

Company Fact Sheet 2020



ABOUT US

Neural Magic exists to expand what's possible with machine learning. We're building a general purpose engine for machine learning, enabling ML to be practically run in new places, on new kinds of workloads. Neural Magic simplifies deep learning performance by turning x86 Intel CPUs into high performance machine learning compute resources. Neural Magic is a software solution for deep learning acceleration, enabling companies to use ubiquitous and unconstrained CPU resources to achieve performance breakthroughs, at scale, with all the flexibility of software.



COMPANY VISION

No-Hardware AI, or shattering the hardware barriers holding back the field of machine learning. Neural Magic is making the power of deep learning simple, accessible, and affordable for anyone. As a part of this next great unlock of machine learning, data scientists will ask bigger questions, challenge norms, and unleash a new class of AI applications that live at the edge of imagination.



Established: 2018

Employees: 25

Industry: Enterprise Software

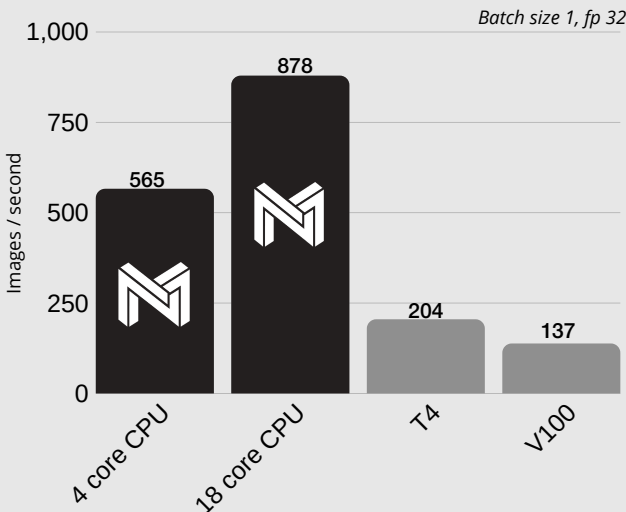
Company type: Machine Learning Infrastructure

OFFERINGS

NEURAL MAGIC INFERENCE ENGINE

Neural Magic Inference Engine software lets data science teams use ubiquitous CPU resources to achieve machine learning performance breakthroughs, at scale, no expensive hardware required. Neural Magic turns the CPUs you already own into high performance machine learning resources.

COMPUTER VISION: Image Recognition



USE CASES

Image Classification

Object Detection

Recommendations

NEWS

“An inference engine that data scientists can run on computers running CPUs, rather than specialized chips like GPUs or TPUs. Greatly reduce the cost associated with machine learning projects by allowing data scientists to use commodity hardware.”

- **TC**
TechCrunch

“Neural Magic proves that high performance execution of deep learning models is ... a systems engineering problem that can be solved with the right algorithms in software.”

- **Venture Beat**

SIMPLIFY PERFORMANCE

Start with Neural Magic pre-optimized models or put our recalibration tools to work on your own models to maximize performance with little effort.

REDUCE COSTS

Say goodbye to expensive GPUs and hardware accelerators. Deploy models on cheaper and more commonly available CPU servers. Run model pipelines on a single CPU instance.

RUN ANYWHERE

Integrate seamlessly into existing deployment pipelines. Deploy in containers or virtual machines. Manage with Kubernetes like any modern software application. Run models anywhere: on premise, in the cloud, or at the edge.

**MULTICORE
COMPUTING
SOFTWARE
PIONEERS**

9 PhDs
60+ patents
200+ published papers
20+ startups
1 NeurIPS Spotlight paper
Gödel & Dijkstra prizes

INVESTORS

COMCAST
VENTURES

pillar

NEA

ANDREESSEN
HOROWITZ

a. amdocs