

Edge-AI Enabled Text Processing System using PYNQ-Z2 FPGA

Sagar Kumar Singh, Nikhil K M, RishiPushkar P, Tarun Hurulikuupi

Dr H R Shashidhara¹

¹Department of Electronics and Communication Engineering, The National Institute of Engineering, Mysuru 570008, India

Emails: 2023ec_sagarkumarsingh_b@nie.ac.in, 2023ec_nikhilm_b@nie.ac.in, 2023ec_rishipushkarp_b@nie.ac.in, 2023ec_b@nie.ac.in, hrshashidhara@nie.ac.in

Abstract—This project presents the design and deployment of a lightweight text-processing web application using the PYNQ-Z2 FPGA development board. The board, running PYNQ Linux, provides an embedded Python environment capable of hosting micro-webservers and performing local computations. Due to hardware limitations (512 MB RAM), the PYNQ board cannot run large machine learning models or transformer-based language systems. Therefore, a lightweight rule-based text-processing system was implemented using Python and Flask, with a simple web interface. The project demonstrates that the PYNQ-Z2 can act as a low-power microserver capable of hosting web applications, processing user input, and serving real-time responses. Future enhancements include converting ONNX machine-learning models to TensorFlow Lite and deploying them on the PYNQ board for offline AI text conversion and speech-based systems.

Index Terms—AI-ML, PYNQ-Z2, Jupyternotebooks, VLSI, Text conversion, Data Collection, Harsh Environments.