LF AI & DATA

Streamlining Neurodivergence Diagnosis with OPEA

OPEA

SJSU SAN JOSÉ STATE UNIVERSITY

ORGANIZATION

This case study is led by Hashem Jaber, a Master's candidate in Artificial Intelligence at San José State University, who is developing comprehensive end-to-end systems designed to diagnose, evaluate, and provide personalized treatments for individuals with neurodivergence. The project integrates various technologies, including FuriosaAI's advanced hardware for real-time inference, Intel's audio stack, Coze's orchestration tools, and facial-language modeling expertise.

At the core of this ecosystem is the **Open Platform for Enterprise AI (OPEA)**, an LF AI & Data project that serves as the orchestration backbone for the system, enabling modular, traceable, and hardware-optimized AI services.

CHALLENGE

Traditional neurodivergence screening and follow-up care, which involves lengthy evaluations by specialists, often delays diagnosis and treatment, especially for children, teens, and adults with conditions like ADHD or Autism Spectrum Disorder (ASD). This case study addresses the challenge of shortening this pathway using an open-source, multi-modal AI agent. By passively analyzing language, facial cues, and interaction patterns, the AI flags individuals who may benefit from earlier diagnosis and personalized treatment plans.

The system depends on real-time, low-latency coordination across multiple AI services. To meet clinical-grade response times while remaining auditable, OPEA's orchestration capabilities are leveraged to manage service chaining, performance monitoring, and hardware-aware optimization.

SOLUTION

The system uses Human-Robot Interaction (HRI) for the diagnosis of neurodivergence, analyzing post-interaction data from conversations. Real-time systems for HRI demand enhanced modularity, traceability, and optimized performance across the diagnostic pipeline. OPEA orchestrates modular services, such as facial metrics extraction and text-to-speech (TTS) generation, ensuring system performance is both scalable and adaptable.

The AI analyzes natural dialogue, facial micro-expressions, eye-gaze patterns, and vocal prosody, comparing these with control corpora to flag potential behaviors of interest. For instance, if a 15-year-old's vocabulary is more typical of a 10-year-old, the AI may identify a potential developmental delay and recommend personalized activities.

Beyond language, the platform tracks blink rate, gaze direction, and facial expressions, offering deeper insights into possible ASD traits. The system validates these methods by benchmarking tweets from neurodivergent and control users and evaluating large language models (LLMs) in an actor-critic framework.

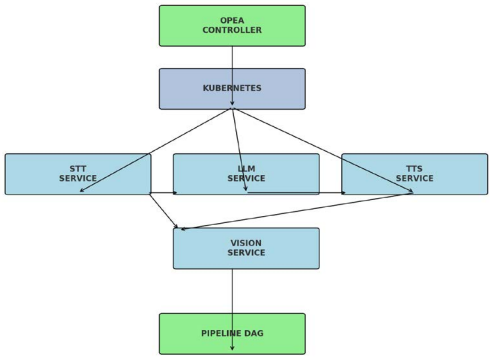

Figure 1: *Diagram of the flow between API/services and the system controller, with LLM services provided by FuriosaAI.*

SOLUTION *(continued)*

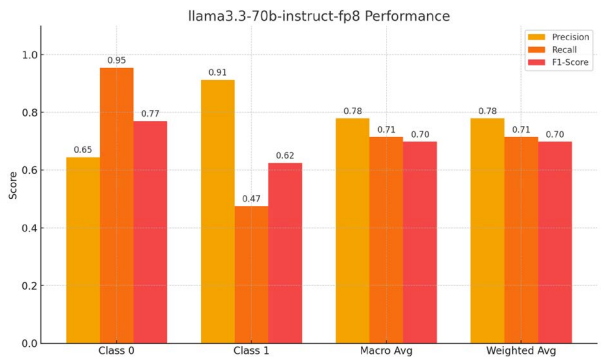

Figure 2: *Bar chart showing the precision, recall, and F1 scores for the Llama 3.3-70b-instruct-fp8 model.*

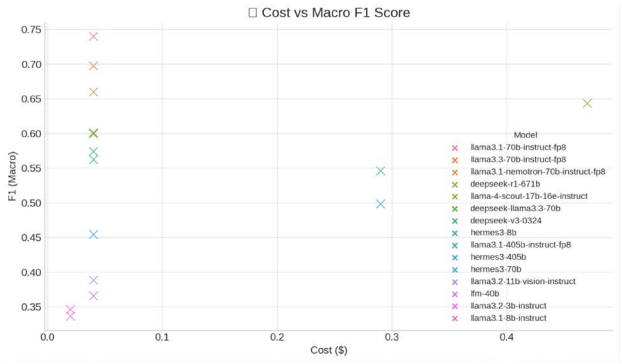

Figure 3: *Scatter plot comparing the inference cost against the macro-averaged F1 score for 13 open-source LLM variants.*

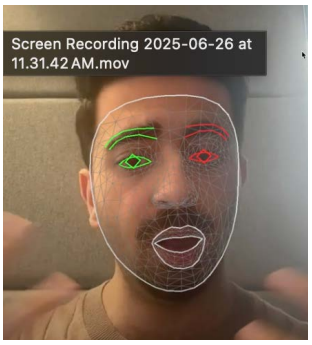

Figure 4: *Screenshot of real-time expression tracking, illustrating facial metrics like blink rate, gaze direction, and lip movement, captured by MediaPipe Face Mesh for neurodivergence analysis.*

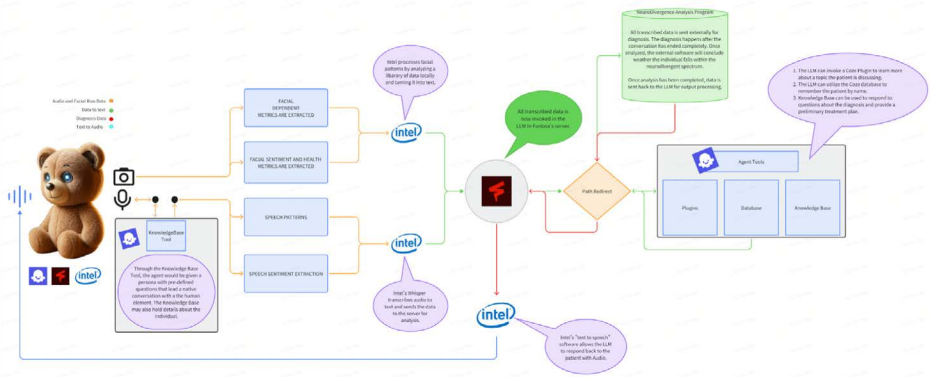

Figure 5: *A plush, camera-equipped companion captures audio and video, which are processed by four Intel modules to extract facial metrics, gauge sentiment, analyze speech patterns, and derive speech sentiment. The data is then sent to the central FuriosaAI LLM for response generation.*

LESSONS LEARNED

- Cost Efficiency: Small models such as Llama 3 (fp8) provided the best cost-to-accuracy ratio, outperforming larger, costlier models.
- Latency Requirements: The system must maintain conversational latency under 300 milliseconds to feel natural, posing a constraint on the ASR > LLM > TTS pipeline.
- Data Privacy: Data-privacy measures were critical for Institutional Review Board (IRB) approval without compromising data utility.

FUTURE PLANS

- Benchmarking: The team plans to release a multi-modal benchmark for tweets, short-form videos, and annotation guidelines to encourage community contributions.
- Expanded Diagnostics: The classifier will be extended to detect sensory-sensitivity cues (e.g., light, sound, texture), mapped directly to DSM-5 criteria, broadening diagnostic coverage.
- Clinical Alignment: The **DocLing** framework (another project under LF AI & Data) will be used to structure DSM-5 criteria into machine-readable components, improving the system's diagnostic accuracy and alignment with clinical standards.

FUTURE PLANS *(continued)*

- Non-Verbal User Support: Future plans include supporting sign-language gesture recognition based on MediaPipe Holistic.
- Edge Deployment: The system aims to reduce cloud dependency by porting both vision and lightweight language models onto FuriosaAI's edge NPU, reducing latency.
- Clinical Pilot: The team plans a clinical pilot in partnership with the San José State University psychology department to validate the system's performance against gold-standard diagnostic interviews.
- By leveraging OPEA for orchestration and performance monitoring, this initiative aims to revolutionize neurodivergence diagnostics, enabling faster interventions and personalized care at scale.