledge Distillation (SKD) for LLMs.
- Developed a taxonomy of SKD approaches:
  - Direct Distillation
  - Multilevel Distillation
  - Reinforcement Learning-based Distillation
- Identified open challenges in scalability, knowledge consistency, interpretability, and trustworthy AI.
- Established SKD as a bridge between neural foundation models and symbolic reasoning.

Large Language Models

$\Downarrow$

Symbolic Knowledge Distillation

$\Downarrow$

Rules • Knowledge Graphs
Logical Reasoning

$\Downarrow$

Interpretable Trustworthy Neurosymbolic AI

Take-Home Messages

1. High-stakes AI requires more than predictive accuracy, it requires trustworthy, interpretable, and human-aligned decision support.
2. Neurosymbolic AI combines data-driven learning with explicit symbolic knowledge to improve reasoning and interpretability.
3. Robustness, uncertainty quantification, and intervenability are key pillars for developing trustworthy AI systems.
4. Symbolic knowledge integration enhances predictive decision support by combining neural learning with human-understandable reasoning.
5. Symbolic knowledge distillation extends these principles to foundation models and LLMs, enabling more efficient, interpretable, and trustworthy AI.

Thank You