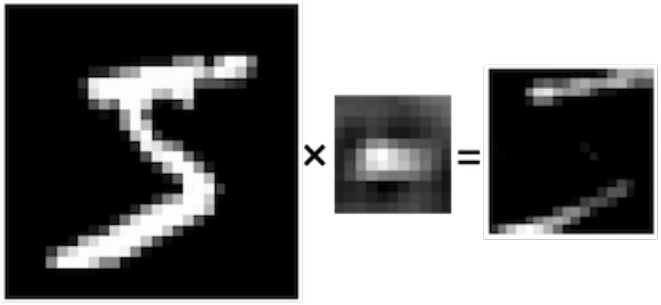
Universität Potsdam

Introduction to Deep Learning

Regularization & Practical Methodology



Andreas Krug, M.Sc.

ankrug@uni-potsdam.de

17. December 2018

Exercise

Tiny exam questions

Tiny exams questions

1. How do you observe underfitting of your model and what can you do to prevent this from happening?
2. How do you observe overfitting of your model and what can you do to prevent this from happening?
3. What is the difference between training, validation and test set and what can low or high error on these sets tell you?
4. What are L1 and L2 regularization and how do they affect the network?
5. What are possible ways of augmenting your data?
6. How can you use noise to regularize your model?
7. How does dropout relate to model averaging?
8. Which typical architectures use parameter sharing and how does this improve generalization?

worksheet questions 1&2

1 Is it reasonable to evaluate a (rare) disease predictor by its accuracy?
For the example below: Discuss, which evaluation metric would be most suitable and why!

		prediction		
		disease	healthy	
prediction	disease			5
	healthy			95

2 A researcher trains a deep neural network.
The model performs great on the train set, but rather poorly on the test set.
To improve generalization the researcher:
adds dropout, adds L1 and L2 regularization, tries different optimizers and learning rates.
Finally there is a configuration where the test error gets close to the train error.
Now the model is ready to be published!
Do you agree? Discuss whether the researcher followed the good practice.
What could he/she have done better?

accuracy

- % of correct predictions = accuracy

		prediction		
		disease	no disease	
Truth (data)	disease	0	5	5
	no disease	0	95	95
		0	100	100

$$\text{accuracy} = \frac{0+95}{100} = 95\%$$

recall

- % true events that were detected = recall

		prediction		
		disease	no disease	
Truth (data)	disease	0	5	5
	no disease	0	95	95
		0	100	100

$$\text{recall} = \frac{0}{5} = 0\%$$

recall

- % true events that were detected = recall

		prediction		
		disease	no disease	
Truth (data)	disease	5	0	5
	no disease	45	50	95
		50	50	100

$$\text{recall} = \frac{5}{5} = 100\%$$

precision

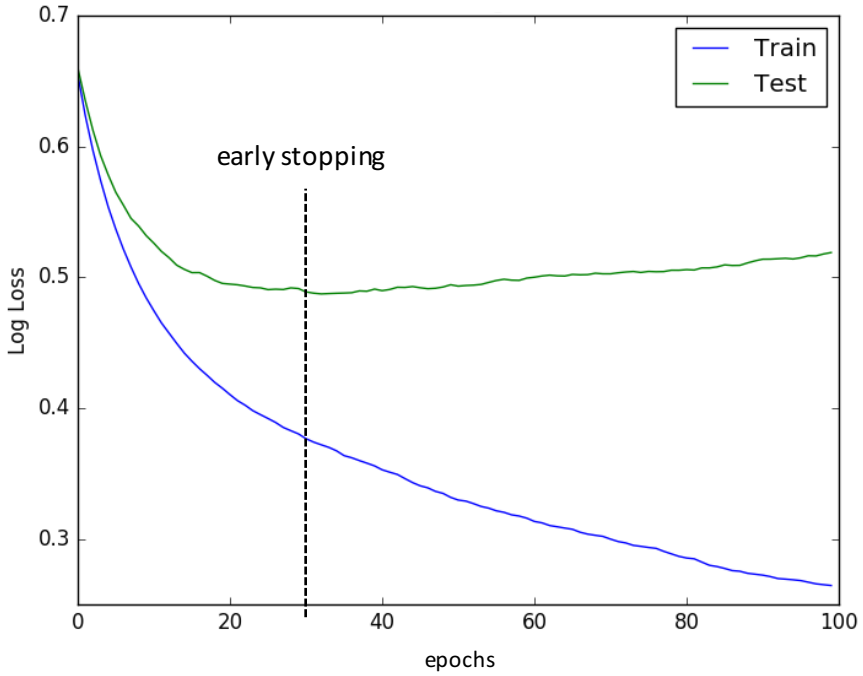
- % detections that were correct = precision

		prediction		
		disease	no disease	
Truth (data)	disease	5	0	5
	no disease	45	50	95
		50	50	100

$$\text{precision} = \frac{5}{50} = 10\%$$

worksheet question 3

3 In some model, we observe that the test loss increases after 30 epochs, therefore we use early stopping at 30 epochs as regularization technique. Discuss whether this is a good approach! If not, think about how to fix the problem!



Metrics

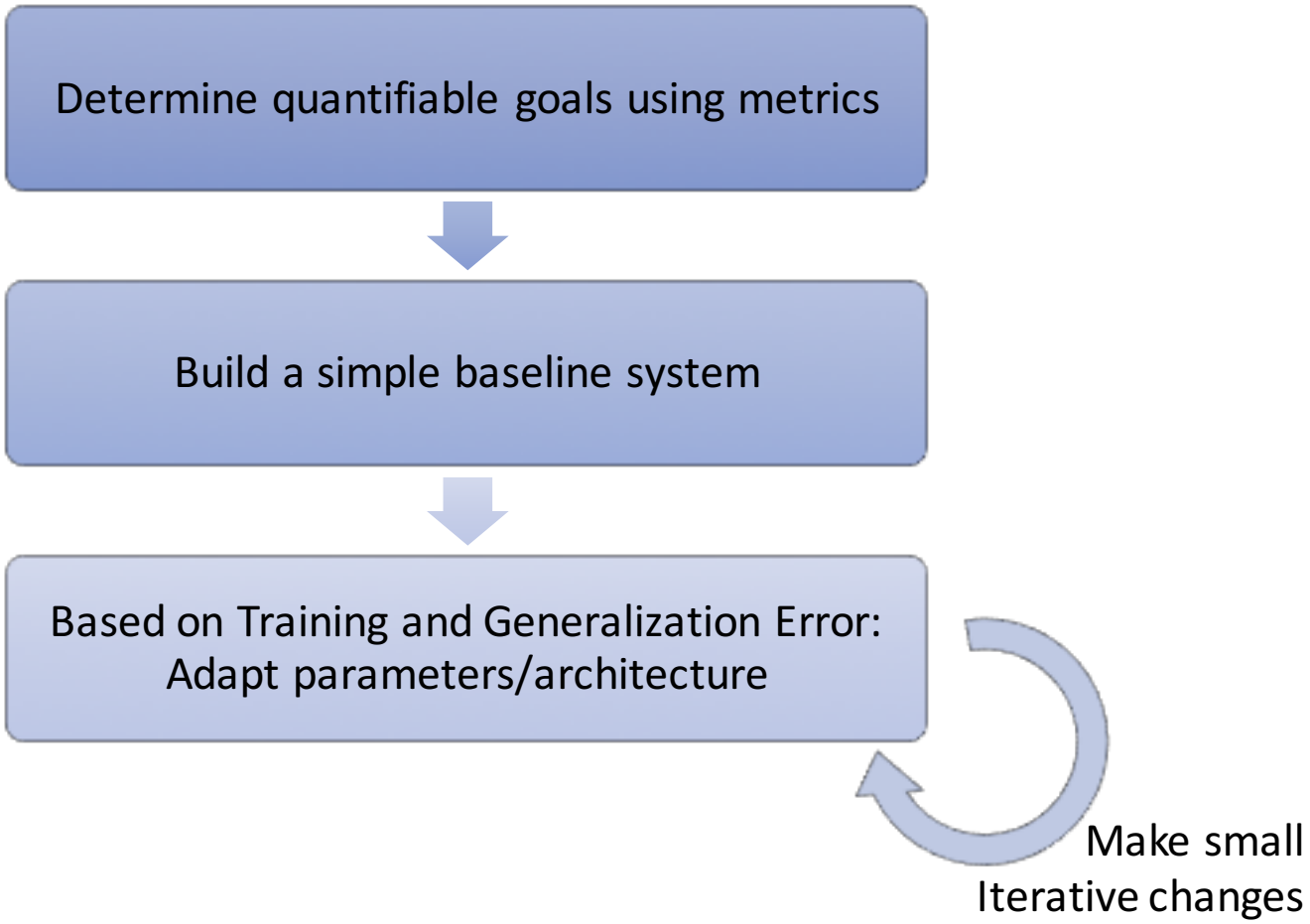
Metric	What it determines	When to use	Example
Accuracy	Correct predictions in all classes	All classes should be similarly abundant	OCR
Recall	Correct predictions of a particular class	If detecting a particular class is of importance	Disease detection
Precision	How specific the prediction is for a particular class	If detecting a particular class has to be correct	Spam filter (non-spam mails)
Coverage	How many examples are processed	If the system can refuse prediction if it is not confident	Street View digit detection
Prob.-distr. divergence	Similarity of data and model probability distributions	If the system is generative	MNIST with RBM
Amount of error	Distance between data and predictions	Regression problems	Linear regression

Training, Validation, Test set

Data subset	What a small loss tells:	Decrease loss by:
Training set	The system learns something (is not underfitting)	- Fixing data errors & code bugs - Adjusting learning parameters - Increasing model capacity
Validation set	The system is not overfitting	- Data augmentation - Hyperparameter tuning - Regularization - Gathering more data
Test set	The system generalizes well	- decreasing training loss and validation loss

Don't change your setting to decrease the test loss directly!

Model refinement



Assignments

- Responsible for recap: Freya & Andrew
- Next session: Jan 07 (in 3 weeks)
- Reading on Optimization
- (no programming task)
- Time for your project!

Slides & assignments on: https://mlcogup.github.io/idl_ws18/schedule