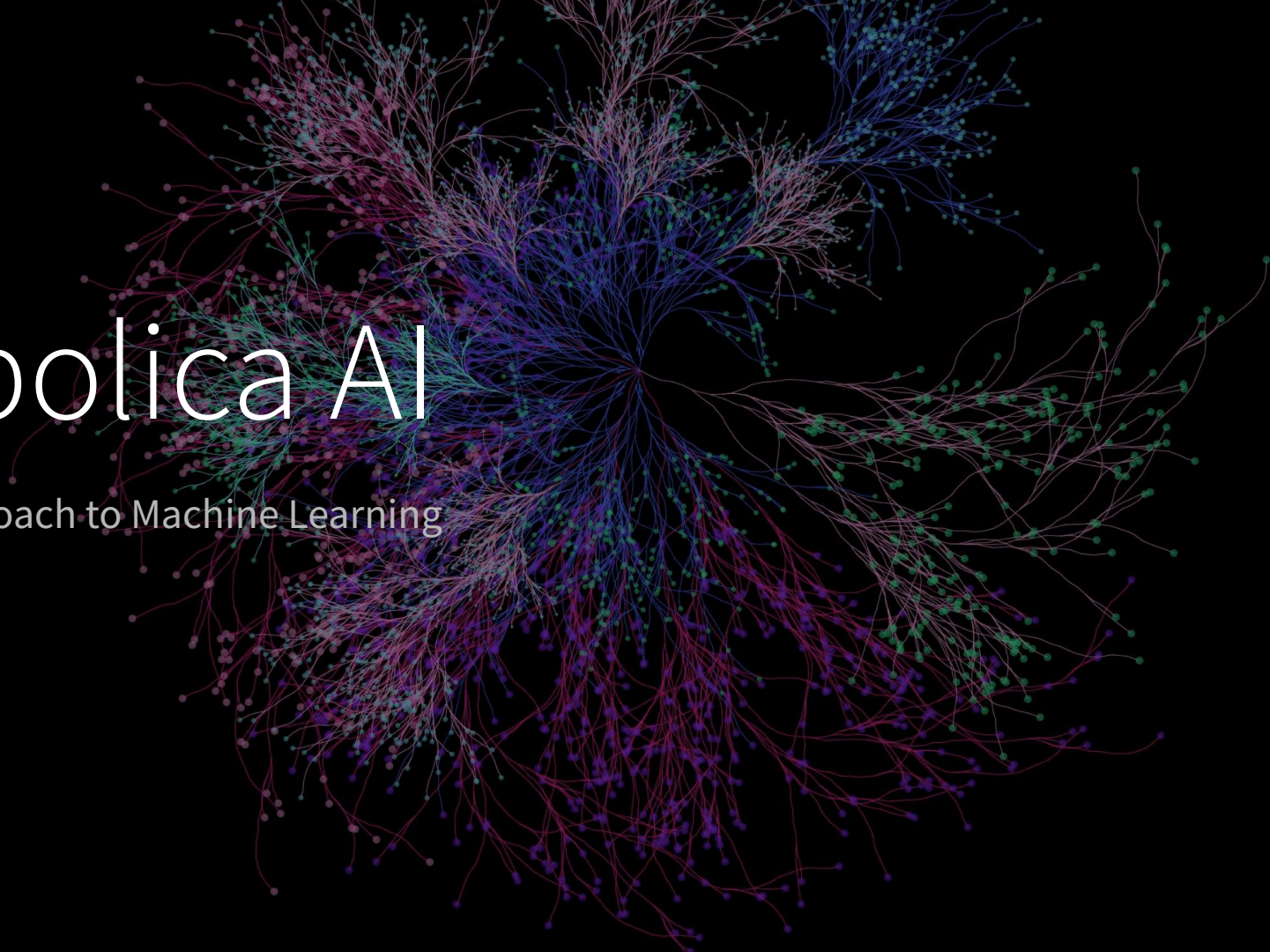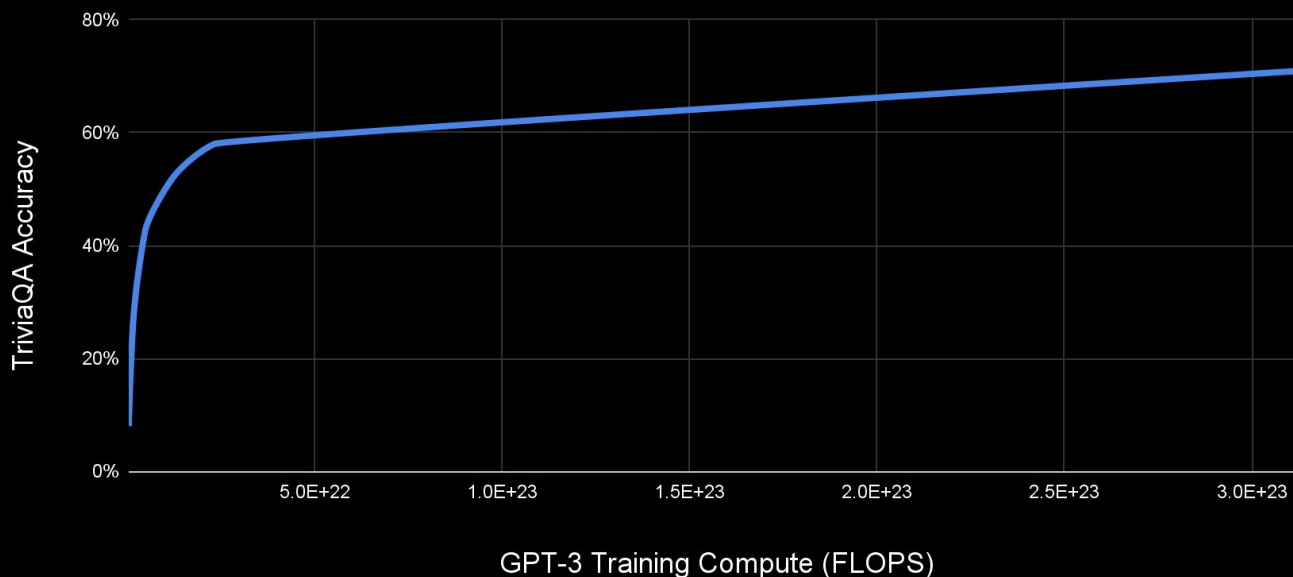# Symbolica AI

A Symbolic Approach to Machine Learning

Compute and Data in ML

Neural networks have diminishing returns on accuracy as training cost increases

Training GPT-3

- State-of-the-art deep learning neural network (DLNN) for generating human-like speech

- 175,000,000,000 hyperparameters

- Model takes > 700 GB of storage

- Training has been estimated* to take:

  - 3.14×10^{23} FLOPS to train

  - 355 GPU years on an Nvidia Tesla V100

  - $4.6M on competitively priced cloud compute platforms
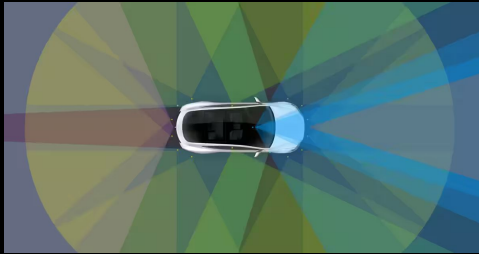
GPT−3

Neural Network Strengths and Weaknesses

- Neural networks are great at data extrapolation
    - Finance
    - Market data
    - Medical records

- Neural networks struggle with decision making and reasoning
    - Autonomous transportation
    - Robotics
    - Natural language processing

Total Addressable Markets

Autonomous Transportation

$94.46B in 2020
~$1.8T by 2030, projected

Robotics

$41.7B in 2020
~$81.4B by 2028, projected

NLP

$11.6B in 2020
~$35.1B by 2026, projected

https://www.precedenceresearch.com/autonomous-vehicle-market
https://www.zionmarketresearch.com/inquiry/industrial-robotics-market
https://www.marketsandmarkets.com/Market-Reports/natural-language-processing-nlp-825.html

Problems in AI requires ML + Distributed Systems Experience

George Morgan

- 3 years of system software at Meta Reality Labs and Apple
- 4 years of AI/ML for Tesla Autopilot
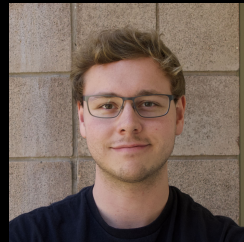
Taylor Wrobel

- 10 years of building scalable distributed systems and tooling at Palantir, Apple, GitHub, and OpenSea

+ A Ph.D. in Theoretical CS with 5+ years of industry experience and a senior ML engineer at industry-leading autonomous systems company to join after funding

Symbolic AI: ML's Past and Future

Symbolic Approaches Aren't New

- Symbolic artificial intelligence was first explored ~60 years ago

- Have struggled to gain widespread adoption

- Traditional approaches focus on *assigning meaning* to symbols
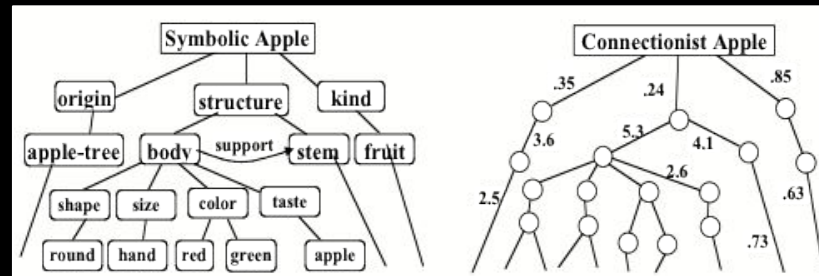
Symbolica's Approach is New

- Symbolica learns by transforming data into abstract terms which obey a formal logic

- As these terms are reduced, a model architecture **emerges** naturally

- Common features then exhibit topological clustering in the emergent space

Symbolica Better Leverages GPUs

- Statistical approaches require floating point arithmetic (e.g. TF32)

- Symbolica is *discrete* and can leverage simpler data types (e.g. INT4)

- GPUs are ready for Symbolica, with INT4 offering **8x** the performance of TF32

Nvidia Ampere A100

Peak TF32 Tensor Core	312 TFLOPS
Peak INT4 Tensor Core	2,496 TOPS

https://developer.nvidia.com/blog/nvidia-ampere-architecture-in-depth/

	Neural Networks	Symbolica
Data	Require **clean, labeled data** to train	Fully **unsupervised**, no labels are required to train
Scale	Adjusting network weights requires **large amounts of training data**	Symbols are constructed using relatively **little training data**
Compute	Uses the **most expensive operation** (TF32)	Uses the **least expensive operation** (INT4)
Model	Must be **hand engineered** from priors about the training data	Emerges **automatically** from the training data
Online learning	**Cannot learn** online due to dependence on labels	Can **continue to learn** online due to being fully unsupervised

Symbolica by the Numbers

Trained on MNIST, a standard ML benchmark, Symbolica models:

- utilize **8x the performance** of modern GPUs

- use **20% of the memory** footprint on average

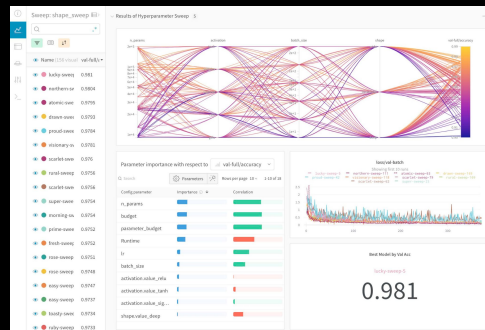- require **99.7% less data** to achieve equivalent accuracy

Extrapolating Symbolica to GPT-3

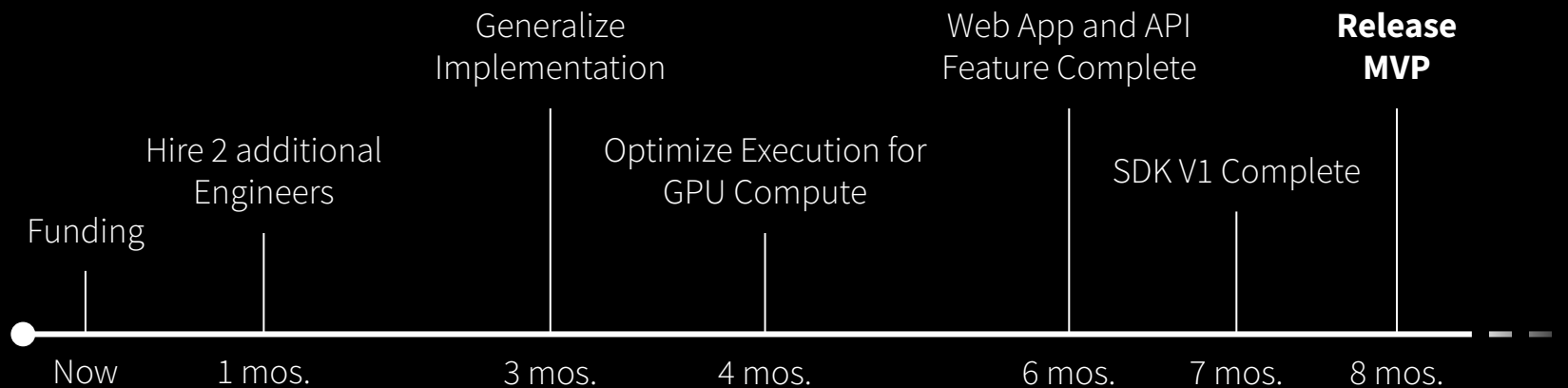How would a GPT-3 scale Symbolica model compare to state of the art neural networks?

Model	Training Duration	Model Size	Cost
GPT-3 at FP32	355 V100 years	700 GB	**$4,600,000**
Symbolica at INT4	7 V100 weeks	140 GB	**$1,725**

MVP Functionality

- Web App and API for training, labeling, and testing Symbolica models

- SDK, similar to PyTorch/TensorFlow

- Free to use for small scale; enterprise licensing for large models and extra features

  - Pricing tiers based on model size and request rate

  - Ability to export models for use in edge computing

  - Personalized models through online fine-tuning

MVP Roadmap

Generalize
Implementation

Web App and API
Feature Complete

**Release
MVP**

Hire 2 additional
Engineers

Optimize Execution for
GPU Compute

SDK V1 Complete

Funding

Now   1 mos.   3 mos.   4 mos.   6 mos.   7 mos.   8 mos.

Looking to raise a $1.5M - $2M seed

Annual Burn Rate

Expense	Cost / Unit / Year	Units	Total Annual Cost
Engineer - AI/ML	$225,000	1	$225,000
Engineer - Math	$200,000	1	$200,000
Office Space	$115,000	1	$115,000
Founder Salaries	$90,000	2	$180,000
Compute Infra	$200,000	1	$200,000
Employee Benefits	$10,000	4	$40,000
Vendors & Licenses	$10,000	1	$10,000
IP Protection	$30,000	1	$30,000
Total			**$1,000,000**