Lecture 4: Backpropagation and Neural Networks (part 1)

Tuesday January 31, 2017



Announcements!

- If you are adversely affected by immigration ban, please talk to me about accommodations
- Send in paper choices by **tonight**
- Should be able to run Jupyter server on Tufts was and network machines now
 - (deep-venv)> pip install --upgrade jupyter
- hw1 deadline in two days Thurs Feb 2: Don't forget to read the course notes.
- Redo calculation of dL/dW for hinge loss



Python/Numpy of the Day

- y_pred = scores.argmax(axis=1)
- inds = np.random.choice(X.shape[0],batch_size)
 - randomly select N numbers in a range,
 - useful for subsampling
- -[:,np.newaxis]
 - reshapes matrices of size (N,) to size (N,1)



Where we are...

$$egin{aligned} s &= f(x;W) = Wx & ext{scores function} \ L_i &= \sum_{j
eq y_i} \max(0, s_j - s_{y_i} + 1) & ext{SVM loss} \ L &= rac{1}{N} \sum_{i=1}^N L_i + \sum_k W_k^2 & ext{data loss + regularization} \end{aligned}$$





Optimization





Gradient Descent

$$rac{df(x)}{dx} = \lim_{h o 0} rac{f(x+h) - f(x)}{h}$$

Numerical gradient: slow :(, approximate :(, easy to write :) **Analytic gradient**: fast :), exact :), error-prone :(

In practice: Derive analytic gradient, check your implementation with numerical gradient



Hinge Loss Gradient wrt Weights W

margin size, usually 1.0

$$w_i = \sum_{j \neq y_i} \left[\max(0, w_j^T x_i - w_{y_i}^T x_i + \Delta) \right]$$

- We want the Jacobian Matrix of all gradients
 - partial derivatives of all output dimensions by all input dimensions

For all rows of dW where the row corresponds to the GT value for that training instance, i.e. $\mathbf{j} = \mathbf{y}_i$

$$\nabla_{w_{y_i}} L_i = -\left(\sum_{j \neq y_i} \mathbb{1}(w_j^T x_i - w_{y_i}^T x_i + \Delta > 0)\right) x_i$$

$$\nabla wL = \begin{bmatrix} \nabla w_1 L_1 & \dots & \nabla w_1 L_N \\ \vdots & \nabla w_j L_i & \ddots & \vdots \\ \nabla w_k L_1 & \dots & \nabla w_k L_N \end{bmatrix}$$

For all rows of dW where $j \neq y_i$

$$\nabla_{w_j} L_i = \mathbb{1}(w_j^T x_i - w_{y_i}^T x_i + \Delta > 0)x_i$$



Softmax Loss Gradient wrt Score S

* note change of subscripts from last slide

$$\begin{array}{c|c} a_{j} = w_{j}^{T} x_{j} \\ S_{j} = \frac{e^{a_{j}}}{\sum_{k=1}^{N} e^{a_{k}}} \quad \forall j \in 1..N \\ \frac{\partial S_{i}}{\partial a_{j}} = \frac{\partial \frac{e^{a_{i}}}{\sum_{k=1}^{N} e^{a_{k}}}}{\partial a_{j}} \end{array} \quad \forall j \in 1..N \\ \begin{array}{c|c} \hline \nabla a_{j}S_{i}, \text{ when } i = j \\ \frac{\partial \frac{e^{a_{i}}}{\sum_{k=1}^{N} e^{a_{k}}}}{\sum 2} \\ = \frac{e^{a_{i}}\sum - e^{a_{j}}e^{a_{i}}}{\sum 2} \\ = \frac{e^{a_{i}}\sum - e^{a_{j}}}{\sum 2} \\ = S_{i}(1 - S_{j}) \end{array} \qquad \begin{array}{c|c} \hline \nabla a_{j}S_{i}, \text{ when } i \neq j \\ \hline \nabla a_{j}S_{i}, \text{ when } i \neq j \\ \hline \nabla a_{j}S_{i}, \text{ when } i \neq j \\ \hline \nabla a_{j}S_{i}, \text{ when } i \neq j \\ \hline \nabla a_{j}S_{i}, \text{ when } i \neq j \\ \hline \nabla a_{j}S_{i}, \text{ when } i \neq j \\ \hline \partial \frac{e^{a_{i}}}{\sum_{k=1}^{N} e^{a_{k}}}}{\partial a_{j}} = \frac{\partial -e^{a_{j}}e^{a_{i}}}{\sum 2} \\ = -\frac{e^{a_{j}}e^{a_{i}}}{\sum 2} \\ = -S_{j}S_{i} \end{array}$$

Skipping some steps for space, please see original notes.

$$\nabla a_j S_i = S_i(\mathbb{1}(i=j) - S_j)$$

а.



J

Softmax Loss Gradient wrt Score S

$$\begin{aligned} a_{j} &= w_{j}^{T} x_{j} \\ S_{j} &= \frac{e^{a_{j}}}{\sum_{k=1}^{N} e^{a_{k}}} \quad \forall j \in 1..N \\ \nabla S_{i}L &= \frac{\partial}{\partial S_{i}} - \log(S_{i}) = S_{j} - \mathbb{1}(i = j) \\ \nabla W_{j}L &= \frac{\partial L}{\partial S_{i}} * \frac{\partial S_{i}}{\partial W_{j}} = (S_{j} - \mathbb{1}(i = j))x_{i} \end{aligned}$$

Skipping some steps for space, please see original notes.

Computational Graph









$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4





$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4
$$q = x + y \qquad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want:
$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$$



$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$
Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

$$\frac{\partial f}{\partial f}$$
Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

$$\frac{\partial f}{\partial z}$$
Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

$$\frac{\partial f}{\partial z}$$
Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$
Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4
$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$
$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$
$$\frac{\partial f}{\partial q} = df \quad \frac{\partial f}{\partial q}$$

Want:
$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$$



$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$
Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$



$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \qquad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

$$\overset{x - 2}{y - \frac{1}{4}}$$



$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4
$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}$$

Want:
$$\frac{\partial y}{\partial x}, \frac{\partial y}{\partial y}, \frac{\partial y}{\partial z}$$



























$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$





$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$





$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$





$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$





$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$





$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$





$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$




$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$





$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$





$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$





$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$





$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$













$$\begin{aligned} f(w,x) &= \frac{1}{1 + e^{-(w_0 x_0 + w_1 x_1 + w_2)}} & \sigma(x) &= \frac{1}{1 + e^{-x}} & \text{sigmoid function} \\ \frac{d\sigma(x)}{dx} &= \frac{e^{-x}}{(1 + e^{-x})^2} = \left(\frac{1 + e^{-x} - 1}{1 + e^{-x}}\right) \left(\frac{1}{1 + e^{-x}}\right) = (1 - \sigma(x)) \sigma(x) \end{aligned}$$





$$\begin{aligned} f(w,x) &= \frac{1}{1 + e^{-(w_0 x_0 + w_1 x_1 + w_2)}} & \sigma(x) &= \frac{1}{1 + e^{-x}} & \text{sigmoid function} \\ \frac{d\sigma(x)}{dx} &= \frac{e^{-x}}{(1 + e^{-x})^2} = \left(\frac{1 + e^{-x} - 1}{1 + e^{-x}}\right) \left(\frac{1}{1 + e^{-x}}\right) = (1 - \sigma(x)) \sigma(x) \end{aligned}$$





Patterns in backward flow

add gate: gradient distributor
max gate: gradient router
mul gate: gradient... "switcher"?





Gradients add at branches





Implementation: forward/backward API

cl



Graph (or Net) object. (Rough psuedo code)

| ass Co | <pre>mputationalGraph(object):</pre> |
|--------|--|
| # | |
| def | <pre>forward(inputs):</pre> |
| | # 1. [pass inputs to input gates] |
| | # 2. forward the computational graph: |
| | <pre>for gate in self.graph.nodes_topologically_sorted():</pre> |
| | gate.forward() |
| | <pre>return loss # the final gate in the graph outputs the loss</pre> |
| def | backward(): |
| | <pre>for gate in reversed(self.graph.nodes_topologically_sorted()):</pre> |
| | <pre>gate.backward() # little piece of backprop (chain rule applied)</pre> |
| | return inputs_gradients |
| | |



Implementation: forward/backward API



(x,y,z are scalars)





Implementation: forward/backward API



| class Mu | ltiplyGate(object): |
|----------|---|
| def | forward(x,y): |
| | z = x*y |
| | <pre>self.x = x # must keep these around!</pre> |
| | self.y = y |
| | return z |
| def | backward(dz): |
| 0 | dx = self.y * dz # [dz/dx * dL/dz] |
| | dy = self.x * dz # [dz/dy * dL/dz] |
| | return [dx, dy] |

(x,y,z are scalars)





Example: Torch Layers

and a site of the

| ele l'est | Distance the time time. | |
|--|--|--|
| the attention of periods per | 121 | |
| 1100.000 | of some states | |
| Inclusion in Contra | And And Add | And Address of the Ad |
| | Charles Martin F | Card and address from an |
| | | |
| A | forest man manage & located as successively | 10.00 |
| ** | Trees and comparison of the same | 2 million and |
| * | As full survivor | |
| S. alarm | and a general second | 4 months app |
| a second | Participant and a strong test | |
| 8 million | and the in stand | 2 million and |
| 8 Marca | An feet service of \$10 sectors instance and the | inter Theorem |
| E Automation | to her sense of a sector of the | inter Transa |
| 8 Marca | to the set of the test | E |
| a second second | | 1.000.00 |
| a and and a | | 1.000.00 |
| a second second | | |
| | | |
| 1. January 1 | the Comments of Section Reserves | 100.00 |
| 8-10-10 | a little server as darks much with star bars and | |
| 8 Inclusion | | 1 million |
| a contraction of | and provide an | 1.000.00 |
| a contraction | and country in | i er ar |
| 8-Dallania | the strength of the president of patients | d works age |
| di langi se | the other may in the Park out which as Dava | 1 months age |
| E Institution or | And furniture process of function frame | 17 million and |
| 8 instead | Rocks and and service | 1.11.11.14 |
| 8 local lot in | Resultance Constitutes and the second result. | 1 million and |
| di lanta la | And and appropriate the second s | 1 months age |
| a tara | a property for a stand many stands of the standard and | 1 |
| a lance | | 1.11.11.14 |
| | | 1.000.00 |
| a contract of the second | | 1.00.00 |
| | the second se | |
| a last sector | | |
| a local sector | | |
| a batch for any lot | in such the local sector is the fermione | 1. mar. an |
| S Monthly | | 1 |
| # Investor | Faller Parts | 100000 |
| different for the second secon | An free service of the locality is adjusted and the | internet internet |
| di Indiana ini | the pile in comparison when the second of the | 1.00.00 |
| C. Latina in | is highly preserving studies (#17), all a black and | A sublicity |
| 8 haire | Carrier la colo | 1.0000.000 |
| & free lateria | which are not here with P.C. all these ad- | 1.000 |
| E Internetion | Age (September of Age | 4 auto ag |
| a subscript | All functions provides of functions | -1 - Har - Apr |
| a service of | | 1.00.00 |
| a sub- | | 1.000.00 |
| | | 1.000 |
| | the second se | |
| a selector | | 1 |
| Surfaces. | in much "Weight on high from the | |
| a real residence. | Interpretation marries | 1 million |
| distance in | hart. Paulo concern marries | |
| Combine . | An Test service et dist insufficie instances instations | internet Transme |
| A comment | Annual sector ratio for a line | 1 |

| Servoid los | Add THNN convention of (ELU, LexityReLU, LogSigmoid, LogSofMax, Looku | 7 days ago |
|--|--|---|
| iotMax.lue | Add THNN conversion of (ELU, LeekyReLU, LogSigmoid, LogSofMax, Looku | 7 days ago |
| up Table Aus | Harmonize LookupTable signature with oursn impl | 5 days ago |
| lua . | Rename unpack to table unpack for Lus 5.2 | 8 months ago |
| Criterion lua | Add SizeAverage to oriterions in the constructor | 2 months ago |
| inCriterion.lua | modernized MarginCriterion | a year ago |
| inRankingCriterion.lua | Fix batch mode in MarginRankingCriterion | 4 days ago |
| lue | Merge pull request #464 from vgirelmaster | 2 months ago |
| n lua | Add support for negative dimension and both balch and non batch input | 2 months ago |
| ua - | Merge pull request #464 from vgirelmaster | 2 months ago |
| areTable.kus | cancel unused variable and useless expression | 29 days ago |
| ule lua | Revert "Don't re-flatten parameters if they are already flattened" | 15 hours ago |
| ua . | removing the requirement for providing size in m.Mul | a year ago |
| Constant.lua | Ignore updateGradinput if self.gradinput is nil | 3 months ago |
| Criterion.lue | asserts in MultiCriterion and ParallelCriterion add | 2 months ago |
| LabelMarginCriterion.lua | initial revamp of torch7 tree | 4 years ago |
| MarginCriterion.lus | multimargin supports p=2 | 11 months ago |
| per Auto | typeAs in Namow not done in place. | 6 months ago |
| owTable.lua | NarrowTable | 6 months ago |
| nalize lue | Remove brem and baddherm from Normalize, because they allocate memory, | 20 days ago |
| LUJua | Buffers for PReLU cude implementation. | 8 months ago |
| Sing lua | fixed broken m.Padding: input was returned in backprop | 5 months ago |
| viseDistance Jua | Merge pull request #532 from xegeng/master | 29 days ago |
| fel.lua | fix a bug in conditional expression | a month ago |
| RelCriterion lua | asserts in MultiCriterion and ParallelCriterion add | 2 months ago |
| lle/Table Jua | Parallel optimization. ParallelTable inherits Container. unit tests | a year ago |
| er Jua | Use UNIX line endings | 7 months ago |
| CME.md | doc readthedocs | 5 months ago |
| LU Aaa | Add randomized leaky rectified linear unit (RReLU) | 3 months ago |
| U Nua | adds in-place ReLU and fixes a potential divide-by-zero in nn.Sigt | 9 months ago |
| icate Jua | Replicate batchMode | 8 months ago |
| tape lua | Added more informative pretty-printing. | a year ago |
| ct.lue | initial revamp of torch7 tree | 4 years ago |
| d'Table Aua | nn.Module preserve type sharing semantics (#167); add nn.Module apply | 4 months ago |
| sential Jua | fxing Sequentiat.remove comer case | 6 months ago |
| eid.lua | initial revamp of torch7 tree | 4 years ago |
| othL1Criterion.lua | Add SizeAverage to criterions in the constructor | 2 months ago |
| Rex. Sub | Fix various unused variables in m | a year ago |
| in ua | Fix vanous unused vanables in m | a year ago |
| 10.08 | toted a numerical issue in the contrists module (it breaks for input g | |
| shink.ua | india revamp or toron/ tree | 2 years ago |
| sign.ua | | 2 years ago 4 years ago |
| | Initial revenue of torch? tree | 2 years ago 4 years ago 4 years ago |
| seJacobian lua | Initial revenp of forch? two Fix various unused variables in m Union exame initiamentation of examples discussed on the force of the second loss of | 2 years ago 4 years ago 4 years ago a year ago |
| seJacobian lua seLinear lua | Initial revents of torch? the Fix various unused variables in m Using spanse implementation of zeroOndPratmeters for SpanseLinear Model Concellingenetiation Provider | 2 years ago 4 years ago 4 years ago a year ago a month ago |
| seJacobian lua seLinear lua laikdaptiveMaxPooling | Initial ensemp of stand? two Fire various smalled line in the Using spanse implementation of zero/Geo/Parameters for SpansaLinear Added Spansialishighterbolkale/boding Schröderbauerboling | 2 years ago 4 years ago a years ago a month ago a year ago a year ago |
| seJacobian lua seLinear lua ialAclaptiveMaxPooling ialAveragePooling lua ialBateNicorratination long | Initial research and the New Yes and Section 2014 Section | 2 years ago 4 years ago a years ago a month ago a year ago 29 days ago 7 daos ar- |
| sulacobian lua sel,inear lua ialAdaptiveMaxPooling ialAveragePooling lua ialBatchNormalization lua ialBatchNormalization lua | hild many fitted two Financial stands of endlose in the Using survey inspirementation of zeroGoadPursenteurs for SpanicLinear Added ExplanicAuge-Mind Prucing Standson-Spanic Prucing Standson and exclude gaid div Add Cheptenetweistere of Experimized/Standsonautorion Made Search Standsonautorion | 2 years ago 4 years ago 4 years ago a year ago a morth ago 29 days ago 7 days ago 9 mortha am |
| seJacobien lus seLinear lus istAdaptiveMaxPooling istRiveragePooling lus istBetchNormelization lus istBetchNormelization lus istConstativeNormeliz istConstativeNormeliz | Jalid energy fund has Tanda unset all saids in the Using query approximation of anti-Dardhaneses for Spanichaer Mark Equitablepartmethology Spatisheergy harding comotion and existing and dr All Componenter of Disputchertometation Mark Egio(1) syntaxise. | 2 years ago 4 years ago 4 years ago a year ago a morth ago 20 days ago 7 days ago 9 morths ago 3 morths ago |
| sularoohien kus sellinear kus altikdaptivelilaa Pooling aldiversgePooling kus altiloth Normelization kus altiloth Normelization kus altiloth Normelization kus altiloth Normelization kus | Statisticang market set Statisticang and Statisticang Sta | 2 years ago 4 years ago 4 years ago a year ago a morth ago 29 days ago 7 days ago 9 morths ago 3 morths ago 3 morths ago |
| sedecobien has sed.inear. has all/disptive/MaxPooling all/overagePooling.has all/ontrastiveNormaliz all/onvolution.has all/onvolution.has all/onvolution!Masa | Nationary and a second | 2 years ago 4 years ago 4 years ago a year ago a north ago 29 days ago 7 days ago 9 montha ago 3 montha ago 3 montha ago |
| sulacobien has sel.inear.ha al/Adaptive/MaxPooling al/AveragePooling.has al/OntrastiveNormaliz al/Onvolution.has al/Onvolution.has al/OnvolutionMit.has al/OnvolutionMit.has al/OnvolutionMit.has al/OnvolutionMit.has | Sala kenang Junio Shao Mang Salaman Salaman Salaman Salaman Salaman Salaman Salaman Salaman Salaman Salaman Salaman Salaman Kenang Salaman Sal | 2 years ago 4 years ago 4 years ago a war ago a morth ago 20 days ago 7 days ago 9 morths ago 3 morths ago 3 morths ago 8 morths ago 11 dags ago |
| subsochen has sekinear kaa alkdaptivelikarPooling alkkerspePooling.kas allidisch-Normalization kas allidisch-Normalization kas allidisch-Normalization kas allidisch-Normalization allidisch-Normalization allidisch-Normalization allidisch-Normalization allidisch-Normalization | Nationary and a second | 2 years ago 4 years ago 4 years ago a year ago a year ago 29 days ago 7 days ago 9 montha ago 3 montha ago 8 montha ago 8 montha ago 18 montha ago 18 days ago |
| salatobin ha salatobin ha alidapting Pooling La alidarsop Pooling La alidarsop Pooling La alidarson kalaton ha alidarsoh kalaton alidarsoh kalaton alidarsoh kalaton alidarsoh kalaton alidarsoh kalaton alidarsoh kalaton alidarsoh kalaton alidarsoh kalaton | Statisticating "United Tead Composition Control Contr | 2 years ago 4 years ago 4 years ago a year ago a year ago 29 days ago 7 days ago 9 months ago 9 months ago 3 months ago 8 months ago 18 days ago 8 months ago 18 days ago 6 months ago 6 months ago |
| salasobin ka selanar ka akikersgePoling akikersgePoling ka allatoNomalastion ka allomotulomMika allomotulomMika allomotulomMika allomotulomMika allomotulomMika allomotulomMika allomotulomMika allomotulomMika allomotulomMika allomotulommalastis allomotulom | Nationary and a second | 2 yean ago 4 yean ago 4 yean ago a moth ago a moth ago 20 fony ago 7 dayn ago 9 mothe ago 3 mothe ago 8 mothe ago 18 dayn ago 8 mothe ago 18 dayn ago |
| nakatoho ka seliner ka akkespinkakatoho akkespinkakatoho akkespinkakatoho aktorekatoha aktorekatoha aktorekatoha aktorekatoha aktorekatoha aktorekatoha aktorekatoha aktorekatoha aktorekatoha aktorekatoha aktorekatoha aktorekatoha aktorekatoha | Stati kenarg (1997) Met Stati kenarg (1997) Met Statig kanarg kenarg kenarg kenarg kenarg kenarg Statishang kenarg kenarg kenarg kenarg Statishang kenarg Statish | 2 yean app 4 yean app 4 yean app 4 yean app a north app a north app 29 days app 20 days app 7 days app 9 morths app 9 morths app 8 morths app |
| skäröhön ka sekinar ka akkersyeftöding ka akkersyeftöding ka akkersyeftöding ka aktorektörön ka aktorektöröhligt ka | State source printed the State | 2 yean app 4 yean app 4 yean app a morth app a morth app 3 year app 2 d dyn app 7 dnyn app 8 morth app 3 morths app 11 dnyn app 8 morths app 3 morths app |
| skatobin ka sel, inar, ha sel, inar, ha alkerge Pooling ka alkerge Pooling ka alkerge Pooling ka alkerse kan kan alkerse kan alkerse kan kan alkerse kan kan alkerse ka | Statisticang Turk Tead Dead Security Control Tead Dead Security Control Tead Security Security Security Security Tead Security Security Security Security Security Security Security Tead Security Security Security Security Security Security Security Security Tead Security Secur | 2 yean ap 4 yean ap a wordt ap a wordt ap a wordt ap a wordt ap a wordt ap 2 dropt ap 7 dont ap 9 morths ap 9 morths ap 9 morths ap 9 morths ap 8 morths ap 9 morths ap |
| salasibina ha ashara ha as | Stati security affection from Konstantian Statistica St | 2 yuun ap 2 yuun ap 4 yuun ap 8 yeer ap a moth ap 1 yeer ap 20 days ap 7 days ap 20 days ap 7 days ap 20 days ap 7 days ap 20 |
| sakatoha ka eskinet ka elkdepixelkarPooleg.a elkdenivelkarPooleg.a elkdenivelkarPooleg.a elkdenivelkarPooleg.a elkdenivelkarPooleg.a elkdenivelkarPooleg.a elkdenivelkarPooleg.a elkdenivelkarPooleg.a elkdenivelkarPooleg.a elkdenivelkarPooleg.a elkdenivelkarPooleg.a | Sala kenang Junot Paul Canada analar dendara kenang Sala Kenang Kenang Kenang Kenang Kenang Kenang Ken | 2 yuun app 4 yuun app 4 yuun app a yuun app a yuun app 20 dayn ap |
| sakatohan ka eshara ka esh | Sala kanagi Jandri Yao K Managi Sanaha Jandri Mahamanan Ka Janakiawa Mang Sanaha Jandri Mahamana Ka Janakiawa Mang Sanahang Mang Sanahan Jang Jandri Mahamana Ka Mang Janahan Mat Cangona Sanahan Jang Jandri Mahamana Ka Mahamana Mang Sanahang Mahamana Janahan Ka Sanahang Jandri Sanahan Jandri Mahamana Ka Sanahan Janahan Jandri Sanahan Jandri Sanahan Ka Sanahan Jang Sanahan Jandri Sanahan Jandri Sanahan Ka Sanahan Jandri Sanahan Jandri Sanahan Jandri Jandri Sanahan Jandri Sanahan Jandri Jandri Sanahan Jandri J | 2 yuun ago 4 yuun ago 4 yuun ago 8 yuun ago 20 daya ago 20 daya ago 20 daya ago 20 daya ago 20 daya ago 20 daya ago 3 nontha ago 8 nontha ago 8 nontha ago 8 nontha ago 8 nontha ago 8 nontha ago 5 daya ago 3 Junatha ago 8 nontha ago 5 daya ago 3 Junatha ago 10 nontha ago 6 nontha ago 6 nontha ago 6 nontha ago 6 nontha ago 6 nontha ago |
| sakatohan ka sakatoan ka sakatohan ka Mungun akkenge Poolng ka sakatohan ka Mungun akkenge Poolng ka sakatohan ka Mungun aktorek ka Mungun | Nak neuropi netro from 6 Van de same ander en de sector de la sector | 2 years app 4 years app 4 years app 8 year app 8 year app 20 days app 21 days app 23 worths app 23 worths app 24 worths app 23 worths app 23 worths app 23 worths app 23 worths app 23 worths app 24 days app |
| | Sala karang Jandra Sala Sala karang Jandra Sala Sala giana karang Sala Karang Sala Sala giana karang Sala Sala Sala Mar Capasa karang Sala Sala Sala Sala Sala Sala Sala Sala | 2 years app 2 years app 4 years app 4 years app 4 years app 4 years app 2 days app 2 days app 2 days app 2 days app 2 days app 3 months app 6 months app 6 months app 6 months app 5 days app 6 months app 5 days app 6 months app 5 days app 5 days app 6 months app 5 days app 6 months app 6 months app 6 months app 6 months app 6 months app 4 months app 3 months app 4 months app |
| skatebioka karea ka uskateja karea ka uskateja karea karea päveleja karea uskatejä karea uskatebioka karea uskatebioka ka uskatebioka ka uskatebioka ka uskatebioka ka uskatebioka ka uskatebioka ka uskatebioka ka uskatebioka ka uskatebioka ka uskatebioka ka | Skil kenning Print Print Verstern aussell er understellen er fragenschwert Stephensemperatieren der aufschaft werenten fragenschwert Stephensemperatieren der aufschaft werten er fragenschwert Stephensemperatiere der aufschaft werten er fragenschwert Stephensemperatiere der aussellt werten er fragenschwert Stephensemperatiere der aufschaft werten er fragenschwert Stephensemperatiere der aussellt werten er fragenschwerten er Stephensemperatiere der aussellt werten er Stephensemperat | 2 years app 4 years app 4 years app 4 years app 4 years app 4 years app 2 didn's app 2 didn's app 5 norths app 5 norths app 8 norths app 8 norths app 8 norths app 8 norths app 8 norths app 9 norths app 8 norths app 9 norths |
| Makababan Kana Ja Makapa Karan Ja Makapa Kana Jang Makapa Kana Jang Makababan Makababan Makababaha | Stati sensing Transformation Stati Composition Statistics Statisti | 2 years app 2 years app 4 years app 4 years app 4 years app 4 years app 4 years app 2 days app 2 days app 5 months app 5 months app 5 months app 5 months app 6 months app 10 months app 1 |
| skatelskatel katerak ka utakapiskatelskatelska utakapiskatelskatelska utakapiskatelskatelska utakapiskatelska utakapiska | Skil kennig Pullet Skil Version Lander Underständ Skill Version Lander Underständ Skill Version Lander Underständ Skill Version Lander Underständ Skill Version Lander Version Lander Underständ Skill Version Lander Version Lan | 2 years app 4 years app 7 onys app 9 norths app 13 norths app 13 norths app 13 norths app 14 norths app 15 norths app 10 norths app |

| E Max.lua | Merge pull request #464 from vgire/master |
|-------------------------------|---|
| E Mean.lua | Add support for negative dimension and both batch and non batch input |
| E Min.lua | Merge pull request #464 from vgire/master |
| MixtureTable.lua | cancel unused variable and useless expression |
| E Module.lua | Revert "Don't re-flatten parameters if they are already flattened" |
| E Mullus | removing the requirement for providing size in nn.Mul |
| MulConstant.lua | Ignore updateGradinput if self.gradinput is nil |
| MultiCriterion.lue | asserts in MultiCriterion and ParallelCriterion add |
| MultiLabelMarginCriterion.lua | initial revamp of torch7 tree |
| MultiMarginCriterion.lua | multimargin supports p=2 |
| E Narrow.kus | typeAs in Narrow not done in place. |
| NarrowTable.lus | NarrowTable |
| Normalize.lua | Remove brim and baddomm from Normalize, because they allocate memory, |
| PReLU.lua | Buffers for PReLU cude implementation. |
| E Padding.lua | fixed broken nn.Padding: input was returned in backprop |
| PainelseDistance.lua | Merge pull request #532 from xwgeng/master |
| Parallellua | fix a bug in conditional expression |
| ParallelCriterion.lua | asserts in MultiCriterion and ParallelCriterion add |
| Paralle/Table.lua | Parallel optimization. ParallelTable inherits Container. unit tests |
| Power.lua | Use UNIX line endings |
| E README.md | doc readthedocs |
| RReLUJua | Add randomized leaky rectified linear unit (RReLU) |
| ReLU Jue | adds in-place ReLU and fixes a potential divide-by-zero in nn.Sgrt |
| Replicate Jua | Replicate batchMode |
| E Reshape.lua | Added more informative pretty-printing. |
| E Selectius | initial reviamp of torch? tree |
| SelectTable.lua | nn.Module preserve type sharing semantics (#187); add nn.Module.apply |
| Sequential lua | fixing Sequential.remove corner case |
| Sigmoid lua | initial revamp of torch7 tree |
| SmoothL1Criterion.lua | Add SizeAverage to oriterions in the constructor |
| E SofMax.lua | Fix various unused variables in nn |
| E Sollin ha | Fix various sourced variables in on |

* Original slides borrowed from Andrej Karpathy and Li Fei-Fei, Stanford cs231n

comp150dl

Example: Torch Layers

| 1000 C | | tent. a give to your of |
|--|--|-------------------------------|
| the lines of | Distance of the state of the | |
| | | |
| second a press of | 10000 | |
| 1100-000 | Product Street | 0.000 |
| 100 | | and the second second |
| | Canada Canad | |
| state and state | All has an interest the same in | Units and public of Cross and |
| | A set water beginning to be | |
| | typed the mage I failed annual the | - 5- 50 |
| - | THE REPORT AND A DESCRIPTION | 1.00 |
| | As full sprang | |
| and the second s | and a descent state | |
| | | |
| | | 1.000 |
| | an fees service ethic conduct configurat coloring con- | 1.00 |
| | an pass to serve a spectral compact collegest collegest con- | 1.000.00 |
| | A AN AN AN AN AN AN | |
| | | 1.0000.00 |
| Automatical State | Record presents, ratios for \$20 details | 1.000.00 |
| Section 1. | to another that | 1 million and |
| 100 000 000 | they are excluded when preservative of particular | 1 |
| | tray and south a dust contracted of problem | 1.000.00 |
| Contract of the | April oppression a lipstallocitiente | 1.00.00 |
| Children of Childr | a high prove prototy music (FT), all a high and | |
| 100 million (100 | they are easily a star president of pathods | - and the second |
| | AND DESCRIPTION OF | 1 10110 14 |
| CP-MOT A | all county for | 1.000.00 |
| Confront on | they also easily a more contracting of galaxies | i sutto ap |
| (market) | the pattern angle interfact and watching three | 1 months age |
| Section 1. | Add for address of some and for all and the second se | 1.000.00 |
| Location I. | Reading is antified transition | 1.11.00.00 |
| Constitute (an | Regulation Control State Longitude | * 1011 A |
| Construction of the | And a set of the set o | 1 months ago |
| Carrier Contraction | to find prove prototy music of the strain and | A success op |
| Letteria . | Acquire land | 1.11.11.14 |
| Landing on La | to a regime state a landau s'understighters | 2 months and |
| and set of second | In the story one wanted in Constitution of Con | 1 months ago |
| (and the second | to find prove protectly security of C., and the field and | A success of |
| Contract State State | Rear and the set of the set of | 1 |
| Indexed Service | charter activate and account to maximize | 1.0000.000 |
| institution of the | alling the industrie local limiting income | 1 million age |
| Seek (Sel) dense for | the second field builded and it stight device space | -1 mar 14 |
| Advance. | And and the Contract of the | 1 coltante |
| Instation | Faller Page | 1.0000.00 |
| Surger Contraction | An Test service et 211 Sectors Support Support Supports Sec. | Trans. |
| | the plan increasing the structure in any local from | 1.00.00 |
| Contract into | is find prevering during security of \$1, all a finite gale | 4 months age |
| bala. | Ear room lacross | 1 million and |
| free free la | a finit server or derive motio 247, all a finite and | 1 million (1 |
| Index Strength of | An Deliver of the | A success of |
| and the second se | Addressing provide of Reddink | 1.00.00 |
| test for the | Addression provide all hard and | 1.00.00 |
| the state of the state of | more that with a contract and not | 1.000.00 |
| and a second | Recent preside likely or independent | 2 months and |
| and a second sec | Ingling at the distance of the | 1 worth op |
| | An other in the late of the second second second | 1 |
| and internal | to Real preservation and a Philade Annual and | 1 million (10 |
| ciania. | in much here and a set of the | 1.00.00 |
| makes and so and | New galaxy marks | i sutto ap |
| Charles into | hard Paralle presents apprent | |
| and the local diversion of the local diversio | to first service effect institute instance calculate term. | Trans. |
| interior. | Annual sector relation from an inter- | |

| Sigmoid lug | Add THNN conversion of (ELU, LeakyReLU, LogSigmoid, LogSofMax, Looku | 7 da |
|--|--|---------|
| gSotMax.lua | Add THNN conversion of (ELU, LeakyReLU, LogSigmoid, LogSofMax, Looku | 7 da |
| okupTable Aus | Harmonize LookupTable signature with ourn impl | 5.64 |
| £lua . | Rename unpack to table unpack for Lua 5.2 | 8 mont |
| SECriterion.lua | Add SizeAverage to oriterions in the constructor | 2 mont |
| erginCriterion.lua | noderrized MarginCriterion | a ye |
| erginRankingCriterion.lua | Fix batch mode in MarginRankingOnterion | 4.64 |
| ex.lue | Merge pull request #464 from vgire/master | 2 mont |
| en.lue | Add support for negative dimension and both batch and non batch input | 2 mont |
| niua | Merge pull request #464 from vgire/master | 2 mont |
| itureTable.kua | cancel unused variable and useless expression | 29.68 |
| odule lua | Revent "Don't re-flatten parameters if they are already flattened" | 15 hou |
| (Jue | removing the requirement for providing size in nn.Mut | a ye |
| #Constant.lua | Ignore updateGradinput if self.gradinput is nil | 3 mont |
| diCriterion.lua | asserts in MutiCriterion and ParallelCriterion add | 2 mont |
| ALabelMarginCriterion.lua | initial revemp of torch7 tree | 4 yea |
| #MarginCriterion.lus | multimargin supports p+2 | 11 mont |
| mpe lus | typeAs in Namow not done in place. | 6 mont |
| mowTable.lua | NanowTable | 6 mont |
| rmalize ka | Remove brem and baddomm from Normalize, because they allocate memory, | 20 da |
| leCU lua | Buffers for PReLU cude implementation. | 8 mont |
| dding.lua | fixed broken nn.Padding: input was returned in backprop | 5 mont |
| inviseDistance.lua | Merge pull request #532 from xegengimaster | 29.64 |
| rafiel.lua | fix a bug in conditional expression | a mor |
| rallelCriterion.lua | asserts in MutiCriterion and ParallelCriterion add | 2 mont |
| ralle/Table.lua | Parallel optimization. ParallelTable inherits Container. unit tests | a ye |
| wer.lua | Use UNIX line endings | 7 mont |
| ADME.md | doc readthedocs | 5 mont |
| ReCU Aux | Add randomized leaky rectified linear unit (RReLU) | 3 mont |
| CU Jua | adds in-place ReLU and fixes a potential divide-by-zero in nn.Sqrt | 9 mont |
| plicate Jua | Replicate batchlilode | 8 mont |
| shape lua | Added more informative pretty printing. | a ye |
| lect.lue | initial reversp of torch? tree | 4 yes |
| iectTable Jaa | nn.Module preserve type sharing semantics (#187); add nn.Module.appty | 4 mont |
| quential lua | tixing Sequential.remove comer case | 6 mont |
| pricid.lua | initial revemp of tarch? tree | 4 yes |
| roothL1Criterion.lus | Add SizeAverage to oriterions in the constructor | 2 mont |
| PMax.Jua | Fix various unused variables in m | a ye |
| PMAC LA | Pix vanous unused vanables in m | a ye |
| iphus.lua | fixed a numerical issue in the bottPlus module (it breaks for input g | 2 yes |
| /tShrink.lua | initial revemp of tarch? tree | 4 yes |
| ASign Jua | initial revemp of tarch? tree | 4 yea |
| erseuecobian tua | r or vancus unuseo vanacies in III Unice exerce instrumentation of source Record Instrumentation for Record Instru | a ye |
| anactical loa | Using sparse imprementation of zeroGradParameters for SparseLinear | a mor |
| aba0AdaptiveMaxPooling | Added SpelasAdaptiveMaxPooling | aye |
| anarwerageroceng lua | openantwenger-sonny supports pedang, cel mode and exclude_pad dv | 29.68 |
| Contraction lus | Hole keeli kekenatia | 7.68 |
| anar. on resolution has | Make type,) may recursive. | 9 mont |
| anan unvolution tus | r in system og anders server 10000 | |
| and the state of t | r in 1999) in approximation with the second se | a more |
| anan-onvolutionship ka | remove unuses and experience initialization logic from initiapatiation vol | s mont |
| anarci ossmapcion us | Control Control of Con | 18 64 |
| Conception and the | oparaquioneraeova,uvvava,ouceaceva/vormanzeoñ work with bal | 8 mont |
| anano-opoutrua | arran in on error message | e nore |
| statistic di Consultation ? | Add adjustment term in StraticE-IP resolution to control the size of | 2 101 |
| and a constant for the | And approximate arms to open amount of condition the 5/28 of | 0.00 |
| and Photos by | President features Broking divides to 1995/44 | 3 yes |
| and Pooling Aus | operative energy of 5285 EV KHT APT | to more |
| attaliant location b | Add Seatisfiched Incodes | 0 more |
| ation Conference and | rear operational powers in and in analytic mode. | 20.00 |
| and the second s | upperer overenen av more of hjøller miner Manne hennels for fast mener | - 100 |
| and a cost of the second second | Marine more available | Jye |
| | the set of the second | |
| and bull sources that have not | Spatial(Constrastive,Divisive,Subtractive)Normalization work with bat | |
| atiaUpSamplingNearest I | Spatial/Constrative.Durave.Subtractive/Normalization work with bit Use UNX line endings | 7 mont |
| atial.lpSampingNearest.l atialZeroPadding.lua | Spatial(Constantive,Drave,Subtactive)Normalization work with bat Use UNIX line endings Added nonce informative pretty origing. | 7 mont |









```
local MulConstant, parent = torch.class('nn.MulConstant', 'nn.Module')
function MulConstant:__init(constant_scalar,ip)
  parent.__init(self)
  assert(type(constant_scalar) == 'number', 'input is not scalar!')
  self.constant_scalar = constant_scalar
  -- default for inplace is false
   self.inplace - ip or false
  if (ip and type(ip) -= 'boolean') then
      error('in-place flag must be boolean')
   end
function MulConstant:updateOutput(input)
  if self.inplace then
    input:mul(self.constant_scalar)
    self.output = input
  0150
    self.output:resizeAs(input)
    self.output:copy(input)
   self.output:mul(self.constant_scalar)
  end
 return self.output
function MulConstant:updateGradInput(input, gradOutput)
 if self.gradInput then
    if self.inplace then
     gradOutput:mul(self.constant_scalar)
     self.gradInput - gradOutput
      -- restore previous input value
      input:div(self.constant_scalar)
    else
      self.gradInput:resizeAs(gradOutput)
     self.gradInput:copy(gradOutput)
     self.gradInput:mul(self.constant_scalar)
    end
    return self.gradInput
```

Example: Torch MulConstant

$$f(X) = aX$$

initialization

forward()





Example: Caffe Layers

| DescriptionDescripti | And Address of the local division of the loc | an i coffe Tayana i | And And And | E-10101.00.000 (8 | donarile sur feature | 1.000 10 |
|---|--|--|--|--|--|---|
| Company< | Company of the local data | nonder or | 1001 1011 00 Mill 1000 00 | England (spin, inc.) | dependent of the second of the | a month ag |
| Note of the section | | | | \$100,000 (B) | denarity teat feature | 1.000 4 |
| Norm | 1 | Result of Sales | | Elena, base or | discussion in an insulation | a month ag |
| Notional Notati | 8-10-10-10-10 | Annual and Assess | 1.111.00 | Science and the | domains tour housen | 1.000 0 |
| Annual Annua | Energy and one | Broath age teams | + | Electric (see as | denarie is a heaters | a menti ap |
| NumberNumberNumberNumberStandardStand | General and an | Annual and Annual An | 1.111.00 | Sciences and in | demantie over heaters | 1.000 0 |
| NoteNoteNoteNoteSamperSam | Energy and particular | waits their investments | 1.000.000 | Energy in the set | dependencies institution | a method |
| Image <td>\$100.00.00 at</td> <td>Bright by Sellin</td> <td>1.111.00</td> <td>Annual states in</td> <td>And the second second</td> <td></td> | \$100.00.00 at | Bright by Sellin | 1.111.00 | Annual states in | And the second second | |
| kittingkittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingkittingkittingkittingkittingkittingKittingki | Eren, majoren | Recards and Institute | 1.000 | | | |
| Lat. Applie and and and applied and a | Science and the | Barrache bair feallan | 111110 | E 1000, 1000 000 | and the states | 1.000 4 |
| InteractioninteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteractionBittanameinteractioninteractioninteractioninteractioninteraction <tr< td=""><td>Construction of</td><td>Borardo auto materi</td><td>1 1 1 1 1 1 1 1</td><td>Biner the D</td><td>demantic tear heaters</td><td>1.000.0</td></tr<> | Construction of | Borardo auto materi | 1 1 1 1 1 1 1 1 | Biner the D | demantic tear heaters | 1.000.0 |
| NumberN | 1 101 1 10 10 10 10 10 10 10 10 10 10 10 | Barracile bar factors | 1.111 | E-months, sign on | documents again managers | 1 4 |
| Interfact <td>\$100,000,000 L</td> <td>Read of the second</td> <td>1.000.00</td> <td>Endedon And to</td> <td>demantitie lauer feasters</td> <td>1.0000.00</td> | \$100,000,000 L | Read of the second | 1.000.00 | Endedon And to | demantitie lauer feasters | 1.0000.00 |
| NumberNumberNumberNumberNumberNumberNumberNumberNumberStandard <t< td=""><td>and an end</td><td></td><td>1100.00</td><td>Employee and</td><td>dense for any treatment</td><td>a could ap</td></t<> | and an end | | 1100.00 | Employee and | dense for any treatment | a could ap |
| NoticeNotic | | | | Ermanne to | derivative last features | 1.000 |
| Norm Norm <th< td=""><td>Contractor of</td><td>Annual to and the second</td><td>+ 100 m</td><td>E-man, non-see</td><td>decards and loaders</td><td>a maril ag</td></th<> | Contractor of | Annual to and the second | + 100 m | E-man, non-see | decards and loaders | a maril ag |
| Normal Normal< | \$1000 million and \$100 | served and hear | 111110 | Suprementary and super- | And a few from the second | 1000 |
| Number Note Note Note Note Number Sold Nation Sold Nation< | Contractor, inc. inc. | derards and leader | 1 | Report and show in | description of the local division | |
| \mathbf{u} \mathbf{u} \mathbf{u} \mathbf{u} | \$100,000 (B | and require the processing of the second sec | 11010 | | | |
| k = k = k = k = k = k = k = k = k = k = | Concernant of the local division of the loca | Street in the local sector. | a month space | | Bridde Gar Namer | 1.000 0 |
| Kath,AreaNationNationNationNationNationKath,AreaK | \$1000 AND 100 | Stratte bet feater | 1. | E-specific and | and a first state. | 1.000 |
| NAMAGE NAME < | E-10,000,000 0 | Automotive and an end of the | 1.000.000 | B. 1993. See 10 | donarile tear feature | 1.000.0 |
| k = k + k + k + k + k + k + k + k + k + | \$1000 C. 100 C. | Benarie bar hasses | 1.111.00 | E dest, borns | discards by a leader. | a marite ap |
| Bit Audit Bit Audit <t< td=""><td>Contraction of the local distance of the loc</td><td>Bound and them.</td><td>1.111.00</td><td>\$100,000 DB</td><td>densitie teer heaters</td><td>1.0007-0</td></t<> | Contraction of the local distance of the loc | Bound and them. | 1.111.00 | \$100,000 DB | densitie teer heaters | 1.0007-0 |
| Number (Note) Number (| \$1000, PL 800 (B) | And a first the second second | 1.000.000 | Entra Jacobia | disearch later healters | a manife age |
| Internation | E-mail of the second | Annalis and Index | 1.000.000 | Sectors. Not 20 | demander team features | 1.107.0 |
| Lat. Mathew Mathw Mat | CONTRACTOR OF STR | Bridde bar faller | 1.000.00 | Endow Jack of | donarile just institut | a method |
| Alternation | Contraction of the local division of the loc | | 1 | Scotten and April 10 | destantine lower housest | 1.000 0 |
| Mathematical Strategy | Train at land | | | Riveland into internet | And and the second | |
| Non-special control Non-specian control Non-specian control | Annual Agency and the | Briefs for Seller | 1007.00 | | And the second second | |
| Point Sarting Point Value | Contra spectra and a | description and testing | a month space | | | 1100 |
| Note State State State | | Strate and Second | 1000 | Electronic and an and | And and the second | 1000.0 |
| Schladig Anstein Schladig | Contraction, and or | stratic up tam. | 1.000 (00) | 2 m. mm to | durights and features | 1000.0 |
| Nation State Nation State< | | And and a second second | 1.111.00 | Eren, son op | denotes and teachers | 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - |
| Entropy Matrix (MI) Matrix (MI) Matrix (MI) Address (MI) | Contra (100, 100 - 10 | Recardle approximation | 1 | Grant, search | donarily but formers | 1000.0 |
| Entropy And Annual Factors Annual (Constanting) Annual | \$100.000 to | Annual and Annual | 1.000 | Charles, son on | denotes again tradema | 1 |
| Extra Lation North Years Little Diffusion Little Extra Lation North Years North Years </td <td>E-man and a</td> <td>and the second second</td> <td>1.000-000</td> <td>Concession, size of</td> <td>domarks part hadnes</td> <td>1.107</td> | E-man and a | and the second second | 1.000-000 | Concession, size of | domarks part hadnes | 1.107 |
| Deput (100) Mark to table 1 mile to table to table 1 mile to table to table to table to tab | Grant and | Reverse and Automatic | 1.111.00.000 | \$100,000 mm | documents in an instance | a month ag |
| Extension State of the stat | Contract Name and | Recently and resident | 1.000 mga | Service of the local division of the local d | strain her here | |
| Description match of the second | Contraction in | Benadia kan faasan | 1107.00 | The second se | | |
| Entra Anna anna anna anna anna anna anna an | Course, and the state | develop the leader | 1.000 | 1. mar. (| and the state of t | 1000 0 |
| Tana a su | 8-000 (M | Barracha basi hasilari | 1.000.00 | | | |
| | 1. mar. 100 - 11 | Borards and Hallen | 1 101 101 | | | |

* Original slides borrowed from Andrej Karpathy and Li Fei-Fei, Stanford cs231n

Sec. 10.0







finclude <vector> finclude <nondoms finclude <itorstor> finclude "coffe/layers/shuffle_layer.hpp" finclude "coffe/util/math_functions.hpp"

namespace caffe {

// ORECLEQ(bottom[0]->shape().size(), 2); shuffle_seed_ = rand();

// calculate count of each item in batch
batch_item_size_ = battom[@]->count(1, battom[@]->CononicalAxisIndex(-1));

vectorkint> shuffle_order; // Make a vector of ordered inds for (int i=0; i<bottom[0]->shape(1); i++) shuffle_order.push_back(i); // Mersenne twister initialized with input seed std::stuffle(shuffle_order.begin(), shuffle_order.end(), gen);

// copy randomized shuffle order to layer member variable shuffle_order_ shuffle_order_.Reshape(shuffle_order.size(), 1, 1, 1);

for (int i = 0; i < shuffle_order.size(); i++) {
 shuffle_order_.mutable_opu_data()[shuffle_order_.offset(i)] = shuffle_order[i];</pre>

const int count = top(0)=>num(); std::cout << bottom(0) =>shape(0) << std::end1; std::cout << bottom(0) =>shape(0) << std::end1; std::cout << bottom(0) =>shape(1) << std::end1; vector<int> new_shape; new_shape(); vector<int> new_shape; new_shape(); bottom(0) =>Reshape(new_shape); top(0) =>Reshape(new_shape(new_shape); top(0) =>Reshape(new_shape(new_shape(new_shape); top(0) =>Reshape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_shape(new_sh

bool forward = true;

#ifdef CPU_ONLY
STU8_GPU(ShuffleLayer);
#endlf

INSTANTIATE_CLASS(ShuffleLoyer); REGISTER_LAYER_CLASS(Shuffle);

} // namespace coffe



101





Vectorized operations













Vectorized operations





Assignment: Writing SVM/Softmax Stage your forward/backward computation!





Summary so far

- neural nets will be very large: no hope of writing down gradient formula by hand for all parameters
- backpropagation = recursive application of the chain rule along a computational graph to compute the gradients of all inputs/parameters/ intermediates
- implementations maintain a graph structure, where the nodes implement the forward() / backward().
- **forward**: compute result of an operation and save any intermediates needed for gradient computation in memory
- **backward**: apply the chain rule to compute the gradient of the loss function with respect to the inputs.



(**Before**) Linear score function:

f = Wx



(**Before**) Linear score function:

(Now) 2-layer Neural Network

$$egin{aligned} f &= Wx \ f &= W_2 \max(0, W_1 x) \end{aligned}$$



(**Before**) Linear score function:

(Now) 2-layer Neural Network

$$egin{aligned} f &= Wx \ f &= W_2 \max(0, W_1 x) \end{aligned}$$









(**Before**) Linear score function:

$$f = Wx$$

$$f=W_2\max(0,W_1x)$$

 $f=W_3\max(0,W_2\max(0,W_1x))$



Full implementation of training a 2-layer Neural Network needs ~11 lines:

| 01. | <pre>X = np.array([[0,0,1],[0,1,1],[1,0,1],[1,1,1]])</pre> | • 🔒 😡 |
|-----|--|---------------------------|
| 02. | y = np.array([[0,1,1,0]]).T | |
| 03. | syn0 = 2*np.random.random((3,4)) - 1 | |
| 04. | syn1 = 2*np.random.random((4,1)) - 1 | |
| 05. | for j in xrange(60000): | |
| 06. | 11 = 1/(1+np.exp(-(np.dot(X,syn0)))) Forward pass | |
| 07. | <pre>12 = 1/(1+np.exp(-(np.dot(l1,syn1))))</pre> | |
| 08. | 12_delta = (y - 12)*(12*(1-12)) | |
| 09. | <pre>l1_delta = l2_delta.dot(syn1.T) * (l1 * (1-l1))</pre> | Backward |
| 10. | syn1 += 11.T.dot(12_delta) | |
| 11. | <pre>syn0 += X.T.dot(l1_delta)</pre> | backprop of derivative |
| | | |

from @iamtrask, http://iamtrask.github.io/2015/07/12/basic-python-network/



Assignment: Writing 2layer Net Stage your forward/backward computation!

| <pre># receive W1,W2,b1,b2 (weights/biases), X (data)</pre> | | | | |
|--|--|--|--|--|
| # forward pass: | | | | |
| <pre>h1 = # function of X,W1,b1</pre> | | | | |
| <pre>scores = # function of h1,W2,b2</pre> | | | | |
| <pre>loss = # (several lines of code to evaluate Softmax loss)</pre> | | | | |
| # backward pass: | | | | |
| dscores = # | | | | |
| dh1,dW2,db2 = # | | | | |
| dW1,db1 = # | | | | |
| | | | | |
















* Original slides borrowed from Andrej Karpathy and Li Fei-Fei, Stanford cs231n







Be very careful with your Brain analogies:

Biological Neurons:

- Many different types
- Dendrites can perform complex nonlinear computations
- Synapses are not a single weight but a complex non-linear dynamical system
- Rate code may not be adequate



[Dendritic Computation. London and Hausser]







Neural Networks: Architectures





Example Feed-forward computation of a Neural Network

```
class Neuron:
# ...
def neuron_tick(inputs):
    """ assume inputs and weights are 1-D numpy arrays and bias is a number """
    cell_body_sum = np.sum(inputs * self.weights) + self.bias
    firing_rate = 1.0 / (1.0 + math.exp(-cell_body_sum)) # sigmoid activation function
    return firing_rate
```

We can efficiently evaluate an entire layer of neurons.



Example Feed-forward computation of a Neural Network



forward-pass of a 3-layer neural network: f = lambda x: 1.0/(1.0 + np.exp(-x)) # activation function (use sigmoid) x = np.random.randn(3, 1) # random input vector of three numbers (3x1) h1 = f(np.dot(W1, x) + b1) # calculate first hidden layer activations (4x1) h2 = f(np.dot(W2, h1) + b2) # calculate second hidden layer activations (4x1) out = np.dot(W3, h2) + b3 # output neuron (1x1)



Setting the number of layers and their sizes



more neurons = more capacity

* Original slides borrowed from Andrej Karpathy and Li Fei-Fei, Stanford cs231n



Do not use size of neural network as a regularizer. Use stronger regularization instead:



(you can play with this demo over at ConvNetJS: <u>http://cs.stanford.edu/people/</u> karpathy/convnetjs/demo/classify2d.html)



Summary

- we arrange neurons into fully-connected layers
- the abstraction of a **layer** has the nice property that it allows us to use efficient vectorized code (e.g. matrix multiplies)
- neural networks are not really neural
- neural networks: bigger = better (but might have to regularize more strongly)



Next Lecture:

More than you ever wanted to know about Neural Networks and how to train them.

