Deep Learning for Fun and Profit, Pt. 2.25
Alexander CS Hendorff
Alexander C. S. Hendorf

- Partner & Senior Consultant Information Technology Consulting & Training on Data Science Matters
- Program Chair EuroPython, EuroSciPy, PyConDE & PyData, PSF Managing Member
- Speaker Europa & USA MongoDB World New York / San José, PyCon Italy, CEBIT Developer World, BI Forum, IT-Tage FFM, PyData, PyParis,...

- ah@koenigsweg.com
- @hendorf
We think digital excellence. We do digital excellence.

- STRATEGY & INNOVATION
- DATA & ARTIFICIAL INTELLIGENCE
- CREATIVE & COMMUNICATION
- SCALABILITY & GROWTH
- OPEN ECOSYSTEMS & CONNECTIVITY
The Story So Far...

How I Made My Computer Write it's First Short Story
*PyData Berlin 2018*

[https://www.youtube.com/watch?v=jCl02teu1k](https://www.youtube.com/watch?v=jCl02teu1k)

Deep Learning with PyTorch for Fun and Profit
*EuroPython 2018*

[https://www.youtube.com/watch?v=_uNgNhII87U](https://www.youtube.com/watch?v=_uNgNhII87U)
The Inspiration

Gene Kogan
Picasso's terminal
data science and AI in the visual arts

PyData London 2017

https://www.youtube.com/watch?v=JXx2MYtAU3o
The lab: eGPU

- External Box, 500W, carries a GPU
- Thunderbolt 3 Connection
- NVIDIA GTX-1080
- It’s not *plug-and-play*
- great source for troubleshooting and advice [egpu.io](http://egpu.io)
The Journey #1: Style Transfer

A Neural Algorithm of Artistic Style
2015 Leon A. Gatys, Alexander S. Ecker, Matthias Bethge
https://arxiv.org/abs/1508.06576
Zusammen mit anderen freien Menschen schaffen Valerian und Laureline den Weg zur Hauptstadt.

Ein Praktisches Märchen? Der Ort ist die berühmte Kapselnation und ein Land im Halbmond.

Ohne Übersetzung, die Wörter klingen sind menschlich nicht, obwohl sie aus einem fremden Sprach nicht auf.

Die Steine, die sich für die Menschheit verwenden sind. Zäube- raum aber wurden laufen nicht ich mit siehern übersprechen nicht.

Einer mal Valerian!

Einem, mit dem Reisen, ein undendes, so sehr, einem nicht zu gehen.

Selten, im düsteren Museum die Galaxie ein, versteckte sich die deird.
I'll be talking about some Deep Learning use cases for content creation at 10:30
#EuroPython Edinburg #PyData

ep2018.europython.eu/conference/tal ...

Thomas Rackow @TRackow · 27. Juli
Antwort an @hendorf

Thomas Rackow @TRackow · 27. Juli
I took this picture a few days ago and was wondering when you posted that drawing.

Thomas Rackow @TRackow · 27. Juli
Oh, I see. Even more amazing.

Alexander CS Hendorf @hendorf · 27. Juli
It's not a drawing, it was generated with artificial intelligence
L'Empire des PyParis
The Journey #2: The Vision
Die drei ???

- American juvenile detective book series „The Three Investigators“
- Investigations include mysteries of baffling phenomena (e.g. whispering mummy)
- 1964

- Very popular as books and radio dramas in Germany
- ~200 taped radio dramas (Hörspiele) published 1979 – today
- Radio dramas have sold more than 45 million, the books ~ 16 million copies in Germany
Hörspiel?

- Taped radio drama
- Started in radio in the 1920ies
- Features:
  - Voice recordings
  - Natural or electroacoustic sounds,
  - Noises as well as silence
  - Cuts
  - Mixes
  - Music
Die drei ??? und die flüsternde Mumie

- Hörspielskript von FlukeSkywater –

Audio

Anfangsmusik

Im Hintergrund Arbeitslärm und Hundegebell.

Blacky Hilfe! Hilfe! Rettet mich! Hilfe!
Mathilda Justus! Du hast den Vogel zu viel fernsehen lassen. Er redet ja wie im Krimi. (Papageiengekrächze)
Justus (schnauft) Ganz recht, Tante Mathilda. Mmh, wo soll ich die Tür hin tun?
Bob (widerwillig) Ja ...
Justus Die Arbeit ist schwer, Tante Mathilda. Wir könnten eine Pause gebrauchen.
Mathilda Aach, ich verstehe - ihr habt gesehen, dass die Post gekommen ist, und ihr seid neugierig. Ihr glaubt wohl, dass ein Brief für euch dabei ist.
Justus (keucht) Ist einer dabei, Tante Mathilda?
Ingredients to be Synthesised

**Story / Plot**
- What’s happening?

**Dialogues**
- Dynamic human speech

**Cover**
- Picture

**Spoken Word**
- Spoken Word by new and returning characters
Questions to Be Explored

- Resources required for individual steps
  - AI/Neural Network
  - Data acquisition / processing
- Costs (& time)
- Which real problems can we potentially solve in the future?
The Journey #3: Realisation

- transcripts / text sources to learn from
- recordings
Technology Selection

- Local Lab eGPU
- Ability to scale to cloud
- Reproducibility & documentation of experiments
- PyTorch
The Process

- Data acquisition
- Data cleansing
- Research a paper / working solution
- Verify it’s working
- Adapt the solution for our use case
- Maximize quality
Text

- Fabricate a text to be spoken by a character
The Unreasonable Effectiveness of Recurrent Neural Networks

May 21, 2015

There’s something magical about RNNs. I’ve always had a recurrent network for Image Captioning and have witnessed their power and robustness, but I’ve in fact reached the opposite conclusion. Sometimes, intuition turns out of it blows past your expectations. Time was that the common wisdom was that there is no structure in the data. I’ve in fact reached the opposite conclusion. The power and robustness of RNNs have been witnessed. This post is about sharing some of this.

http://karpathy.github.io/2015/05/21/rnn-effectiveness/

Shakespeare

It looks like we can learn to spell English words. But how about if there is more structure and style in the data? To examine this I downloaded all the works of Shakespeare and concatenated them into a single (4.4MB) file. We can now afford to train a larger network, in this case let’s try a 3-layer RNN with 512 hidden nodes on each layer. After we train the network for a few hours we obtain samples such as:

PANDARUS:
Alas, I think he shall be come approached and the day
When little strain would be attain’d into being never fed,
And who is but a chain and subjects of his death,
I should not sleep.

SECOND SENATOR:
They are away this miseries, produced upon my soul,
Breaking and strongly should be buried, when I perish
The earth and thoughts of many states.
Process

- Fan-made transcripts from PDF
- Pre-processing with Textract
- Identify characters via names other fan page
  296 unique names identified
- German language, 1,521,382 words, 56,674 unique
Process

- Verified produced something *Shakespeary*
- Fed with the dialogue corpus
"wildes der prücks und nun bei sie redaste sich es hinden struch wenn hatten spielte schlokonianer lüst!"

"Wild of the pruecks and now they redaste themselves it had hunched when Schlokonianer lüst had played!"
Results Character Based

wie sein bericht justus fusspielte missam hächten durch das bob doch wir paren ermob war der mathilbek

Harald Schmidt

Sprecht ihr vong?

Das ist ein Neuronales Netzwerk, das
Results: Word Based

“nicht zitrone dran schlammgrünen welcher peter ins all will”

"Not lemon on the mud green which peter wants in all"

“morton ist sich noch nicht einig einen wollt kerlen müssten”

“Morton is not yet agreed one would like to have Kerlen”
Chat Bot To The Rescue?

https://arxiv.org/abs/1708.00818
Artwork

- *Fabricate artwork meeting the style of the series*
A Neural Algorithm of Artistic Style
Die drei ???
und der Super-Papagei
1

Die drei ???
und die flüsternden Puppen
180

Die drei ???
und der seltsame Wecker
12

Die drei ???
und die flüsternde Mumie
10

Die drei ???
Das Grab der Inka-Mumie
163

Die drei ???
und der verschollene Pilot

Die drei ???
und der riskante Ritt
51

Die drei ???
und der Karpatenhund
Python Version

- Python 2.7 support ends by 2020
- All projects should really use Python 3 now
PyRand

Python Code Quality

- Keywords often used for variables, e.g.
  - input
  - in
  - list
  - > use _, e.g. _input

- Closures:
  functions reference variables outside the function

```python
In [1]:
1 a = 2
2 b = 3

5 def mysuperfunction():
 6 a = 100
 7 input = a * b
 8 return input
9
10 mysuperfunction()

Out[1]: 300
```
Alex @hendorf has learnt that the internet includes
  - untruths
  - cherry picking
  - bad code.

Use with care 😐.

#EuroPython
Updates!

-
Speech

- Fabricate spoken language
- Make it sounds like a known character
Process: Fabricate

- LJ Speech Dataset (English)
  https://keithito.com/LJ-Speech-Dataset

- Natural TTS Synthesis by Conditioning WaveNet on Mel Spectrogram Predictions, Jonathan Shen, Yuxuan Wang and Zongheng Yang et al.

- Tacotron2 implementation used:
  NVIDIA/tacotron2
Training the Network

- Requires:
  - audio snippets and
  - text of the audio snippet
Once upon a time there was a little mermaid named Siren, who lived with her step mother under the sea, She didn't get to go out of the Sea like any other.
Let’s Do This in German!

- Requires:
  - audio snippets and
  - text of the audio snippet
  - In GERMAN
We Can Do This in German!

- **Task:** building a German corpus to learn with
  - audio snippets
  - text of the audio snippet
- **Options**
  - audiobooks
  - newspapers with audio
  - ...
- **Requirements**
  - humans to prepare the dataset
Semi-automation of Dataset Generation

Audio version of newspaper article

Word: time index

Transcribe audio via Cloud service

Sentencees

audio – text snippet

audio – text snippet

audio – text snippet
Assumptions

1. Transcription service will deliver a very good quality
2. Transcription service a generally available at all providers and very cheap
3. It should be easy to generate a new corpus
Assumptions

1. Transcription service will deliver a very good quality
2. Transcription service a generally available at all providers and very cheap
3. It should be easy to generate a new corpus

Findings German

Quality is ok, but transcription is not accurate, many artefacts

(+) Google (N/A) Azure (N/A) IBM (N/A) AWS
Offline libs (-) Apple (-)

Not so simple...
The Dataset

- Newspaper article read by ~25 speakers
- Article text scraped from website
Semi-automation of Dataset Generation

Audio version of newspaper article

Word: time index

Transcribe audio via Cloud service

Sentencees

audio – text snippet

audio – text snippet

audio – text snippet
Deutschland darf kein Talent vergeuden, so heißt es immer wieder. Wo aber bleiben in unseren Schulen die Angebote für die besonders Talentierten?

Germany must not waste any talent, it is said again and again. But where are the offers for the most talented in our schools?
Challenges

Lessons learned
Assumptions

1. Neural Networks are hard to set up
2. Neural Networks are hard to train
3. Neural Networks are smart
4. Data is like the new Oil
5. The singularity is coming 2045 (Ray Kurzweil)
Assumptions

1. Neural Networks are hard to set up. Knowledge is very accessible, open-source software
2. Neural Networks are hard to train
3. Neural Networks are smart
4. Data is like the new Oil
5. The singularity is coming 2045 (Ray Kurzweil)
Assumptions

1. Neural Networks are hard to set up. Knowledge is very accessible, open-source software.
2. Neural Networks are hard to train. Parameters are known or can be auto-tuned.
3. Neural Networks are smart.
4. Data is like the new Oil.
5. The singularity is coming 2045 (Ray Kurzweil).
As a result, the total number of the members of the tribe of the sons of Gershom was one hundred fifty thousand.
Passing a Chicken through an MNIST Model

14 Mar 2018

When you put a picture of a chicken through a model trained on MNIST, the model is 99.9% confident that the chicken is a 5. That’s not good.

This problem does not just relate to chickens and digits but the fact that a neural net makes very confident predictions on data that does not come from the same distribution as the training data. While this example is artificial, it is common in practice for a machine learning model to be used on data that is very different from the data it was trained on. A self-driving car, for example, may encounter an unusual environment that was never seen during training. In such cases, the system should not be overly confident but instead let the driver know that it is not able to make a meaningful prediction.

Images from MNIST and a chicken.

https://emiliendupont.github.io/2018/03/14/mnist-chicken/
Breaking Linear Classifiers on ImageNet

Mar 30, 2015

You’ve probably heard that Convolutional Networks work very well in practice and across a wide range of visual recognition problems. You may have also read articles and papers that claim to reach a near “human-level performance”. There are all kinds of caveats to that (e.g. see my G+ post on Human Accuracy is not a point, it lives on a tradeoff curve), but that is not the point of this post. I do think that these systems now work extremely well across many visual recognition tasks, especially ones that can be posed as simple classification.

Yet, a second group of seemingly baffling results has emerged that brings up an apparent contradiction. I’m referring to several people who have noticed that it is possible to take an image that a state-of-the-art Convolutional Network thinks is one class (e.g. “panda”), and it is possible to change it almost imperceptibly to the human eye in such a way that the Convolutional Network suddenly classifies the image as any other class of choice (e.g. “gibbon”). We say that we break, or fool ConvNets. See the image below for an illustration:

Figure from Explaining and Harnessing Adversarial Examples by Goodfellow et al.

http://karpathy.github.io/2015/03/30/breaking-convnets/
Assumptions

1. Neural Networks are hard to set up. Knowledge is very accessible, open-source software.
2. Neural Networks are hard to train. Parameters are known or can be auto-tuned.
3. Neural Networks are smart. Neural Networks are blinkered specialists.
4. Data is like the new Oil.
5. The singularity is coming 2045 (Ray Kurzweil).
Opinion: Data isn’t the new oil — it’s the new nuclear power

Data is a valuable, powerful commodity — but unlike oil, it is unlimited in quantity and in its capacity for harm, says technology thinker James Bridle.

James Bridle

https://networkingnerd.net/2017/12/20/data-is-not-the-new-oil-its-nuclear-power/

Managing Expectations on Resources Required
Reality Resources Required
What do machine learning practitioners actually do?

Written: 12 Jul 2018 by Rachel Thomas

This post is part 1 of a series. Part 2 is an opinionated introduction to AutoML and neural architecture search, and Part 3 looks at Google’s AutoML in particular.

There are frequent media headlines about both the scarcity of machine learning talent (see here, here, and here) and about the promises of companies claiming their products automate machine learning and eliminate the need for ML expertise altogether (see here, here, and here). In his keynote at the TensorFlow DevSummit, Google’s head of AI Jeff Dean estimated that there are tens of millions of organizations that have electronic data that could be used for machine learning but lack the necessary expertise and skills. I follow these issues closely since my work at fast.ai focuses on enabling more people to use machine learning and on making it easier to use.

In thinking about how we can automate some of the work of machine learning, as well as how to make it more accessible to people with a wider variety of backgrounds, it’s first necessary to ask, what is it that machine learning practitioners do? Any solution to the shortage of machine learning expertise requires answering this question: whether it’s so we know what skills to teach, what tools to build, or what processes to automate.
Assumptions

1. Neural Networks are hard to set up. Knowledge is very accessible, open-source software
2. Neural Networks are hard to train up. Parameters are known or can be auto-tuned
3. Neural Networks are smart. Neural Networks are blinkered specialists
4. Data is like the new Oil. Please stop saying this
5. The singularity is coming 2045 (Ray Kurzweil)
Assumptions

1. Neural Networks are hard to set up
2. Neural Networks are hard to train
3. Neural Networks are smart
4. Data is like the new Oil
5. The singularity is coming 2045 (Ray Kurzweil)

"Maschinen ist alles egal" – Interview with Margaret Boden

ZEITmagazin: Many contemporaries see these advances with concern. They are frightened of a world where machines dominate people. The inventor and futurologist Ray Kurzweil, now director of technical development at Google, even claims that superhuman artificial intelligence is near. Around the year 2045, it will become reality. Do you believe him?

Boden: I asked him if he actually believes what he writes.

ZEITmagazin: And?

Boden: He really believes it.

ZEITmagazin: Can you dismiss Kurzweil as a spinner?

Boden: no. He has done an extraordinary job for the advancement of technology. I use the word genius very rarely. With Kurzweil I would be ready for it. Maybe someday there will be an artificial intelligence superior to us. But certainly not in this century. Things are a lot more complicated than people imagine. Kurzweil's prediction is just crazy.
AI vs. Marketing B-Bingo

Alexander - hast recht 😜 wir entwickeln Software, welche in unterschiedlichen Systemen implementiert oder embedded wird.
Next Steps

- Look into story generation
- Training a better model for German
  ???-like text
- Improving speech synth by adding
  more data
- Multi-style transfer

- Auto selection of good models
- Data pre-tuning to optimize output
- Scene agnostic (sub)-style selection
Summary

Are Neural Networks great and useful?
Summary

Are Neural Networks great and useful?

Of course they are!
Thank you!

Q & A

We're hiring

koenigsweg.com  ah@koenigsweg.com
@hendorf
Kontakt

Königsweg GmbH

- Musikpark Mannheim
  Hafenstraße 49
  68159 Mannheim

- Mafinex Technologiezentrum
  Julius-Hatry-Straße 1
  68163 Mannheim

Telefon: +49 621 43 74 10 22
Telefax: +49 621 43 74 10 25
E-Mail: info@koenigsweg.com
Web: www.koenigsweg.com