Shuo Liu
202-378-4386   shuoliu9566@gmail.com

Dear hiring manager,

I am a PhD student in Computer Science at Georgetown University, and I am highly interested in your position of Data Governance, Associate - Controllers (Jersey City, NJ). I am attracted to the potential and opportunities offered by the position and the creative environment at Goldman Sachs. I believe there is a match between the requirements and my qualifications, motivating me to apply for the position.

My work experiences in industry consists of two software engineering internships at Google and Meta, which provided me with exposure to engineering in industrial standards, and experience in handling work in unfamiliar topics and tasks. My project at Google was on creating a metric of staleness for evaluating recommendation systems using machine learning and mathematical methods, and serving an internal tool to replace the current crowdsourcing-based method. This project was a whole lifecycle of applied machine learning, from data collection to training to evaluation and model selection, also involving fine-tuning of a BERT-like LLM model. It also included an implementation of internal tool that allows using of the metric signals with Python. My project at Meta involved building a verification step for the transfer learning workflow to speed up the overall procedure and enable early detection of error in inputs. The work was done on an internal version of PyTorch. These experiences complement my PhD research experience which is more theoretical, and would come in handy if any of my projects require implementations in the future.

My PhD research topic is distributed algorithms, specifically, resilient distributed optimization algorithms, including fault-tolerant and asynchrony in the distributed system. For example, machine learning is usually an optimization problem. Consider a system solving machine learning tasks involving multiple computers working collaboratively – federated learning, for example, and some of the participants might be faulty, due to hardware failures or adversarial attacks. My research focuses on designing resilient algorithms against these faulty participants so that they can still output usable outcomes, and showing how resilient the algorithms can be. The problems we are solving can also be appreciated from a security point of view, considering the faulty participants as malicious. Research in this line of work involves mathematical analysis of algorithms, as well as simulations of the algorithms with Python, and a list of my publications can be found in my resume and on Google Scholar.

I had a seminar during my PhD on data governance. We discussed data governance guidelines in different parts of the world like GDLP, CPRA, HIPAA, etc., and the applications of certain techniques like differential privacy to data governance and protection.

My other research experience is my graduate thesis on social media privacy. I studied a type of privacy breach on social media websites, where an attacker could discover a user's hidden information through public information of the users that he/she is connected to. This research involved data collection, data mining, analysis, and creating a demo website for users to check their possible social media privacy issues. It is a part of a larger project in the lab and we used Python, Java, PHP, and other coding languages in the whole project.

In summary, my research and work experiences involve extensive engineering tasks. My training and experience during my pursuit of graduate degrees and work experiences allow me to positively contribute to the position. Thank you in advance for your consideration of my application. I look forward to hearing from you soon and I would appreciate an opportunity for an interview.