FINANCE + DEEP LEARNING

SKYMIND / DEEPLEARNING4J 2015

WHICH OUTCOMES MATTER?

BASIC APPLICATIONS

Anomaly Detection for Compliance

- Rogue Traders, Fat Fingers
- Fraud, Money Laundering Detection

Trading Strategies

- Risk and Reward Prediction

Personnel and Customer Management

- Recruiting / Turnover prevention
- CRM for trading platforms

THE COSTS OF FRAUD / YEAR

- **Merchants : $190 billion**
- **Healthcare : $80 billion**
- **Seniors : $36 billion**
- **Banks : $11 billion**

(Huge, & probably understated)

ANOMALY DETECTION ACCURACY

- **Traditional Decision Trees : 70-75%**
  - Former state of the art
- **Deep Neural Networks**
  - Setting records above 90%
  - Cuts the costs of fraud in half

HIGH-LEVEL WORKFLOW

- **Train neural net with labels**
  - Users labeled “risk” or “not risk”
- **Pass real-time data through neural net**
  - Voice, text, transaction logs
- **Neural net surfaces high fraud scores**
- **Act on fraud scores**

SUPERVISED MACHINE LEARNING

- **Labels: Human judgments on historical data; e.g. *fraud* or *not_fraud***
- **Statistical analysis of training data**
- **Model finds correlations between input data and human-applied labels**
  - 1,000s of features
  - Millions of fraud patterns

FEATURE EXTRACTION & CLASSIFICATION

1) FEATURE EXTRACTION

- **Deep Neural Networks**
  - Automatic
  - Real-time

2) CLASSIFICATION

- **Logistic Regression**
- **Naïve Bayes**

WHAT ARE FEATURES?

IP Address: 109.189.24.260

Billing Name: Chris Nicholson

Billing Address: San Francisco, CA

Email Address: chris@skymind.io

Credit Card: 5590xxxxxxxxxxx

Item Purchased: Bookshelf

Cost: 150.00 USD

Authorization Result: Success

TYPES OF FEATURES

- Event features
- State features
- Temporal features
- Graph features

NOISE IS EVERYWHERE

- Wrong labels
- Duplicate labels
- Bad integrations
- Incomplete integrations
- Missing fields
- Bugs
- System downtime

HOW DOES DEEP LEARNING WORK?

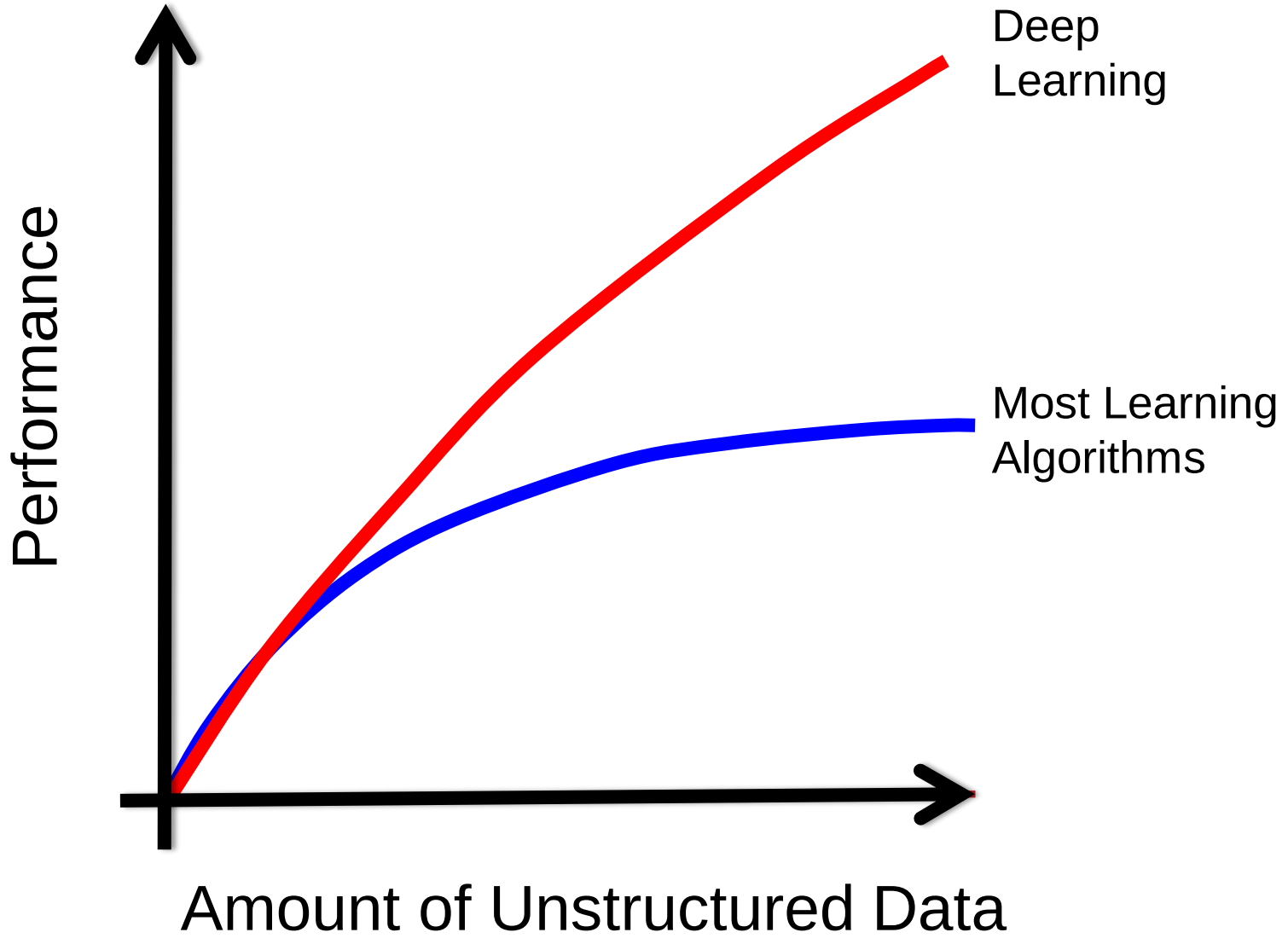
DEEP LEARNING IS MACHINE PERCEPTION FOR…

IMAGES • FACES • HANDWRITING	**TEXT** • WORD CHOICE • WEB LOGS • SOCIAL GRAPH
SOUND • VOICE • LATENT	**TIME SERIES** • TRANSACTIONS • TRADE TIMING

RECORD-BREAKING ACCURACY

- FACIAL RECOGNITION = 97% accuracy
- GENERAL IMAGE RECOG. = 95%
- SPEECH RECOGNITION = 81%
- VIDEO ACTIVITY RECOG. = 52% - 94% (Varies by dataset)
- TEXT CLASSIFICATION = 94%

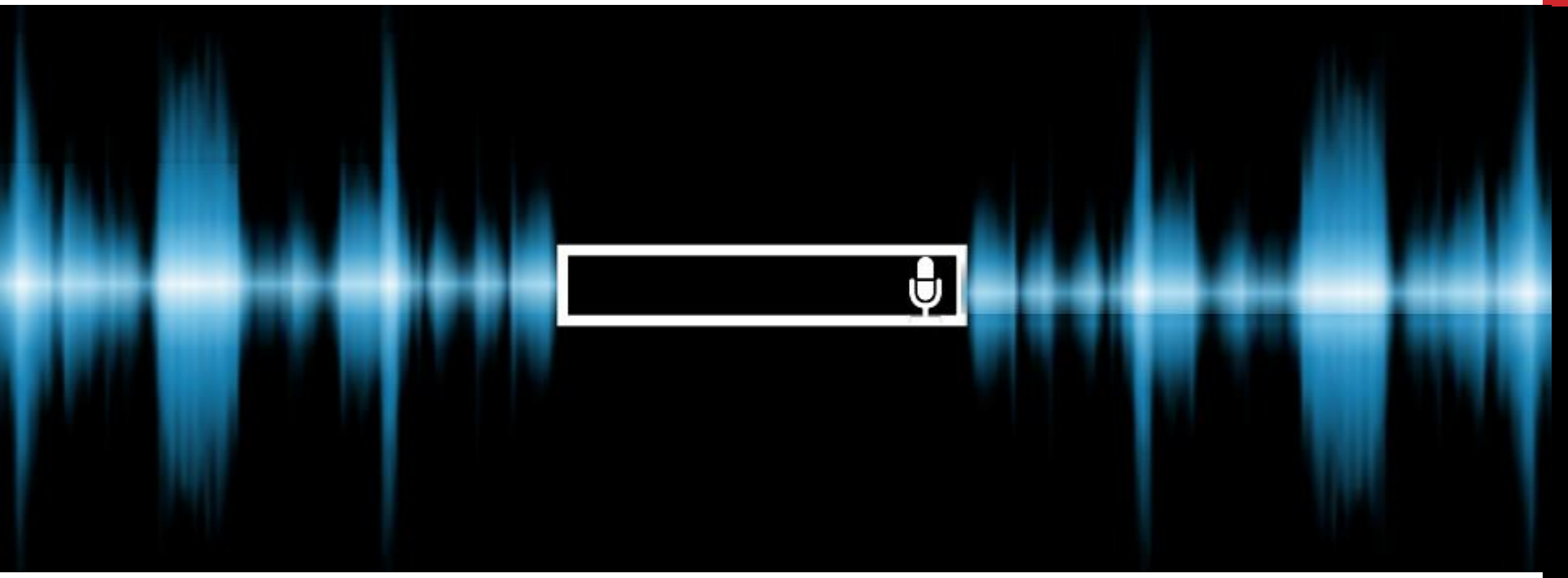
BIG DATA & DEEP LEARNING



VOICE AUTHENTICATION

DL DETECTS LATENT FACTORS

SOUND HAS SUBTLE FEATURES

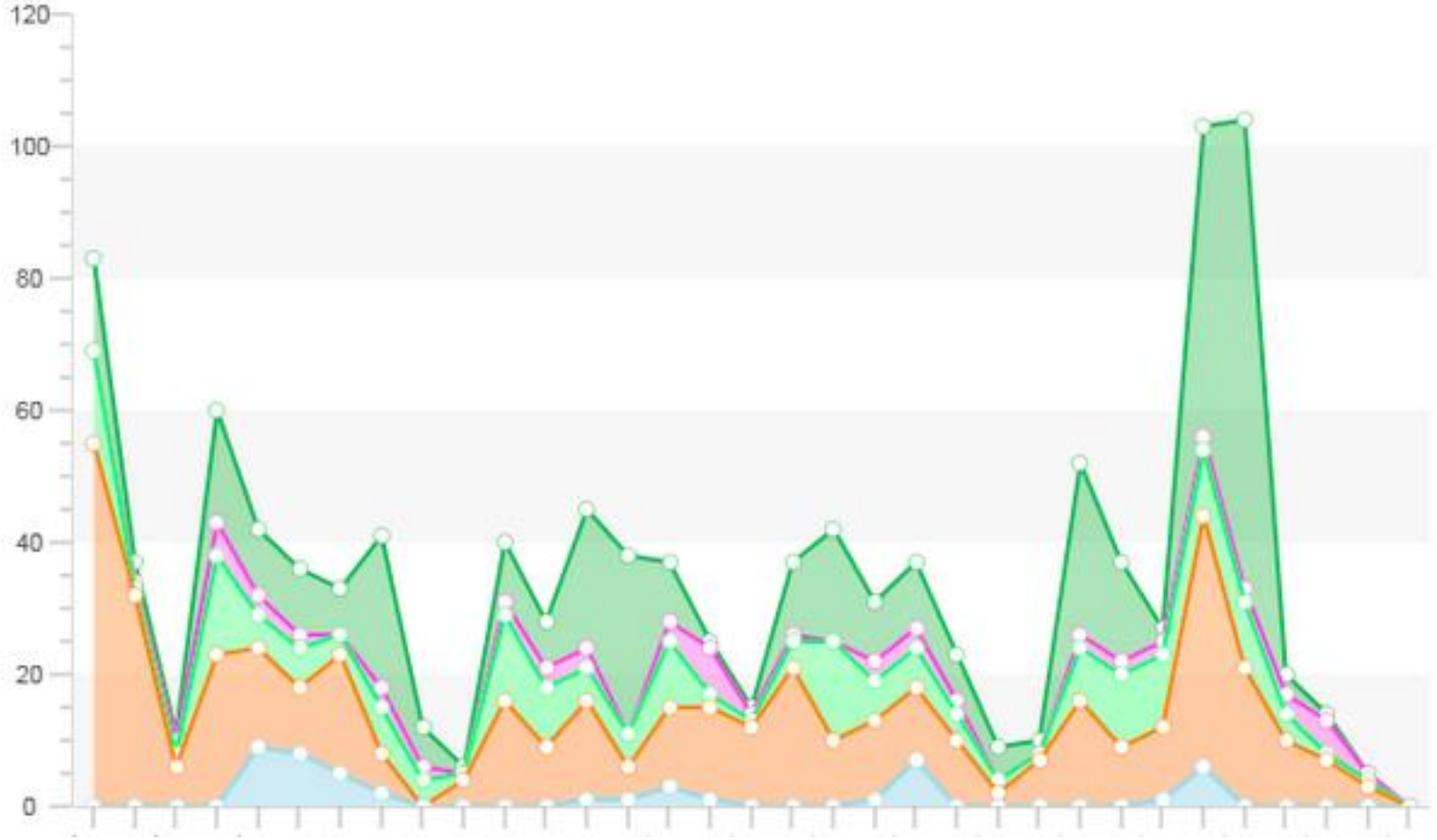


TIED TO LOCATION, DEVICE, WIRES

MACHINE TRANSCRIPTION MAKES SOUND SEARCHABLE

TEXT

ROGUES USE CODE WORDS



AND RARELY EMAIL GROUPS

TRADING + DEEP LEARNING

BASIC APPLICATIONS

Natural Language Processing

- Sentiment Analysis on Trad. & Social Media related to Companies and Products
- Similarity Detection to Anticipate Market Moves

Reinforcement Learning

- DeepMind

Options Pricing With Neural Networks

Global Image Analysis (Markets, Parking Lots)

- Estimating real-time demand

BEHAVIORAL ANALYTICS

BASIC APPLICATIONS

Customer Relationship Management

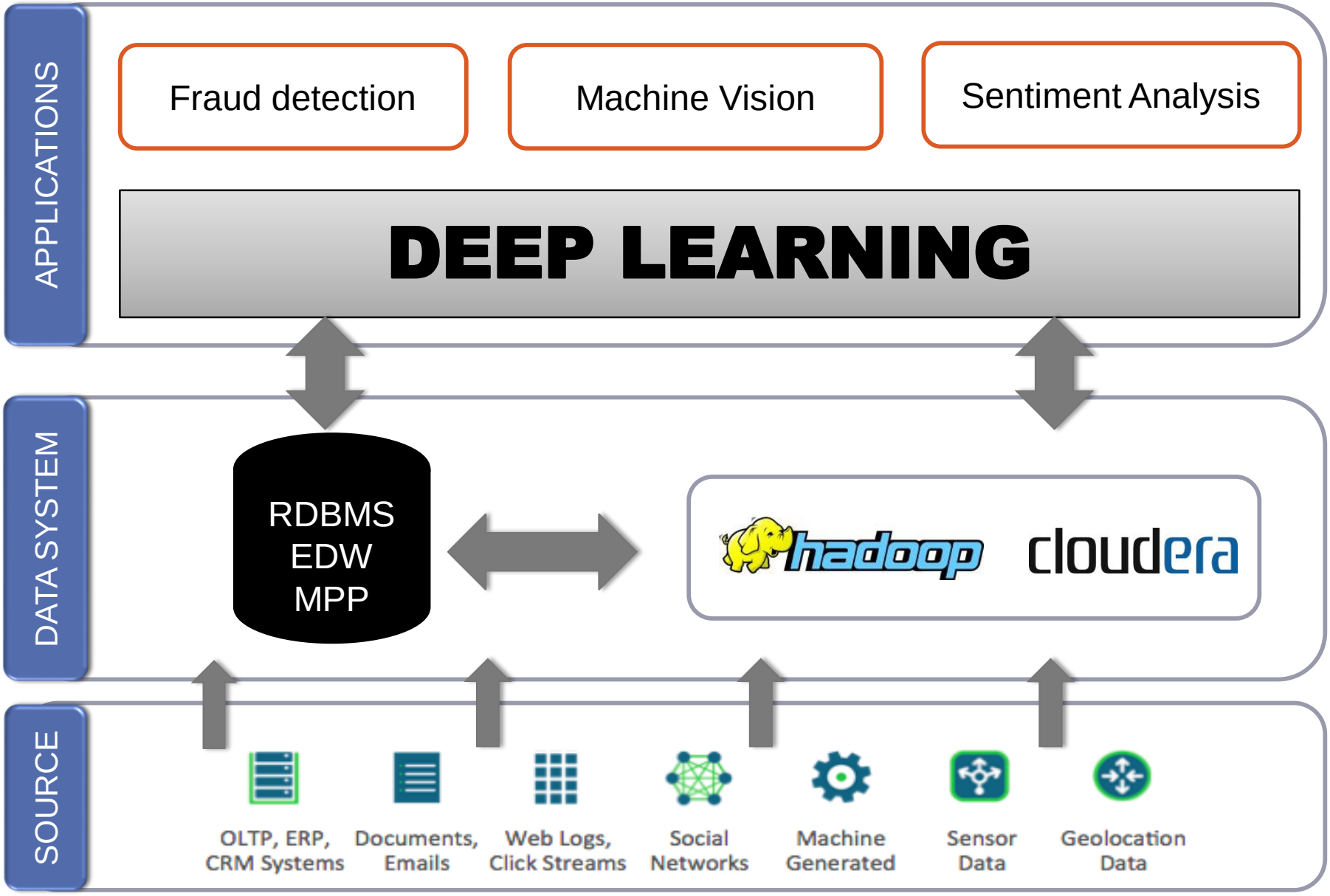
- Clickstream, web log and trading analysis of clients to predict when they will churn, when they can be upsold

Personnel Management

- Anomaly detection for staff trading patterns
- Predict and prevent employee churn

OPEN-SOURCE DEEP LEARNING

A "GOOGLE BRAIN" FOR INDUSTRY



WHAT TO LOOK FOR

- **All major neural nets**
  - RBM, ConvNet, LSTM, RNTN, DBN, SDA, Deep autoencoder
- **Composable framework**
- **Hadoop, Spark, Akka integrations**
- **Parallelized GPUs and CPUs**
- **Cross-platform**
- **Apache 2.0 License**

CROSS-PLATFORM

- LINUX (SERVERS)
- WINDOWS (DESKTOP)
- OSX
- ANDROID (MOBILE)

HOW WILL AI STANDARDIZE?

An open-source framework.

Open source dominates OS with Linux, and big data with Hadoop.

Open source will win AI with Deeplearning4j:

- **Distributed deep learning on GPUs**
- **Serving 10M Java/Scala programmers**

QUESTIONS?

HELP@SKYMIND.IO