

Dear Niclas Persson,

We are Yitian Shen and Zixi Wang, master students with a shared background in Information Security and Machine Learning. We are writing to express our joint interest in collaborating on this Master's Thesis project, which focuses on Explainable AI and the use of LLMs as explainers for AI models. We believe that our research background is highly aligned with the requirements of this position, and we are eager to contribute to the exploration and implementation of novel techniques in XAI.

Our research focus has been on the resilience of AI/ML in cyber security, including defending against adversarial attacks on machine learning models. This experience has given us a strong foundation in machine learning theory, adversarial training, and adaptive AI systems. These topics, particularly in enhancing model transparency and resilience, closely intersect with explainability, as both fields aim to enhance the reliability and trustworthiness of AI systems. Moreover, the potential to use LLMs to provide natural language explanations for complex AI decisions resonates with our interest in bridging the gap between technical AI solutions and broader user understanding.

We are confident that our experience with machine learning and our enthusiasm for tackling challenging research problems will allow us to excel in this project. Additionally, we are proficient in Python and PyTorch, have experience with Git, and are motivated to explore new methods such as retrieval-augmented generation and consistency evaluation in the context of XAI.

Thank you for considering our application. We look forward to the possibility of discussing our qualifications further.

Best regards,
Yitian Shen and Zixi Wang