# Lecture 10: Training Neural Networks (Part 2)

Justin Johnson

Lecture 10 - 1



### Due Friday, February 11



Lecture 10 - 2

### Midterm

- Wednesday, February 23
- Will be remote as a Canvas quiz (most likely)
- Exam is 90 minutes
- You can take it any time in a 24-hour window
- We will have 3-4 "on-call" periods during the 24-hour window where GSIs will answer questions within ~15 minutes
- Open note
- True / False, multiple choice, short answer
- For short answer questions requiring math, either write LaTeX or upload an image with handwritten math



#### **1.One time setup**

Activation functions, data preprocessing, weight initialization, regularization

### **2.**Training dynamics

Learning rate schedules;

hyperparameter optimization

### 3. After training

Model ensembles, transfer learning

**Last Time** 

Today

Justin Johnson

Lecture 10 - 4

### Last Time: Activation Functions



Justin Johnson

# Leaky ReLU $\max(0.1x, x)$





**GELU**  $\approx x\sigma(1.702x)$ 

Lecture 10 - 5



### Last Time: Data Preprocessing



#### Justin Johnson

#### Lecture 10 - 6

### Last Time: Weight Initialization





Glorot and Bengio, "Understanding the difficulty of training deep feedforward neural networks", AISTAT 2010

#### Justin Johnson

#### Lecture 10 - 7

### Last Time: Dropout Regularization

In each forward pass, randomly set some neurons to zero Probability of dropping is a hyperparameter; 0.5 is common





Srivastava et al, "Dropout: A simple way to prevent neural networks from overfitting", JMLR 2014

#### Justin Johnson

Lecture 10 - 8

# **Training**: Add some kind of randomness

$$y = f_W(x, z)$$

**Testing:** Average out randomness (sometimes approximate)

$$y = f(x) = E_z[f(x,z)] = \int p(z)f(x,z)dz$$

**Training**: Add some kind of randomness

$$y = f_W(x, z)$$

**Testing:** Average out randomness (sometimes approximate)

$$y = f(x) = E_z[f(x,z)] = \int p(z)f(x,z)dz$$

**Example**: Batch Normalization

**Training**: Normalize using stats from random minibatches

**Testing**: Use fixed stats to normalize

# **Training**: Add some kind of randomness

$$y = f_W(x, z)$$

For ResNet and later, often L2 and Batch Normalization are the only regularizers! **Example**: Batch Normalization

**Training**: Normalize using stats from random minibatches

**Testing:** Average out randomness (sometimes approximate)

$$y = f(x) = E_z[f(x,z)] = \int p(z)f(x,z)dz$$

**Testing**: Use fixed stats to normalize

### Data Augmentation



#### Justin Johnson

Lecture 10 - 12

### Data Augmentation



#### Justin Johnson

Lecture 10 - 13

### Data Augmentation: Horizontal Flips





#### Justin Johnson

Lecture 10 - 14

### Data Augmentation: Random Crops and Scales

**Training**: sample random crops / scales ResNet:

- 1. Pick random L in range [256, 480]
- 2. Resize training image, short side = L
- 3. Sample random 224 x 224 patch



#### Justin Johnson

#### Lecture 10 - 15

### Data Augmentation: Random Crops and Scales

**Training**: sample random crops / scales ResNet:

- 1. Pick random L in range [256, 480]
- 2. Resize training image, short side = L
- 3. Sample random 224 x 224 patch

### **Testing**: average a fixed set of crops

ResNet:

- 1. Resize image at 5 scales: {224, 256, 384, 480, 640}
- 2. For each size, use 10 224 x 224 crops: 4 corners + center, + flips



#### Justin Johnson

### Data Augmentation: Color Jitter

# Simple: Randomize contrast and brightness





### More Complex:

- Apply PCA to all [R, G, B] pixels in training set
- Sample a "color offset" along principal component directions
- 3. Add offset to all pixels of a training image

(Used in AlexNet, ResNet, etc)

#### Justin Johnson

### Data Augmentation: RandAugment

Apply random combinations of transforms:

- **Geometric**: Rotate, translate, shear
- Color: Sharpen, contrast, brightness, solarize, posterize, color

```
transforms = [
'Identity', 'AutoContrast', 'Equalize',
'Rotate', 'Solarize', 'Color', 'Posterize',
'Contrast', 'Brightness', 'Sharpness',
'ShearX', 'ShearY', 'TranslateX', 'TranslateY']
```

```
def randaugment(N, M):
"""Generate a set of distortions.
```

```
Args:
    N: Number of augmentation transformations to
        apply sequentially.
    M: Magnitude for all the transformations.
"""
```

```
sampled_ops = np.random.choice(transforms, N)
return [(op, M) for op in sampled_ops]
```

Cubuk et al, "RandAugment: Practical augmented data augmentation with a reduced search space", NeurIPS 2020

#### Justin Johnson

#### Lecture 10 - 18

### Data Augmentation: RandAugment

- Apply random combinations of transforms:
- **Geometric**: Rotate, translate, shear

Justin Johnson

 Color: Sharpen, contrast, brightness, solarize, posterize, color



February 9, 2022

Cubuk et al, "RandAugment: Practical augmented data augmentation with a reduced search space", NeurIPS 2020

### Data Augmentation: Get creative for your problem!

### Data augmentation encodes **invariances** in your model

# Think for your problem: what changes to the image should **not** change the network output?

### May be different for different tasks!

**Training**: Add some randomness **Testing**: Marginalize over randomness

### **Examples**:

Dropout Batch Normalization Data Augmentation

Wan et al, "Regularization of Neural Networks using DropConnect", ICML 2013

### **Regularization**: DropConnect

**Training**: Drop random connections between neurons (set weight=0) **Testing**: Use all the connections

#### **Examples**:

Dropout Batch Normalization Data Augmentation DropConnect





Wan et al, "Regularization of Neural Networks using DropConnect", ICML 2013

Justin Johnson

Lecture 10 - 22

### **Regularization:** Fractional Pooling

**Training**: Use randomized pooling regions **Testing**: Average predictions over different samples

### **Examples**:

Dropout Batch Normalization Data Augmentation DropConnect Fractional Max Pooling



February 9, 2022

### Regularization: Stochastic Depth

**Training**: Skip some residual blocks in ResNet **Testing**: Use the whole network

### **Examples**:

Dropout Batch Normalization Data Augmentation DropConnect Fractional Max Pooling Stochastic Depth



Justin Johnson

### Regularization: Stochastic Depth

**Training**: Skip some residual blocks in ResNet **Testing**: Use the whole network

### **Examples**:

Dropout

**Batch Normalization** 

Data Augmentation

DropConnect

Fractional Max Pooling Stochastic Depth

# Starting to become common in recent architectures!

- Pham et al, "Very Deep Self-Attention Networks for End-to-End Speech Recognition", INTERSPEECH 2019
- Tan and Le, "EfficientNetV2: Smaller Models and Faster Training", ICML 2021
- Fan et al, "Multiscale Vision Transformers", ICCV 2021
- Bello et al, "Revisiting ResNets: Improved Training and Scaling Strategies", NeurIPS 2021
- Steiner et al, "How to train your ViT? Data, Augmentation, and Regularization in Vision Transformers", arXiv 2021



Justin Johnson

### Regularization: CutOut

**Training**: Set random images regions to 0 **Testing**: Use the whole image

### **Examples**:

Dropout Batch Normalization Data Augmentation DropConnect Fractional Max Pooling Stochastic Depth Cutout / Random Erasing



Replace random regions with mean value or random values

DeVries and Taylor, "Improved Regularization of Convolutional Neural Networks with Cutout", arXiv 2017 Zhong et al, "Random Erasing Data Augmentation", AAAI 2020

#### Justin Johnson

Lecture 10 - 26

### Regularization: Mixup

**Training**: Train on random blends of images **Testing**: Use original images

### **Examples**:

Dropout Batch Normalization Data Augmentation DropConnect Fractional Max Pooling Stochastic Depth Cutout / Random Erasing Mixup







Target label: cat: 0.4 dog: 0.6

Randomly blend the pixels of pairs of training images, e.g. 40% cat, 60% dog

Zhang et al, "mixup: Beyond Empirical Risk Minimization", ICLR 2018

#### Justin Johnson

#### Lecture 10 - 27

#### February 9, 2022

**CNN** 

### Regularization: Mixup

**Training**: Train on random blends of images **Testing**: Use original images

### **Examples**:

Dropout Batch Normalization Data Augmentation DropConnect Fractional Max Pooling Stochastic Depth Cutout / Random Erasing Mixup







Sample blend probability from a beta distribution Beta(a, b) with a=b≈0 so blend weights are close to 0/1



Target label: cat: 0.4 dog: 0.6

Randomly blend the pixels of pairs of training images, e.g. 40% cat, 60% dog

Zhang et al, "mixup: Beyond Empirical Risk Minimization", ICLR 2018

#### Justin Johnson

#### Lecture 10 - 28

#### February 9, 2022

**CNN** 

### Regularization: CutMix

Yun et al, "CutMix: Regularization Strategies to Train Strong Classifiers with Localizable Features", ICCV 2019

**CNN** 

**Training**: Train on random blends of images **Testing**: Use original images

### **Examples**:

Dropout Batch Normalization Data Augmentation DropConnect Fractional Max Pooling Stochastic Depth Cutout / Random Erasing Mixup / CutMix







Target label: cat: 0.6 dog: 0.4

Replace random crops of one image with another: e.g. 60% of pixels from cat, 40% from dog

#### Justin Johnson

#### Lecture 10 - 29

Justin Johnson

### **Regularization**: Label Smoothing

#### **Training**: Change target distribution **Testing**: Take argmax over predictions

### **Examples**:

Dropout **Batch Normalization** Data Augmentation DropConnect **Fractional Max Pooling Stochastic Depth** Cutout / Random Erasing Mixup / CutMix Label Smoothing



### Target Distribution

Standard Training	Label Smoothing
Cat: 100%	Cat: 90%
Dog: 0%	Dog: 5%
Fish: 0%	Fish: 5%

Set target distribution to be  $1 - \frac{K-1}{K}\epsilon$  on the correct category and  $\epsilon/K$  on all other categories, with K categories and  $\epsilon \in (0,1)$ . Loss is cross-entropy between predicted and target distribution.



Szegedy et al, "Rethinking the Inception

### Regularization: Summary

**Training**: Add randomness **Testing**: Marginalize over randomness

### **Examples**:

Dropout **Batch Normalization** Data Augmentation DropConnect **Fractional Max Pooling Stochastic Depth** Cutout / Random Erasing Mixup / CutMix Label Smoothing

- Use DropOut for large fully-connected layers
- Data augmentation always a good idea
- Use BatchNorm for CNNs (but not ViTs)
- Try Cutout, MixUp, CutMix, Stochastic Depth,
   Label Smoothing to squeeze out a bit of extra performance



#### **1.One time setup**

Activation functions, data preprocessing, weight initialization, regularization

#### 2. Training dynamics

Learning rate schedules;

hyperparameter optimization

### **3.After training**

Model ensembles, transfer learning

Today

# Learning Rate Schedules

Justin Johnson

Lecture 10 - 33

# SGD, SGD+Momentum, Adagrad, RMSProp, Adam all have **learning rate** as a hyperparameter.



#### Justin Johnson

#### Lecture 10 - 34

# SGD, SGD+Momentum, Adagrad, RMSProp, Adam all have **learning rate** as a hyperparameter.



# Q: Which one of these learning rates is best to use?

#### Justin Johnson

#### Lecture 10 - 35

# SGD, SGD+Momentum, Adagrad, RMSProp, Adam all have **learning rate** as a hyperparameter.



Q: Which one of these learning rates is best to use?

A: All of them! Start with large learning rate and decay over time

#### Justin Johnson

#### Lecture 10 - 36
### Learning Rate Decay: Step



**Step:** Reduce learning rate at a few fixed points. E.g. for ResNets, multiply LR by 0.1 after epochs 30, 60, and 90.



#### February 9, 2022

#### Justin Johnson

Lecture 10 - 37

### Learning Rate Decay: Cosine



Loshchilov and Hutter, "SGDR: Stochastic Gradient Descent with Warm Restarts", ICLR 2017 Radford et al, "Improving Language Understanding by Generative Pre-Training", 2018 Feichtenhofer et al, "SlowFast Networks for Video Recognition", ICCV 2019 Radosavovic et al, "On Network Design Spaces for Visual Recognition", ICCV 2019 Child at al, "Generating Long Sequences with Sparse Transformers", arXiv 2019

**Step:** Reduce learning rate at a few fixed points. E.g. for ResNets, multiply LR by 0.1 after epochs 30, 60, and 90.

Cosine:

$$\alpha_t = \frac{1}{2}\alpha_0 \left(1 + \cos\left(\frac{t\pi}{T}\right)\right)$$



#### February 9, 2022

#### Justin Johnson

Lecture 10 - 38

### Learning Rate Decay: Linear



**Step:** Reduce learning rate at a few fixed points. E.g. for ResNets, multiply LR by 0.1 after epochs 30, 60, and 90.

Cosine:

$$\alpha_t = \frac{1}{2} \alpha_0 \left( 1 + \cos\left(\frac{t\pi}{T}\right) \right)$$

Linear:

$$\alpha_t = \alpha_0 \left( 1 - \frac{t}{T} \right)$$

Devlin et al, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", NAACL 2018 Liu et al, "RoBERTa: A Robustly Optimized BERT Pretraining Approach", 2019

Yang et al, "XLNet: Generalized Autoregressive Pretraining for Language Understanding", NeurIPS 2019

#### Justin Johnson

Lecture 10 - 39

### Learning Rate Decay: Inverse Sqrt



**Step:** Reduce learning rate at a few fixed points. E.g. for ResNets, multiply LR by 0.1 after epochs 30, 60, and 90.

Cosine:

$$\alpha_t = \frac{1}{2}\alpha_0 \left(1 + \cos\left(\frac{t\pi}{T}\right)\right)$$

Linear:

$$\alpha_t = \alpha_0 \left( 1 - \frac{t}{T} \right)$$

Inverse sqrt:

$$\alpha_t = \alpha_0 / \sqrt{t}$$

Vaswani et al, "Attention is all you need", NIPS 2017

Justin Johnson

#### Lecture 10 - 40

### Learning Rate Decay: Constant!



**Step:** Reduce learning rate at a few fixed points. E.g. for ResNets, multiply LR by 0.1 after epochs 30, 60, and 90.

Cosine:

$$\alpha_t = \frac{1}{2} \alpha_0 \left( 1 + \cos\left(\frac{t\pi}{T}\right) \right)$$

Linear:

$$\alpha_t = \alpha_0 \left( 1 - \frac{t}{T} \right)$$

Inverse sqrt:

 $\alpha_t = \alpha_0 / \sqrt{t}$ 

**Constant:** 

 $\alpha_t = \alpha_0$ 

Brock et al, "Large Scale GAN Training for High Fidelity Natural Image Synthesis", ICLR 2019 Donahue and Simonyan, "Large Scale Adversarial Representation Learning", NeurIPS 2019

#### Justin Johnson

#### Lecture 10 - 41

### How long to train? Early Stopping



Stop training the model when accuracy on the validation set decreases Or train for a long time, but always keep track of the model snapshot that worked best on val. **Always a good idea to do this!** 

#### Justin Johnson

#### Lecture 10 - 42

Justin Johnson

Lecture 10 - 43

### Choosing Hyperparameters: Grid Search

Choose several values for each hyperparameter (Often space choices log-linearly)

#### Example:

Weight decay: [1x10<sup>-4</sup>, 1x10<sup>-3</sup>, 1x10<sup>-2</sup>, 1x10<sup>-1</sup>] Learning rate: [1x10<sup>-4</sup>, 1x10<sup>-3</sup>, 1x10<sup>-2</sup>, 1x10<sup>-1</sup>]

Evaluate all possible choices on this hyperparameter grid

Justin Johnson

Lecture 10 - 44

### Choosing Hyperparameters: Random Search

Choose several values for each hyperparameter (Often space choices log-linearly)

#### Example:

Weight decay: log-uniform on [1x10<sup>-4</sup>, 1x10<sup>-1</sup>] Learning rate: log-uniform on [1x10<sup>-4</sup>, 1x10<sup>-1</sup>]

Run many different trials

### Hyperparameters: Random vs Grid Search



Bergstra and Bengio, "Random Search for Hyper-Parameter Optimization", JMLR 2012

Justin Johnson

Lecture 10 - 46

### Choosing Hyperparameters: Random Search



Radosavovic et al, "On Network Design Spaces for Visual Recognition", ICCV 2019

February 9, 2022

#### Lecture 10 - 47

Justin Johnson

(without tons of GPUs)

Justin Johnson

Lecture 10 - 48

Step 1: Check initial loss

Turn off weight decay, sanity check loss at initialization e.g. log(C) for softmax with C classes

Step 1: Check initial loss
Step 2: Overfit a small sample

Try to train to 100% training accuracy on a small sample of training data (~5-10 minibatches); fiddle with architecture, learning rate, weight initialization. Turn off regularization.

Loss not going down? LR too low, bad initialization Loss explodes to Inf or NaN? LR too high, bad initialization

Step 1: Check initial lossStep 2: Overfit a small sampleStep 3: Find LR that makes loss go down

Use the architecture from the previous step, use all training data, turn on small weight decay, find a learning rate that makes the loss drop significantly within ~100 iterations

Good learning rates to try: 1e-1, 1e-2, 1e-3, 1e-4

Step 1: Check initial loss
Step 2: Overfit a small sample
Step 3: Find LR that makes loss go down
Step 4: Coarse grid, train for ~1-5 epochs

Choose a few values of learning rate and weight decay around what worked from Step 3, train a few models for ~1-5 epochs.

Good weight decay to try: 1e-4, 1e-5, 0

**Step 1**: Check initial loss

- Step 2: Overfit a small sample
- Step 3: Find LR that makes loss go down
- **Step 4**: Coarse grid, train for ~1-5 epochs
- **Step 5**: Refine grid, train longer

Pick best models from Step 4, train them for longer (~10-20 epochs) without learning rate decay

Step 1: Check initial loss

- Step 2: Overfit a small sample
- Step 3: Find LR that makes loss go down
- **Step 4**: Coarse grid, train for ~1-5 epochs
- Step 5: Refine grid, train longer
- Step 6: Look at learning curves

### Look at Learning Curves!



Losses may be noisy, use a scatter plot and also plot moving average to see trends better

#### Justin Johnson

#### Lecture 10 - 55



	isti	in I	hn	SO	n
J U	13C			30	

Lecture 10 - 56



Justin Johnson

Lecture 10 - 57



Justi	in J	0	hn	SO	1

Lecture 10 - 58



time

J	lu	st	i	n	J	0	h	n	S	0	n
		<u> </u>		•••				•••	$\sim$	$\sim$	

Lecture 10 - 59



#### Justin Johnson

Lecture 10 - 60

No or small gap between train / val means underfitting: train longer, use a bigger model, maybe higher LR Train

time

Ju	stin	Jo	hns	son

Accuracy

Lecture 10 - 61

Step 1: Check initial loss

- Step 2: Overfit a small sample
- Step 3: Find LR that makes loss go down
- **Step 4**: Coarse grid, train for ~1-5 epochs
- Step 5: Refine grid, train longer
- Step 6: Look at loss curves
- Step 7: GOTO step 5

## Hyperparameters to play with:

- network architecture
- learning rate, its decay schedule, update type
- regularization (L2/Dropout strength)

neural networks practitioner music = loss function



#### Justin Johnson

Lecture 10 - 63

### Cross-validation "command center"

Conservation Conservation												
accelere 3 tant menerit, HL1-Dr	workers 2 more later. later wateries 2010-00-	Service 1 Instant and And generaly (2) 7-21-	marking 4 marking and	Sector 1 Constitute out	Angelene & State State	Secritory 7 Last spinters (01)-11-	Sector 1 Percentions Percention 201-01-	angler B mariners Lab antered (021-21-	anglas 11- anglas anglasis 2012-11-		analyry 12 man and har winten (0.+-1	workers 1.3 mart 1.000 Dec montato 2020-01-
manda Lik. HL/HD Fine: SHORA-CHE/THP	month 29-33/10 First 189787/302004	Marrie B. BADBI	manuto 20. 12/100 View 12/114/20/2010	ments 5. 17/00 piner 1000-00000	Numb 12, HJ/W	ments 10, TUIN Lines 40,040,040,700	mpanh 12. 16/34 Line 2018/11/0/10	Agent 12 12/10 Line Tribulation	apath 7 Ab/Hi blas DORTINITO		manh 3. 62/30 hise 20070122300	Inguist 14. Ph/NU Inguist 14. Ph/NU Information 2008051 223000
ten, rela annales", 5, Sanaest's Soull', "lange estadou alte": 20, "soul participati" de	Press, replay according to it. Automated in National Contraction and the State of the Automatical State of the State of the Automatical State of the Automatical State State of the Automatical State of the Automatical State of the Automatical State of the Automatical State of the Automatical State of the Automatical Stat	Anny paint property in Appropriate Property Tampy providing score?) (201, "score paint) and ()	New role members, i. Names's New?, "Samp seconds, and"; and, "sould party magnification	Print rolls with merch 1, Parametric Print Parametric Print Print, Parametric Print, 2010,	Press, relia, annotatif's 31, "annores") Printella, "Senae annotation start"; 2011, "Senae (Senae)	Peter only manager?'s 1, Pariager?'s Percept, "Single Accessing start"; 201, "Accessing and Single?" -4;	Prote price processor's 1, "approach", Prote', "press According 1997's 1991, "Note that conduct", 4,	Party self-a mandeter 1 (h. Farrantel- Party), Campa Secondary (1997) 201, Nation (2007), 241	Print rolls described in . Nanoann's Print/, "Swapt monthing start"; print/start/start/starter"; dr. Noted Start/Starter"; dr.		Processing recently in Parameter's Provid, Charge recording (0.97) Phys. Review (0.07) (1)	Press, and a manufacture is a Parameter's Press P. Transp. metericity with P. Pris, "Same data (magar") sta
terre excellence excellence of the	Name production and the second	Party sectors (sector) 10.	Parage processing accessing to the later of	Print Har Jacobi's 'S Ja-	Participanti anno 1971 - 1974 Participanti anno 1977 - 1974 Participanti anno 1977	Party and another to a sur-	Annual Annual Control of Control	Party production of the line o	Parage processing store 1 to 1.		Parameters and an and a second	Party professional and the second sec
Salar, "generation") Tater, New Jer, H. Henner, M.	Plant, "provident": "late", "ent_pla_st_atort": my	Plater, "provident", "later, "Son, St., Lt., and "L. By Analist Marches",	Time", "process": "Los", "en.go., r., act " 0, "and concern".	Area Jords and 1 (3), "Imperial"	Tand, Tananatat's Fanal, San an an and the day	Tanat, forestation (state), have been as another the	Tand, "percents" Comp., San an an annu an	Planet, "president for "inter", New Jos, et., and "r way Planet, water of the	Planet, "anterior", "late", National and anter the		Tanif, Yarmanator's Cold's National Action in the	"Land", "promine"( "land") "Land", examine" ( "land") "Land", examine" ( R)
Contraction and the second of	Number of the second se	Nucley Decision (a) December(*) 14. "direct sectors (2.40) New Decisions", 1. Transmission, (*)	Netter (Netherland, part Alterior)P1 24, "Mary network U.S. New Seed analy 1, Names analy	Paulo recent 1.	Parties (method and interestor) (1), "direct and" (1) (1), (2), "direct and" (1) (1), (3), "direct and" (1), "states and")	Partie chevapoint and threesed?"; 11, "diag ment" 1/12, Part lines exactly 1, "house, south	Partie declaration par declaration?"	"stills declarate ant deservation":	Party (metastel, pp. threshold), 1, "Marg (2007) 1, 400, The long number, 1, Weiner, 2107)		Native discharged and discount Pri- by Anna Sarah 1, Station and	Printer, description, pp1, monotories, etc., "Markay result", 2, 21, Print, head result, 2, 75, 76, 86, 11, 27,
<ol> <li>"Network" I.L.</li> <li>Status, et.al., saturation, descenter, Pro- templity, Protecting, conv. Security, 199</li> </ol>	Fig. "Something") A. Saminar shares yourse, doing things (doing of Neuroscience States).	All, Network 1 5 Network (Series, Stream, Blackery) Possil, Schemering (Series, Sample)	<ol> <li>Second Annual States (Second States)</li> <li>Second States (Second States)</li> <li>Second States (Second States)</li> </ol>	Pain real score 1, "Anima and") hat, "Animarca", aj Peterer charar science disectory".	<ol> <li>Nesserver, S., Nesserver, Hand, Steppin, Nesserver, P. P. Nesserver, Nesserver, Stepping, Nesserver, Nesserv</li></ol>	11. Network 1, Network (Method, 1979) American Press, elsever, second and the Press, and the second and the second second and the second se	(1), "Anisotrop": 1, "Anisotrop (Market, Salings), Anisotrop," "Anisotrop (Market, Salings), Anisotrop, " "Anisotrop (Market, Salings), Salings,	<ul> <li>"Non-traffic Linear Annual States", Construction (Construction), Construction, Construc</li></ul>	<ol> <li>"Non-color", 5,</li> <li>Normer, charter, school, documents/n.</li> <li>"Non-color contents, new lower</li> </ol>		Mr. Network, 5, Nether Joans, setter directory Constantion Magnification (Magnification)	101, "Reserve": 1, Parsar starts, inipat, Armitely": Press, whereas inipat, armitely
erthet provides and de providente anti-	Family stand, study 1, 44, Family stand, study 1, 44, Flamming, samphy 0, 30077,	Aunoration and the contract of	Provide Distant, **	Parent Action (1997) 101-10404	Provide and a second state	Print Man (Mar) H.	Parate and Analysis at	Provide and a state of the stat	Plant mint, but's w.		Planet many hole?	Point mant, traph Pr. Point mant, traph Pr. Pintering start, 1, 10100,
send on the set of the	"Condition Conditional (Const. Magnetic States Const. The second scheme "Con- Parent, young Advantage (PC ),	<ul> <li>Available of the part of the strapped description of the strapped of the strapped of the Part of the strapped of the strapped of the strapped of the Strapped of the strapped of the strapped of the strapped of the strapped of the strapped of the strapped of the strapped of</li></ul>	<ul> <li>Sensitive Memory Proc. Heads, Beneficiary, Ann. Anal. Sensitive Ta- Parent Start Memory Proc.</li> </ul>	Constraint, strain, networks pro-	Press and stated in the state	" transporter and a resident to the second s	Contract of the party of the other party of the party of	Provide the second providence of the second	President of the party of the object of the second		Process of the process magning of the second	Printer, with model (1995, 1999), fermities (1995, 1995, 1993), short 1, 1, Print cost, Wanter, 1975, 1,
And staffs side frage's science frances age's level, forces to appropriate and pattern science	Participant's de Participantes Participantes de la Participante Participantes Participantes Participantes	Pageod, analysis, "sape"s films, Pageod, sape"s level, "science"s Pageode, sapes, second and	Participanti da Assorti concela Secon esta con concela Secon esta con concela	Participation (Constituted to be a second state of the second stat	Participation of the super-	Sarah and and she have a second	Participant's set, "separations," - "anness sents set, "second", - "anness, second set, "second", - "anness, "second", second", 5.1,	Preside starts the "super- schedule Preside sprite trib, "subsche President", Neuer particit", hol-	Paris, and Link, Super Link, 5, Smooth and Link, Society, Support, New periods 115.		Participant's inc. "Sapir's Calanti, Security and a local, Soliter's Streamp's, Neuropathon's 1.0,	Parent, search in, "sepret both to feedbal apple and, "second- temping", "search second-
er bere bestehnt? "Ulfin offen Treete gabt VL.STW	tion per introducts. Sile pinnis Table 1971 - 34-2100	trans pro. heteroical a difference and a second a same and a second a secon	tion are introductly blir internet to the second state and a second state and a second	These processing and the second secon	Lane pre lands (L.12hr angele 1944), ppr 117,000	tion per historich', 142r - mentic track ppr. (1.00)	Down you handwide, 20, 30 kg	tion pro Antoinill, 1900 manufat, 1844 (pp) - 57.478	ther are herein't, blir		tion pro interview. 1404- tion pro interview. 1404- statility (1444-000) 14,440	Law and Anthony III.
51 11	-		-1 1111	117.812	-1.00.00	-1	-1					-1 11
-	2	5		- Unceller	-	1	-		-		3	4
1				-			-				-	
	Landau II	IL.	IL	1. and the theory of the track		Marchael 11	IL	Contract 11		J.	Marchard II.	Marine Str.
An annexi atta-ta-	barty Land. barty destants, 2011-00-	And writes in the second	Part anness (http://	inny late	tert mittert (ND-D-	Last manuals (001-0)-	Printy Junio Terri Multinis 2012-01-	tanta ante lasta meneral (001-0)-	tary apprends 2010-12-	Contraction of the contraction o	Last appress 200-12-	Land under Child
mit 2.24/80 mit 98295/30000 mit 98295/30000	mponin 12-24,000 Fines 12+8201/30/0044 1 **/rist, r1-0/** (75)	Specify 20. 84/00 21ms 00951/045/90 1 ********	Report 22, 94/00 View 70011/240/700 5 **print_st_0**0_0**0	Investi 24, 74,781 View 149555/20080 1. *(2401,111)*) 121	Samoit LD. 52/00 Liber 20050/06/200	Apart 18.42/80 Line (*1960)40/90 *plat_111(*: 1	International Control of Control	Appende 20. Million Appende 20. Million Applied, 2010/2010	April 1, 14740 Line 2740/142700	Manufa h. 242/88 Line: 40-02/88801 - *2540, cruzt*1 100	Applicate 20 Applied to 10 App	apach id. NOW Line MERCIESINI 1. "plat. 114271 101
No. 1012 Normal Print, N. Narrashiri, 2017, "Since exception star": 20, "Note, NaCobate?", 21, 21, "Sole, NaCobate?", 21,	Press, result descenter's to Pressent's Press, "Press, descenter and the state", 2011, "Annual and the second state of the	Trans, "Same wooding size") and, "Same balling size") and, "same ball the set	Protection of the second secon	Prove record provident's by Parameter's Prover', "Sease securities stat's http://www.securities.com	North Annual Annual State (1997)	Proof with the second second staff,	Proof, "Tange Accelling and", Proof, "Tange Accelling and", Bill, "Accelling and ",	Population and an appropriate and a second s	Provide and the second state of the second sta	Plink Act, manually is, Passan's Plink TP, Plank Action (1994)	"Claim Dir, "Claim Dir, "Claim Dir, "Chain, attenuit, actar", may	Provide, "lange secondary size"; http://www.usc.com/organ/"; http://www.usc.com/argan/"; http://wwww.usc.com/argan/"; http://www.usc.com/argan/"; http://www.usc.com/argan/"; http://www.usc.com/argan/"; http://www.usc.com/argan/"; http://www.usc.com/argan/"; http://www.usc.com/argan/"; http://www.usc.com/argan/"; http://www.usc.com/argan/"; http://wwww.usc.com/argan/"; http://wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww
nal decidera starte tito, las aportes foi religion estarte tito fragmete	Next Accelling 1994's 101, Next Accelling 1994's 101, Next (Action (Accel 1 101), Next 1997)	Parat, section, cost, etc.	Personal and a second state of the second stat	Parent exceding screet, etc. Parent incomercy for Perent incomercy screet's list, "Segment's	Party and an article start with the second start of the second sta	Party sectors and 1 Mile Party and a control for Party and a control of the "Standard"	President and Color Color Part Appendent for President and the Programmer's	Provident States and a state of the states o	Paral sectors read ( 0), Paral (active) (b), Paral (active) (c), (c), (c), (c), (c), (c), (c), (c)	Prices, prof. (montes?), 5 (b), Accel, accelling, cont?), 50, Page, (accelleng, cont?), 50,	Parage peril merchel (* 1.5.), Parage peril merchel (* 1.5.), Parage perilet (* 1.5.)	Party modeling starts (1997) Party Jacob (1997) (1997) Party Jacob (1997) (1997) (1998)
tern", "personal"; "late"; not per in parater on table viewint"; 1,	Plant, Special Control (1996), Plant (1997), Address (1997), Plant, Presidents (1997)	Plant, Spectral States, Plant, Spectral States, Spectral Plant, Plant, Spectral States, Spe	Paral, "providers", "Lord, "encipe of work", "Cont. Practice encodes to	Party of Article 1	Cardel, "percentaria funda", Sena Jan, er andre funda	Party Construction (1998)	Party Parents Print Street, Parents Print Barry, Parents Parents Print Barry, Street, Parents Print Barry, Parents Parents Print Barry, Parents Pa Parents Parents Par	Parents, Parentset's Cateria New Jos of America's Ma Parent Constant's 1,	Part provide and a second state	Providence of the second secon	Presidential and a second second second	Parent Parents of Active States
C. National Joseph T. 2005, non-tend secondly, 22, "Relation-period"s 12, "Networkship 12,	11, "Million (MARCA S. 1995) New press works S. Weinser, scient's err, "manufactory L.	<ol> <li>"direct stars" (L-dit),</li> <li>"rest, incom", L. "Inside cost",</li> <li>Int. "Assessed", L.</li> </ol>	<ol> <li>"Many netw" U-Mit.</li> <li>Ten networks L. Nickler, post's 101, "Statement", 1.</li> </ol>	11, "Alter card" 1, 475 Free deep server 2, "Ender card"s Alt: "Seesand" 5,	11, "Bits provide to the first man seen"s 5, "Instances"s 101, "Name of S.	11, "Miles (1972") 1, 1975. The loss starts (), "Science, conth. 201, "Ambieval") 1,	11, "Million State" 1, 18, Team State Acad", 1, "Science, acad", 101, "Number of St.	1. "Ming car", 'Child, 'ne line cor's L. "Same cor's its, 'Same of L.	<ol> <li>Maray (2017) 1.401</li> <li>Solar (2017) 1. Vision (2017) 1.</li> <li>Solar (2017) 1.</li> </ol>	"Averal Antenna (pt. Scontor)" 10, Mean Art's 1, 40, "Sea, deel, Scot", 1, "Scotter, cont"	Partia, Antopoli, pp. terminal*; 10, Minay and*; UHL Pain And Anto*; 1, "States, star";	14, "Birds, Straff, L. 201, Print, Intel Straff, L. Burghen, March, 101, "Networks, L.
Animal Andreas (Santas in April 1977) Animal Andreas (Santas Internet American Animal Animal Anim Animal Animal	Provide products (second second secon	Annual scheme, solare, and solar scheme (	Nervice, shared, single, determine(*) * result is formation or chapt determine (invice)	Parriar, Charles, Ladyar, Branchar, Pr. Printing, C. Stanger, Barrier, Charger, Parrier, Stanger, St	Particle (Mariae , Letter), Robert (Mariae), P. (1997),	Person plana, ranget, Arectary, "	Provide phartic integral detection (***	Provide printing being an and an and a second secon	Provide phases, pages, and an organized and a second secon	Particles, Platter retent, Southern's, Particles, Platter retent, Southern's, Platter, Status, Southerness,	Provide and states distances	Partial pharts (integer, Arminager) Press, whereas integer Rest(integer)
Description and a L. Stephin, Description, and and dispersion of the	Comparing sample Laboration (*)	Presenting approved to Milling	Concerning activity 0.2007000	Conversion, and "1 - below, incomparing the second second second	Conversion assorb 1, 20171, 191711, 19171, 19171, 19171, 19171, 19171, 19171, 1	Company and A. L. SOUN, "Description, company, and and company"	Conversion sector's 1, 50000's	Committee server's 2, 500 mm, "Committee", origin, Alternative", Committee, origin, Alternative",	"Secondariant and and a second and a second	"Depending sport", L.MONG, "Consequently, surger, monotory") Consequently, surger, monotory")	"increasing serve", is billion, "increasing of server, increasing	"interior cone", 1,10100, "interior in output products")
energies (er., No. and second) is, and state dimension (*) N. appl, state (b, "taget (s, state)).	Manufaction, National Action, 5, National Accession, 5, National Accession, 5, National Accession, 5, National Accession, 5, National Acc	Section 107, The and storts, L. Name man showing the Name and States and States	Manufaction", "No post ment": 5, Nation must departicip?" N. Papel, and P. 118, "Apply 1, Jun-	menutation of the analysis of the Paral State of the Paral State of the Association of the Paral State of th	Participation of the seal desire in Sector case, description in Nature and in the sector of the	President (or C. Not, and court's la Noted court, dissociative N. Naryh, essent on Nary's Losso.	management of the part should be be be been as a straight the straight of the	Party in the set of th	Party man disantial's 5. Particle scales (10, Pargets 1) Pa- ty Particle, gars) and Pargets	Parent control of the parent control of the Parent count, this and off of the Parent, county comp, the parent, strength	Manufacture, No. and Smith 5, Note coast description, 5, Parent, and 5 and 5 and 5 and 5	Managine (1976) Pala anali sherif ta fa- Palat costa (tombalaffa 'b) Parat (tombalaffa (b), 'anget (tobach),
America Apple 1975, Package 1971	Company's terms participation (	H. "Among app", 2018, Particular Pringersp", "Among particul", 2 (2), Paragraphic annual Particul", 2 (2).	L. Canadita and S. Seria, Physics 7, Computer 7, Neural parameter 5, 5, 5 Parameter production of the second system Parameter system (second system), parameter 5, 5 Parameter 5, 5 Par	ParagongP, Netal participer (1), ParagongP, Netal participer (1), ParagongP, Netal participer (1), Netal (1), ParagongP, Netal (1), ParagongP, Netal (1),	Property and according to the second	<ul> <li>"Second county lards, function").</li> <li>"Second production of the second production of the second production of the second product of</li></ul>	Paragraphy, and paragraphy in a	Proping (and parts lark, fully and the second secon	Property", "New period", Life, "Stag processing", 0,01 Tam per second 2.155	<ul> <li>Company P., News, pattern V. P. 1997, News, pattern V. 1997, News, pattern V. 1997, News, News,</li></ul>	<ol> <li>"Sector apply level, fuller/s released, Neurophysics, Proceeding, Neurophysics, Sector 2, 1971</li> <li>Patrick, Print, Sector 2, 1971</li> <li>Patrick, Print, Sector 2, 1971</li> </ol>	Paragraph, Annual Annual (
12.454	13.013	22.458	23.650	12. CHA	TIL 483	51.678	THE REAL AND ADDRESS OF ADDRESS OF	DI THE	50.931	215.484	14-145	THE SUS
\$L 1111	1	11	51 11	211 11 1	711111	31 F	211 1 1	-1	E Van		-4	-1111111
-		8				3	2		-			
S	the search and a		Accession						Tenengine in	Farmenter	يديد مرموني و	set at a set
Arber 13	worker 28	worker 28	socker 38	sortar 31	worker 32	sociar 11	socker 34	sorbar 35	worksz 36-	worker 37	scolor 38	sorier and
roin, in. to vertice parts to the term	aporto 14, 21, 20,400 aporto 18, 26,700 anas 19120,000000	Anna Hanalan	spint to taits and the taits	apach 15. 85/10/100	Sport an tarte	aparts 5, 55,751	appende all'Antonio	apoin 7 MUSELSON	again in animi	Aparts of the second second	aparts & Storts Aparts & Storts	Amer Medicing Law
Name	1 "April 120" D. True rate account?" In "Account", Name, Account account of the Name, Account of the Name, Account of the Name, Name, N	<ol> <li>Same clarity to</li> <li>Same calls, annual devices on the same "i</li> <li>Same Control of the same calls of the same "i</li> <li>Same Control on the same calls of the same call same calls of the same c</li></ol>	<ol> <li>Agried, ed. (P. 10).</li> <li>Perror, and a maximum Proc. Ps. Photoscill<sup>19</sup>1.</li> <li>Perror<sup>1</sup>, "Simple, second on products."</li> <li>Perror<sup>1</sup>, "Simple, second on products."</li> </ol>	<ol> <li>Appel (197) 5.</li> <li>Press, pala associativity in "factoria": Pressil, "Seath microling starts".</li> <li>Pressil, "Seath microling starts".</li> </ol>	<ol> <li>Append (Charles Science), No. Substanting Proceeding and approximation of the Processive Science and Appendix of the Science Science Science and Appendix of the Science Science and Appendix of the Science Science Science Science and Appendix of the Science Science Science Science and Appendix of the Science Science Science</li></ol>	<ul> <li>New Control (1),</li> <li>Provide Control (1),</li> <li>Provide Control (1),</li> <li>Provide Control (1),</li> <li>Provide Antipolitical Control (1),</li> <li>Provide Antipolitical Control (1),</li> </ul>	[2] Apost schifts 1. Pres. (all: another/rt u., "spinster," Presil, "Image providing tasts", Presil, Comp. providing tasts", 2011. Apost and Internal Conf.	C. Annual and Anticipation of the Physics of the Processing Anticipation of the Physics of the Processing Anticipation of the Physics Processing Anticipation of the Physics of the Phy	<ol> <li>April (1)271 27.</li> <li>Transition and appropriate to the Statistical Statistics (Statistics), Comparison and Statistics (Statistics), Comparison and Statistics), Comparison and Statistics, Comparison and Statisties, Comparison and Statistics, Comparison and Statistics, C</li></ol>	<ol> <li>Annel (1277) 1;</li> <li>Pasa, calls printerior(1) 1; Papeagite(1) Print(2), Pasaga printerior(1) 1;</li> <li>Paraga printerior(1), Papeagite(1), Pap</li></ol>	<ul> <li>Next clarts to</li> <li>Peter of a second circle to</li> </ul>	<sup>1</sup> Speed (Carl) 20, <sup>1</sup> String particular description, <sup>1</sup> String and <sup>1</sup> String <sup>1</sup> String <sup>1</sup> String (String String (String <sup>1</sup> )) <sup>1</sup> String <sup>1</sup> String (String String <sup>1</sup> )
the print action of a day, and according at a day, an appendix by	Party procession starts they been souther starts the	Party production starts they been produced as a starts they been products in the	Party plan Andrew's to by Party mending starfs 100, Party another 10,	Party print had and a re-	Name proc. And Adv. 1. A.	Parat das page 1 - 1, Parag productor ( 1, 5, Parat productor ( 1, 2, 5)	Party processing scores was been provided to be	Party production with the state	Party proclassing score's min-	Parage prior Section ** 1974	Parage print (and had Print in A. Parage providing strends ( 100), Parage print ( 10),	Paragenet, and an Provide Tar. 6, 5 Paragenet monitory strends with the paragenet of the pa
net same unart on, "separate net, "personal"; "land", in pol in short"; N.	Provid Lanced, and an article of the state o	Plant Art Stranger 1 100 Plant -	Print and the sector's steep, "Rapping", Print and "Spatiation and "Strange", "Non-policy and advective the	Press Laster, press of the Parent's the	Placed, Spreader's State, Spreader's State, Spreader's State, Spreader's State, Spreader State	Paral Jacob and 1 201, "Engend") Paral Jacob and 1 201, "Engend") Paralleline, "A spectrum", "South	Provide Control and Provide Control Processory - Procesory - Processory - Processory - Processor	Press, barrier, trans-r and, Preparate- Press, Spectroscoper Fonder, Press and on short's 25,	Planet, batter, and " the Planet", Planet, "selectors", "cate", Planet, is short", To.	Provid Roboth, 20067, 200, "Papparent", Planof, "permittant", Ruda", Pate apt, or disert", 20, Provide conditions, 1	Paral Justin, study 100, "Pagasor" Fared, "pasacarar"s flands, Palo dat, or dent?) 35,	Print, Sales, and Trans. Programmers Print, Parlamenter's Rootf's Print, Mr. Nr. America, Str. Print, Mr. Marketter, St.
Disconstruction and support of the second se	Another transmission and transmission of the second	Partic contacts and thread Pr- 11, "Area and a", 1, 490, "And manufacture", 1, "Areas, proof."	Parata disebut and threadout? 11, "Secur cond": 2,905, "An man, load", 1, "Scane, cond".	Person compared and compared in the second s	Parties constantiat and threased Pri- bit, "Brief ratio" U. Dr. Not. Sant, Store 4, "Standing rate"	Name station's 1, Notice characterized part threatering's re, "Million characterized to 400,	Partice conductor and threaders/Pi- th, Neural condition (C), Neural Joseff 2, Neural condi-	Person interference and streams off- it, "Solver bear", 7, 55, "etc., dash, Doort, A, "Solates, proof-	Period Antipactal and Antipactal Pro- bly Mercan and Colli- Party and Dearth 1, Minima and	Percin connects and connectPrint, Microscoper, 1997.	Particle characteristic and constants?"- 15, "Neuron Const?", 1, 983, "The Association", 1, 983, "The Association", 1, "Science, 1997"	Percip Anotherate path threathad the D. Marce another D.M. Percipant International Control of the
L. "Newsylver's L. diller, Marcan, estanti, al revenue"r anal i furnargabity cost, mass.	ADL, "Appendings"s, L. "Appendix allocation operation of execution"s - consect (according over consects)	ADL, "Restance", L. Perceted, Analise, Science, Miller South C. Provide Conference of	ESL, "expected": In Sector statute added, directors": "reveal for temperature and statute."	Product of Alling Street Stree	Contractory L. Contractory and the second se	President study L. Solans, starts OL. Schleright L. Sector, starts, starts, assessed	Anthe Mater (1998) Resident's La Contraction (1998) Anthe Contraction (	<ul> <li>Period Anna Antara Strategic</li> <li>Period Anna Antara Strategic</li> <li>Period Anna Antara Strategic</li> </ul>	Porte Alexandre St. St. Streetwart-	No. "execute": 5 Souther, status, school, 6(sectors)": "assession to be \$2.50 millionage.	N.L. "Nearcost": L. Neartes status satural interitoria": " analytic languages tell single.	Problem and the subject distances "
CO. Autori, Drug" 1, 77, autority, 10147 1, 1010000, Designer, 10147 1, 2010000,	Parts Andel (1997) **.	Party April Don't The	Proprietory (2007) 7%, Providence (2007) 1, minute, Description and an approximate.	Plant Armit (Car") **, *saading (Car") * Linearin, Postalari (Cargor Cardinari)	Contracted Annalis P.	Printing shart linear	Panel Andel (1947) **, Samilare (1947) S. marti, Northering of science alcosters?;	"suit sout fra" . ". "Lasting time, is more,	Party Armed Store": P., Planting Store": C. Smith, Provident of second store (arms)	Front gradet from*1. **. Frankling Gradet s. minute, Frankling Gradet states	"Analitication and an anti-	Picci milet Non?' M. Samilton sile?' Lambais, Possignist migns tipester?'
and trendently from them- optimation, which prior reactive is, we present threaders of a se-	<ul> <li>Antal Conference (1) Theory and a standard start (1) and a standard start (1) and a start (1) and (1) and</li></ul>	President and the second secon	Press (1976) and a line couple- many many only. And plan, many by pre- parent, press, the balance Pro-	Press of the second sec	Press Contract of the state of the second se	Particle or bigod to be could be and the second control of the part of the second could be a second of the second could be a second could	Provide the second seco	Provide the second seco	Provide the second state of the second state o	Provide the second seco	Print cost dimension in the	"curation (Tagethal Incompa- matrixian over, "an year character in- Press report formation of the K
region, some till, "begins i den regione, gant som frammer. mannet", "rest, period"s i da.	Provide an arrival and a reason of a second at a secon	Particle and the second	Parallel and the State of the S	Compared, and Compared and Comp	Parents approved and American Sciences, Parents approved and American Sciences, Parents and American Sciences, Sciences, Parents and American Sciences, Scie	Particular and such such the	Participation (1997), 1997, 1998, 19	Particle 12 (2017) 100, 7 (2017) 1 (2017), Patronic 2017 1 (2017) 7 (2017) 1 (2017) Patronic 2017, 7 (2017) 7 (2017) 1 (2017) Patronic 2017) 2 (2017) 1 (2017)	Participation (1997) Temperature (1) Manual (1997) Texts, Manuarthi Participation, Manual (1997) Texts Participation (1997) Texts	Paramit, apper 1 on A. Camper's Lotter-1, Paramit, apper 1 on A. Annual *1 Parameters', New Appendix 1 of A.	Parent start's life, Varge's Life- by Amilian gar's land, Anisone's Presentation and an antipology biology	Partick school 1000, Sound's 1,000- b, Applicate apply lack, Applicate Press, Applicate application of the Press, and Application of the Application Press, and Applic
a fast identicit office with scenes and a 21-017	train for advocut, ten munic systemstry 34,370	tran has manufactured and period train relation	tran for threads the breath transport and 11,000	tree for sinch in 1984 mouth resp. ppls 11.00 2.5.861	from the science of the	term for better and the	tran for services. the martin service main 11.000	tem for advance, 14a manual texts mits 31,700	tion per since on the	Stein Bar Schward, 12, 403	Lank Dar Matherit. Asta menerit Assess halo 14,222. 32 343	tim far sind /t. 1911 month orter mile 12,000
<b>ZI I I</b>	-11 1 1		-1	-111		-3 1 I I	71	21 1	21	511		×1
2	2			5	-	- New Contraction	11 A		5	8		2
	-	- L				·	- <u> </u>	-	-		-	
the second	malass 41.	auter ti	sealors 43	author 44	michae 45	anabar 44	status 41	statur 11	auguar 13	niphar M	miniter St.	elater 12
a manager with the	Loss spectrals replaced	Seat manager think-th- build and the second	see optimized officiality of the second seco	rent general rich-tu- critice ric to Aug ri word F. 62/31	And spinster mercine Address in the surger	and approximate processing and a second seco	And applicated michaels, and an applicated michaels, and a second seco	Anne agrandant investorio- burganesis inconcesso membra 25. Materi	Party spectrum minarts- community, schools-	and polytain molecula- protoning LL-language	saat updataat mith.ta- titurain in. purpu upen 38. Mith	and Spin-tail Strik-11- Latin-relation thats
West right, Mr.	1 April 11475 11 From order windows ** 5, *incaset*s	The second secon	And Test Aren's Line (Construction of the Construction of the Cons	Print risers in the "statest"	1. April 12 provide 1. Apr	<pre>bit frame.com Append bing*/ is Provide and appendix for the frame.com Provide and appendix</pre>	Provide and a state of the stat	<ul> <li>Anna (1971) 21,</li> <li>Part of a standard (17, 7), Takasan's,</li> </ul>	<ul> <li>Appl, starting</li> <li>Para and starting's h. Typesant's</li> </ul>	Print and a standard in the Patron of the	Para (Construction)	Protection and the second seco
<ul> <li>"text dat begen", da and end begener, t.t.</li> </ul>	101. "true pair manufacture", G., "Store attain decoder", T. K. "secon consistent structure", N. K.	http://www.com/committee/	<ul> <li>Provide president and an and a state of the state of the</li></ul>	Fill, "truit par images": -L. "Gran arrive devices" / U. Ny Second constants of the State and Constants of the State State of the State of the State State of the State of the State State of the State of the State of the State State of the State of the State of the State State of the State of the State of the State State of the State of the State of the State of the State State of the State of the State of the State of the State State of the State of the State of the State of the State State of the State of the State of the State of the State State of the State of the State of the State of the State of the State State of the State o	Pil, "well, mail manner", it.	[14] Yorki and Manuff A.A. Name and American's Adv. Name and American's Adv.	Print, "Arrait pair (stages)"," (1), "Print and the second of 1 (1), "Array and the second of 1 (1),	No. "eral pat immed" in Second and an and the second seco	tick, formi par jumper? 1,	Petage advance starts and	"Start Striking courts 100, "Start Start Start", - 1, "Start Start Start", - 1,	Paral and analysis and a start and
er, Special's In. el, Sarol Harls, 105, NappenPr. 1097, "Splacetty": "Janp",	Part Barris 10, 10, Pagesters 10, Pagesters	Past, Receipting 10, Past, Speck, alors 1, 201, "Depend") "color", "palacator": "color",	Pain America's Dis- Paul, Santa and S. 10, "Summer"s Paul ", "polisiants": "Lang",	Page Barrier's St. Presi and scar's US, Tamont's range, spalacator's mater.	Para special (0), Presi, secol month, 20, "Inspecial") Presid, "palarizati", "Large (")	Peril Inter starts III, "Ingent?"	Past investor's Dr. Perel inter scients III, "Convert"s Paster, "palletation"s matters	Pass dependent (10), President aller (10), "Especific President (10), "Especific	Provident of the start of the s	Party monthing year's 200, Party Spectra 10, Party Solution, 2004"5, 201, "Pagesort"	Party meeting start's till, Date spectra 1 (b), Party picture, start's till, "rappent"	Persy musicar scarts mis- bar lenses, 10, real parts, ators in, reagants
A BELIEF BERTY BE	Practice and an added to the	Print ppl, 10 (dept) 1 (b) Printle, weather?) (c) Printle, weather?)	Provide and the second	Paul and a second to the	Prantic receiver, 1, Prantic receiver, 1, Prantic receiver, 1, Prantic receiver, 1, 1, Prantic receiver, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,	Preside and an address of the second	Press, pr., or dearfy the Press, management to Press, management to the	Party and the party of the part	Plant and the address of the second s	Party and American States	Parties and the state of the st	Anter and the store of the stor
n Deal and 1 (2019) (2019) (1990) And 1 (2019) (1990) And 1 (2019)	New Pert Loophi L. "Autors with"	The Part Scott 1, "Sides or "I	Plan Peel and 1 L. Nobles admin and readering? L. Nobles admin Patients status respect distances?	Print Peril, South 1, "Elitter, print") Print, "established" 1, Printer, Peril, Status, Status, St.	The lost conty L. Scherolder's	"Lin, family, local"), S., Tophini, sone"), http://www.science.com	Prov. Sand. 2010', C., "Silden, 2147', prot. "Analytication", S., Protect Marcin, Salarsi, Sciencer's	Party, Bard, Scott, S., Tarrier, Carth. Sec., Westernam, S., Party, Manager, S., Party, Manager, S.,	Print, Seed, Joseff, Sr. William, etc., 51 (1997) Automotive St. William, etc., 51 (1997) Automotive St.	1. Marin cards", T. Ha, 199, Sant Jinn?", J. Wanner, 1994. Mr. Namerat's L.	15. Neise (etc?) 1. Hit. 15. Neise (etc?) 1. Hit. 16. Neise (etc?) 1. Hit. 16. Neise (etc?) 1.	<ol> <li>Mercey and C. Mills, "Star Aust. District in "Content and", Mill, "Instances", 1,</li> </ol>
how I is here every the theet-	Constant Sector Sector Sector	Construction of the constr	TANKITCHINGHING/WH-Dimp-	Catholic Schopelly 185-Dage-	Constitution and an and	(hog)//m/hogosta/mi-chape-	Control of the part of the chapter	Contractor and a second second	President and an effect on the second	Poche stana copit kiestraffi	Permit status sarpti intertes/"-	Perfect Martin School Sciences/1-

#### Justin Johnson

#### Lecture 10 - 64

### Track ratio of weight update / weight magnitude

```
# assume parameter vector W and its gradient vector dW
param_scale = np.linalg.norm(W.ravel())
update = -learning_rate*dW # simple SGD update
update_scale = np.linalg.norm(update.ravel())
W += update # the actual update
print update_scale / param_scale # want ~le-3
```

ratio between the updates and values: ~ 0.0002 / 0.02 = 0.01 (about okay) want this to be somewhere around 0.001 or so

```
Justin Johnson
```

### Overview

### 1. One time setup

Activation functions, data preprocessing, weight initialization, regularization

### **2.**Training dynamics

Learning rate schedules;

hyperparameter optimization

### **3.After training**

Model ensembles, transfer learning,

large-batch training

Justin Johnson

Lecture 10 - 66

### **Model Ensembles**

- 1. Train multiple independent models
- 2. At test time average their results (Take average of predicted probability distributions, then choose argmax)

## Enjoy 2% extra performance

Justin Johnson

# Model Ensembles: Tips and Tricks Instead of training independent models, use multiple snapshots of a single model during training!



Loshchilov and Hutter, "SGDR: Stochastic gradient descent with restarts", arXiv 2016 Huang et al, "Snapshot ensembles: train 1, get M for free", ICLR 2017 Figures copyright Yixuan Li and Geoff Pleiss, 2017. Reproduced with permission.



Cyclic learning rate schedules can make this work even better!

#### Justin Johnson

#### Lecture 10 - 68

### Model Ensembles: Tips and Tricks

Instead of using actual parameter vector, keep a moving average of the parameter vector and use that at test time (Polyak averaging)



Polyak and Juditsky, "Acceleration of stochastic approximation by averaging", SIAM Journal on Control and Optimization, 1992. Karras et al, "Progressive Growing of GANs for Improved Quality, Stability, and Variation", ICLR 2018 Brock et al, "Large Scale GAN Training for High Fidelity Natural Image Synthesis", ICLR 2019

Justin Johnson

#### Lecture 10 - 69

# Transfer Learning

Justin Johnson

Lecture 10 - 70

### Transfer Learning

# "You need a lot of a data if you want to train/use CNNs"

Justin Johnson

Lecture 10 - 71

### Transfer Learning





Lecture 10 - 72
1. Train on ImageNet

FC-1000
FC-4096
FC-4096
MaxPool
Conv-512
Conv-512
MaxPool
Conv-512
Conv-512
MaxPool
Conv-256
Conv-256
MaxPool
Conv-128
Conv-128
MaxPool
Conv-64

Conv-64

Image

t 2. Use CNN as a

feature extractor



Donahue et al, "DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition", ICML 2014

#### Justin Johnson



Image

1. Train on ImageNet

FC-1000 FC-4096 FC-4096 **MaxPool** Conv-512 Conv-512 MaxPool Conv-512 Conv-512 **MaxPool** Conv-256 Conv-256 MaxPool Conv-128 Conv-128 MaxPool Conv-64

Conv-64

Image



Donahue et al, "DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition", ICML 2014

February 9, 2022

#### Justin Johnson

1. Train on ImageNet

FC-1000 FC-4096 FC-4096 **MaxPool** Conv-512 Conv-512 MaxPool Conv-512 Conv-512 MaxPool Conv-256 Conv-256 MaxPool Conv-128 Conv-128 MaxPool Conv-64

Conv-64

Image

et 2. Use CNN as a

Image

feature extractor



Donahue et al, "DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition", ICML 2014

#### Justin Johnson

Lecture 10 - 75

1. Train on ImageNet

FC-1000 FC-4096 FC-4096 **MaxPool** Conv-512 Conv-512 MaxPool Conv-512 Conv-512 MaxPool Conv-256 Conv-256 MaxPool Conv-128 Conv-128 MaxPool Conv-64

Conv-64

Image

### t 2. Use CNN as a

feature extractor



### Bird Classification on Caltech-UCSD



Donahue et al, "DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition", ICML 2014

February 9, 2022

#### Justin Johnson

### 1. Train on ImageNet

FC-1000 FC-4096

FC-4096

**MaxPool** 

Conv-512

Conv-512

MaxPool

2. Use CNN as a

feature extractor





MaxPool Conv-64 Conv-64

Image

Justin Johnson

#### Lecture 10 - 77

### 1. Train on ImageNet

2. Use CNN as a

feature extractor

FC-1000
FC-4096
FC-4096
MaxPool
Conv-512
Conv-512
MaxPool
Conv-512
Conv-512
MaxPool
Conv-256
Conv-256
MaxPool
Conv-128
Conv-128
MaxPool
Conv-64
Conv-64

Image

Justin Johnson



#### Lecture 10 - 78



#### Justin Johnson

Lecture 10 - 79

1. Train on Imagenet

FC-1000	
FC-4096	
FC-4096	
MaxPool	
Conv-512	
Conv-512	
MaxPool	
Conv-512	
Conv-512	
MaxPool	
Conv-256	
Conv-256	
MaxPool	
Conv-128	
Conv-128	
MaxPool	
Conv-64	
Conv-64	

Image

2. Use CNN	l as a
feature ext	ractor
FC-4096	
FC-4096	Remove
MaxPool	last layer
Conv-512	
Conv-512	
MaxPool	
Conv-512	
Conv-512	
MaxPool	Freeze
Conv-256	
Conv-256	these
MaxPool	
Conv-128	
Conv-128	
MaxPool	
Conv-64	
Conv-64	)
Image	-

### 3. Bigger dataset: Fine-Tuning

FC-4096	←
FC-4096	
MaxPool	
Conv-512	
Conv-512	
MaxPool	
Conv-512	
Conv-512	
MaxPool	
Conv-256	
Conv-256	
MaxPool	
Conv-128	
Conv-128	
MaxPool	
Conv-64	
Conv-64	
Image	

Continue training CNN for new task!

Some tricks:

- Train with feature extraction first before fine-tuning
- Lower the learning rate: use ~1/10 of LR used in original training
- Sometimes freeze lower layers to save computation
- Train with BatchNorm in "test" mode

#### Justin Johnson

Lecture 10 - 80



#### Justin Johnson

Lecture 10 - 81

### Transfer Learning with CNNs: Architecture Matters!

### ImageNet Classification Challenge



Improvements in CNN architectures lead to improvements in many downstream tasks thanks to transfer learning!

#### Justin Johnson

#### Lecture 10 - 82

### Transfer Learning with CNNs: Architecture Matters!

Object Detection on COCO



Ross Girshick, "The Generalized R-CNN Framework for Object Detection", ICCV 2017 Tutorial on Instance-Level Visual Recognition

Justin Johnson

#### Lecture 10 - 83



#### Justin Johnson

Lecture 10 - 84



#### Justin Johnson

Lecture 10 - 85



#### Justin Johnson

Lecture 10 - 86



#### Justin Johnson





Karpathy and Fei-Fei, "Deep Visual-Semantic Alignments for Generating Image Descriptions", CVPR 2015

Girshick, "Fast R-CNN", ICCV 2015 Figure copyright Ross Girshick, 2015. Reproduced with permission.

#### Justin Johnson

#### Lecture 10 - 88



Girshick, "Fast R-CNN", ICCV 2015 Figure copyright Ross Girshick, 2015. Reproduced with permission.

Justin Johnson

Karpathy and Fei-Fei, "Deep Visual-Semantic Alignments for Generating Image Descriptions", CVPR 2015

Lecture 10 - 89





### 1. Train CNN on ImageNet

- 2. Fine-Tune (1) for object detection on Visual Genome
- 3. Train BERT language model on lots of text
- 4. Combine (2) and (3), train for joint image / language modeling
- 5. Fine-tune (5) for image captioning, visual question answering, etc.

Zhou et al, "Unified Vision-Language Pre-Training for Image Captioning and VQA", arXiv 2019

#### Justin Johnson

Lecture 10 - 91

Transfer learning is pervasive! Some very recent results have questioned it

**COCO** object detection



Training from scratch can work as well as pretraining on ImageNet!

... If you train for 3x as long

He et al, "Rethinking ImageNet Pre-Training", ICCV 2019

February 9, 2022

Justin Johnson

Transfer learning is pervasive! Some very recent results have questioned it

COCO object detection



Pretraining + Finetuning beats training from scratch when dataset size is very small

Collecting more data is more effective than pretraining

He et al, "Rethinking ImageNet Pre-Training", ICCV 2019

February 9, 2022

#### Justin Johnson

Transfer learning is pervasive! Some very recent results have questioned it



My current view on transfer learning:

- Pretrain+finetune makes your training faster, so practically very useful
- Training from scratch works well once you have enough data
- Lots of work left to be done

He et al, "Rethinking ImageNet Pre-Training", ICCV 2019

#### Justin Johnson

Recap

### **1.One time setup**

Activation functions, data preprocessing, weight initialization, regularization

### 2. Training dynamics

Learning rate schedules;

hyperparameter optimization

### **3.After training**

Model ensembles, transfer learning,

**Last Time** 

Today

Justin Johnson

# Next Time: More CNN Architectures

Justin Johnson

Lecture 10 - 96