

AUSTIN TX AI Inference Software Development Manager

Flex Logix develops industry-leading AI inference engines and we are the #1 provider of eFPGA solutions. Our InferX X1 is the industry's fastest and most-efficient AI edge inference accelerator that brings AI to high-volume applications, surpassing competitor's performance at 1/7th size and much lower price. InferX X1 is available as a chip, PCIe board and M.2 board. InferX X1 is working and running YOLOv3 today and starts production shipments this summer.. Our Inference Compiler is easy to use (we take in neural network models in TensorFlowLite and ONNX) and our APIs for Infer X1 allow rapid integration by the customer. Our architecture is covered by dozens of patents and is highly differentiated giving us a sustainable competitive advantage.

We recently completed a \$55M funding round.

Flex Logix is seeking a Manager to build a team of Inference Software Developers in Austin developing state-of-the-art compilers and run time software for InferX X1. This role reports to our Senior Director of AI Inference Software Development. Since January 40% of our new hires are located in Austin including the VP of Inference Hardware and the VP of Inference Sales/Marketing/Applications; and also our Director of eFPGA Software Development. Our Austin office will be opening shortly. This is an exceptional opportunity to develop the technology that breaths life into AI inferencing solutions targeting systems in medical, industrial, automotive and other Enterprise edge applications.

RESPONSIBILITIES

Manage about a third of our growing Inference Compiler team, reporting to the Senior Director, which is responsible for our Model Compiler for our Reconfigurable Tensor Processor: a DNN Model-to-binary flow:

- Build a team in Austin to supplement our existing team in Mountain View CA
- Lead the Austin team in expanding functionality of our Model Compiler, written in modern C++, for support of additional capabilities and computation optimizations
- Understand the required feature content and delegate implementation of parts to team members
- Own program management, create schedules for designs and implementations and track the team progress on a weekly basis with updates to management
- Conduct design and code reviews, promote communication and good design practices across the team
- Work with related functional groups: architecture, solutions/applications, Softlogic, etc.

This is a hands-on management role, so in addition to understanding software you need to have basic understanding of computer architecture, digital logic and Deep Neural Networks.

Read more about nnMAX Reconfigurable Tensor Processor and InferX on our Inference page at www.flex-logix.com

EXPERIENCE AND SKILL REQUIRED

8+ years of relevant industry experience and at least 3 years as a manager of a team of 3-8 software developers. AI/NN expertise is not required but desired.

Previous experience of leading the team or individual contribution into development of system-level software or embedded systems in Windows or Linux in C++ is required.

Excellent understanding of software/hardware development engineering practices, workflow, schedule dependencies.

Must be passionate about being part of an aggressive, venture-backed startup team that is changing the way chips are architected, designed, and programmed

Must be entrepreneurial, innovative problem solver and willing to work hard.

Must live in Austin TX. Strong preference for US citizenship or permanent residency (“green card”); will consider candidates with current H1-B visas who are willing to transfer promptly.