Statement of Purpose

(IIT Madras – Summer Fellowship Programme 2026)

I am interested in understanding how large-scale data systems and machine learning pipelines are designed, evaluated, and improved for real-world problems. I am especially drawn to work that brings together machine learning, data engineering, and systems thinking particularly in areas like sustainability, geospatial data, and scientific applications.

I am currently pursuing a B.Tech in Computer Science and Engineering at Amity University Mumbai, with a minor in Economics. During my undergraduate studies, I have focused on building strong fundamentals in core computer science while also looking for chances to apply what I've learned through projects, internships, and working with others. These experiences have shown me the importance of teamwork, careful experimentation, and learning from people who have more experience.

Through internships and project work, I have been exposed to problems in remote sensing, geospatial analytics, natural language processing, and energy-aware machine learning. At the Journal of AI & Knowledge Engineering, I worked with researchers on technical evaluations of research platforms and workflows. In industry internships, I collaborated with teams to build data pipelines using tools like Pandas and NumPy, worked with cloud-based data processing systems, and helped create interactive dashboards for analysing large datasets.

A few academic projects really shaped my interests: hyperspectral image classification, rooftop solar potential mapping using computer vision, sustainability-aware recommender systems, and simulation studies related to planetary defence. Working on these (mostly in teams) taught me how decisions about data representation, preprocessing, and storage formats can have a big impact on model performance, scalability, and interpretability especially when dealing with large or messy datasets.

Lately, I've been thinking about how modern data systems could do a better job of integrating metadata like schema, provenance, and basic statistical information to make analytics more reliable and reproducible. I don't see this as a solved problem, but as an open research question. Through the Summer Fellowship, I'd like to explore these ideas together with others, study existing data formats and system designs, and learn how researchers at IIT Madras evaluate the practical trade-offs in real data-heavy systems.

During the fellowship, I hope to contribute to a clear, well-defined research problem under faculty guidance whether that involves building prototypes, running experiments, or doing empirical analysis. I am particularly keen to learn how research questions are refined, how experiments are designed at scale, and how results are clearly communicated in a research group. I see this as a chance to learn from fellow students, PhD scholars, and faculty while also contributing something useful to ongoing work.

IIT Madras has a strong culture of collaborative research in computer science, data science, and interdisciplinary areas, which makes it a great fit for what I'm looking for. The chance to work closely with researchers on campus, have technical discussions, and be part of a serious academic environment is what really attracts me to this programme.

In the long term, I want to do research in ML systems and data engineering to build scalable, responsible AI. This fellowship would help me develop a stronger research approach and prepare me for future work.