

CSE 473: Introduction to Artificial Intelligence

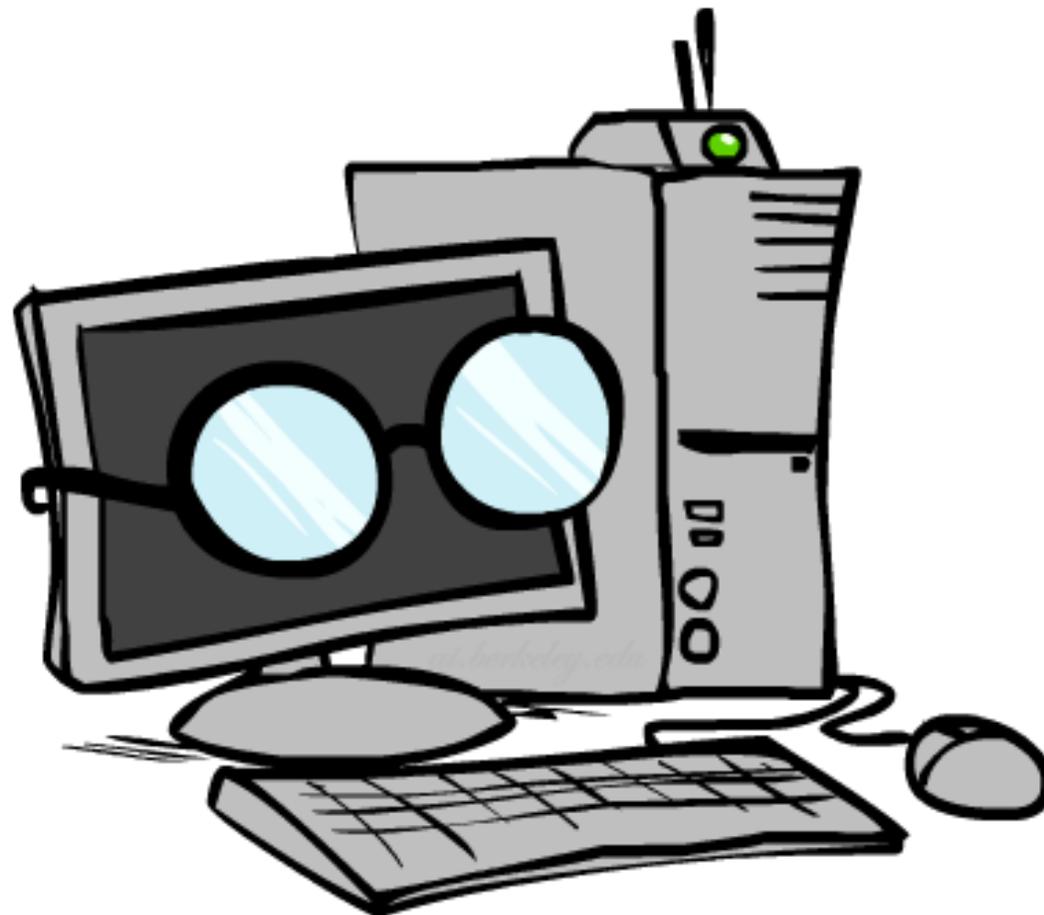
Hanna Hajishirzi
Applications

slides adapted from
Dan Klein, Pieter Abbeel ai.berkeley.edu
And Dan Weld, Luke Zettlemoyer

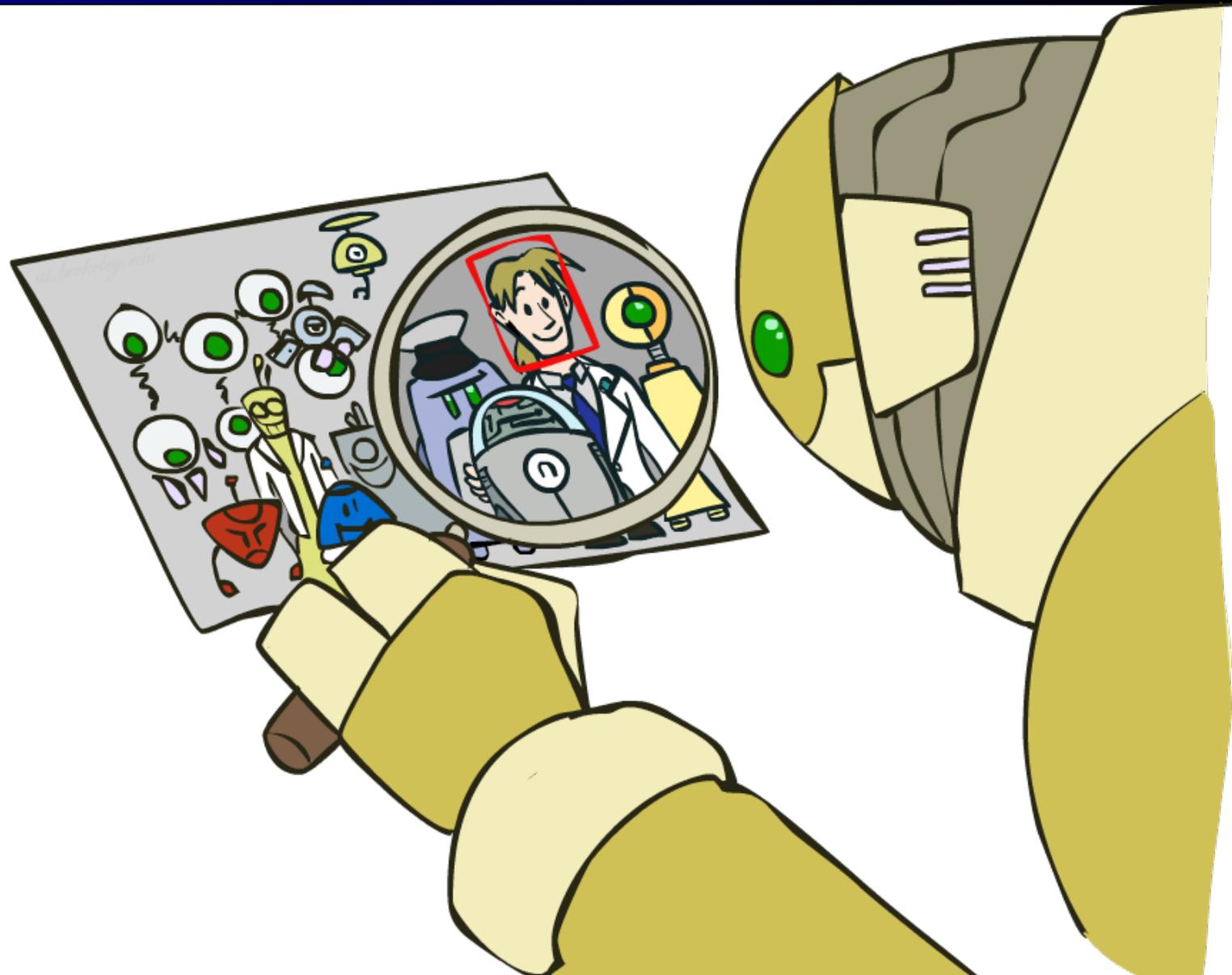


How well does ML work?

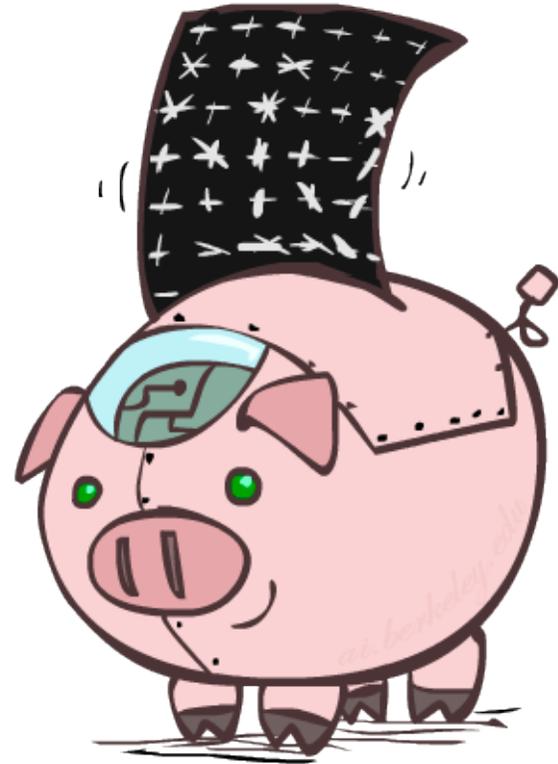
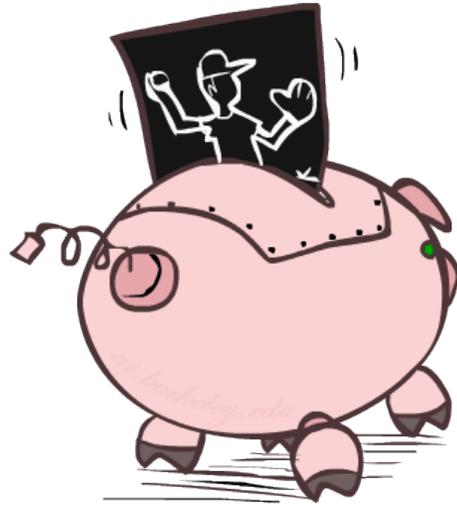
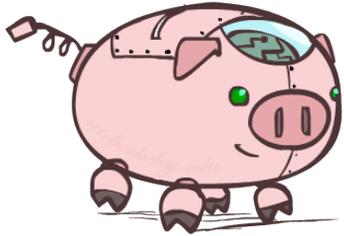
Computer Vision



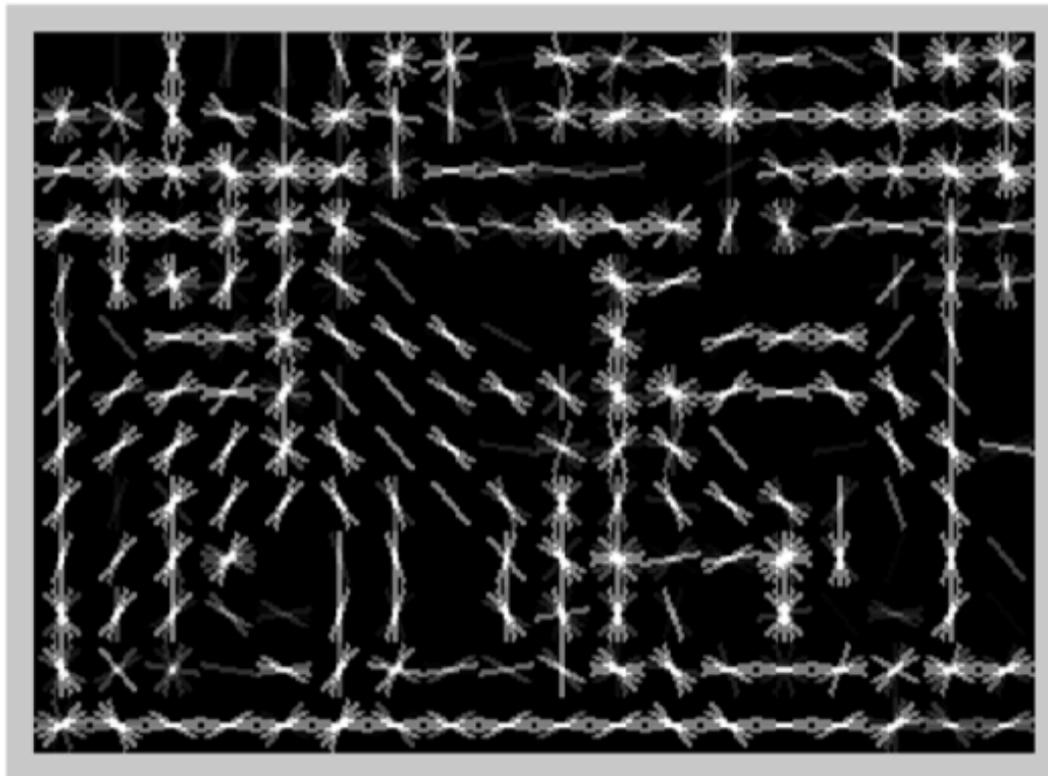
Object Detection



Manual Feature Design



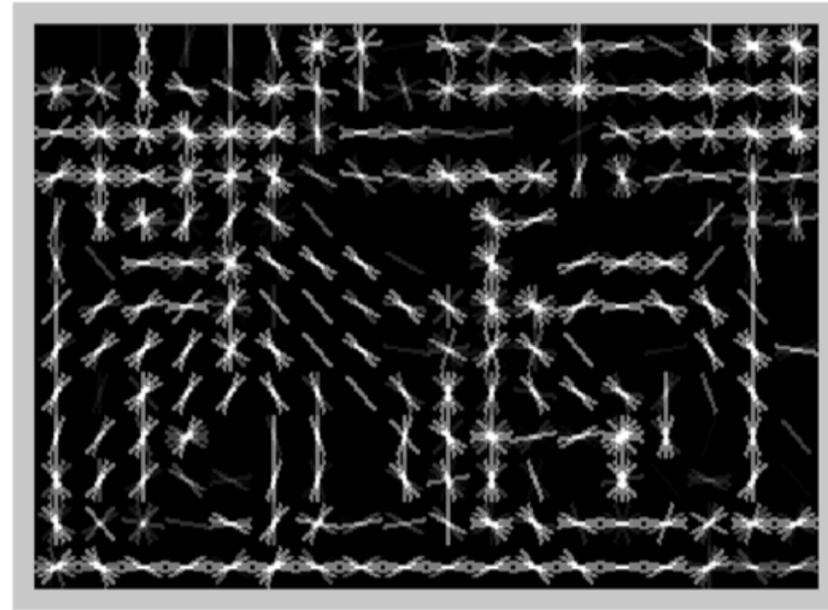
Features and Generalization



Features and Generalization



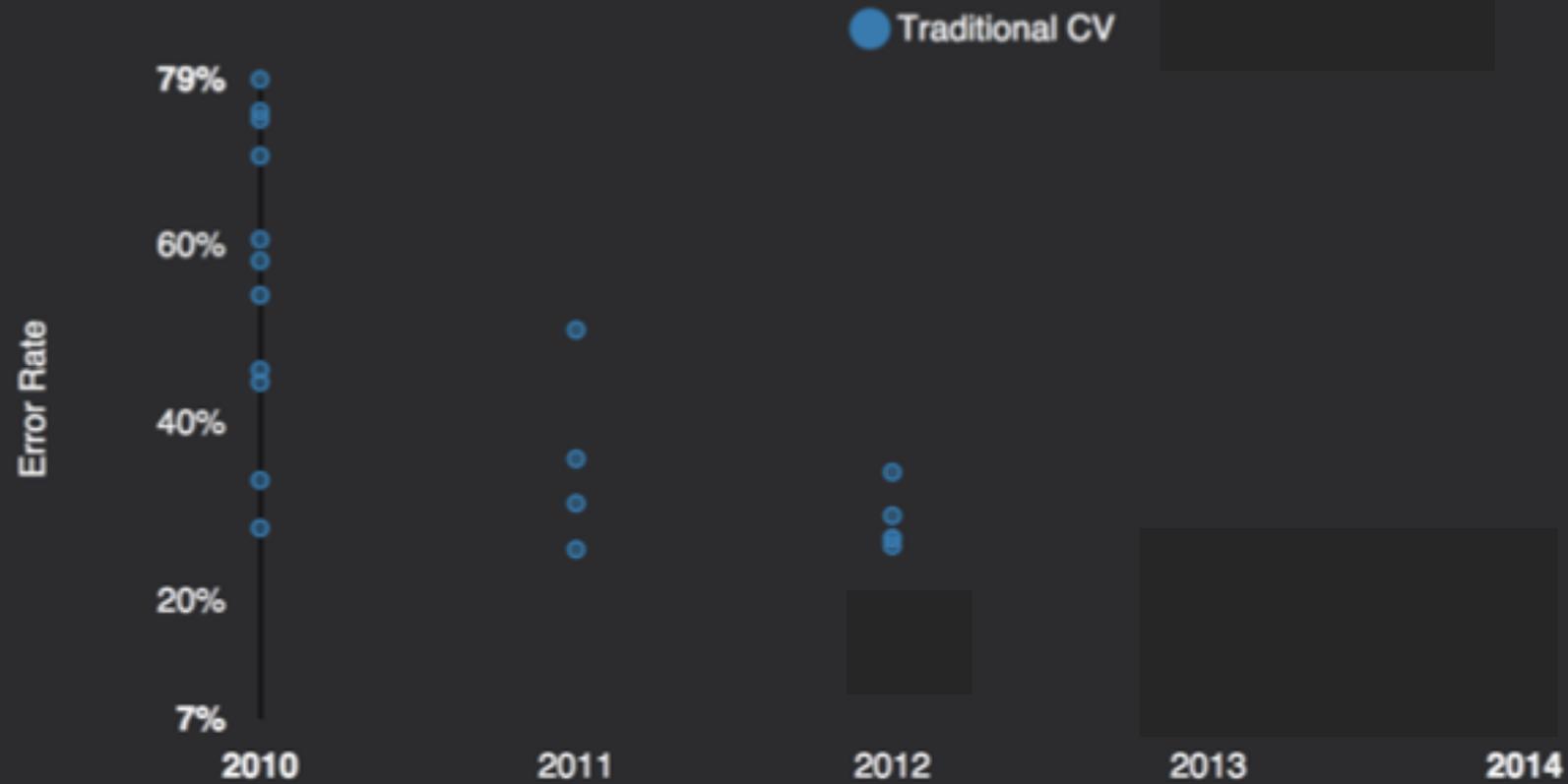
Image



HoG

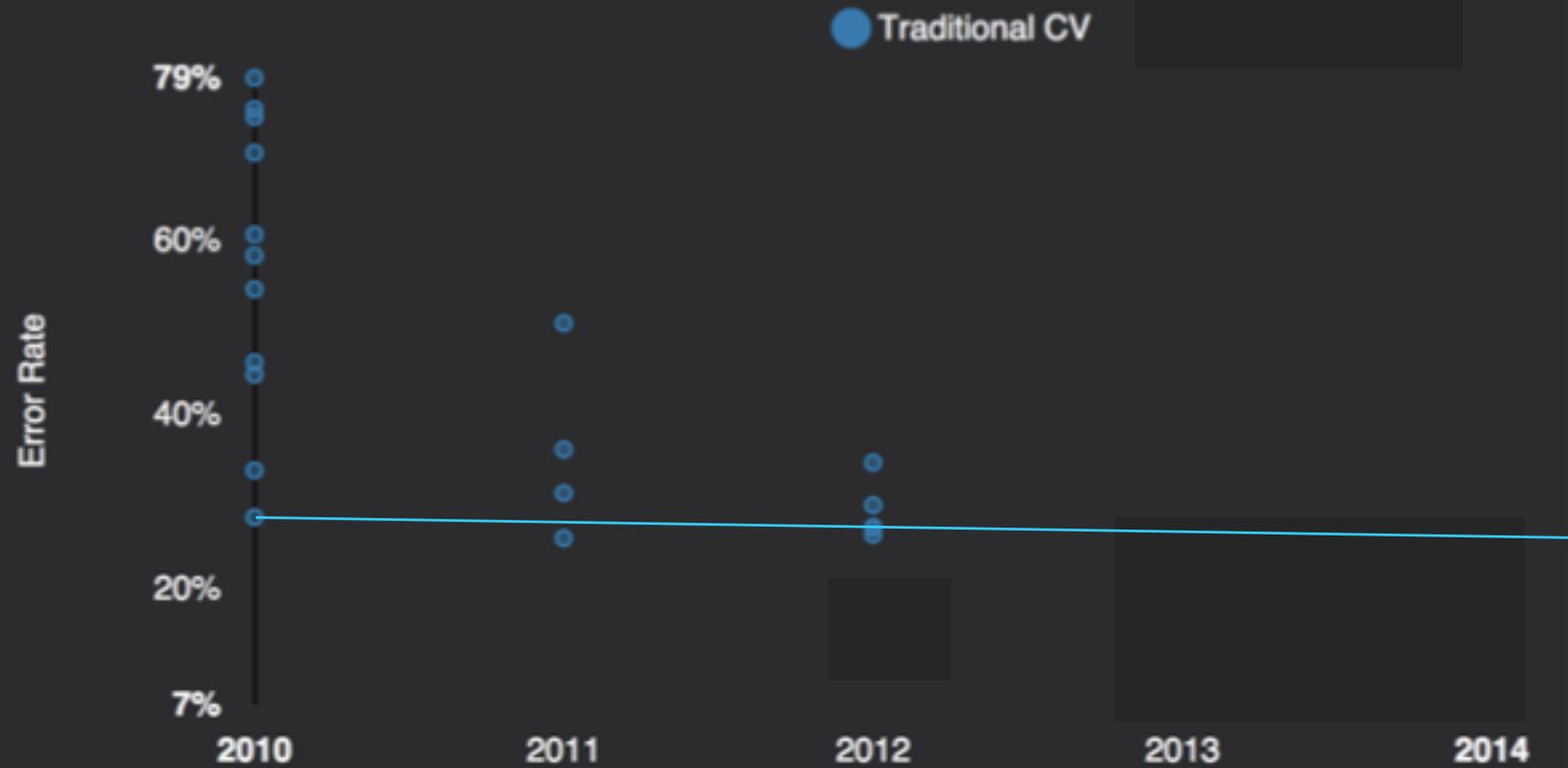
Performance

ImageNet Error Rate 2010-2014



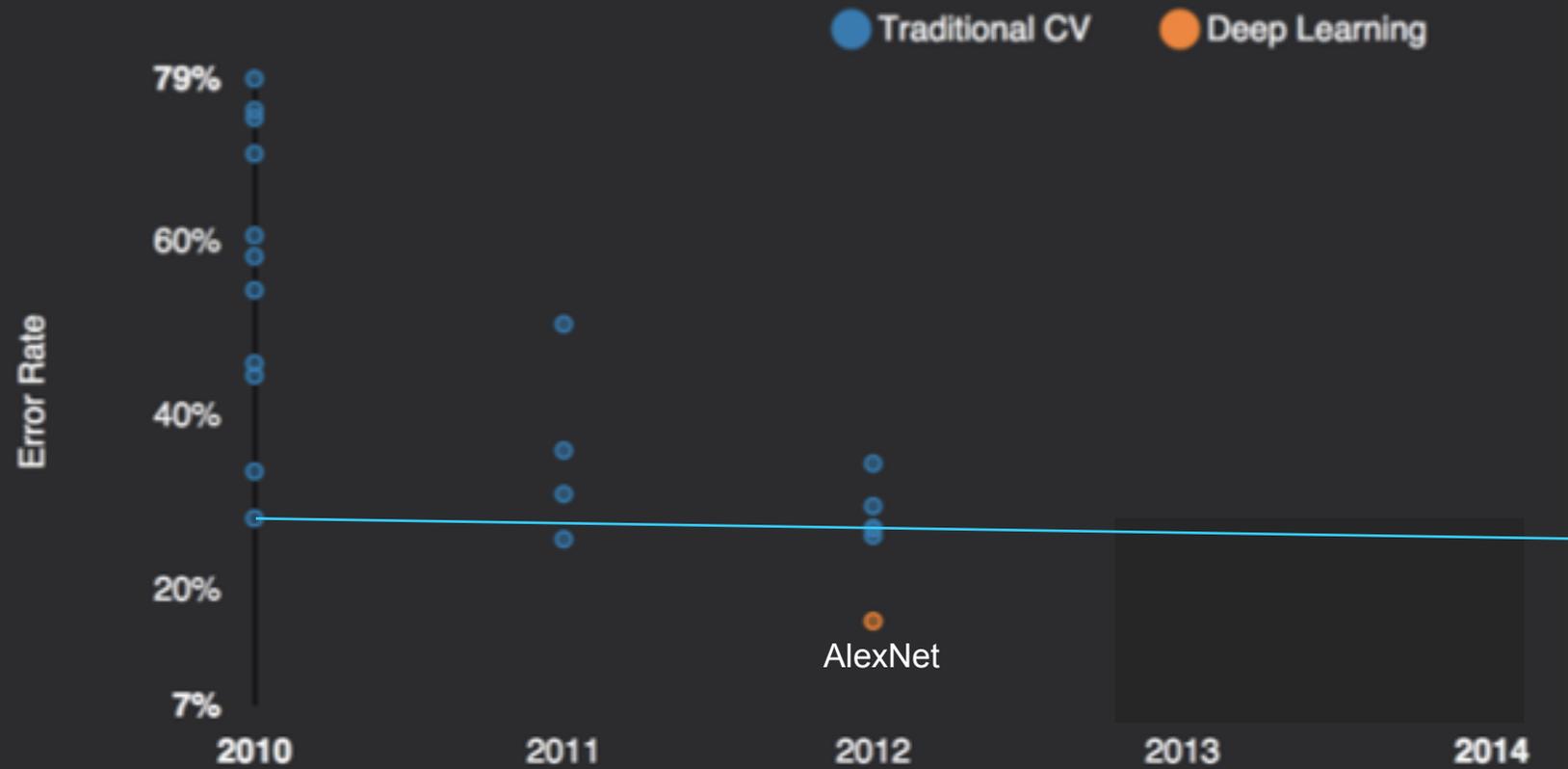
Performance

ImageNet Error Rate 2010-2014



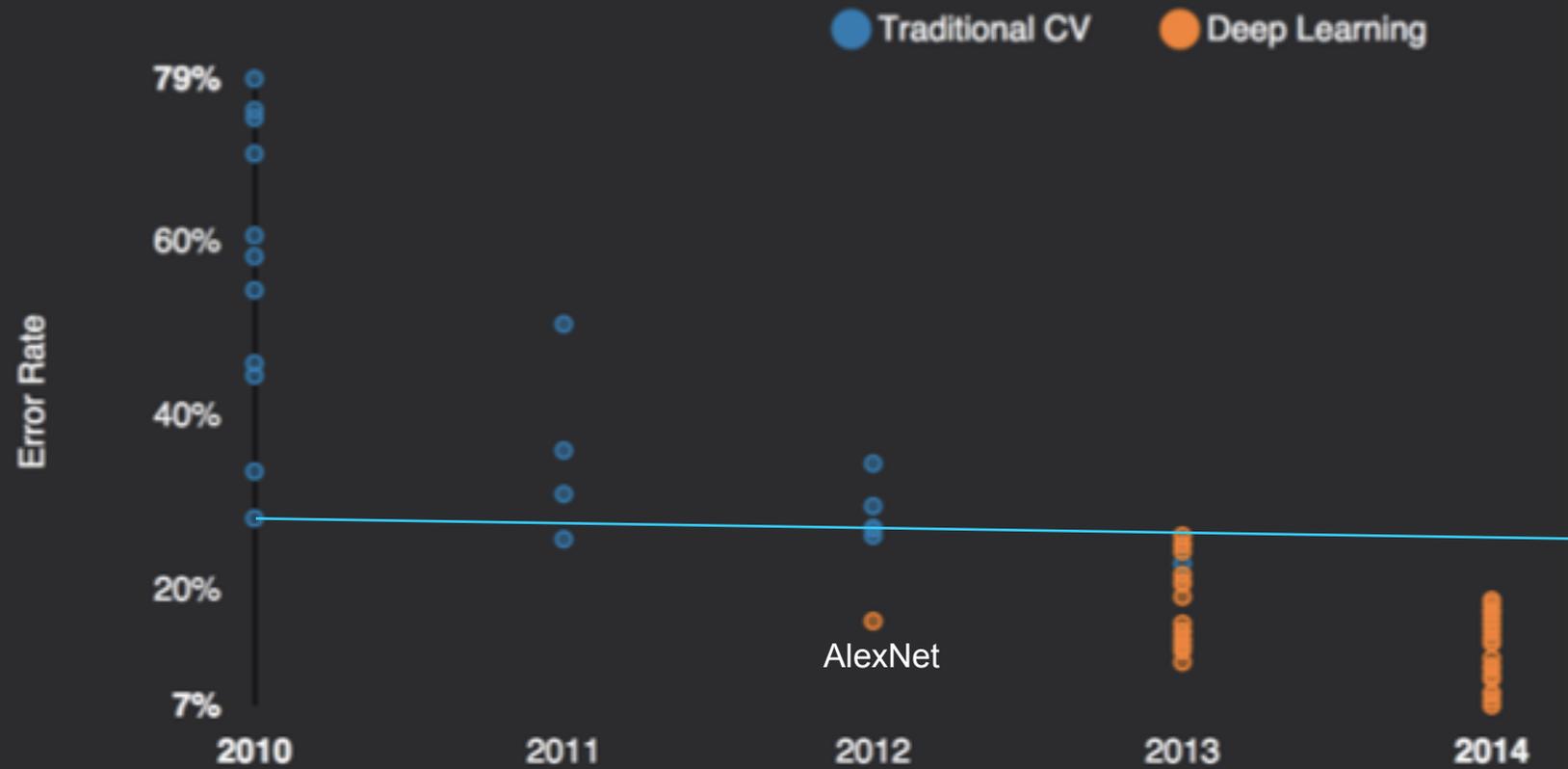
Performance

ImageNet Error Rate 2010-2014



Performance

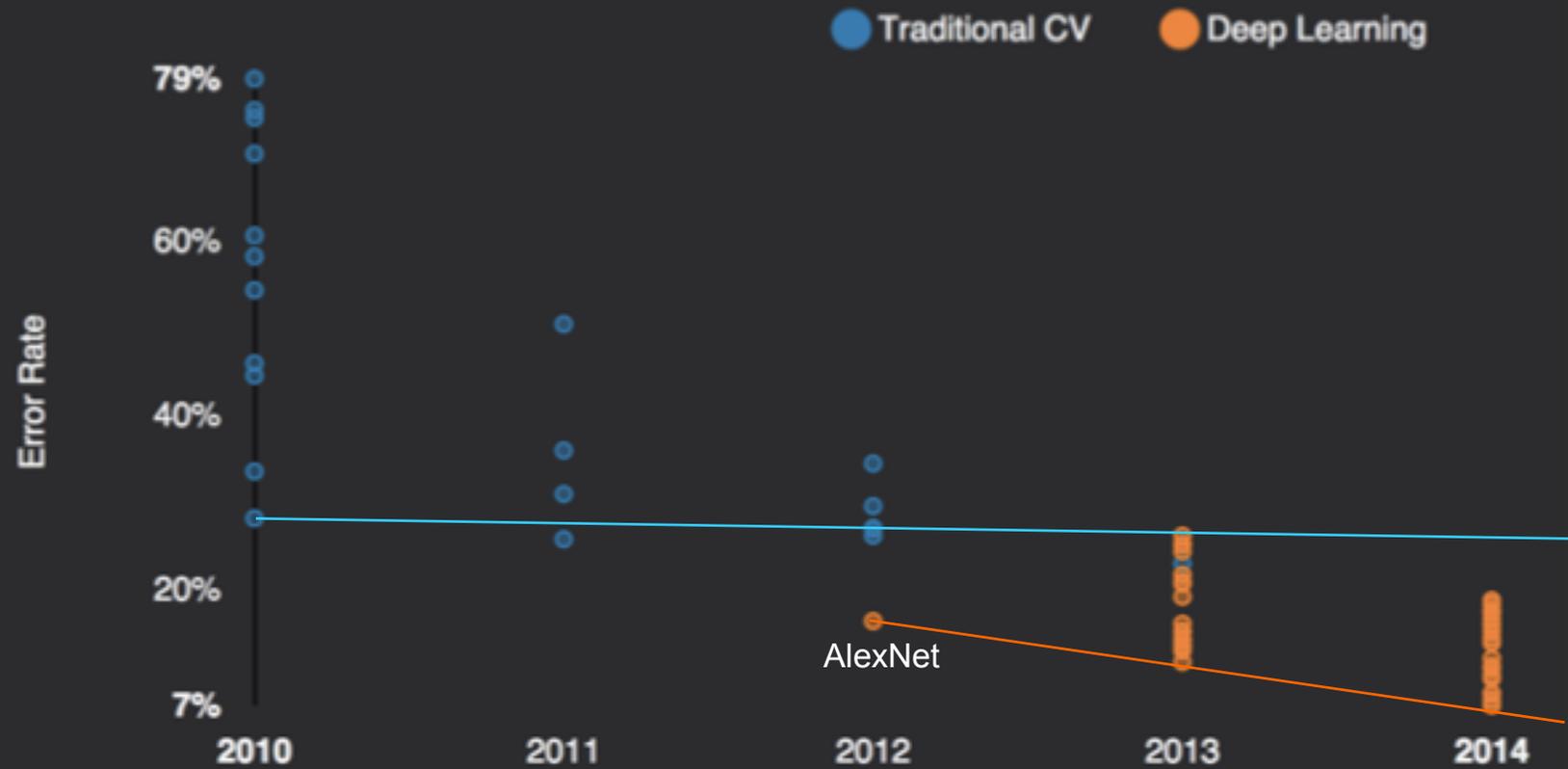
ImageNet Error Rate 2010-2014



graph credit Matt Zeiler, Clarifai

Performance

ImageNet Error Rate 2010-2014



MS COCO Image Captioning Challenge



"man in black shirt is playing guitar."



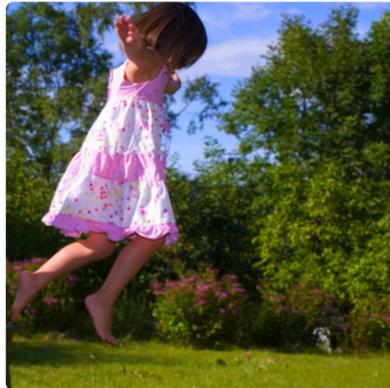
"construction worker in orange safety vest is working on road."



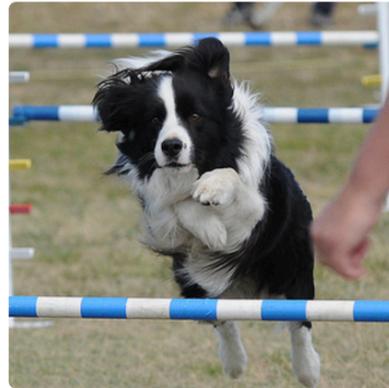
"two young girls are playing with lego toy."



"boy is doing backflip on wakeboard."



"girl in pink dress is jumping in air."



"black and white dog jumps over bar."



"young girl in pink shirt is swinging on swing."



"man in blue wetsuit is surfing on wave."

Karpathy & Fei-Fei, 2015; Donahue et al., 2015; Xu et al, 2015; many more

Visual QA Challenge

Stanislaw Antol, Aishwarya Agrawal, Jiasen Lu, Margaret Mitchell, Dhruv Batra, C. Lawrence Zitnick, Devi Parikh



What vegetable is on the plate?

Neural Net: broccoli

Ground Truth: broccoli



What color are the shoes on the person's feet ?

Neural Net: brown

Ground Truth: brown



How many school busses are there?

Neural Net: 2

Ground Truth: 2



What sport is this?

Neural Net: baseball

Ground Truth: baseball



What is on top of the refrigerator?

Neural Net: magnets

Ground Truth: cereal



What uniform is she wearing?

Neural Net: shorts

Ground Truth: girl scout



What is the table number?

Neural Net: 4

Ground Truth: 40



What are people sitting under in the back?

Neural Net: bench

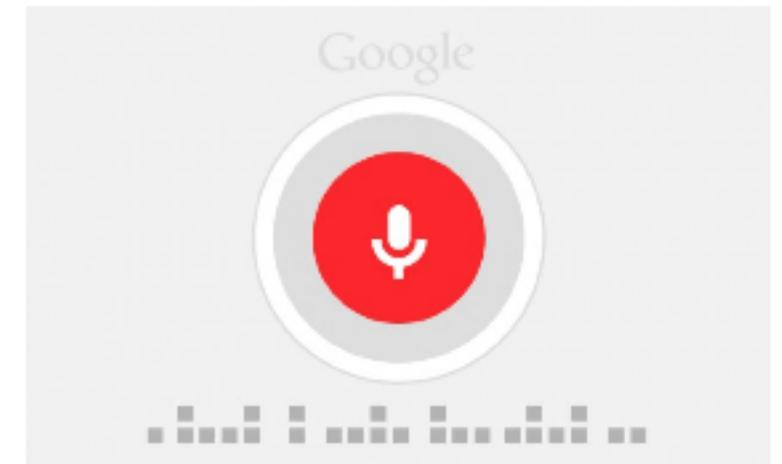
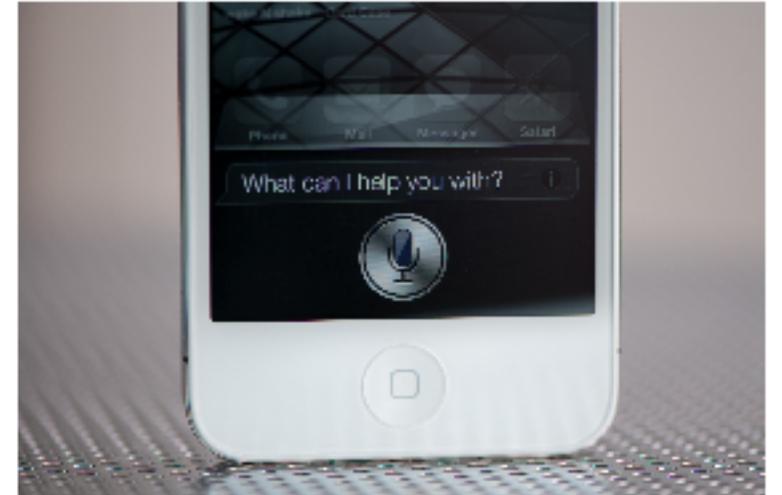
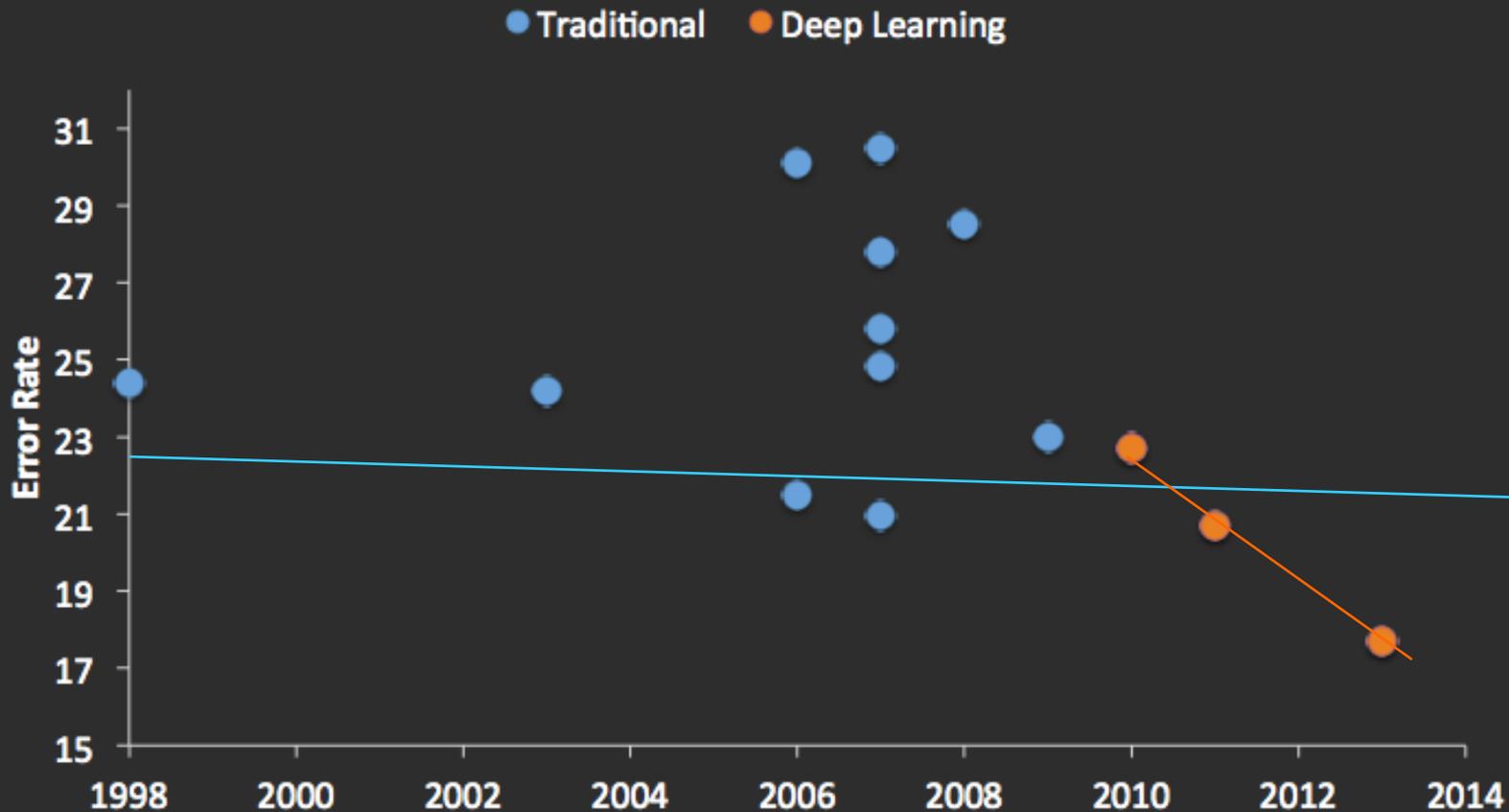
Ground Truth: tent

Speech and Natural Language Processing

- Different approaches to:
 - Modeling sequences of tokens
- Language Modeling: $P(x_t | x_{t-1})$
- Applications:
 - Machine Translation
 - Document Classification
 - Sentiment
 - Document types
 - Question Answering
 - etc

Speech Recognition

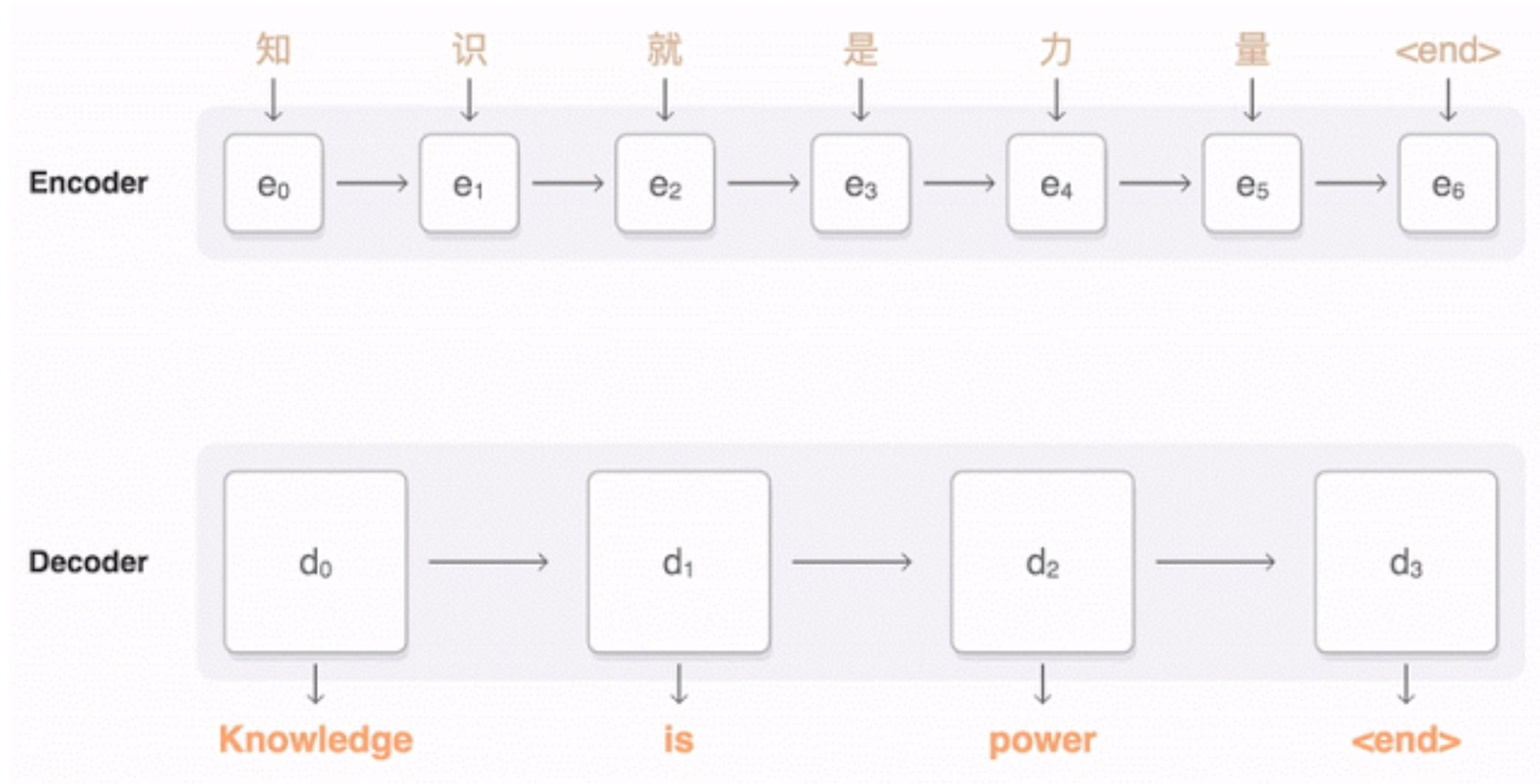
TIMIT Speech Recognition



graph credit Matt Zeiler, Clarifai

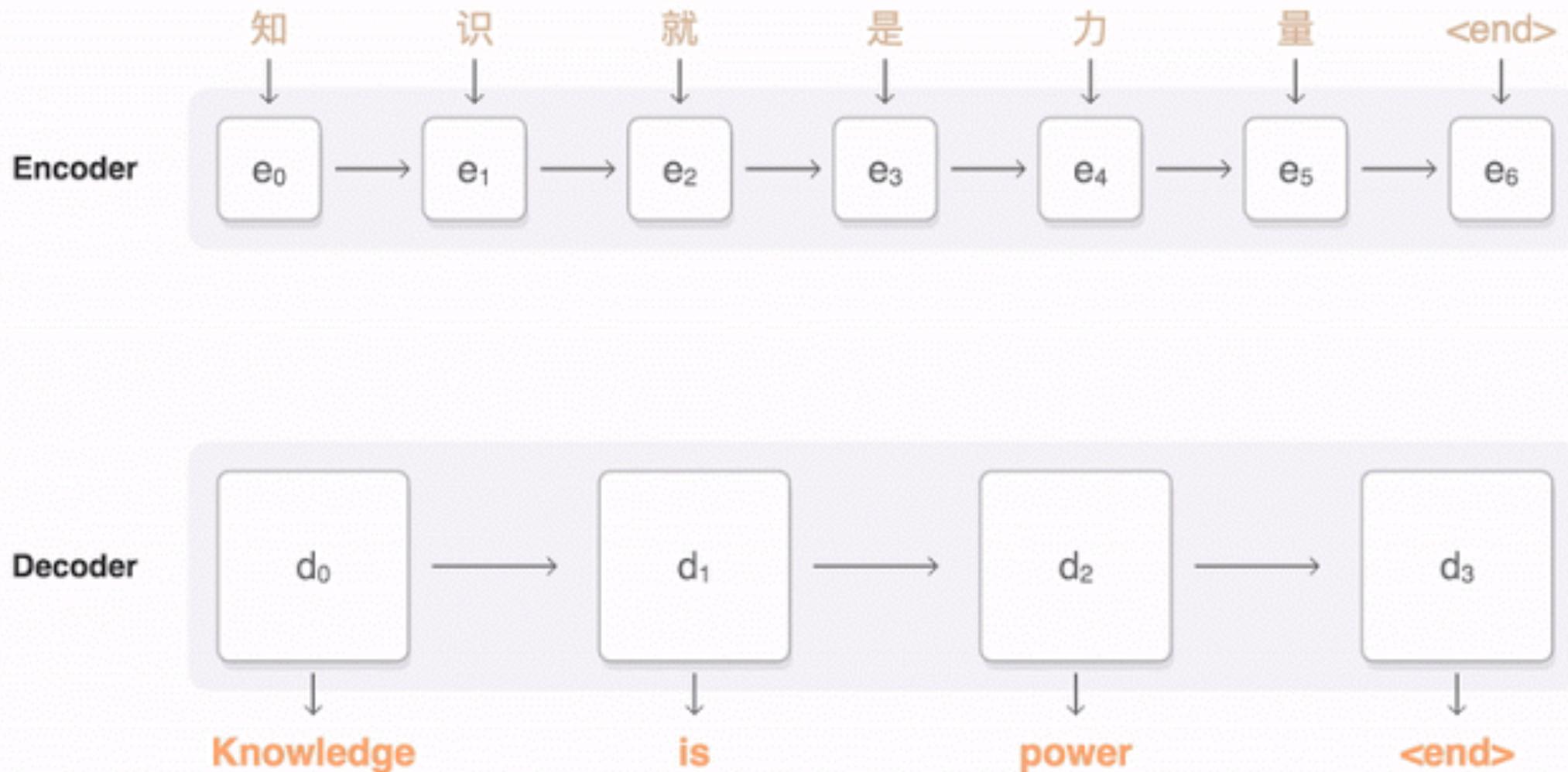
Machine Translation

Google Neural Machine Translation (in production)



Machine Translation

Google Neural Machine Translation (in production)



Question Answering

Context

Super Bowl 48 was an American football game to determine the champion of the National Football League (NFL) for the 2013 season. The National Football Conference champions Seattle Seahawks defeated the American Football Conference champions Denver Broncos. The Seahawks defeated the Broncos 43–8, the largest margin victory for an underdog and tied the third largest point differential overall (35) in Super Bowl history with Super Bowl XXVII (1993). It was the first time the winning scored over 40 points, while holding their opponent to under 10.

Question

Which NFL team represented the NFC at Super Bowl 48?

Answer

Seattle
Seahawks

Pipeline Approach for Question Answering

- Feature engineering
- Classifying phrases

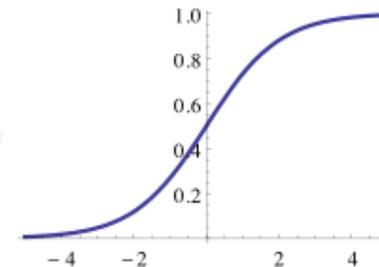
Super Bowl 48 was an American football game to determine the champion of the National Football League (NFL) for the 2013 season. The National Football Conference champions Seattle Seahawks defeated the American Football Conference champions Denver Broncos. The Seahawks defeated the Broncos 43–8, the largest margin victory for an underdog and tied the third largest point differential overall (35) in Super Bowl history with Super Bowl XXVII (1993). It was the first time the winning scored over 40 points, while holding their opponent to under 10.

Which NFL team represented the AFC at Super Bowl 48?

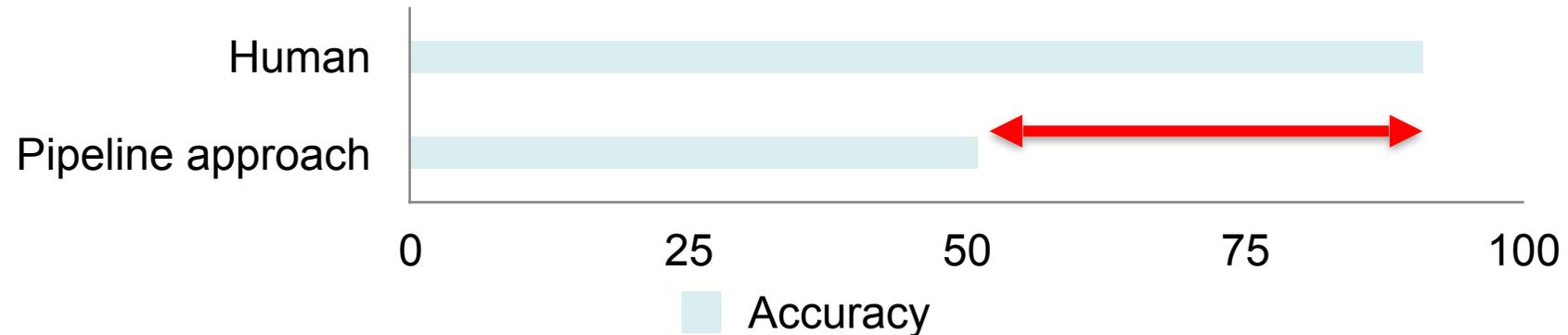
words, types, frequencies
dependency relations



f_1, f_2, \dots, f_n →



Pipeline Approach Results



- Dataset: Stanford Question Answering Dataset (SQuAD) [Rajpurkar et al 2016]:
 - 100k Wikipedia documents with question
- Accuracy: percentage of correctly predicted phrases

Neural Approach

[ICLR'17]

Find a function that assigns a high score to the the correct answer given the context and question



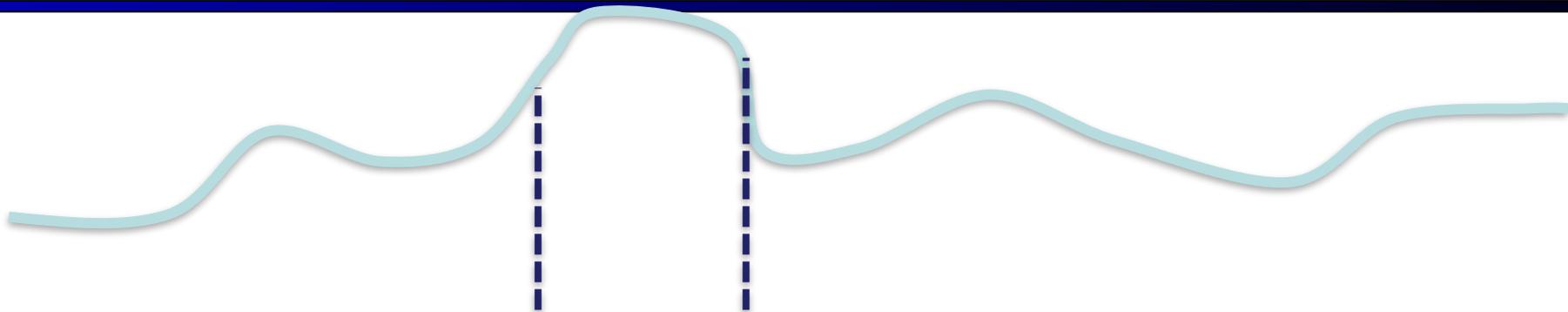
Super Bowl 48 was an American football game to determine the champion of the National Football League (NFL) for the 2013 season. The National Football Conference champions Seattle Seahawks defeated the American Football Conference champions Denver Broncos. The Seahawks defeated the Broncos 43–8, the largest margin victory for an underdog and tied the third largest point differential overall (35) in Super Bowl history with Super Bowl XXVII (1993). It was the first time the winning scored over 40 points, while holding their opponent to under 10.

Context

Which NFL team represented the NFC at Super Bowl 48?

Question

Seattle Seahawks



The National Football Conference champions Seattle Seahawks defeated the American Football Conference champions Denver Broncos.



$f($ Encoding , Encoding)



Super Bowl 48 was an American football game to determine the champion of the National Football League (NFL) for the 2013 season. The National Football Conference champions Seattle Seahawks defeated the American Football Conference champions Denver Broncos. The Seahawks defeated the Broncos 43–8, the largest margin victory for an underdog and tied the third largest point differential overall (35) in Super Bowl history with Super Bowl XXVII (1993). It was the first time the winning scored over 40 points, while holding their opponent to under 10.

Context



, Which NFL team represented the NFC at Super Bowl 48?

Question

Question Answering Leaderboard

Jan 1, 2017

Test Set Leaderboard

Since the release of our dataset ([and paper](#)), the community has made rapid progress! Here are the ExactMatch (EM) and F1 scores of the best models evaluated on the test and development sets of v1.1.

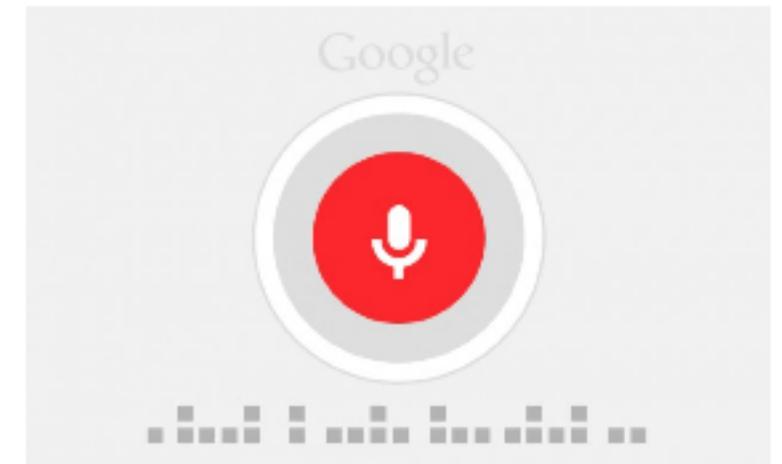
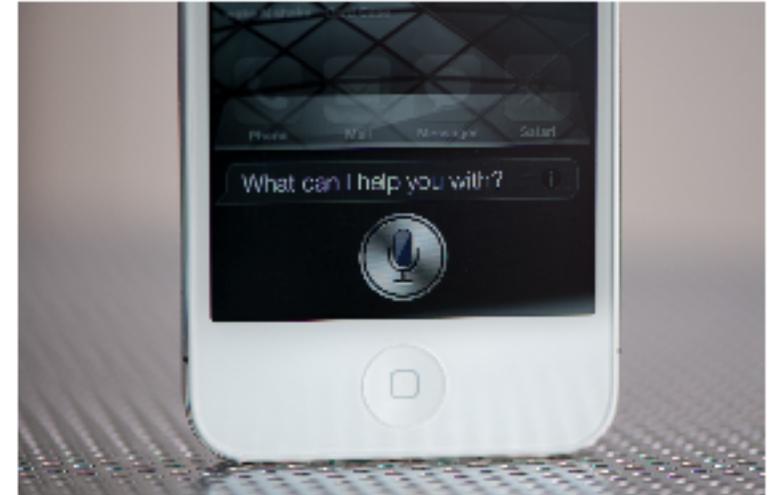
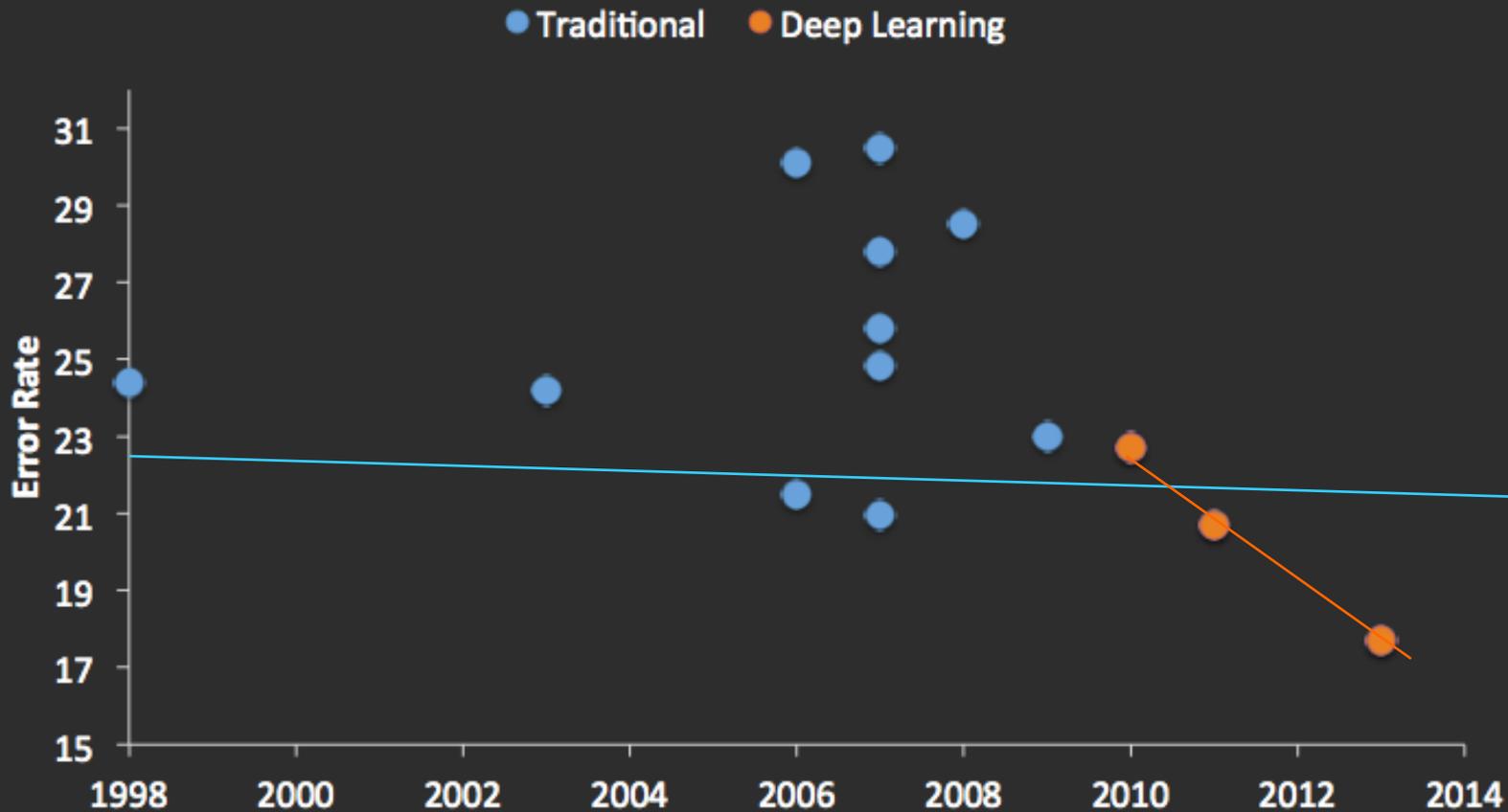
Rank	Model	Test EM	Test F1
1	BiDAF (ensemble) Allen Institute for AI & University of Washington (Seo et al. '16)	73.3	81.1
2	Dynamic Coattention Networks (ensemble) Salesforce Research (Xiong & Zhong et al. '16)	71.6	80.4
2	r-net (ensemble) Microsoft Research Asia	72.1	79.7
4	r-net (single model) Microsoft Research Asia	68.4	77.5
5	BiDAF (single model) Allen Institute for AI & University of Washington (Seo et al. '16)	68.0	77.3
5	Multi-Perspective Matching (ensemble) IBM Research	68.2	77.2

March 8, 2021

Rank	Model	EM	F1
	Human Performance Stanford University (Rajpurkar & Jia et al. '18)	86.831	89.452
1	FPNet (ensemble) Ant Service Intelligence Team Feb 21, 2021	90.871	93.183
2	IE-Net (ensemble) RICOH_SRCB_DML Feb 24, 2021	90.758	93.044
3	SA-Net on Albert (ensemble) QIANXIN Apr 06, 2020	90.724	93.011
4	SA-Net-V2 (ensemble) QIANXIN May 05, 2020	90.679	92.948
4	Retro-Reader (ensemble) Shanghai Jiao Tong University http://arxiv.org/abs/2001.09694 Apr 05, 2020	90.578	92.978
4	FPNet (ensemble) YuYang Feb 05, 2021	90.600	92.899
5	EntitySpanFocusV2 (ensemble) RICOH_SRCB_DML Dec 01, 2020	90.521	92.824
5	ATRLP+PV (ensemble) Hithink RoyalFlush Jul 31, 2020	90.442	92.877
5	ELECTRA+ALBERT+EntitySpanFocus (ensemble) SRCB_DML May 04, 2020	90.442	92.839

Speech Recognition

TIMIT Speech Recognition



graph credit Matt Zeiler, Clarifai