YAML-Driven Evaluation Pipeline

