Project Title: Data-analysis tool to ingest high-velocity and high-fidelity system metrics and sensor data from the field-deployed Waggle Sensors

Abstract: The main goal of this project was to design a data-analysis tool to ingest high-velocity and high-fidelity system metrics and sensor data from the field-deployed Waggle Sensors and to provide useful insights on the functioning of a field-deployed sensor array. The goal included designing a pipeline for live ingest and failure analysis of the data and the creation of datasets for future AI/ML tasks for predictive maintenance algorithms. We explored algorithms for predicting future system component failures, and ways to integrate the feedback into standard DevOps procedures to proactively schedule maintenance or change system load to safeguard against failures. A wide variety of system metrics (~70 parameters) and sensor values (~10 physical environments) values were collected from each node every 30-60 seconds. In this effort, we studied the metrics trends in combination with the weather, physical environment, and sensing/computing objectives assumed by the sensor nodes. This project is part of a larger initiative called Sage, a project funded by the National Science Foundation to design and build a new kind of national-scale reusable cyberinfrastructure to enable AI at the edge.