Driven by a personal conviction to eliminate repetitive work, my journey into artificial intelligence began in my Machine Learning course at NTU, where I realized AI was not just a theoretical curiosity but a practical tool for freeing human potential from mundane tasks. This ambition raised an essential question that has since guided my academic and professional pursuits: **how can these powerful techniques be transformed into efficient, reliable, real-world applications with tangible impact?** To find the answer, I immersed myself in both academic research and industry, focusing specifically on **ML and Large Language Model applications**. My career goal is to become an **AI R&D Engineer** at a leading tech company, where I can build intelligent systems that automate complex workflows and streamline routine tasks for people. However, to transition from building novel applications to architecting robust and scalable systems, I recognize the need for a more rigorous foundation in software engineering and system design. Therefore, I am convinced that the Master of Science in Electrical and Computer Engineering in (MSECE) program at Carnegie Mellon University is the essential next step to bridge this gap and achieve my full potential.

My foundational **research experience** progressed from applying traditional ML models to enhance network reliability at WMNLab to using advanced GNNs for security prediction at Academia Sinica, consolidating my belief in ML's power to solve impactful, real-world problems. As LLMs began to proliferate, I saw immense opportunities for optimization and pivoted my focus to this domain in the SPMLLab and RLLab, where **equipped me with LLM architecture concepts and LLM efficiency for real-world applications**. My works centered on a core question: how can we make LLMs more efficient and effective for assisting people in creative tasks? My first project was a multi-agent discussion framework that successfully enhanced LLM creativity, published at COLM 2024. However, I soon identified the practical limitations of this approach, which were high latency and computational cost. This led me to a more refined question: how can we achieve these creative gains efficiently? My subsequent research tackled this directly by exploring representation engineering, achieving a similar creative boost within a single model while drastically reducing computational costs by 95%. This journey from one solution to a more refined one taught me a critical lesson: the best approach is often a trade-off between performance, cost, and scalability under various circumstances. This realization has ignited my desire to master the fundamental principles of AI systems to build truly efficient and reliable AI, a pursuit I aim to deepen through CMU's 18-662 Principles and Engineering Applications of AI .

To gain practical insights into **industrial AI products**, I began with an internship at Compal Electronics, which led to an opportunity to join the AI startup Genibuilder as a **founding AI R&D Engineer**. During my year there, my primary project was building an AI agent to replace aesthetic clinic consultants. This was a multifaceted challenge that encompassed ensuring factual accuracy, maintaining legal compliance, and enabling sales functions and appointment scheduling via robust API and system design. Driven by my principle of automating inefficient workflows, I also tackled a critical internal bottleneck: the time-consuming manual evaluation of our agent. To solve this, I designed an AI system that authentically simulated client-customer dialogues, reducing the team's manual review time by over 80%. These projects taught me that building commercial AI requires a holistic view beyond model performance, and AI could indeed solve real-world problems in many aspects. This industry experience provides a practical perspective that I am keen to enhance in the Master's program at CMU, and I am eager to deepen my knowledge of designing elegant systems through the 18-763 Systems & Toolchains for AI Engineers.

Beyond my strong research and work experience, I apply the same drive for innovation to my academic **coursework**. Recognizing the vast potential of modern AI, I proactively enrolled in courses that could enhance my ability in utilizing LLM in different aspects. For instance, my final project in Programming for Business Computing, where I led the development of an Exchange Student Matching Platform. As project leader, I designed the full-stack system and implemented a KNN matching algorithm to deliver personalized recommendations. My role also involved mentoring teammates through unfamiliar technologies, a leadership experience that solidified our collaboration and resulted in our project earning the **top score** in the class. This project strengthened my skills in end-to-end system design and proved that these concepts are genuinely impactful in real-world applications.

The same **consistency** that drives my passion for AI reflected in my decade-long commitment to basketball. Serving as a point guard on school basketball team from elementary school to college has honed my skills that directly translate to my technical pursuits. Point guard is analogous to that of an engineer: I was tasked with analyzing complex, real-time situations on the court, breaking down problems, and leading my teammates to execute the optimal strategy-acting as the "coach on the court." This experience was essential in developing my abilities in **rapid analytical problem-solving and leadership** under pressure. I am confident that these skills will be invaluable as I tackle complex research questions and collaborate within the diverse community of Master's program.

Combining my hands-on experience with my desire to advance my expertise in building robust AI applications across various domains, I believe the Master's program in Electrical and Computer Engineering at Carnegie Mellon University is the ideal fit for my academic and career goals. I aim to take courses focused on real-world system design and large-scale data processing, which will strengthen my ability to architect the impactful applications I am passionate about. I am particularly drawn to **Professor Sarah Cen's** research on designing and governing AI in diverse domains, which perfectly aligns with my journey of building applications to solve distinct problems in fields like legal tech, medical consulting, and sports analytics. I am excited to be able to contribute to her lab and expand on this vision. I am equally interested in **Professor Beidi Chen's** work on foundation model understanding and reasoning enhancement. This directly connects with my recent research, where I sought to improve LLM creativity but with method inside the model's activations. I believe her guidance would be invaluable in deepening my understanding of model internals and advancing my focus on LLM interpretability. Finally, after my year as an AI engineer, I deeply appreciate the importance of efficient resource allocation, a challenge faced by any company without a dedicated GPU cluster. For this reason, I am interested in **Professor Greg Ganger's** research on systems and cloud computing, which I am confident will provide the foundational knowledge I need to design scalable and cost-effective AI solutions.

In conclusion, I am confident that the training at Carnegie Mellon University will help me reach my full potential and support me in achieving my career goals. My enthusiasm and strong foundational knowledge make me well-suited for this program, and I can contribute to the Master's program in Electrical and Computer Engineering.