

ELEC 677:  
Recurrent Neural Network Applications &  
Recurrent Neural Network Language Models  
Lecture 9

**Ankit B. Patel, CJ Barberan**

*Baylor College of Medicine (Neuroscience Dept.)*

*Rice University (ECE Dept.)*

11-8-2016

# Latest News

# Facebook works on TorchCraft

- StarCraft is the next battleground for AI to master
- Test deep learning models on a real-time strategy game
  - Code will be using Torch
  - Code to be released soon

# DeepMind and Blizzard to let AI learn from StarCraft II

- Make StarCraft II a new frontier of competitive gaming AI research
- Release early next year



# ICLR 2017 Paper Submissions

- More than 500 papers were submitted
  - Conference will be in France
- List of Submissions
  - ICLR 2017

# Bored Yann Lecun & Election



**Bored Yann LeCun** @boredyannlecun · 8h

Pumped up for this election. Refreshingly, it'll be the first competition since 2012 that won't be won by a ConvNet [#feelthelearn](#)

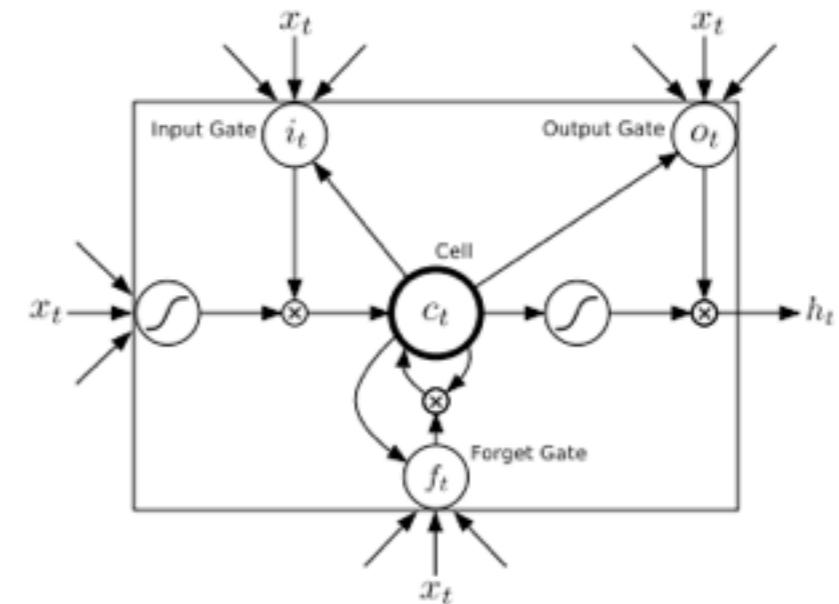


...

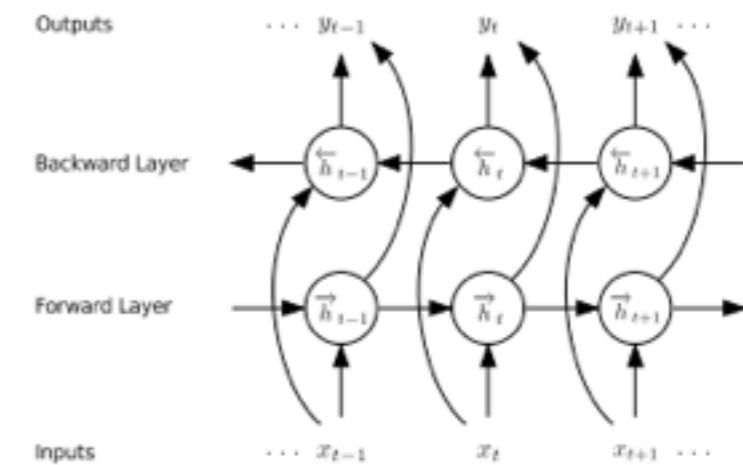
# Applications

# Phoneme Recognition

- To learn about context, future has information as much as the past ==> bidirectional LSTM
- Use bidirectional LSTM/HMM using Viterbi training

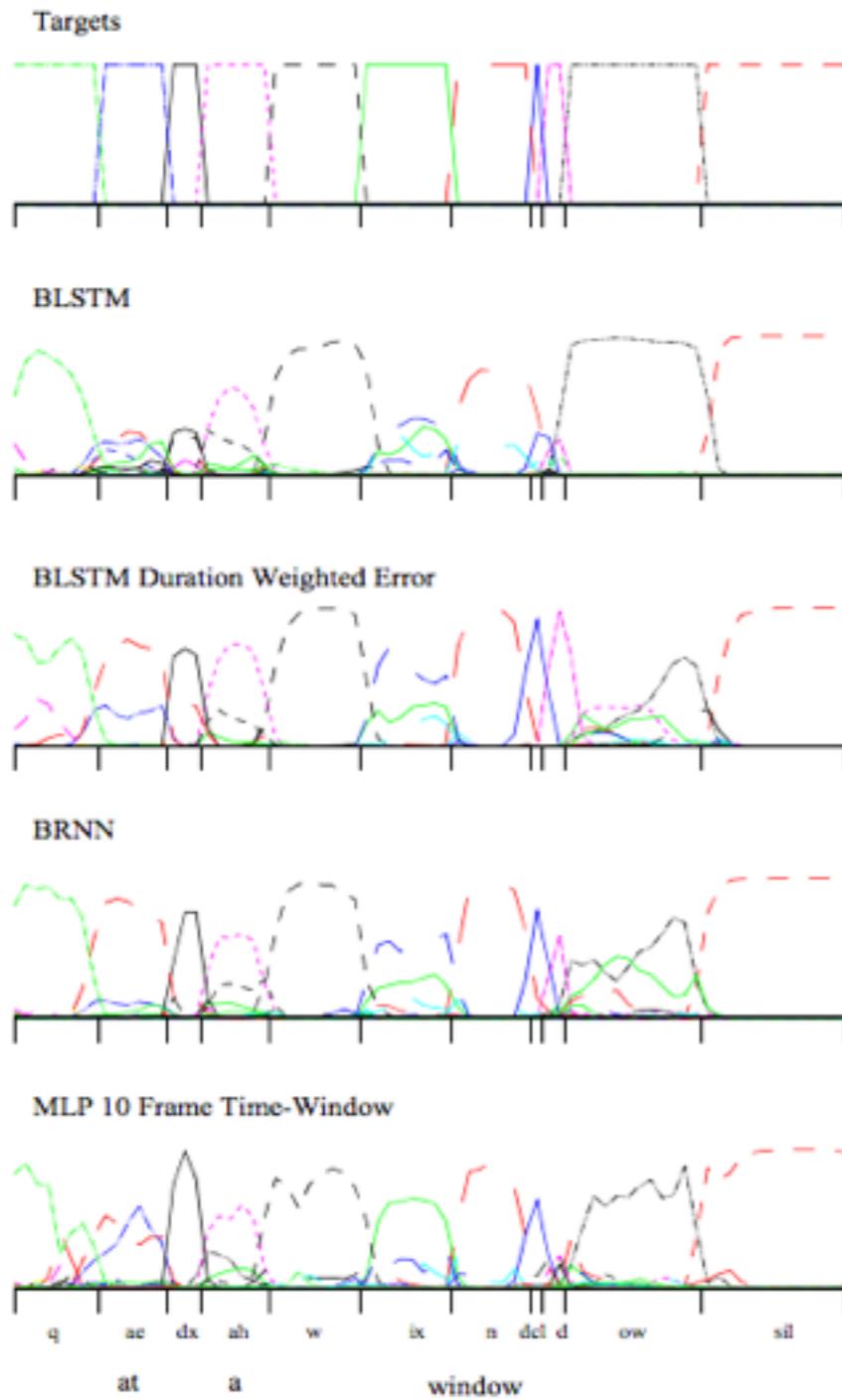


**Fig. 1.** Long Short-term Memory Cell

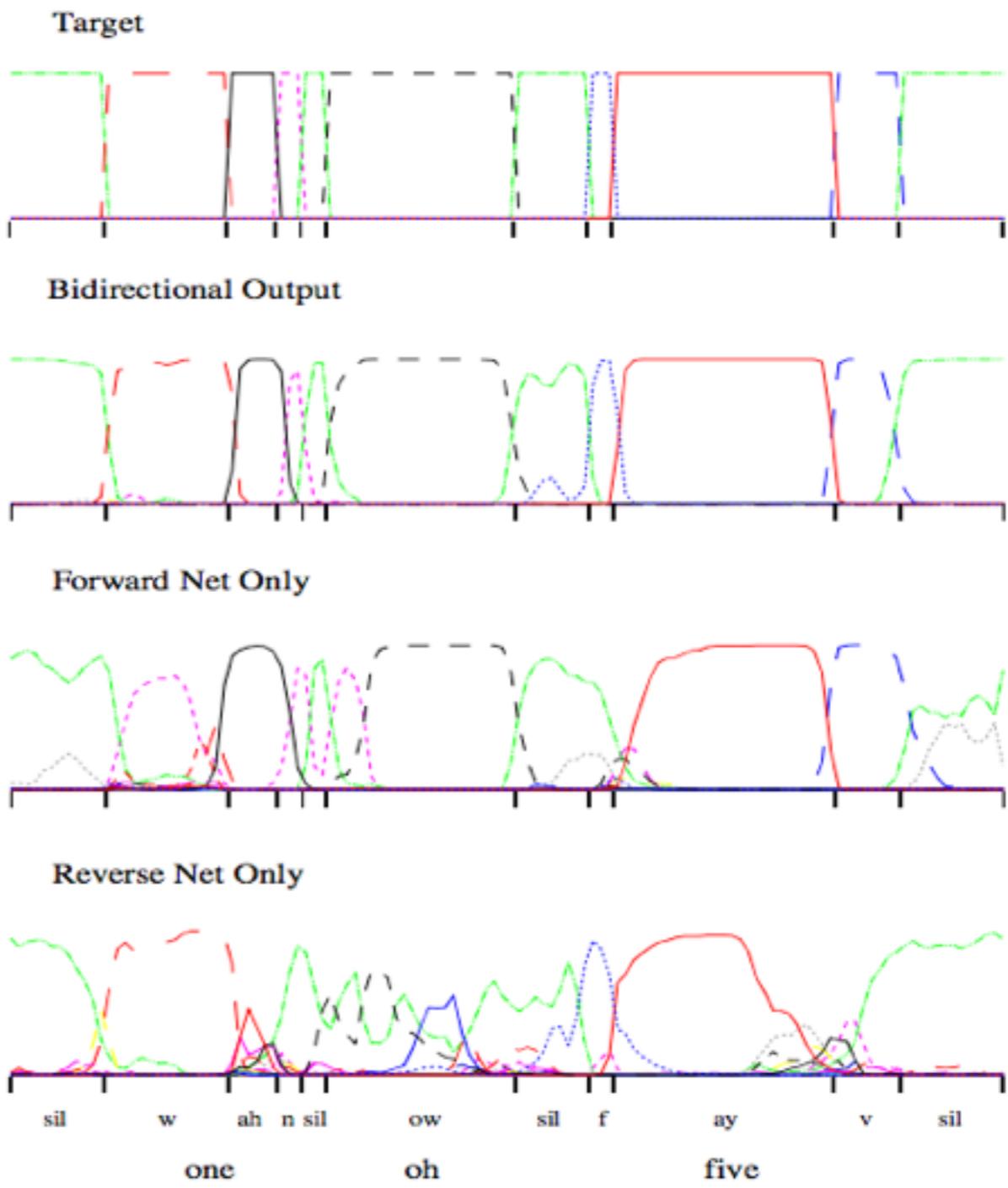


**Fig. 2.** Bidirectional RNN

# Phoneme Recognition

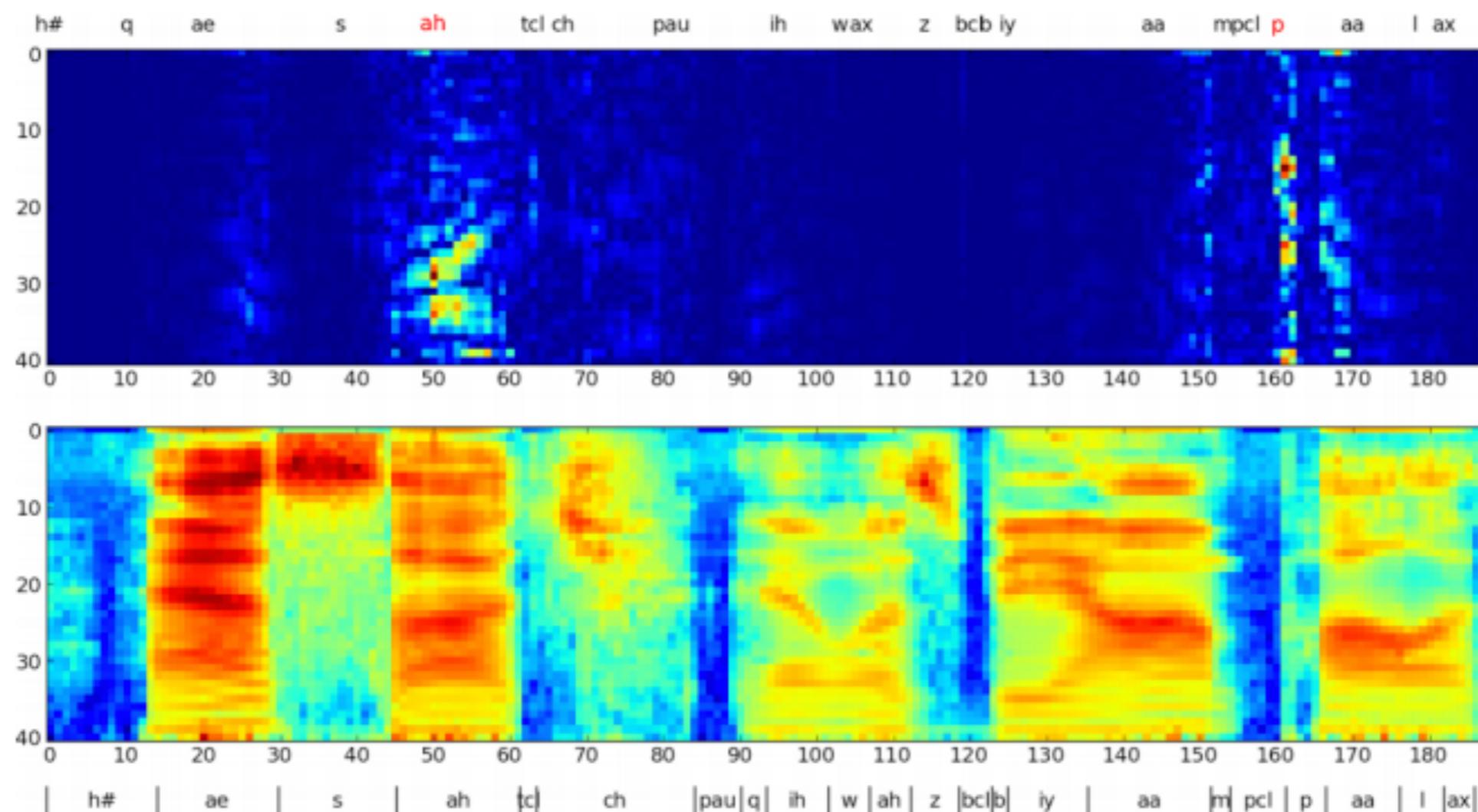


Earlier



Later

# Visualizing the Features



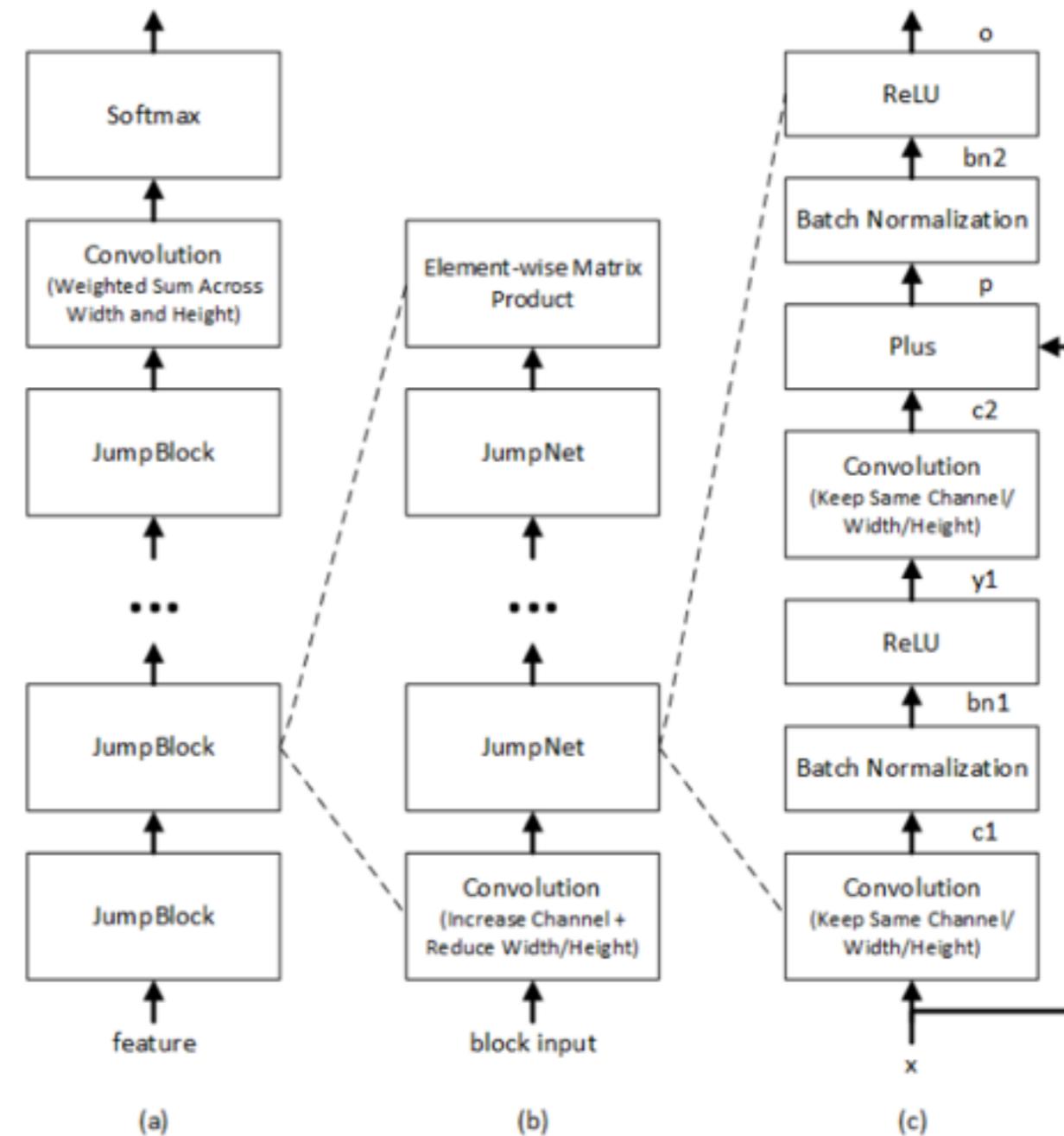
# Phoneme Recognition Results

**Table 1.** Framewise Phoneme Classification

Network	Training Set	Test Set	Epochs
BLSTM	77.4%	69.8%	21
BRNN	76.0%	69.0%	170
BLSTM Weighted Error	75.7%	68.9%	15
LSTM (4 frame delay)	77.5%	65.5%	33
RNN (4 frame delay)	70.8%	65.1%	144
LSTM (0 frame delay)	70.9%	64.6%	15
RNN (0 frame delay)	69.9%	64.5%	120

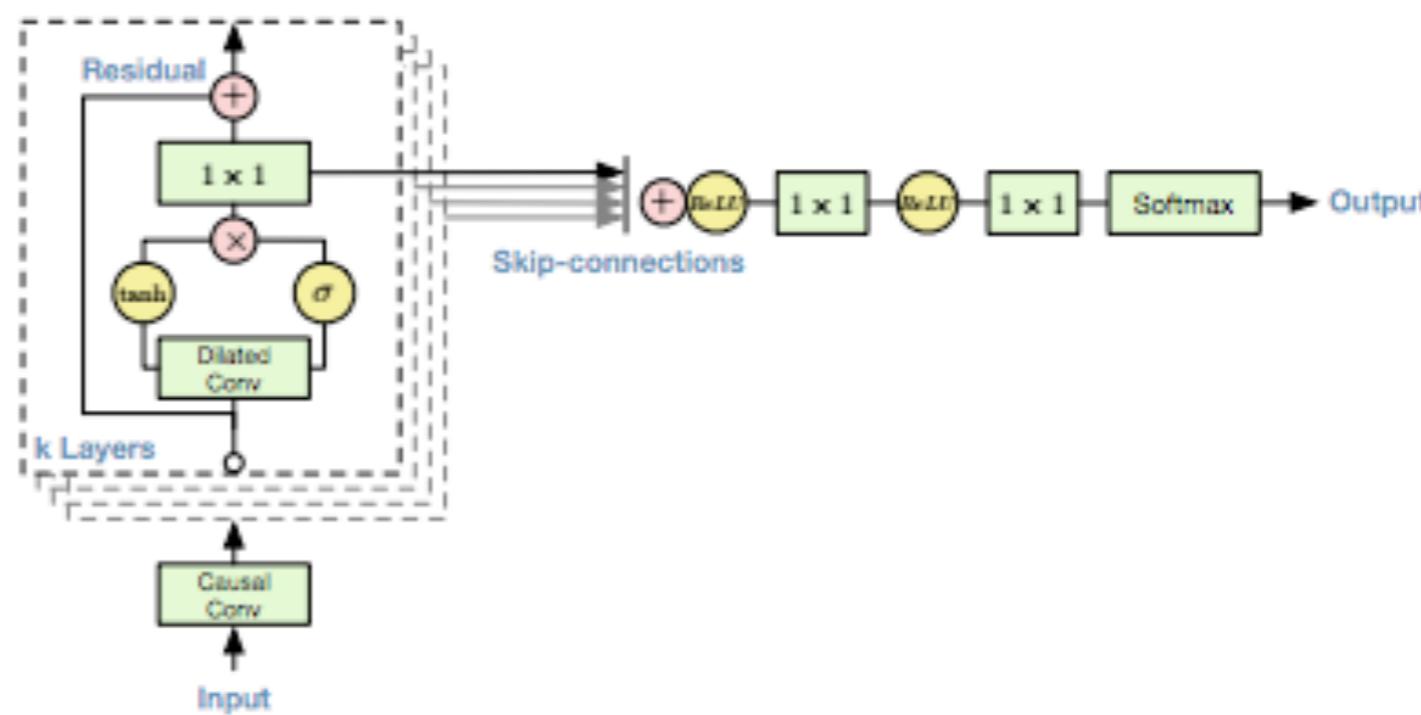
# Conversational Speech Recognition

- Achieving human parity
- 3 CNNs
- 6 layer LSTM



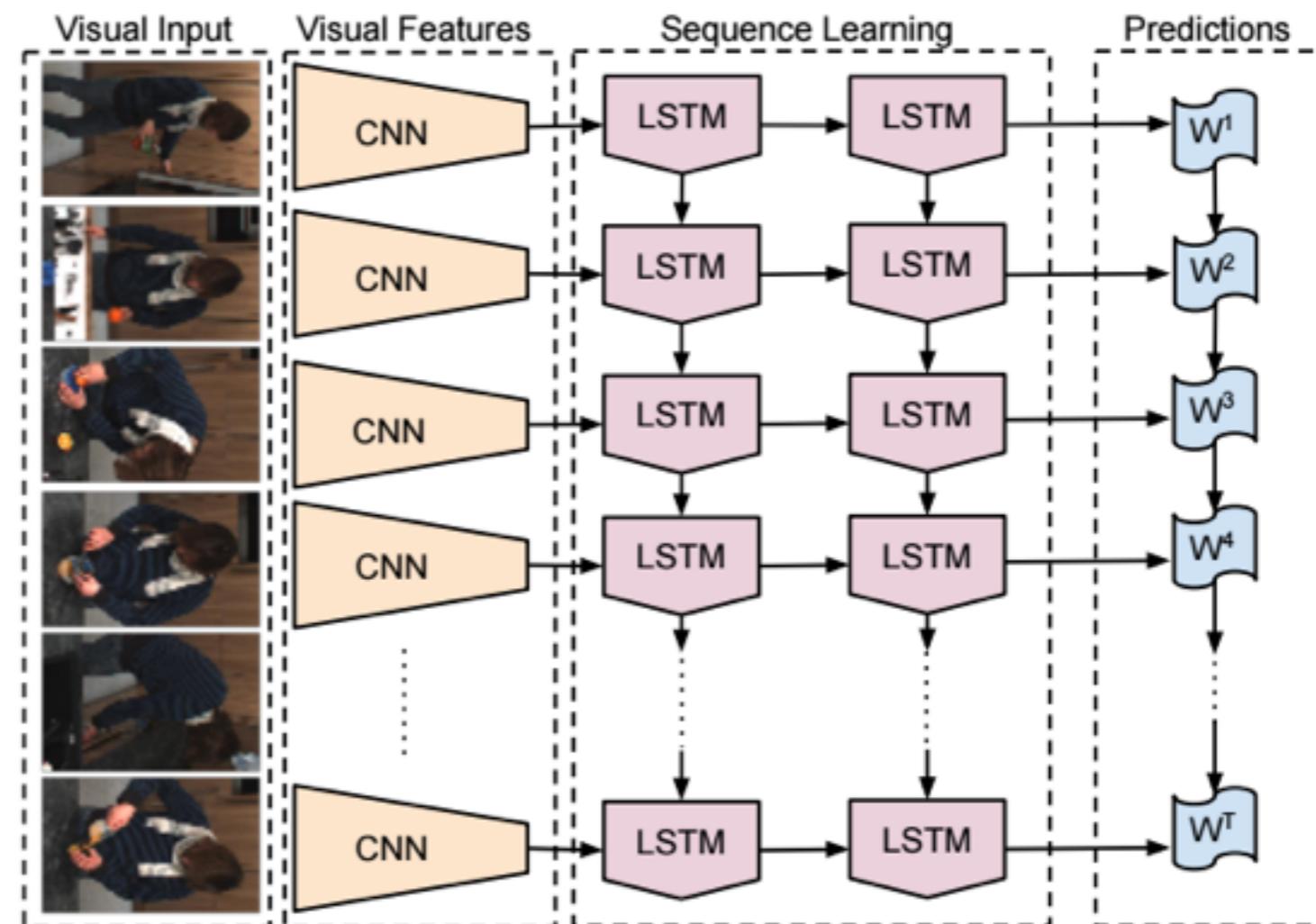
# WaveNet

- Using stack of diluted layers
- To generate next sample, it models conditional probability given previous samples



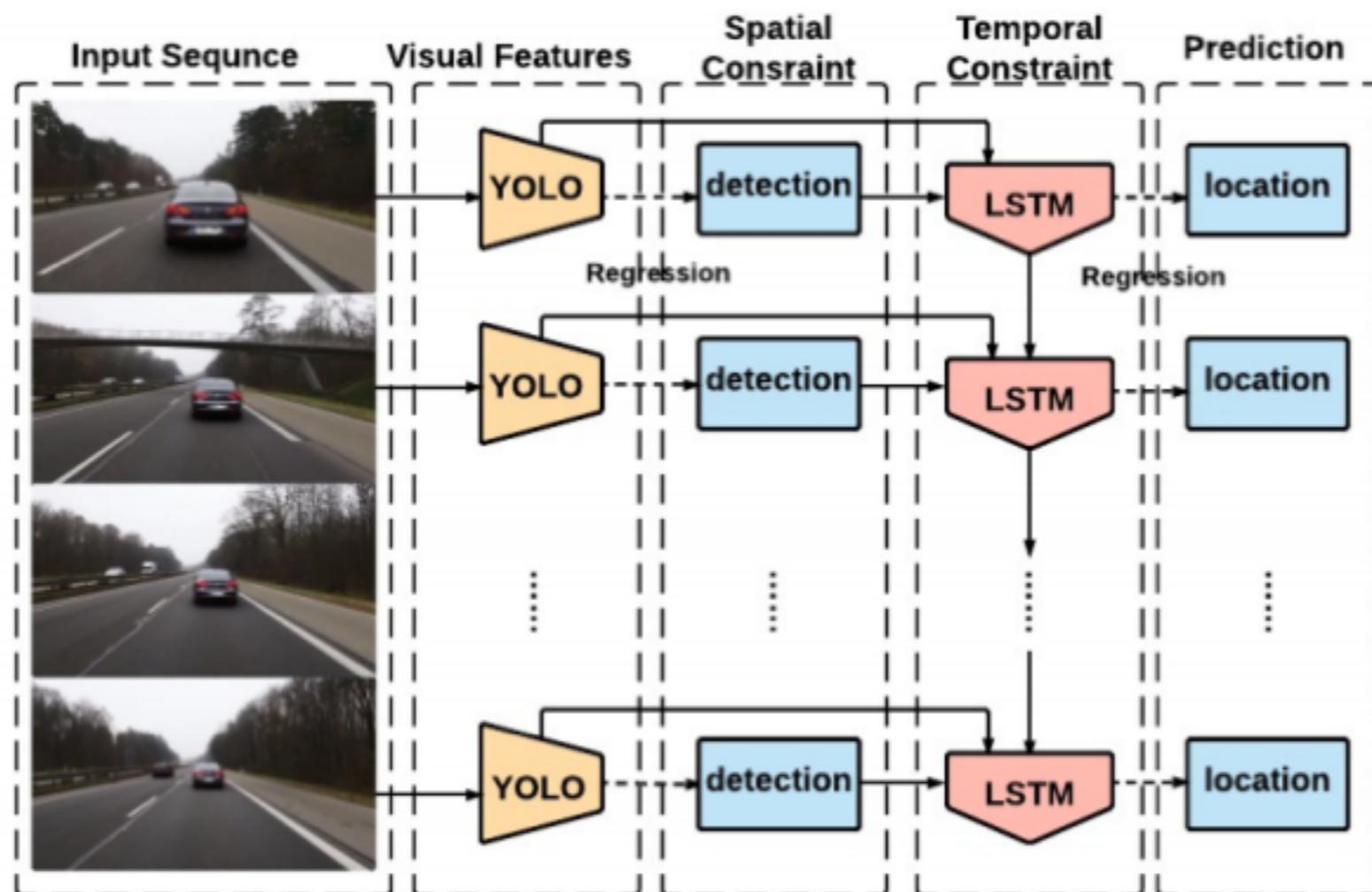
# Action Recognition

- Using LSTMs and CNN for videos
- CNN creates a feature vector that is feed into LSTM

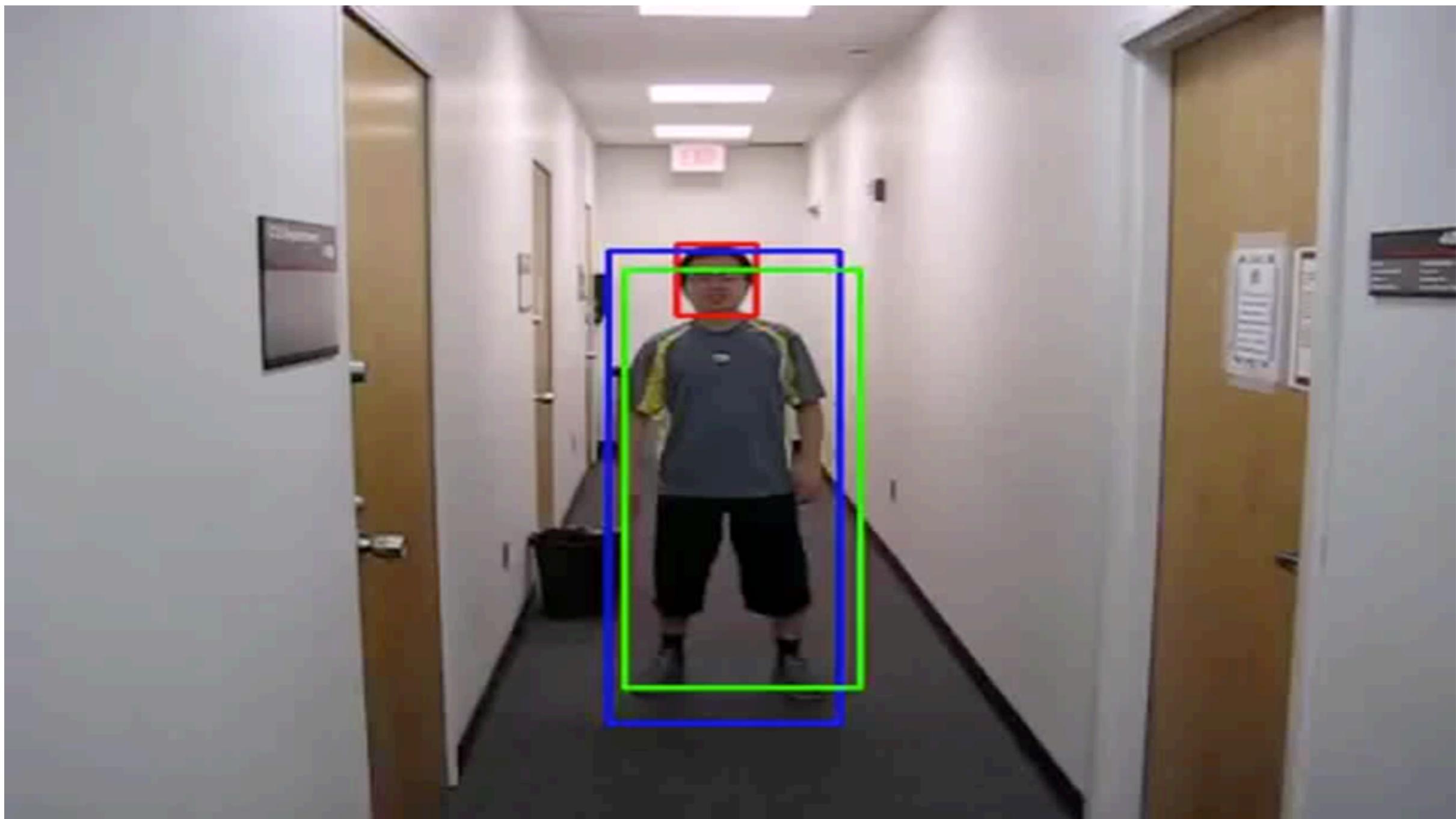


# Object Tracking

- Using an object detector with an LSTM to track objects
- Model the dynamics of video

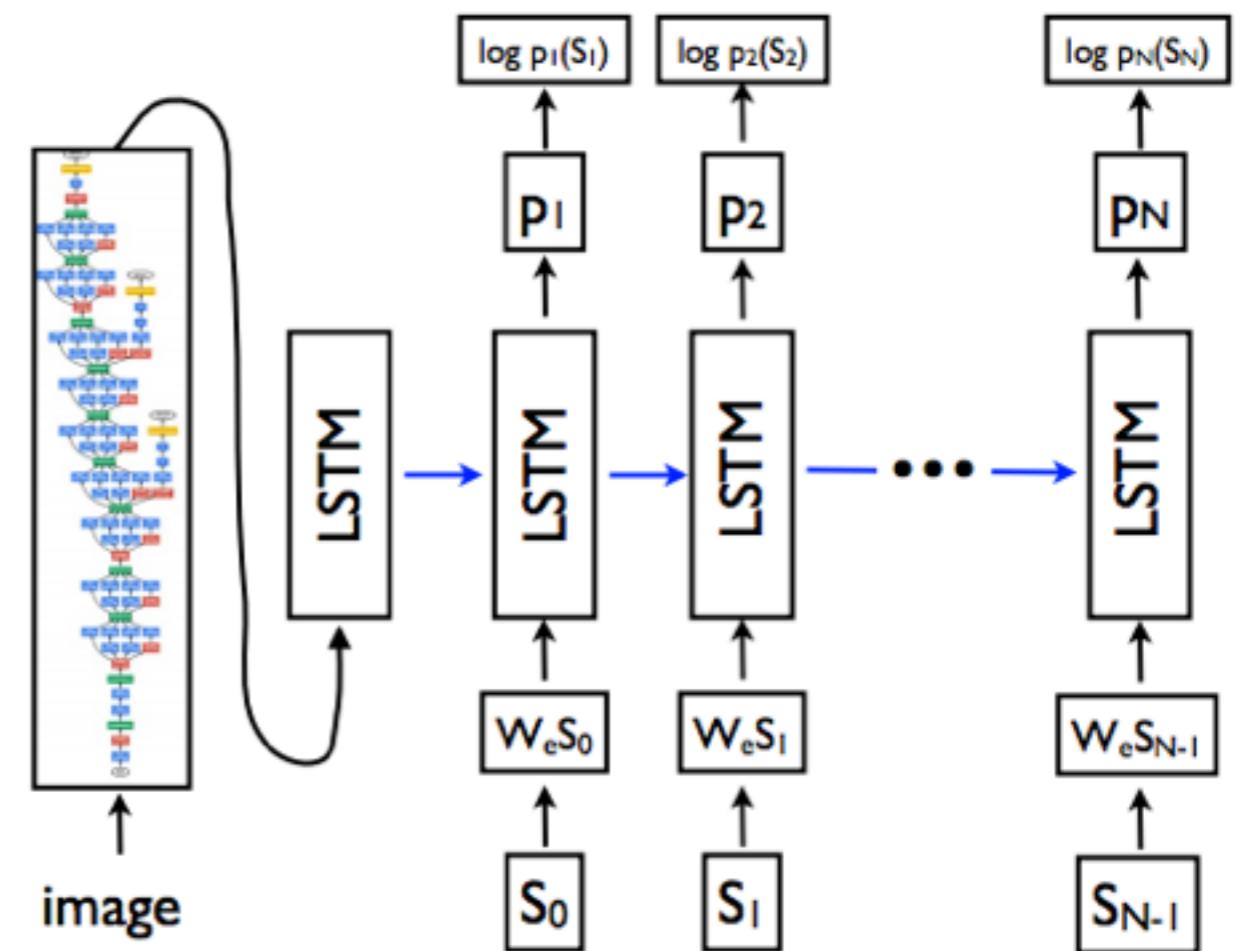


# ROLO



# Image Captioning

- Combination of CNN and LSTM to caption images
  - Using a pretrained CNN for visual features



# Image Captioning

A person riding a motorcycle on a dirt road.



Two dogs play in the grass.



A skateboarder does a trick on a ramp.



A dog is jumping to catch a frisbee.



A group of young people playing a game of frisbee.



Two hockey players are fighting over the puck.



A little girl in a pink hat is blowing bubbles.



A refrigerator filled with lots of food and drinks.



A herd of elephants walking across a dry grass field.



A close up of a cat laying on a couch.



A red motorcycle parked on the side of the road.



A yellow school bus parked in a parking lot.



Describes without errors

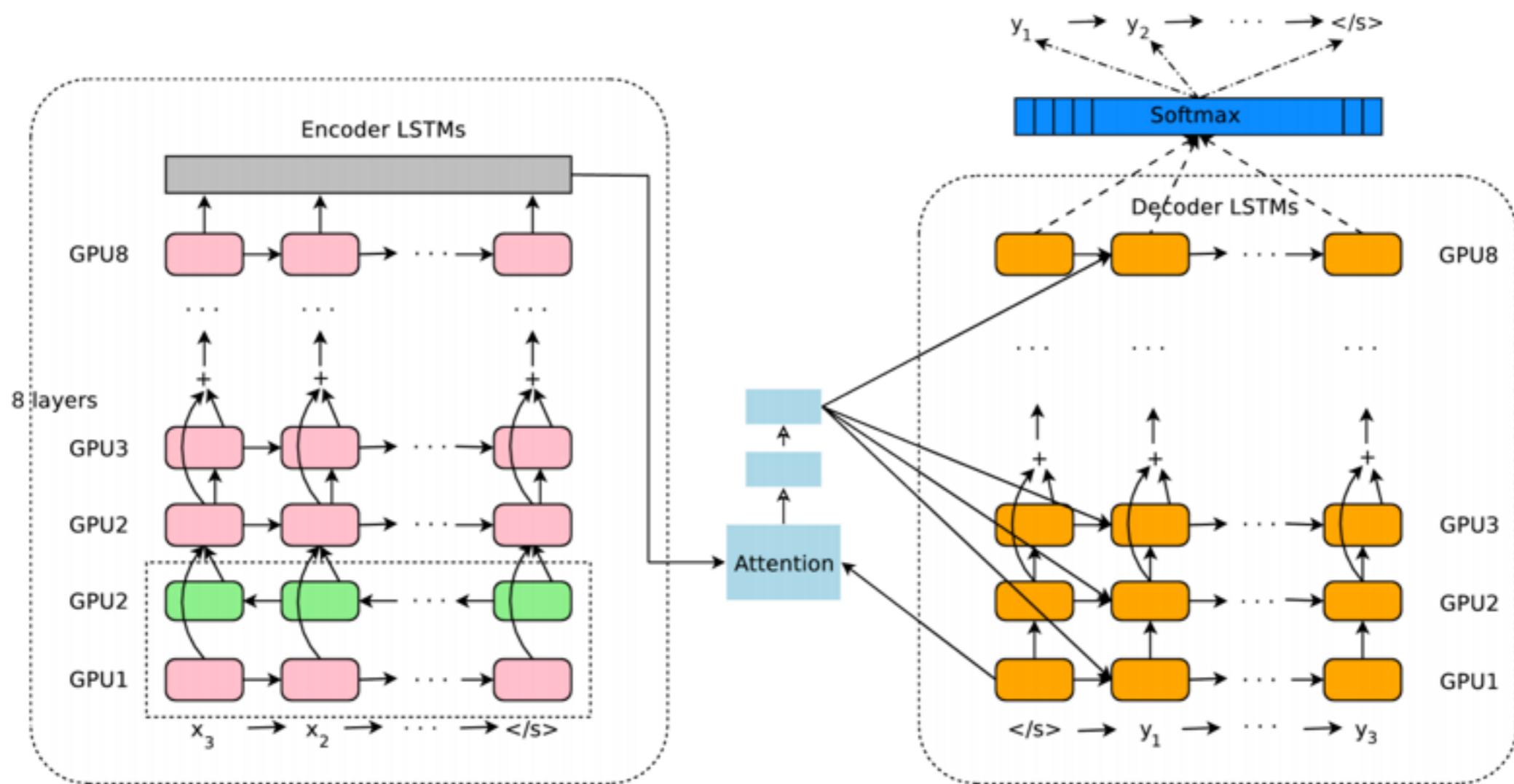
Describes with minor errors

Somewhat related to the image

Unrelated to the image

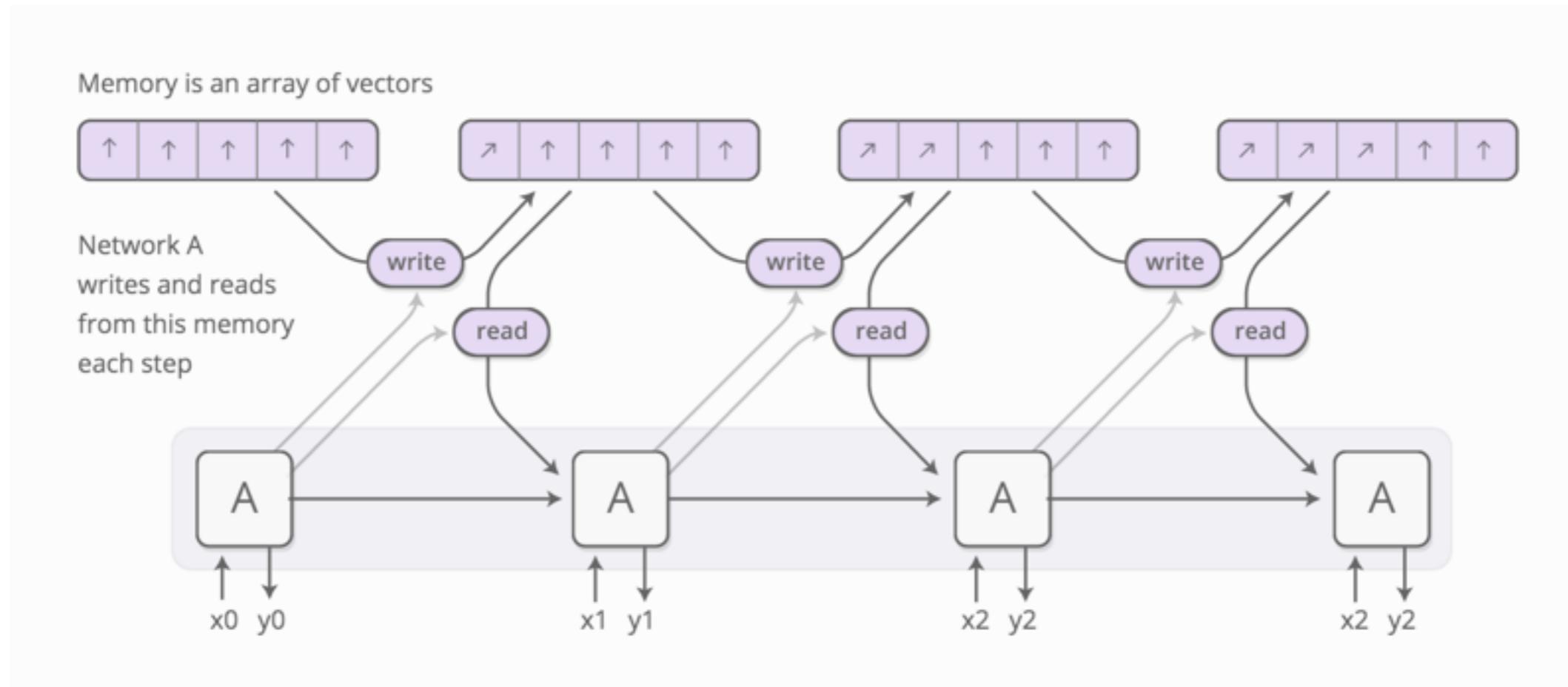
# Google's Neural Machine Translation System

- Encoder and Decoder LSTMs
- Attention model



# Neural Turing Machines

- LSTM with external memory
- Analogous to a Turing Machine



# DoomBot

- Doom Competition
  - Facebook won 1st place (F1)
  - <https://www.youtube.com/watch?v=94EPSjQH38Y>

# Character-level RNN Language Models

# Goal

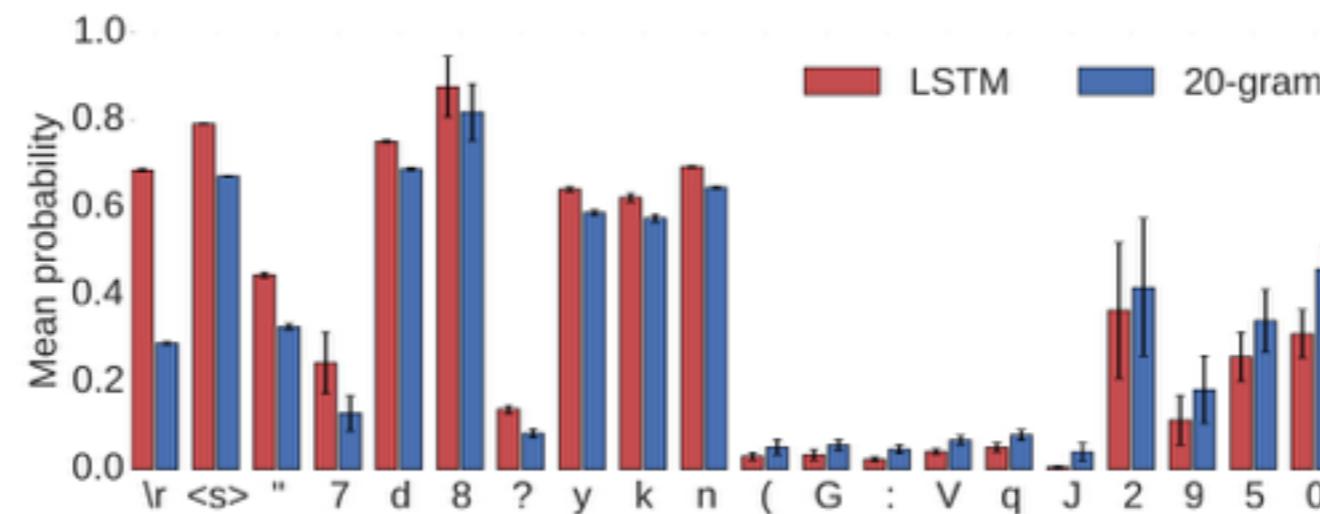
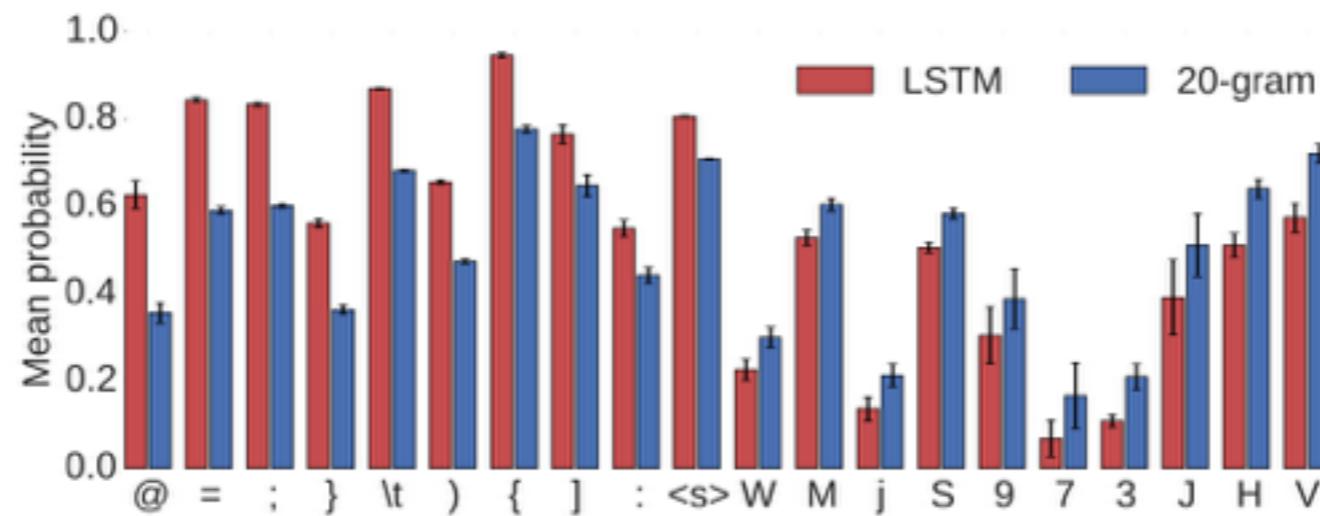
- Model the probability distribution of the next character in a sequence
  - Given the previous characters

$$P(x_t = k | x_{1:t-1}) = \frac{\exp(w_k h_t)}{\sum_{j=1}^{|V|} \exp(w_j h_t)}$$

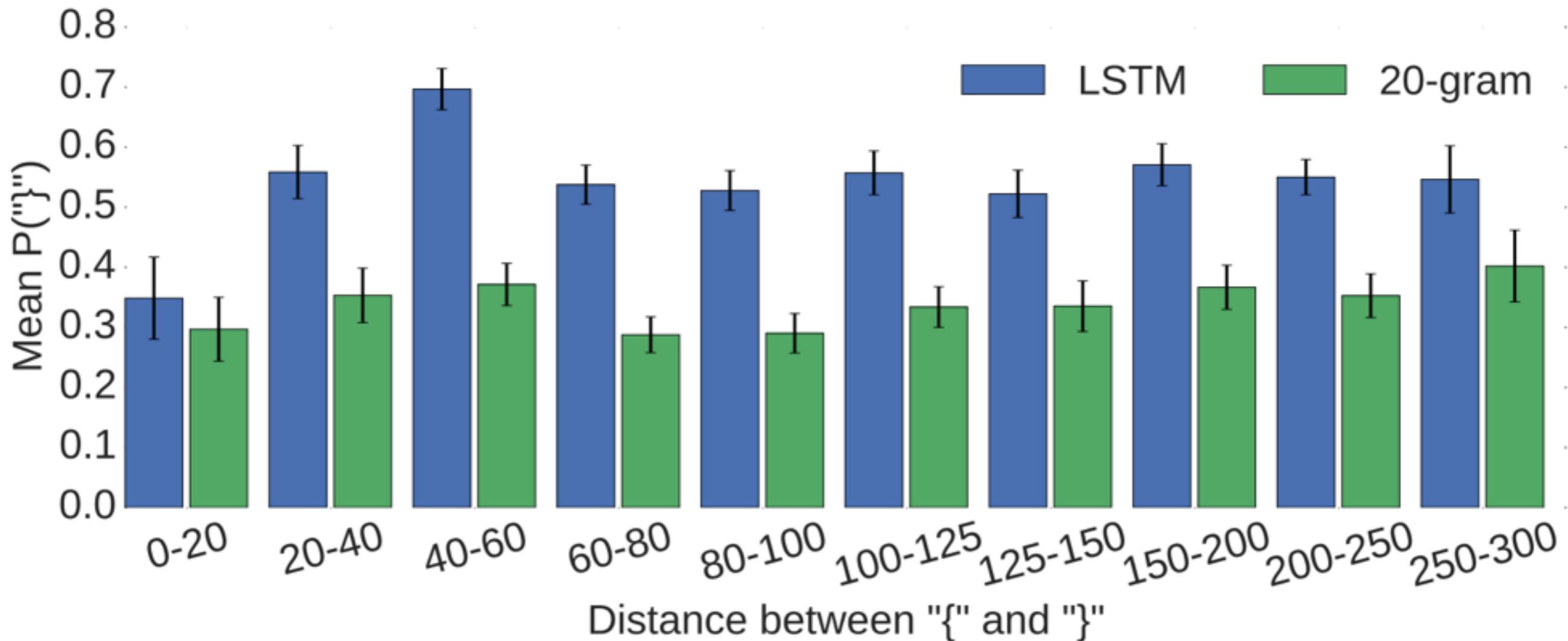
# N-grams

- Group the characters into n characters
  - n=1 unigram
  - n=2 bigram
- Useful for protein sequencing, computational linguistics, etc.

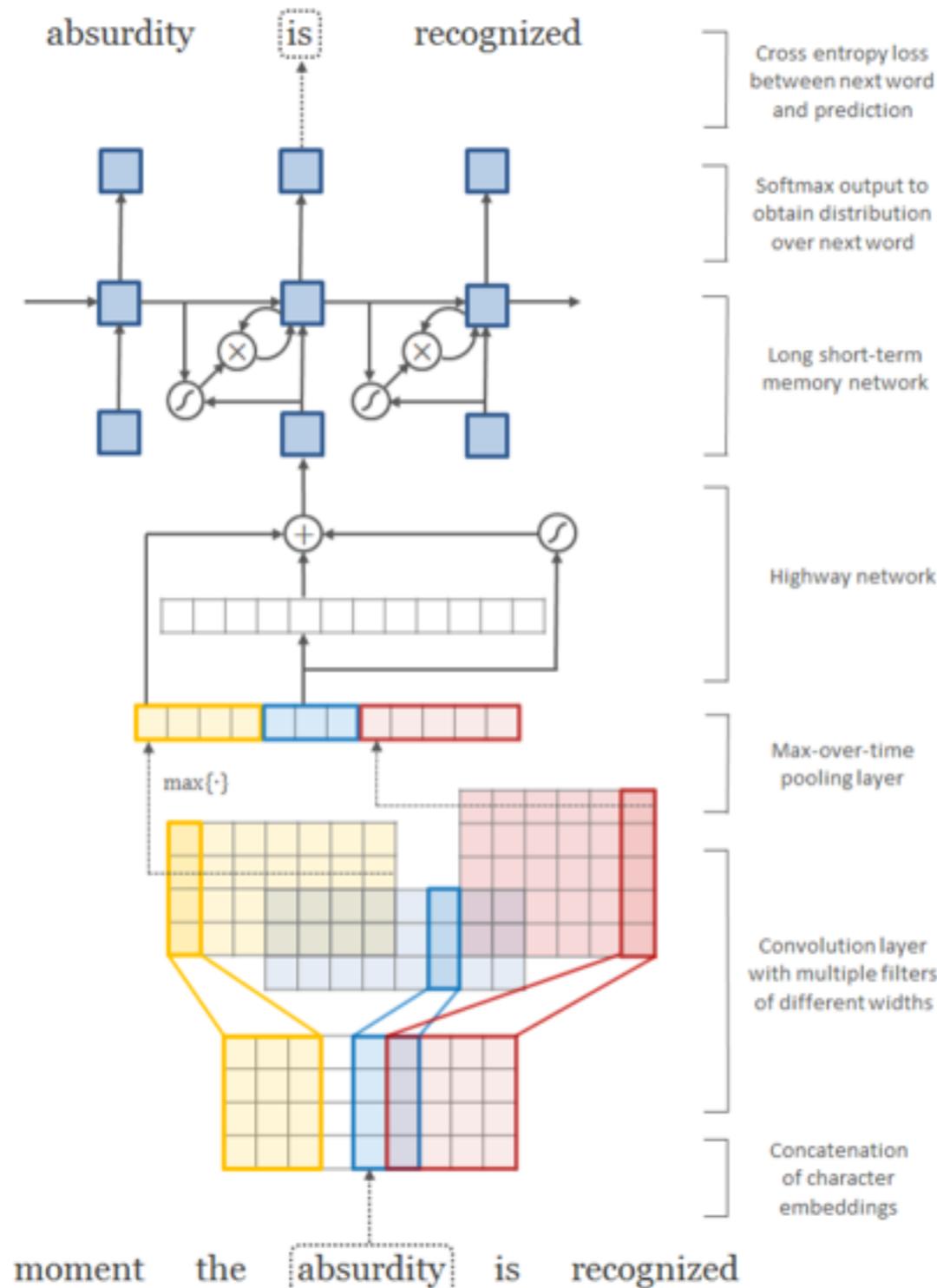
# Comparing Against N-Grams



# Remembering for Longer Durations



# Character-Aware Neural Language Models



[Kim, Jernite, Sontag, Rush]

# The Effectiveness of an RNN

```
#define REG_PG      vesa_slot_addr_pack
#define PFM_NOCOMP   AFSR(0, load)
#define STACK_DDR(type)    (func)

#define SWAP_ALLOCATE(nr)      (e)
#define emulate_sigs() arch_get_unaligned_child()
#define access_rw(TST)  asm volatile("movd %%esp, %0, %3" : : "r" (0)); \
    if (__type & DO_READ)

static void stat_PC_SEC __read_mostly offsetof(struct seq_argsqueue, \
    pC>[1]);

static void
os_prefix(unsigned long sys)
{
#ifdef CONFIG_PREEMPT
    PUT_PARAM_RAID(2, sel) = get_state_state();
    set_pid_sum((unsigned long)state, current_state_str(),
        (unsigned long)-1->lr_full; low;
}
```

# The Effectiveness of an RNN

Naturalism and decision for the majority of Arab countries' capitalide was grounded by the Irish language by [[John Clair]], [[An Imperial Japanese Revolt]], associated with Guangzham's sovereignty. His generals were the powerful ruler of the Portugal in the [[Protestant Immineners]], which could be said to be directly in Cantonese Communication, which followed a ceremony and set inspired prison, training. The emperor travelled back to [[Antioch, Perth, October 25|21]] to note, the Kingdom of Costa Rica, unsuccessful fashioned the [[Thrales]], [[Cynth's Dajoard]], known in western [[Scotland]], near Italy to the conquest of India with the conflict. Copyright was the succession of independence in the slop of Syrian influence that was a famous German movement based on a more popular servicious, non-doctrinal and sexual power post. Many governments recognize the military housing of the [[Civil Liberalization and Infantry Resolution 265 National Party in Hungary]], that is sympathetic to be to the [[Punjab Resolution]] (PJS)[<http://www.humah.yahoo.com/guardian.cfm/7754800786d17551963s89.htm>] Official economics Adjoint for the Nazism, Montgomery was swear to advance to the resources for those Socialism's rule, was starting to signing a major tripad of aid exile.]

# The Effectiveness of an RNN

*Proof.* Omitted.  $\square$

**Lemma 0.1.** Let  $\mathcal{C}$  be a set of the construction.

Let  $\mathcal{C}$  be a gerber covering. Let  $\mathcal{F}$  be a quasi-coherent sheaves of  $\mathcal{O}$ -modules. We have to show that

$$\mathcal{O}_{\mathcal{O}_X} = \mathcal{O}_X(\mathcal{L})$$

*Proof.* This is an algebraic space with the composition of sheaves  $\mathcal{F}$  on  $X_{\text{étale}}$  we have

$$\mathcal{O}_X(\mathcal{F}) = \{\text{morph}_1 \times_{\mathcal{O}_X} (\mathcal{G}, \mathcal{F})\}$$

where  $\mathcal{G}$  defines an isomorphism  $\mathcal{F} \rightarrow \mathcal{F}$  of  $\mathcal{O}$ -modules.  $\square$

**Lemma 0.2.** This is an integer  $\mathcal{Z}$  is injective.

*Proof.* See Spaces, Lemma ???.  $\square$

**Lemma 0.3.** Let  $S$  be a scheme. Let  $X$  be a scheme and  $X$  is an affine open covering. Let  $\mathcal{U} \subset X$  be a canonical and locally of finite type. Let  $X$  be a scheme. Let  $X$  be a scheme which is equal to the formal complex.

The following to the construction of the lemma follows.

Let  $X$  be a scheme. Let  $X$  be a scheme covering. Let

$$b : X \rightarrow Y' \rightarrow Y \rightarrow Y \rightarrow Y' \times_X Y \rightarrow X,$$

be a morphism of algebraic spaces over  $S$  and  $Y$ .

*Proof.* Let  $X$  be a nonzero scheme of  $X$ . Let  $X$  be an algebraic space. Let  $\mathcal{F}$  be a quasi-coherent sheaf of  $\mathcal{O}_X$ -modules. The following are equivalent

- (1)  $\mathcal{F}$  is an algebraic space over  $S$ .
- (2) If  $X$  is an affine open covering.

Consider a common structure on  $X$  and  $X$  the functor  $\mathcal{O}_X(U)$  which is locally of finite type.  $\square$

This since  $\mathcal{F} \in \mathcal{F}$  and  $x \in \mathcal{G}$  the diagram

$$\begin{array}{ccccc}
 S & \longrightarrow & & & \\
 \downarrow & & & & \\
 \xi & \longrightarrow & \mathcal{O}_{X'} & \nearrow & \\
 \text{gor}_x & & & & \\
 & & & & \\
 & & = \alpha' \longrightarrow & & \\
 & & \downarrow & & \\
 & & = \alpha' \longrightarrow \alpha & & \\
 & & \uparrow & & \\
 \text{Spec}(K_\phi) & & & & X \\
 & & \text{Mor}_{\text{sets}} & & \downarrow \\
 & & & & \text{d}(\mathcal{O}_{X_{\text{étal}}}, \mathcal{G})
 \end{array}$$

is a limit. Then  $\mathcal{G}$  is a finite type and assume  $S$  is a flat and  $\mathcal{F}$  and  $\mathcal{G}$  is a finite type  $f_*$ . This is of finite type diagrams, and

- the composition of  $\mathcal{G}$  is a regular sequence,
- $\mathcal{O}_{X'}$  is a sheaf of rings.

$\square$

*Proof.* We have see that  $X = \text{Spec}(R)$  and  $\mathcal{F}$  is a finite type representable by algebraic space. The property  $\mathcal{F}$  is a finite morphism of algebraic stacks. Then the cohomology of  $X$  is an open neighbourhood of  $U$ .  $\square$

*Proof.* This is clear that  $\mathcal{G}$  is a finite presentation, see Lemmas ???.

A reduced above we conclude that  $U$  is an open covering of  $\mathcal{C}$ . The functor  $\mathcal{F}$  is a field

$$\mathcal{O}_{X, \pi} \rightarrow \mathcal{F}_\pi \rightarrow \mathcal{O}_{X, \pi}^{-1} \mathcal{O}_{X, \lambda}(\mathcal{O}_{X, \lambda}^\vee)$$

is an isomorphism of covering of  $\mathcal{O}_{X, \lambda}$ . If  $\mathcal{F}$  is the unique element of  $\mathcal{F}$  such that  $X$  is an isomorphism.

The property  $\mathcal{F}$  is a disjoint union of Proposition ?? and we can filtered set of presentations of a scheme  $\mathcal{O}_X$ -algebra with  $\mathcal{F}$  are opens of finite type over  $S$ .

If  $\mathcal{F}$  is a scheme theoretic image points.  $\square$

If  $\mathcal{F}$  is a finite direct sum  $\mathcal{O}_{X, \lambda}$  is a closed immersion, see Lemma ???. This is a sequence of  $\mathcal{F}$  is a similar morphism.

# The Effectiveness of an RNN

Trained on *War & Peace*

Iteration: 100

tyntd-iafhatawiaoahrdemot lytdws e ,tfti, astai f ogoh eoase rrranbyne 'nhthnee e  
plia tkldrgd t o idoe ns,smtt h ne etie h,hregtrs nigtike,aoaenns lng

Iteration: 300

"Tmont thithey" fomesscerliund  
Keushey. Thom here  
sheulke, anmerenith ol sivh I lalterthend Bleipile shuwy fil on aseterlome  
coaniogennc Phe lism thond hon at. MeiDimorotion in ther thize."

Iteration: 2000

"Why do what that day," replied Natasha, and wishing to himself the fact the  
princess, Princess Mary was easier, fed in had oftened him.  
Pierre aking his soul came to the packs and drove up his father-in-law women.

# Visualize the Neurons of an RNN

```
static int __dequeue_signal(struct sigpending *pending, sigset_t *mask,
    siginfo_t *info)
{
    int sig = next_signal(pending, mask);
    if (sig) {
        if (current->notifier) {
            if (sigismember(current->notifier_mask, sig)) {
                if (!!(current->notifier)(current->notifier_data)) {
                    clear_thread_flag(TIF_SIGPENDING);
                    return 0;
                }
            }
        }
        collect_signal(sig, pending, info);
    }
    return sig;
}
```

# Visualize the Neurons of an RNN

Cell sensitive to position in line:

The sole importance of the crossing of the Berezina lies in the fact that it plainly and indubitably proved the fallacy of all the plans for cutting off the enemy's retreat and the soundness of the only possible line of action--the one Kutuzov and the general mass of the army demanded--namely, simply to follow the enemy up. The French crowd fled at a continually increasing speed and all its energy was directed to reaching its goal. It fled like a wounded animal and it was impossible to block its path. This was shown not so much by the arrangements it made for crossing as by what took place at the bridges. When the bridges broke down, unarmed soldiers, people from Moscow and women with children who were with the French transport, all--carried on by vis inertiae--pressed forward into boats and into the ice-covered water and did not, surrender.

Cell that turns on inside quotes:

"You mean to imply that I have nothing to eat out of.... On the contrary, I can supply you with everything even if you want to give dinner parties," warmly replied Chichagov, who tried by every word he spoke to prove his own rectitude and therefore imagined Kutuzov to be animated by the same desire.

Kutuzov, shrugging his shoulders, replied with his subtle penetrating smile: "I meant merely to say what I said."

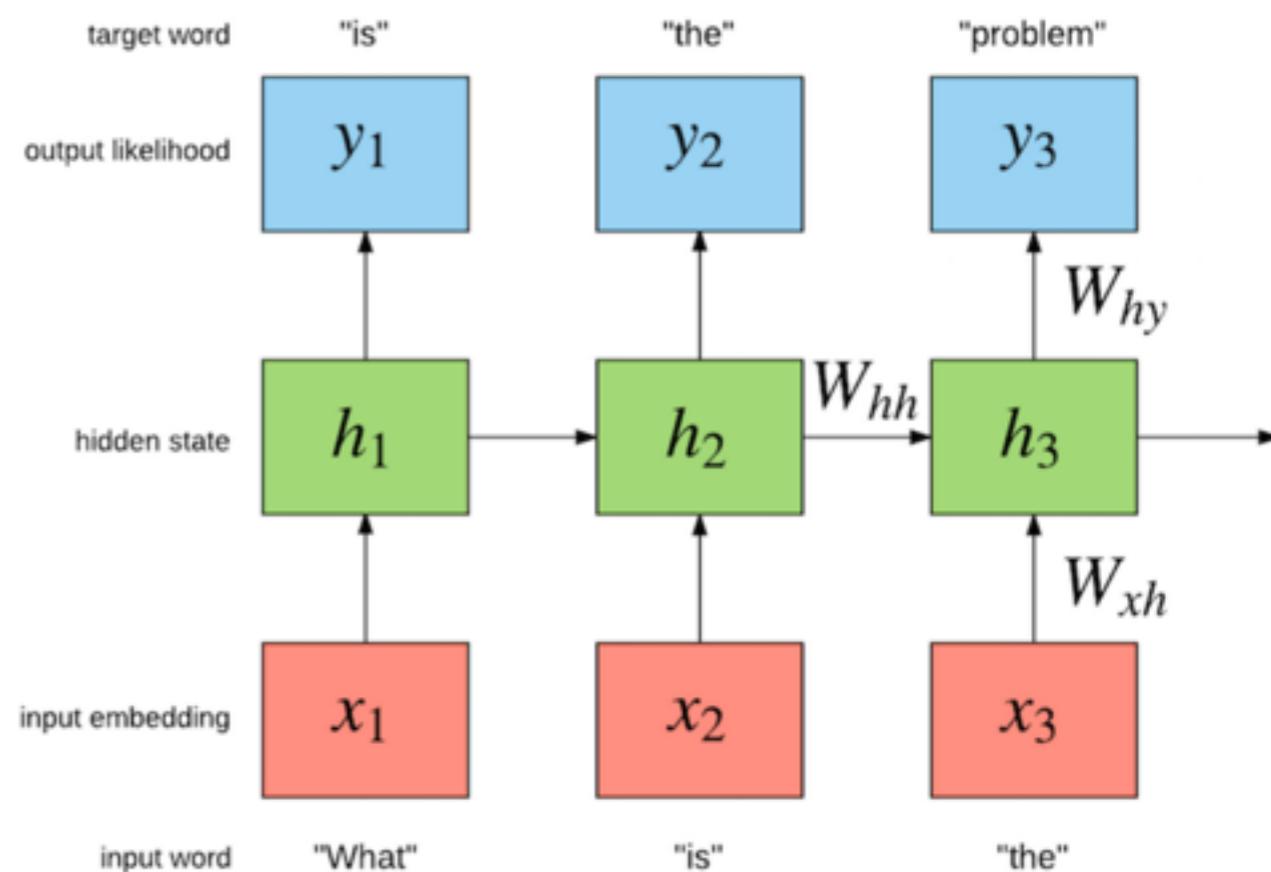
# Important Theoretical Questions

- How do LSTMs encode the expression tree for the code that they generate in their hidden states?
- With the programming languages with well specified grammar, does the LSTM implicitly learn the full generative grammar?
- Can we train LSTMs to generate valid programs that solve tasks?

# Word-level RNN Language Models

# Motivation

- Model the probability distribution of the next word in a sequence, given the previous words
- Words are the minimal unit to provide meaning
- Another step to a hierarchical model

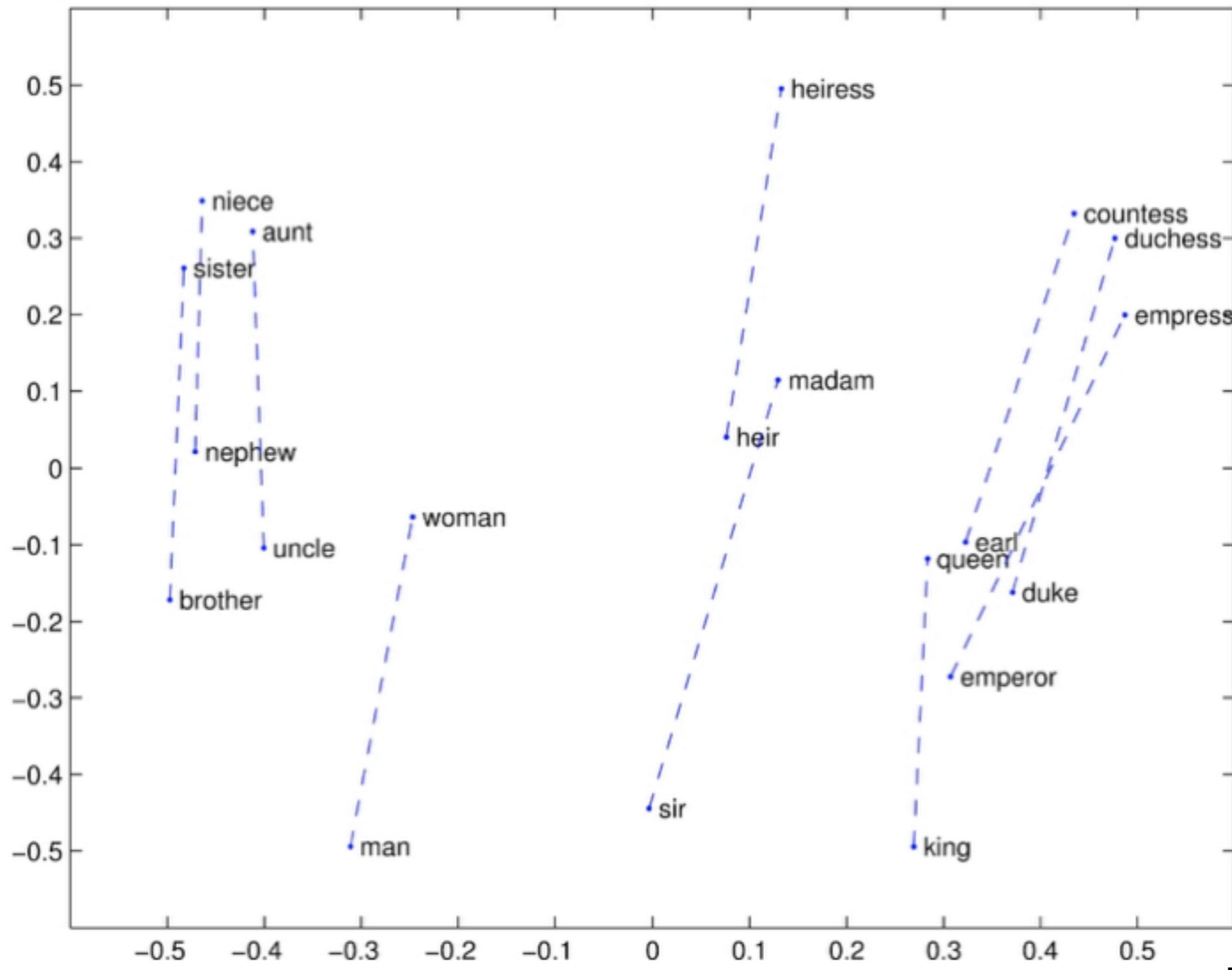


# Global Vectors for Word Representation (GloVe)

- Provide semantic information/context for words
- Unsupervised method for learning word representations

$$J(\theta) = \frac{1}{2} \sum_{i,j=1}^W f(P_{ij})(u_i^T v_j - \log P_{ij})^2$$

# Glove Visualization



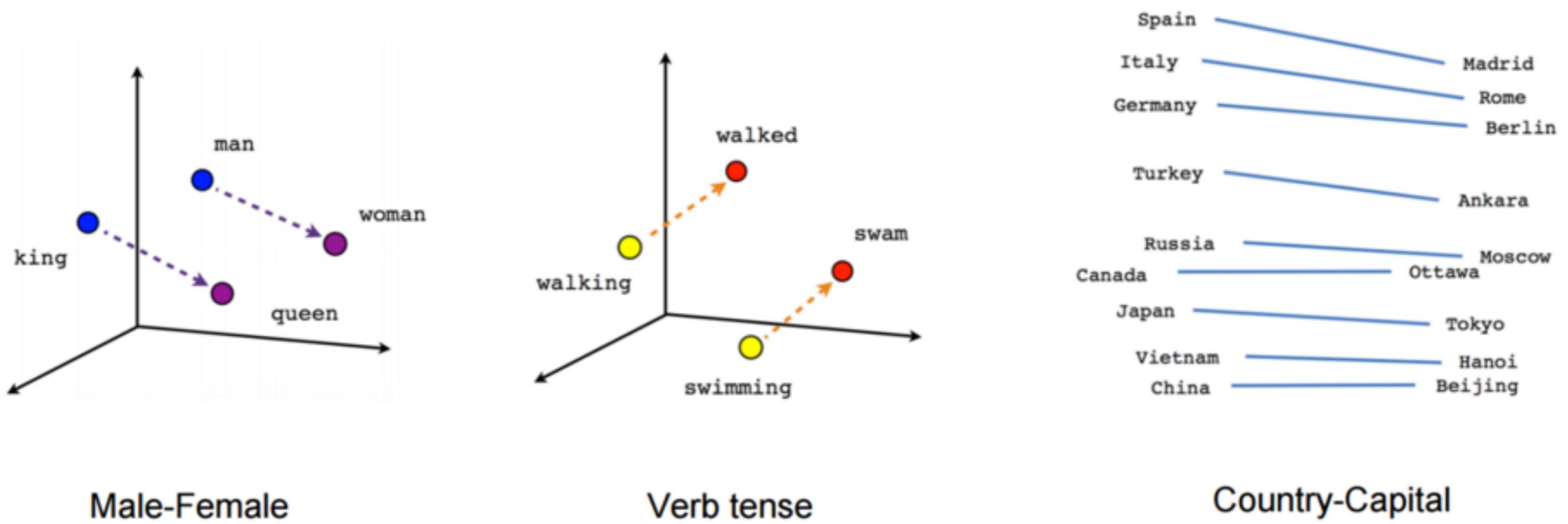
[Richard Socher]

# Word2Vec

- Learn word embeddings
- Shallow, two-layer neural network
  - Trained to reconstruct linguistic context between words
  - Produces a vector space for the words

$$= \arg \max_{\theta} \sum_{(w,c) \in D} \log \sigma(v_c \cdot v_w) + \sum_{(w,c) \in D'} \log \sigma(-v_c \cdot v_w)$$

# Word2Vec Visualization



[Tensorflow]

# Question Time

- What are the difference(s) between word2vec and GloVe?

# Word2vec with RNNs

Method	Adjectives	Nouns	Verbs	All
LSA-80	9.2	11.1	17.4	12.8
LSA-320	11.3	18.1	20.7	16.5
LSA-640	9.6	10.1	13.8	11.3
RNN-80	9.3	5.2	30.4	16.2
RNN-320	18.2	19.0	45.0	28.5
RNN-640	21.0	25.2	54.8	34.7
<b>RNN-1600</b>	<b>23.9</b>	<b>29.2</b>	<b>62.2</b>	<b>39.6</b>

Table 2: Results for identifying syntactic regularities for different word representations. Percent correct.

# Word RNN trained on Shakespeare

LEONTES:

Why, my Irish time?  
And argue in the lord; the man mad, must be deserved a spirit as drown the warlike Pray him, how seven in.

KING would be made that, methoughts I may married a Lord dishonour  
Than thou that be mine kites and sinew for his honour  
In reason prettily the sudden night upon all shalt bid him thus again. times than one from mine unaccustom'

LARTIUS:

O,'tis aediles, fight!  
Farewell, it himself have saw.

SLY:

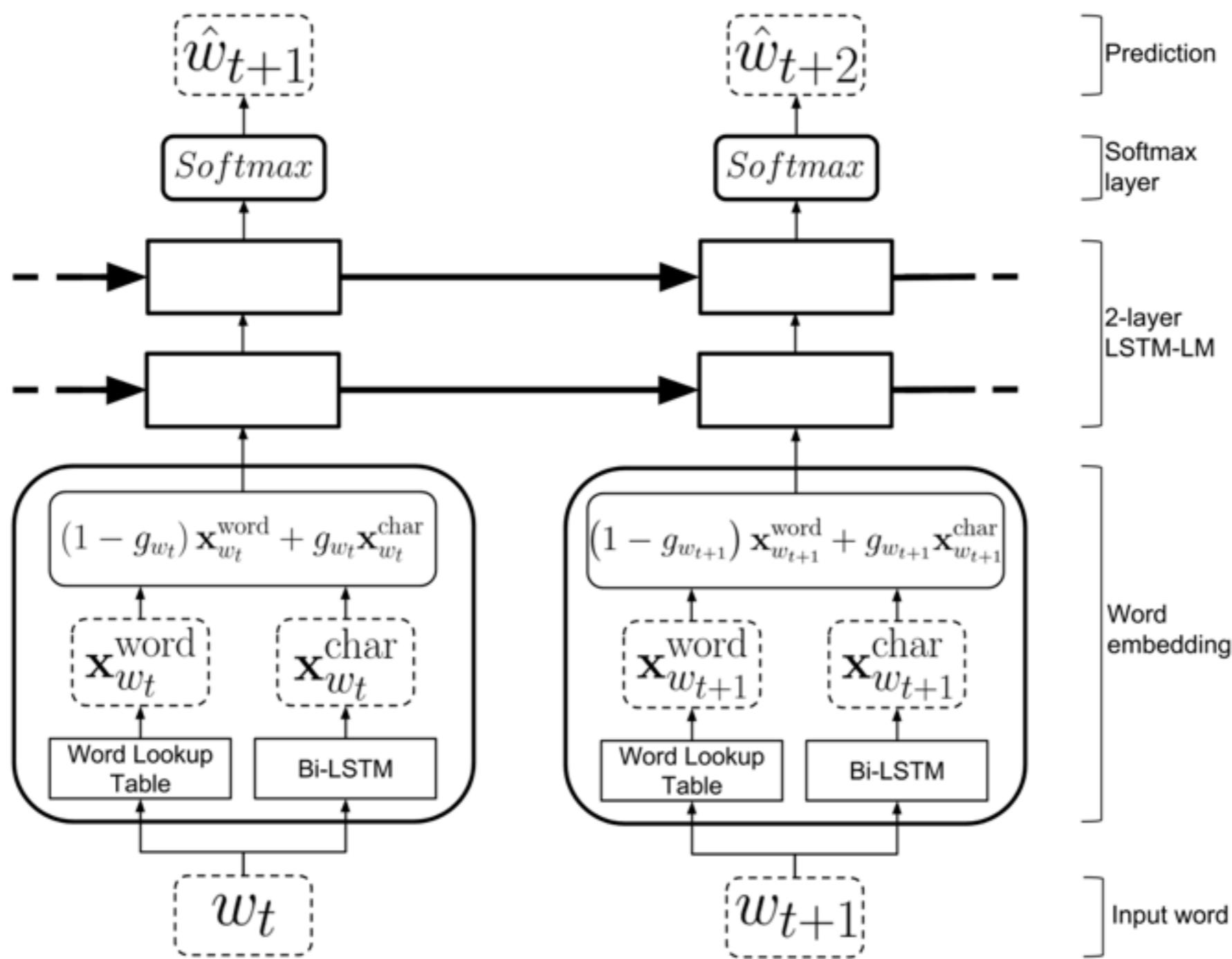
Now gods have their VINCENTIO:  
Whipt fearing but first I know you you, hinder truths.

ANGELO:

This are entitle up my dearest state but deliver'd.

DUKE look dissolved: seemeth brands  
That He being and  
full of toad, they knew me to joy.

# Gated Word RNN

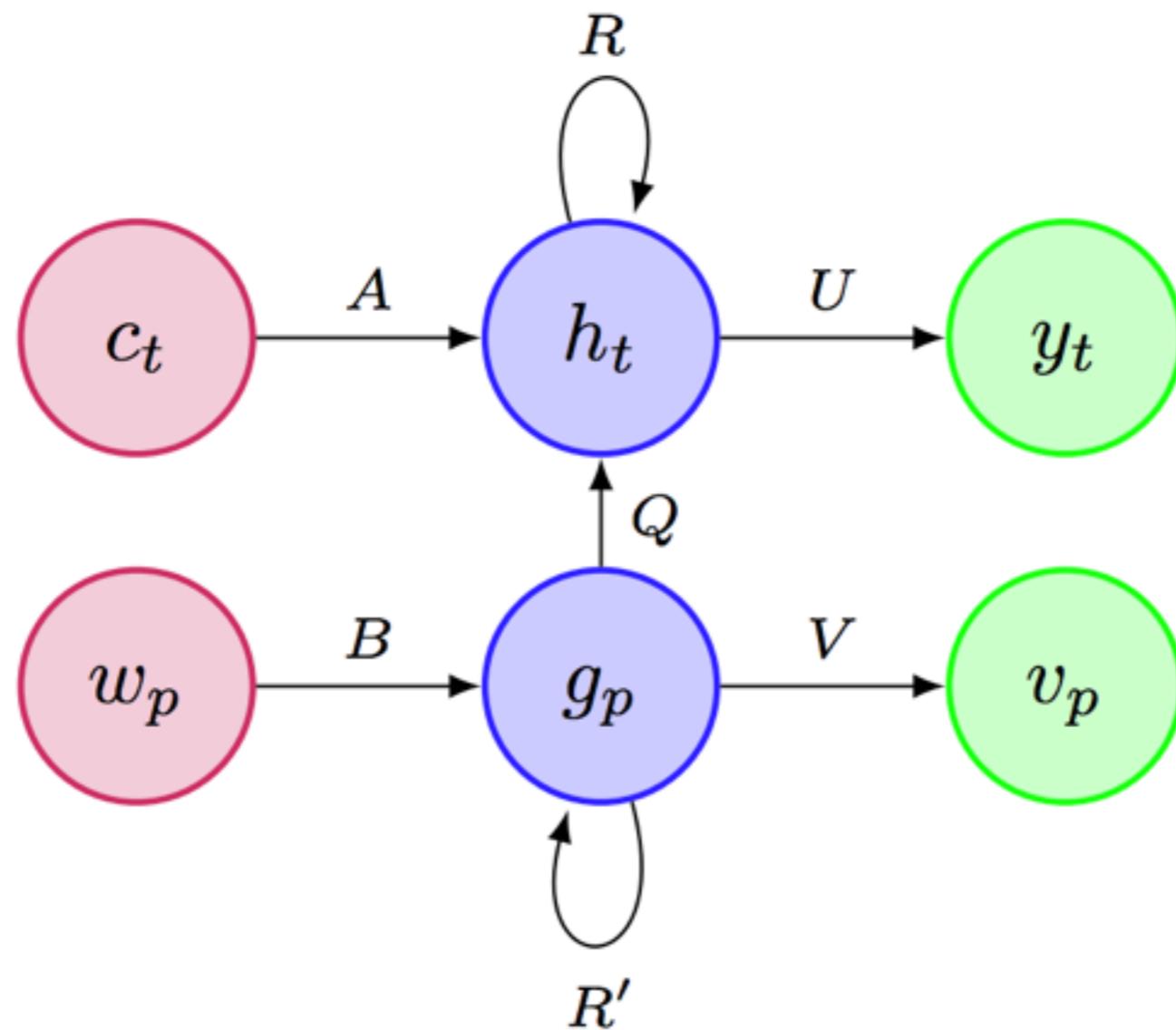


# Gated Word RNN Results

Model	PTB		BBC		IMDB	
	Validation	Test	Validation	Test	Validation	Test
Gated Word & Char, adaptive	117.49	113.87	<b>78.56</b>	<b>87.16</b>	71.99	72.29
Gated Word & Char, adaptive (Pre-train)	117.03	112.90	80.37	87.51	71.16	71.49
Gated Word & Char, $g = 0.25$	119.45	115.55	79.67	88.04	71.81	72.14
Gated Word & Char, $g = 0.25$ (Pre-train)	<b>117.01</b>	<b>113.52</b>	80.07	87.99	<b>70.60</b>	<b>70.87</b>
Gated Word & Char, $g = 0.5$	126.01	121.99	89.27	94.91	106.78	107.33
Gated Word & Char, $g = 0.5$ (Pre-train)	117.54	113.03	82.09	88.61	109.69	110.28
Gated Word & Char, $g = 0.75$	135.58	135.00	105.54	111.47	115.58	116.02
Gated Word & Char, $g = 0.75$ (Pre-train)	179.69	172.85	132.96	136.01	106.31	106.86
Word Only	118.03	115.65	84.47	90.90	72.42	72.75
Character Only	132.45	126.80	88.03	97.71	98.10	98.59
Word & Character	125.05	121.09	88.77	95.44	77.94	78.29
Word & Character (Pre-train)	122.31	118.85	84.27	91.24	80.60	81.01
Non-regularized LSTM (Zaremba, 2014)	120.7	114.5	-	-	-	-

**Table 1:** Validation and test perplexities on Penn Treebank (PTB), BBC, IMDB Movie Reviews datasets.

# Combining Character & Word Level



# Question Time

- In which situation(s) can you see character-level RNN more suitable than a word-level RNN?

# Generating Movie Scripts

- LSTM named Benjamin
  - Learned to predict which letters would follow, then the words and phrases
- Trained on corpus of past 1980 and 1990 sci-fi movie scripts
- "I'll give them top marks if they promise never to do this again."
- <https://www.youtube.com/watch?v=LY7x2lhqjmc>

# Character vs Word Level Models

# Character vs Word-Level Models

	EN-Wikipedia				EN-WSJ			
	Acc.	P	R	$F_1$	Acc.	P	R	$F_1$
Word-based Approach								
LM ( $N = 3$ )	94.94	89.34	84.61	86.91	95.59	91.56	78.79	84.70
LM ( $N = 5$ )	94.93	89.42	84.41	86.84	95.62	91.72	78.79	84.77
CRF-WORD	96.60	94.96	87.16	<u>90.89</u>	97.64	93.12	90.41	<u>91.75</u>
Chelba and Acero (2006)	n/a				97.10	-	-	-
Character-based Approach								
CRF-CHAR	96.99	94.60	89.27	91.86	97.00	94.17	84.46	89.05
LSTM-SMALL	96.95	93.05	90.59	91.80	97.83	93.99	90.92	<b>92.43</b>
LSTM-LARGE	97.41	93.72	92.67	<b>93.19</b>	97.72	93.41	90.56	91.96
GRU-SMALL	96.46	92.10	89.10	90.58	97.36	92.28	88.60	90.40
GRU-LARGE	96.95	92.75	90.93	91.83	97.27	90.86	90.20	90.52

# Word Representations of Character & Word Models

	In Vocabulary					Out-of-Vocabulary		
	<i>while</i>	<i>his</i>	<i>you</i>	<i>richard</i>	<i>trading</i>	<i>computer-aided</i>	<i>misinformed</i>	<i>loooooook</i>
LSTM-Word	<i>although</i>	<i>your</i>	<i>conservatives</i>	<i>jonathan</i>	<i>advertised</i>	—	—	—
	<i>letting</i>	<i>her</i>	<i>we</i>	<i>robert</i>	<i>advertising</i>	—	—	—
	<i>though</i>	<i>my</i>	<i>guys</i>	<i>neil</i>	<i>turnover</i>	—	—	—
	<i>minute</i>	<i>their</i>	<i>i</i>	<i>nancy</i>	<i>turnover</i>	—	—	—
LSTM-Char (before highway)	<i>chile</i>	<i>this</i>	<i>your</i>	<i>hard</i>	<i>heading</i>	<i>computer-guided</i>	<i>informed</i>	<i>look</i>
	<i>whole</i>	<i>hhs</i>	<i>young</i>	<i>rich</i>	<i>training</i>	<i>computerized</i>	<i>performed</i>	<i>cook</i>
	<i>meanwhile</i>	<i>is</i>	<i>four</i>	<i>richer</i>	<i>reading</i>	<i>disk-drive</i>	<i>transformed</i>	<i>looks</i>
	<i>white</i>	<i>has</i>	<i>youth</i>	<i>richter</i>	<i>leading</i>	<i>computer</i>	<i>inform</i>	<i>shook</i>
LSTM-Char (after highway)	<i>meanwhile</i>	<i>hhs</i>	<i>we</i>	<i>eduard</i>	<i>trade</i>	<i>computer-guided</i>	<i>informed</i>	<i>look</i>
	<i>whole</i>	<i>this</i>	<i>your</i>	<i>gerard</i>	<i>training</i>	<i>computer-driven</i>	<i>performed</i>	<i>looks</i>
	<i>though</i>	<i>their</i>	<i>doug</i>	<i>edward</i>	<i>traded</i>	<i>computerized</i>	<i>outperformed</i>	<i>looked</i>
	<i>nevertheless</i>	<i>your</i>	<i>i</i>	<i>carl</i>	<i>trader</i>	<i>computer</i>	<i>transformed</i>	<i>looking</i>

**Table 6:** Nearest neighbor words (based on cosine similarity) of word representations from the large word-level and character-level (before and after highway layers) models trained on the PTB. Last three words are OOV words, and therefore they do not have representations in the word-level model.

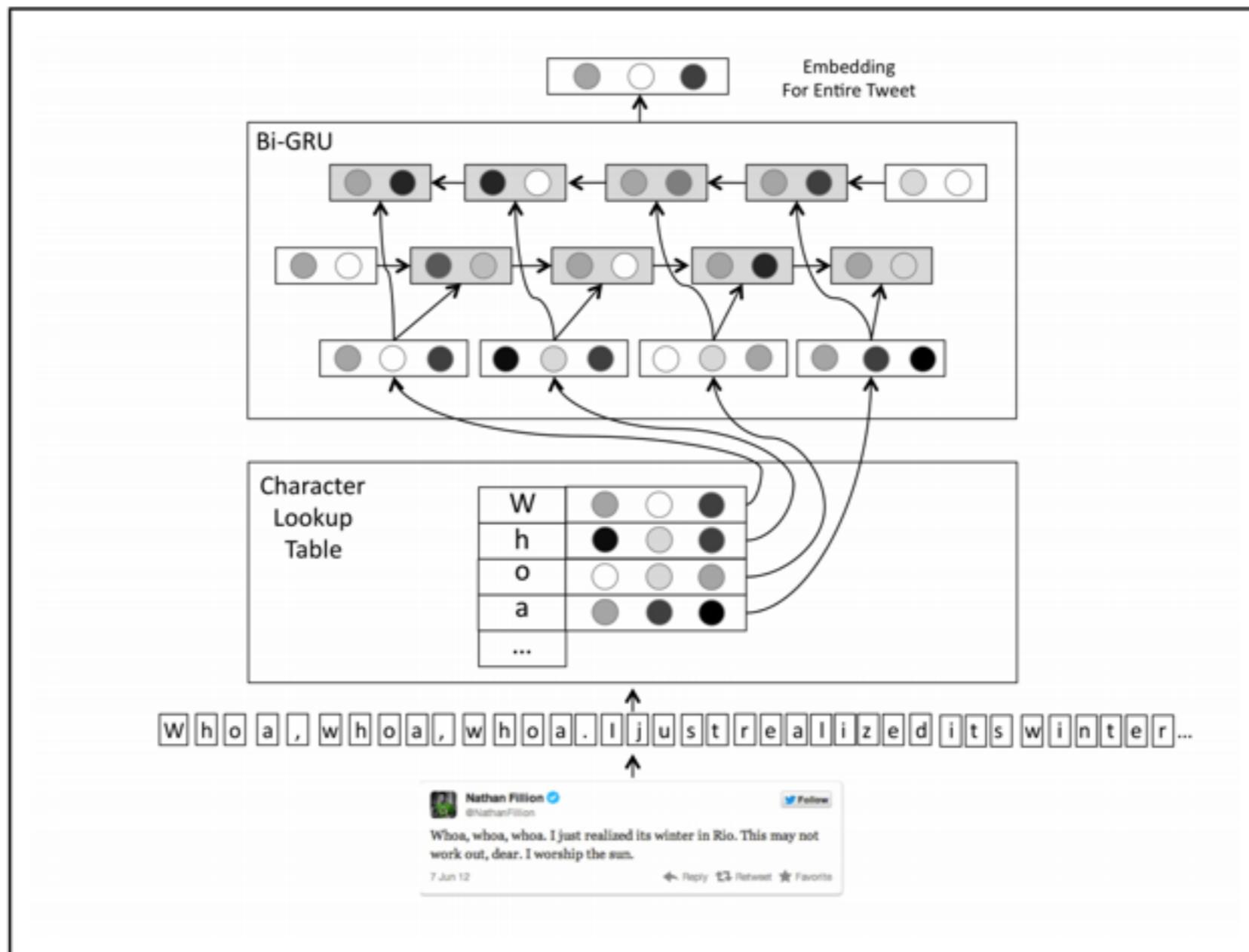
# Other Embeddings

# Tweet2Vec

$$J = \frac{1}{B} \sum_{i=1}^B \sum_{j=1}^L -t_{i,j} \log(p_{i,j}) + \lambda \|\Theta\|^2. \quad (3)$$

Here  $B$  is the batch size,  $L$  is the number of classes,  $p_{i,j}$  is the predicted probability that the  $i$ -th tweet has hashtag  $j$ , and  $t_{i,j} \in \{0, 1\}$  denotes the ground truth of whether the  $j$ -th hashtag is in the  $i$ -th tweet. We use L2-regularization weighted by  $\lambda$ .

# Tweet2Vec Encoder



[Dhingra, Zhou, Fitzpatrick, Muehl, Cohen]

# Tweet2Vec Results

Tweets	Word model baseline	<i>tweet2vec</i>
ninety-one degrees. *♥️😊	#initialsofsomeone.. #nw #gameofthrones	#summer <b>#loveit</b> #sun
self-cooked scramble egg. yum!! !url	#music #cheap #cute	<b>#yummy</b> #food #foodporn
can't sleeeeeep	#gameofthrones #heartbreaker	#tired <b>#insomnia</b>
oklahoma!!!!!!!!!! champions!!!!	#initialsofsomeone.. #nw #lrt	<b>#wcws</b> #sooners #ou
7 % of battery . iphones die too quick .	#help #power #money #s	#fml #apple #bbl <b>#thestruggle</b>
i have the cutest nephew in the world !url	#nephew <b>#cute</b> #family	#socute <b>#cute</b> #puppy

Table 1: Examples of top predictions from the models. The correct hashtag(s) if detected are in bold.

# Gene2Vec

- Word2Vec performs poorly on long nucleotide sequences
- Short sequences are very common like AAAGTT

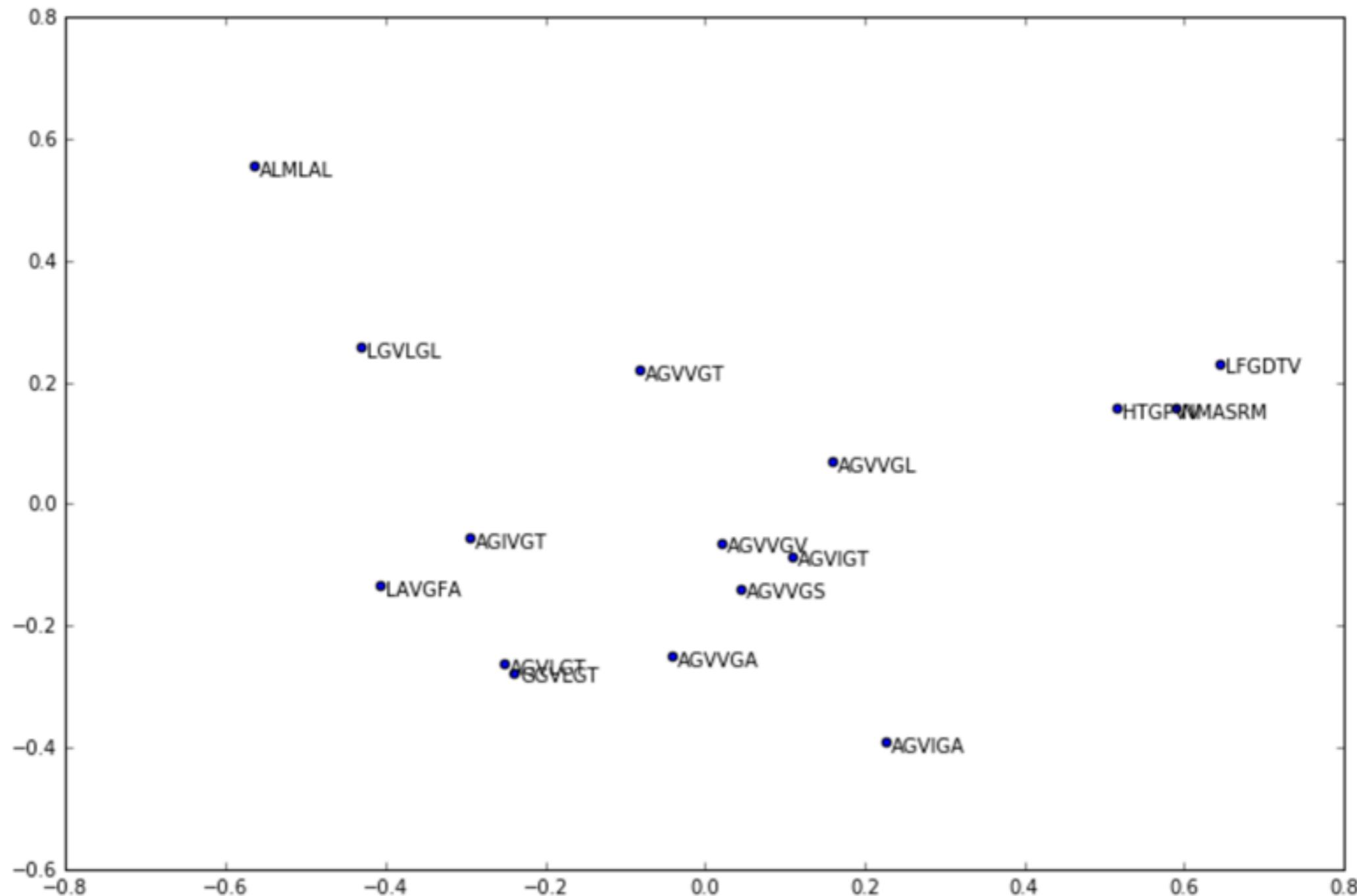
AAAAAATAGTATAAAAGTTGCCAAAAG -> *Triticum aestivum* chromosome 3B  
||||| | | ||| |

AAAAAACATGCAACAAACAGGAACCTGGC -> *Triticum aestivum* chromosome 3B  
||||||| | | | |

AAAAACAGAACATCTGTCTAAAACAGAAC -> *Triticum aestivum* chromosome 3B  
||||||| | | | | |

AAAAACAGAGACATTACTTGCACACA -> *Ovis canadensis* isolate 43U chromosome 26

# Gene2Vec Visual



Hydrophobic Amino Acids

[David Cox]

# Doc2Vec

- Similar to Word2Vec but to a larger scale
  - Sentences & Paragraphs

TARGET (72927): «this is one of the best films of this year . for a year that was fueled by controversy and crap , it was nice to finally see a film that had a true heart to it . from the opening scene to the end , i was so moved by the love that will smith has for his son . basically , if you see this movie and walk out of it feeling nothing , there is something that is very wrong with you . loved this movie , it's the perfect movie to end the year with . the best part was after the movie , my friends and i all got up and realized that this movie had actually made the four of us tear up ! it's an amazing film and if will smith doesn't get at least an oscar nom , then the oscars will just suck . in fact will smith should actually just win an oscar for this role . ! ! ! i loved this movie ! ! ! everybody needs to see especially the people in this world that take everything for granted , watch this movie , it will change you !»

SIMILAR/DISSIMILAR DOCS PER MODEL Doc2Vec(dm/m,d100,n5,w10,mc2,t8):

MOST (2046, 0.7372332215309143): «i thought this movie would be dumb , but i really liked it . people i know hate it because spirit was the only horse that talked . well , so what ? the songs were good , and the horses didn't need to talk to seem human . i wouldn't care to own the movie , and i would love to see it again . 8/10»

# Applications of Document Models

- Discovery of litigation e.g. CS Disco
- Sentiment Classification e.g. movie reviews