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

		i iid i idii i estii/	ii. / Comprehensi		- 3		Month Highlight		Plan Duration	Actual Star	rt % (	Complete Actual (be	yond plan)	% Comp	lete (beyond	plan)
Research Modules	Research Questions	Ressearch Objectives	Tasks	PLAN START	PLAN DURATION	ACTUAL START	ACTUAL DURATION	PERCENT COMPLETE				MONTHS				
Test case generation	How can we generate highly effective test cases that ensure complete	Efficient input sampling: Develop a sampling approach that effectively	Reading Literature on Test Cases	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 ing Literature on Test Cases 1 18 1 18 50%											29 30 31 32	2 33 34 35 36
rest case generation	coverage?	identifies and prioritizes corner cases. It can be used to guide best	Exploring Libraries for Test Case Generation	1	2	1	2	80%								
		selection of inputs for test case generation.  Effective test case generation with interpretability analysis: Implement		3	1	3	1	80%								
		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.	Apply the exisiting test case genrtaion methods to the benchmark datasets and analyze the results.	13	1	0	0	0%								
			Automate my proposed test generation module	15	1	0	0	0%								
			ristoniate my proposed test generation models		_											
Interpretablility	Can interpretability analysis aid in effective test case generation?	Effective test case generation with interpretability analysis: Implement	Reading liiterature on interpretability analysis	4	24	4	24	30%								
, ,		interpretability analysis to identify and prioritize key influential features in the DNN testing process, exploring the use of high influential features for	Implementation of SHAP tool	4	2	4	2	80%								
		effective test case generation.	Explore other interpretability analysis techniques,	4 15	3	0	6	70%								
			such as LIME, to identify key features that can guide the generation of optimal test cases for evaluating model robustness.													
			Integrate Interpretability approach in test case generation module	16	1	0	0	0%								
			generationmodule													
Coverage Criteria	How can we ensure complete test coverage for deep learning models?	Systematic robustness evaluation: Integrate advanced probabilistic methods to evaluate both local coverage and global coverage, providing	Reading papers and identifying gaps	6	4	6	4	50%								
			Apply the exisiting coverage criteria to the benchmark datasets and analyze the results.	13	1	0	0	0%								
Probabilistic	How can we systematically evaluate the robustness both at local (property-	Systematic robustness evaluation: Integrate advanced probabilistic methodsto evaluate both local coverage and global coverage, providing	Reading Papers related to problog	6	2	6	2	50%								
Programming Language Sampling	specific) and global (overall system) levels within the framework?	comprehensive error summaries and systematically assessing robustness at different levels within the framework.														
			Undertsand the problog language and compiler	9	1	9	1	80%								
			Run examples on problog and try to understand	9	1	9	1	70%								
			inputs and outputs													
	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%								
		selection of impots for test case generation.	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	9	1	9	1	30%								
			Integrate this module into the framework	9	1	0	0	0%								
Error Summarization	How can error summarization be employed to quantify the impacts on model robustness?	Quantifying model robustness through error summarization: Innovate error summarization techniques that identify and quantify model weaknesses, utilizing error summarization to quantify the impacts on model robustness	Find ways to proper summarize the counter  Best viusuals to represent erorrs report	17		0	0	0%								
			Integrate error summarization module in frame	18 19	1	0	0	0%								
Specification  Initial Proposed	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	Find a way to define specification, how to formalize it	20	3	0	0	0%								
		and specify the properties of the system and its associated data for testing purposes	How to pass specifications to ProbLog	20	3	0	0	0%								
	How can we design a comprehensive framework to test system	Design framework: Create a framework that test the DNN under a variety	Integrate this module into the framework  Designing a conceptual framework	20 8	3	0 8	0 4	0% 70%								
Methodology	How can we ensure complete test coverage for deep learning models?	of conditions, ensuring it meets performance standards even in edge cases	Differentiating local and global coverage	8	2	8	2	70%								
		and adverse scenarios	Implementing the Simple Adder Example Exploring ProbLog	9	2	9	1	100% 60%								
			Implementing ProbLog for calculating global	9	2	9	1	80%								
			robustness Integrating ProbLog with python	9	2	9	1	100%								
			Applying the framework to different datasets	9	2	9	1	60%								
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 interpretability analysis aid in 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-	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	specific) and global (overall system) levels within the framework?  How can we clearly define and specify the properties of the system?	Prepare and submit a comprehensive journal paper based on your	Prepare and submit a detailed manuscript based	24	4	0	0	0%								
200	Can interpretability analysis aid in effective test case generation?  How can error summarization be employed to quantify the impacts on model	extended research findings.	on the extended findings and analysis of the research conducted.	24	4			070								
Thesis Muiting	robustness?	Propae thesis draft	Focus primarily on writing thesis	24	2	0	0	206								
Thesis Writing Thesis Submission	How can we design a comprehensive framework to test system robustness?	Prepae thesis draft Submit thesis	Focus primarily on writing thesis. Finalize and submit thesis for review.	24 28	3	0	0	0%								
Defense and Final	How can we clearly define and specify the properties of the system?  How can we sample inputs efficiently?	Prepare final comments	Defense and submit the finalized thesis.	36	1	0	0	0%						,,,,,,,,,		
Submission	How can we generate highly effective test cases that ensure complete															
	coverage? Can interpretability analysis aid in effective test case generation?															
	How can we ensure complete test coverage for deep learning models?															
	How can error summarization be employed to quantify the impacts on model															
	robustness?															