Amir Pouya Moeini *pmoeini@ce.sharif.edu* pooyamoini2000@gmail.com

Statement of Purpose

of Amir Pouya Moeini (Computer Engineering student @ SUT)

I am Pouya Moeini and I'm currently a third-year undergrad in **Computer Engineering at the Sharif University of Technology**. I also have **minoring in Mathematics**. Recently, I have finished my fifth semester which was kinda my most important semester so far. If I want to talk briefly about myself, I'm going to first mention my interests. It seems that at least a bachelor's degree is needed for a person to have the experience to select his/her future way and path. But I knew my path (at least most of it) from the very first semester. Before entering university I was involved with the Olympiad (**International Olympiad Astronomy and Astrophysics**) and I managed to win the gold medal there (Actually, this International competition had overlap with my first semester and cause my first semester very difficult and challenging). So unlike my peers, I studied further than usual from high school, and since there, I was crazy about Physics and Mathematics. So by winning the gold medal I was able to select the most wanted major and university in my country which was perfectly fit for my interests. By knowing my way, I mean I see myself as an **academic person**, and I willing to want to work in this space for the future, not in an industrial space. Because I want the (relative) freedom you will have by working in an academic space rather than industrial to spent time on your problems. I always want to be a **researcher and a person which studies his entire life**. The best pleasure I have ever experienced in life is the joy of learning and problem-solving. Therefore, **this research internship experience would be one of the best first experiences that I will have the chance to try during my way and it perfectly fits my dreams and goals**. Thus, due to the above explanations, I've decided that my SOP be mostly about my academic experiences.

Experiences - Academic courses:

Now, I want to briefly explain my experiences so far. Currently, I am working on two research projects under the supervision of **Prof H.R Rabiee** at the Data Science and Machine learning lab. I start to become interested in the data science area approximately more than one year ago when I was studying **statistics and linear algebra** in the third semester (Statistics is still my most favorite topic). These two courses were amazingly interesting to me and so far, I decided to be teaching assistant at each semester for each of these courses, to always involved with them. After these two courses, like many other people I started learning machine learning aspects by Andrew Ng's Coursera course. But this course wasn't so much satisfying to me so I continue this course by watching its Stanford course rather than Coursera's one. After the start of the fourth semester, I decided to audit the deep learning course. The course was taught by Prof. Solyemani. This was one of the benefits of Covid that I was able to watch video lectures. I completely took the course seriously and did all the homework and the project. I also watched the cs231n video lectures alongside this course. Combining these two courses, I was passionate about computer vision during there. In particular, I was so interested in GANs and read a lot of papers about it, and implemented their codes to increase my literacy and mastery of this subject. You can see these codes here. My interest in GANs was such that I gave a presentation on it. I used this survey of Prof. Goodfellow as the main reference. Also, I used many other courses that I don't think it's necessary to mention. It was then that I decided to contact Prof. Rabiee to start my research experience under his supervision. I knew him and his group from a few of my friends that were working with him. As my research experience began, I decided to take more serious lessons. So I took two graduate courses this semester. One of them was the **theory of machine learning course** where was taught by Prof. Maddah-Ali from the Electrical Engineering department and the other one was **Bayesian statistics in Machine learning** where was taught by Prof. Alishahi from the Mathematics department. Also, I had **Artificial Intelligence** from my department and you can see the syllabus of this course here. I was the only person in the class who got the grade 20 out of 20. I introduce the first course I mentioned (theory of machine learning) as **the most beautiful and effective course I will experience**. You can see the syllabus here. This course made me realize that my main interest is in theoretical work (in particular Machine Learning and Information theory) and I want to work more in the fields of theory and Also, I realised that to work in theoretical fields, **I need to have strong mathematical background so I chose Mathematics as my second major (i.e minoring)**. Also, I am the youngest student in this course. This course has a paper presentation project and I have decided to present the

paper "Communication-Efficient Algorithms for Federated Learning" by Prof Mohri. The other graduate course I had this semester is Bayesian statistics in Machine learning. You can see the syllabus here. I am the youngest person in this course too. Also, this course has a paper presentation project too and I decided to present the paper "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning" by Prof. Zoubin Ghahramani. You can see my presentation here. For the next semester, I'm willing to take two graduate High Dimensional probabilities and Network-coding and Also Stochastic process. Also, I will be teaching assistant for **Machine learning, Statistical Machine learning, Artificial Intelligence and Regression Analysis courses**. It's mentionable that other than the theory of the machine learning course and Its lecturer (Prof. Maddah-Ali) there are other reasons that I become interested in learning theory. I am a huge fan of Prof. John duchi and in particular, his lecture notes was the main reason that I become interested in Information theory. Also, I'm participating in TTIC weekly webinars and those presentations are another main reasons for me to become interested in learning theory. Also, reading Manfred K. Warmuth's papers was another pleasure I have experienced since his works are responsible for many topics of the course. Also, I'm following Pof. Ilias Diakonikolas' works and courses because I believe that he is the leader of the learning theory community of the young generation. I'm following Prof. Ryan Tibshirani and Prof. Akshay Krishnamurthy works and courses too. There are many other people that I can mention but I don't think that's necessary.

Experiences - Research:

Now, I want to talk about my research experiences. **My primary research objective and interest are in the area of Machine Learning theory and Statistics**. I am actively involved in two research projects that the first one is **"Fake news mitigation in Network"**. The second one is **"A Fair Private Meta-Learning Approach for Distributed Data Personalization and User Profiling"**. In the first project, we want to find a trustable solution for mitigation the spread of fake news. The solution we intend to implement is like a vaccination process. We are implementing a way to **predict the next infected users** in an active cascade. Then, we will send the true news to them and mitigate the further spread of fake news. For this purpose, we need to find a way to predict a **batch of the next infected users** (rather than the next person, As it's already approximately solved). But the main challenge that we got was we wanted to find a "practical" solution. In a real situation, you don't have access to any social network of the cascade (meaning that you don't know about edges, you don't know who infected whom and which users are the candidate for getting infected) and you have partial observations. Meaning that you won't have the complete history of the cascade (which user and when got infected), Instead, you will have a very few observation of the cascade and Also, each user won't participate in more than a few cascades. To solve these challenges, we find an embedding for each user by its properties that are public (Like profile, Followers and etc). So basically we used a representation learning approach to learn this embedding space and represent each user in this space rather than using just its ID. After that we use a RL framework to predict next batch infected users. The currently state of this project is getting ready to write the initial draft of the paper and try to find a way to retrieve the missing information. The other project that I'm attending is to find a fair and private meta-learning approach for personalization tasks for distributed Data. This project has started recently and I'm taking mostly the theoretical parts, as my personal interest. These days I'm mostly studying (local) differential privacy and its applications in meta-learning and federated learning. To challenge my knowledge and literacy about differential privacy, I gave a presentation about it. Also, as I mentioned before I chose one of my course projects, related to the federated learning to be close to this project. In the first step, we found an application for our project (meaning that prepare dataset and etc) and now we are studying related works and necessary background for the topics. Currently, I am the only student in the lab who is working on two projects simultaneously.