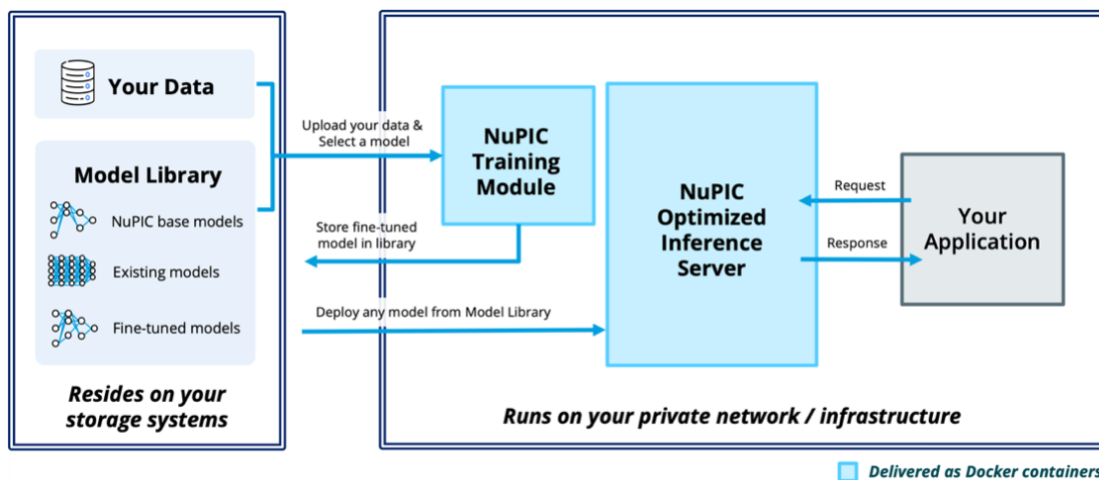




Numenta Platform for Intelligent Computing (NuPIC™)

EASILY DEPLOY AI MODELS ON CPUS WITH NUPIC

NuPIC™ is an AI software platform that draws on Numenta's decades of neuroscience research to enable the deployment of Large Language Models (LLMs) on CPUs with simplicity, scalability, and speed. NuPIC comes with a highly optimized inference server and production-ready pretrained models that can be customized to specific NLP use cases. Deployed as Docker containers, NuPIC operates entirely within a customer's infrastructure, ensuring 100% privacy, reduced latency, lower costs, and easier compliance. It is designed to allow any developer to get up and running easily, making it an ideal solution for LLM deployment.



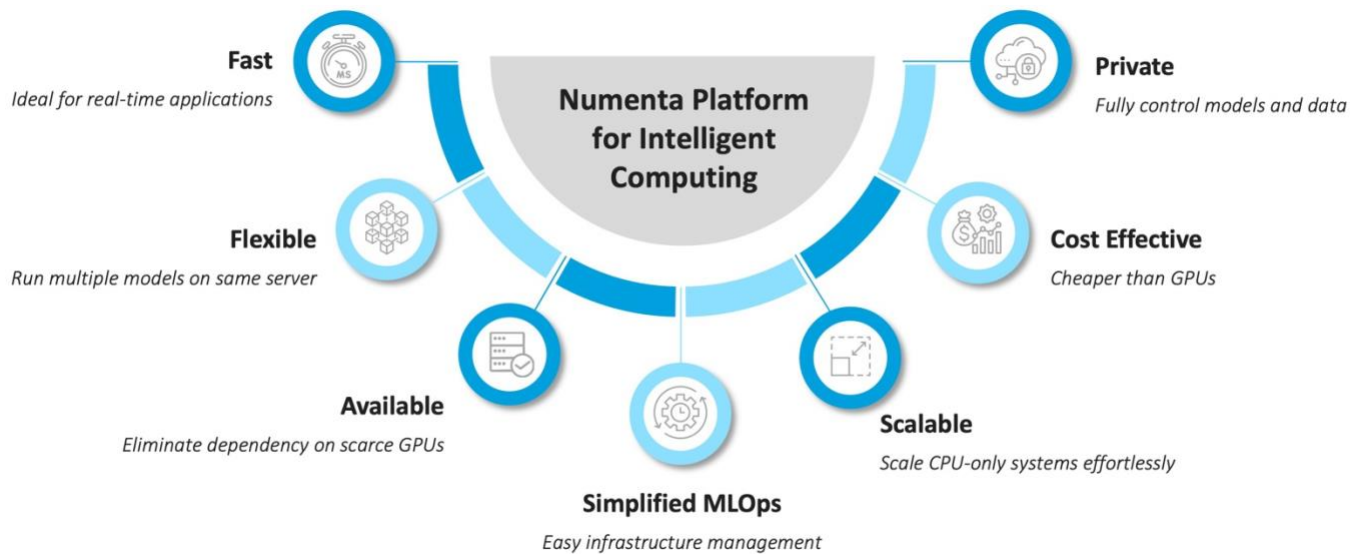
NuPIC can run entirely on CPUs

Use Case	Description	Applications
Sentence Similarity	Compare semantic similarities between different sentences and how they relate to each other.	Contract analysis, recommendation systems, plagiarism detection, etc.
Document Classification	Sort and categorize documents into predefined groups based on different subject matters.	Organizing large datasets, email filtering, content management, etc.
Sentiment Analysis	Analyze opinions or emotions within text and classify as positive, neutral, or negative.	Market research, social media monitoring, customer feedback analysis, etc.
Question-Answer	Provide accurate responses to user queries based on a knowledge base or previous interactions.	Automated chatbot responses, technical support, information systems, etc.
Retrieval Augmented Generation	Enhance text generation by retrieving and summarizing relevant data from a database.	Market analysis, customer service enhancement, resource integration etc.
Summarization	Condense lengthy documents and articles into concise, informative summaries.	Information retrieval, customer support, content curation, etc.

EXAMPLE IMPLEMENTATION

A business with thousands of daily customer inquiries can use NuPIC BERT models for general inquiries to automate accurate, context-aware responses to common questions such as “Where is my order?” or “How do I return this item?” They can also use sentiment analysis to categorize inquiries based on emotional tone, where escalations can then be directed to the relevant support staff. Subsequently, NuPIC GPT models can provide the Operations team with brief summaries of flagged tickets, allowing them to focus on complex cases and streamline customer support.

BENEFITS OF NUPIC ON CPUS



TECHNICAL CONFIGURATIONS

Models:	NuPIC BERT, NuPIC GPT (similar to LLAMA2)	Operating System:	Recent version of Linux (Ubuntu 22.04) is required to ensure kernel support for AMX
Minimum Processor:	Any Intel-compatible server with AVX512/AVX2/AMX instruction set support	Software:	Docker, Python 3.8 or later
Recommended Processor:	AMX enabled server such as AWS m7i.4xlarge	Training:	GPU with at least 12GB RAM
Memory & Storage:	16GB RAM, 200GB of storage space		

GET STARTED TODAY

Discover the power of CPU-based AI with NuPIC. Contact us for a demo to discuss how our solutions can meet your business needs: numenta.com/demo



889 Winslow Street, 4th Floor
Redwood City, CA 94063

info@numenta.com

As a world leader in deploying large AI models on CPUs, Numenta has mapped its neuroscience-based advances to modern CPU architectures to redefine what’s possible in AI. Numenta’s AI platform NuPIC, the Numenta Platform for Intelligent Computing, helps businesses leverage the flexibility of CPUs to build robust AI applications that are efficient, scalable, and secure.