29th Vienna Deep Learning Meetup

24th September 2019

#VDLM

Raiffeisen Software

Vienna Deep Learning Meetup

The Organizers:

Thomas Lidy
Musimap

Alex Schindler
AIT & TU Wien

Jan Schlüter
OFAI & UTLN

René Donner
contextflow

Topics for Today

Vienna Deep Learning Meetup

- **Welcome & Introduction**
- **Deep Learning for Recommender Systems**

  Jakub Mačina, Machine Learning Engineer, Exponea
- **The Fastai Deep Learning Library**

  Michael Pieler, Data Scientist
- **<Break>**
- **ACM Recommender Systems Conference Review**

  Jakub Mačina, Machine Learning Engineer, Exponea
- **Hot Topics**

  Michael Pieler, Data Scientist

Vienna Deep Learning Meetup

Announcements

Vienna
Deep Learning
Meetup

VDLM on Github

https://github.com/vdlm/meetups

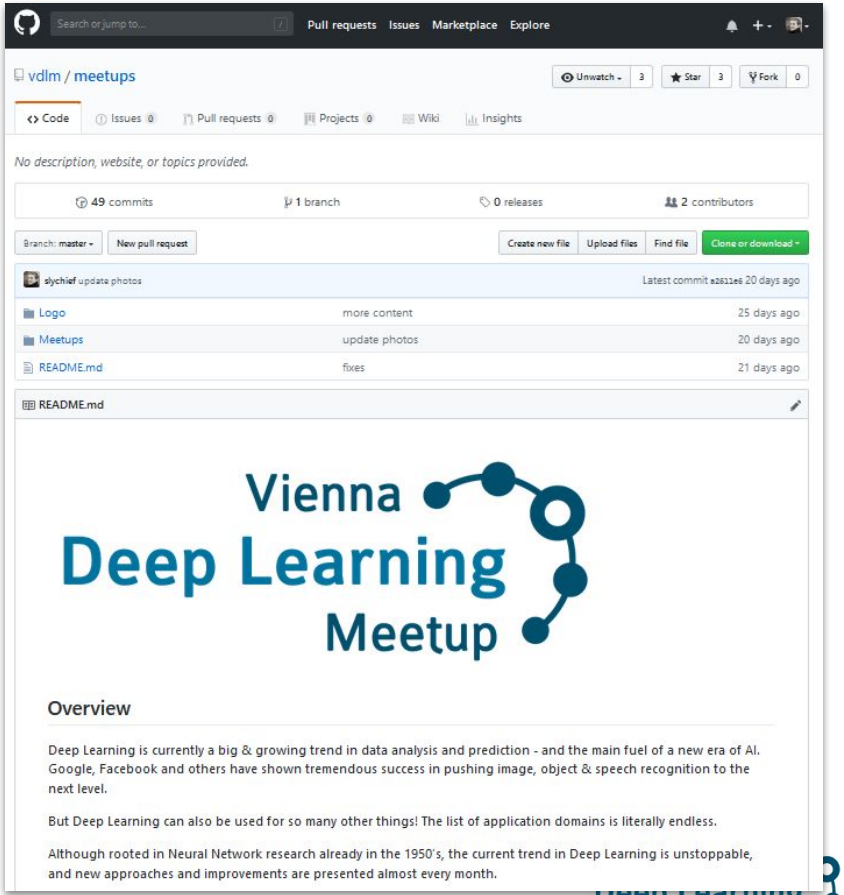
- all talks
- slides
- photos
- videos
- Wiki

Meetups

#	Date	Place	Topic	Link	Video	Meetup.com
1	2016-04-07	Sector 5	intro	more		link
2	2016-05-09	Sector 5		more		link
3	2016-06-06	Sector 5		more		link
4	2016-07-07	TU Wien		more		link
5	2016-09-22	Automic Software GmbH		more		link
6	2016-10-12	Sector 5		more		link
7	2016-12-01	Agentur Virtual Identity		more		link
8	2017-01-17	TU Wien Informatik		more		link
9	2017-02-21	bwin.party services (Austria) GmbH		more		link

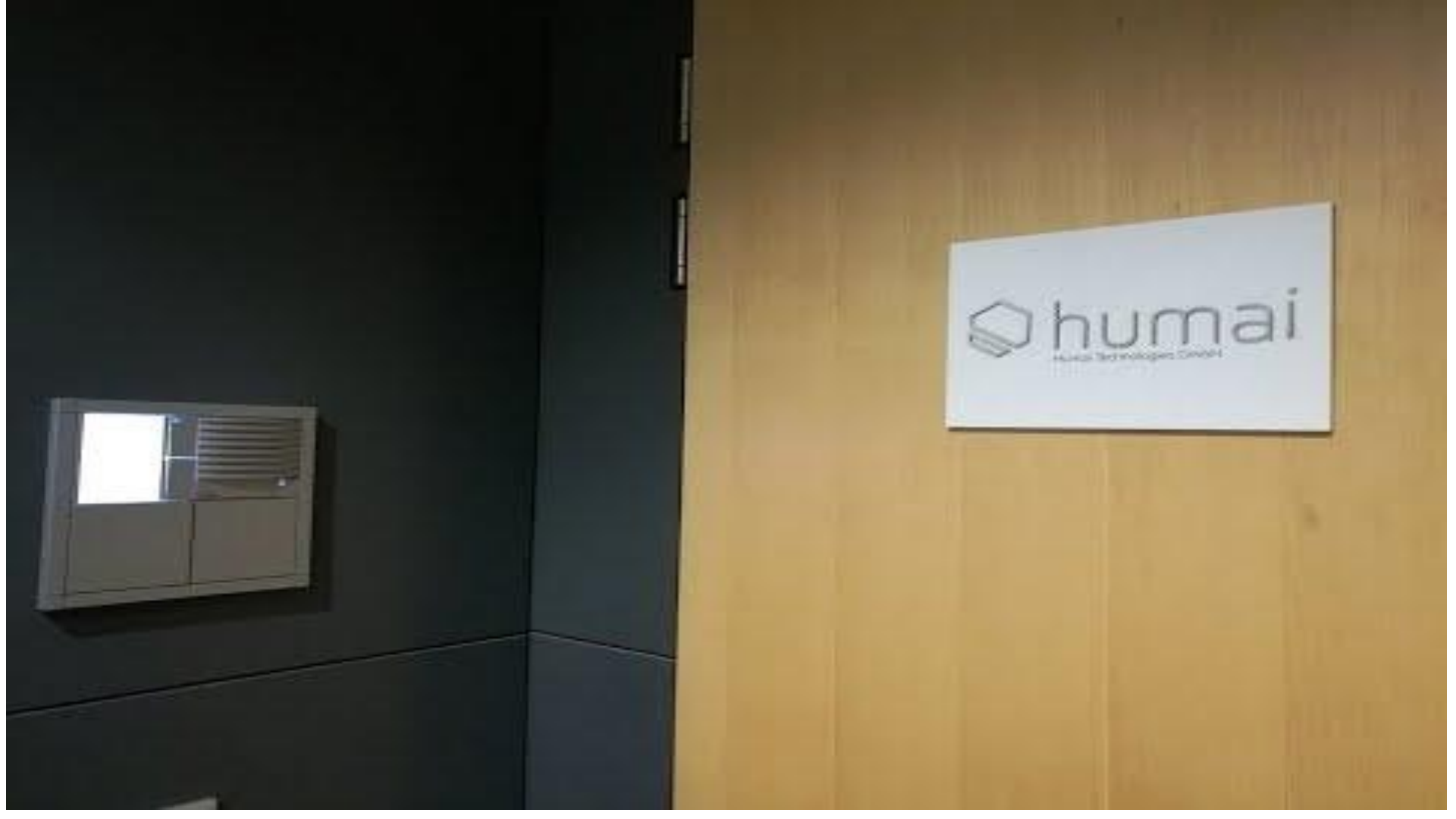
Talks

Date	MU#	Speaker	Topic	Slides
2016-04-07	1	Thomas Lidy	An overview presentation of Deep Learning	pdf
2016-04-07	1	Jan Schlüter	History, Approaches, Applications	pdf
2016-05-09	2	Alex Champandard	Neural Networks for Image Synthesis	
2016-05-09	2	Gregor Mitscha-Baude	Recurrent Neural Networks	pdf
2016-06-06	3	Jan Schlüter	Open-source Deep Learning with Theano and Lasagne	pdf
2016-09-22	5	Josef Puchinger	Deep Learning & The Future of Automation	
2016-09-22	5	Christoph Körner	Going Deeper with GoogLeNet and CaffeJS	pdf



Deep Learning
Meetup


humai
Looking for a new Career Challenge
@Humai?


humai

humai

We are looking for:

EXPERIENCED COMPUTER VISION ENGINEER (M/W/D)

FULL-STACK WEB DEVELOPER (M/W/D)

Reference for applications:

Stephan Ganzinger , BA , MBA
jos@humai.tech
https://www.humai.tech/karriere/
+43 660 9127 849

At Austrian Airlines, we are looking for a **Strategic Data Engineer** to join our emerging Digital Delivery Team

What will you be doing?

Drive **automation and digitalization** of business from a data point-of-view

Evaluate, develop, prepare & test data sets of the entire organization

Discover **opportunities of data acquisition** for us

Develop **processes/pipelines for data modeling**, mining & production

Put machine learning/deep learning **data models in production**

Why join Austrian?

You will **pioneer and drive forward digitalization topics**, thereby having a lasting impact on our organization

You will interact with **top management** and be situated in Austrian's **strategy department**

You will have the ability to work on a **variety of use cases** & different data sets (Ops, Finance, HR, …)

You will become part of the Austrian Airlines family with many benefits and attractive travel discounts

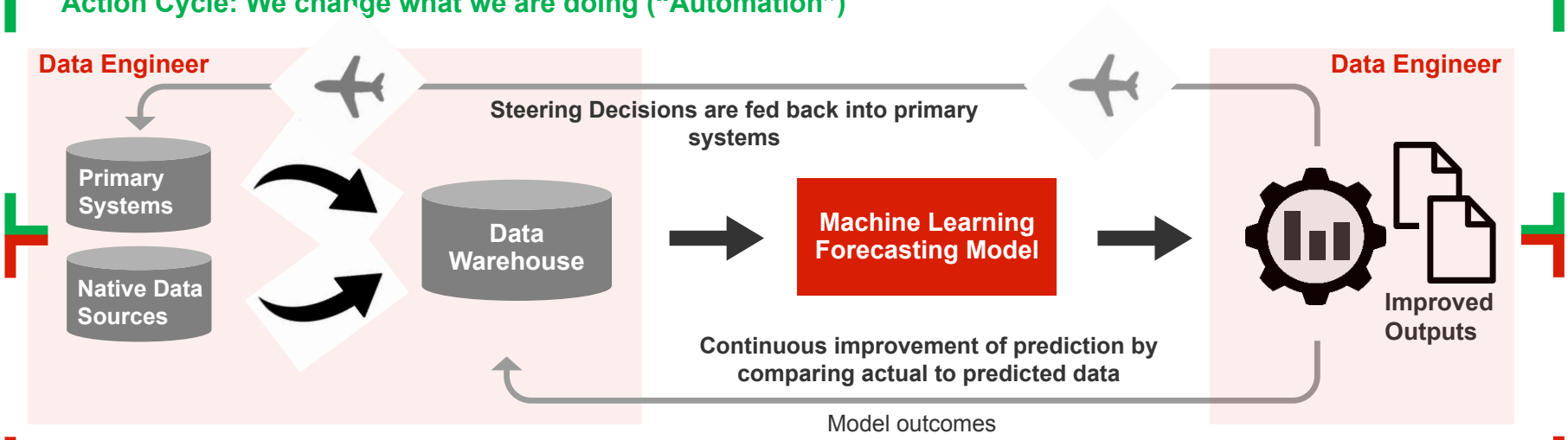
Interested in joining us?

Please apply via
career.be-lufthansa.com

Austrian

As Strategic Data Engineer at Austrian Airlines, you will be essential in supporting our business units to implement use cases

Action Cycle: We change what we are doing ("Automation")



Learning Cycle: We further improve ("Machine Learning")

SBY Crew Use Case: Using a machine learning model to forecast SBY levels

Weather, sick rates & historical SBY level data is saved → Data is aggregated and cleaned in central data warehouse → Machine learning model accesses data and creates forecast → Prediction is used for planned SBY levels, actual requirement is fed back into data warehouse → **Actual Value Generation:** Optimized crew planning with improved SBY crew levels

Austrian


The
BRAINSTORMS
FESTIVAL
September 27th & 28th, 2019
Palladion XXI Andreas Hofer Straße 28, 1210 Vienna, Austria

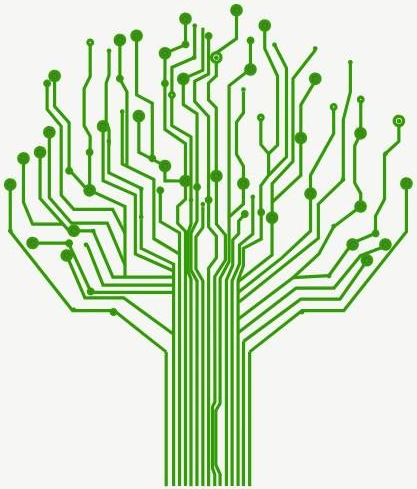
A.I. highlights

Day 1
11:00-11:30 Vince Madai, Cognitive Biases in AI
14:00-14:30 Lushi Chen, AI and mental disorders
14:30-15:30 PANEL DISCUSSION Transformation of diagnostics in the next 5 years

Day 2
11:30-12:00 Peter Schlecht, How to radically improve our brains and save humanity
11:30-12:00 (stage 2) Kira Kempinska, Making new drugs with a dose of AI
13:30-14:40 (stage 2) Daniel Spichtinger, Horizon Europe - an intro to the new EU R&I programme
14:00-14:30 Diana Deca, Robots inspired by the brain
14:30-15:30 PANEL DISCUSSION The Augmented Men (Isabella Hillmer, Peter Schlecht, Diana Deca)
14:30-18:00 Liad Magen, Using AI to fight Climate Change
16:00-16:30 Cosima Prahm, The new generation prosthetics and computer-assisted rehabilitation
17:30-18:00 Hackathon grand prize reveal

IN PARALLEL : BRAIN-COMPUTER INTERFACE HACKATHON

USING AI TO FIGHT CLIMATE CHANGE



Brainstorming workshop
@the Brainstorms Festival

Colosseum XXI - Colosseum 21 Palladion 21
Andreas Hofer Straße 28, 1210 Vienna, Austria

Developer Circles
from facebook

Saturday, 28.9 - 14:30 -18:00

Ready2build your own Startup?

- Game that increases emotional intelligence
- Real-estate recommendation system
- Custom clearance automatic form-filling
- Conference Schedule RecSys
- PodCasts Speech Detection (With Dialect!)
- And more...

READY2ORD3R HACKATHON VIENNA '19

October 4-6th

https://ready-2-build-your-mvp.com

Hot Topics & Latest News

a short block at every meetup
to briefly present recent papers and news in Deep Learning

Send us contributions (tom.lidy@gmail.com)
or come with slides to do a short block yourself!

Vienna
Deep Learning
Meetup

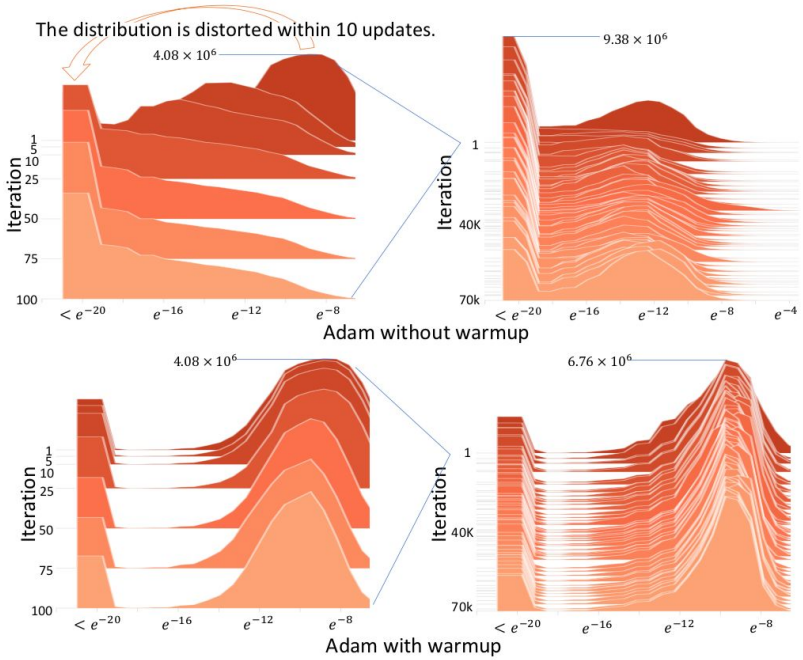
RAdam - Rectified Adam

- learning rate warmup stabilizes training, accelerates convergence and improves generalization

- problem of adaptive learning rates:
  - problematically large variance in the early stage of training

  → *suggests warmup works as a variance reduction technique!*

Absolute gradient histogram:



- RAdam introduces a term to rectify the variance of the adaptive learning rate!

https://arxiv.org/pdf/1908.03265 & https://medium.com/@lessw/new-state-of-the-art-ai-optimizer-rectified-adam-radam-5d854730807b

Vienna Deep Learning Meetup

RAdam - Update

Algorithm 2: Rectified Adam. All operations are element-wise.

Input: $\{\alpha_t\}_{t=1}^{T}$: step size, $\{\beta_1, \beta_2\}$: decay rate to calculate moving average and moving 2nd moment, θ_0: initial parameter, $f_t(\theta)$: stochastic objective function.

Output: θ_t: resulting parameters

$m_0, v_0 \leftarrow 0, 0$ (Initialize moving 1st and 2nd moment)

$\rho_\infty \leftarrow 2/(1-\beta_2) - 1$ (Compute the maximum length of the approximated SMA)

while $t = \{1, \cdots, T\}$ **do**

  $g_t \leftarrow \Delta_\theta f_t(\theta_{t-1})$ (Calculate gradients w.r.t. stochastic objective at timestep t)

  $v_t \leftarrow \beta_2 v_{t-1} + (1-\beta_2) g_t^2$ (Update exponential moving 2nd moment)

  $m_t \leftarrow \beta_1 m_{t-1} + (1-\beta_1) g_t$ (Update exponential moving 1st moment)

  $\widehat{m_t} \leftarrow m_t / (1-\beta_1^t)$ (Compute bias-corrected moving average)

  $\rho_t \leftarrow \rho_\infty - 2t\beta_2^t/(1-\beta_2^t)$(Compute the length of the approximated SMA)

  if *the variance is tractable, i.e.,* $\rho_t > 4$ **then**

    $\widehat{v_t} \leftarrow \sqrt{v_t/(1-\beta_2^t)}$ (Compute bias-corrected moving 2nd moment)

    $r_t \leftarrow \sqrt{\frac{(\rho_t-4)(\rho_t-2)\rho_\infty}{(\rho_\infty-4)(\rho_\infty-2)\rho_t}}$ (Compute the variance rectification term)

    $\theta_t \leftarrow \theta_{t-1} - \alpha_t r_t \widehat{m_t} / \widehat{v_t}$ (Update parameters with adaptive momentum)

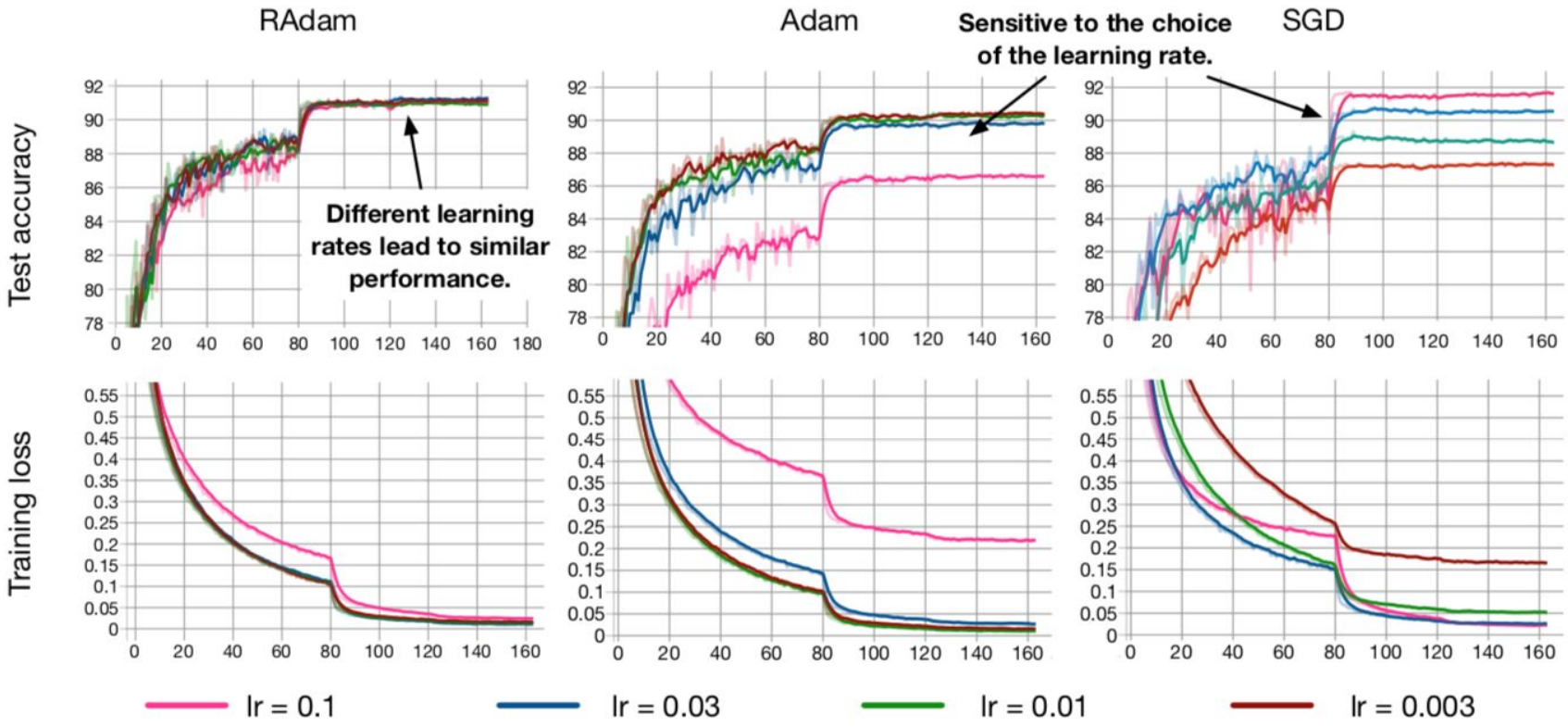
  else

    $\theta_t \leftarrow \theta_{t-1} - \alpha_t \widehat{m_t}$ (Update parameters with un-adapted momentum)

return θ_T

https://arxiv.org/pdf/1908.03265

RAdam, Adam, and SGD

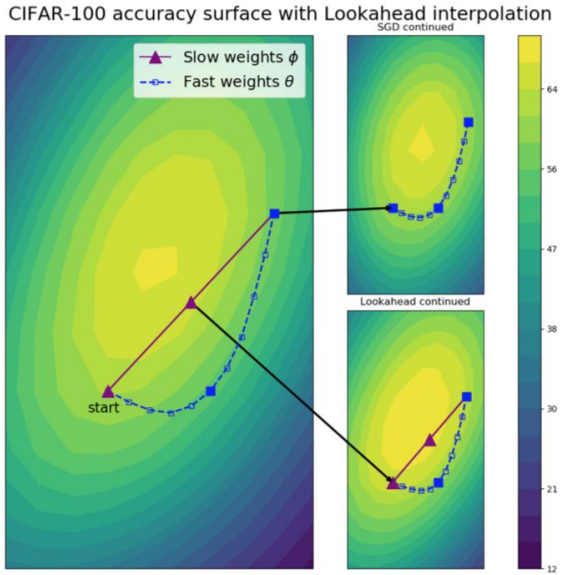


https://arxiv.org/pdf/1908.03265

Vienna Deep Learning Meetup

LookAhead

- Iteratively updates two sets of weights:
  - “slow weights” get updated by looking ahead at the sequence of “fast weights” generated by another optimizer

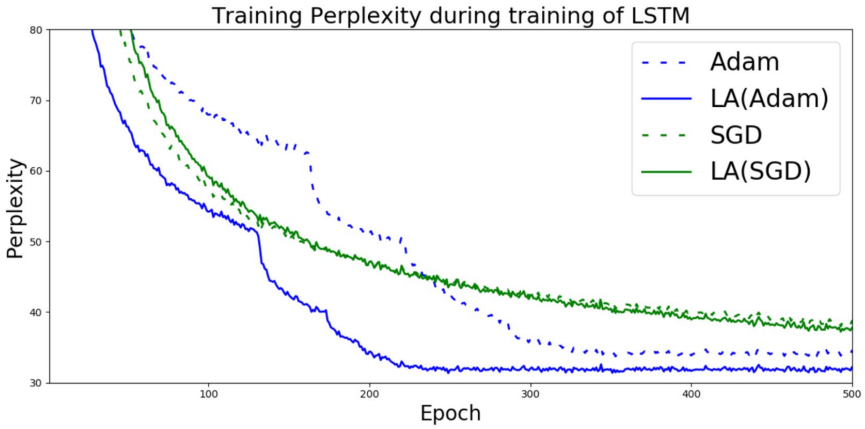


Algorithm 1 Lookahead Optimizer:

Require: Initial parameters ϕ_0, objective function L
Require: Synchronization period k, slow weights step size α, optimizer A
for $t = 1, 2, \ldots$ **do**
  Synchronize parameters $\theta_{t,0} \leftarrow \phi_{t-1}$
  for $i = 1, 2, \ldots, k$ **do**
    sample minibatch of data $d \sim \mathcal{D}$
    $\theta_{t,i} \leftarrow \theta_{t,i-1} + A(L, \theta_{t,i-1}, d)$
  end for
  Perform outer update $\phi_t \leftarrow \phi_{t-1} + \alpha(\theta_{t,k} - \phi_{t-1})$
end for
return parameters ϕ

https://arxiv.org/pdf/1907.08610.pdf

LookAhead + ?



A better combination?

→ Ranger = LookAhead + RAdam

https://arxiv.org/pdf/1907.08610.pdf &
https://medium.com/@lessw/new-deep-learning-optimizer-ranger-synergistic-combination-of-radam-lookahead-for-the-best-of-2dc83f79a48d

Next Meetup:
Late October

www.meetup.com/Vienna-Deep-Learning-Meetup

Vienna
Deep Learning
Meetup