The Team

Jeremy D. Martinez, M.D. is an addiction psychiatrist and the CEO of Neurotrophic Labs, which develops data innovations for healthcare. His work has included the management of financial data and analysis of clinical outcomes.

Jair Wuilloud, Ph.D. is a particle physicist and data scientist CTO of Neurotrophic Labs. He has extensive experience solving complex problems with data and algorithms. His professional work includes extensive use of NLP, neural networks, machine learning (ML) methods, "AI-based chatbots" and he has built real systems and libraries from scratch. His research focus is on high dimensional computing.

Problems to be Solved

The Data Problem

- Large and unwieldy datasets
- Sparse and noisy historical data
- Data security

The Usefulness Problem

- Producing data that is *useful* to M.D.
- Giving results that are *timely*

The Physician User-experience (UX)

- Intuitive and easy usage
- Transparency (How did the system get these results?)
- The opportunity to gain physician knowledge and feedback

Our Approach - Light and Fast

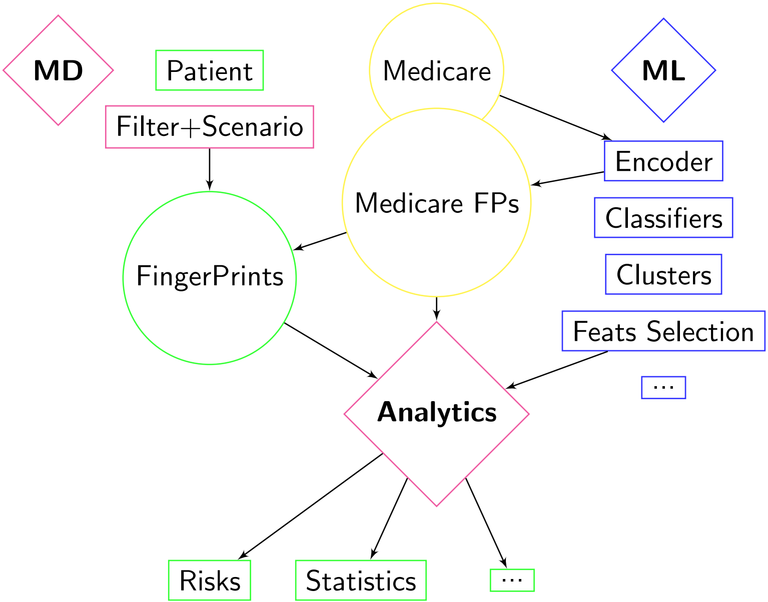
Fingerprints for Sparse, Historical Personal Data

The Model: A novel approach, modeling patient data with *High Dimensional Fingerprints* combined with established ML techniques. Fingerprints are compact, binary representations of data that allow fast computing, data security, and remain robust to noise and superposition. The model is a *Cognitive System* that learns from the interactive user interface (UI) and data.

Advantages

- Binary patient data is compact, with computationally efficient *encoding* and *retrieval*
- Intuitive operations with simple notions of similarity, aggregation, filtering, and simulations
- Learning from MD's for optimization, while creating trust

AI Strategy Overview: Patients as Fingerprints



A Cognitive System

HD Computing: large ($N \geq 10000$), binary vectors (HDV's)
We build a Symbolic Vector Architecture with simple algebra:

- Sets: $\| \sum_i v_i^{HDV} \|$
- Records: $\| \sum_i v_i^{HDV} \star w_i^{HDV} \|$
- Sequences: $\| \sum_i \Pi_i \star w_i^{HDV} \|$

with $\| \ldots \|$ the normalization to binary vectors, $\star$ the XOR product and Π_i random permutations.

Fingerprints: Unlike vector embedding, HD embeddings are robust under superposition, allowing use through multiple NN layers.

Using HD Fingerprints, we can compress historical, sparse, multidimensional data while maintaining data security

We Exploit the following Features of HD Computing:

- Compression and embedding into fingerprints
- One-shot Learning
- Faster learning benchmarked against classical ML accuracy
- High-performance and economical computing
- Can be combined with standard ML (NNets, ...)

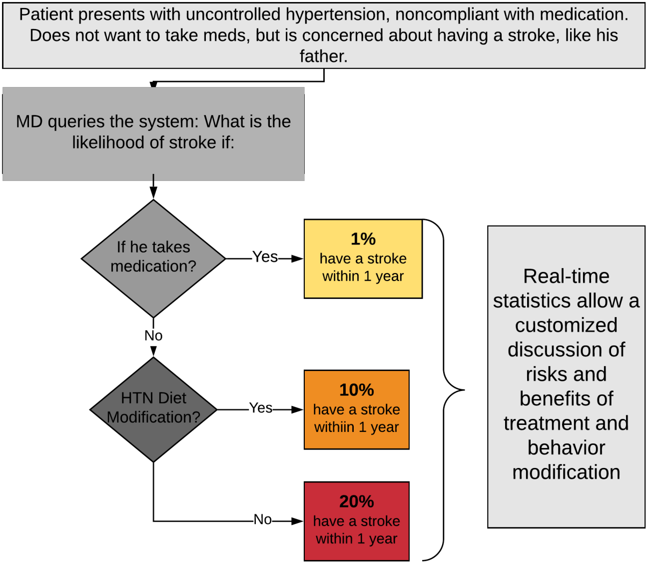
Support for HD Computing techniques (a selection):

Kanerva, P. (2009). Hyperdimensional computing: An intro. to computing in distributed representation with high-dim. random vectors. *Cognitive Computation*, 1(2), 139-159.

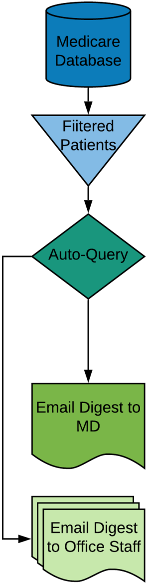
Rahimi, A., et al. (2017). "High-Dim. Computing as a Nanoscalable Paradigm." *IEEE Transac on Cir and Sys I: Reg Papers* 64(9): 2508-2521.

Mitrokhin, A., et al. (2019). "Learning sensorimotor control with neuromorphic sensors: Toward hyperdim. active perception." *Science Robotics* 4(30).

M.D. Workflow 1 - Real-time Patient Care



M.D. Workflow 2 - Automated Alerts



Ongoing Updates to System from Medicare Database

Each MD has a list of patients under treatment

Custom Filters

Outcomes of Interest
- ✓Hospitalization
- ✓New Diagnosis
- ✓Surgery
- ✓Other Outcome

Timing of Outcomes
- ✓Next 30 Days
- ✓Next 90 days
- ✓Next Year, 5 Years
- ✓Lifetime Risk

Custom Notifications

Patients at Risk for Adverse Outcome
- ⚠Hospitalization
- ⚠ Surgery
- ⚠ Death, etc.

Patient Adverse Outcome Reported
- ⚡ Hospitalized
- ⚡ New Diagnosis
- ⚡ Procedure/Surgery

Patients needing New Appointment
Customized email alerts are sent to Office Staff to call patient to set an appointment. For example, if Risk of Hospitalization is >50% in the next 30 days and has not been seen in 3 months

Training The Model

Strategy

- Start with random HDVs and gradually improve model
- MD's full query histories are stored and encoded
- MD gives recs on queries for improvement feedback loop

Data Processing

- Clustering, feature selection, better embedding, NNs, ...
- Multi-turn, multi-dimensional; gradually automates queries

Identified Data Sources

- Launch Stage: 2008-2010 CMS Public Files (DE-SynPUF)
- Stage I, II: CMS Limited Data Set (LDS) Files
- Future: CMS Research Identifiable Files (RIFs)

Development Timeline

Launch Stage

1. Warm up and benchmarks with light fingerprint library
2. Comparison of HD Fingerprints vs. standard ML

Stage I

1. Modeling, database, hashmaps, similarities, and superficial analytics performed on LDS data set
2. Optimize fingerprint code for search, clustering, security
3. Multi-turn, multi-dimensional recs with one-shot learning

Stage II

1. Select features, enrichment with clusters and predictors
2. Integrate MD interaction data and larger datasets