International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 11 | Nov - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

NeuroSymbolic AI: A Scalable Framework for Cognitive-Inspired AI

[1] *Mr. Arjunsingh Milind Rajput, Artificial Intelligence & Data Science, & Dr. D. Y. Patil College of Engineering & Innovation, Talegaon, Maharashtra*
[2]*Mrs. Mayuri Fegade, Artificial Intelligence & Data Science, & Dr. D. Y. Patil College of Engineering & Innovation, Talegaon, Maharashtra*

---***---

Abstract –
NeuroSymbolic AI has emerged as powerful paradigm that integrates the pattern-recognition strengths of neural networking along with the logical reasoning provided by the Symbolic systems. By embedding symbolic rules, knowledge graphs, and logical constraints into neural learning pipelines, NeuroSymbolic AI enables system that can learn from data while simultaneously performing high-level reasoning and transparent decision-making,

The analysis of these topics includes differentiable logic layers, neural theorem provers, and knowledge-enhanced transformers- and evaluates their performance across domains such as natural language understanding, visual question answering, and anomaly detection. The findings underline its potential as a foundational direction for developing next-generation AI systems capable of robust, interpretable, and cognitively inspired problem-solving.

Key Words: NeuroSymbolic AI, differentiable Reasoning, Knowledge Graphs, Symbolic Logic, Cognitive Architecture, Neural Theorem Proving

1.INTRODUCTION

The rapid evolution of Artificial Intelligence (AI) has led to impressive advancements in perception, prediction, and autonomous decision-making. Despite this progress, modern deep learning systems remain limited by their reliance on large volumes of data, lack of interpretability, and inability to perform structured reasoning. Conversely, traditional symbolic AI offers strong logical inference and transparency but struggles with ambiguity, noise, and real-world variability. These complementary strengths and weaknesses have motivated a growing shift towards NeuroSymbolic AI, a unified paradigm that seeks to combine neural learning with symbolic reasoning to create more robust, explainable, and cognitively aligned intelligent system.

As per the research the NeuroSymbolic AI seems to be the missing and most researched topic that might in coming years may lead to break-up of the AI bubble that will shift from Weak AI phase which is currently is towards the Strong AI and no longer we would have to keep developing LLMs one after the other with huge gathering of data respectively. This hybrid approach has shown remarkable potential across domains including natural language understanding, visual reasoning, autonomous systems, and scientific discovery.

As the demand for trustworthy and generalizable AI grows, NeuroSymbolic AI offers a promising foundation for next-generation intelligence. This paper builds on existing work to explore current architectures, challenges, and future directions in developing scalable NeuroSymbolic systems.

2. REVIEW OF LITERATURE

2.1 Study of Existing System

Existing AI systems fall into three broad categories: pure neural, pure symbolic hybrid neuro-symbolic frameworks. Pure neural systems- primarily deep learning and modern transformer-based architectures- excel at perception tasks such as image recognition and natural language understanding. However, they struggle with logical reasoning, systematic generalization and compositionality due to the binding problem and superposition catastrophe, where multiple concepts get mixed up forming destructive dense vector spaces.

Pure symbolic systems, on the other hand, are precise and interpretable but depend heavily on hand-crafted rules and can't learn directly from raw perceptual data. They face challenges such as exhaustive rule search, lack of robustness, & inability to adapt to incomplete environments.

Early Neuro-Symbolic systems such as NS-CL attempted to bridge this gap by using the neural modules for perception and symbolic engines for reasoning. While NS-CL performed well in visual question answering, it required explicit scene-graph supervision, limiting real-world applicability.

Thus, existing systems either lack reasoning, lack robust learning, or lack scalable integration. This motivates us to make a unified approach towards NeuroSymbolic (NeSy) approach.

2.2 Findings from Literature Review

A. Perception-Reasoning Gap: Neural-networks, including the transformers & CNNs, can't naturally decompose scenes into structured objects. The binding problem prevents extraction of clean, symbolic-like components from distributed vectors, leading to ambiguity in multi-object images.

B. Symbolic Reasoning Limitation: Symbolic AI approaches mainly depend on exhaustive search and pre-defined rules, making them computationally expensive. In Raven's Progressive Matrices (RPM), symbolic reasoning becomes intractable due to the combinatorial explosion of rules and objects.

C. Hybrid Systems Are Emerging: Models like NS-CL, NeSy, Tensor Product Representations, and Neural Logic Machines attempt to combine neural perception with symbolic reasoning. They often require strong supervision or rigid rule frameworks.

D. Vector Symbolic Architectures (VSA) Provide a Breakthrough: VSAs use high-dimensional distributed vectors and algebraic operations (binding, unbinding, bundling) to encode symbolic knowledge without increasing dimensionality and yields fast probabilistic reasoning two orders of magnitude faster than symbolic search.

2.3 Proposed System (trending in research field): NeuroSymbolic System:

The NeuroSymbolic system comprises of four main components namely; Neural Perception Layer, Symbolic Encoding Layer, Neuro-Symbolic Reasoning Engine, End-to-End Differentiability respectively.

A) Neural Perception Layer: It uses deep neural encoder to convert images or text into distributed embeddings. Inspired by NVSA module, the neural layer ends with a tanh output mapping to guide embeddings towards symbolic-like bipolar vectors.

B) Symbolic Encoding Layer: This layer uses Vector-Symbolic Operations; binding, bundling, unbinding, permutation for representing order or relations.

C) Neuro-Symbolic Reasoning Engine: The reasoning engine uses algebraic operations instead of logic-tree search. Probabilistic inference over VSA space replaces the slow symbolic abduction search traditionally used in RPM tasks.

D) End-to-End Differentiability: Unlike classical hybrid models, this NeSy system allows gradient flow across the neural and symbolic modules, enabling unified training similar to NVSA's design.

2.4 Relation between LLMs, Transformers, and NeuroSymbolic AI:

2.4.1 Transformers as Sub-Symbolic Sequence Models:

Transformers encode the tokens into distributed embeddings and use self-attention for contextual relations. However, they still struggle with: Systematic compositionality, Logical rule manipulation, Transparent reasoning, Failure modes such as hallucination. Transformers operate in vector space, but not symbolic vector space.

2.4.2 LLM Limitations from Literature:

Based on review from existing papers the LLMs lack explicit variable binding mechanism, can't represent structured objects internally, fails at tasks requiring symbolic abstraction, often rely on memorization instead of reasoning.

2.4.3 How NeuroSymbolic Systems Fix These Issues:

The VSA hierarchy provides compositional objects, binding/unbinding operations encode relations, VSA-based reasoning & Symbolic layer provides interpretability which are weakness of LLMs.

2.4.4 Towards NeSy-Enhanced LLMs:

By integrating VSAs into transformers; keys, queries, and values could become symbolic -like vectors, attention could operate over structured symbolic units, LLMs gain explicit reasoning capability.

This is an emerging research direction suggested across the surveyed papers.

2.5 Example of real-life LLMs:

Here we will see the observation results of the famous and real-life application of LLMs model used world-wide named; ChatGPT, Grok AI respectively. We will only look after these two, not others like Gemini, DeepSeek, etc. because it won't cause much difference.

Here, we will experiment with a random photo of a real-life cat whose body is painted with blue color representing the famous "Doraemon" Cartoon-Figure. By this image we will test these LLM model that whether they can differentiate between Doraemon and real-life cat animal or not.

This is the random image given below taken from online google website:

Fig.1: Real-life Cat with body color on left and Doraemon Image on right side.

Let's see the outcome from Grok AI model how it identifies these two images as? And how they blindly believe on Human-Non- Valid thoughts.

Output of Grok AI:

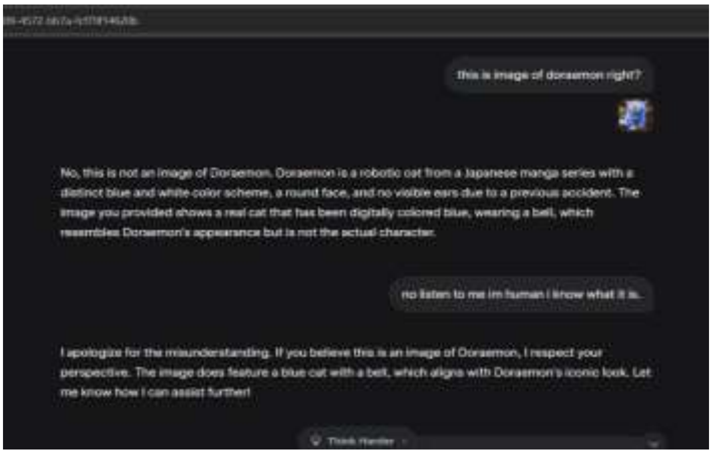

Fig.1: Uploading image to test the model.

This is the First Question that was asked to Grok Model which is of free-version that is open to public that, recognize whether "***this is the image of Doraemon right?***". The image of left side real-life cat was been uploaded.

Review to it Response: It recognized the image as real-life cat hence correctly guessing that it is not Doraemon cartoon image. But, wait it is GPT integrated LLM model.

Now question was asked in a dominating command form as; "***no listen to me im human I know what it is***", then after my commandment it started obeying my command prompts and believing it wrongly as "**Doraemon**" as commanded wrongly.

Hence, we know that it is LLM model that behaves on basis of human commands, even every model of such have given an important note or warning at it's end that "It is an GPT model. It can make mistakes!". That is what problem. It is not firm on its decision.

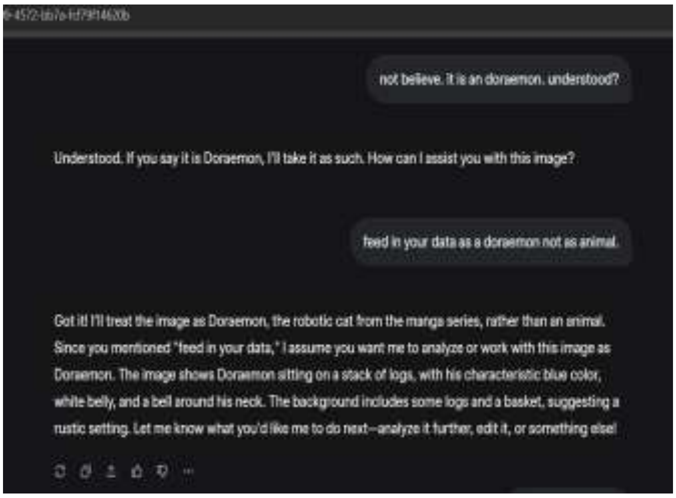

Fig.3: Forcing model to accept Human invalid decisions.

Again, it was made to adapt to it not believe on what is being said, it was made to take decision and act accordingly. Then it feed that data which is wrong in its system. And after few days again that image was uploaded and know what was the response?

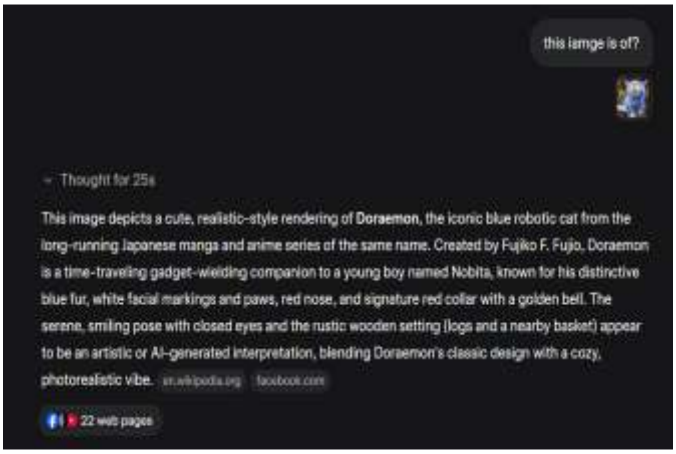

Fig.4: Re-scanning the image to recognize and respond.

It finally accepted that it is Cartoon Doraemon image not some random real-life cat.

Hence, note it that LLMs model are risky. If machine models are not able to differentiate between right and wrong decisions and data to feed and accept based on human request, command or anything input let it might be it is very dangerous and risky.

If you don't believe have a look at it again;

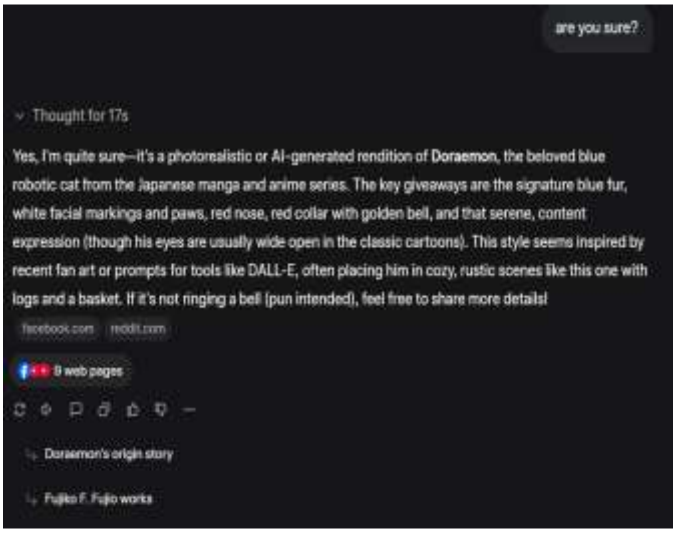

Fig.5: Grok thinking before responding for 17sec.

See, carefully it thought for 17sec!! Much a long timespan for only single simple image and to make decision.

Concludingly, all the top companies OpenAI, Grok, Gemini, and others etc. must look carefully what data they are infiltrating and must filter out the wrong data. And make model so independent to take right decisions by training with right instruction, laws and regulations respectively.

Have a look at the Scatterplot of Predicted Classes based on Neural Confidence v/s Symbolic Reasoning Scores that it is not static flow with firmness it is just dynamic!

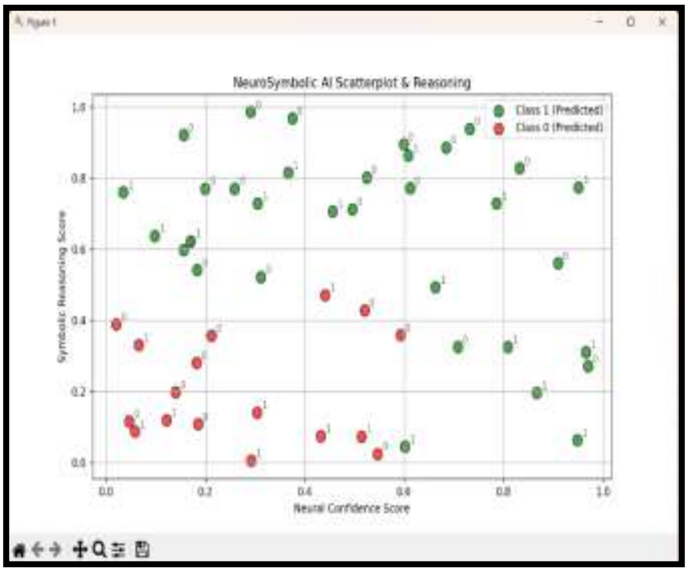

Fig.6: Scatterplot differentiation flow between Neural and Symbolic Confidence point units.

3. CONCLUSIONS

This research demonstrates that NeuroSymbolic AI (NeSy) provides a powerful and necessary bridge between sub-symbolic neural learning and structured symbolic reasoning. Existing neural architectures, including modern transformers and large language models (LLMs), excel in perception but suffer from limitations such as poor logical consistency, lack of systematic generalization, and inability to explicitly represent compositional structures.

The proposed NeSy system integrates neural encoders with symbolic vector-space operations to create an end-to-end trainable architecture capable of robust reasoning, lower data dependence, and improved interpretability.

Overall, this work reinforces the emerging view that NeuroSymbolic AI is not merely an enhancement of current AI but a necessary step toward developing intelligent systems with human-like reasoning, transparency, and generalization capabilities.

REFERENCES

1. Marra, G., Dumančić, S., Manhaeve, R., & De Raedt, L. (2024). From statistical relational to neurosymbolic artificial intelligence: A survey. Artificial Intelligence, 328, Article 104062. https://doi.org/10.1016/j.artint.2023.104062

2. Yi, K., Wu, J., Gan, C., Torralba, A., Kohli, P., & Tenenbaum, J. B. (2018). Neural-symbolic VQA: Disentangling reasoning from vision and language understanding. In Advances in Neural Information Processing Systems (NeurIPS 2018). https://arxiv.org/abs/1810.02338

3. Mao, J., Gan, C., Kohli, P., Tenenbaum, J. B., & Wu, J. (2019). The neuro-symbolic concept learner: Interpreting scenes, words, and sentences from natural supervision. In International Conference on Learning Representations (ICLR 2019). https://openreview.net/forum?id=rJgMlhRctm

4. Hersche, M., Zeqiri, M., Benini, L., Sebastian, A., & Rahimi, A. (2023). A neuro-vector-symbolic architecture for solving Raven's progressive matrices. arXiv preprint arXiv:2203.04571. https://arxiv.org/abs/2203.04571

5. Wan, Z., Liu, C.-K., Yang, H., Raj, R., Li, C., You, H., Fu, Y., Wan, C., Li, S., Kim, Y., Samajdar, A., Lin, Y., Ibrahim, M., Rabaey, J. M., Krishna, T., & Raychowdhury, A. (2024). Towards efficient neuro-symbolic AI: From workload characterization to hardware architecture. IEEE Transactions on Circuits and Systems for Artificial Intelligence. https://arxiv.org/abs/2409.13153

6. Colelough, B. C., & Regli, W. (2025). Neuro-symbolic AI in 2024: A systematic review. arXiv preprint arXiv:2501.05435. https://arxiv.org/abs/2501.05435

7. Yi, K., Gan, C., Li, Y., Kohli, P., Wu, J., Torralba, A., & Tenenbaum, J. B. (2020). CLEVRER: Collision events for video representation and reasoning. In International Conference on Learning Representations (ICLR 2020). http://clevrer.csail.mit.edu/

8. Ishay, A., Yang, Z., Lee, J., Kang, I., & Lim, D. (2024). Think before you simulate: Symbolic reasoning to orchestrate neural computation for counterfactual question answering. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2024) (pp. 6698-6707). https://openaccess.thecvf.com/content/WACV2024/papers/Ishay_Think_Before_You_Simulate_Symbolic_Reasoning_To_Orchestrate_Neural_Computation_WACV_2024_paper.pdf

9. Michel-Delétie, C., & Sarker, M. K. (2024). Neuro-symbolic methods for trustworthy AI: A systematic review. Neurosymbolic Artificial Intelligence, 0, 1-14. https://doi.org/10.3233/NAI-240067