DEEP LEARNING IN 540 MINUTES

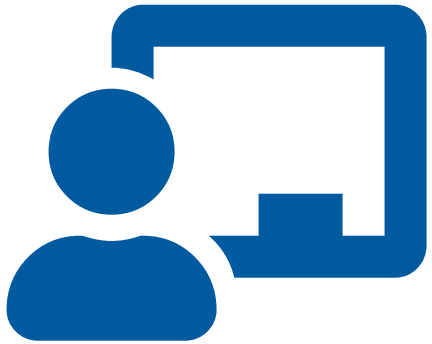
Peter Steinbach
(HZDR)

psteinb psteinb_

Somewhere, Mar 1, 2021

AN INTRODUCTION TO ...

WHAT?



data science,
machine learning,
deep learning

DATA SCIENCE

Data science is an inter-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from many structural and unstructured data. Data science is related to data mining, machine learning and big data.

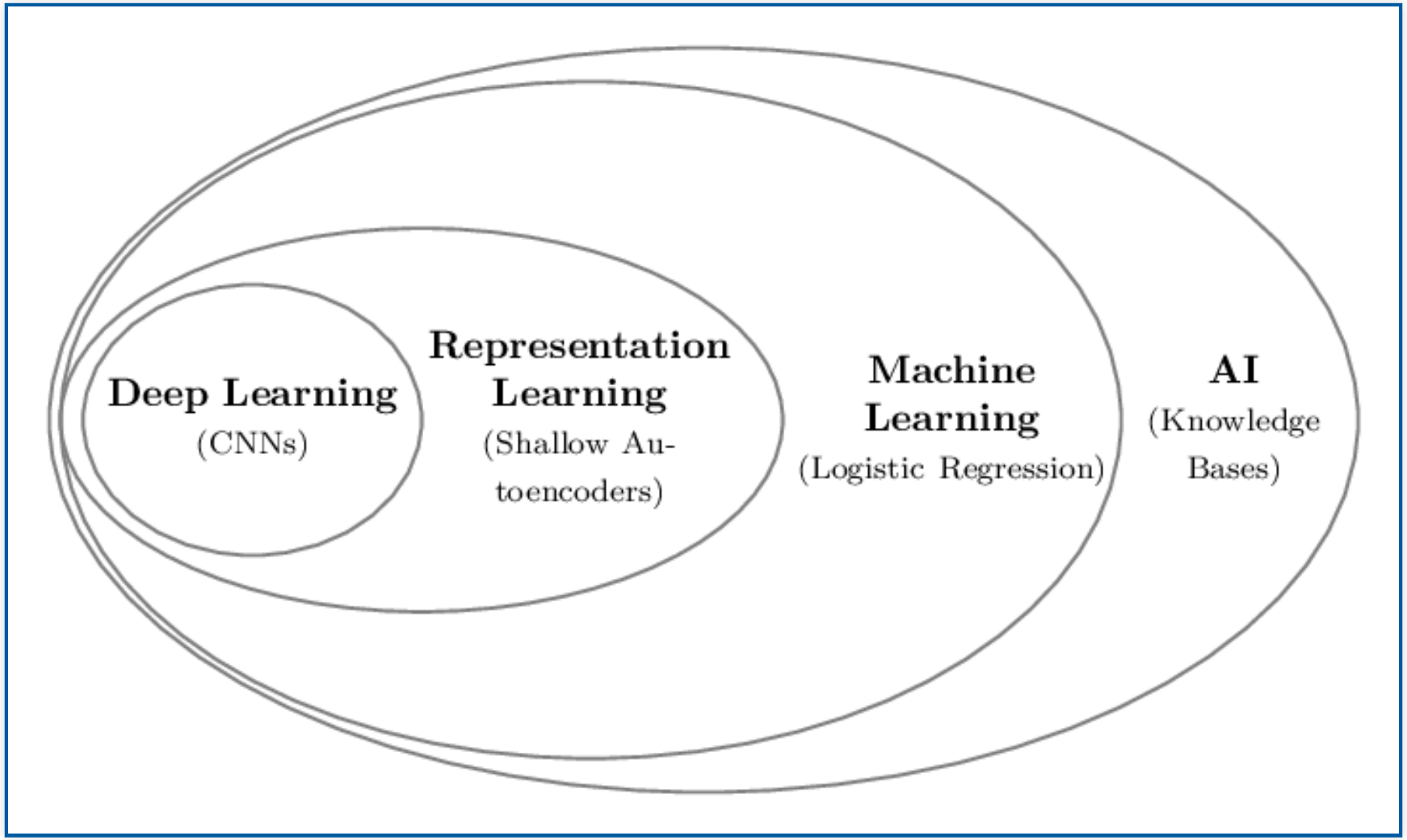
MACHINE LEARNING

Machine learning (ML) is the study of computer algorithms that improve automatically through experience. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so.

DEEP LEARNING

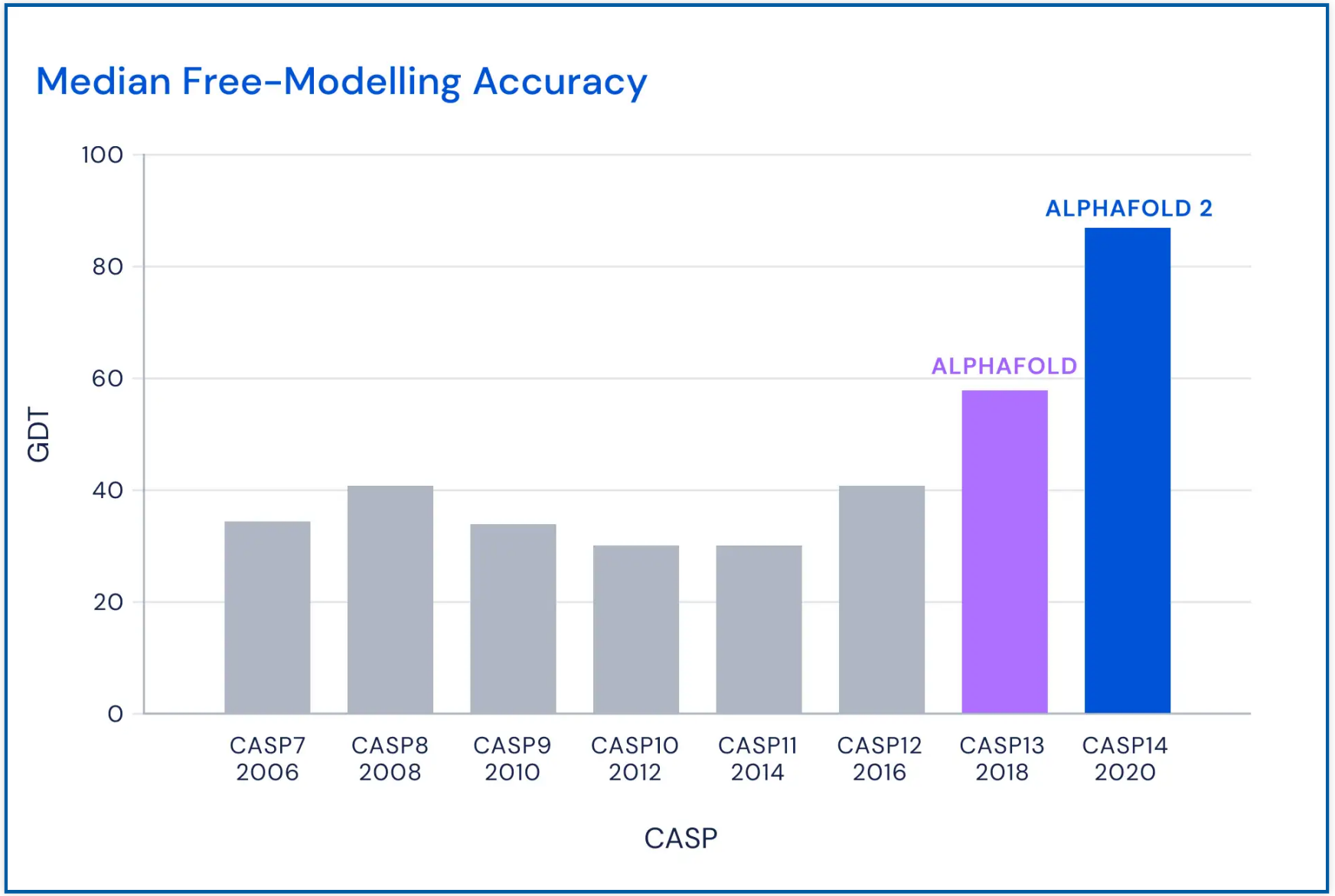
Deep learning (also known as deep structured learning) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised.

PUTTING IT ALL TOGETHER



ML TODAY

ALPHAFOLD2


Median Free-Modelling Accuracy
100
80
60
40
20
0
GDT
ALPHAFOLD 2
ALPHAFOLD
CASP7 2006
CASP8 2008
CASP9 2010
CASP10 2012
CASP11 2014
CASP12 2016
CASP13 2018
CASP14 2020
CASP

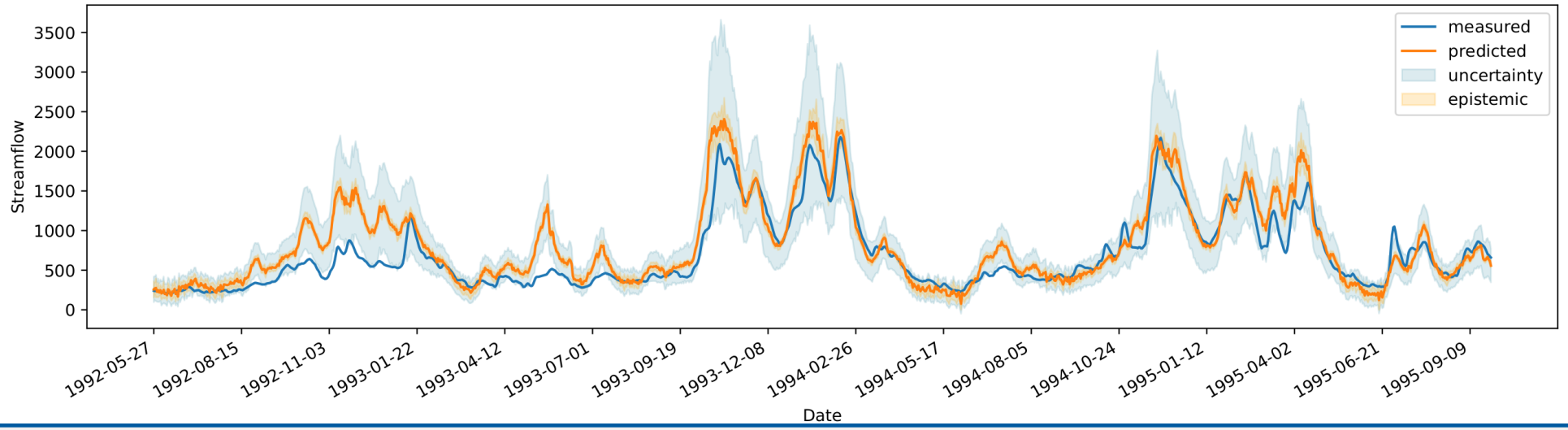
FIGHTING DESEASES WITH DL

f.isensee@dkfz.de
https://github.com/MIC-DKFZ/nnUNet

Test set prediction

- Marginally (if at all) improved upon the nnU-Net baseline
- Ensemble 5 best configurations
- For each configuration: use 5 models from cross-validation + 5 additional models trained on random 80:20 splits
  - 50 models total
  - No external data, models trained from scratch
- Test set Dice: 0.6543 ± 0.2710 (ranked 4th out of 98)

PREDICTING RIVER TIDES



ML HAS BECOME A TOOL FOR SCIENCE

TEACHING THIS INTRODUCTION

THE NO FREE LUNCH THEOREM BY DAVID WOLPERT

Averaged over all possible data generating distributions, every classification algorithm has the same error rate when classifying previously unobserved points.

THE NO FREE LUNCH THEOREM BY DAVID WOLPERT

Averaged over all possible data generating distributions, every classification algorithm has the same error rate when classifying previously unobserved points.

How well you do is determined by how 'aligned' your learning algorithm P(h|d) is with the actual posterior, P(f|d).

IN THE NEXT 3 DAYS …

- 100 people
- not in one place due to pandemic
- all online (including distractions)
- from newbie to pro?

WHAT I CAN HOWEVER ASPIRE TO …

WHAT MOTIVATES ME …

teaching for inclusion

,,Helping the students with less computing background succeed makes a much bigger difference for society long-term than does keeping entertained the most privileged students." (Mark Guzdial, 2020)

FOR THE BORED ...

- IN2P3 School of Statistics
- Sebastian Raschka:
  - Introduction to Deep Learning and Generative Models
  - Introduction to Machine Learning and Statistical Pattern Classification

WHAT YOU CAN EXPECT: ADAPTED INVERTED CLASSROOM

- instructions as video
  - each learner decides the speed
  - can be completed outside class
- learning in teams - together
- many exercises (more than you can complete)

HOW THIS WILL WORK

We set up **teams of about 10** learners!

Each team …

- has one instructor
- has one zoom room
- has one mattermost channel
- has one hackmd pad for notes

There are 8 lessons!

- each lesson designed for 1h
- during each lesson

1. watch the video
2. answer the check-your-learning together
3. conduct an exercise (if time permits)

YOUR INSTRUCTORS

- Erik Buhmann
- Sascha Diefenbacher
- Manuel Sommerhalder
- William Korcari
- Tobias Lösche
- David Brunner
- Mykyta Shchedrolosiev
- Jonas Rübenach
- Oleg Filatov
- Moritz Scham
- Lucas Wiens
- Helene Hoffmann

Thank you!

AGENDA

1. Diving into Regression
2. Enter Clustering
3. From Clustering To Classification
4. Classification Performance ROCs
5. Neural Networks as Code
6. How did we train
7. CNNs
8. Deep Learning is ML too

SUMMARY

LEARN!

- be positive
- help each-other (hackmd)
- communicate (mattermost)
- decide together

Thank you for your Attention!
Questions, Concerns, Feedback?