Phd Plan - TestifAI: A comprehensive testing framework for safe AI

Month Highlight: 36

Research Modules	Research Questions	Objectives	Tasks	PLAN START	PLAN DURATIO	ACTUAL START	ACTUAL DURATION	PERCENT COMPLETE
Test case generation	How can we sample inputs efficiently? Can SHAP values be used for effective test case generation?	**Efficient input sampling:** Develop a sampling approach that effectively identifies and prioritizes corner cases. It can be used to guide best selection of inputs for test case generation. **Effective test case generation with interpretability analysis:** Implement interpretability analysis to identify and prioritize key influential features in the DNN testing process, exploring the use of high influential features for effective test case generation.	Reading Literature on Test Cases	1	18	1	18	50%
			Exploring Libraries for Test Case Generation	1	2	1	2	100%
			Implementing Adversarial Attacks and Semantic Ac	3	1	3	1	100%
			Apply the exisiting test case genrtaion methods to	13	1	0	0	0%
			Automate my proposed test generation module	15	1	0	0	0%
Interpretablility	Can SHAP values be used for effective test case generation?	**Effective test case generation with interpretability analysis:** Implement interpretability analysis to identify and prioritize key influential features in the DNN testing process, exploring the use of high influential features for effective test case generation. **Systematic robustness evaluation:** Integrate advanced probabilistic methods to evaluate both local coverage and global coverage	Reading Literature on interpretabiity analysis	4	24	4	24	30%
			Implementation of SHAP tool	4	2	4	2	80%
			Applying SHAP to Identify Important Pixels	4	3	4	6	70%
			Explore other interpretability analysis techniques, such as LIME, to identify key features that can guide the generation of optimal test cases for evaluating model robustness.	15	3	0	0	0%
			Integrate Interpretability approach in test case generation module	16	1	0	0	0%
Coverage Criteria	How can we design a comprehensive framework to test system	**Systematic robustness evaluation:** Integrate advanced probabilistic methods to evaluate both local coverage and global coverage , providing comprehensive error summaries and systematically assessing robustness at different levels within the framework.	Reading Papers and Identifying Gaps	10	1	10	1	50%
			Apply the exisiting coverage criteria to the benchmark datasets and analyze the results.	13	1	0	0	0%
Proposed Methodology	How can we design a comprehensive framework to test system How can we systematically evaluate robustness at both local and global	**Design framework:** Create a framework that test the DNN under a variety of conditions, ensuring it meets performance standards even in edge cases and adverse scenarios	Designing a Conceptual Framework	8	4	8	4	70%
			Differentiating Local and Global Coverage	8	2	8	2	70%
			Implementing the Simple Adder Example	9	2	9	1	100%
			Exploring ProbLog	9	2	9	1	60%
			Implementing ProbLog for Global Robustness	9	2	9	1	80%
			Integrating ProbLog with Python	9	2	9	1	100%
			Applying the Framework to Different Datasets	9	2	9	1	60%
			Writing the Mini Thesis	9	3	9	3	100%
Probabilistic Programming Language	How can we systematically evaluate the robustness both at local (property- specific) and global (overall system) levels within the framework?	**Systematic robustness evaluation:** Integrate advanced probabilistic methodsto evaluate both local coverage and global coverage , providing comprehensive error summaries and systematically assessing robustness at different levels within the framework.	Reading Papers related to problog	10	2	0	0	0%
Sampling	How can we sample inputs efficiently?	**Efficient input sampling:** Develop a sampling approach that effectively identifies and prioritizes corner cases. It can be used to guide best selection of inputs for test case generation.	Reading Papers related to sampling techniques and identify gaps	9	1	9	1	50%
			Develop efficient sampling technique that will cover all inputs or corner cases	10	2	0	0	0%
			Implement existing sampling techniques to identify corner cases and prioritze that corner cases	9	1	9	1	30%
Error Summarization	How can error summarization be employed to quantify the impacts on	**Quantifying model robustness through error summarization:** Innovate error summarization techniques that identify and quantify model weaknesses, utilizing error summarization to quantify the impacts on	Find ways to proper summarize the counter	17	1	0	0	0%
			Best viusuals to represent erorrs report	18	1	0	0	0%
			Integrate error summarization module in frame	19	1	0	0	0%
Specification	How can we clearly define and specify the properties of the system?	**User-defined specifications:** Formalize specifications by developing templates and a specification language that enable users to clearly define and specify the properties of the system and its associated data for testing	Find a way to define specification, how to	20	3	0	0	0%
			How to pass Specifications to ProbLog	20	3	0	0	0%
			Integrate this module into the framework	20	3	0	0	0%

Milestones	Research Questions	Objectives	Tasks					
Conference Paper 1 (EuroML Conf 2025)	How can we design a comprehensive framework to test system robustness?	Prepare and submit a paper for the International Conference	Consider publishing initial findings or presenting at a conference.	9	3	9	5	50%
Conference Paper 2 (ICSE 2025)	Can SHAP values be used for effective test case generation?	Prepare and submit a paper for the International Conference	Consider publishing initial findings or presenting at a conference.	16	2	0	0	0%
Journal Paper 1	How can we sample inputs efficiently? How can we systematically evaluate the robustness both at local (property- specific) and global (overall system) levels within the framework?	Prepare and submit a comprehensive journal paper based on your extended research findings.	Prepare and submit a detailed manuscript based on the extended findings and analysis of the research conducted.	19	4	0	0	0%
Journal Paper 2	How can we clearly define and specify the properties of the system? Can SHAP values be used for effective test case generation? How can error summarization be employed to quantify the impacts on model robustness?	Prepare and submit a comprehensive journal paper based on your extended research findings.	Prepare and submit a detailed manuscript based on the extended findings and analysis of the research conducted.	24	4	0	0	0%
Thesis Writing	How can we clearly define and specify the properties of the system?	Prepae thesis draft	Focus primarily on writing thesis.	24	3	0	0	0%
Thesis Submission	How can we design a comprehensive framework to test system		Finalize and submit thesis for review.	28	3	0	0	0%
Defense and Final Submissio	robustness? How can we sample inputs efficiently? Can SHAP values be used for effective test case generation? How can we systematically evaluate the robustness both at local (property- specific) and global (overall system) levels within the framework? How can error summarization be employed to quantify the impacts on model	Prepare final comments	Defense and submit the finalized thesis.	36	1	0	0	0%