Adaptive Learning: Evolving Explainable Predictions

Bharatia, Harshal (School: Plano Senior High School)

This project presents a novel adaptive learning approach that not only enables a system to learn and evolve on its own but also explains its predictions, even for expansive problem domains that undergo frequent changes. It is hard to train and adapt learning models to changes when the problem domain is very large. With the adaptive learning approach, a system performs knowledge refinement automatically by determining what can be improved and trains itself without explicit guidance to do so. A system that provides its prediction without a basis to gauge the reliability of its predictions may not appear trustworthy for many applications. Therefore, each prediction is supplemented with additional information providing a model-agnostic explanation for the prediction and additional metadata reflecting the reliability of its predictions. Automatically derived explanation rules and additional possibilities using domain-specific knowledge further improve the utility of these predictions. The system uses novel techniques to automatically create dynamic ensemble learning models and learn meta-parameters that enable optimal predictions and their associated explanations. A reinforcement learning-based model selector identifies optimal models using automatically created hierarchical states. As a result, smart adaptive solutions that go beyond initial training become possible and alleviate the need for big complex monolithic models requiring extensive training. Experimental results show this approach learns quickly, adapts to changes very rapidly, performs quite well against prevalent learning methods, and provides accurate explanations.

Awards Won:

National Security Agency Research Directorate : First Place Award “Principles of Security and Privacy”