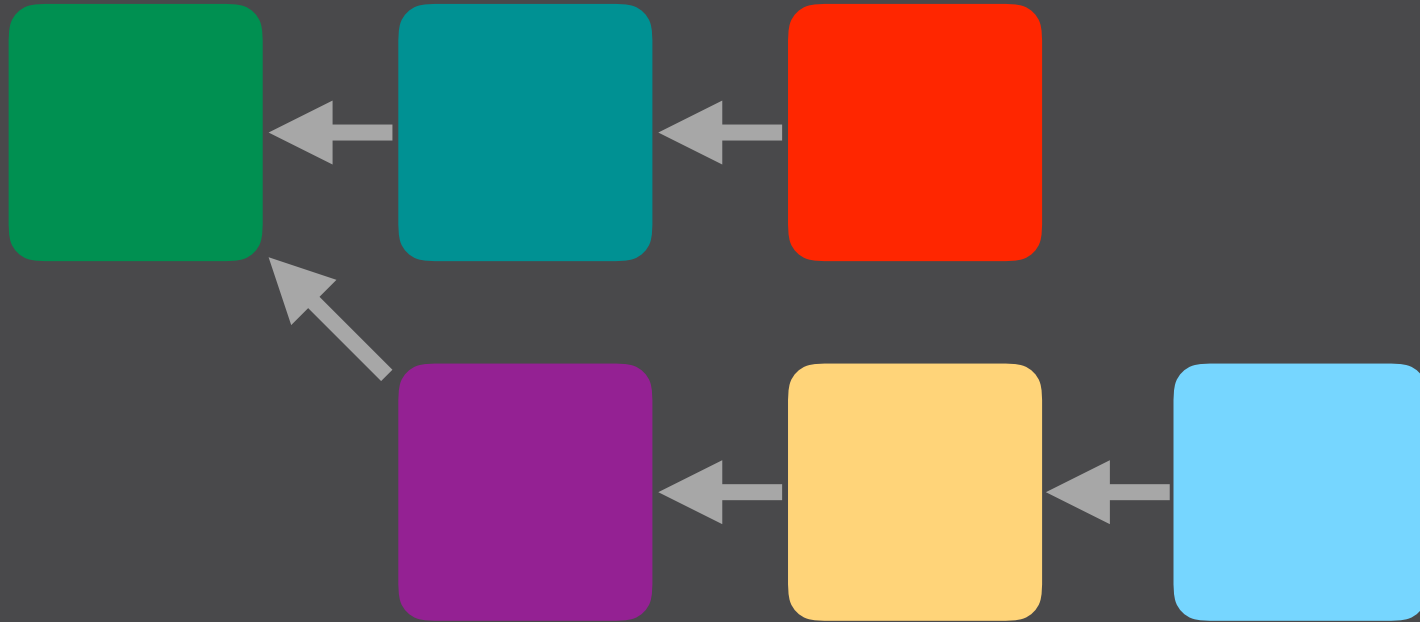
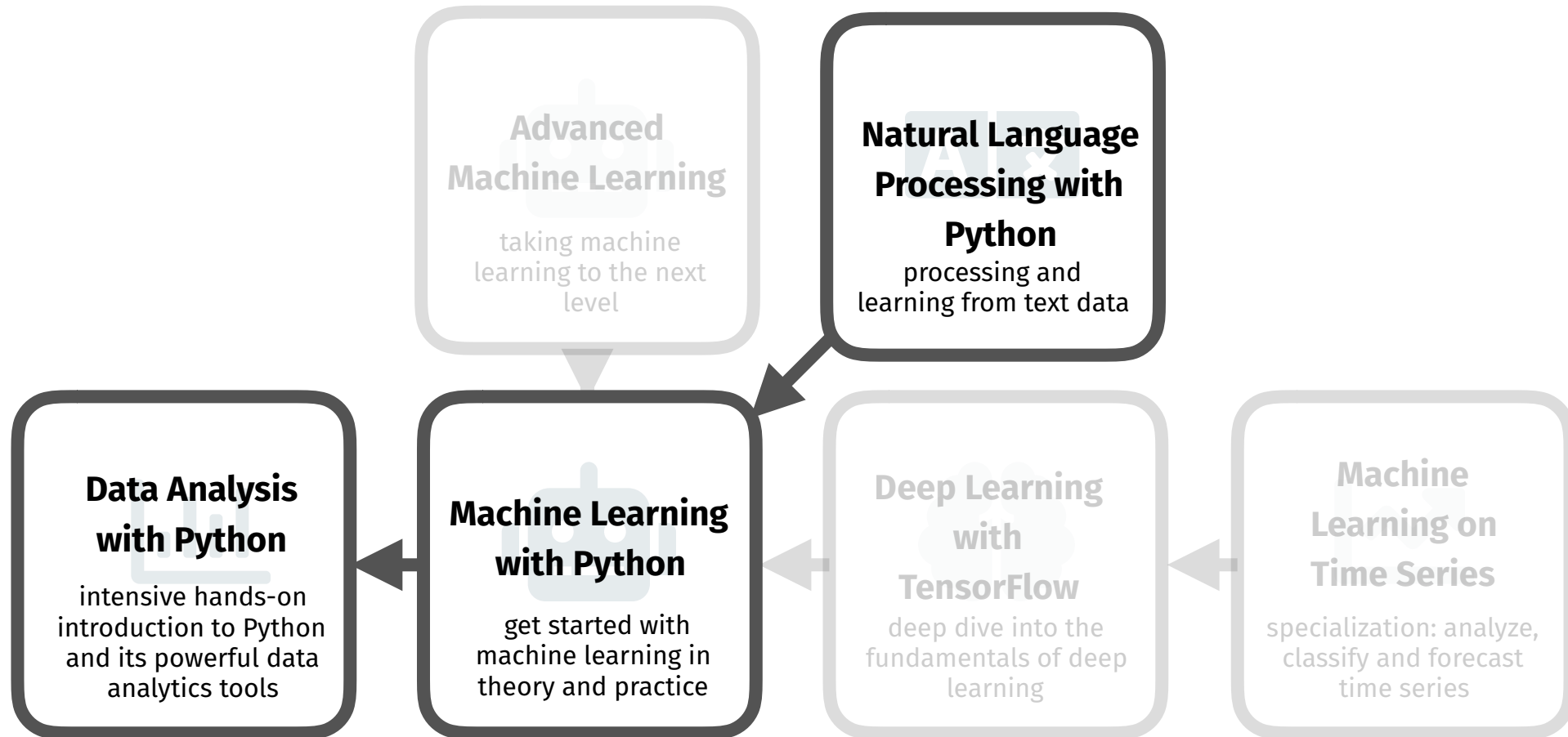


data science 101: skill building blocks for the data-driven business

Dr. Christian Staudt · data scientist · in cooperation with **point8**
data matters.





Natural Language Processing with Python

processing and learning
from text data

Acronym: NLPP
Level: Advanced
Duration: 2 days

This course is an introduction to natural language processing techniques, using Python - a useful specialization for those looking to apply data analysis and machine learning to text data.

Curriculum

- 1. Extracting Text from Documents**
Extracting natural language texts from various sources and formats.
- 2. Natural Language Processing: The Basic Toolkit**
Tokenization, stemming, and other fundamental tools for slicing and dicing text data.
- 3. Document Similarity for Clustering and Recommendation**
How to measure the similarity of texts and do useful things with that.
- 4. Text Classification**
Machine learning for classifying text documents.
- 5. Sentiment Analysis**
How to deal with more comments than you can possibly read.
- 6. Topic Modelling**
An unsupervised ML technique for discovering topics in text documents.
- 7. Word Embeddings**
Mapping words to vectors for fun and profit.

Setting

- up to 16 participants
- we provide a distraction-free workshop location

Didactics

- we aim for a healthy mix of lecture, demonstration, exercises (individual, pair programming and team exercises) and Q&A
- workshops are participatory - we integrate your questions and use cases

Material

- Jupyter Notebooks: interactive documents with instructions, code and graphics
- will be provided free of charge for self-study after the course

Language

- course language: English or German
- materials in English



Dr. Christian Staudt
Freelance Data Scientist

Technical Track Instructor Bio

- **Independent Data Scientist**

July 2016 – now

data science contract roles in multiple industries - analytics applications, big data solutions, machine learning, artificial intelligence

- **Doctoral Researcher at Karlsruhe Institute of Technology (KIT)**

October 2012 – June 2016

algorithms, parallelism, data analysis, software engineering

Education

- Doktor der Naturwissenschaften (Dr. rer. nat.) (= PhD), Computer Science, Karlsruhe Institute of Technology (KIT)
- Diplom (= Master's Degree), Computer Science, 2005 – 2012, Karlsruhe Institute of Technology (KIT)



Experimentelle Physiker mit Innovationsdrang, interessiert an Zukunftstechnologien und immer auf der Suche nach neuen Herausforderungen. Das Point 8 Team bringt Big Data Know-how vom CERN in die Wirtschaft und unterstützen Unternehmen und Organisationen mit Datenanalyse, Machine Learning und Simulationen.

our track record

We continually update and improve **data science 101** according to our experience with previous workshops and industry data science projects. Together, we have so far conducted **over 30 successful workshops for industry leaders in the machine building, telecommunications and energy sectors.**

References (public)



get in touch

.... for bookings and inquiries about data science 101, including exclusive and custom workshops. **Current courses are listed at:** <https://point-8.de/trainings.html>



Dr. Christian Staudt
Freelance Data Scientist
Hamburg

email: mail@clstaudt.me

web: <http://clstaudt.me>

twitter: @C_L_Staudt



Data Science as a Service
Dortmund

email: kontakt@point-8.de

web: <http://point-8.de>

twitter: @point8hq