Memory span of recurrent neural nets

Master Thesis

Jaroslav Ištok

Overview

1. SOM
2. Recurrent SOM
3. Merge SOM
4. SRN - Elman network

SOM

SOM

- Self organizing map
- Biologicaly motivated model
- Unsupervised competitive learning
- Clustering
- Preserving topological features
- Quantization error

Find Winner

$$i^* = argmin_i ||x - w_i||$$

Rule for update weights

$$w_i(t+1) = w_i(t) + \alpha(t)h(i^*, i)([x(t) - w_i(t)]$$

Recurrent SOM

Recurrent SOM

Self organizing map with context from previous steps

- RecSom - context is copy of whole map from previous step
- Many attributes

Weights update

$$w_i(t+1) = w_i(t) + zh_{ik}[s(t) - w_i(t)]$$

$$c_i(t+1) = c_i(t) + zh_{ik}[y(t-1) - c_i(t)]$$

$$y_i = exp(-d_i)$$

Distance

$$d_i(t) = \alpha||x(t) - w_i||^2 + b||r(t) - c_i||^2$$

Recursive context

$$r(t) = [y_i(t-1), ..., y_N(t-1)]$$

Merge SOM

Merge SOM

- In Merge SOM **context is not** copy of whole map from previous step
- Fewer parameters than RecSOM
- γ_1 γ_2 - learning rates
- h_σ - Gaussian shaped function
- d_N - neighborhood function

Weights update

$$\Delta w_i = \gamma_1 \cdot h_\sigma(d_N(i, I_t)) \cdot (x^t - w^i)$$

$$\Delta c_i = \gamma_2 \cdot h_\sigma(d_N(i, I_t)) \cdot (c^t - c^i)$$

Distance

$$d_i(t) = (1 - \alpha) \cdot ||x^t - w^i||^2 + \alpha \cdot ||c^t - c^i||^2$$

Recursive context

$$c^t = (1 - \beta) \cdot w^{I_{t-1}} + \beta \cdot c^{I_{t-1}}$$

SRN - Elman network

SRN - Elman network

- Elman recurrent network
- Training with backpropagation through time
- Supervised learning
- Context layer

Measuring memory span

- Words sequences
- Every neuron will have set of letters
- If this neuron is winner, it will save letter to its set of letters (receptive field)
- Neuron saves *n* previous letters in its receptive field
- We can then find longest common sub-sequence and calculate weighted average
- This will be our measure of memory span

Example of receptive field

T. Voegtlin / Neural Networks 15 (2002) 979–991 983

to	so	lo	o	ho	ie	me	me	me	ee	ne	ne	ne	it	lit	s	rs	aw	dow	tw
to	ro	co	wo	te	ite	e	e	e	lle	ine	ine	were	there		was	ss	sw	w	x
to	do	no	o	te	te	be	be	we	we	ore	re	re	re	is	was	ls		ew	ex
they	they		fo		se	se	are	see	we	here	re	are		his	es	us	rs	ds	ow
y	y	su		ev	pre	ru	pu	one	pe	pe		t	os	ins	es	ms	as	its	as
ey	my	pu	she	v	du	cou	nu	fu	ke	put	rat	that	ous	is	s	ts	ins	cons	as
ly	ly	ple	fe	ge	de	hu	ou	u	u	red	fort	ght	ns	res	s	fa	a	a	q
sh		le	e	the	de	mu	tu	bu	ed	red	ed	art	ot	is	ca	na	ba	ha	wa
sh	h	he	he	ce	ye	beg	hu	bu	ild	cond	ed	tt	att	et	t	ma	na	tha	ra
ch	wh	ple	one	whe	the	ng	ug	ag	g	nd	und	ut	rt	at	st	sc	c	ea	sa
th	h	with	e	ge	the	ng	g	ing	g	nk		but	not	at	nt	t	nc	a	ta
th	h	with	e	the	she	od	and	id	ad	k	k	int	ght	wit	et	an	an	n	wn
th	h	er	be	ve	ad	ted	id	ard	hed	el	out	t	it	it	wan	men	an	n	n
wer	er	er	der	rd	ther	had	ld	od	d	d	wl	t	int	on	when	han	then	ten	wen
ber	for	ter	rr	fr	r			tl		cl	cl	bel	all	on	in	fin	len	nn	un
or	her	tr	ar	str	dr	si	i	l	al	bl	sl	ll	all	lon	in	en	j	men	un
wer	wer	ur	r	br	z		ni	l	l	el	il	il	ul	om	ven	en	then	kn	con
pr	thr	their	ab	b	b	fi	ati	wi	hi	pi	ci	al	al	him	them	alp	up	sn	ep
ther	for	if		b	i	ri	ti	whi	di	bi	mi	al	sm	em	em		op	sp	ap
dr	af	f	of		hi	i	i	thi	li	li	ili	m	m	com	him	dl	ep	rep	p

Fig. 2. Receptive fields of a two-dimensional recursive SOM trained on English text. A receptive field is defined as the intersection of all the sequences that trigger selection of the corresponding unit. Receptive fields are displayed in natural reading order. Topographic organization is observed, principally based on most recent letters.

Questions?

References i

- Jeffrey L.Elman *Finding Structure in Time*. University of California, San Diego, 1990
- H. Ritter and T. Kohonen *Self-Organizing Semantic Maps* Helsinky University of Technology, 1982
- Thomas Voegtlin *Recursive self-organizing maps*, 2002
- Marc Strickert, Barbara Hammer *Merge SOM for temporal data* Technical University of Clausthal, 2005