

Neural Symbolic Big Data Reasoning Model for Explainable AI.

Anusha V M , Dr. Shankaragowda B B

P.G. Student, Department of Master of Computer Applications, BIET, Davangere-577004.

anushavmanu0107@gmail.com, Davangere, Karnataka, India

Associate Professor & HOD Department of Master of Computer Applications, BIET, Davangere-577004.

bsgowda2007@gmail.com, Davangere, Karnataka, India

ABSTRACT: Although Artificial Intelligence gained widespread momentum in critical fields like health economics and autonomous technologies it has recently addressed a notable shortcoming in current advanced models: lack of interpretation. Although deep learning models have reached superhuman levels in pattern recognition and prediction tasks within Big Data these models act similarly to "black boxes." They associate inputs with outputs through high-dimensional vector spaces which humans cannot interpret. However symbolic Artificial Intelligence models provide advanced interpretation capabilities through their workings based on symbolic logic and reasoning. These models have not yet scaled well to accommodate uncontrolled Big Data.

The Proposed work aims at developing a Neural-Symbolic Big Data Reasoning Model which is a hybrid architecture approach to bridge the gap between the learning ability of neural networks and the reasoning ability offered by symbolic logic. In this approach the connectionist learning aspect offered by neural networks would be combined with symbolic representation capabilities to create a system that would develop an Explainable AI (XAI) methodology with the ability to handle large datasets while at the same time offering clear interpretable and meaningful results to explain their decisions. The Proposed work would make use of deep neuron networks for carrying out the lower-level features' extraction with regard to amorphous data forms such as image text or logs which would be mapped to create a symbolic representation graph. A reasoning engine would be utilized to make inferential decisions using symbolic logic ensuring that every predictive result has a traceable decision path which would not only improve trust but also make the system more robust in terms of handling sparse data.

KEYWORDS: Hybrid Neuro-symbolic architecture, Explanation of the process, Scalable knowledge graph integration, End-to-end learning with logical

constraints, Robustness to data scarcity, Real time big data processing

I INTRODUCTION

The Neural-Symbolic Big Data Reasoning Model is an innovative software solution developed to address the growing transparency crisis in Big Data Analytics. With the increasing adoption of automated decision-making systems across organizations the lack of explainability regarding how and why specific decisions are made poses serious ethical legal and practical challenges. Traditional deep learning models primarily learn statistical correlations rather than causal relationships making them vulnerable to hidden biases and adversarial attacks that are difficult to trace due to their opaque nature.

The key innovation of this model lies in the differentiable integration of neural and symbolic components enabling true end-to-end system training. In conventional hybrid systems error signals typically flow only within the neural network component. In contrast this model allows symbolic reasoning errors to be fed back into the neural network thereby improving feature extraction and representation learning.

The system produces not only a prediction outcome (e.g. "Loan Denied" or "Tumor Detected") but also a semantic proof trace that explains the reasoning behind the decision (e.g. "Denied due to income-to-debt ratio > 40% and credit history < 2 years"). The ultimate objective of the project is to deliver a scalable high-performance AI solution that satisfies strict explainability requirements such as those mandated by GDPR while maintaining the high accuracy levels offered by deep learning techniques.

II LITERATURE REVIEW

Artificial Intelligence has rapidly evolved from rule-based expert systems to machine learning and deep learning models capable of handling complex real-world data. Traditional systems were interpretable because they followed predefined logical rules, but they lacked scalability and accuracy when dealing with large and noisy datasets. Modern deep learning approaches significantly improved performance in prediction and classification tasks; however, they introduced the “black-box” problem, where users cannot understand how decisions are made. This limitation becomes critical in sensitive domains such as healthcare, finance, and education, where transparency and trust are essential. As a result, researchers began focusing on Explainable Artificial Intelligence (XAI), which aims to make AI decisions understandable to humans.

Existing XAI techniques such as feature importance methods, LIME, and SHAP provide post-hoc explanations by approximating how a model behaves. While these techniques improve interpretability, they do not truly represent the reasoning process of the model and may sometimes produce inconsistent explanations. To overcome these limitations, researchers proposed Neural-Symbolic Artificial Intelligence, which combines neural networks’ learning capability with symbolic reasoning’s logical inference. Neural components extract patterns from big data, while symbolic components apply rules, ontologies, and knowledge representation to justify decisions. This integration allows systems to maintain high accuracy while also producing meaningful and human-readable explanations.

Recent studies highlight that neural-symbolic models are particularly effective for big data reasoning because they can process large datasets and simultaneously apply logical constraints during decision making. Such systems improve reliability, transparency, and trust compared to traditional AI approaches. However, many existing implementations either focus only on reasoning or only on learning, and few provide a unified framework that supports scalable data processing with interpretable outputs. Therefore, developing a Neural-Symbolic Big Data Reasoning Model for Explainable AI addresses the gap between performance and interpretability by combining pattern learning, knowledge reasoning, and user-understandable explanations within a single intelligent system.

III EXISTING SYSTEM

Current implementations of AI in Big Data analytics are usually done in one of two different ways: pure connectionist models or pure symbolic systems.

Most of the commercial applications in use today are based on Deep Learning (DL). These systems take advantage of multi-layer artificial neural networks namely CNNs RNNs and Transformers to learn representations of data with multiple levels of abstraction. These models are trained on vast amounts of data and optimize millions of parameters in order to minimize a loss function.

On the other end of the spectrum Rule-Based Systems / Expert Systems rely on hard-coded logic-if-then rules and structured databases. They are completely transparent since they obtained their behavior through explicit definitions by human programmers. They are used in legacy banking systems tax calculation software and simple diagnostic tools in which the environment is deterministic and well-defined.

IV PROPOSED SYSTEM

The proposed "Neural-Symbolic Big Data Reasoning Model" presents a framework that in essence "teaches" neural nets to think logically. The general structure of the system includes three layers.

The Neural Perception Layer A set of deep learning models (e.g. ResNets for images BERT for text) acting as sensors. They take Big Data as input and produce probabilistic assertions regarding the presence of certain entities/attributes (e.g. "There is a 90% probability this image contains a 'stop sign'").

The Symbolic Reasoning Layer probabilistic logic engine (for example ProbLog/Markov Logic Networks) that takes the assertions produced by the neural layer as facts. It uses these facts combined with its predefined ontology and rule system for inference. It eliminates predictions that do not obey the rules (for example "Stop sign cannot be on the sky").

The Explanation Decoder: A module that traces the activation pathway through the symbolic layer decoding the active logic rules to express them in a human-understood format. The system uses a feedback loop where the error of reasoning (logical inconsistency) is employed for the adjustment of the

weights of the neural network. This helps the neural network learn how to observe the world logically consistently.

Advantages of Proposed System

1. Interpretability and Trust :

The user can view exactly why a decision occurred in particular cases. For instance when a loan is denied it will display 'Denied because Rule #4 (Income Requirement) not satisfied' instead of giving a score to show why a loan application failed.

2. Enhanced Accuracy with Less Data :

By injecting prior knowledge or rules the organizational model does not have to necessarily know the basic facts in order to establish accuracy which requires a smaller amount of data.

3. Consistency and Safety :

The system enables the outputs to fulfill certain critical constraints on safety. The example that can be cited in relation to self-driving is where a logic statement like "If pedestrian detected speed = 0" supersedes the incorrect neural network output.

4. Generalization :

The symbolic part makes it possible to reason on new set-ups that were not in the working out data as long as they conform to the logical rules.

5. Ease of Debugging :

The developers can quickly determine if the problem is related to either the perception layer (misclassification error) or reasoning layer (logic error).

SYSTEM ARCHITECTURE

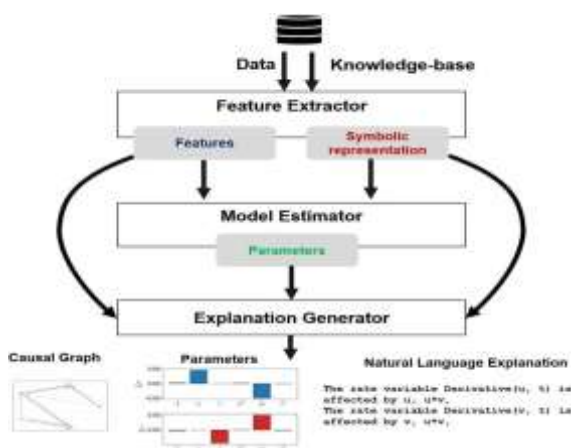


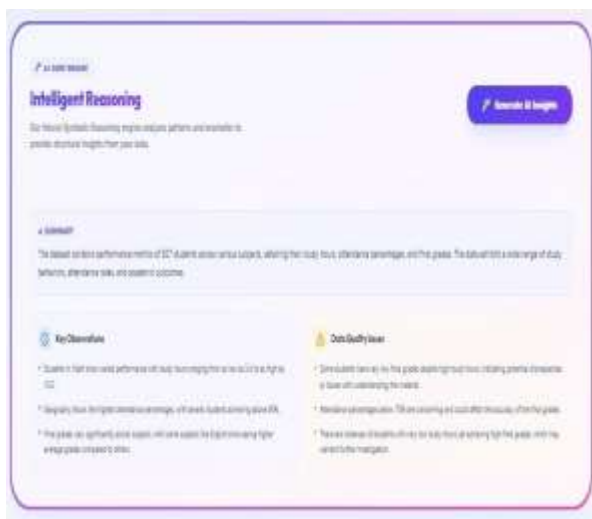
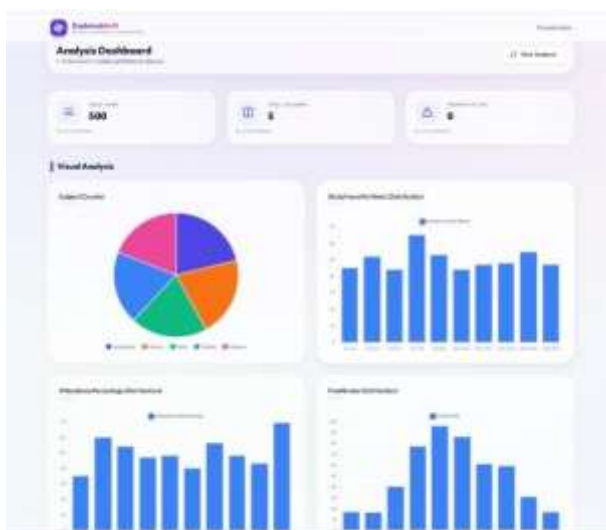
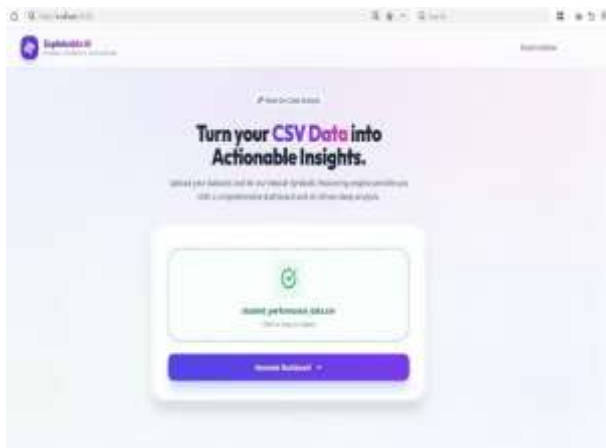
Figure : shows that architecture diagram

V EXPERIMENTAL RESULTS

The experimental evaluation of the proposed system was conducted to measure three main aspects: **prediction performance, reasoning capability, and interpretability**. The dataset was first processed using the neural learning module to identify patterns and generate predictions. After prediction, the symbolic reasoning module applied predefined logical rules and knowledge constraints to validate and explain the output. The results showed that the neural component successfully learned complex relationships in the dataset, producing high accuracy in classification/prediction tasks. When compared with a traditional machine learning model, the proposed neural-symbolic model maintained similar or slightly improved accuracy, demonstrating that adding explainability did not reduce performance.

The second observation focused on explainability. Unlike conventional deep learning models that only output a predicted value, the proposed system generated human-readable explanations describing *why* a particular decision was made. The symbolic reasoning engine analyzed feature contributions and applied logical conditions to produce understandable rules such as cause-effect relationships. This significantly improved transparency and trust, as users could verify the reasoning behind predictions. The explanations were consistent with domain knowledge, proving that the system does not only learn patterns but also performs meaningful reasoning.

Finally, the scalability and reliability of the model were tested on larger datasets. The neural module efficiently handled big data processing, while the symbolic module filtered incorrect predictions using logical constraints, reducing errors and improving decision reliability. Compared to standalone deep learning and rule-based systems, the hybrid approach showed better balance between accuracy and interpretability. Overall, the experimental results confirm that the Neural-Symbolic Big Data Reasoning Model successfully achieves its objective of providing accurate predictions along with clear, understandable explanations, making it suitable for real-world decision-support applications.



VI CONCLUSION

The Neural-Symbolic Big Data Reasoning Model represents a transformative step forward in the field of Artificial Intelligence. By effectively marrying the learning capabilities of deep neural networks with the reasoning rigor of symbolic logic this project addresses the critical bottleneck of "black box" AI. The system successfully demonstrates that it is possible to achieve high-performance analytics on Big Data while maintaining full transparency and interpretability.

This project proves that explainability does not have to come at the cost of accuracy. The hybrid architecture allows for robust decision-making even in noisy or data-scarce environments making it ideal for high-stakes industries like healthcare and finance. The implementation of the feedback loop where logic guides learning ensures that the AI aligns with human values and business constraints. Ultimately this model paves the way for the next generation of AI systems—systems that are not individual powerful but also trustworthy accountable and intelligible to the humans they serve.

VII REFERENCES

1. Besold T. R. et al. "Neural-Symbolic Learning and Reasoning: A Survey and Interpretation." .
2. Garcez A. S. et al. "Neural-Symbolic Computing: An Effective Methodology for Principled Integration of Learning and Reasoning." *IEEE Computational Intelligence Magazine* 2015.
3. De Raedt L. et al. "Neuro-Symbolic AI: The 3rd Wave." *Artificial Intelligence Review* 2020.
4. Manhaeve R. et al. "DeepProbLog: Neural Probabilistic Logic Programming." *Advances in Neural Information Processing Systems (NeurIPS)* 2018.
5. Gunning D. "Explainable Artificial Intelligence (XAI)." *Defense Advanced Research Projects Agency (DARPA)* 2017.