Introduction. Greetings, I'm Taha. I am passionate about science, and I thought that a piece of knowledge could be among the best legacies one can leave behind. For this reason, in 2013, I departed from the Iranian national chess team. My academic journey includes a BSc in Software Engineering, during which influential figures like Prof. Ali Gholami Rudi left a lasting impact. Subsequently, during my MSc, I had the privilege of being supervised by Prof. Saeed Jalili, who was also once a supervisor to Prof. Gholami.

Research Statement. Would be available for the formal application form or alternatively upon request.

Main Research Experience. My primary research spanned two years under Professor Saeed Jalili's guidance. Our work centered on Software Testing, specifically Mutation Testing. Our research process involved reviewing over 80 research papers in three months, during which we identified errors and occasionally engaged with authors or editors to resolve them. We also saw opportunities to enhance certain papers through modest changes, and in cases of significant flaws, we communicated our concerns with journal editors or shared them publicly on PubPeer.

In the next phase, we embarked on a specific thesis project in mutation testing. A fundamental challenge in mutation testing is selecting a subset of mutants for analysis. Recent research has highlighted the effectiveness of focusing on a particular type of mutant called "fault-revealing mutants." Our goal was to reduce the costs of mutation testing by predicting these useful mutants more accurately, primarily relying on data science techniques.

We started by exploring our main dataset, Codeflaws, and we discovered a substantial overlap between non-fault-revealing mutants and their fault-revealing counterparts. This overlap was observed when considering fault-revealing mutants separately as well. We conducted experiments by training a baseline model under two conditions: random oversampling and no resampling. Surprisingly, the model's predictions were highly similar, indicating comparable performance and predictions for any given data point. This observation raised two potential implications. Firstly, it suggested the existence of straightforward patterns in the dataset that could be learned effectively by our models, regardless of the attention given to fault-revealing mutants. However, this implication might not hold true for all patterns in the data. Secondly, the findings indicated an exceptionally high density of non-fault-revealing mutants between any two relatively close fault-revealing mutants, not just around individual ones.

Given these insights, we concluded that extensive effort to identify obvious patterns was unnecessary since our models could already learn them effectively. However, for fault-revealing mutants in close proximity, the machine learning algorithm needed to distinguish between two potential hypotheses: one that separates them into two groups and another that considers them a single group. Despite these findings, the high density of non-fault-revealing mutants in the dataset posed challenges for the algorithm in recognizing and considering these competing hypotheses. To address this, we explored various approaches, including various feature selection methods, resampling techniques, and different classification algorithms, including recent deep learning models. Unfortunately, none of these approaches yielded satisfactory results on our data.

To address these challenges, we devised two heuristic functions for undersampling, effectively removing around 90% of non-fault-revealing mutants from the training dataset. While this method improved model performance, the enhancement was insignificant. Subsequently, we critically assessed our previous method, acknowledging its limitations, such as the potential

creation of artificial patterns and scalability issues. However, it bolstered our confidence in the approach's promise and the possibility of resolving its drawbacks for higher accuracies.

Finally, we introduced the concept of "killability severity" as a previously overlooked feature in this context. Specifically, we designed a non-linear model to estimate "killability severity" as a novel feature. This addition enabled us to identify hidden mutant subclasses and consider alternative hypotheses, ultimately leading to improved fault-revealing mutant predictions. Our findings in this regard were published in a respected journal.

Subsequently, we formulated and explored additional research questions that are currently under review. As my research interests have changed, and I am now applying for a different focus, I will refrain from providing further details on these topics.

Additional Research Experience. I also participated in a few (small-scale) interdisciplinary research projects, assuming various roles to contribute. For instance, in a project related to tourism with two colleagues from the University of Mazandaran, I played a role in developing a web scraper to collect relevant data. In another project involving researchers from multiple universities, including Laurentian University, Lublin University of Technology, University of Alberta, University of Manitoba, and others, focusing on health, I took charge of data gathering. I also collaborated on a project with colleagues from the Babol Noshirvani University of Technology, centered around emotion recognition using deep reinforcement learning on the DEAP dataset. In this endeavor, my involvement spanned various project aspects. Lastly, I joined forces with two colleagues from Kharazmi University for a project in the domain of Business Administration. Our goal was to estimate the Gross Domestic Product using Artificial Intelligence (AI) techniques, and my involvement spanned various project aspects.

Institution, Program, and Supervisor Selection. Would be available for the formal application form or alternatively upon request.

Conclusion. Einstein once mused, 'God does not play dice.' Yet, as the intricacies of the universe unfold, it becomes evident that the Creator has woven the threads of chance into the very fabric of existence. As we conclude this discourse, you stand at the crossroads of choice among prospective candidates. The mantle of decision now rests upon your shoulders. Allow me, in closing, to extend my best wishes. In recognition of your kindness and wisdom, I offer a final gesture in your favor. May luck roll the dice in your favor, bringing forth the equivalent of a winning pair in your endeavors. Warm regards, Taha