The State of Machine Learning in 2014

in 30min

Paris Data Geeks @ Open World Forum
October 2014

Olivier Grisel
@ogrisel
Datageek, contributor to scikit-learn, works with Python / Java / Clojure / Pig, interested in Machine Learning, NLProc, {Big|Linked|Open} Data and braaains!
Paris, France · http://github.com/ogrisel

INVENTEURS DU MONDE NUMÉRIQUE

Content Warnings

This talk contains buzz-words
and highly non-convex objective functions
that some attendees may find disturbing.

The State of ~~Machine~~ Deep Learning in 2014

in 30min

Paris Data Geeks @ Open World Forum
October 2014

Olivier Grisel
@ogrisel
Datageek, contributor to scikit-learn, works with Python / Java / Clojure / Pig, interested in Machine Learning, NLProc, {Big|Linked|Open} Data and braaains!
Paris, France · http://github.com/ogrisel

INVENTEURS DU MONDE NUMÉRIQUE

Outline

- ML Refresher
- Deep Learning for Computer Vision
- Word Embeddings for Natural Language Understanding & Machine Translation
- Learning to Play, Execute and Program

Quick refresher on what is Machine Learning

Predictive modeling ~= machine learning

- Make predictions of outcome on new data
- Extract the structure of historical data
- Statistical tools to summarize the training data into a executable predictive model
- Alternative to hard-coded rules written by experts

type (category)	# rooms (int)	surface (float m2)	public trans (boolean)	sold (float k€)
Apartment	3	50	TRUE	450
House	5	254	FALSE	430
Duplex	4	68	TRUE	712
Apartment	2	32	TRUE	234

features

target

samples (train)

type (category)	# rooms (int)	surface (float m2)	public trans (boolean)
Apartment	3	50	TRUE
House	5	254	FALSE
Duplex	4	68	TRUE
Apartment	2	32	TRUE

sold (float k€)
450
430
712
234

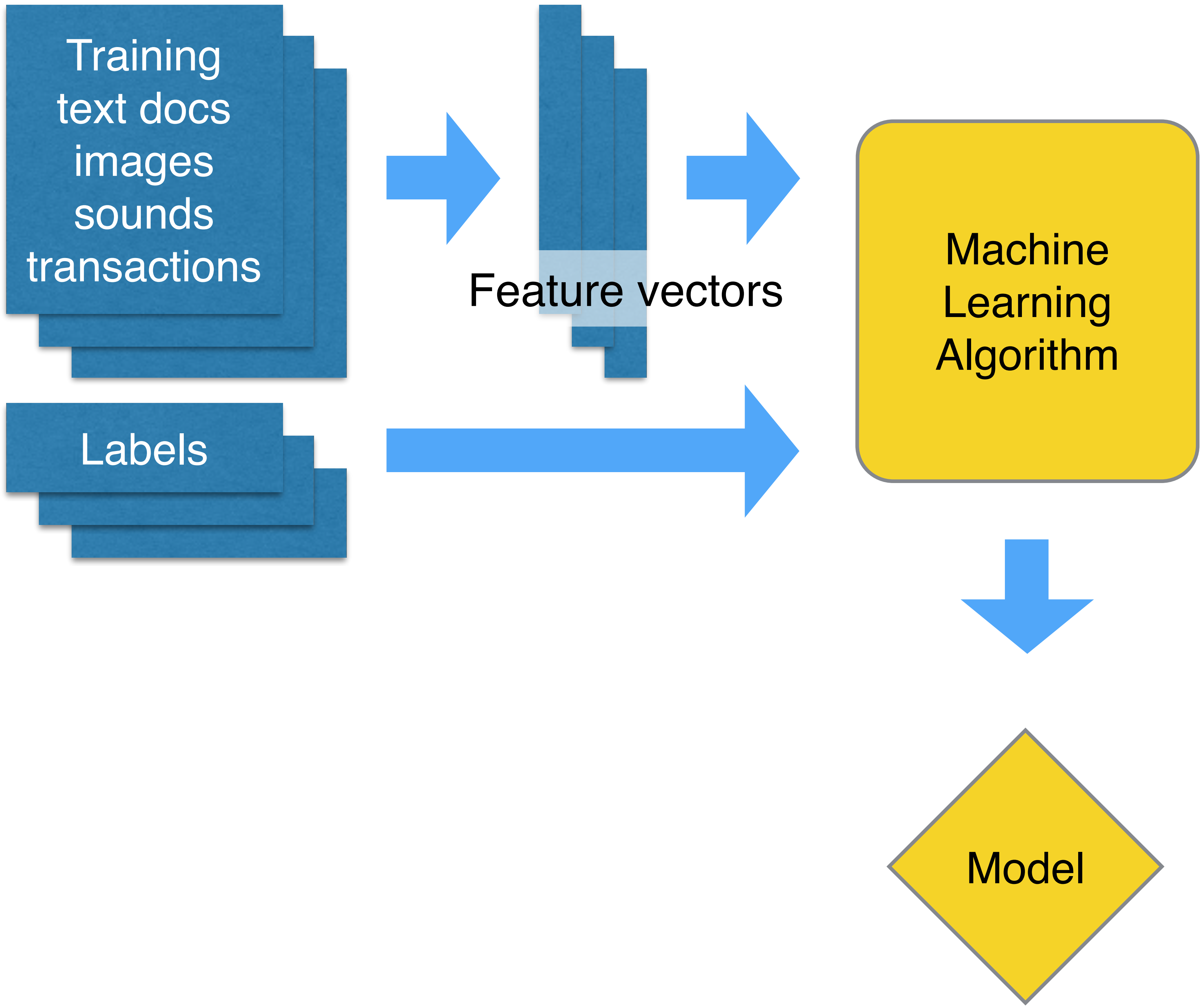
features | target

samples (train)

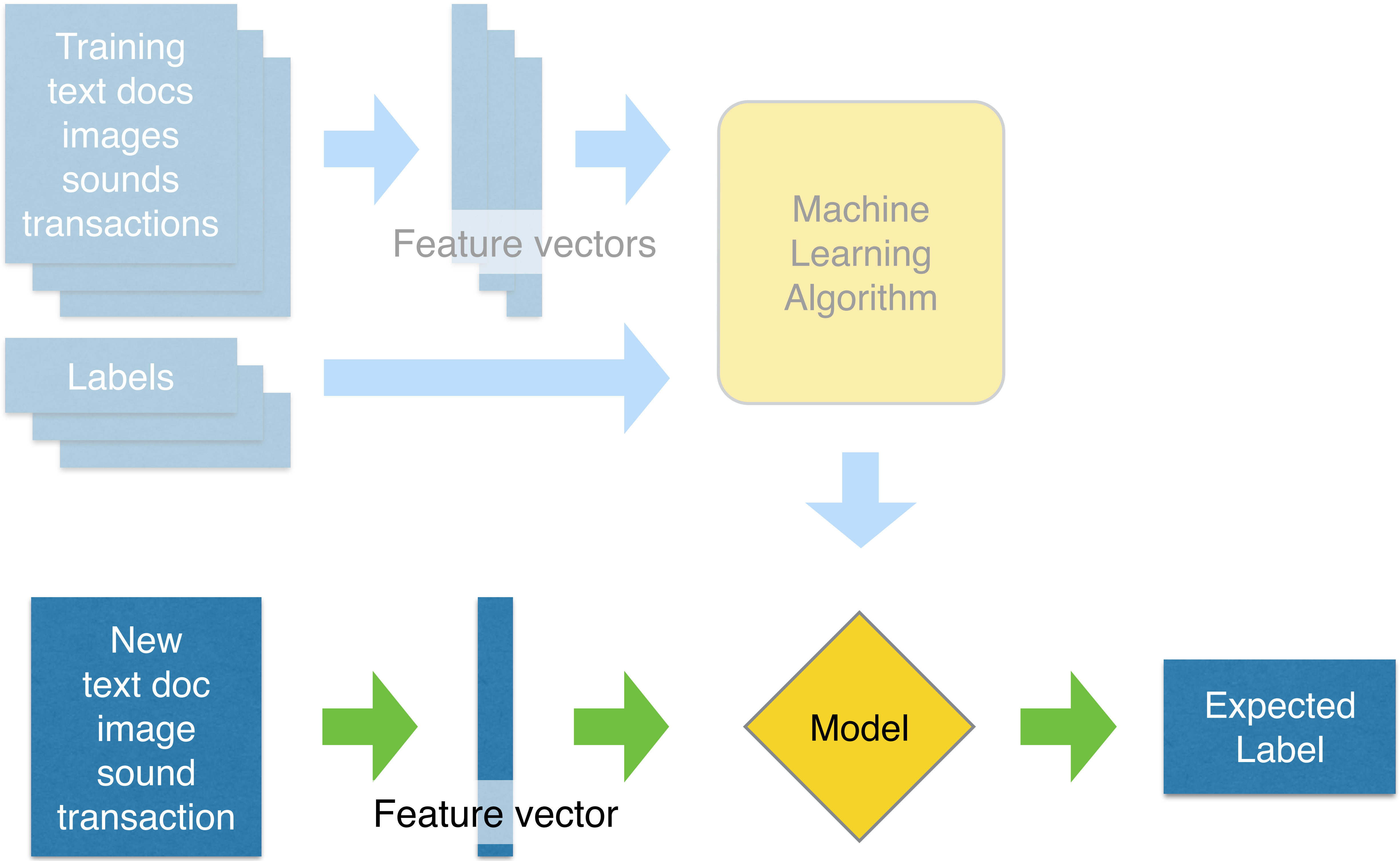
type (category)	# rooms (int)	surface (float m2)	public trans (boolean)	sold (float k€)
Apartment	3	50	TRUE	450
House	5	254	FALSE	430
Duplex	4	68	TRUE	712
Apartment	2	32	TRUE	234

samples (test)

Apartment	2	33	TRUE	?
House	4	210	TRUE	?



Predictive Modeling Data Flow



Predictive Modeling Data Flow

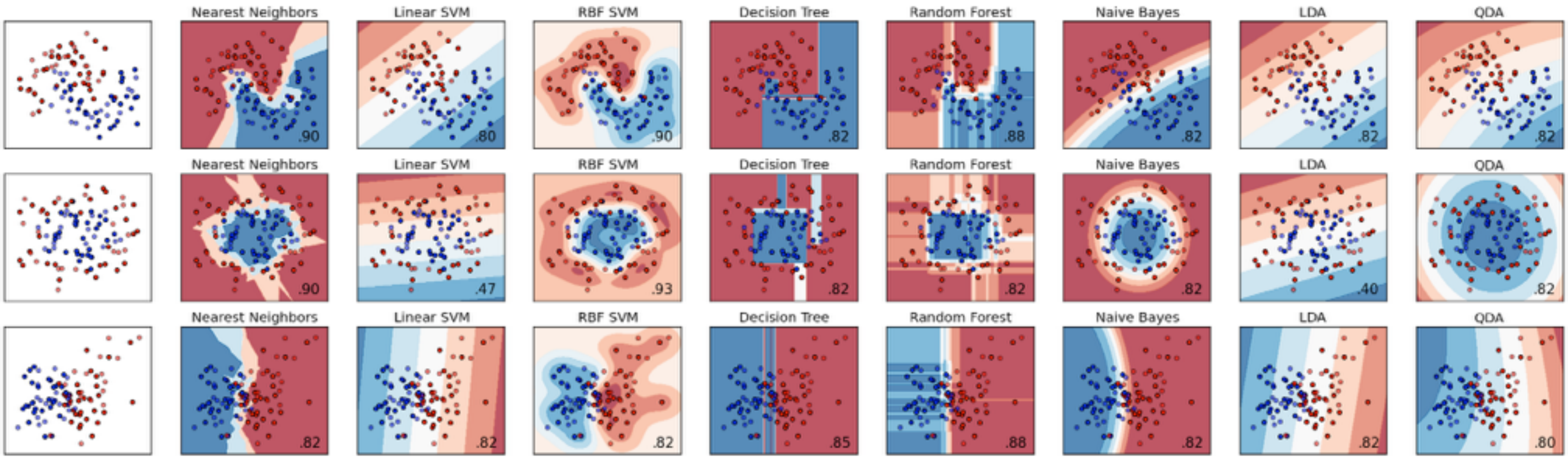
ML in Business

- Predict sales, customer churn, traffic, prices, CTR
- Detect network anomalies, fraud and spams
- Recommend products, movies, music
- Speech recognition for interaction with mobile devices
- Build computer vision systems for robots in the industry and agriculture… or for marketing analysis using social networks data
- Predictive models for text mining and Machine Translation

ML in Science

- Decode the activity of the brain recorded via fMRI / EEG / MEG
- Decode gene expression data to model regulatory networks
- Predict the distance to each star in the sky
- Identify the Higgs boson in proton-proton collisions

Many ML methods

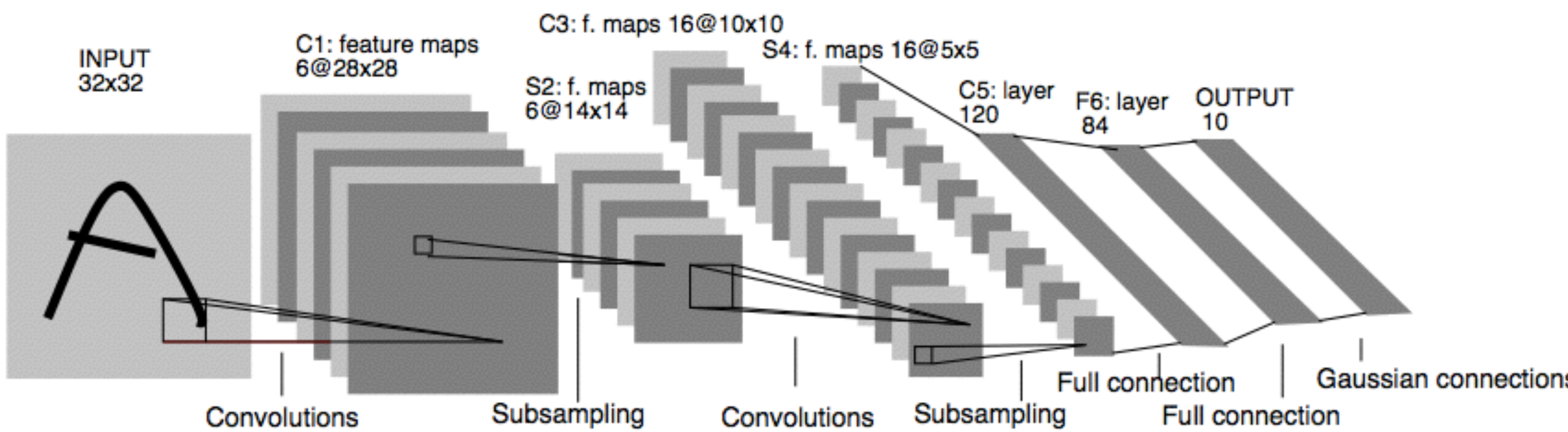


- different assumptions on data
- different scalability profiles at training time
- different latencies at prediction time
- different model sizes (embedability in mobile devices)

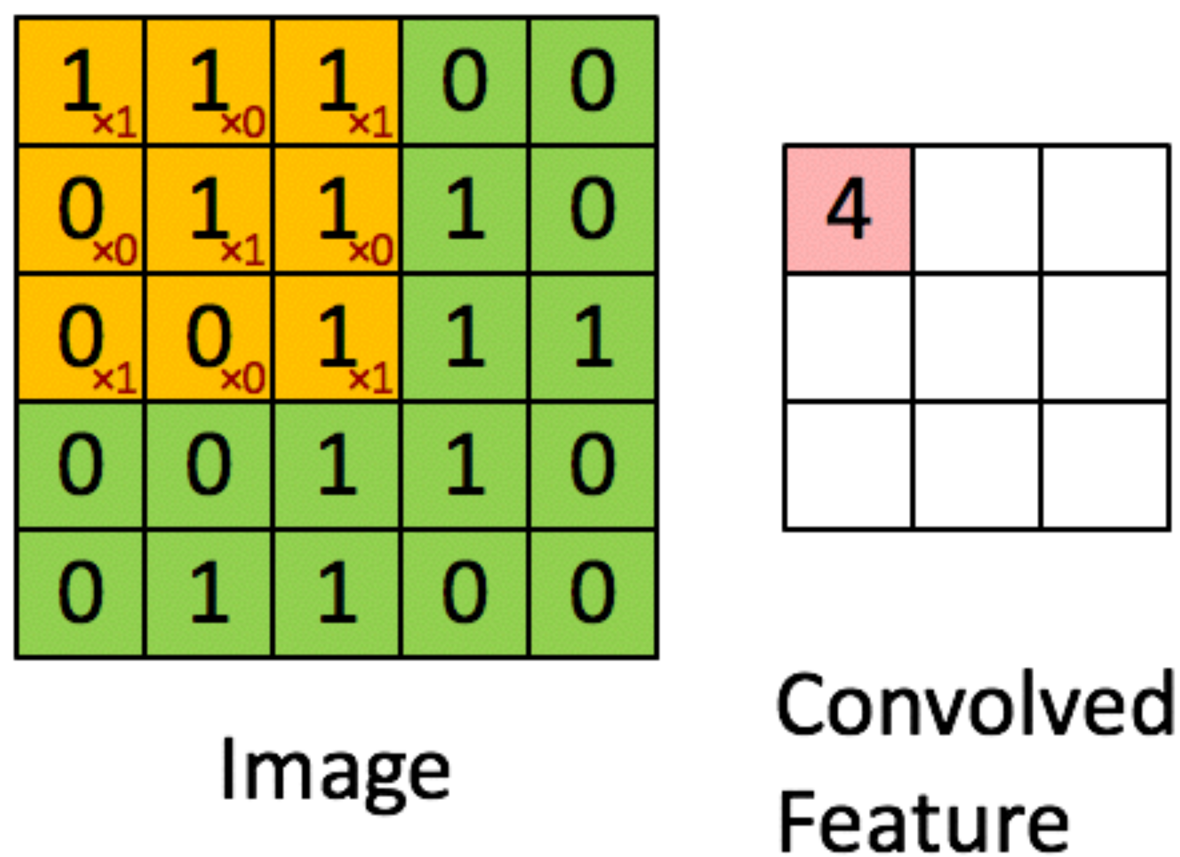
Deep Learning for Computer Vision

Deep Learning in the 90's

- Yann LeCun invented Convolutional Networks
- First NN successfully trained with many layers

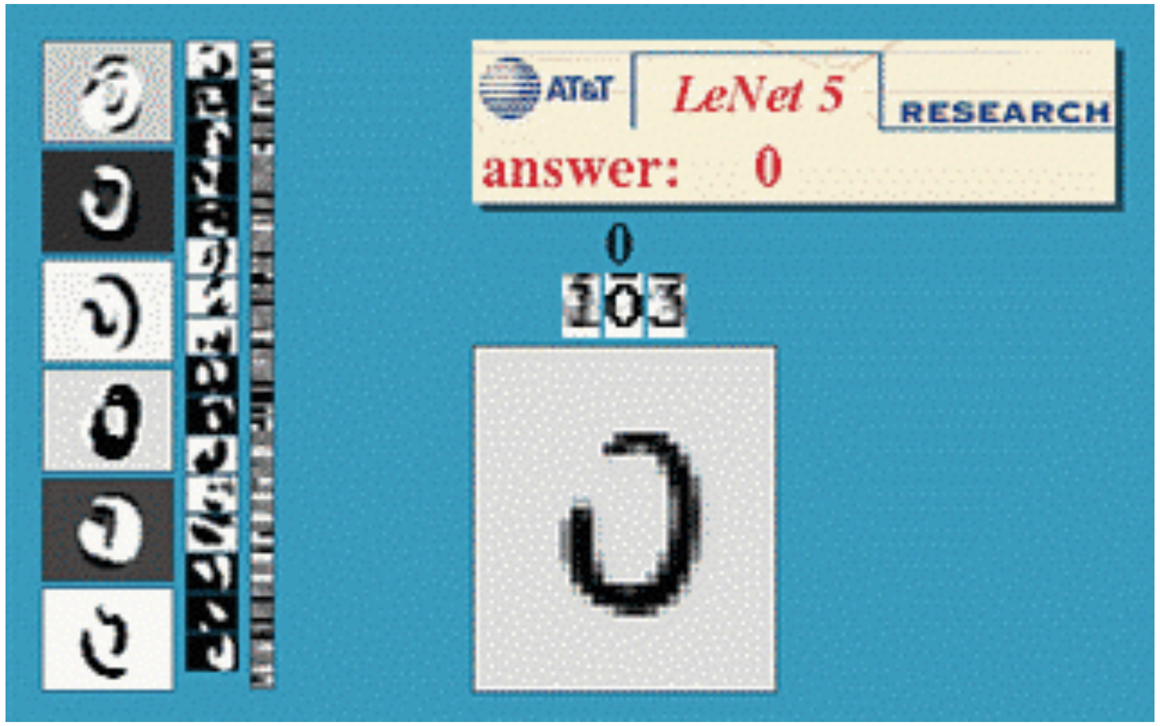


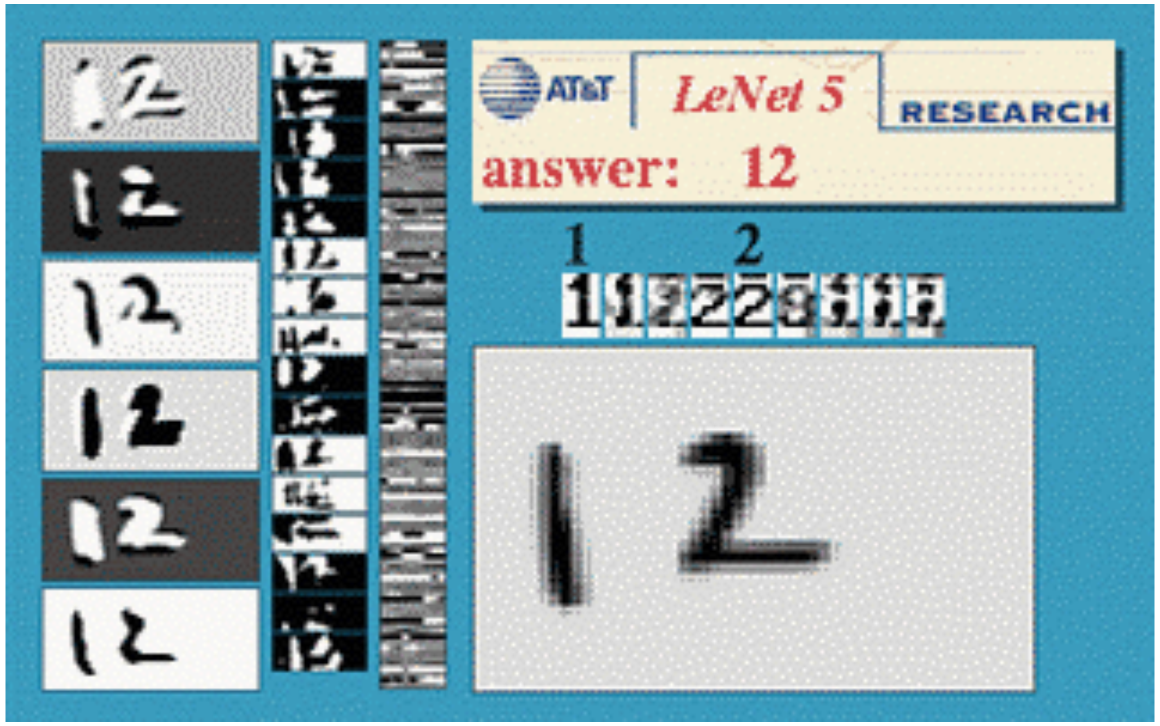
Convolution on 2D input



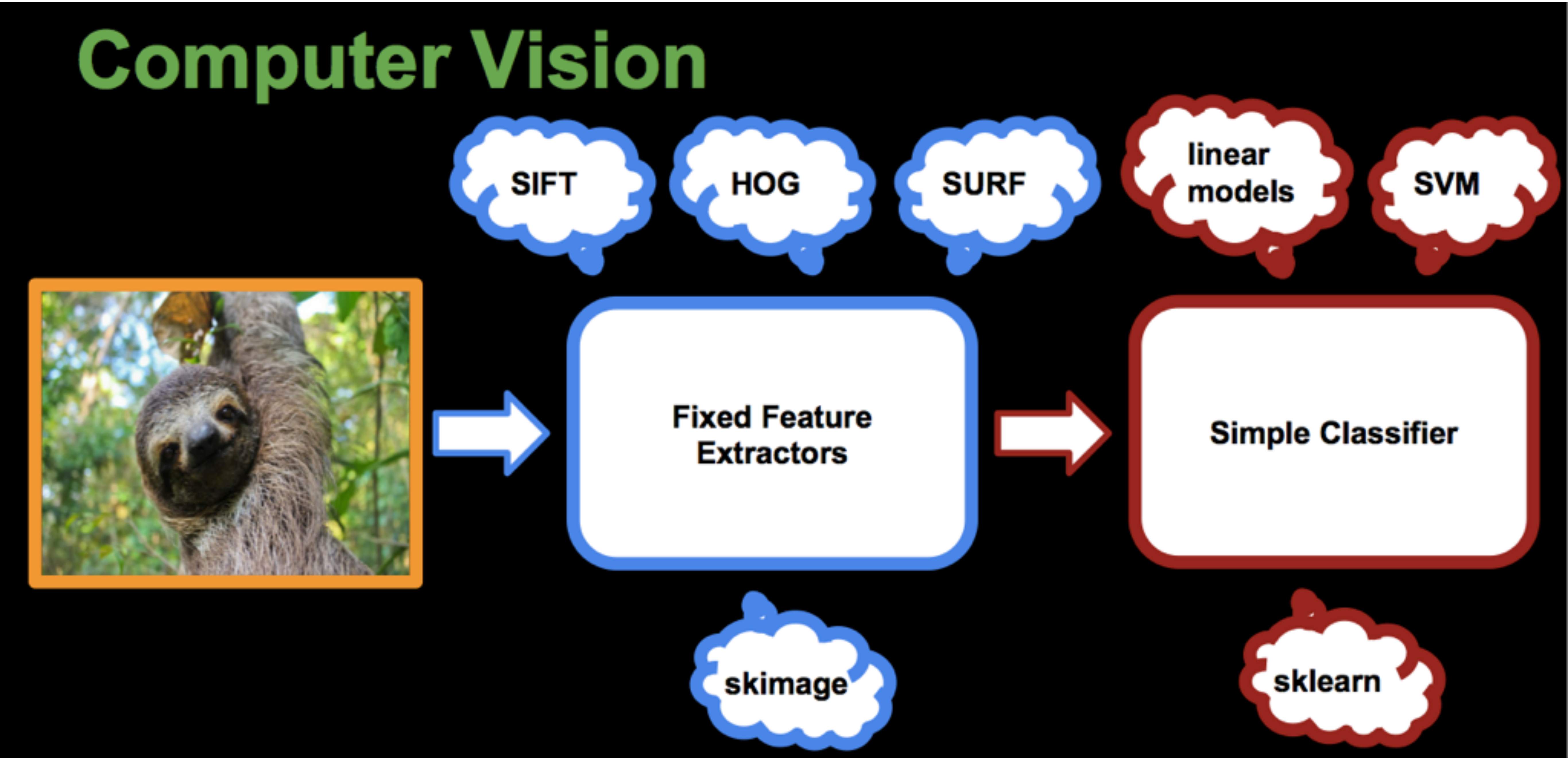
source: Stanford Deep Learning Tutorial

Early success at OCR





Natural image classification until 2012

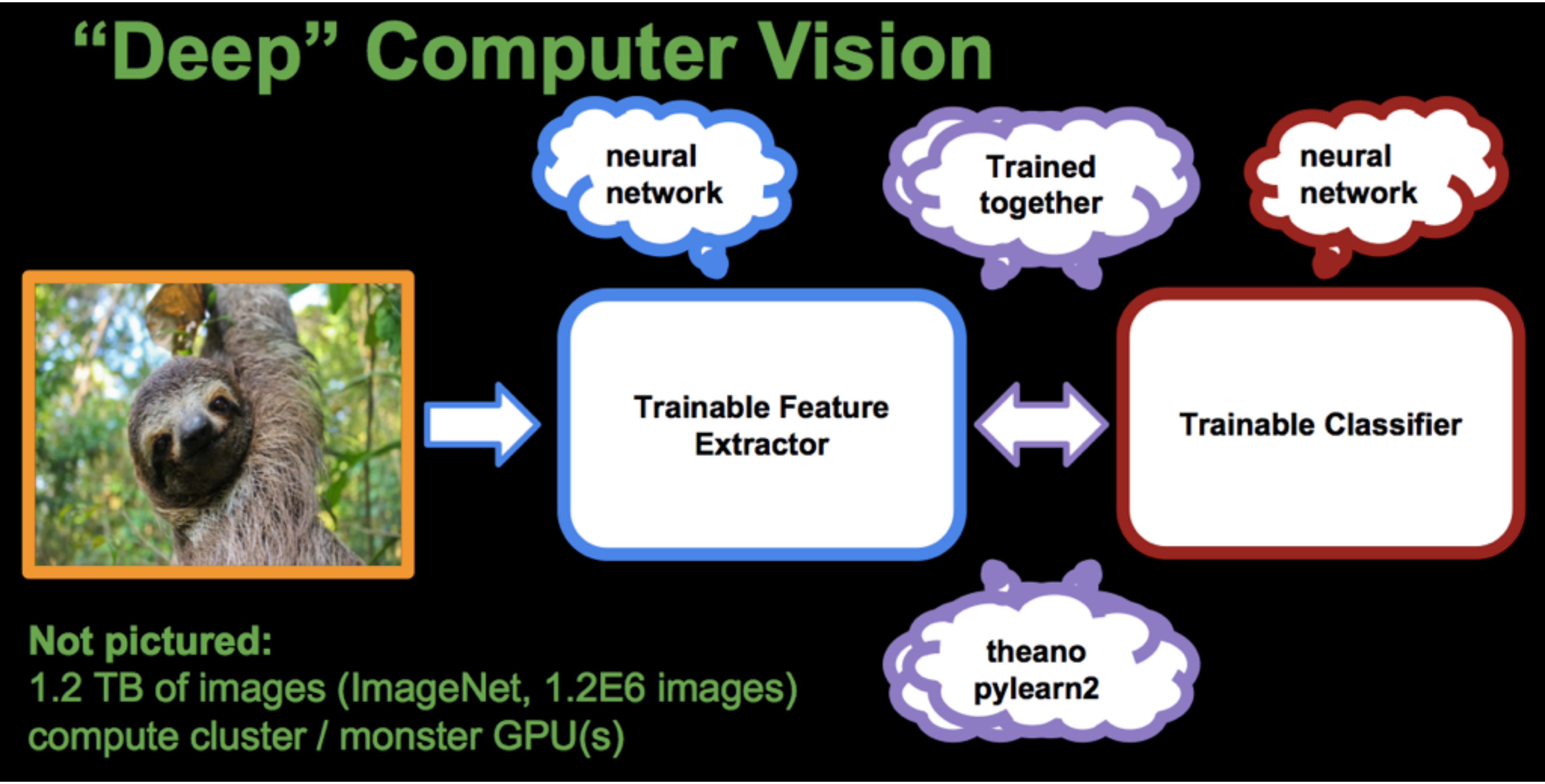


credits: Kyle Kastner

ImageNet Challenge 2012

- 1.2M images labeled with 1000 object categories
- AlexNet from the deep learning team of U. of Toronto wins with **15% error rate** vs 26% for the second (traditional CV pipeline)
- Best NN was trained on GPUs for weeks

Image classification today



credits: Kyle Kastner

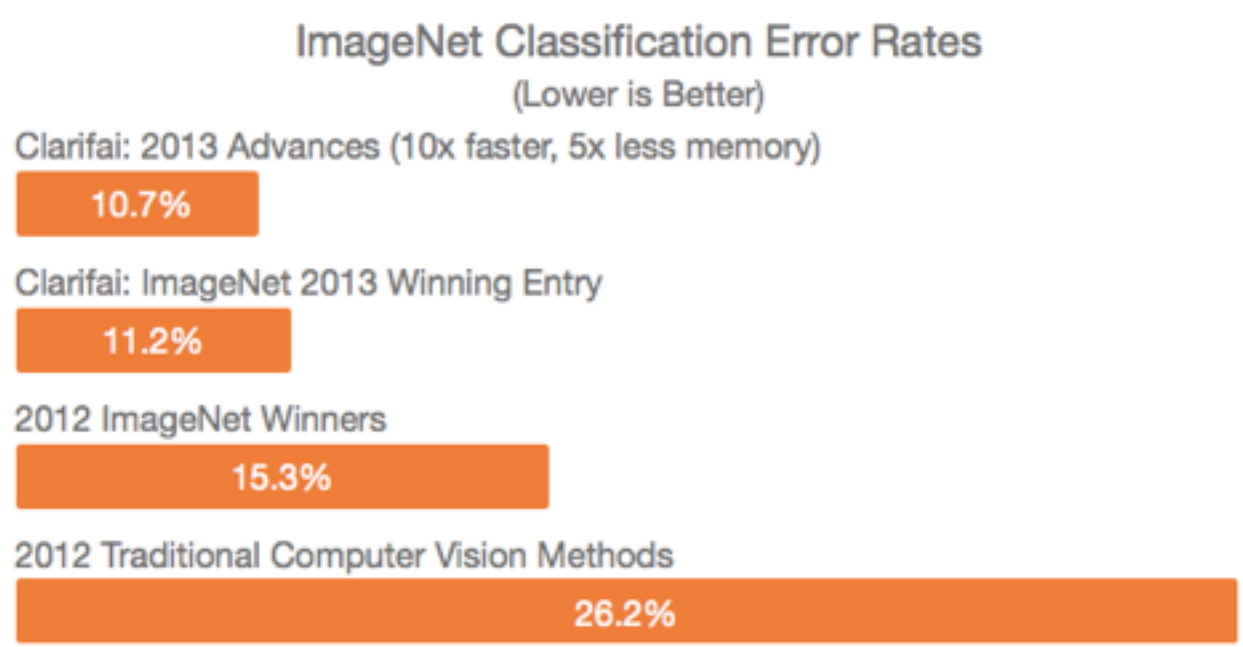
Google Scoops Up Neural Networks Startup DNNresearch To Boost Its Voice And Image Search Tech

Posted Mar 12, 2013 by Rip Empson (@ripemp)

Well, Google's M&A strategy is nothing if not diverse in focus. In November, it acquired package delivery startup Bufferbox. Last month, Google it made its first acquisition of the year, buying eCommerce startup Channel Intelligence. Today, Google dug into the Computer Science department at The University of Toronto to acquire DNNresearch, a young startup founded by professor Geoffrey Hinton and two of his grad students, Alex Krizhevsky and Ilya Sutskever.

ImageNet Challenge 2013

- Clarifai ConvNet model wins at 11% error rate



- Many other participants used ConvNets
- OverFeat by Pierre Sermanet from NYU: shipped binary program to execute pre-trained models

News TCTV Events

Search

Yann LeCun | artificial intelligence | Facebook

NYU "Deep Learning" Professor LeCun Will Head Facebook's New Artificial Intelligence Lab

Posted Dec 9, 2013 by **Josh Constine** (@joshconstine)

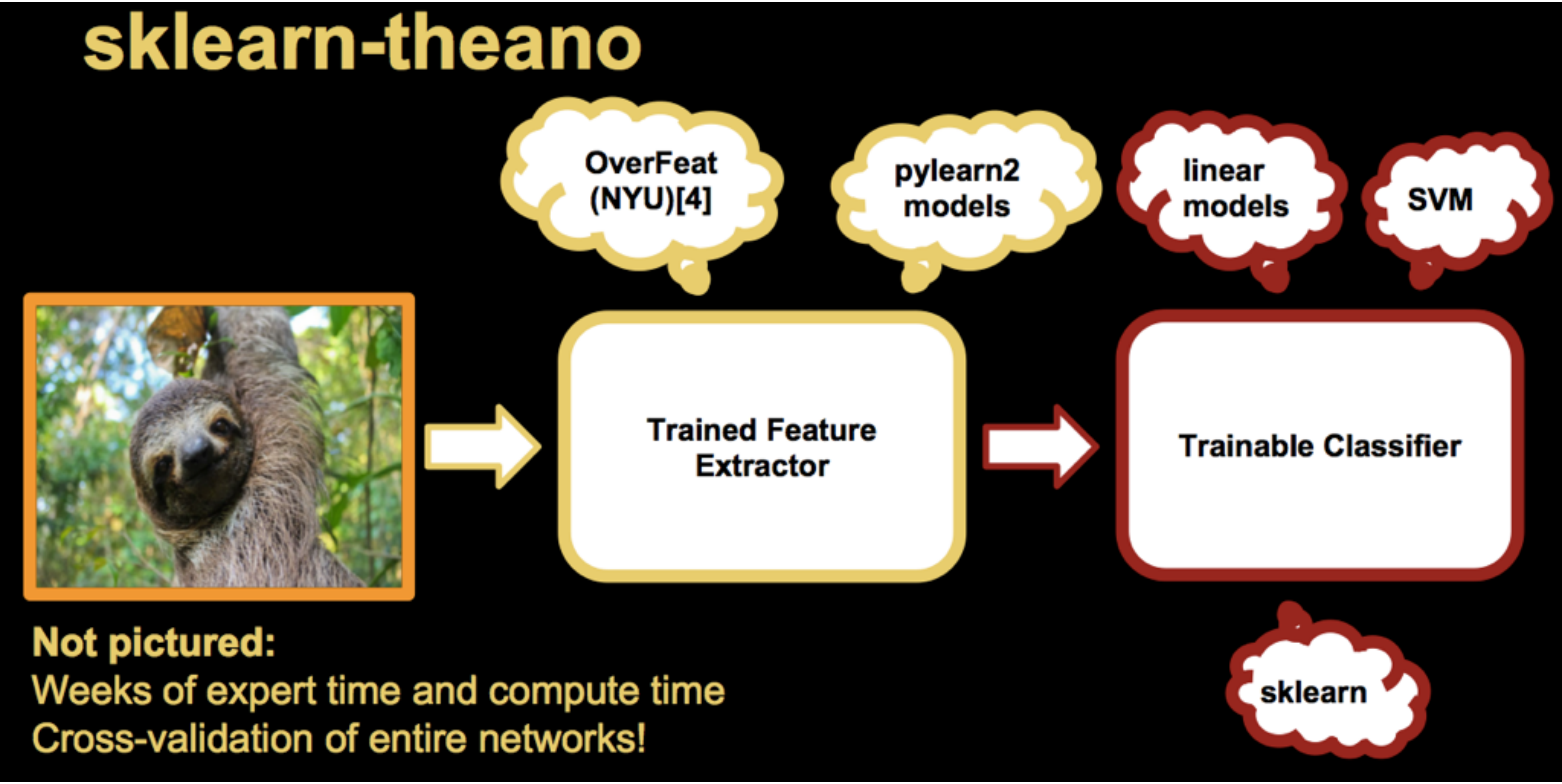
5 Tweet

Next Story

By teaching a computer to think, Facebook hopes to better understand how its users do too. So today the company announced that one of the world's leading deep learning and machine learning scientists, NYU's Professor Yann LeCun, will lead its new artificial intelligence laboratory.

MIT Technology Review first reported that Facebook would launch an Artificial Intelligence lab back in September, but now it has something of a celebrity scientist at its helm. Facebook's AI research will be split across its Menlo Park headquarters, London office, and a new AI lab built just a block from NYU's campus in Manhattan.

Pre-trained models adapted to other CV tasks



credits: Kyle Kastner

Transfer to other CV tasks

- KTH CV team: <u>CNN Features off-the-shelf: an Astounding Baseline for Recognition</u>

"It can be concluded that from now on, deep learning with CNN has to be considered as the primary candidate in essentially any visual recognition task."

Jetpac: analysis of social media photos

- Ratio of **smiles** in faces:
  city **happiness index**
- Ratio of **mustaches** on faces:
  hipster-ness index for coffee-shops
- Ratio of **lipstick** on faces:
  glamour-ness index for night club and bars

HAPPIEST
U.S. CITIES

Rank	City	Smile score
1	SAINT LOUIS, MO	54.7
2	KANSAS CITY, MO	52.0
3	COLUMBUS, OH, US	50.3
4	INDIANAPOLIS CITY, IN	49.6
5	PITTSBURGH, PA	47.9
6	SAN ANTONIO, TX	47.1
7	MINNEAPOLIS, MN	46.4
8	JACKSONVILLE, FL	46.1
9	DETROIT, MI	45.9
10	RALEIGH, NC	45.8
11	NASHVILLE, TN	43.8
12	CHICAGO, IL	43.5
13	CHARLOTTE, NC	43.1
14	BALTIMORE, MD	42.8
15	TAMPA, FL	42.8
16	DENVER, CO	42.5
17	DALLAS, TX	42.1
18	PHOENIX, AZ	41.9
19	BOSTON, MA	41.7
20	HOUSTON, TX	40.9

21	ATLANTA, GA	40.1
22	PHILADELPHIA, PA	39.2
23	SALT LAKE CITY, UT	39.0
24	WASHINGTON, DC	38.8
25	FORT LAUDERDALE, FL	38.5
26	SEATTLE, WA	37.9
27	SCOTTSDALE, AZ	37.2
28	ORLANDO, FL	37.0
29	SACRAMENTO, CA	36.1
30	NEW ORLEANS, LA	35.5
31	AUSTIN, TX	35.4
32	NEW YORK CITY-QUEENS, NY	34.4
33	SAN DIEGO, CA	34.3
34	PORTLAND, OR	33.4
35	NEW YORK CITY-MANHATTAN, NY	33.2
36	HUNTINGTON BEACH, CA	32.9
37	HONOLULU, HI	32.5
38	SAN FRANCISCO, CA	32.1
39	SAN JOSE, CA	31.8
40	MIAMI, FL	31.6
41	LAS VEGAS, NV	31.1
42	LOS ANGELES, CA	30.1
43	OAKLAND, CA	29.7
44	CAMBRIDGE, MA	29.4
45	NEW YORK CITY-BROOKLYN, NY	29.3
46	COSTA MESA, CA	28.6
47	MIAMI BEACH, FL	28.5
48	BAY LAKE, FL	28.4
49	PARADISE, NV	25.6
50	ANAHEIM, CA	24.6

Google Buys Jetpac To Give Context To Visual Searches

Posted Aug 15, 2014 by Sarah Buhr (@sarahbuhr)

ImageNet Challenge 2014

- In the mean time Pierre Sermanet had joined other people from Google Brain

- Monster model: GoogLeNet now at **6.7% error rate**

GoogLeNet vs Andrej

- Andrej Karpathy evaluated human performance (himself): **~5% error rate**
- "It is clear that humans will soon only be able to outperform state of the art image classification models by use of significant effort, expertise, and time.”
- “As for my personal take-away from this week-long exercise, I have to say that, qualitatively, I was very impressed with the ConvNet performance. Unless the image exhibits some irregularity or tricky parts, the ConvNet confidently and robustly predicts the correct label.”

source: What I learned from competing against a ConvNet on ImageNet

Word Embeddings

Neural Language Models

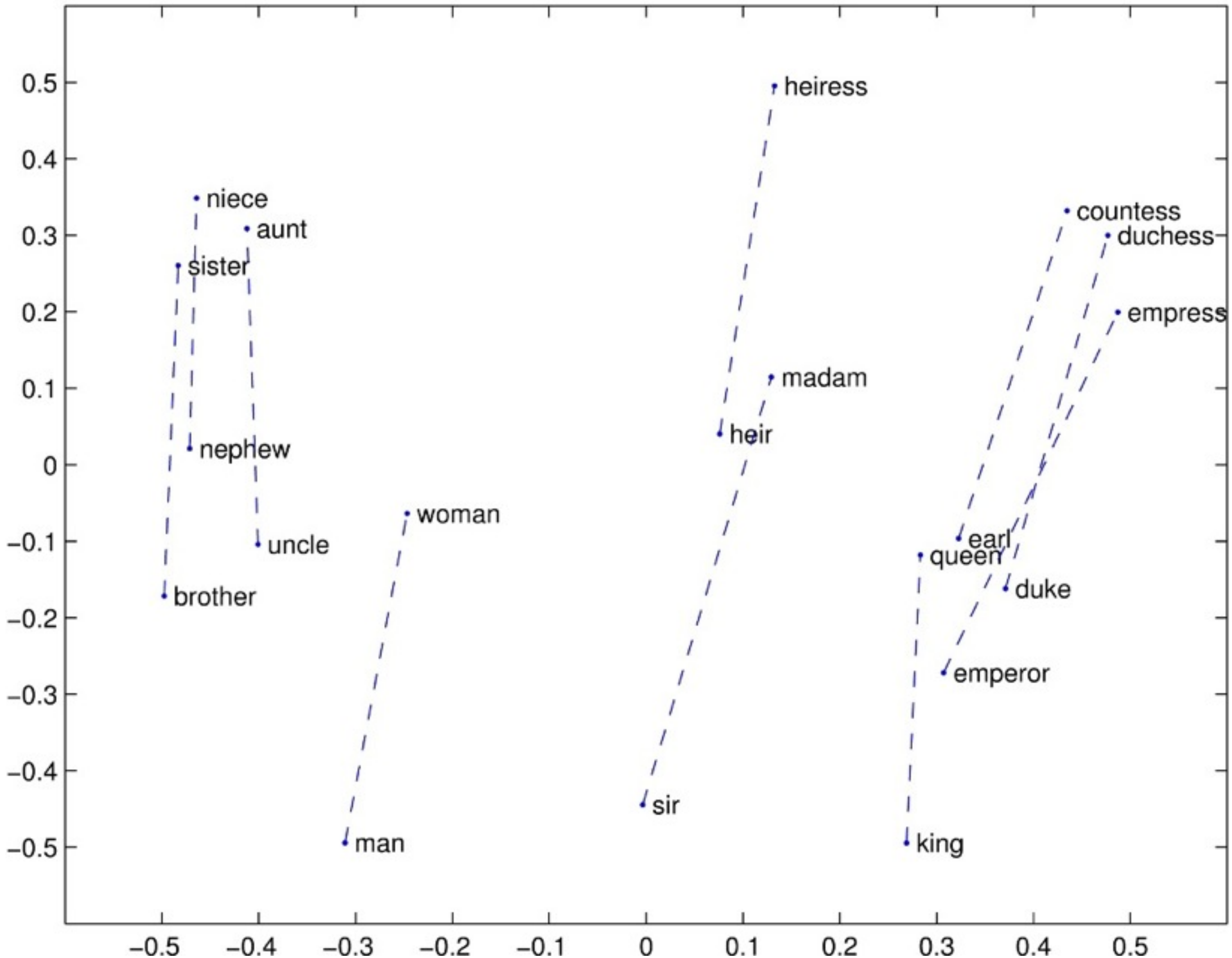
- Each word is represented by a fixed dimensional vector
- Goal is to predict target word given ~5 words context from a random sentence in Wikipedia
- Random substitutions of the target word to generate negative examples
- Use NN-style training to optimize the vector coefficients

Progress in 2013 / 2014

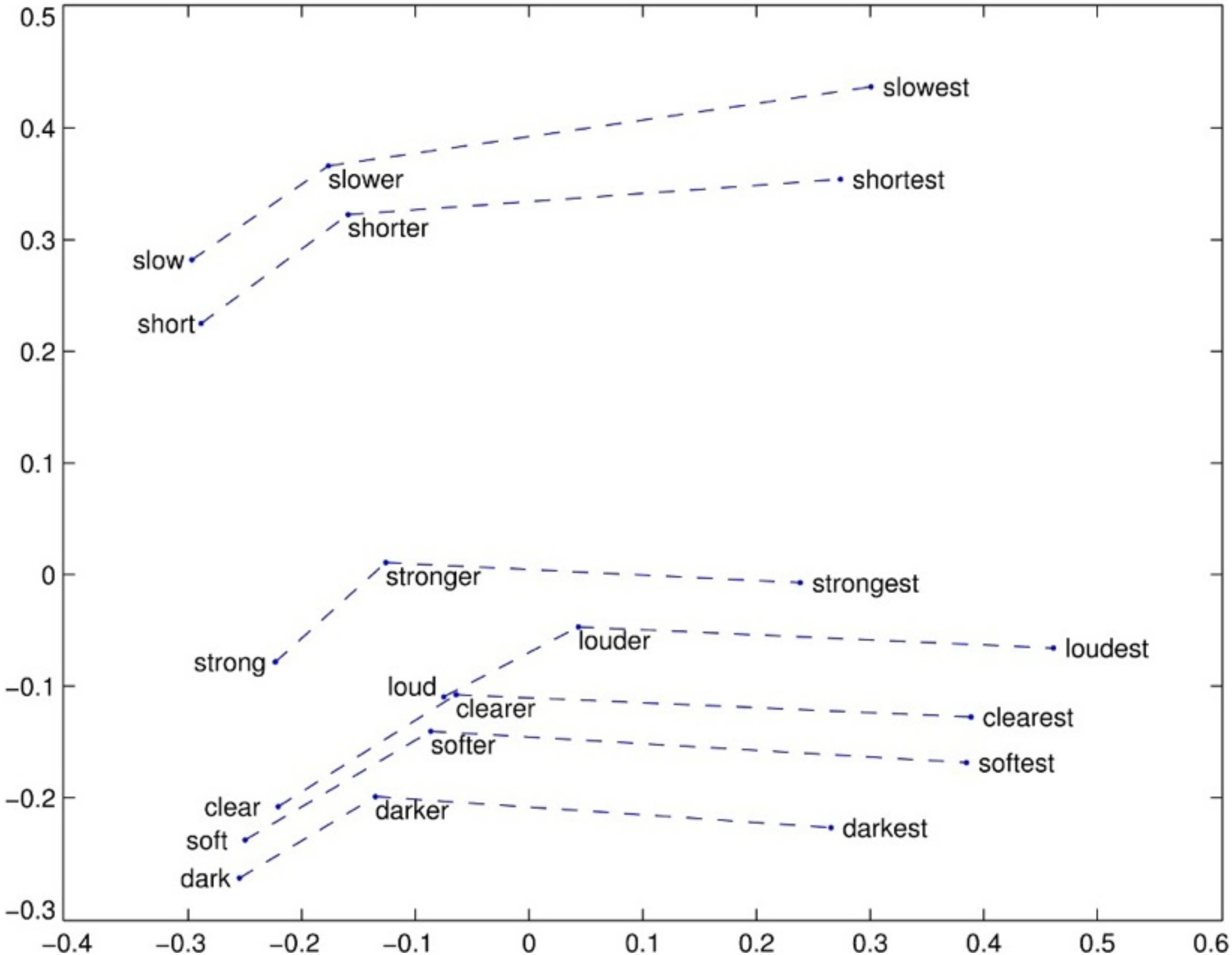
- Simpler linear models (word2vec) benefit from larger training data (1B+ words) and dimensions (300+)
- Some models (GloVe) now closer to matrix factorization than neural networks
- Can successfully uncover semantic and syntactic word relationships, *unsupervised!*

Analogies

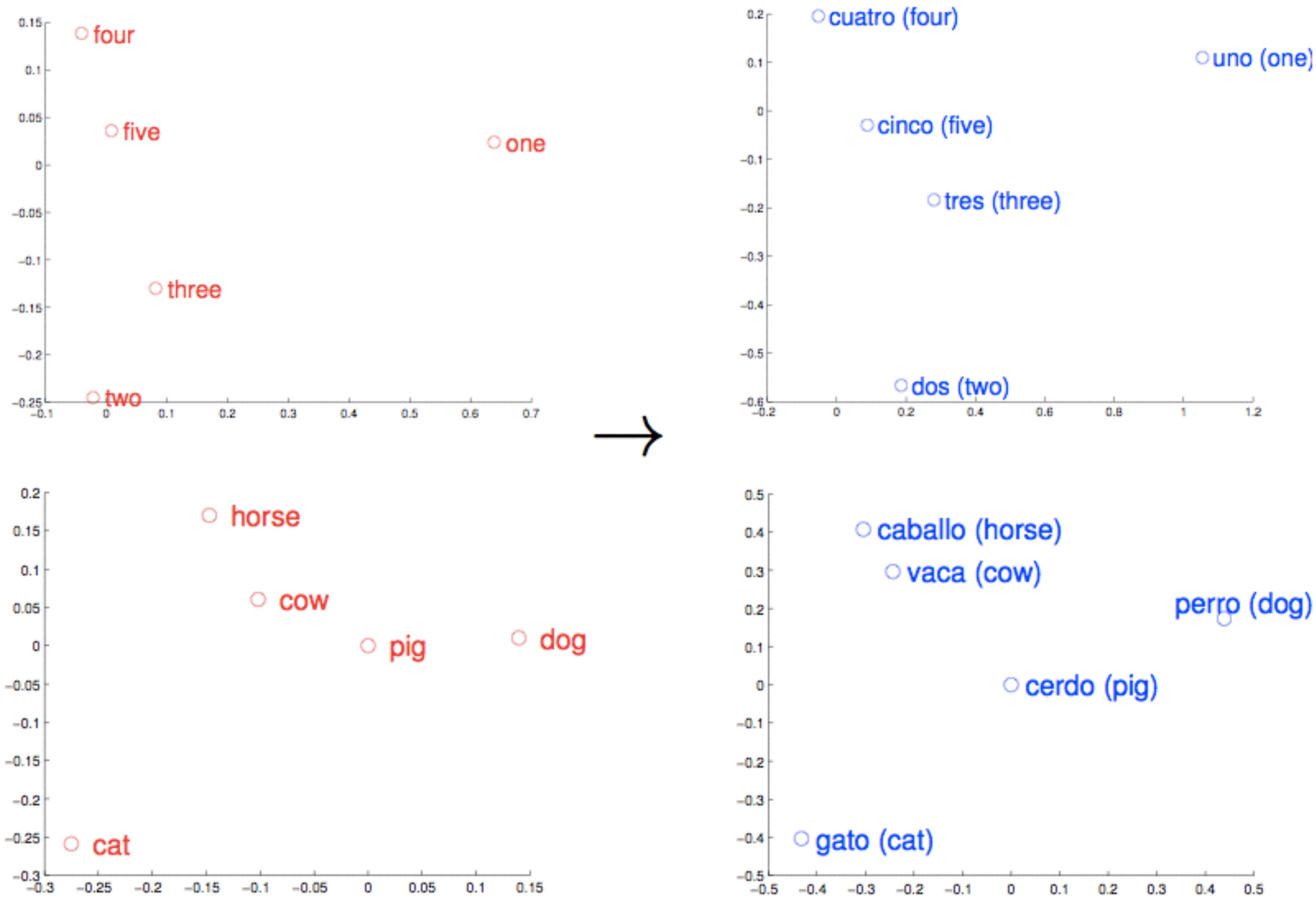
- [king] - [male] + [female] ~= [queen]
- [Berlin] - [Germany] + [France] ~= [Paris]
- [eating] - [eat] + [fly] ~= [flying]



source: http://nlp.stanford.edu/projects/glove/



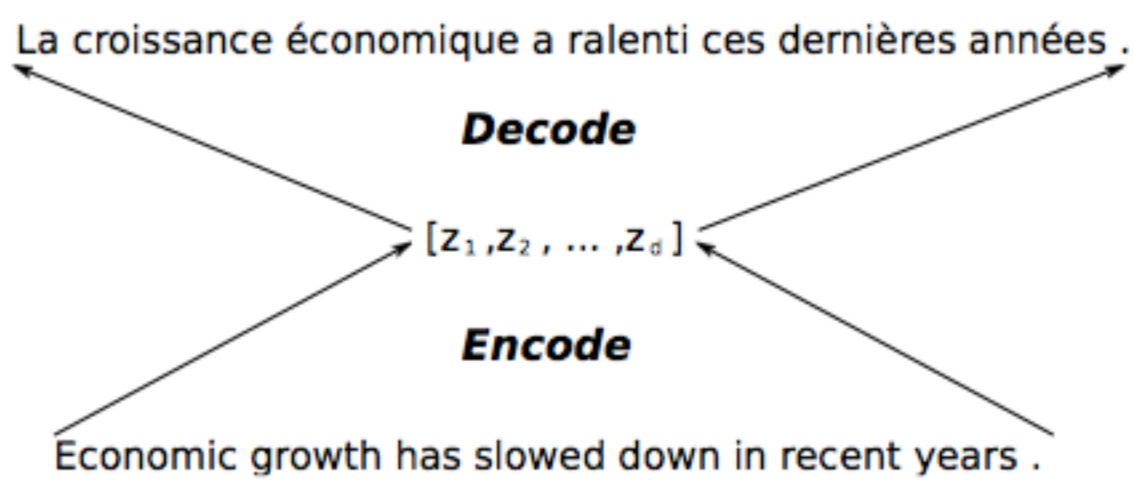
source: http://nlp.stanford.edu/projects/glove/

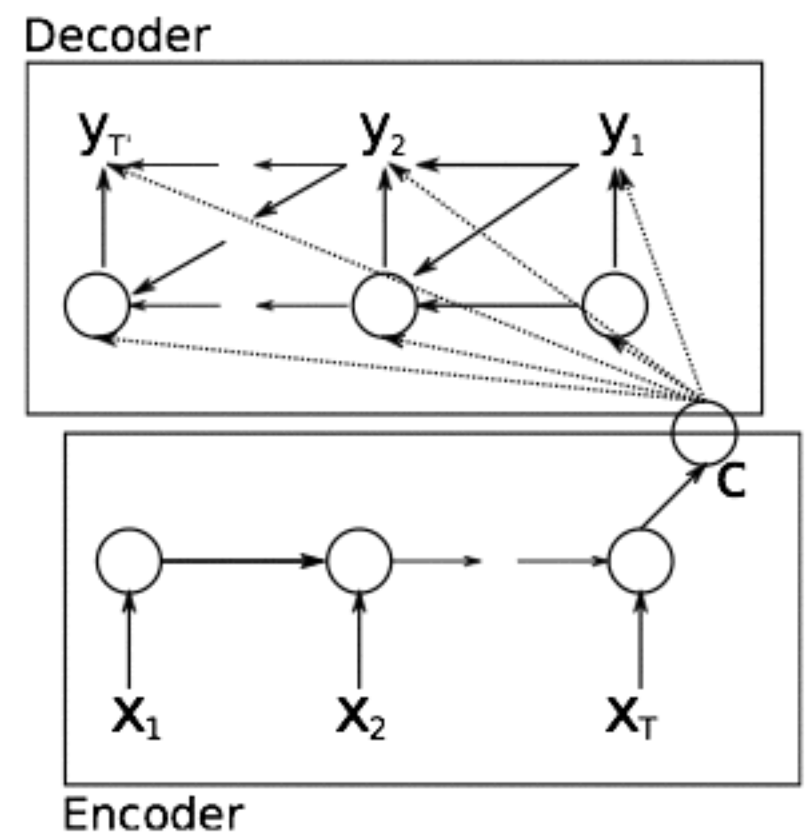


source: Exploiting Similarities among Languages for MT

Neural Machine Translation

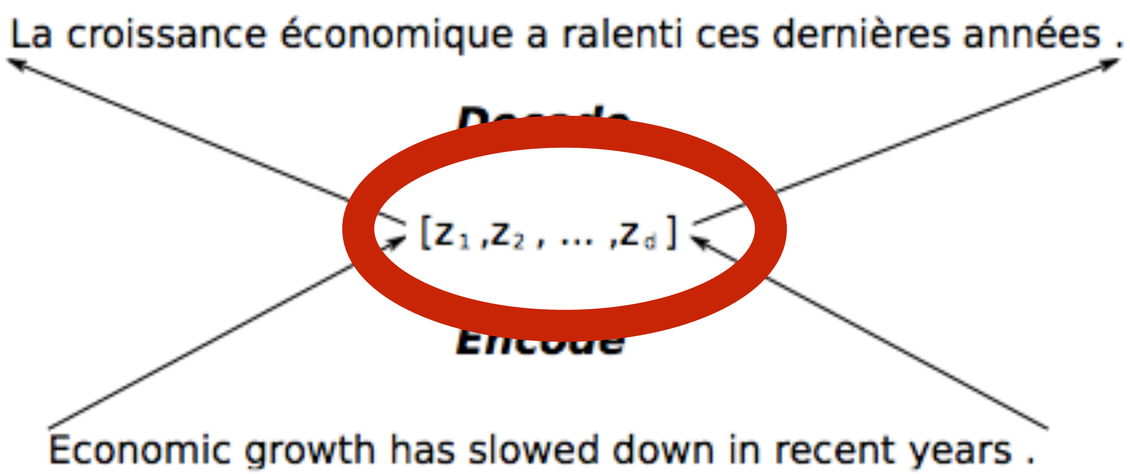
RNN for MT





source: <u>Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation</u>

RNN for MT



Language independent,
vector representation
of the meaning of
any sentence!

Neural MT vs Phrase-based SMT

Model	All	No UNK
U. Montreal	28.45	**36.15**
Google Brain (forward*)	26.17	N/A
Google Brain (backward*)	30.59	N/A
Google Brain (backward*, 5 models)	**34.81**	N/A
Moses (no NN)	33.30	35.63

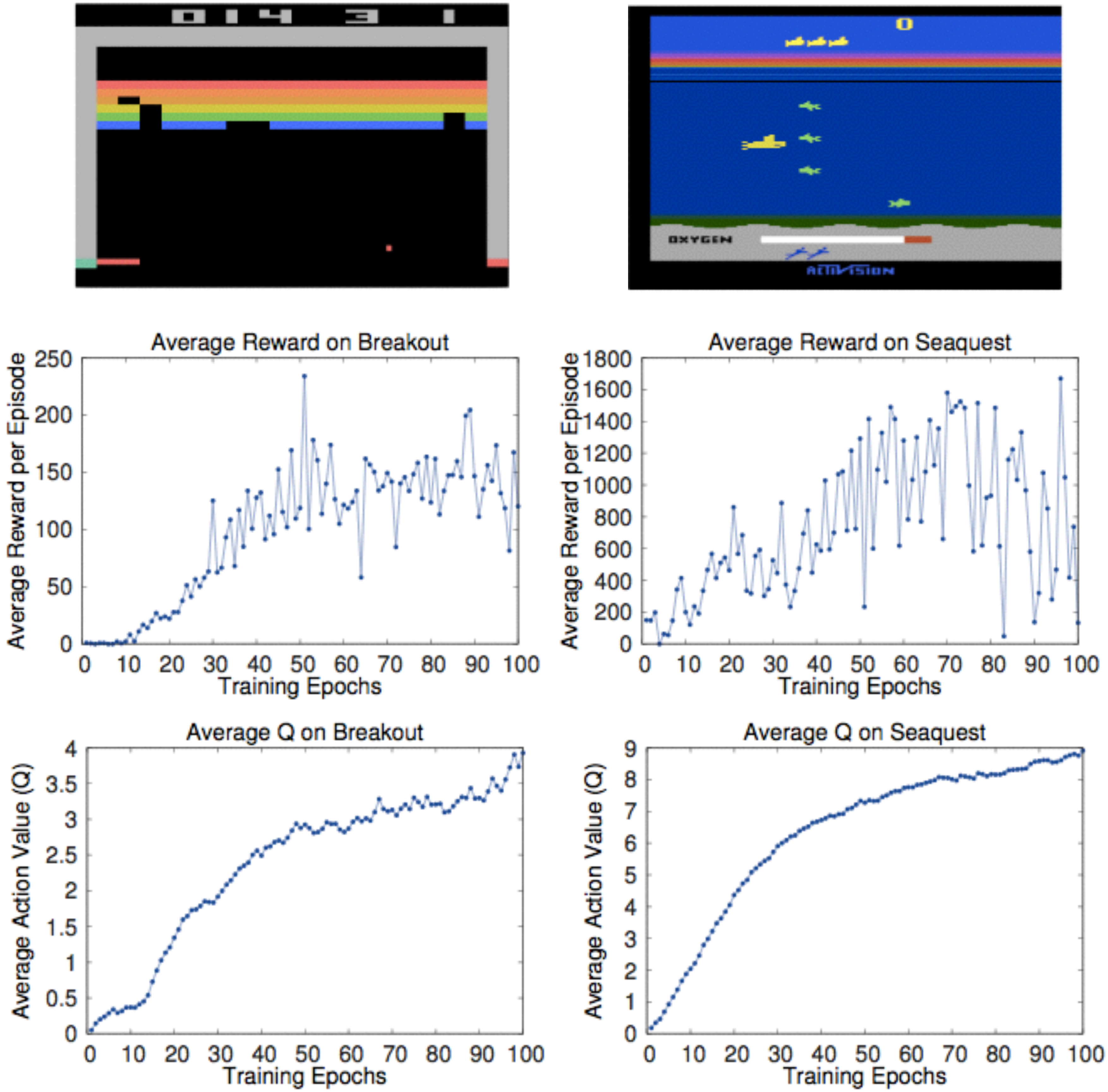
BLEU scores of NMT & Phrase-SMT models on English / French (Oct. 2014)

Deep Learning to Play, Execute and Program

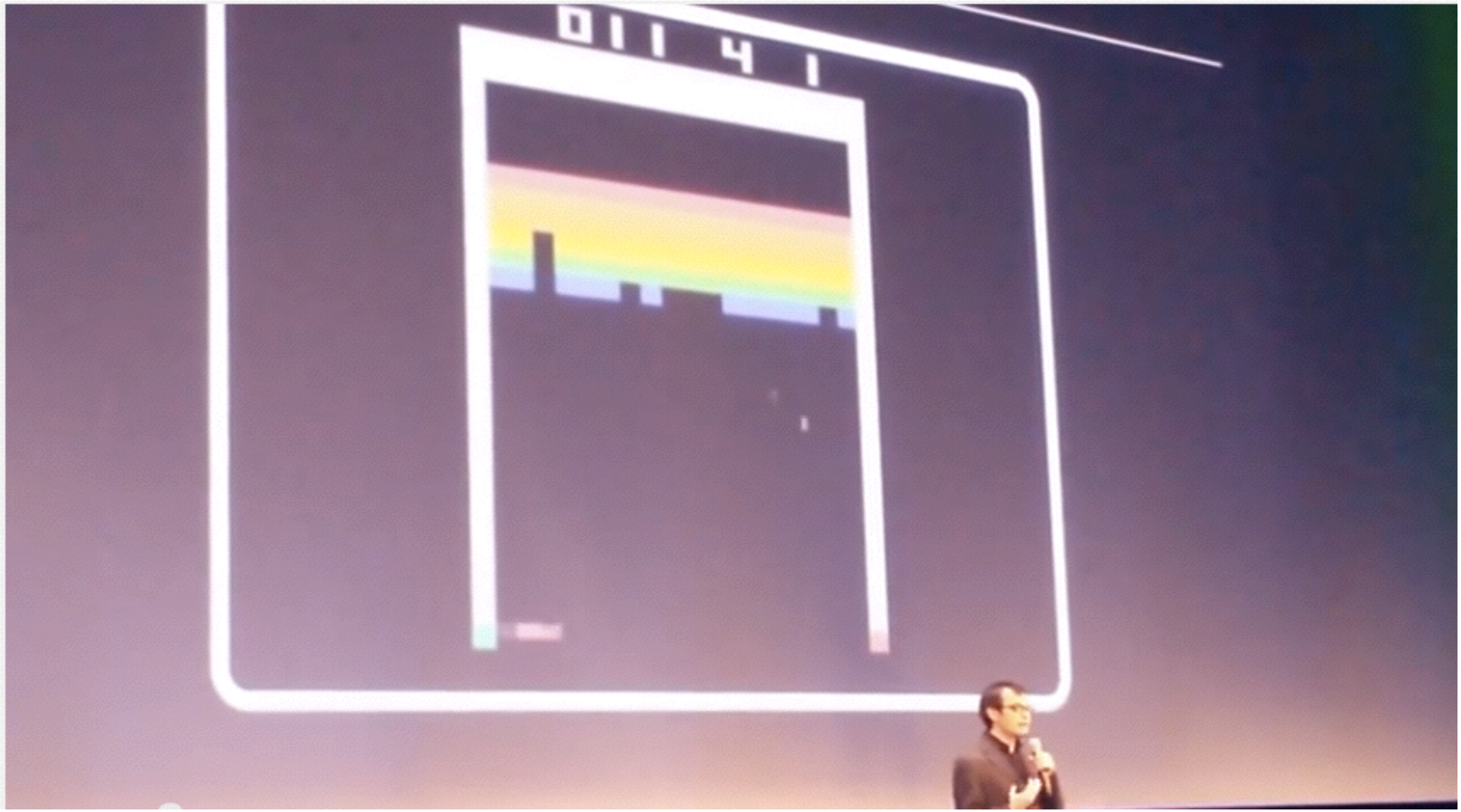
Exploring the frontier of learnability

DeepMind: Learning to Play & win dozens of Atari games

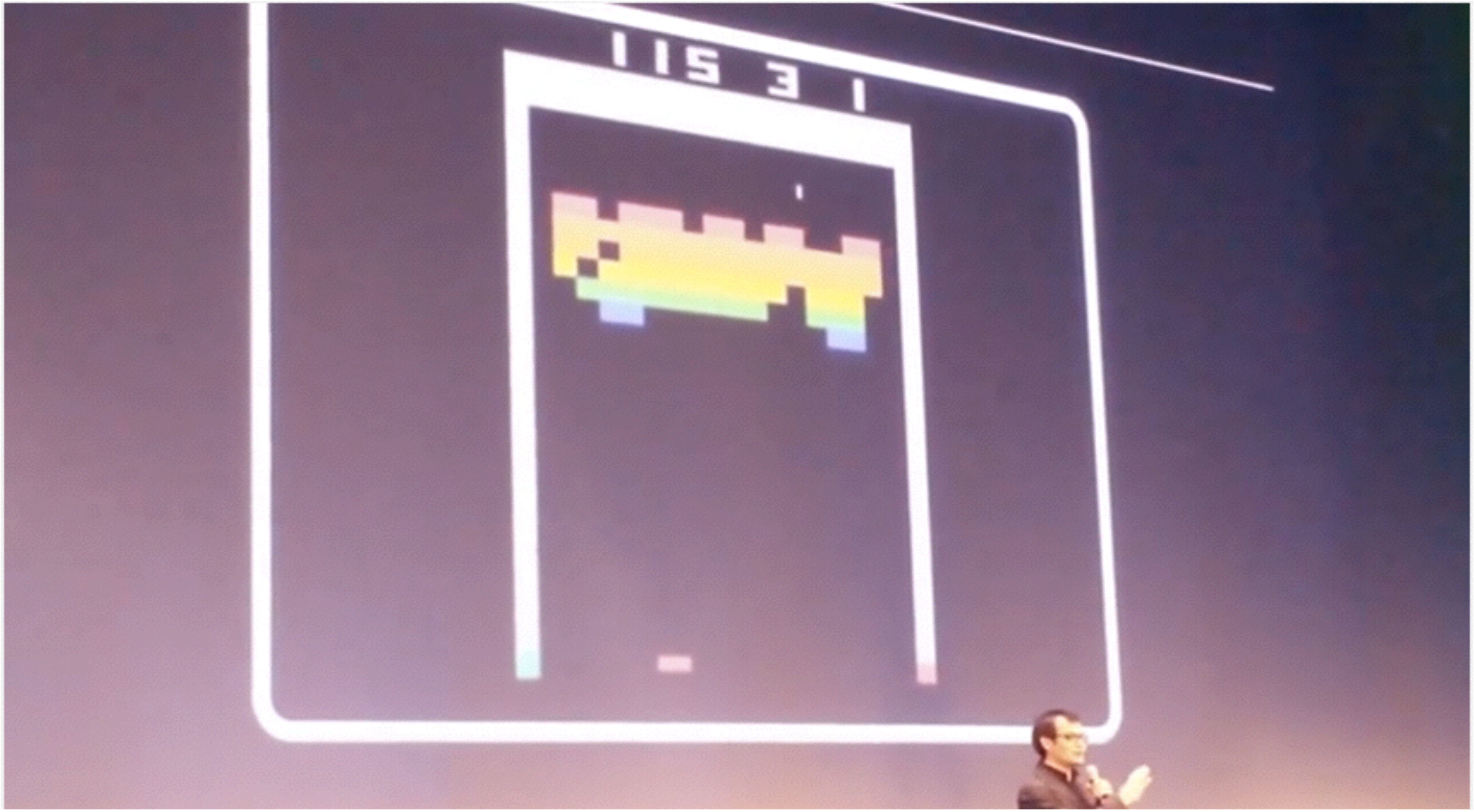
- DeepMind startup demoed at NIPS 2013 a new Deep Reinforcement Learning algorithm
  - Raw pixel input from Atari games (state space)
  - Keyboard keys as action space
  - Scalar signal {“lose”, “survive”, “win”} as reward
  - CNN trained with a Q-Learning variant



source: Playing Atari with Deep Reinforcement Learning



https://www.youtube.com/watch?v=EfGD2qveGdQ



https://www.youtube.com/watch?v=EfGD2qveGdQ

TechCrunch

News TCTV Events

Search

Europe acquisition Google DeepMind Technologies DeepMind

Google Acquires Artificial Intelligence Startup DeepMind For More Than $500M

Posted Jan 26, 2014 by Catherine Shu (@catherineshu)

Tweet

Next Story

Learning to Execute

- Google Brain & NYU, October 2014 (very new)
- RNN trained to map character representations of programs to outputs
- Can learn to emulate a simplistic Python interpreter from examples programs & expected outputs
- Limited to one-pass programs with O(n) complexity

Input:

```
f=(8794 if 8887<9713 else (3*8334))
print((f+574))
```

Target: 9368.

Model prediction: 9368.

Input:

```
c=445
d=(c-4223)
for x in range(1):
  d+=5272
print((8942 if d<3749 else 2951))
```

Target: 8942.

Model prediction: 8942.

source: Learning to Execute

Input:

```
a=1027
for x in range(2):
  a+=(402 if 6358>8211 else 2158)
print(a)
```

Target: 5343.

Model prediction: 5293.

Input:

```
j=8584
for x in range(8):
  j+=920
b=(1500+j)
print((b+7567))
```

Target: 25011.

Model prediction: 23011.

source: Learning to Execute

What the model actually *sees*

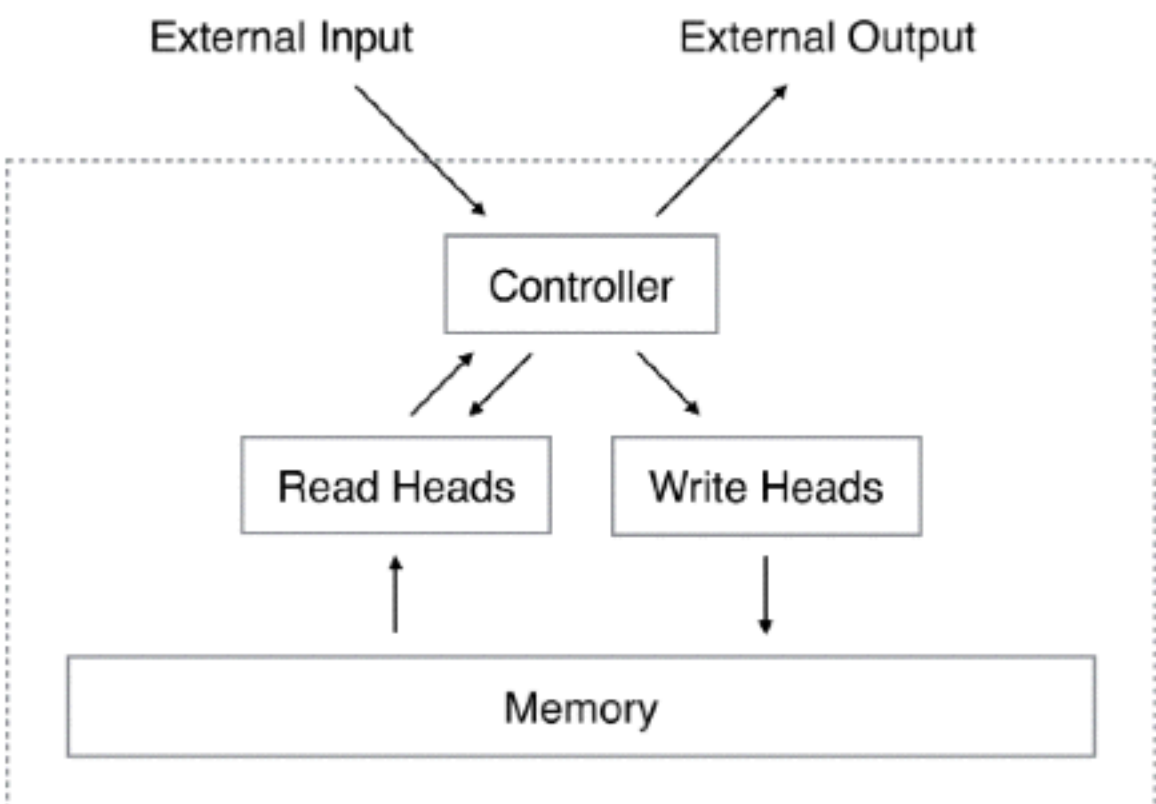
Input:
vqppkn
sqdvfljmnc
y2vxdddsepnimcbvubkomhrpliibtwztbljipcc
Target: hkhpg

source: Learning to Execute

Neural Turing Machines

- Google DeepMind, October 2014 (very new)
- Neural Network coupled to external memory (tape)
- Analogue to a Turing Machine but differentiable
- Can be used to learn to simple programs from example input / output pairs
  - copy, repeat copy, associative recall,
  - binary n-grams counts and sort

Architecture



- Turing Machine:

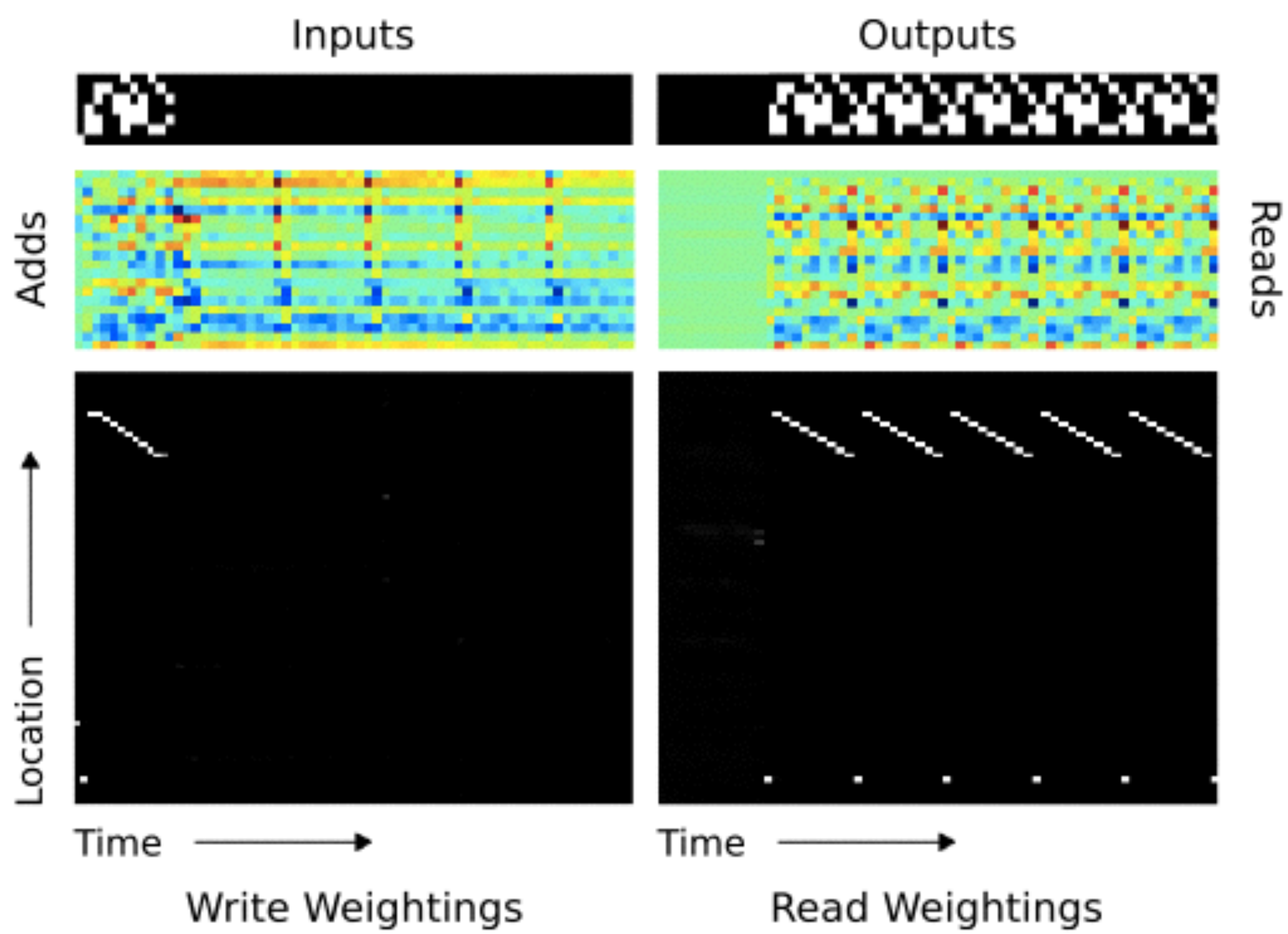
  controller == FSM

- Neural Turing Machine

  controller == RNN w/ LSTM

source: Neural Turing Machines

Example run:
copy & repeat task



source: Neural Turing Machines

Concluding remarks

- Deep Learning now state of the art at:
  - Several computer vision tasks
  - Speech recognition (partially NN-based in 2012, fully in 2013)
  - Machine Translation (English / French)
  - Playing Atari games from the 80's
- Recurrent Neural Network w/ LSTM units seems to be applicable to problems initially thought out of the scope of Machine Learning
- Stay tuned for 2015!

Thank you!

http://speakerdeck.com/ogrisel

http://twitter.com/ogrisel

References

- ConvNets in the 90's by Yann LeCun: LeNet-5

  http://yann.lecun.com/exdb/lenet/

- ImageNet Challenge 2012 winner: AlexNet (Toronto)

  http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks

- ImageNet Challenge 2013: OverFeat (NYU)

  http://cilvr.nyu.edu/doku.php?id=software:overfeat:start

- ImageNet Challenge 2014 winner: GoogLeNet (Google Brain)

http://googleresearch.blogspot.fr/2014/09/building-deeper-understanding-of-images.html

References

- Word embeddings

  First gen: http://metaoptimize.com/projects/wordreprs/

  Word2Vec: https://code.google.com/p/word2vec/

  GloVe: http://nlp.stanford.edu/projects/glove/

- Neural Machine Translation

  Google Brain: http://arxiv.org/abs/1409.3215

  U. of Montreal: http://arxiv.org/abs/1406.1078

  https://github.com/lisa-groundhog/GroundHog

References

- Deep Reinforcement Learning:

  http://www.cs.toronto.edu/~vmnih/docs/dqn.pdf

- Neural Turing Machines:

  http://arxiv.org/abs/1410.5401

- Learning to Execute:

  http://arxiv.org/abs/1410.4615

Thanks to @kastnerkyle
for slides / biblio coaching :)