Automated Machine Learning For Autonomic Computing

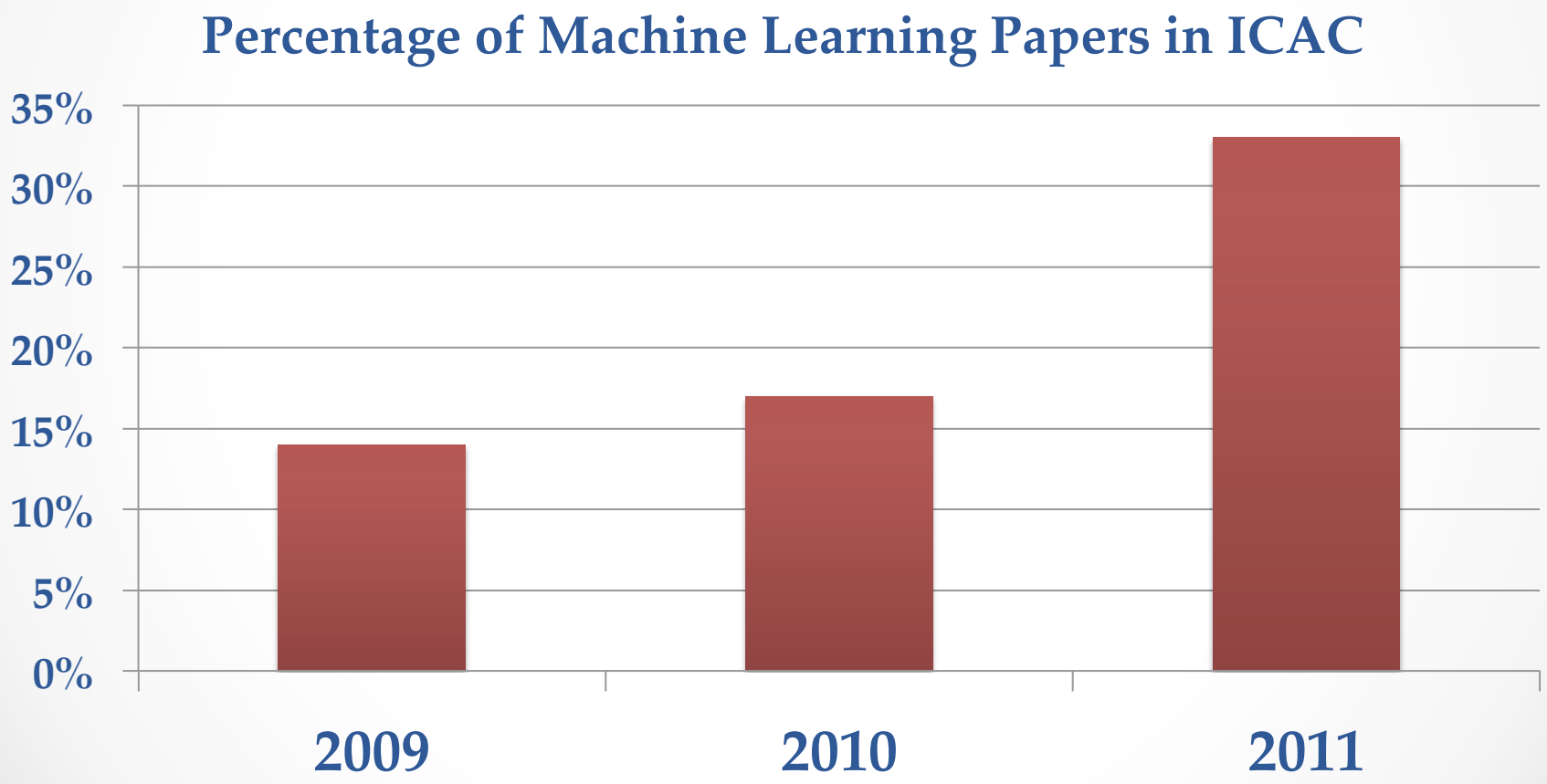
ICAC 2012

Numenta
Subutai Ahmad

Autonomic Machine Learning

ICAC 2012

Numenta
Subutai Ahmad


Percentage of Machine Learning Papers in ICAC
35%
30%
25%
20%
15%
10%
5%
0%
2009
2010
2011

Numenta

Small company, located in Redwood City, CA

About 22 people

Numenta

Created in 2005

Jeff Hawkins, founder of Palm Computing and Handspring

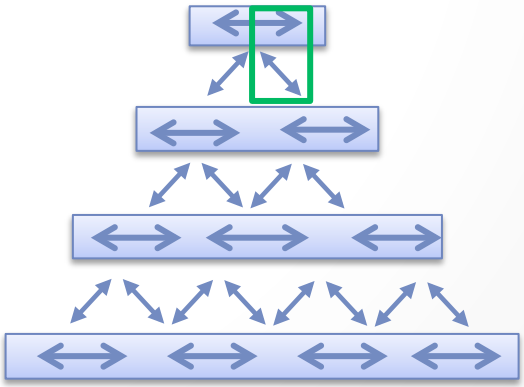
Author of *On Intelligence*

Early Years

2005 – 2010

Focused on developing a new theory of machine intelligence and machine learning

Inspired by biology and specifically neuroscience



2012

Numenta®

Grok (BETA)

A product designed for automated streaming analytics

Data

Dynamically Generated Models

Talk Outline

What are the challenges of traditional machine learning?

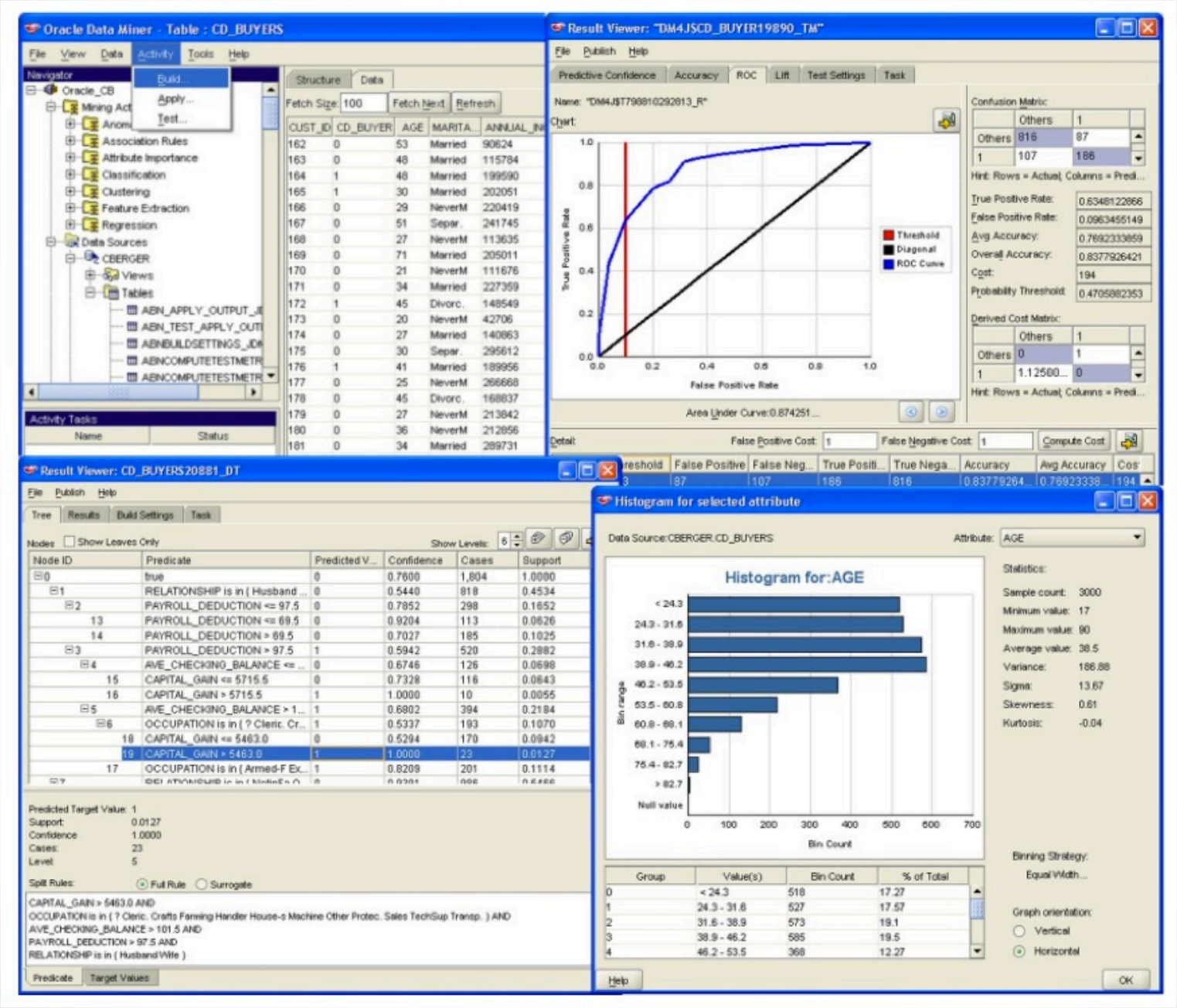
What are the properties of autonomic machine learning?

Numenta's algorithm and Grok product architecture

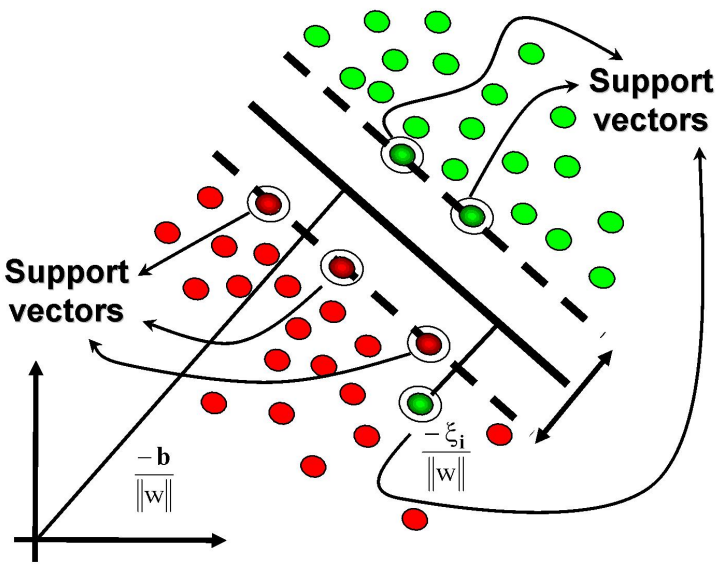
The future – where should we go?

Machine Learning Challenges

Extremely manual

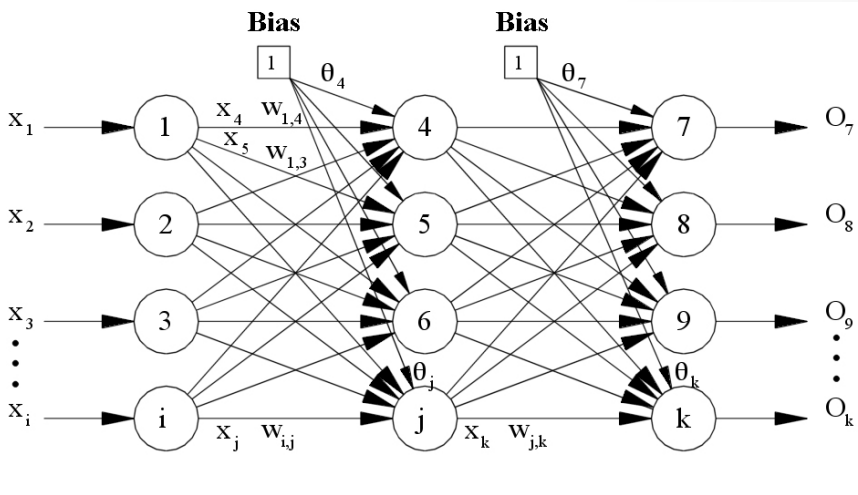

Oracle Data Miner - Table : CD_BUYERS
File View Data Activity Tools Help
Build...
Apply...
Test...
Navigator
Oracle_CB
Mining Act
Anom
Association Rules
Attribute Importance
Classification
Clustering
Feature Extraction
Regression
Data Sources
CBERGER
Views
Tables
ABN_APPLY_OUTPUT_JI
ABN_TEST_APPLY_OUT
ABNBUILDSETTINGS_JD
ABNCOMPUTETESTMETR
ABNCOMPUTETESTMETR
Activity Tasks
Name
Status
Structure
Data
Fetch Size: 100
Fetch Next
Refresh
CUST_ID CD_BUYER AGE MARITA... ANNUAL_IN
162 0 53 Married 90624
163 0 48 Married 115784
164 1 48 Married 199590
165 1 30 Married 202051
166 0 29 NeverM 220419
167 0 51 Separ. 241745
168 0 27 NeverM 113635
169 0 71 Married 205011
170 0 21 NeverM 111676
171 0 34 Married 227359
172 1 45 Divorc. 148549
173 0 20 NeverM 42706
174 0 27 Married 140863
175 0 30 Separ. 295612
176 1 41 Married 189956
177 0 25 NeverM 266668
178 0 45 Divorc. 168837
179 0 27 NeverM 213842
180 0 36 NeverM 212856
181 0 34 Married 289731
Result Viewer: "DM4JSCD_BUYER19890_TM"
File Publish Help
Predictive Confidence
Accuracy
ROC
Lift
Test Settings
Task
Name: "DM4J$T798810292813_R"
Chart:
True Positive Rate
False Positive Rate
Threshold
Diagonal
ROC Curve
Area Under Curve:0.874251...
Confusion Matrix
Others 1
Others 816 87
1 107 186
Hint: Rows = Actual, Columns = Predi...
True Positive Rate: 0.6348122866
False Positive Rate: 0.0963455149
Avg Accuracy: 0.7692333859
Overall Accuracy: 0.8377926421
Cost: 194
Probability Threshold: 0.4705882353
Derived Cost Matrix:
Others 1
Others 0 1
1 1.12500... 0
Hint: Rows = Actual, Columns = Predi...
Detail
False Positive Cost: 1
False Negative Cost: 1
Compute Cost
reshold False Positive False Neg... True Positi... True Nega... Accuracy Avg Accuracy Cos
3 87 107 186 816 0.83779264... 0.76923338... 194
Result Viewer: CD_BUYERS20881_DT
File Publish Help
Tree
Results
Build Settings
Task
Nodes
Show Leaves Only
Show Levels: 6
Node ID Predicate Predicted V... Confidence Cases Support
0 true 0 0.7600 1,804 1.0000
1 RELATIONSHIP is in (Husband ... 0 0.5440 818 0.4534
2 PAYROLL_DEDUCTION <= 97.5 0 0.7852 298 0.1652
13 PAYROLL_DEDUCTION <= 69.5 0 0.9204 113 0.0626
14 PAYROLL_DEDUCTION > 69.5 0 0.7027 185 0.1025
3 PAYROLL_DEDUCTION > 97.5 1 0.5942 520 0.2882
4 AVE_CHECKING_BALANCE <= ... 0 0.6746 126 0.0698
15 CAPITAL_GAIN <= 5715.5 0 0.7328 116 0.0643
16 CAPITAL_GAIN > 5715.5 1 1.0000 10 0.0055
5 AVE_CHECKING_BALANCE > 1... 1 0.6802 394 0.2184
6 OCCUPATION is in (? Cleric. Cr... 1 0.5337 193 0.1070
18 CAPITAL_GAIN <= 5463.0 0 0.5294 170 0.0942
19 CAPITAL_GAIN > 5463.0 1 1.0000 23 0.0127
17 OCCUPATION is in (Armed-F Ex... 1 0.8209 201 0.1114
Predicted Target Value: 1
Support: 0.0127
Confidence 1.0000
Cases: 23
Level 5
Split Rules:
Full Rule
Surrogate
CAPITAL_GAIN > 5463.0 AND
OCCUPATION is in (? Cleric. Crafts Farming Handler House-s Machine Other Protec. Sales TechSup Transp.) AND
AVE_CHECKING_BALANCE > 101.5 AND
PAYROLL_DEDUCTION > 97.5 AND
RELATIONSHIP is in (Husband Wife)
Predicate
Target Values
Histogram for selected attribute
Data Source:CBERGER.CD_BUYERS
Attribute: AGE
Histogram for:AGE
Bin range
< 24.3
24.3 - 31.6
31.6 - 38.9
38.9 - 46.2
46.2 - 53.5
53.5 - 60.8
60.8 - 68.1
68.1 - 75.4
75.4 - 82.7
> 82.7
Null value
0 100 200 300 400 500 600 700
Bin Count
Statistics:
Sample count: 3000
Minimum value: 17
Maximum value: 90
Average value: 38.5
Variance: 186.88
Sigma: 13.67
Skewness: 0.61
Kurtosis: -0.04
Group Value(s) Bin Count % of Total
0 < 24.3 518 17.27
1 24.3 - 31.6 527 17.57
2 31.6 - 38.9 573 19.1
3 38.9 - 46.2 585 19.5
4 46.2 - 53.5 368 12.27
Binning Strategy:
Equal Width...
Graph orientation:
Vertical
Horizontal
Help
OK

Support Vector Machines



or

Back propagation



Machine Learning Is Labor Intensive

Select the right algorithm

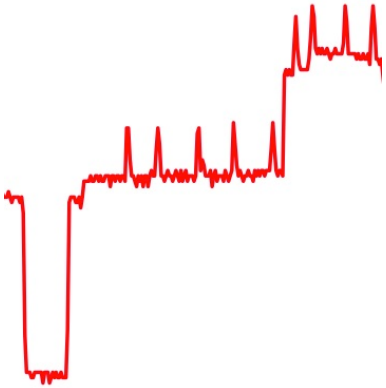
Experiment with parameters

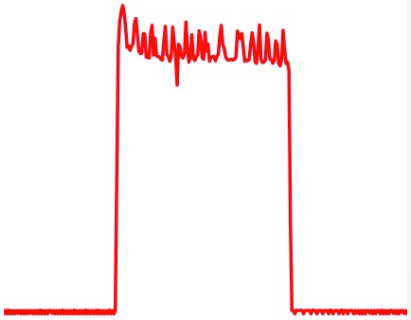
Construct good input features

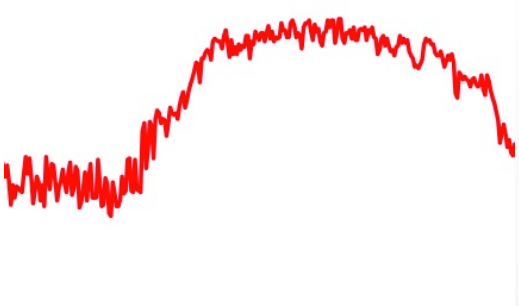
Robust testing and cross-validation

Application: Smart Meters

Application: Smart Meters







Machine Learning Challenges

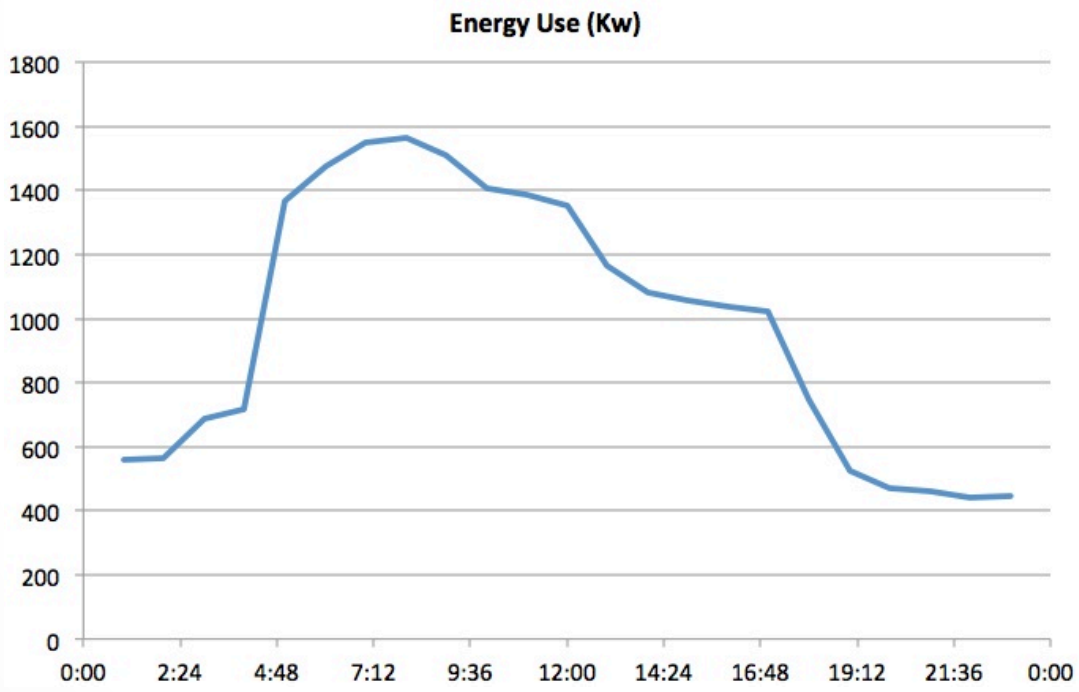
Extremely manual → Automated

Machine Learning Challenges

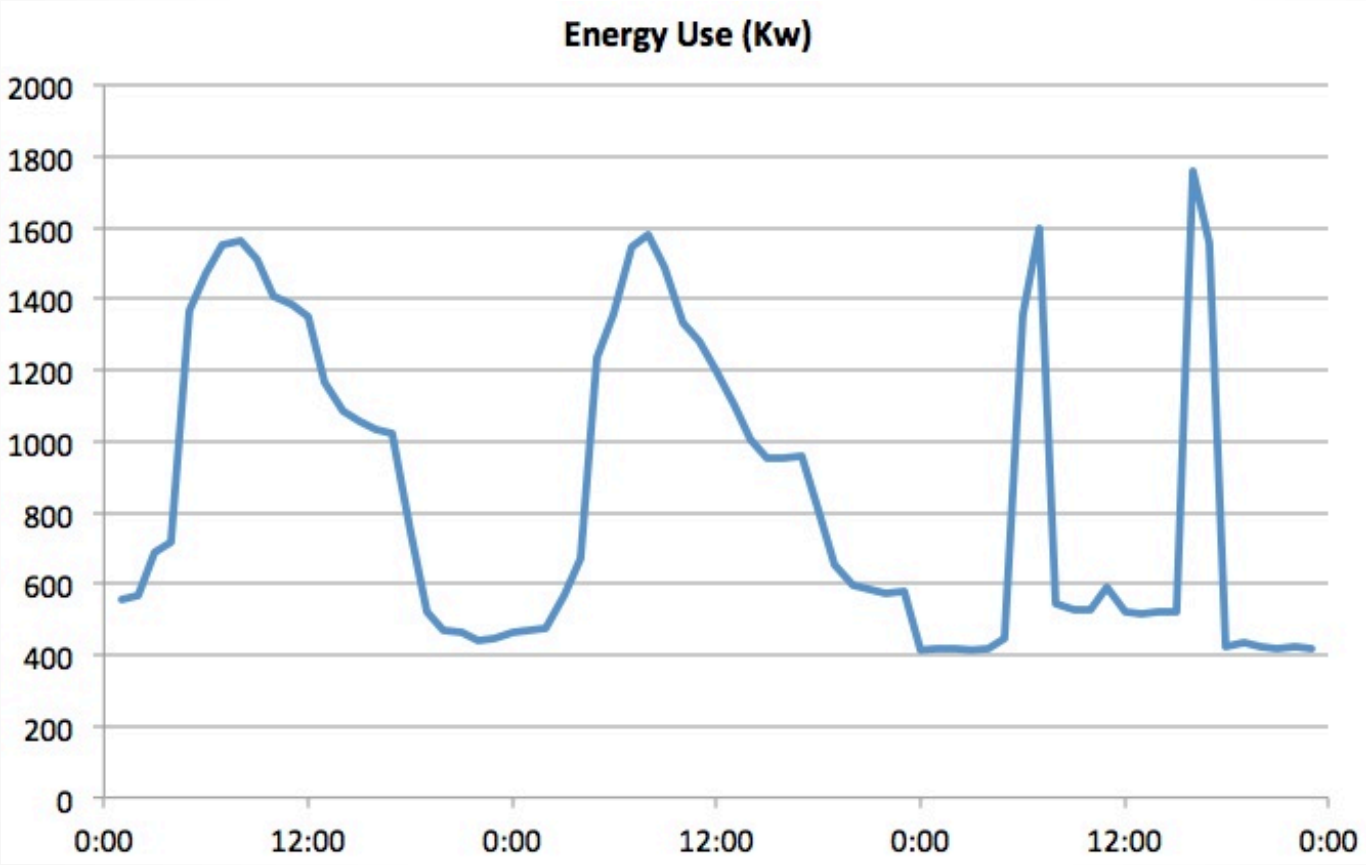
Extremely manual → Automated

Static models

Application: Smart Meters



Application: Smart Meters



Machine Learning Challenges

Extremely manual → Automated

Static models → Continuous learning

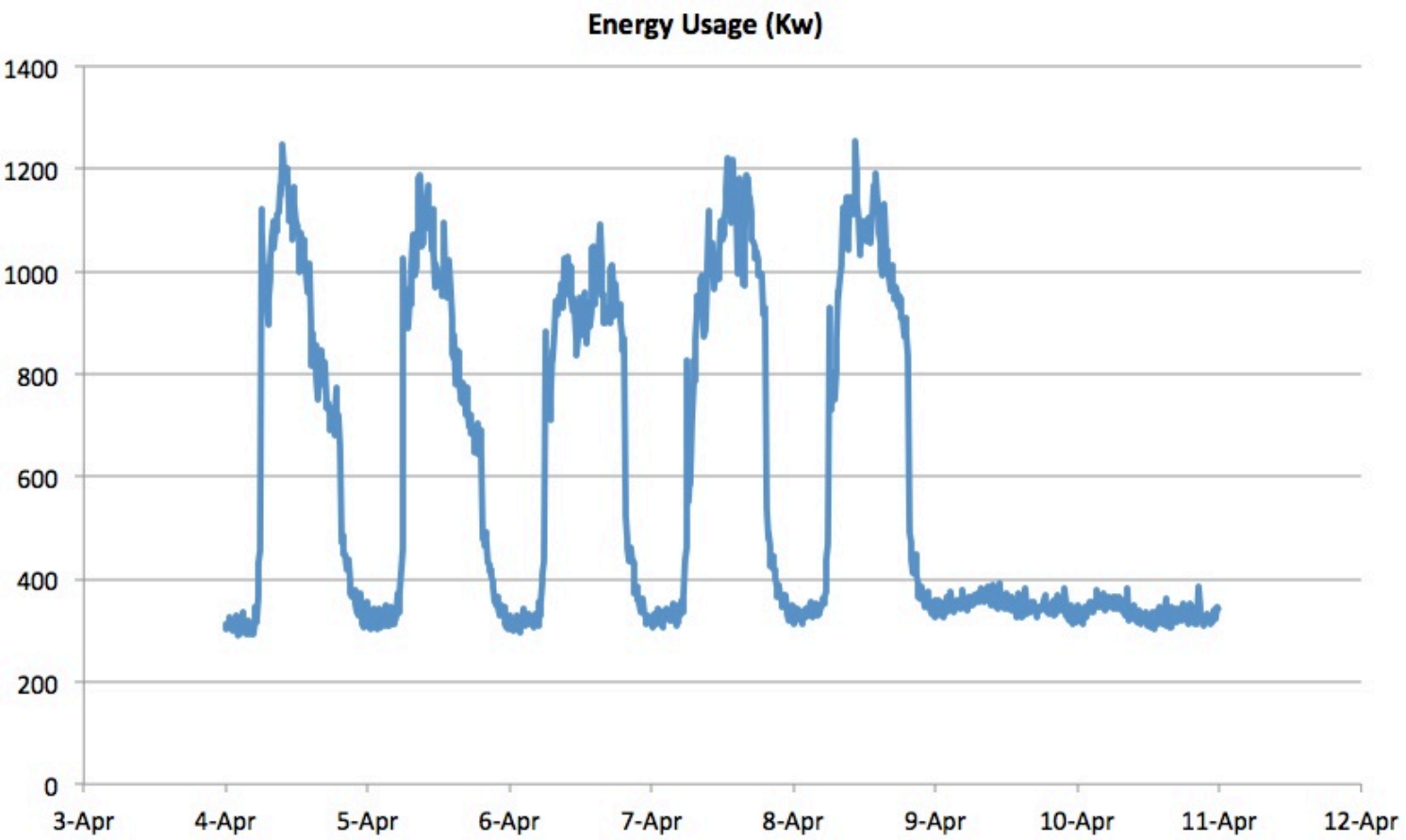
Machine Learning Challenges

Extremely manual → Automated

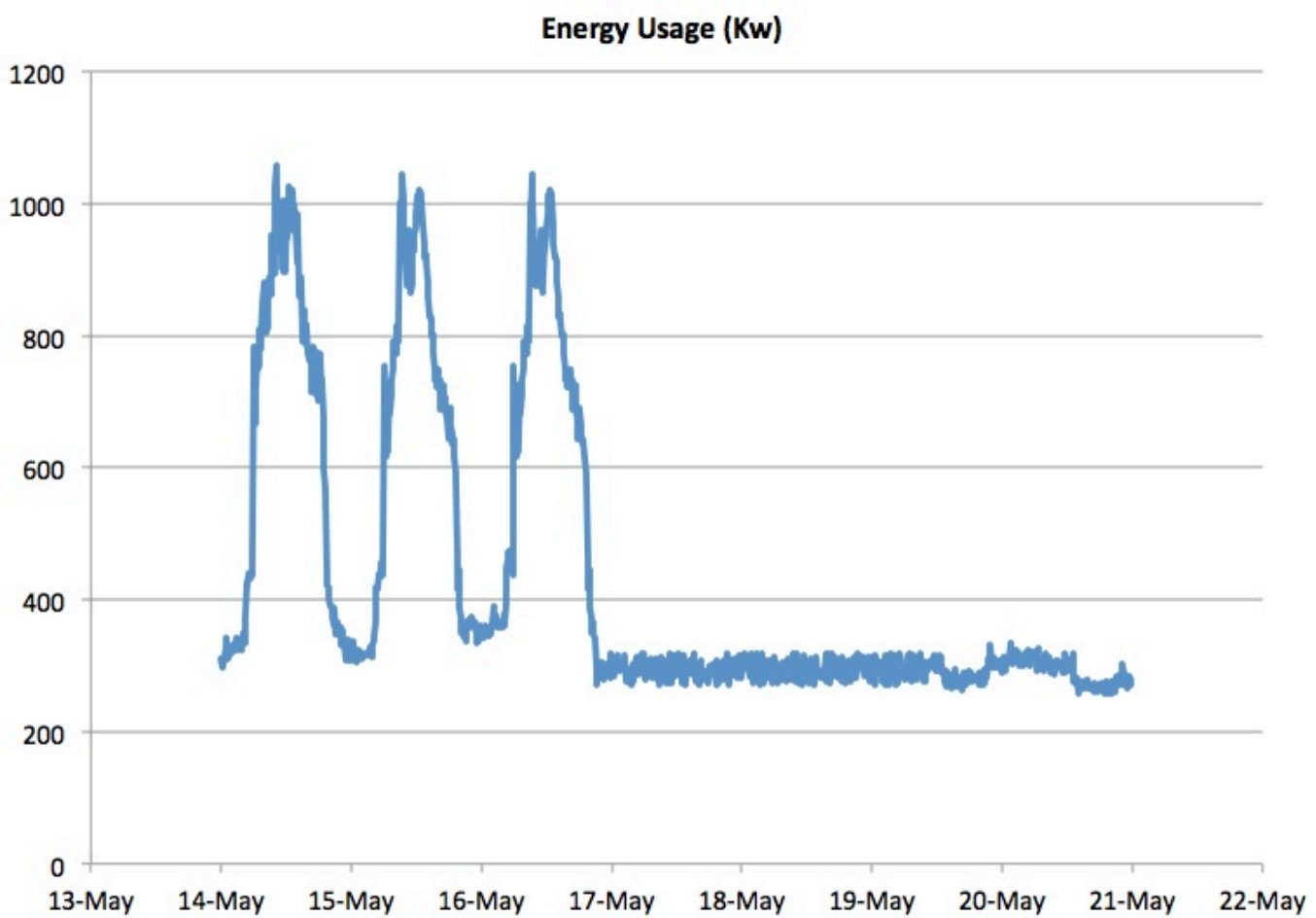
Static models → Continuous learning

Non-temporal

Application: Smart Meters



Application: Smart Meters



Machine Learning Challenges

Extremely manual → Automated

Static models → Continuous learning

Non-temporal → Inherently temporal

Machine Learning Challenges

Extremely manual → Automated

Static models → Continuous learning

Non-temporal → Inherently temporal

Batch focused

Machine Learning Challenges

Extremely manual	→	Automated
Static models	→	Continuous learning
Non-temporal	→	Inherently temporal
Batch focused	→	Streaming architecture

Properties Of Autonomic Machine Learning

Automated

Continuous learning

Inherently temporal

Streaming architecture

Talk Outline

What are the challenges of traditional machine learning?

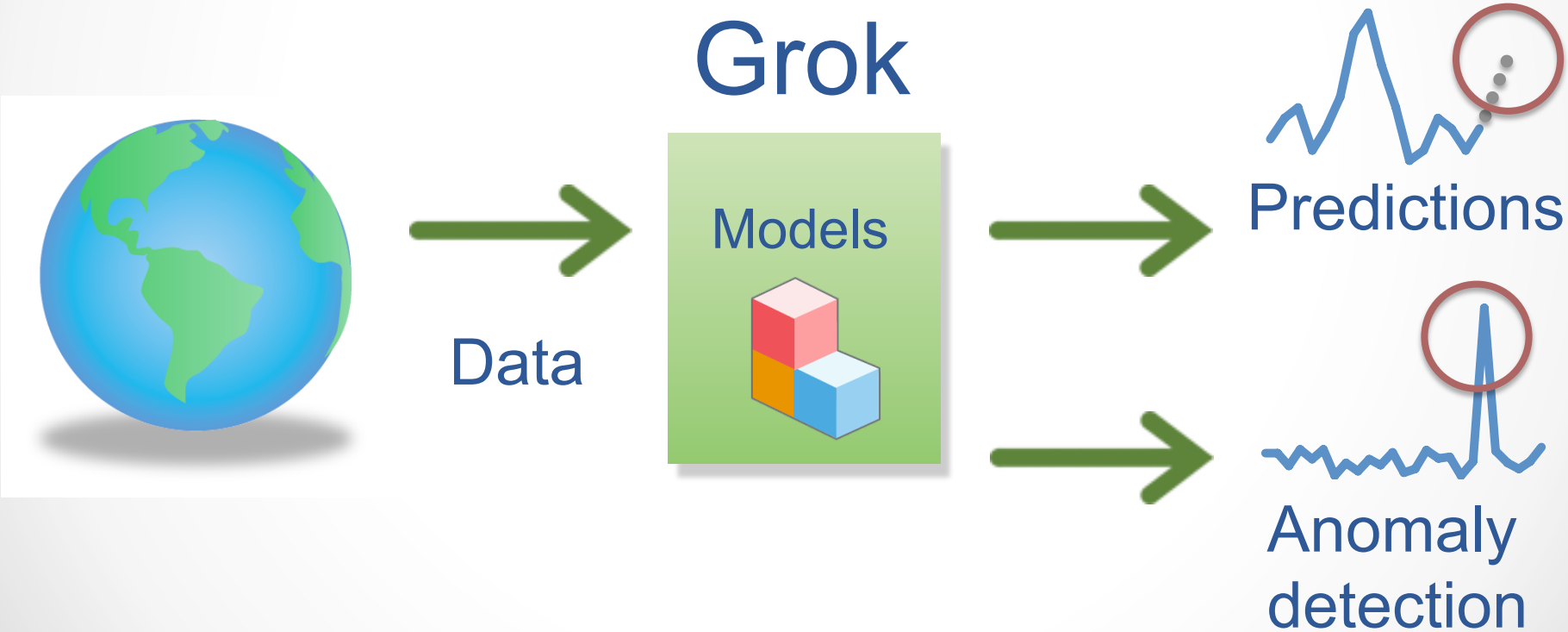
What are the properties of autonomic machine learning?

Numenta's algorithm and Grok product architecture

The future – where should we go?

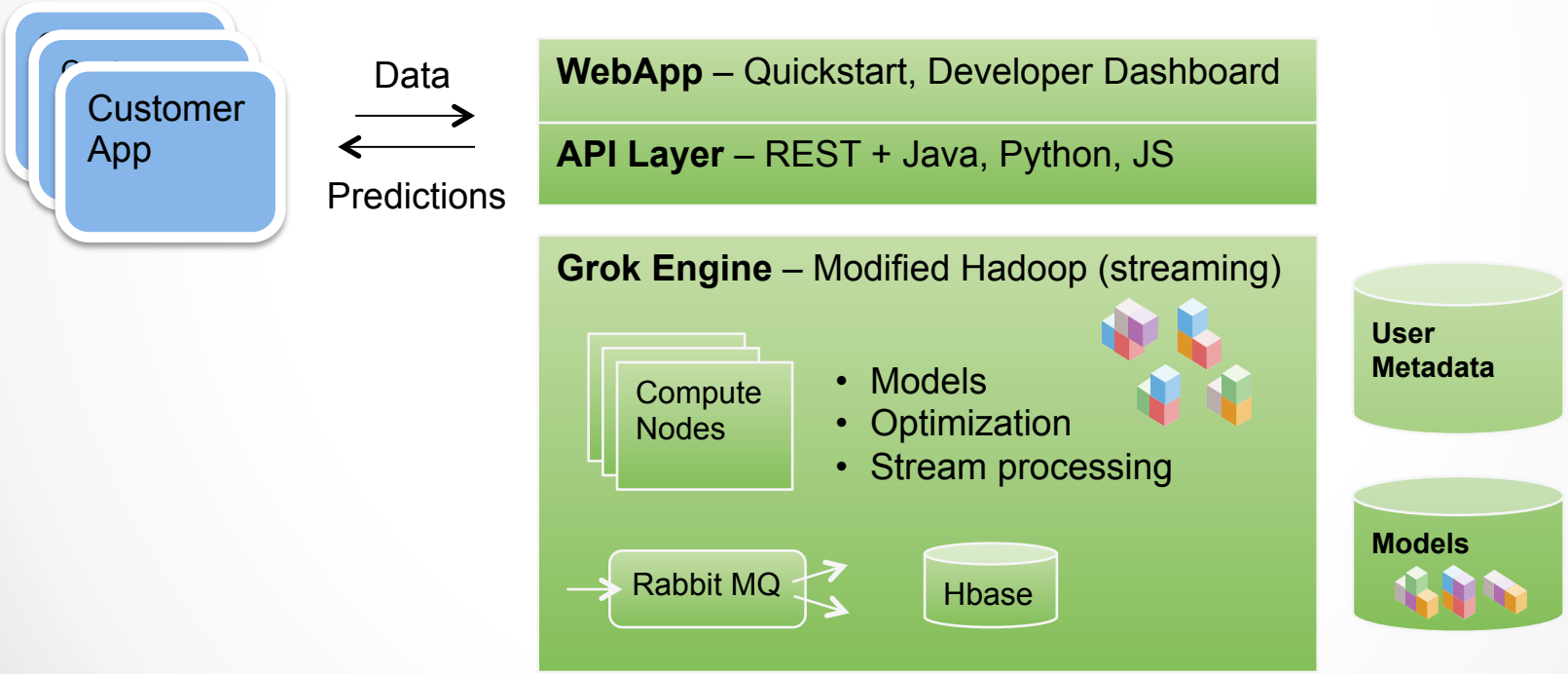
Grok

A prediction engine for automated streaming analytics

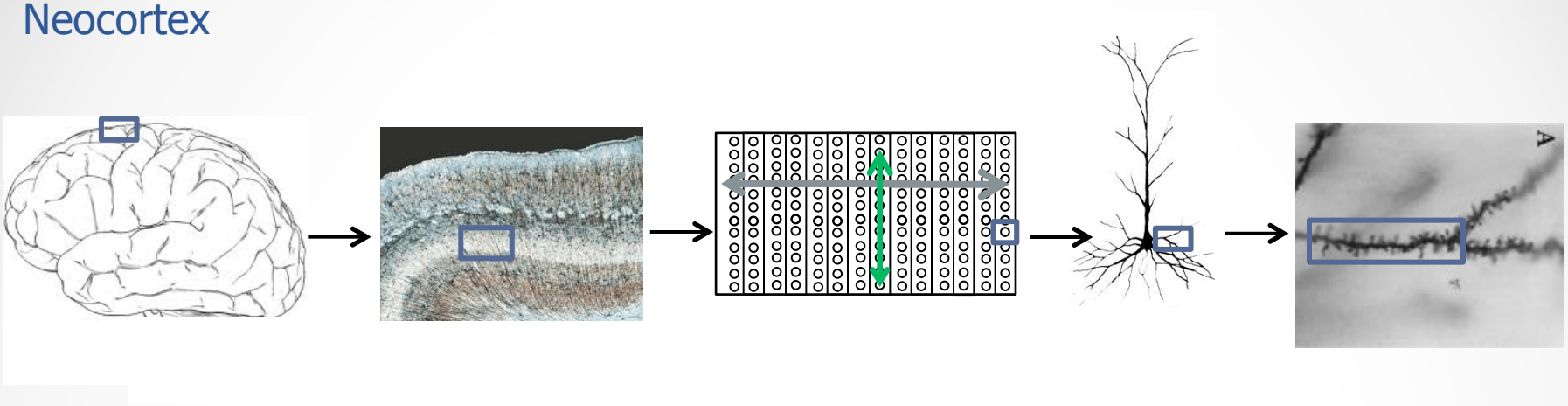


Grok Architecture

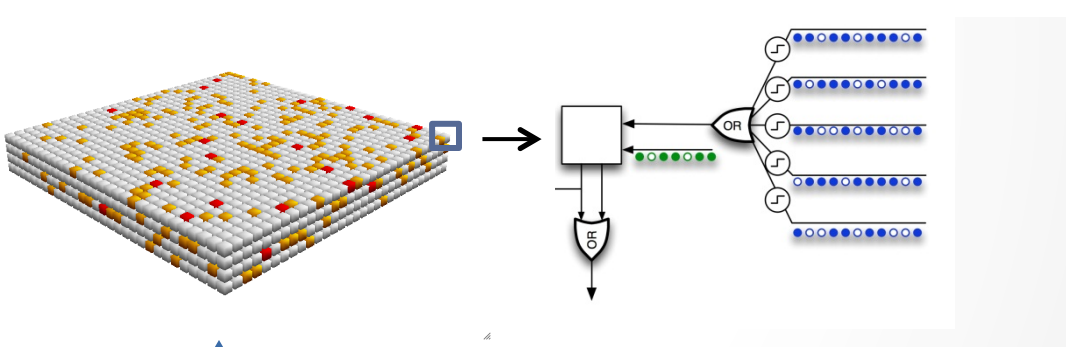
A cloud based architecture running on Amazon AWS



Neocortex



Cortical Learning Algorithm

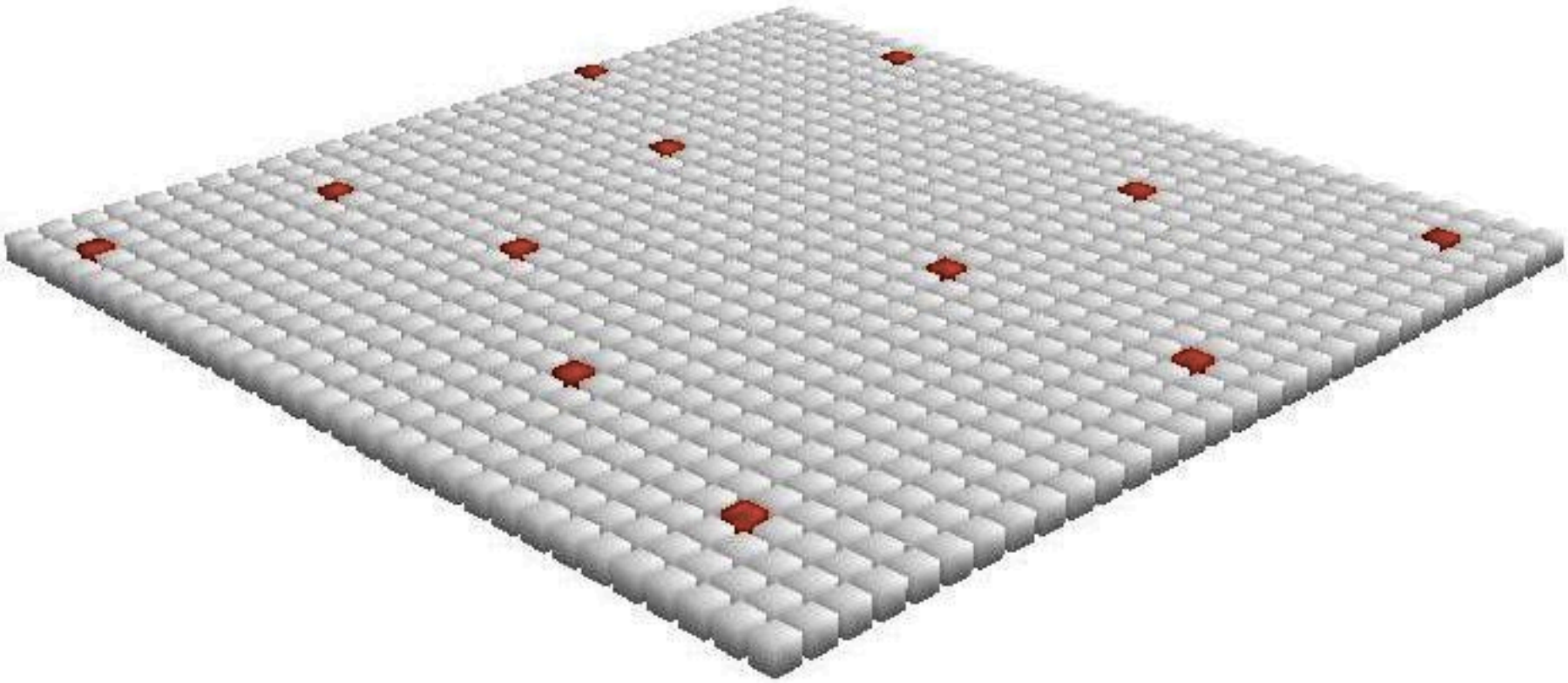


How does this structure learn?

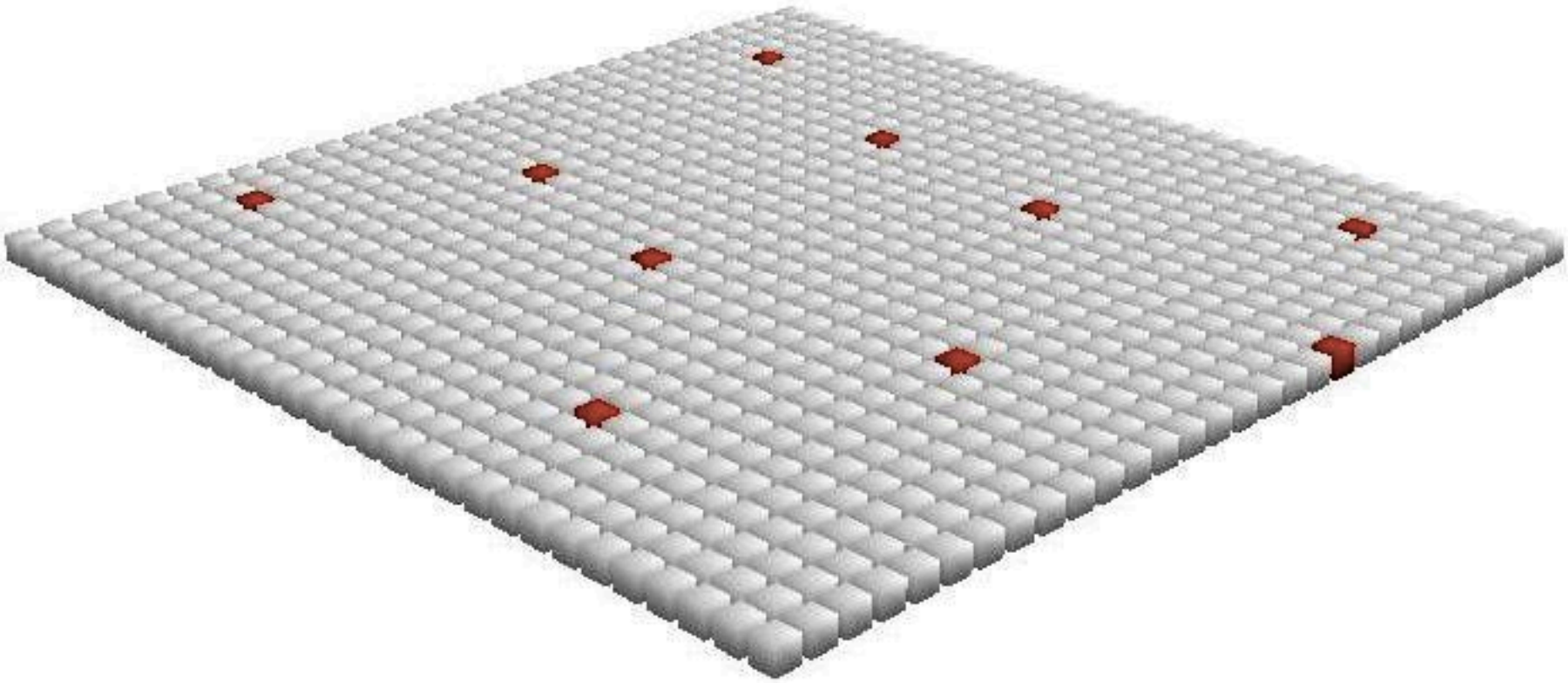
The Cortical Learning Algorithm

1. Learns to encode common static patterns
2. Learns the common sequences of those patterns
3. Makes predictions about future patterns
4. Continuously creates, strengthens and weakens cell connections

Static pattern (time =1)

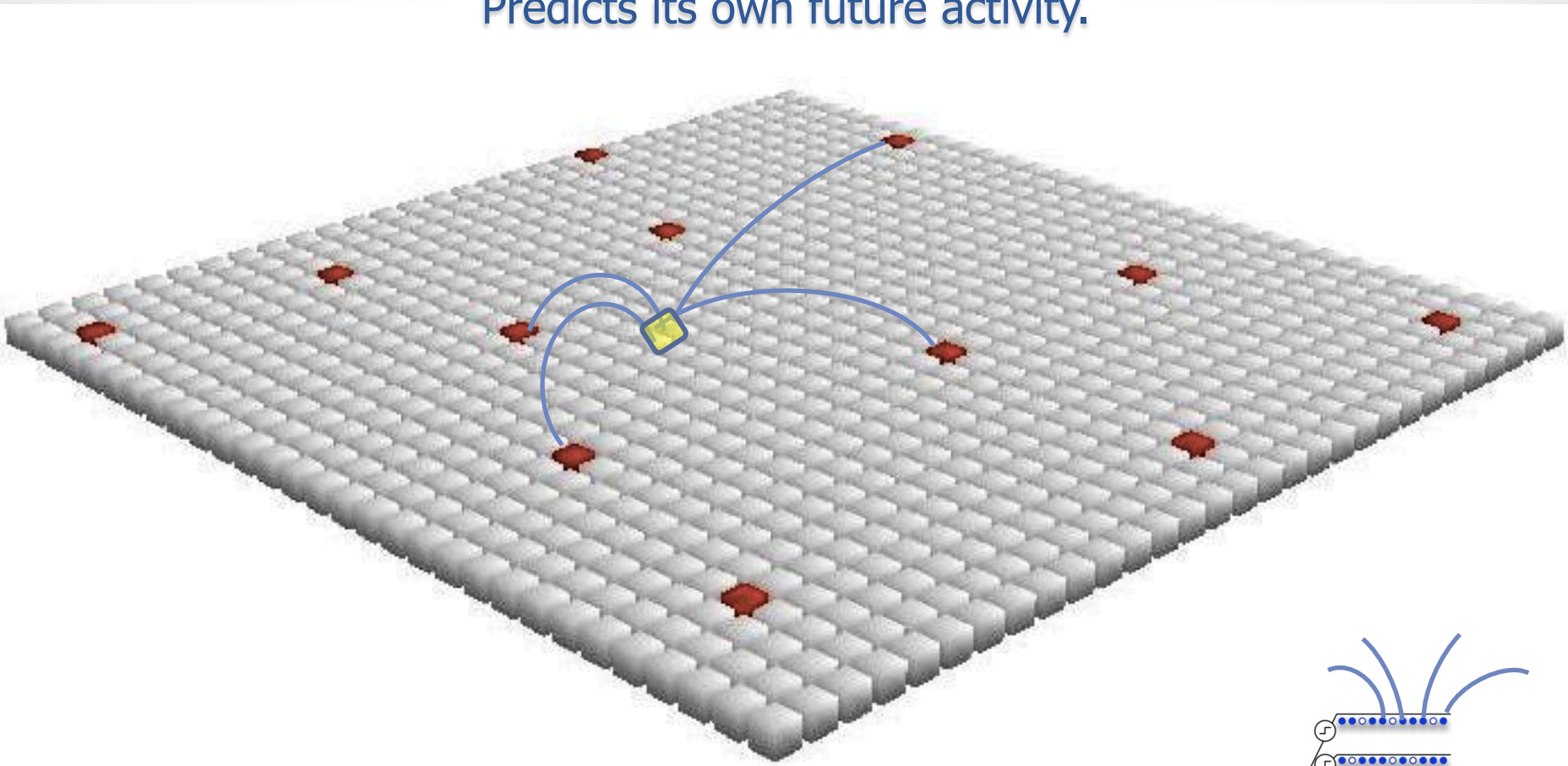


Static pattern (time =2)

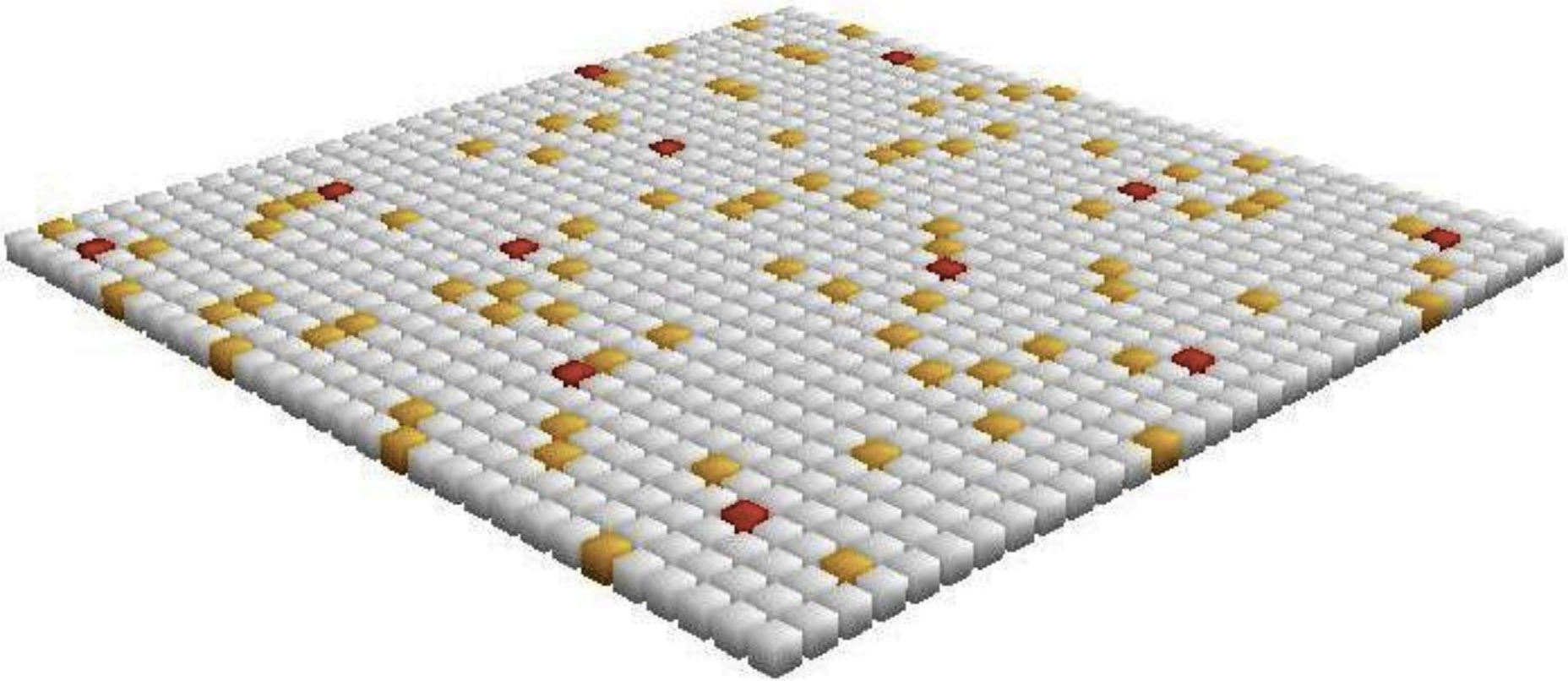


Cells form connections to previously active cells.

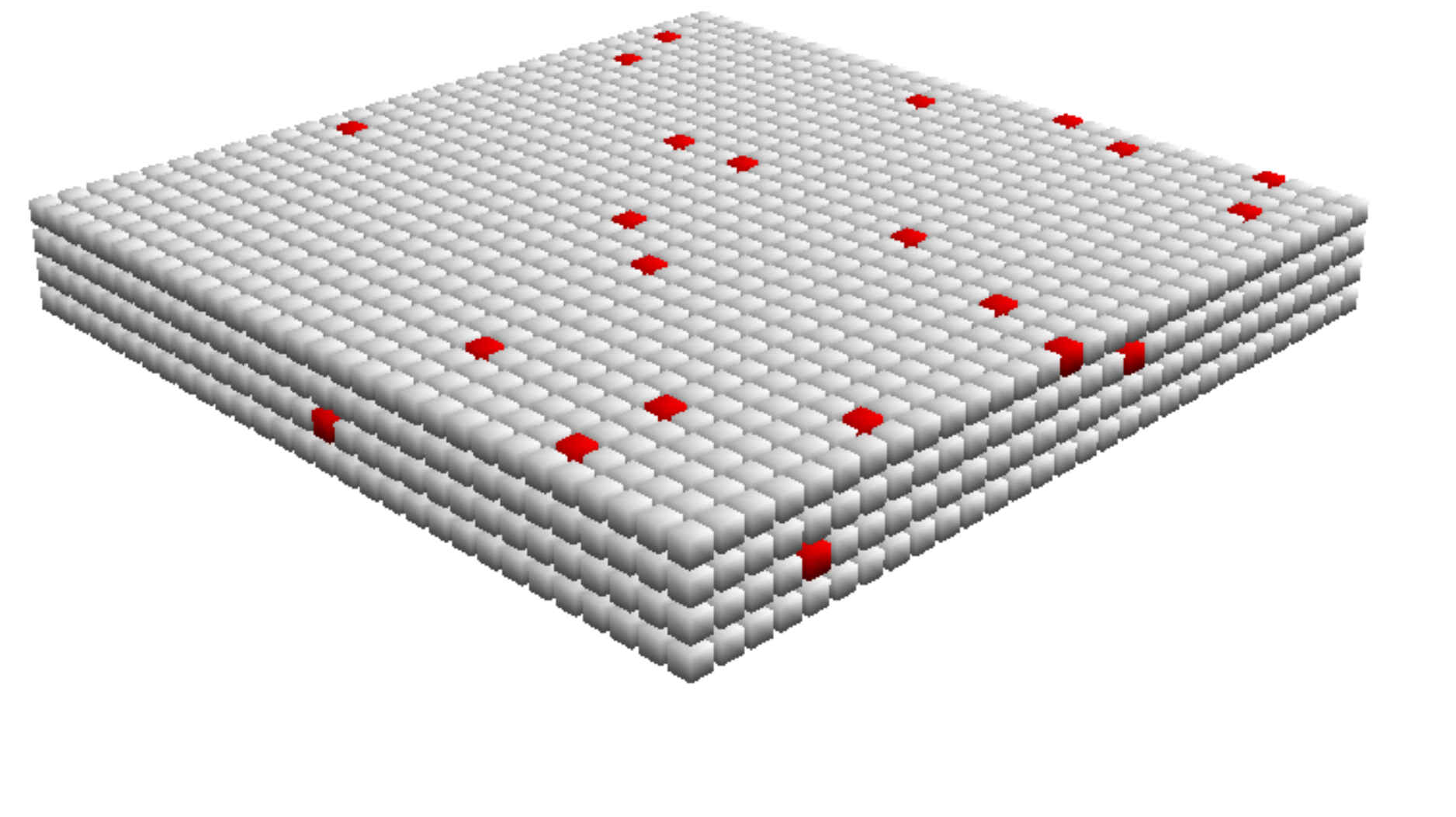
Predicts its own future activity.



Multiple predictions can occur simultaneously.

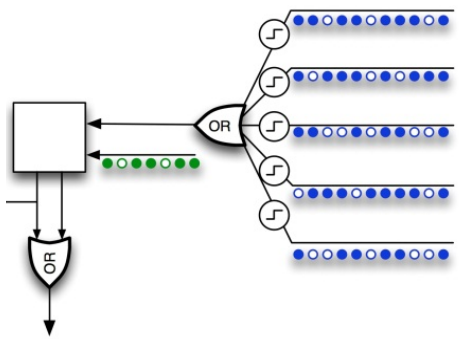


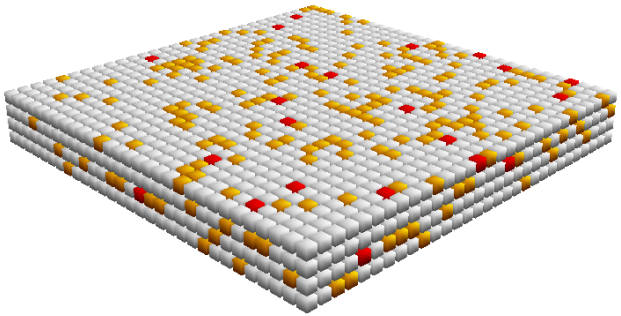
High order sequences are enabled with multiple cells per column.



Typical Implementation Size

Columns	2000
Cells per column	30
Segments per cell	128
Connections per segment	40





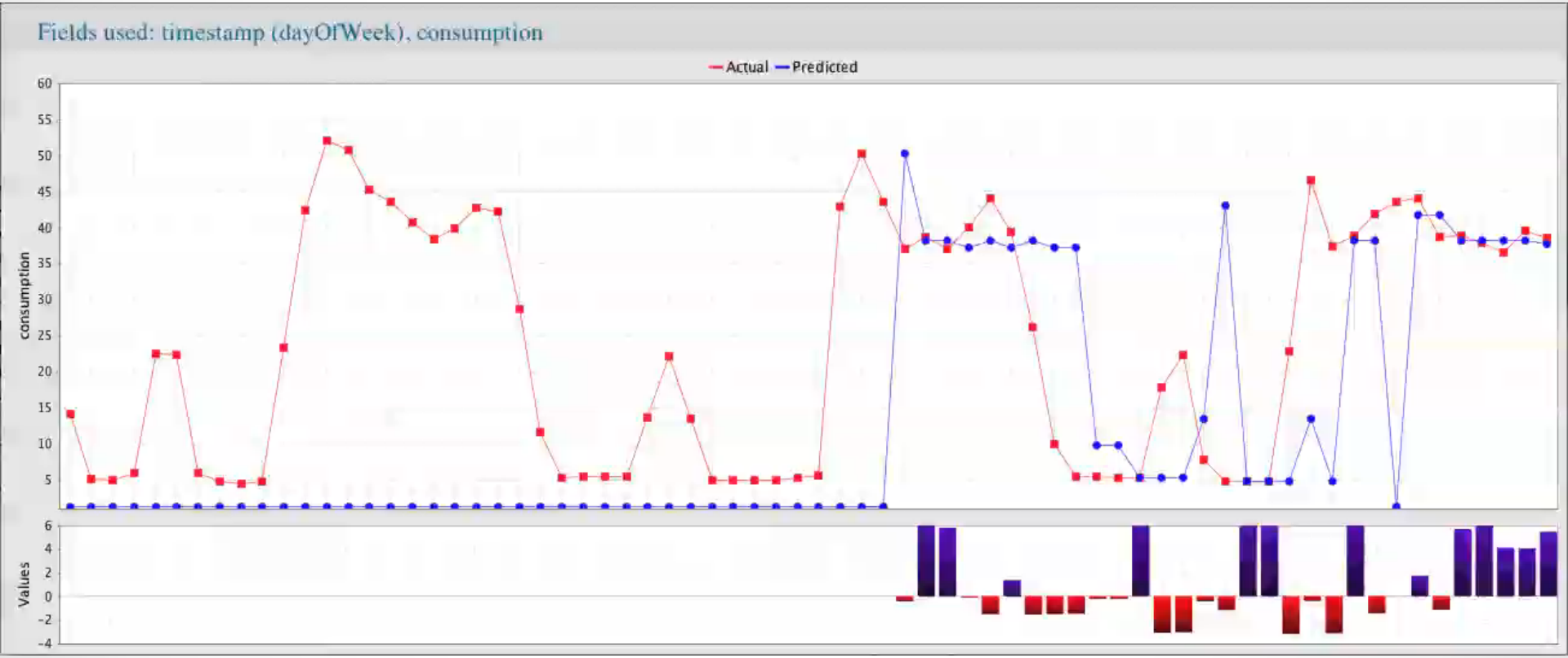
Storage capacity for patterns and sequences is huge

Automated Streaming Architecture

Grok automatically estimates model parameters.

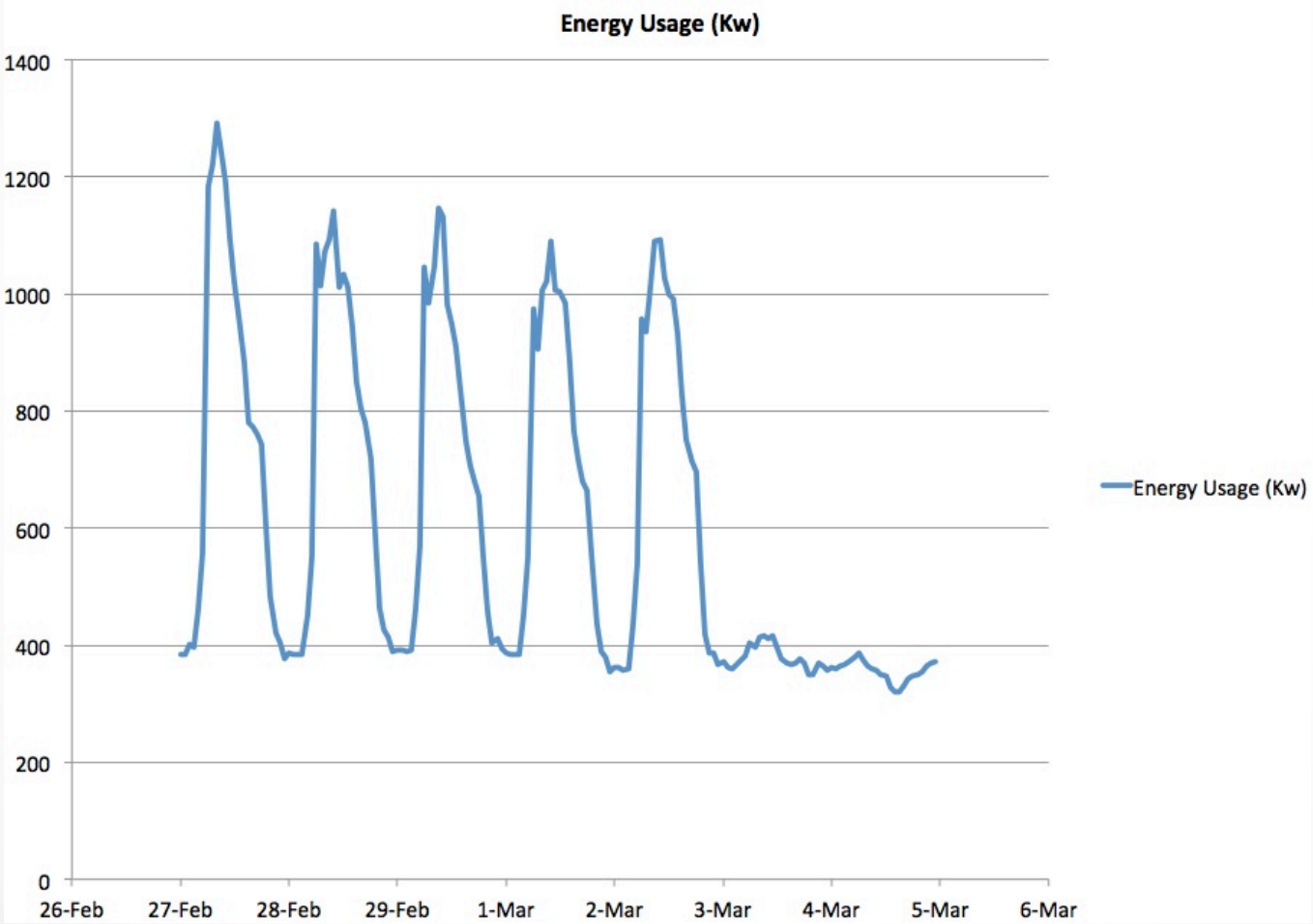
Users stream data and it starts making predictions within minutes.

Grok Learning Demo

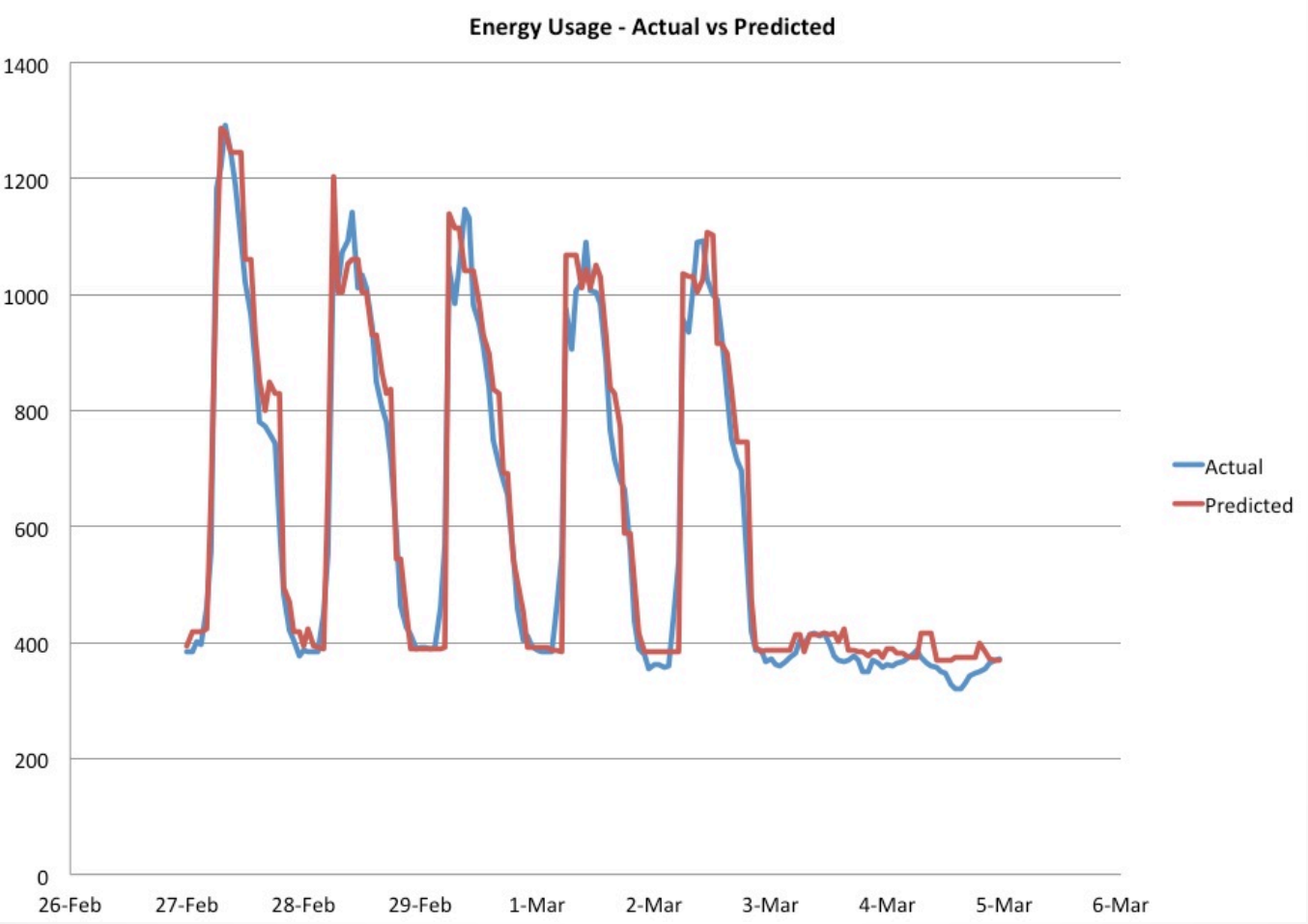


Application: Smart Meters

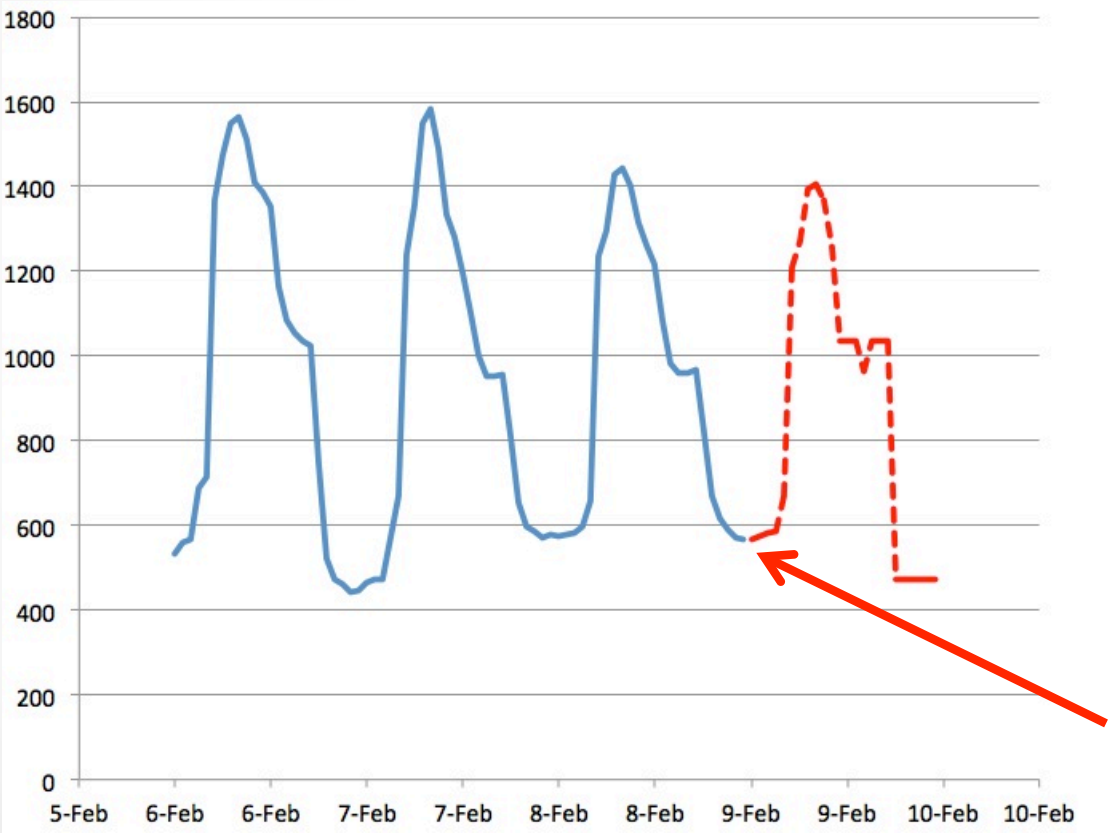
Typical Energy Profile



Grok Model Predictions

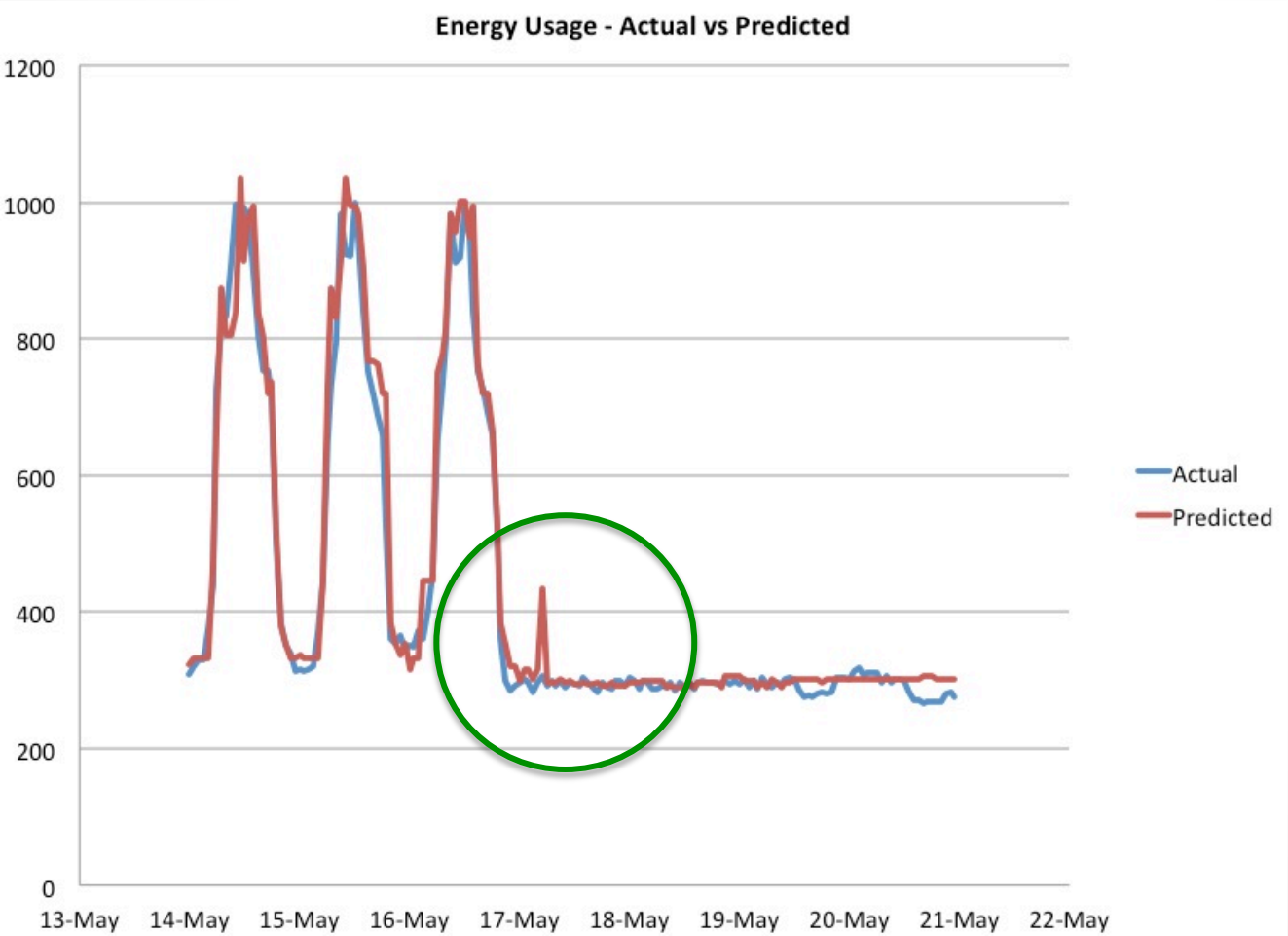


Long Range Temporal Predictions



At midnight, Grok makes predictions for next 24 hours

Sequence Memory

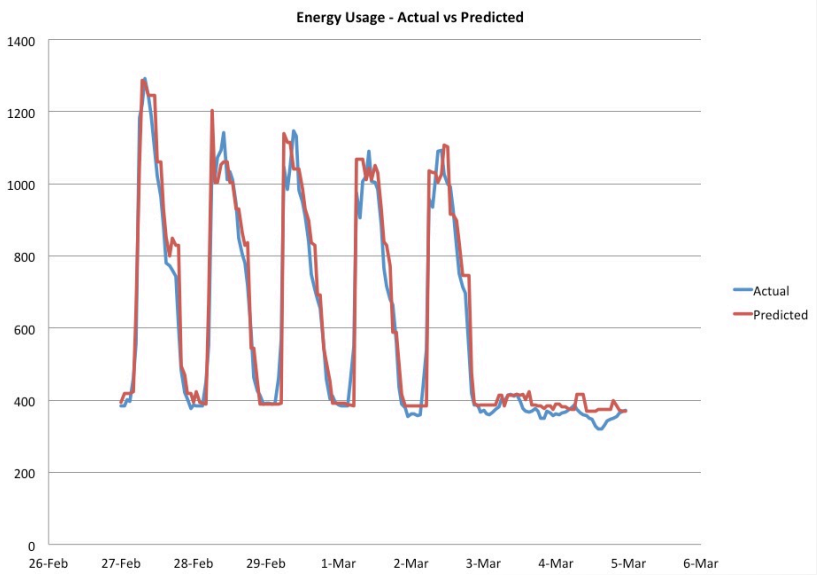


Smart Meter: Results

Model creation time about 60 minutes

Overall prediction accuracy within 2.5% to 5% of actuals

Grok able to make hourly, daily, weekly, and monthly predictions

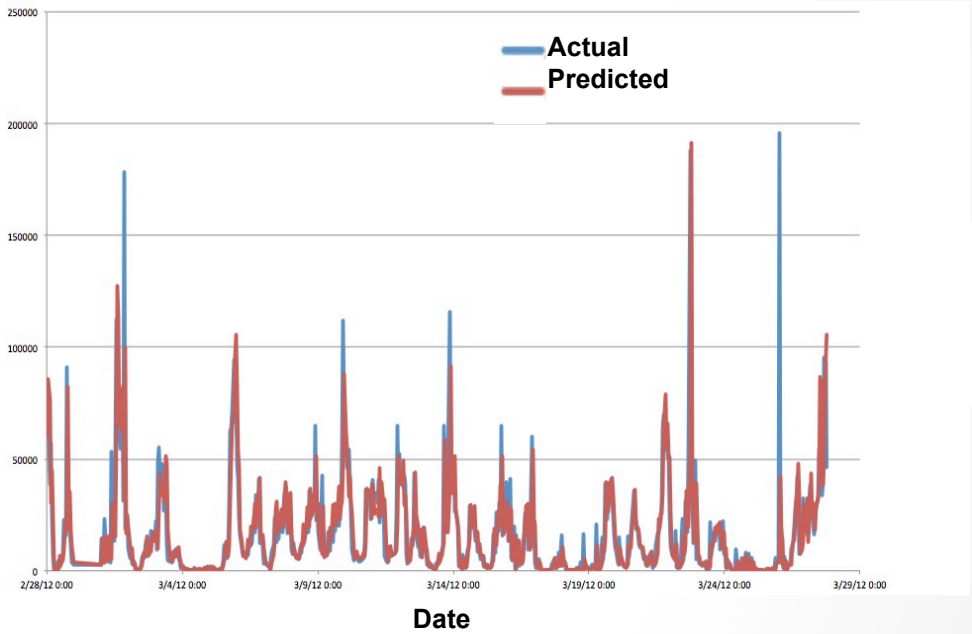


Managing Server Capacity

Grok used to predict server demand

Used to provision instances ahead of time

Results show approximately 15% reduction in AWS cost



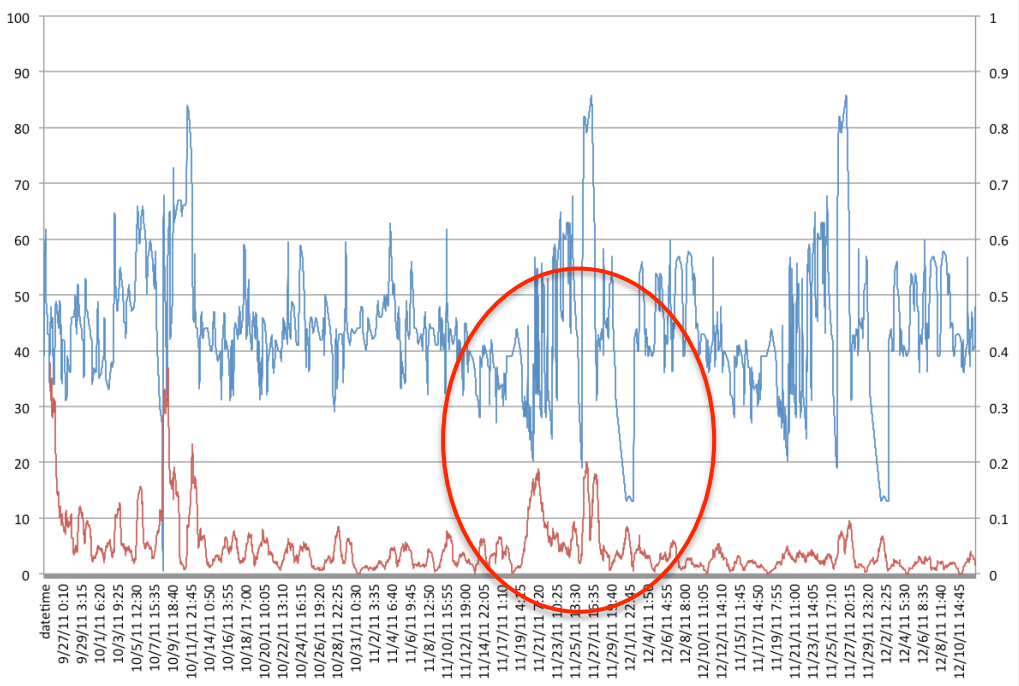
Incoming server demand, Actual vs. Predicted

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

Numerous Applications For Grok & Autonomous Machine Learning

- Energy prediction
- Auto provisioning
- Preventive Maintenance
- Ad network optimization
- Credit card fraud detection
- Power management
- Server monitoring
- Malware detection
- Inventory management

Properties Of Autonomic Machine Learning

Automated

Continuous learning

Inherently temporal

Streaming architecture

What's Next?

Autonomic Machine Learning is inevitable

Machine Learning as a field needs to expand its focus beyond just algorithms, and also address system level issues

Are there lessons from Autonomic Computing that should be applied to machine learning?

Hope: Autonomic Computing community is able to act as a catalyst, working with the Machine Learning community

Oh, and contact me if Grok sounds interesting for your application!

Thank you

Contact Info:
Subutai Ahmad
sahmad@numenta.com