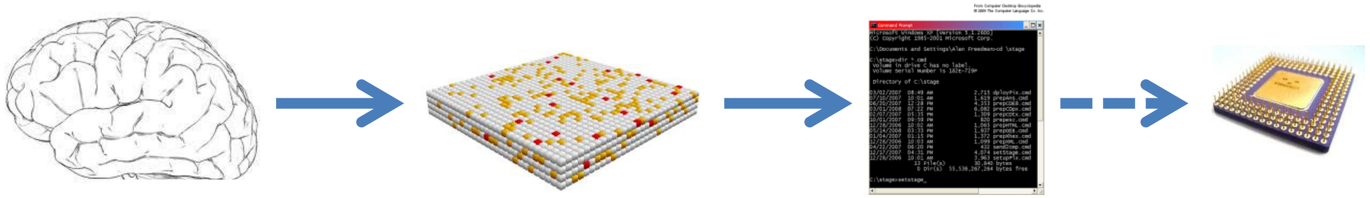
Computing Like the Brain

StrangeLoop 2012

Jeff Hawkins

1) **Discover operating principles of neocortex**

2) **Build systems based on these principles**

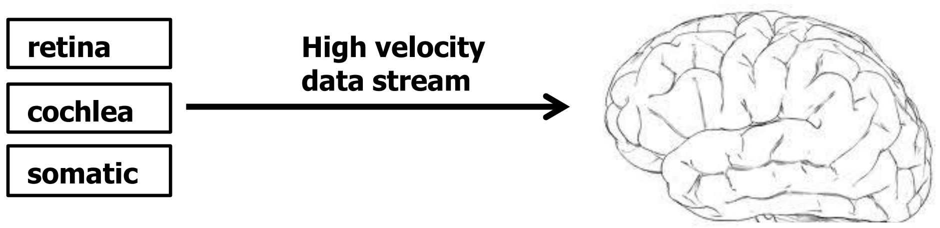


Anatomy, Physiology, | **Theoretical principles** | **Software** | **Silicon**

Talk Agenda

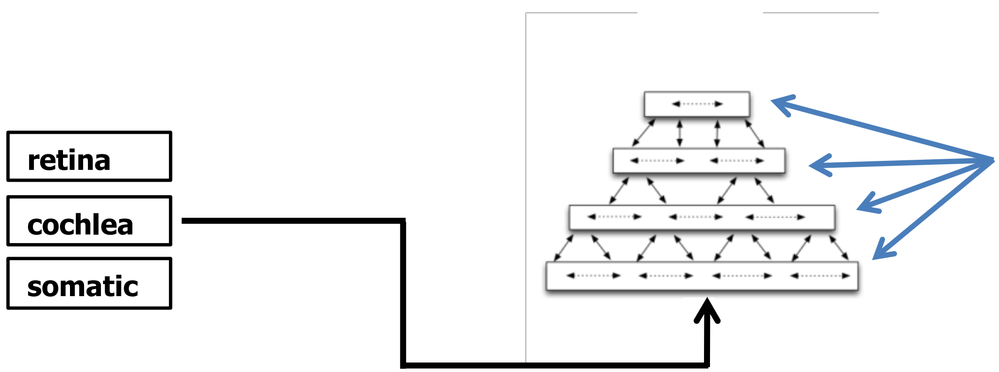
- The brain as a predictive modeling system
    - Sparse Distributed Representations
    - Sequence memory
    - Online learning
- Grok, a predictive modeling product
- Future of machine intelligence

The neocortex is a predictive modeling system



The neocortex builds online models from streaming data

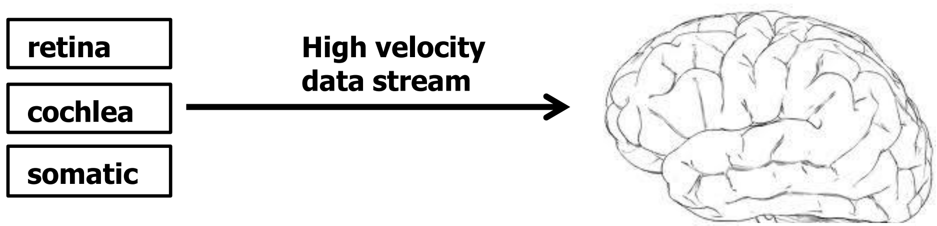
- predictions
- anomaly detection
- actions



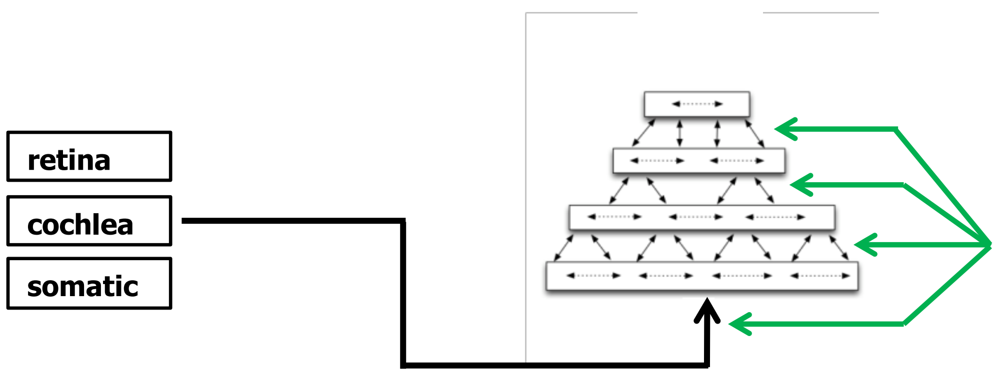
1) Hierarchy

2) Sequence memory

The neocortex is a predictive modeling system



The neocortex builds online models from streaming data
- predictions
- anomaly detection
- actions

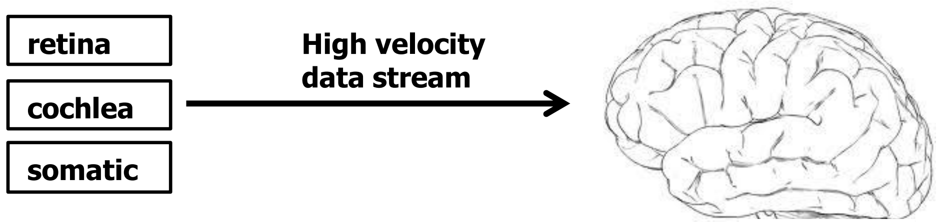


1) Hierarchy

2) Sequence memory

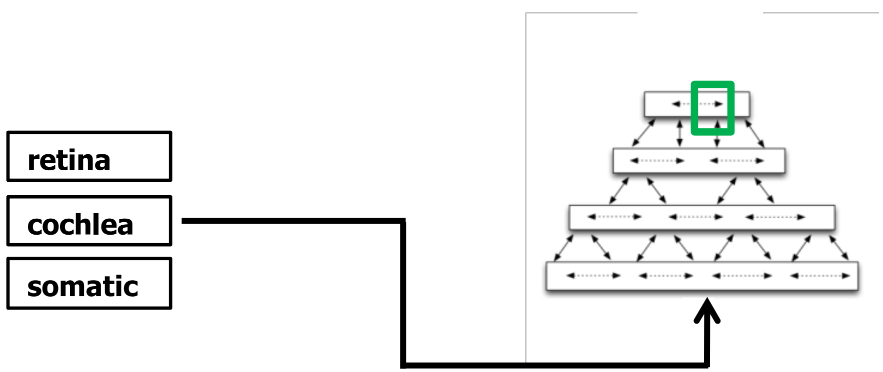
3) Sparse Distributed Representations

The neocortex is a predictive modeling memory system



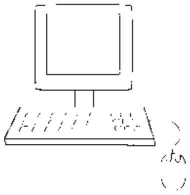
The neocortex builds online models from streaming data

- predictions
- anomaly detection
- actions



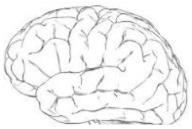
1) **Hierarchy**
2) **Sequence memory**
3) **Sparse Distributed Representations**

Dense Representations



- Few bits (8 to 128)
- All combinations of 1's and 0's
- Example: 8 bit ASCII
  01101101 = m
- Individual bits have no inherent meaning
- Representation is assigned by programmer

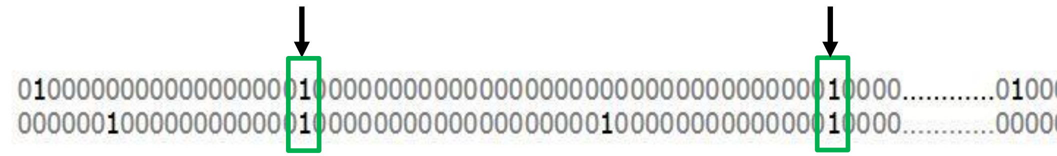
Sparse Distributed Representations (SDRs)



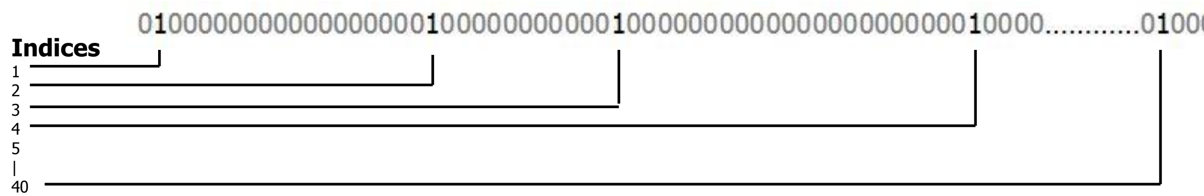
- Many bits (thousands)
- Few 1's mostly 0's
- Example: 2,000 bits, 2% active
  0100000000000000000100000000000000000000000000000000010000…………01000
- Each bit has semantic meaning
- Meaning of each bit is learned, not assigned

SDR Properties

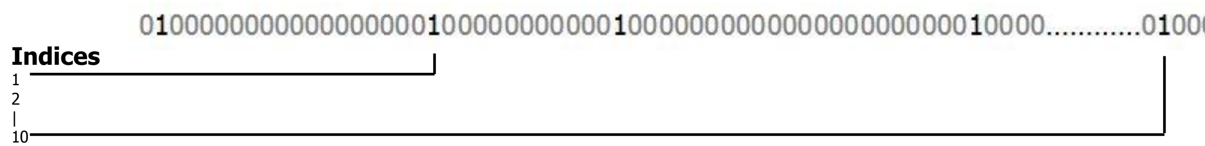
1) **Similarity**:
shared bits = semantic similarity



2) **Store and Compare**:
store indices of active bits



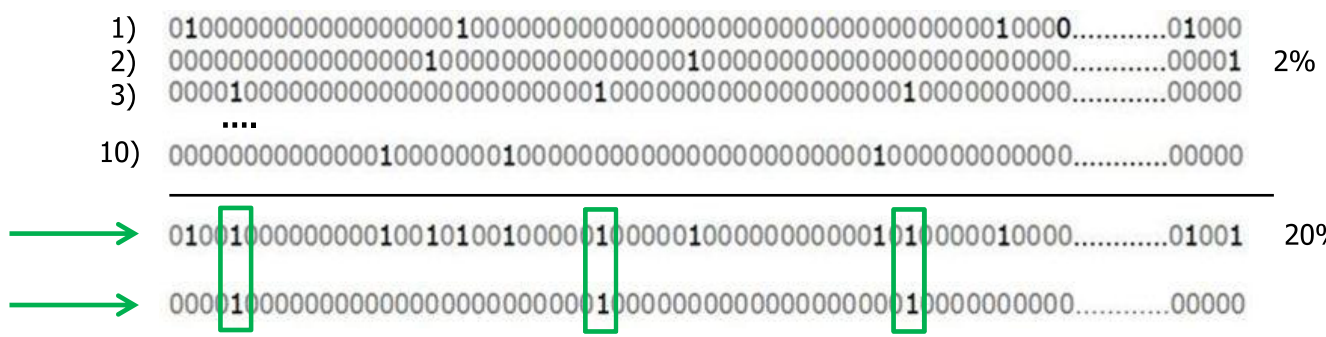
subsampling is OK



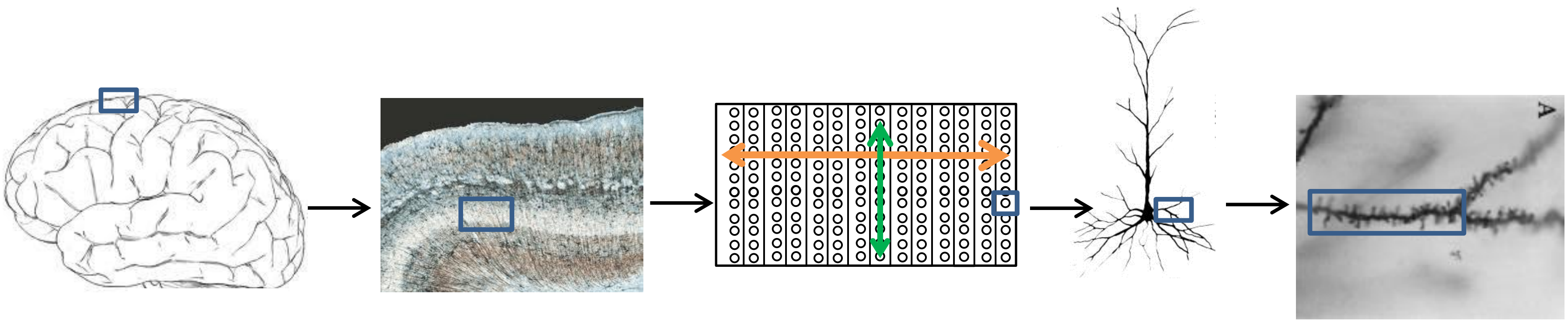
3) **Union membership**:

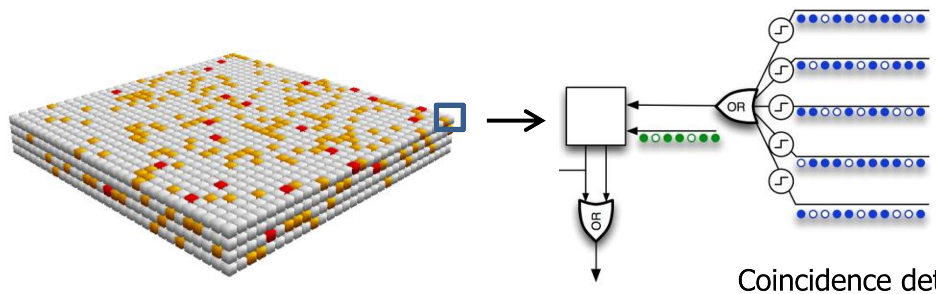
Union

Is this SDR
a member?



Sequence Memory

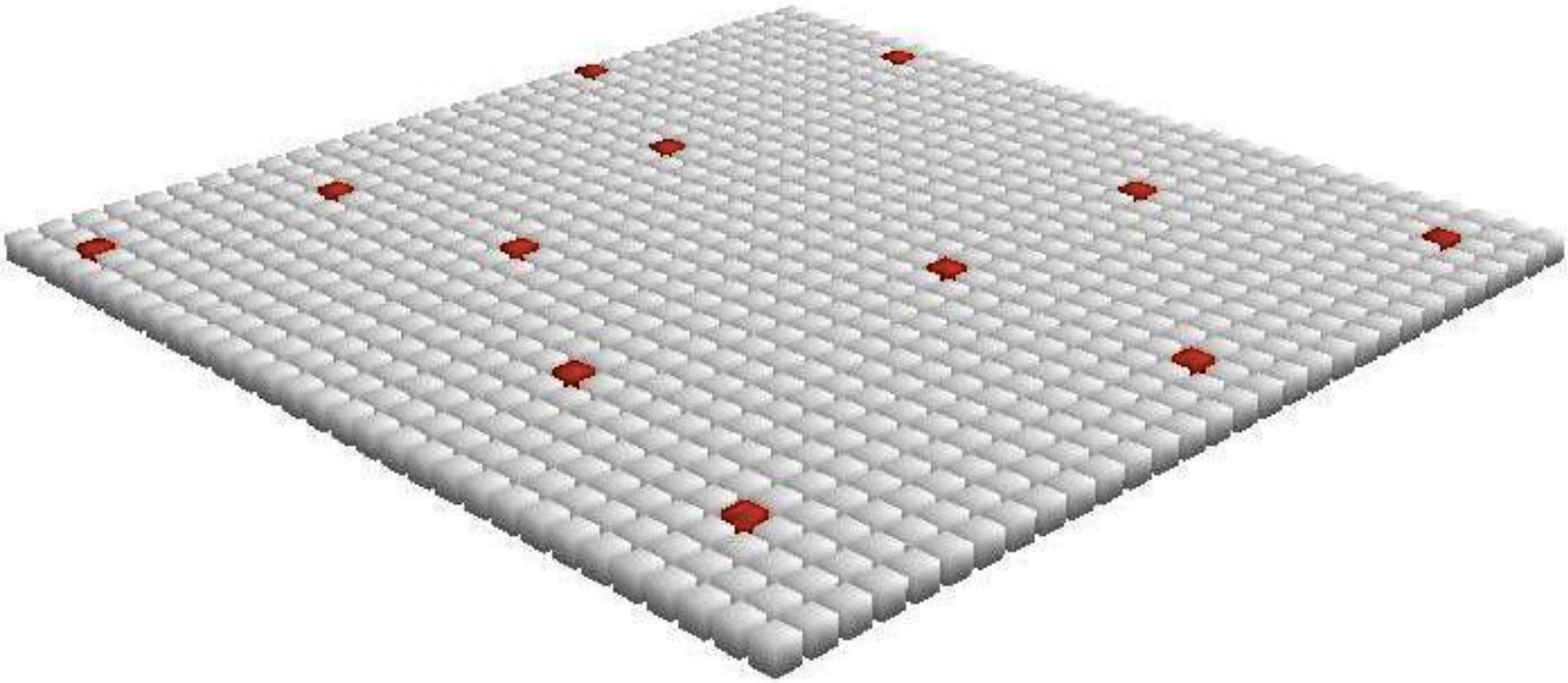




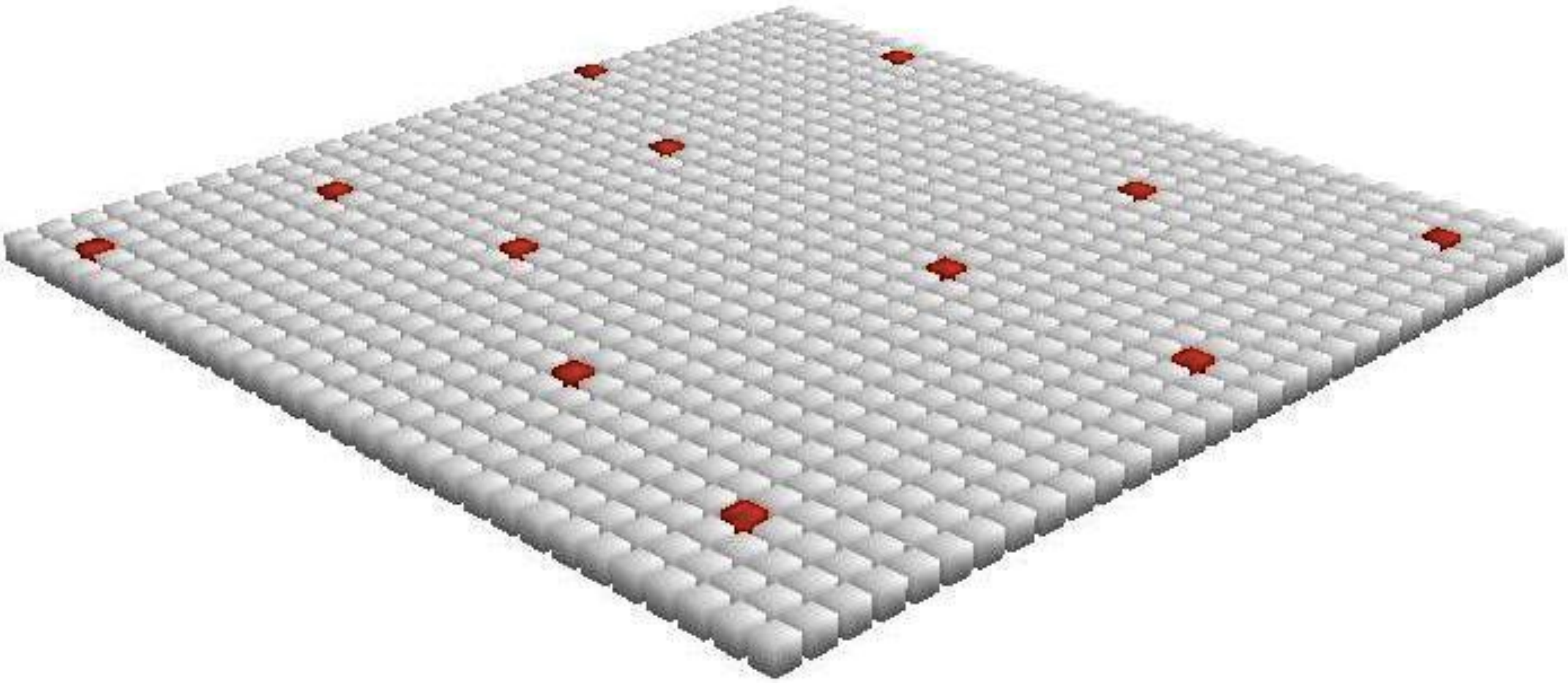
Coincidence detectors

How does this structure learn sequences?

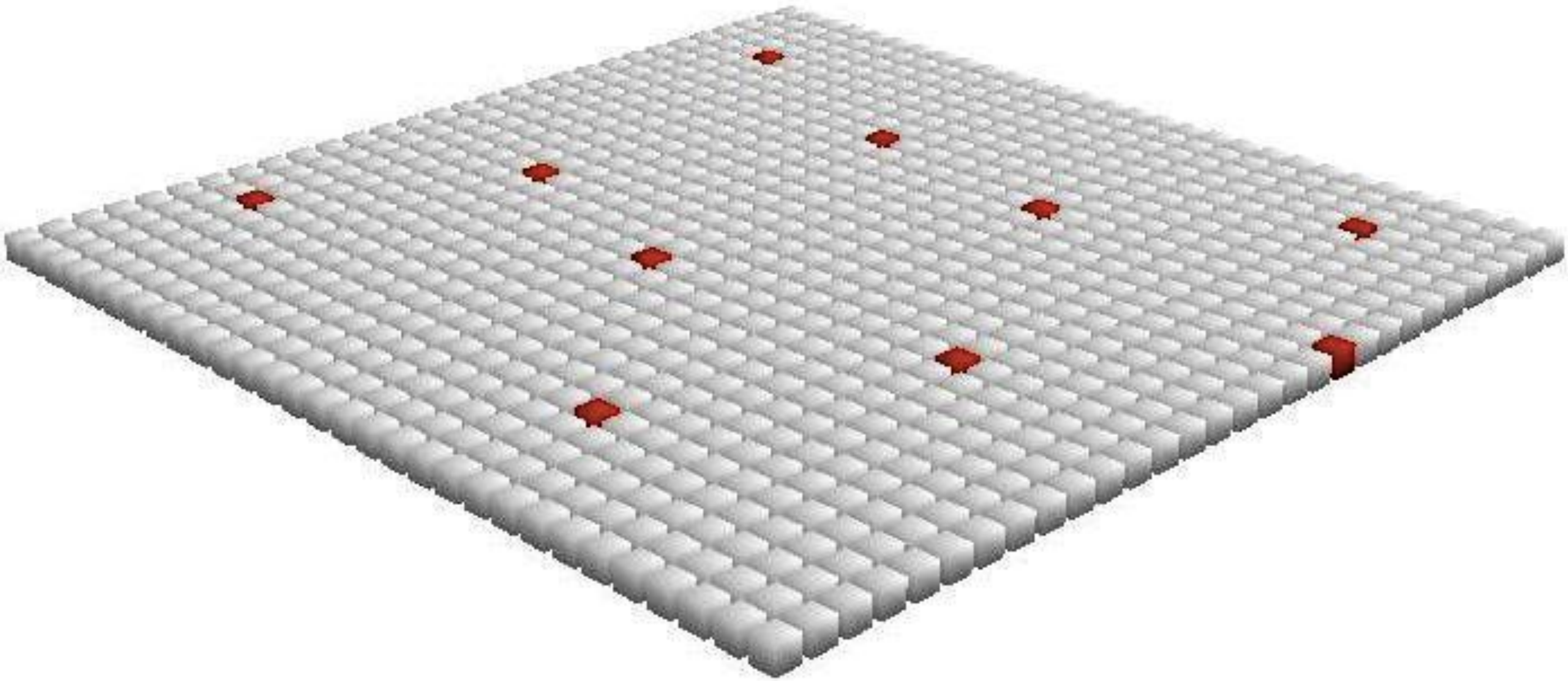
Each cell is one bit in our Sparse Distributed Representation



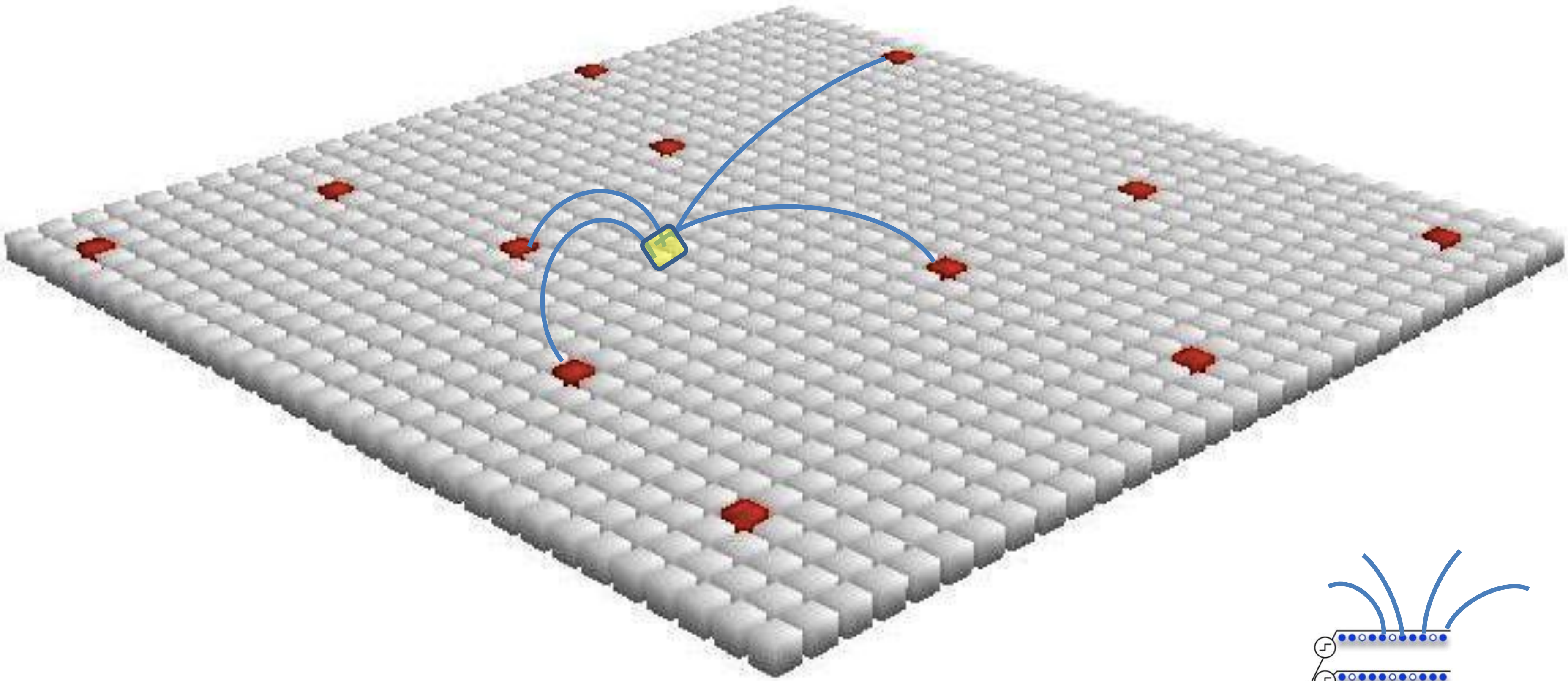
SDR (time =1)

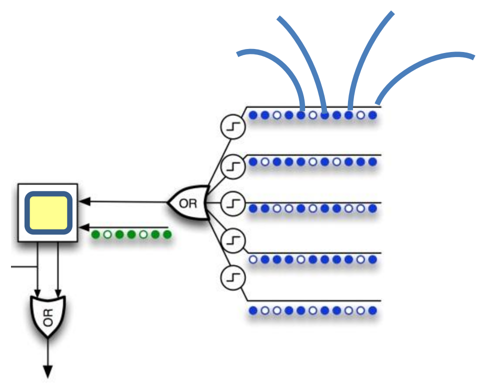


SDR (time =2)

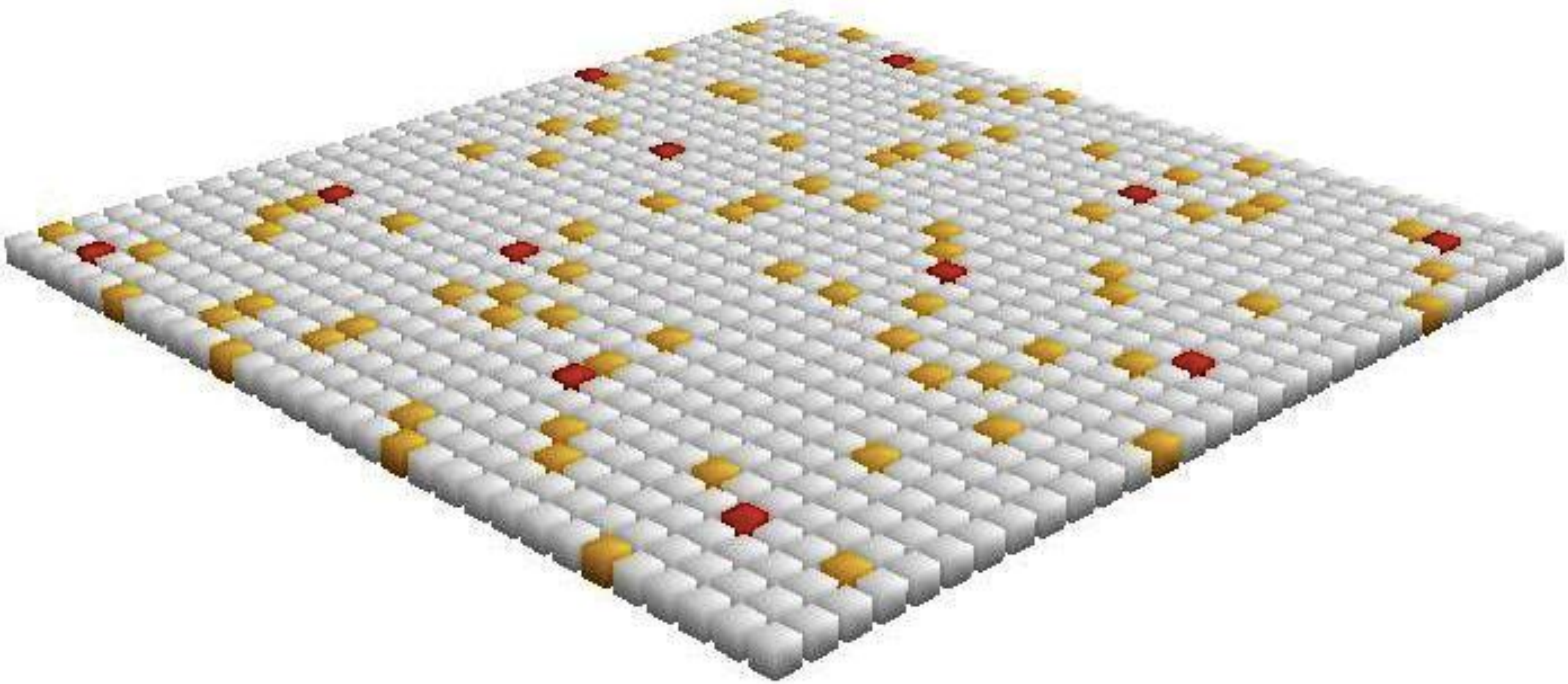


Cells form connections to subsample of previously active cells.
Predicts its own future activity.



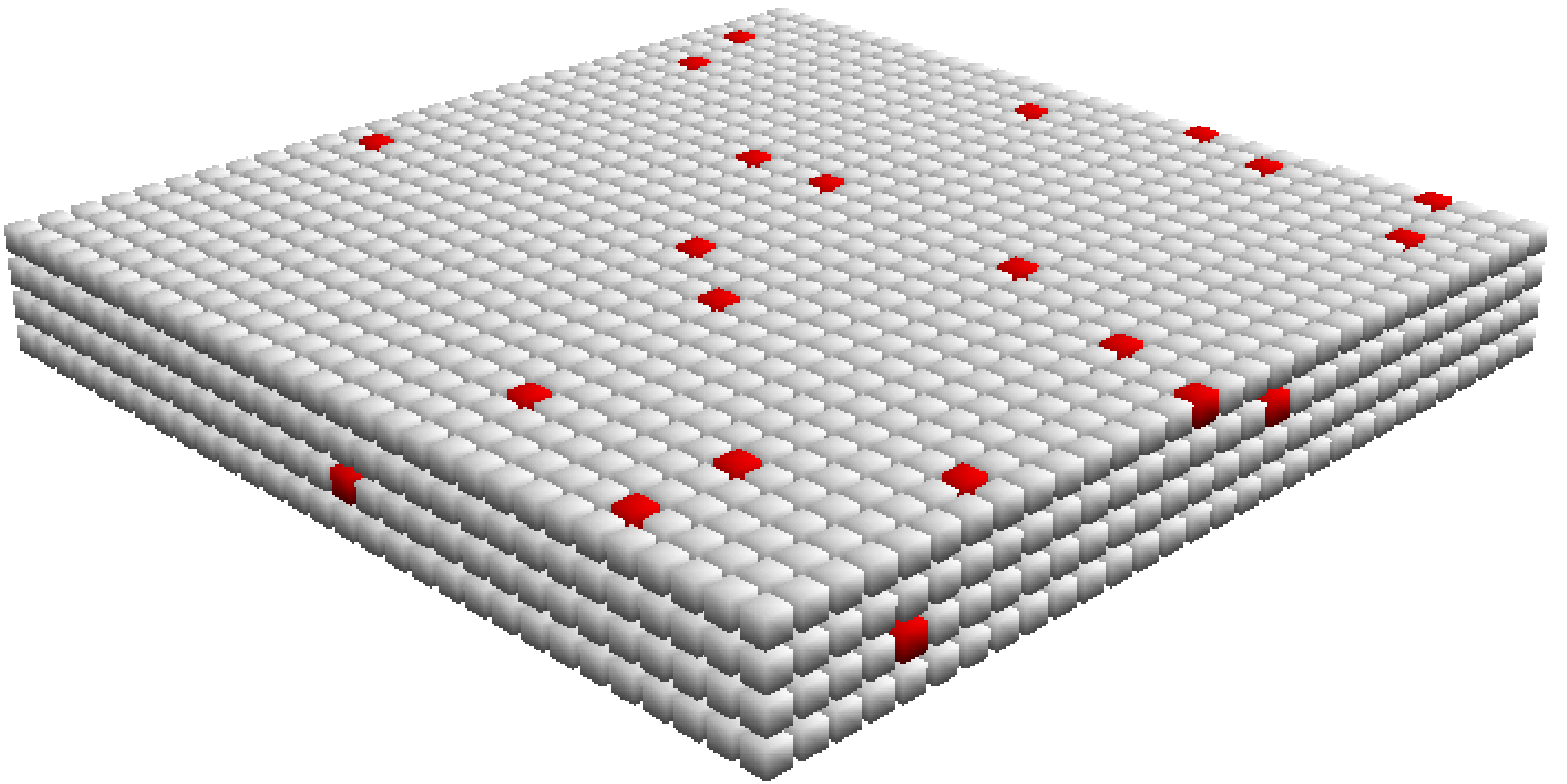


Multiple Predictions Can Occur at Once

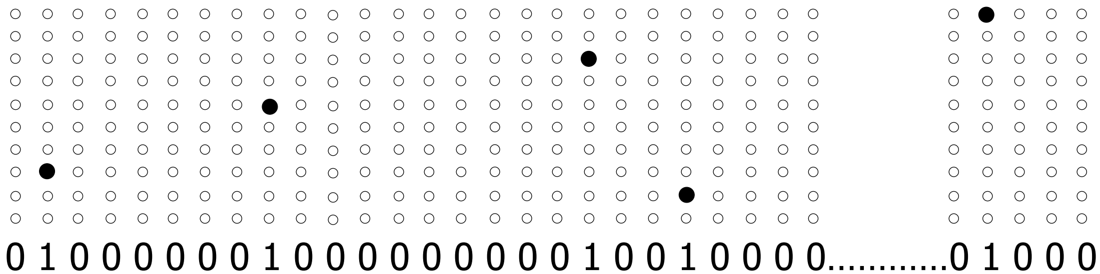


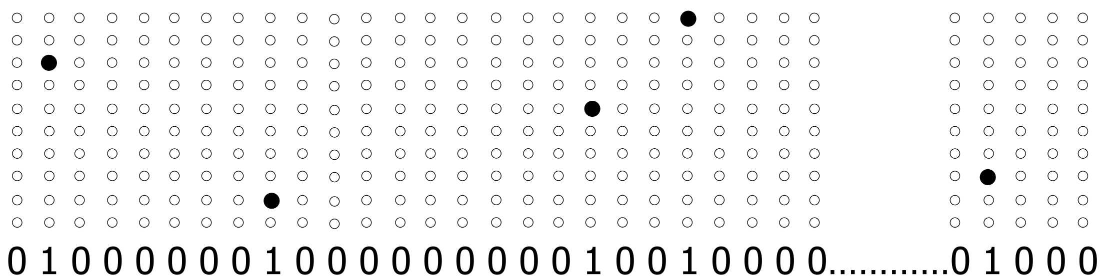
With one cell per column, 1st order memory
We need a high order memory

High order sequences are enabled with multiple cells per column.



Variable Order Sequence Memory

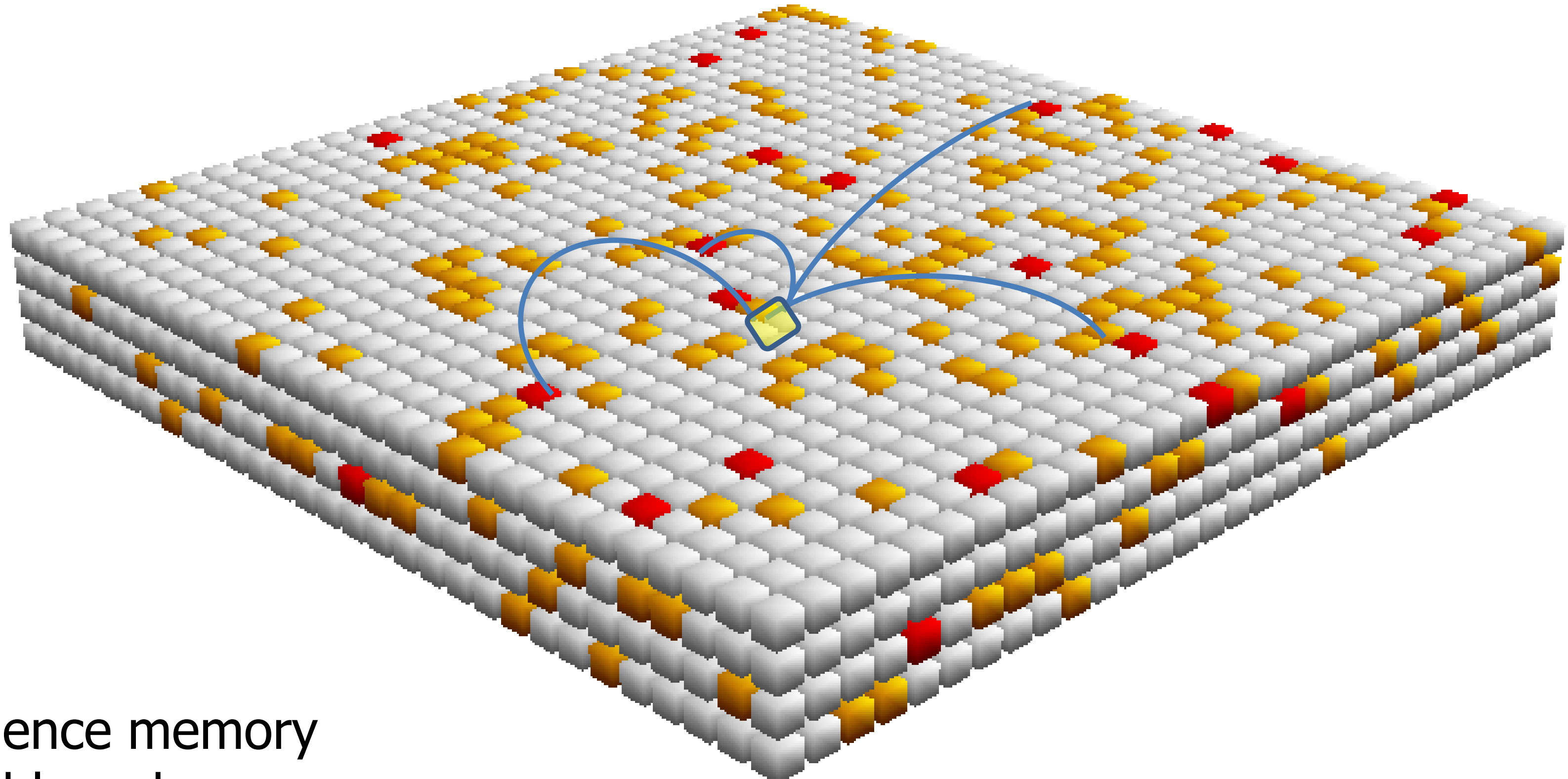




40 active columns, 10 cells per column

= 10^{40} ways to represent the same input in different contexts

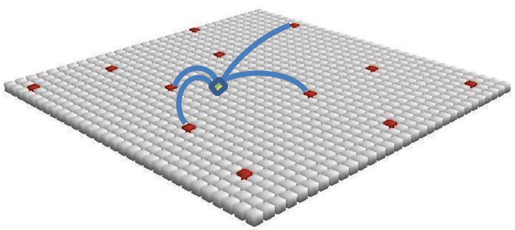
Variable Order Sequence Memory

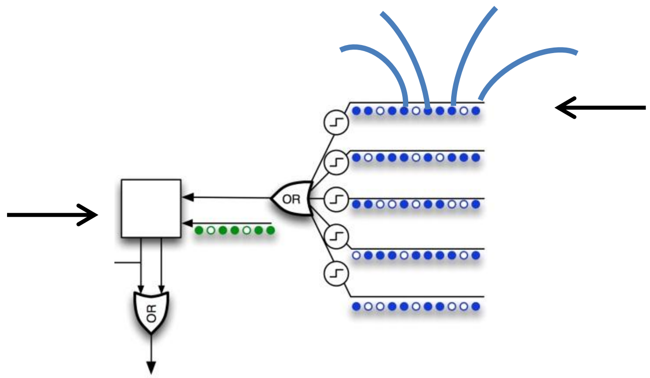


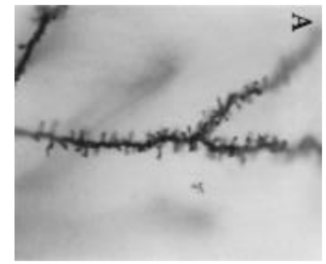
Sequence memory
Variable order
Distributed
Multiple simultaneous predictions
High capacity
Semantic generalization

Requirements for Online learning

- Train on every new input
- If pattern does not repeat, forget it
- If pattern repeats, reinforce it







Learning is the formation of connections

Connection permanence

unconnected — connected

0 — 0.2 — 1

Connection strength/weight is binary

Connection permanence is a scalar

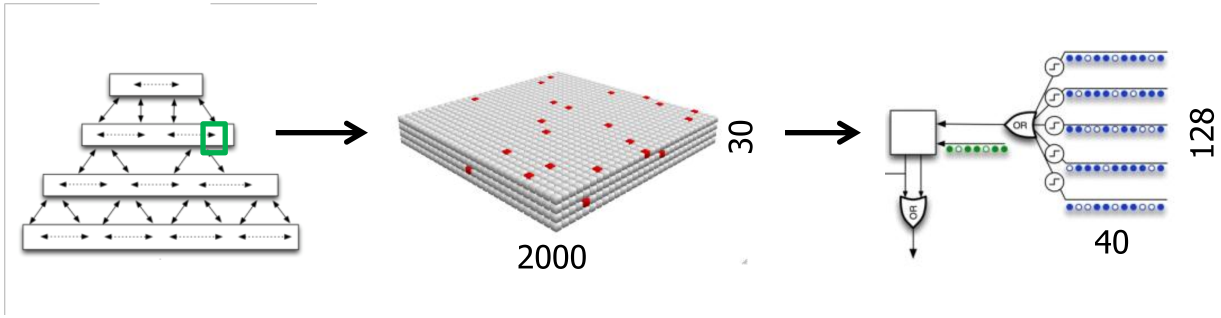
Training changes permanence

If permanence > threshold then connected

Memory Requirements For Minimum Viable Implementation

- Regions (1)
- Columns per region
- Cells per column
- Dendrite segments per cell
- Connections per segment

	Typical values today
Regions (1)	
Columns per region	2,000
Cells per column	30
Dendrite segments per cell	128
Connections per segment	40



300M connections

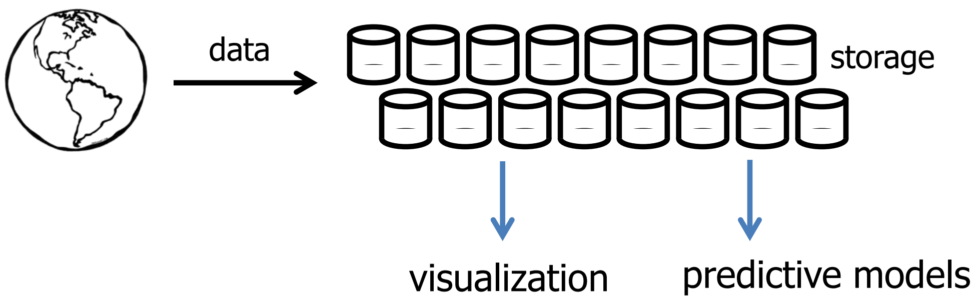
- Connection index
- Connection permanence

Connections are sparse

No single points of failure

Predictive Analytics

Today



<u>Challenges</u>

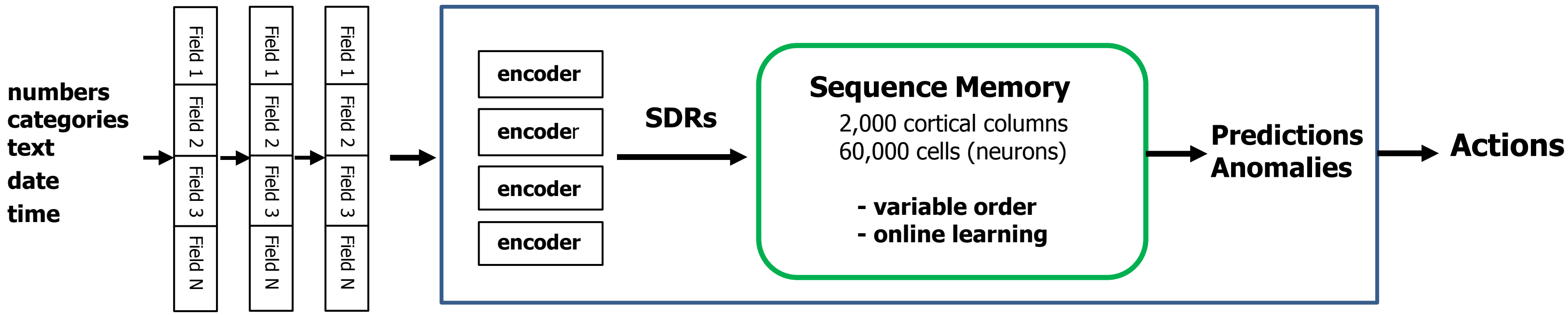
Data prep
Model obsolescence
People

Tomorrow

<u>Key criteria</u>

Automated model creation
Continuous learning
Temporal and spatial models

Grok: A Engine for Acting on Data Streams



User

Create data stream

Define problem
- what to predict
- how often
- how far in advance

Grok

Creates models

Learns continuously

Finds spatial/temporal patterns

Outputs
- predictions
  with probabilities

Customer areas

Energy pricing

Energy demand

Product forecasting

Ad network return

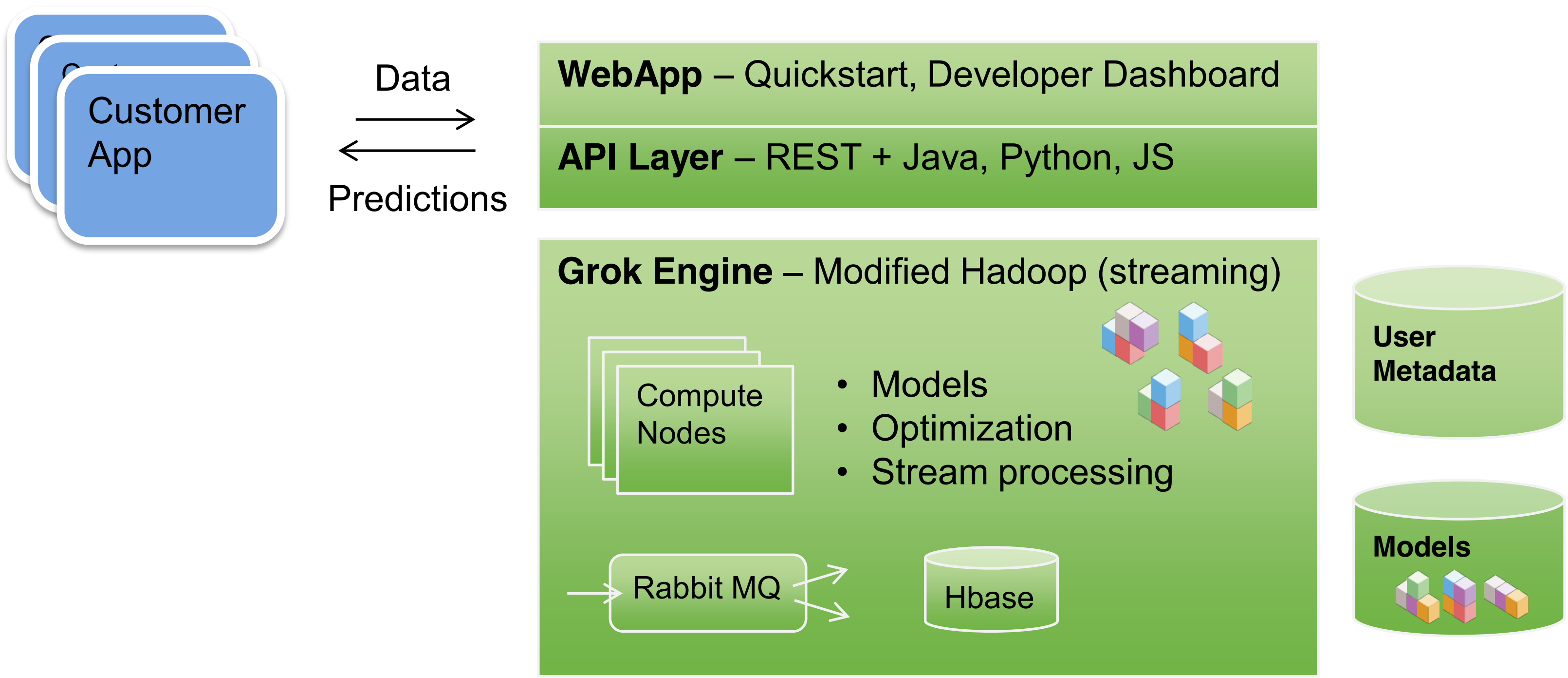
Machine efficiency

Anomaly detection

Server loads

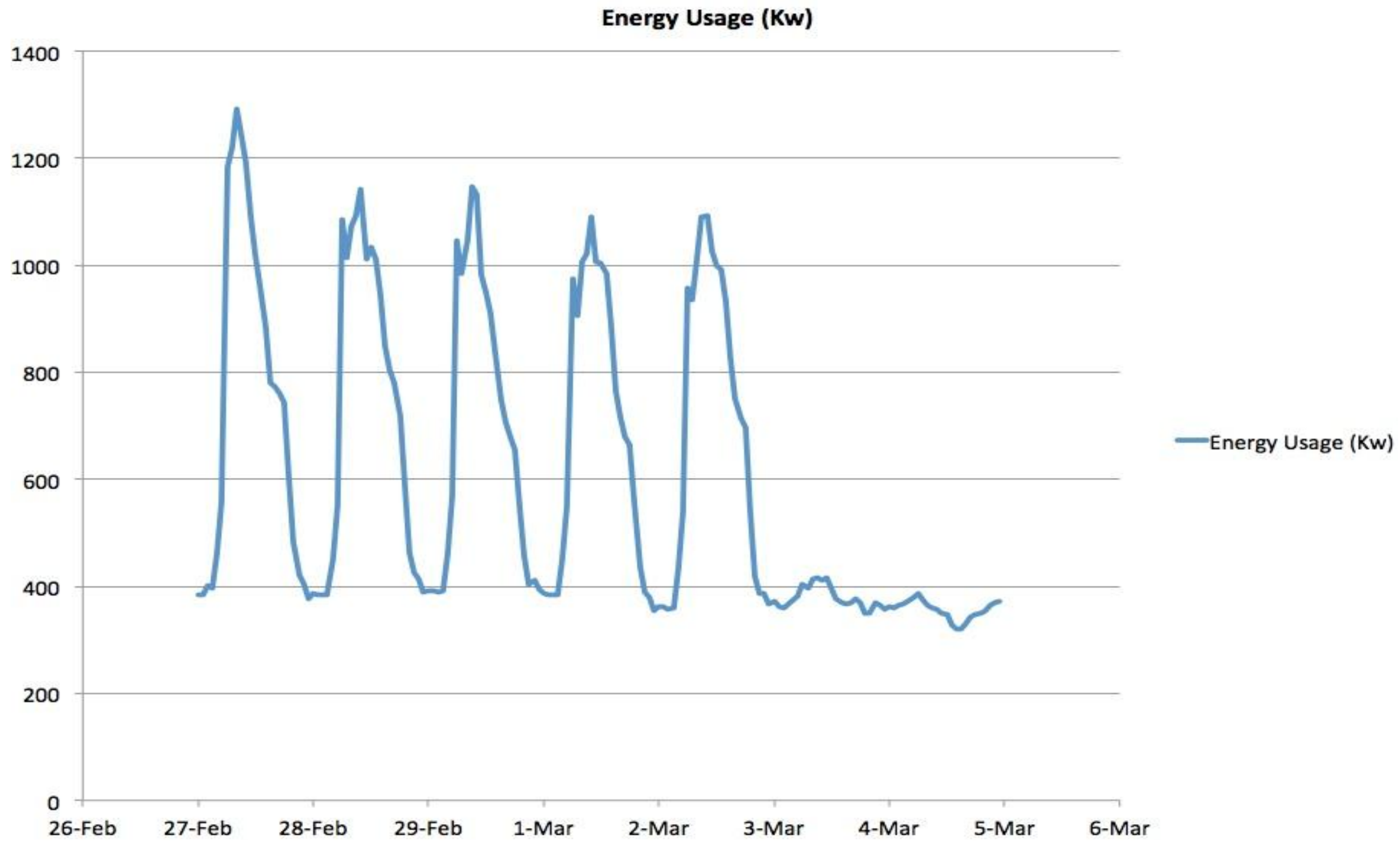
Grok Architecture Today

Cloud based running on Amazon AWS

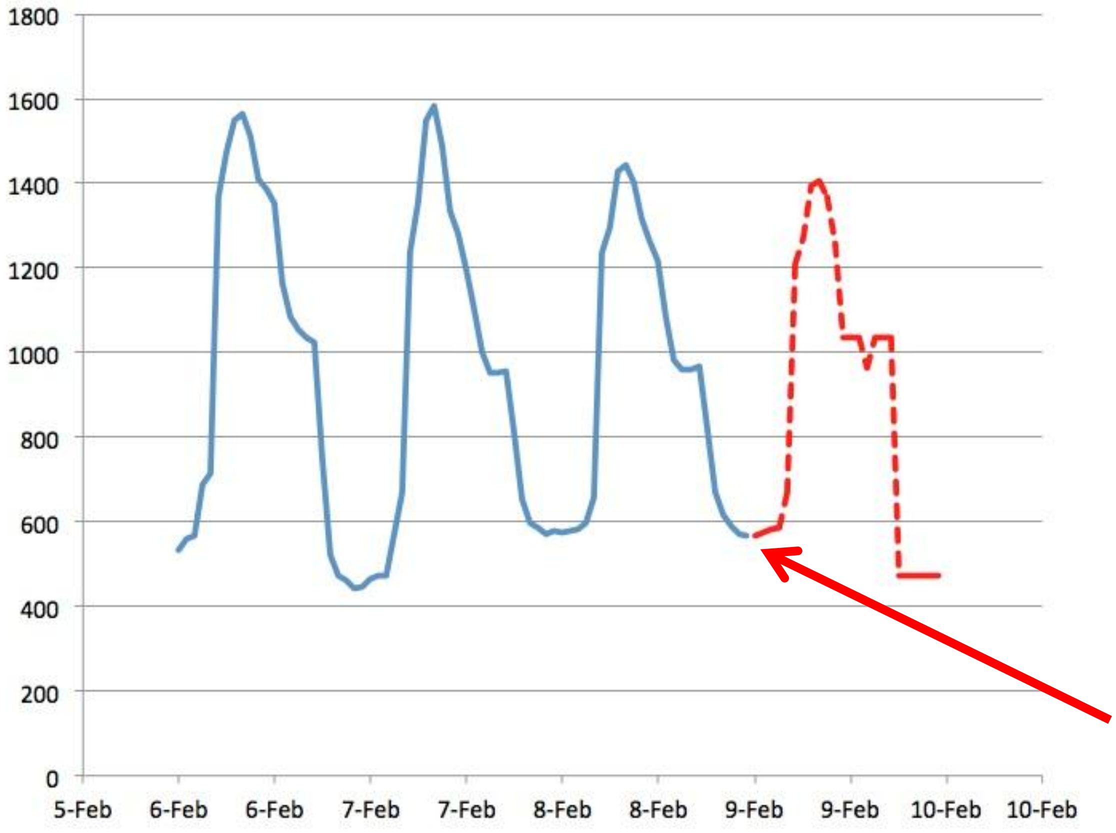


Application: Energy Demand/Response

Factory Energy Profile

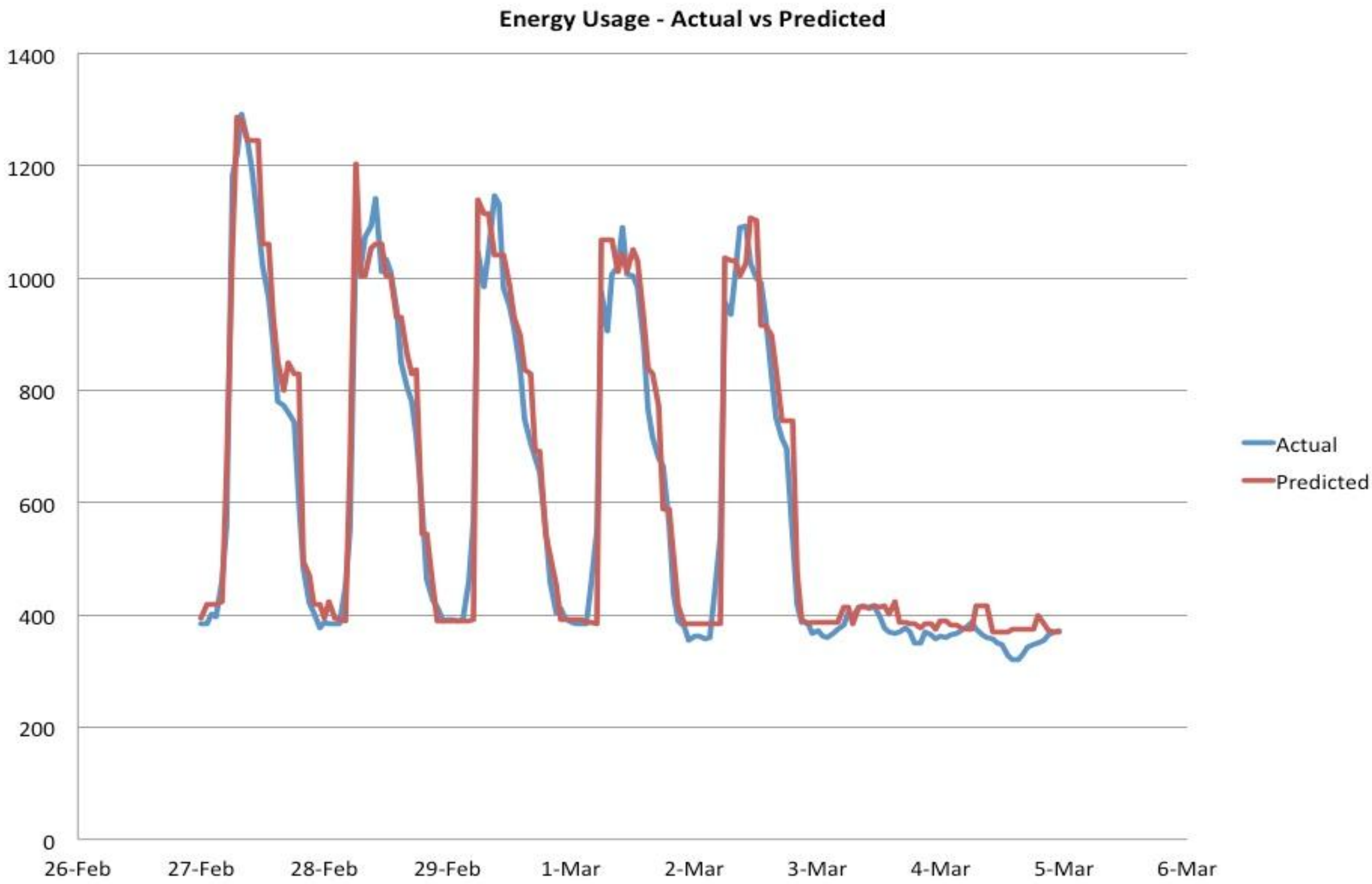


Customer need

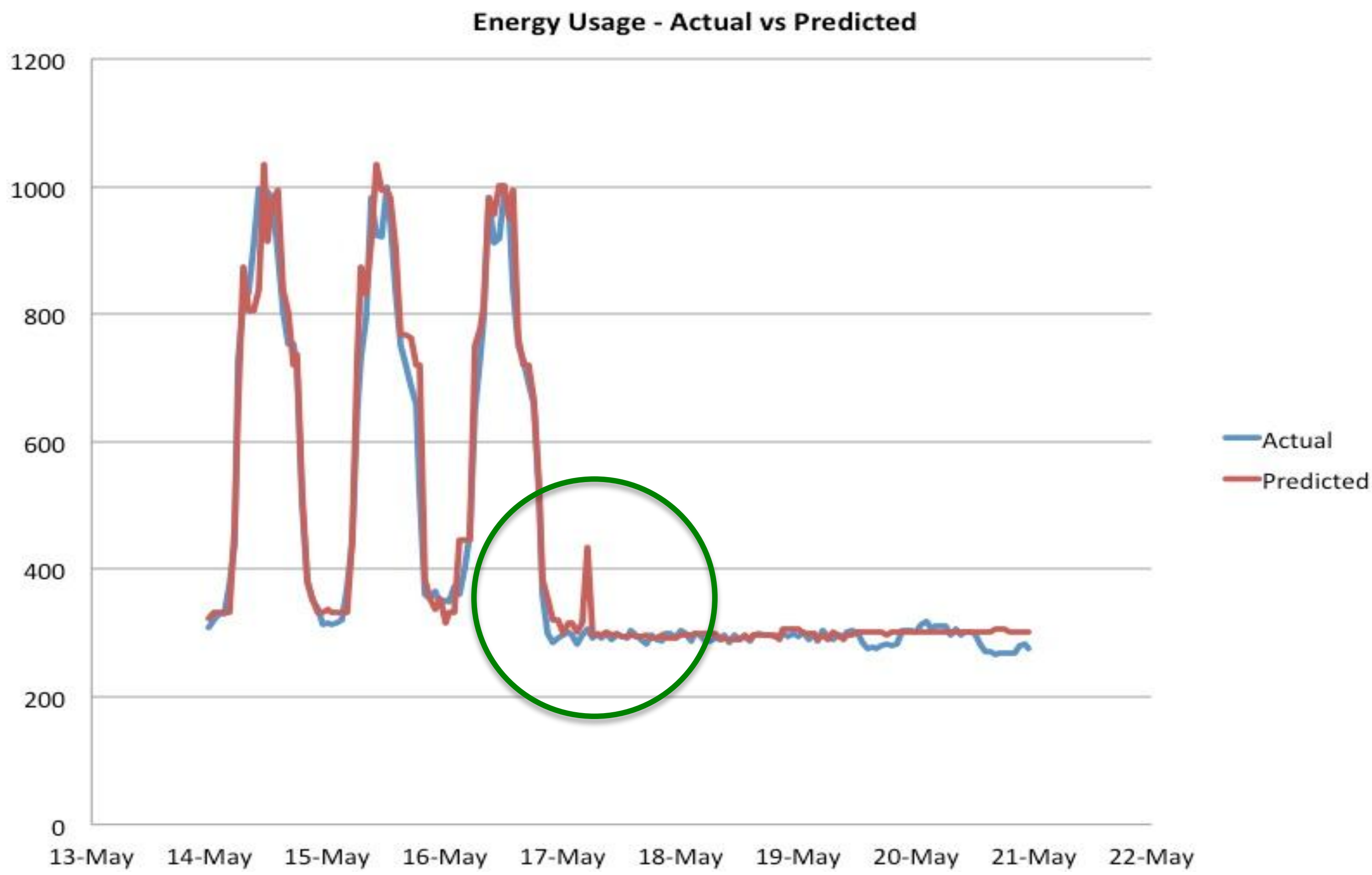


At midnight, make predictions for next 24 hours

Predictions and Actuals



Predictions and Actuals II

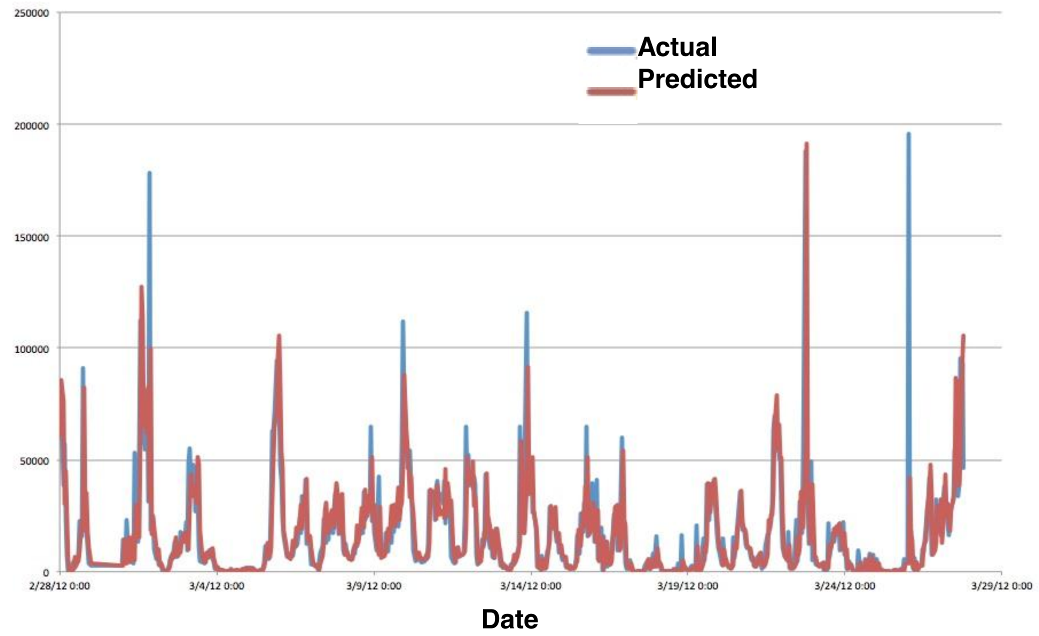


Managing Server Capacity

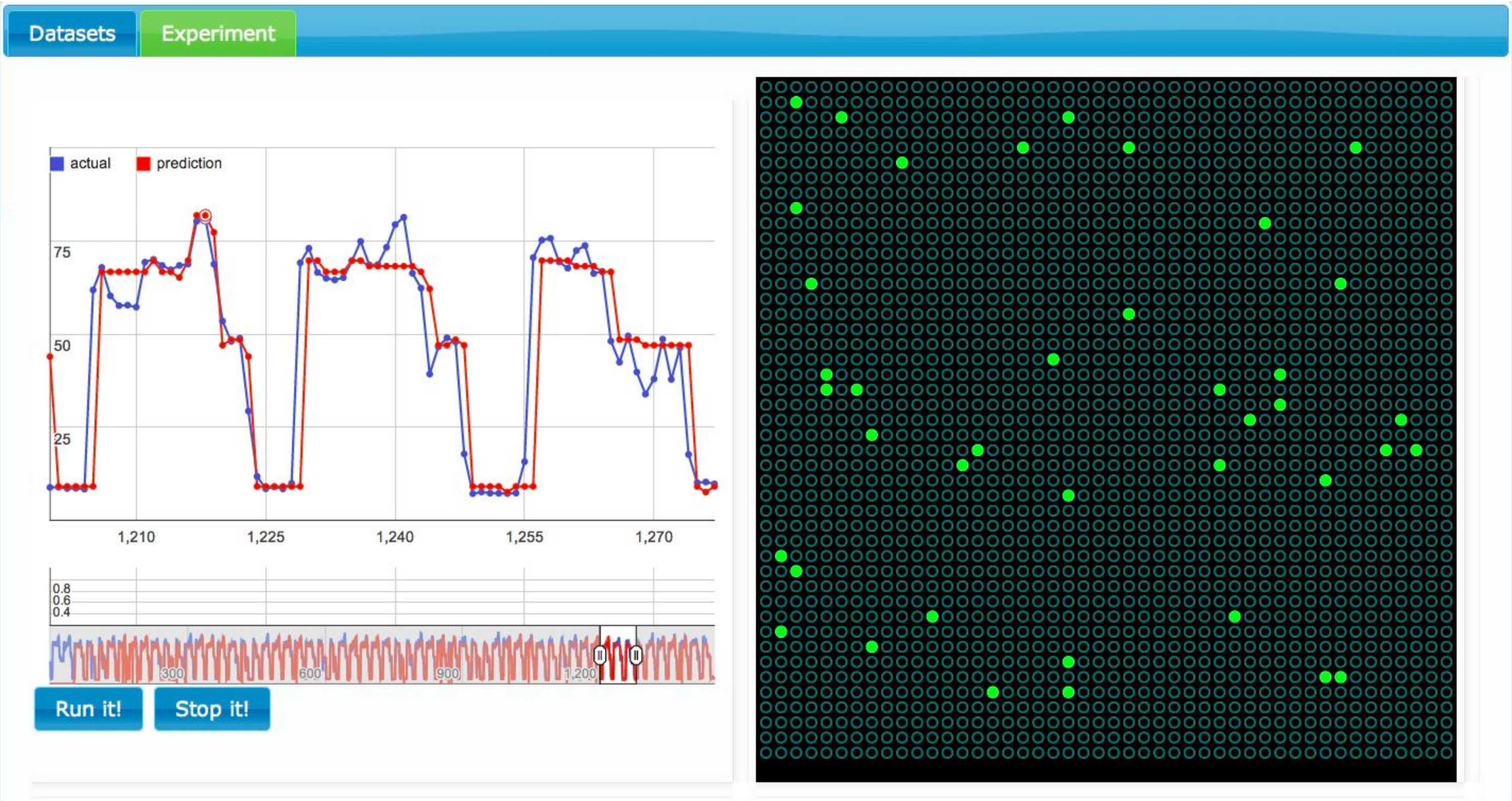
Grok used to predict server demand

Used to provision instances ahead of time

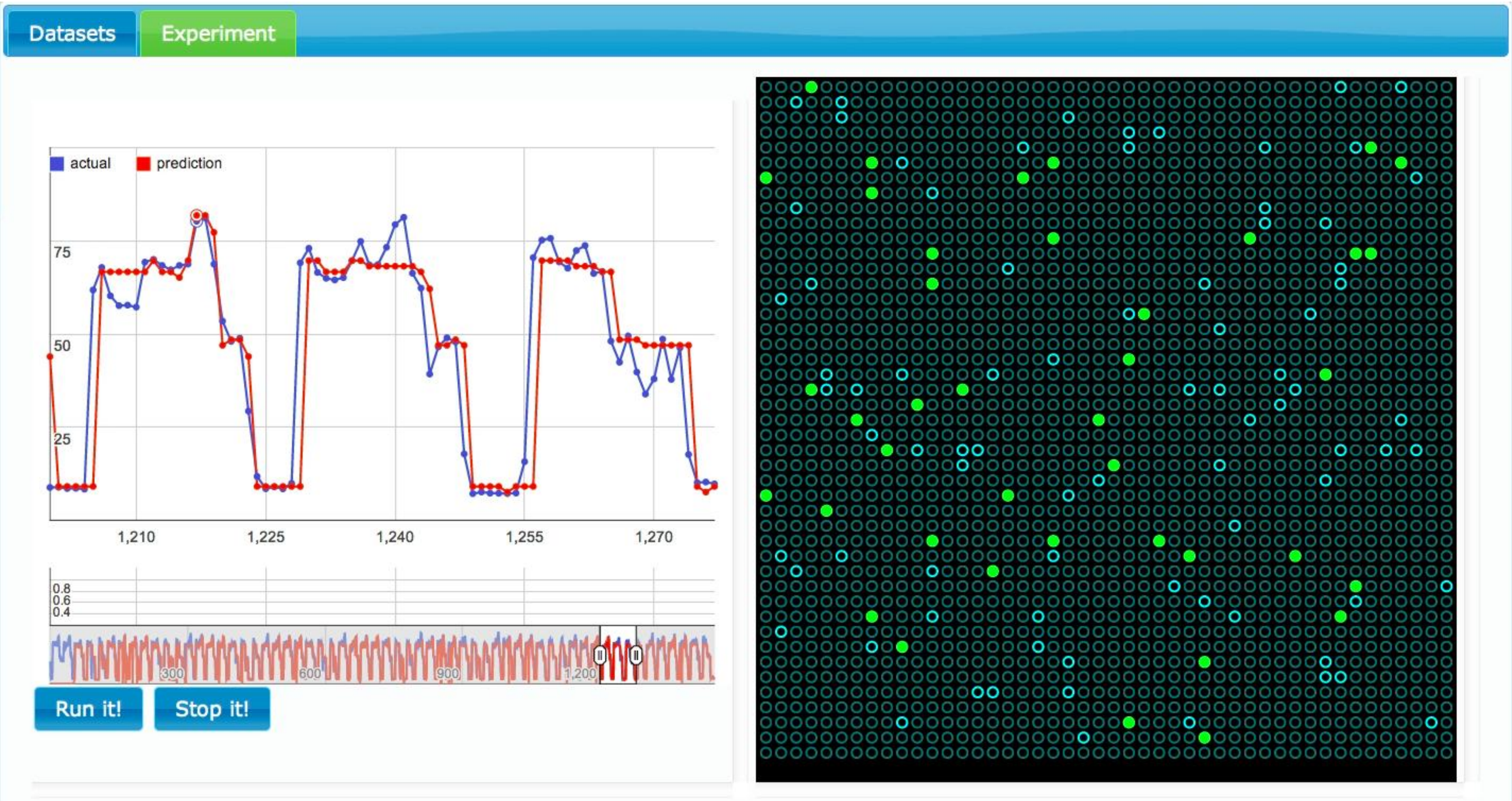
Results show approximately 15% reduction in AWS cost

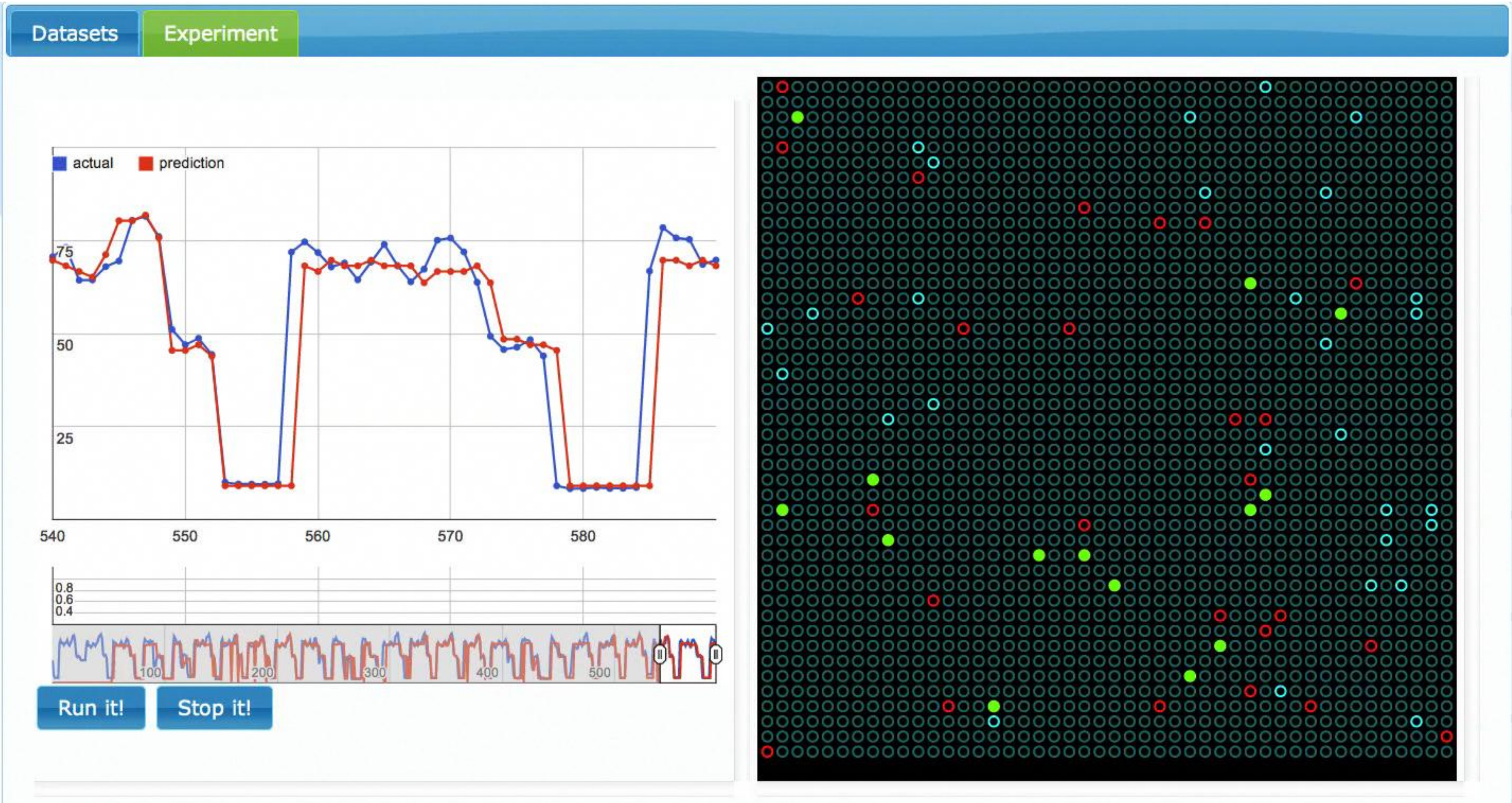


Incoming server demand, Actual vs. Predicted



Datasets
Experiment
actual
prediction
75
50
25
1,210
1,225
1,240
1,255
1,270
0.8
0.6
0.4
300
600
900
1,200
Run it!
Stop it!


Datasets
Experiment
actual
prediction
75
50
25
1,210
1,225
1,240
1,255
1,270
0.8
0.6
0.4
300
600
900
1,200
Run it!
Stop it!

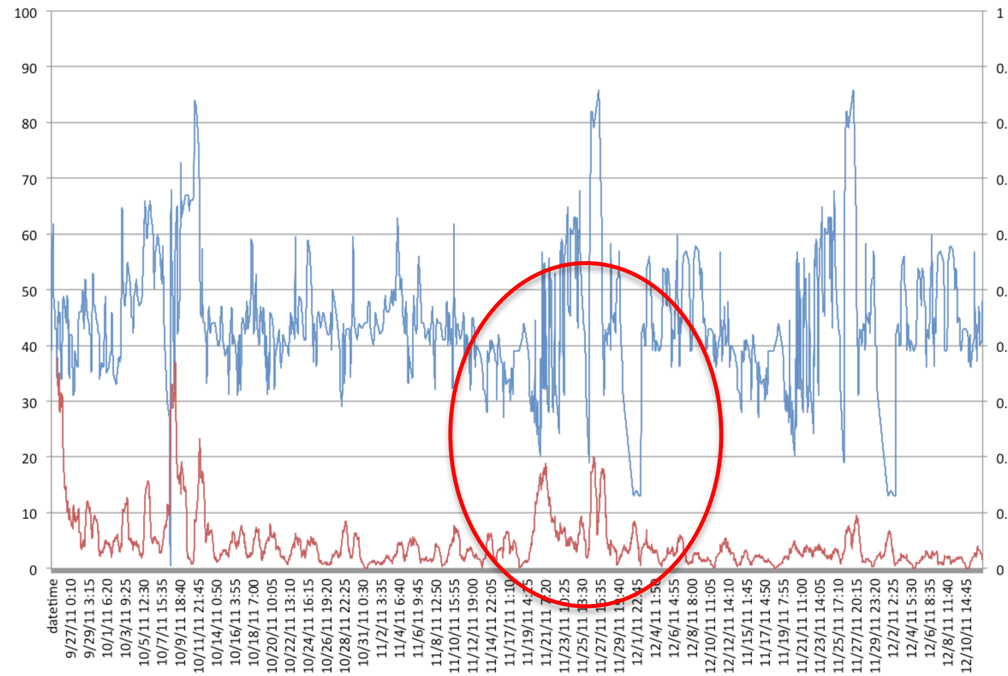

Datasets
Experiment
actual
prediction
75
50
25
540
550
560
570
580
0.8
0.6
0.4
100
200
300
400
500
Run it!
Stop it!

Predictive Maintenance

Grok used to detect anomalies in gear bearing temperature

Can detect anomalies based on temporal characteristics

Can be used to proactively optimize maintenance schedules



Gear bearing temperature & Grok Anomaly Score

Future of Machine Intelligence

More Theory

- Hierarchy
- Sensory-Motor Integration
- Attention

Embodiment

- Today: Cloud-based service
- Embedded
- Distributed: Billions of sensors, models,
  hierarchically connected

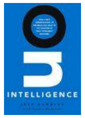
Hardware

- Speed, cost, power
  1) Memory
     - natural fault tolerance
  2) Interconnects
     - sparsity and hierarchy
     - sub-sampling

Applications

- Today: Prediction/anomaly detection
- Classics: vision, language, speech?
- The big wins….

For More Information

- www.Numenta.com (white paper)
- On Intelligence 

- Speak to me

  Joe Hayashi

  Matt Taylor

- jhawkins@numenta.com

Thank You