

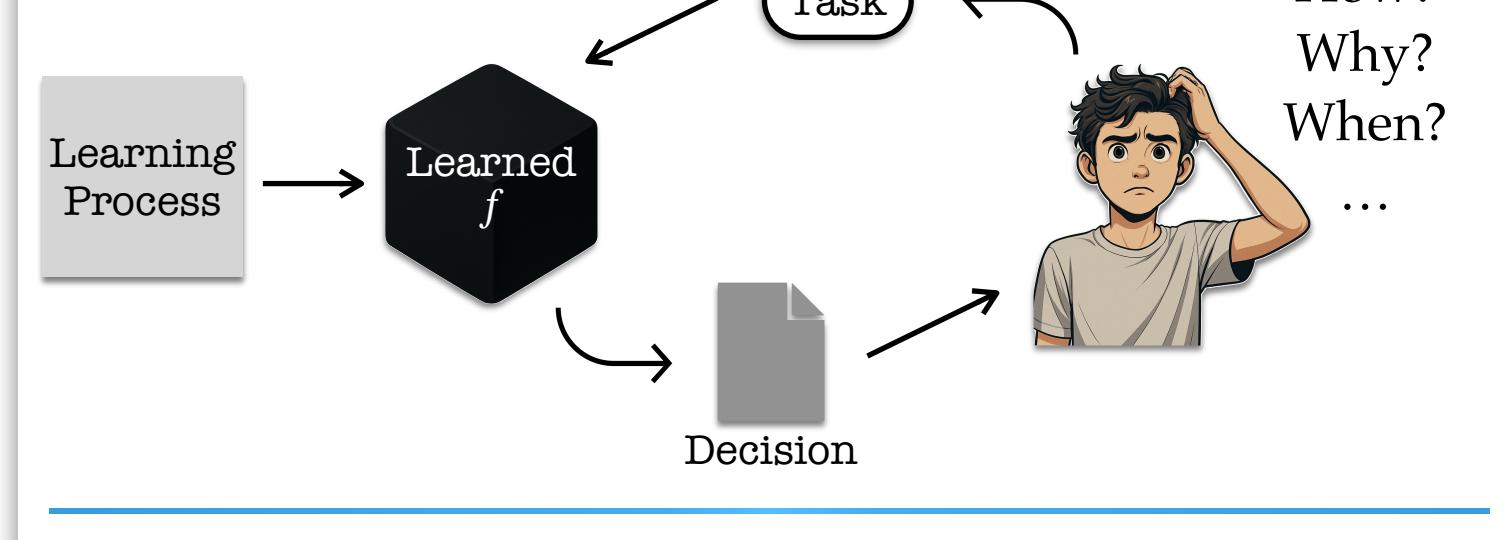
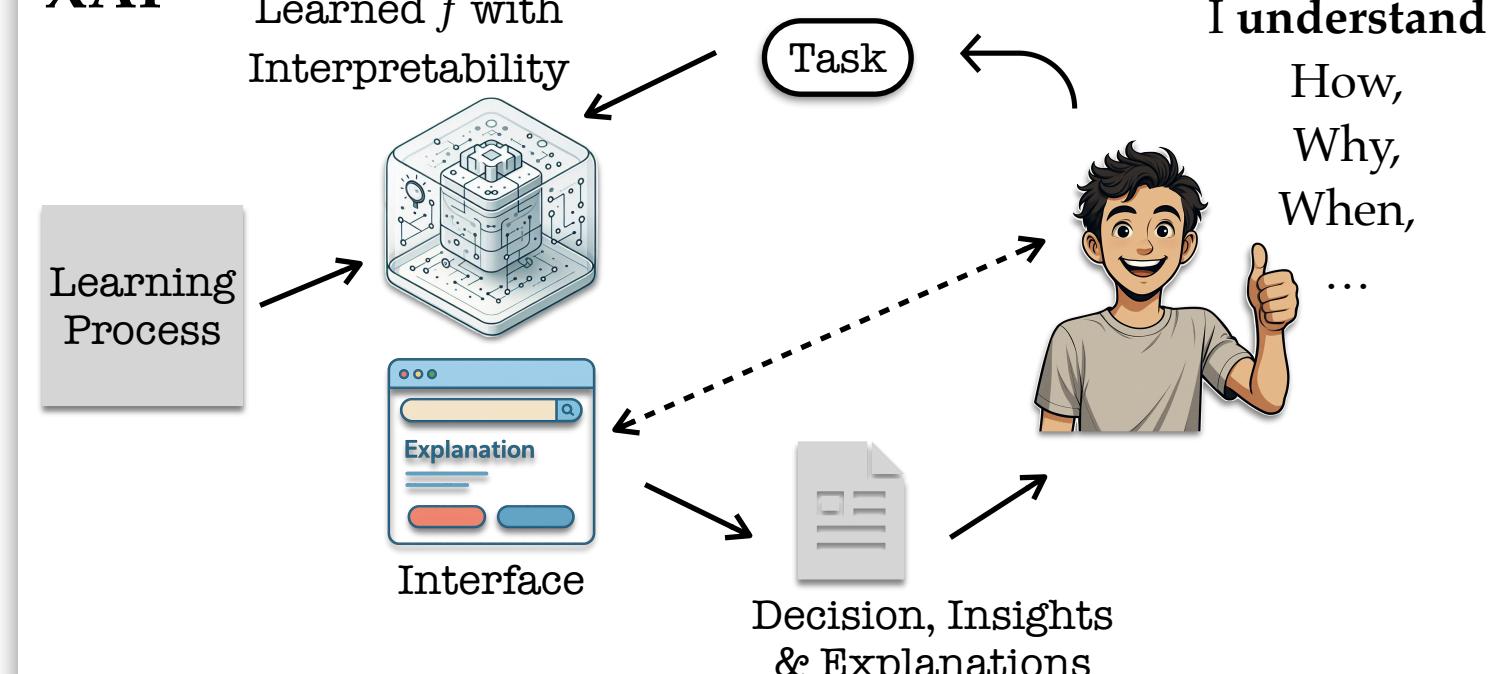


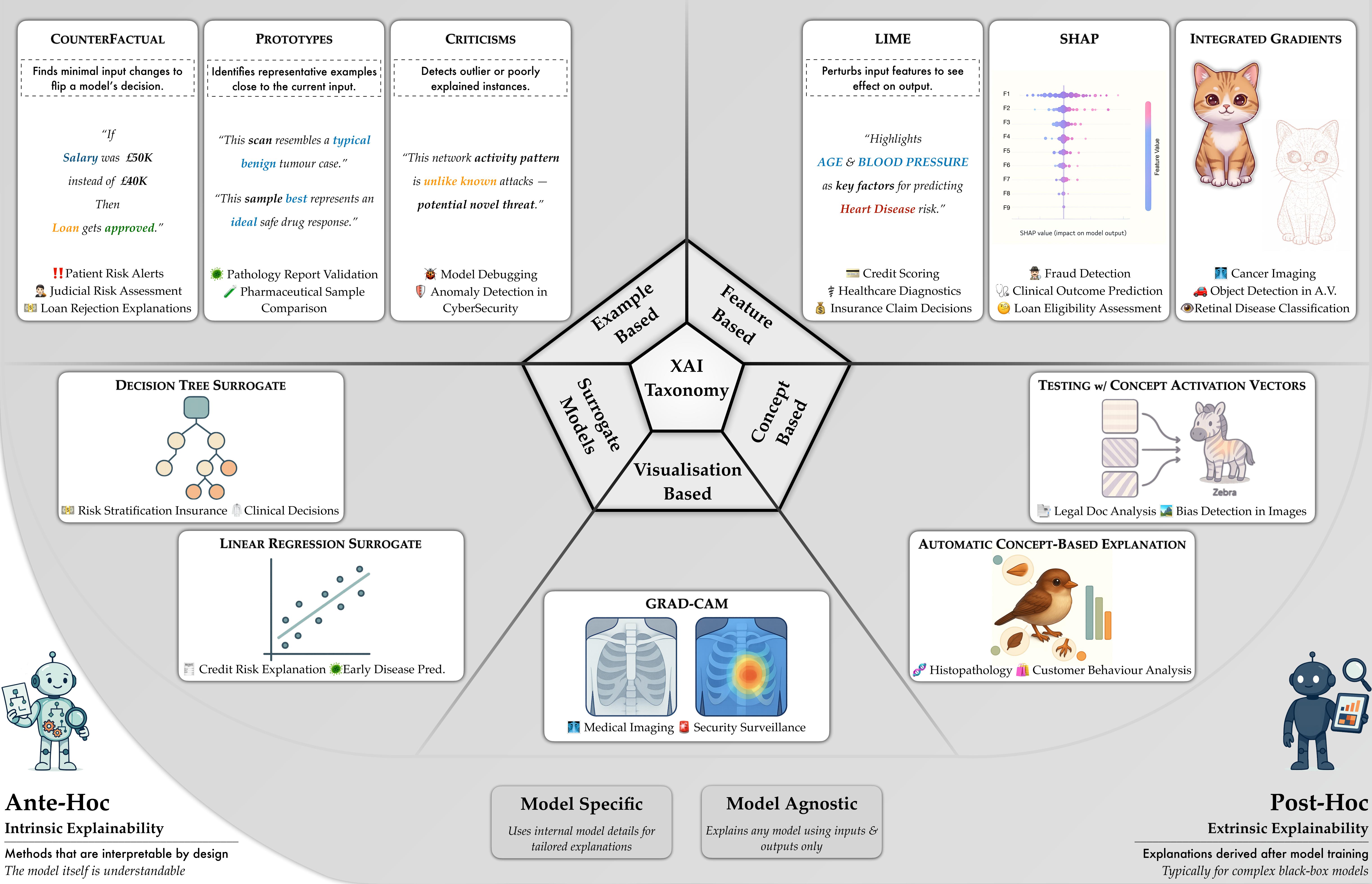
The Future of A.I. depends on

Your Knowledge of Explainability



Mr Samarthya Patel

What is eXplainable AI	Why it is “Need of the Hour”	Challenges & Open Problems	Future Directions
AI  XAI 	<p>Accountability & Responsibility Provides ability to trace & understand AI decisions, ensuring accountability when mistakes or harm occur.</p> <p>Ensuring Responsible AI Ensures that AI decisions are fair, transparent, and aligned with societal values.</p> <p>Regulatory Compliance Crucial for meeting legal requirements, such as the EU's GDPR, which mandates the right to explanation for automated decisions.</p> <p>Trust & Adoption AI systems need to be explainable for users to trust and adopt them in critical sectors like healthcare, finance, and autonomous driving.</p> <p>Managing Risks In high-impact sectors, XAI helps mitigate risks by providing context and reducing uncertainty in AI decisions.</p>	<p>Performance-Explainability trade-off More explainable models often sacrifice performance, while high-performing models tend to be opaque.</p> <p>Causal vs. Correlative Explanations Correlative explanations highlight associations, while causal explanations seek true cause-effect reasoning, which is harder to achieve.</p> <p>Faithfulness vs. Plausibility Explanations may appear intuitive to humans but fail to reflect the model's true reasoning.</p> <p>Domain Knowledge Gap Explanations often lack alignment with domain-specific concepts, limiting their usefulness to experts in real-world settings.</p> <p>Standardised Metrics There is no agreed-upon way to objectively evaluate the quality of explanations.</p>	<p>Causal Explainability Developing methods that provide insights into the cause-and-effect relationships behind model predictions, not just correlations.</p> <p>Temporal Explanations Developing methods that explain model behaviour over time-series, especially for sequential and dynamic data.</p> <p>Human-AI Collaboration Developing interactive explanations where users can query model decisions.</p> <p>Unified Evaluation Framework Creating standardised, comprehensive benchmarks to consistently assess and compare XAI methods.</p> <p>XAI for Emerging AI Paradigms Explaining Reinforcement Learning, Multi-Agent Systems, Foundation Models, Few-Shot Learning, and Large Language Models.</p>



Research Question: Can Combinations of Explainability Methods enhance the Confidence in Explanations?

