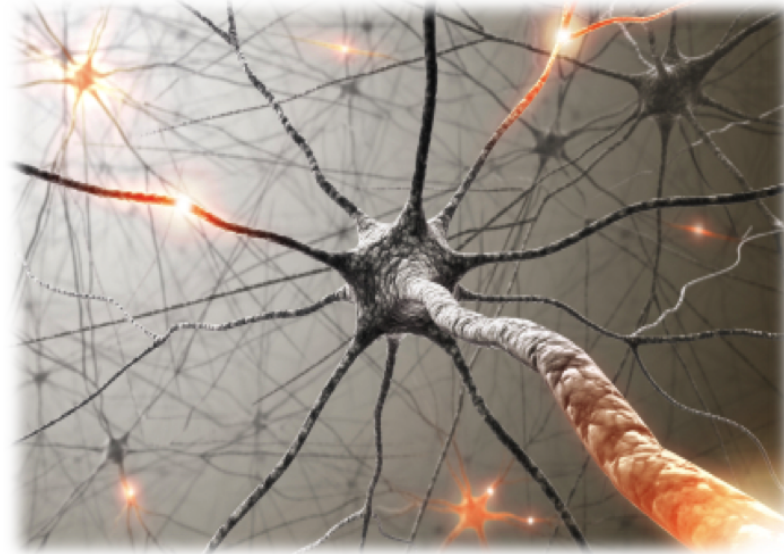# Lecture 18: Deep Learning

A brief history

- 1943: neural networks $\Leftrightarrow$ logical circuits (McCulloch/Pitts)

- 1949: "cells that fire together wire together" learning rule (Hebb)

- 1969: theoretical limitations of neural networks (Minsky/Papert)

- 1974: backpropagation for training multi-layer networks (Werbos)

- 1986: popularization of backpropagation (Rumelhardt, Hinton, Williams)

A brief history

- 1980: Neocognitron, a.k.a. convolutional neural networks (Fukushima)

- 1989: backpropagation on convolutional neural networks (LeCun)

- 1990: recurrent neural networks (Elman)

- 1997: Long Short-Term Memory networks (Hochreiter/Schmidhuber)

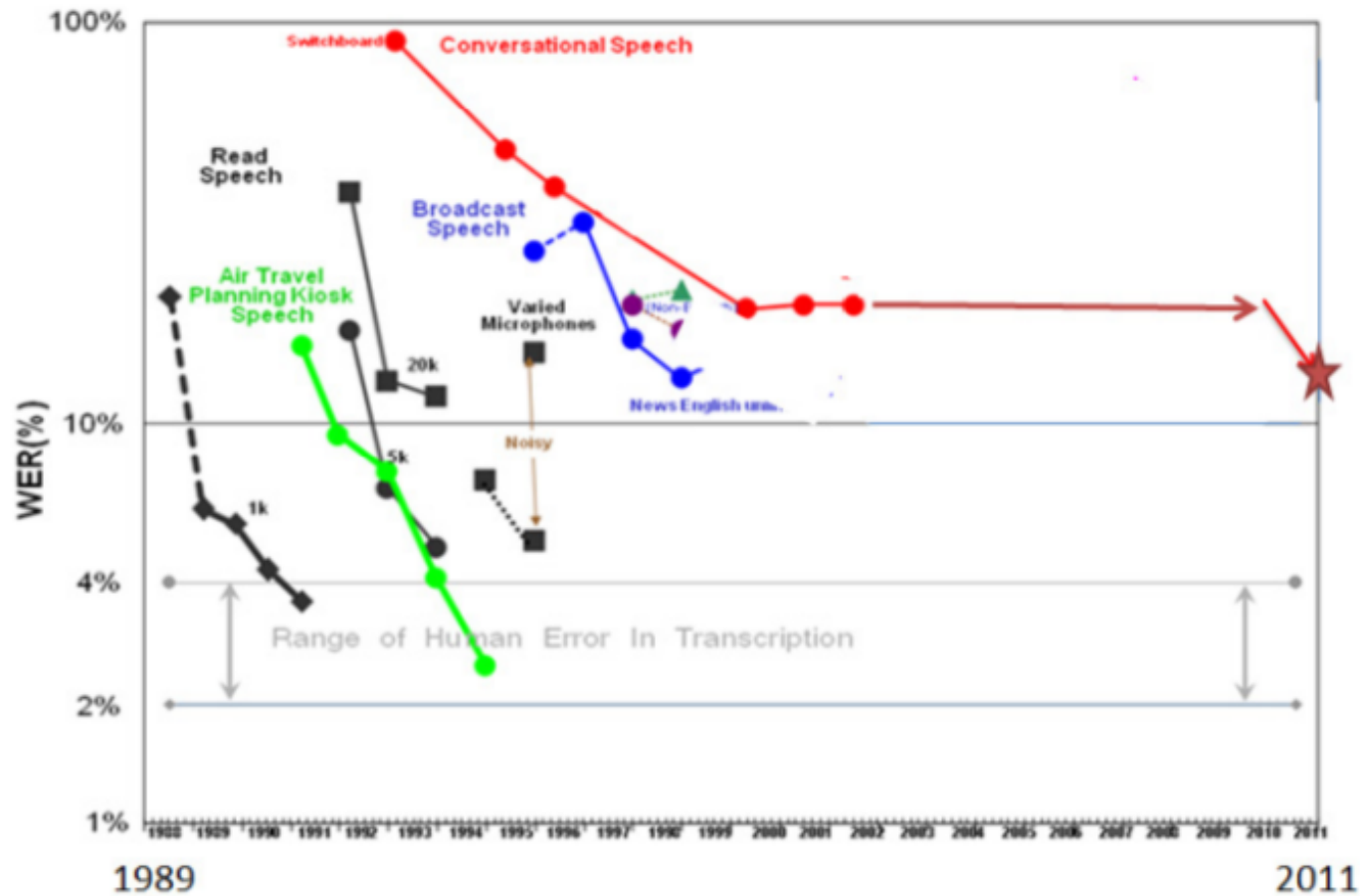- 2006: unsupervised layerwise training of deep networks (Hinton et al.)

Google Trends

Query: deep learning

Interest over time ?

100
75
50
25

Jan 1, 2004   Note   Mar 1, 2010   Note   May 1, 2016
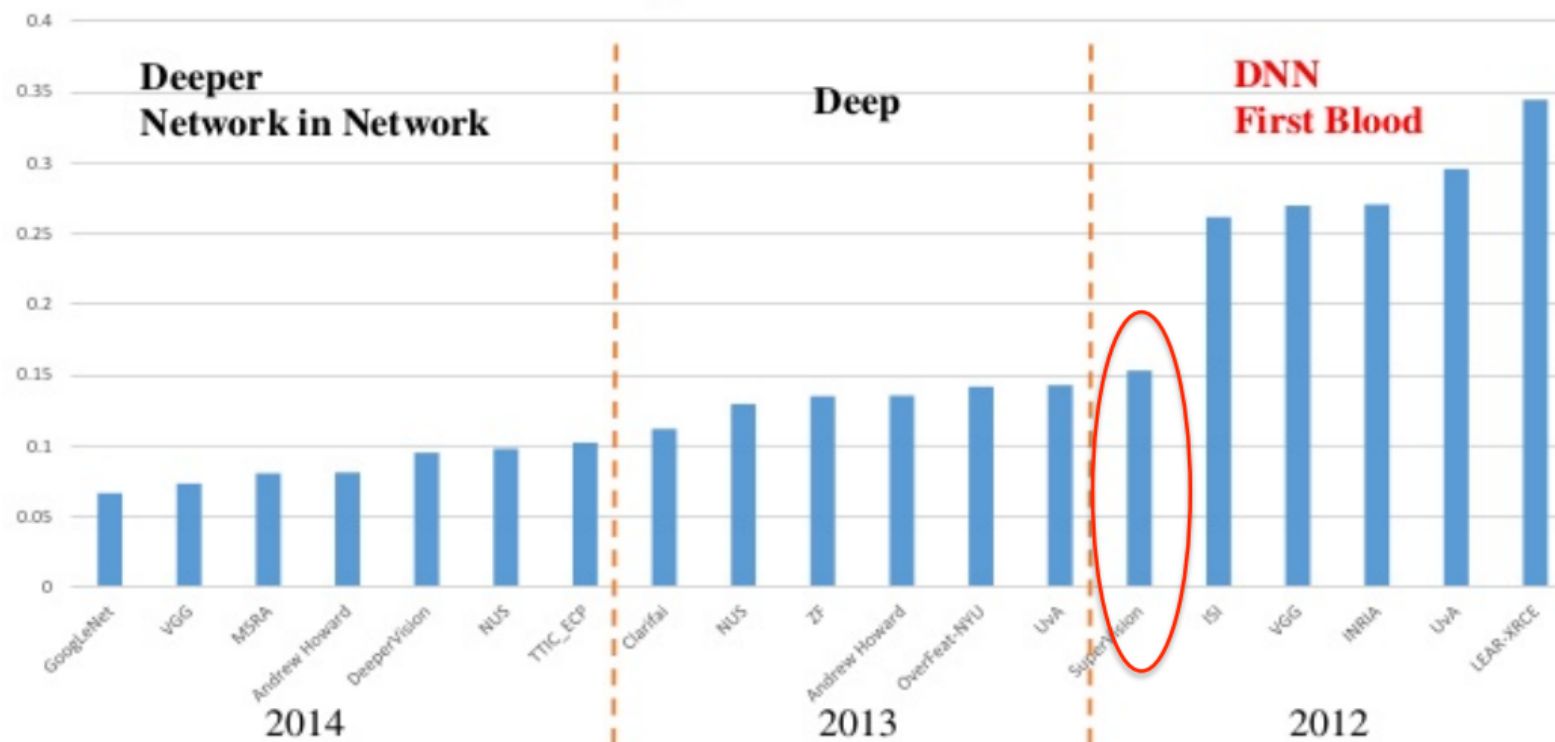
Speech recognition (2009-2011)

- Steep drop in WER due to deep learning

- IBM, Google, Microsoft all switched over from GMM-HMM

Object recognition (2012)

- Landslide win in ILSVRC object recognition competition

- Computer vision community switched to CNNs

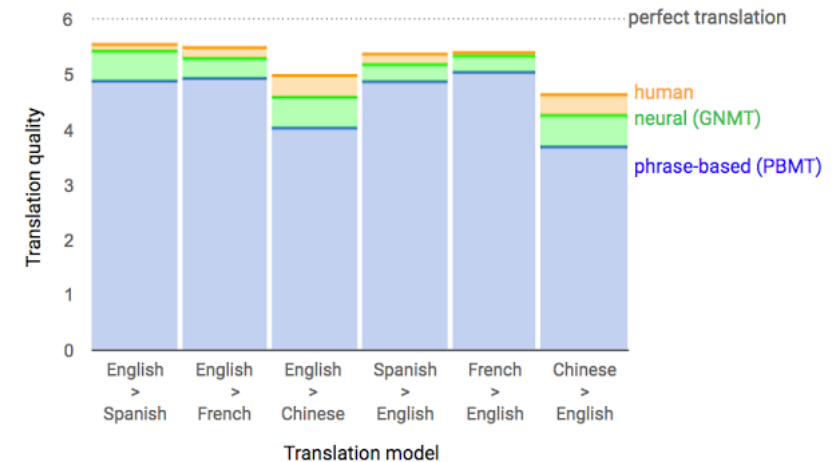- Simpler than hand-engineered features (SIFT)

Machine translation (2016)

Google's Neural Machine Translation System: Bridging the Gap
between Human and Machine Translation

Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi
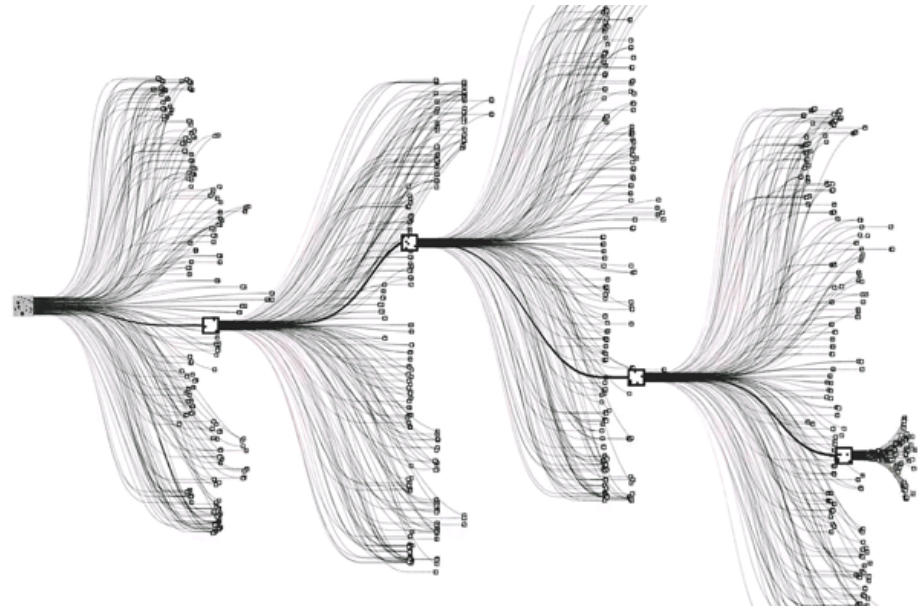`yonghui,schuster,zhifengc,qvl,mnorouzi@google.com`

Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey,
Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Łukasz Kaiser,
Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens,
George Kurian, Nishant Patil, Wei Wang, Cliff Young, Jason Smith, Jason Riesa,
Alex Rudnick, Oriol Vinyals, Greg Corrado, Macduff Hughes, Jeffrey Dean

- Decisive wins have taken longer to achieve in NLP (words are meaningful in a way that pixels are not)

- Current state-of-the-art in machine translation

Go (2016)

- Defeated world champion Le Sedol 4-1

- Simple architecture (in contrast, DeepBlue was search + hand-crafted heuristics)

- 2017: AlphaGoZero does not require human expert games as supervision

What is deep learning?

A family of techniques for learning compositional vector representations of complex data.