AUTOMATED STREAMING DATA ANALYSIS
It's Not The Same Game

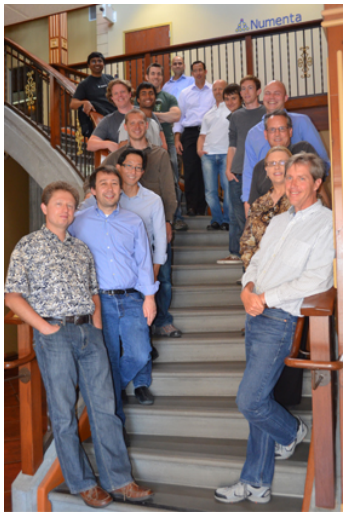
Subutai Ahmad
VP Engineering, Numenta

Numenta

Small startup located in Redwood City

Approximately 24 employees

Founded in 2005 by Jeff Hawkins (Palm, *On Intelligence*)

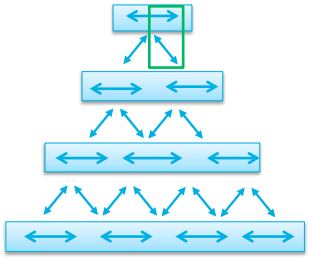


Numenta: Early Years

2005 – 2010:

Focus on creating a new theory of machine intelligence and machine learning

Developed Cortical Learning Algorithm (CLA), inspired by neuroscience



Numenta: Today

2011 – now:

Developed Grok, a cloud based product designed for streaming analytics

Status: in “private beta”, working closely with a number of customers

Talk Outline

Streaming data and its challenges

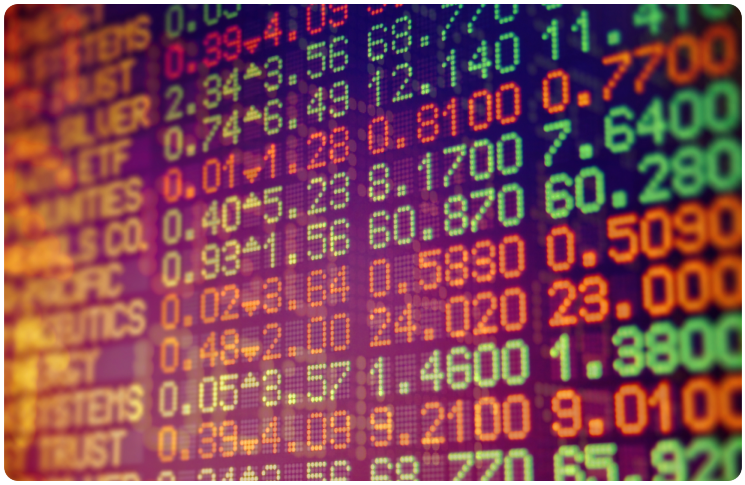
Grok architecture/algorithms

Applications

What Is Streaming Data?

A continuous source of data,
that is sequentially generated,
with no logical beginning or end.

Streaming data applications often involve
dynamically changing temporal patterns,
a very large number of sources,
and often require continuous decisions.



Non-Streaming Applications

- Accept/reject mortgage application
- Face recognition in photo albums
- Clustering collection of documents into topics
- Medical diagnosis based on symptoms and tests
- Post-hoc exploratory analysis

Streaming Applications: What Are The Challenges?

Requires high degree of automation
Temporal patterns are the norm
Non-batch methodologies for training/testing
Models must be continuously adaptive
Need to create very large number of models on the fly

Good mathematical framework for the above??

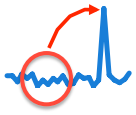
What Is Grok?

Real-Time, Streaming Data

GROK

Automatically Creates & Continuously Updates Models

Predictions

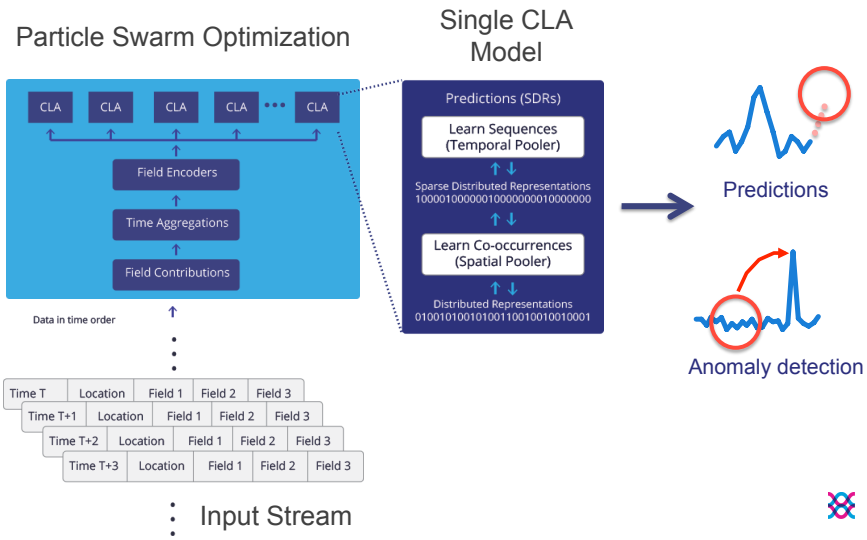


Anomaly detection

Grok User's Viewpoint

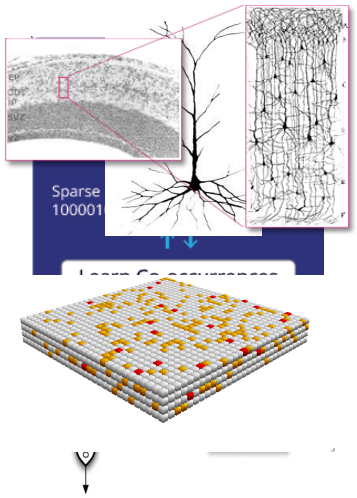
- Straightforward REST based API
- Answer a few simple questions
  - Data model: number, types of each field
  - Which field do you want to predict?
  - How often do you want to make predictions?
  - (optional) What is your cost function?
- User does not need to understand machine learning

Automated Model Building In Grok

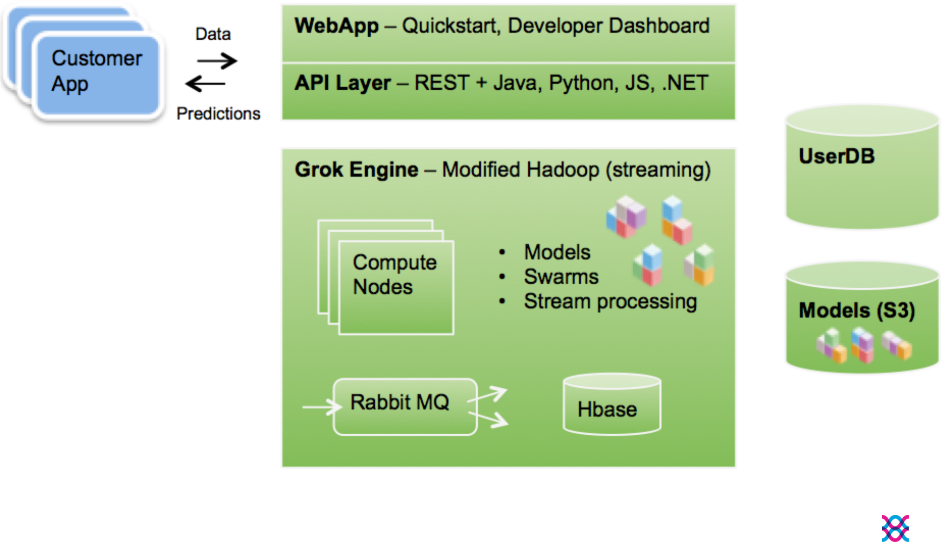


Cortical Learning Algorithm (CLA)

- Models neuronal dendritic functions
- Models laminar structure of cortex
- Sparse distributed representations
- Learns high order temporal sequences
- Makes multiple simultaneous predictions with probabilities
- Learns with every time step



Grok System Architecture

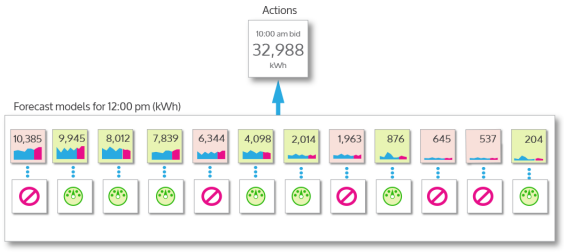


Application: Smart Meter Demand Response

Leading Energy Companies

Accurate energy prediction = $$

Q: Can this customer participate next hour?



Grok

Grok predicts energy usage & recommends participation

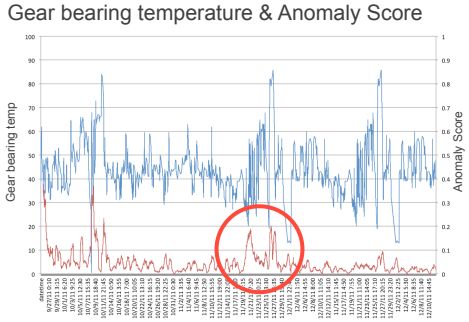
- Hourly predictions and actions
- Highly accurate predictive models
- Dynamically adapts as energy usage evolves
- Initial deployment > 20,000 models

Application: Predictive Maintenance

European Company

800 windmills, 34 sensor points, 1 analyst

Q: When should I repair the windmills?

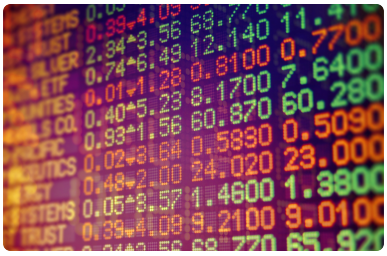


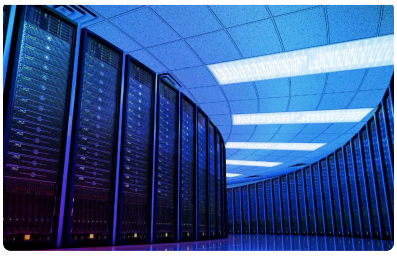
Grok

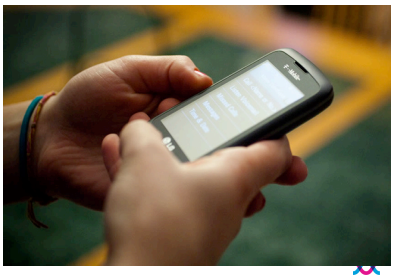
Grok finds anomalies, alerts operators to dispatch teams.

- Automatically learns normal behavior
- Unusual behavior flagged as anomalies
- Detects spatial and temporal anomalies
- Continuously adapts to changes

Other Applications Of Grok







Summary

Streaming analytics

- Streaming data poses interesting challenges for machine learning applications

Grok

- Automated, continuously learning system designed for streaming applications

The future

- Streaming analytics will be the dominant model for ML
- Potentially fruitful arena for research

Thank You!

Contact information:
Subutai Ahmad
sahmad@numenta.com