Universität Konstanz

01 | Logistics & Motivation

Max Pellert (https://mpellert.at)

Deep Learning for the Social Sciences

https://mpellert.at

Universität Konstanz

Max Pellert   Courses & Talks

Max Pellert

Assistant Professor
University of Mannheim

Download CV as PDF

Professor for Social and Behavioural Data Science (interim, W2) at the University of Konstanz

Assistant Professor (Business School of the University of Mannheim)

I worked in industry at SONY Computer Science Laboratories in Rome, Italy

PhD from the Complexity Science Hub Vienna and the Medical University of Vienna in Computational Social Science

Studies in Psychology and History and Philosophy of Science

Msc in Cognitive Science and Bsc in Economics (both University of Vienna)

(Some) Research interests

Universität
Konstanz

- Computational Social Science
- Digital traces
- Affective expression in text
- Natural Language Processing
- Collective emotions
- Belief updating
- **Psychometrics of AI**

Pellert, M., Lechner, C. M., Wagner, C., Rammstedt, B., & Strohmaier, M. (2024). AI Psychometrics: Assessing the Psychological Profiles of Large Language Models Through Psychometric Inventories. *Perspectives on Psychological Science*. https://doi.org/10.1177/17456916231214460

Giordano de Marzo

Universität Konstanz

PostDoc at the University of Konstanz

Junior Research Fellow of the Complexity Science Hub Vienna

Consultant of the International Labour Office

Phd in Physics at Sapienza University, Enrico Fermi Research Center and Sapienza School for Advanced Studies in Rome

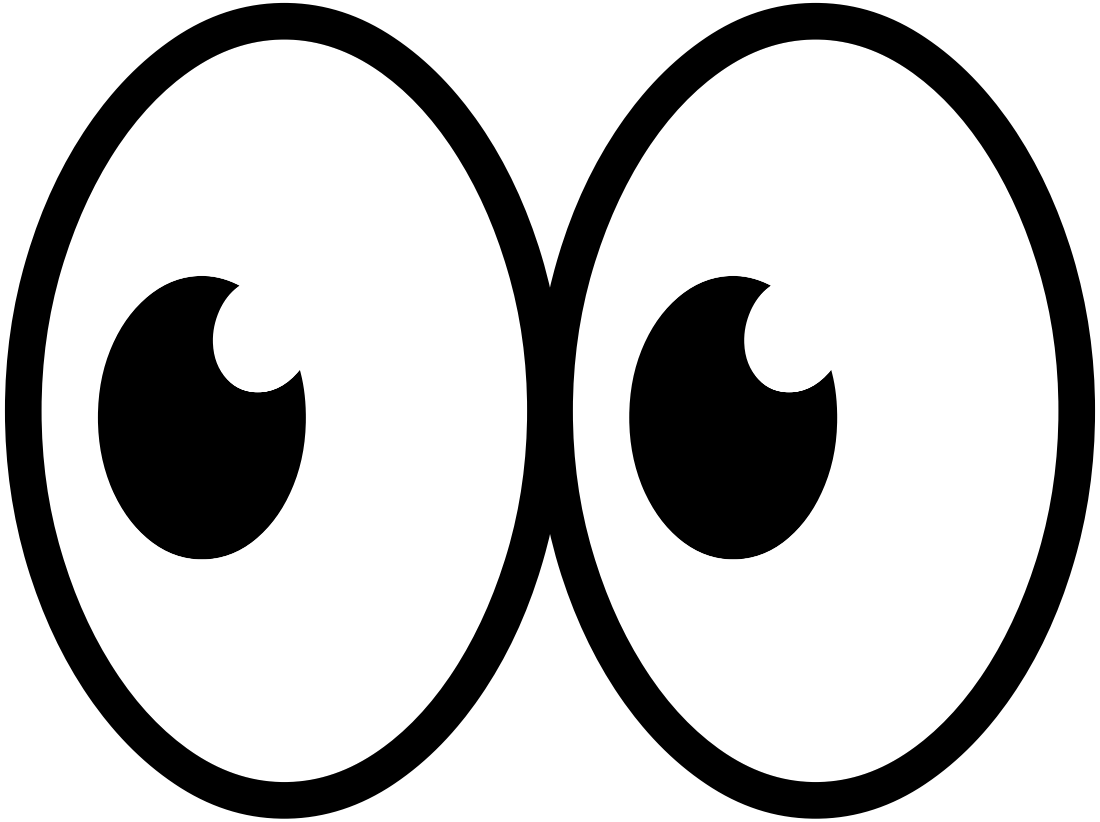
MSc in Theoretical Physics

Giordano de Marzo

Research Interests

- Complex Digital Systems
- Social Networks
- Recommendation Algorithms
- Large Language Models
- AI

Who are you?



Universität Konstanz

Course program

Date	Topic	Who?
9.4.	Logistics & Motivation	Max
16.4.	Supervised Learning	Max
23.4.	Shallow Neural Nets	Max
30.4.	Perceptron and Multi Layer Perceptrons	Giordano
7.5.	Convolutional Neural Networks	Giordano
14.5.	Graph Neural Networks	Giordano
21.5.	NN for Time Series analysis	Giordano

Date	Topic	Who?
28.5.	No class	
4.6.	Generative Deep Learning 1	Giordano
11.6.	NLP 1	Max
18.6.	NLP 2	Max
25.6.	Reinforcement Learning	Giordano
2.7.	Large Language Models	Max
9.7.	Generative Deep Learning 2	Giordano
16.7.	Outlook	Max

Logistics

Lectures on Tuesday 10.00 - 11.30 in C421

Exercises (practical sessions) are provided over the semester on Wednesday 13:30 - 15.00 in D430

Your tutor will be Andri Rutschmann, he will co-teach tutorials with us

Assignments will be released on Tuesday evening or the latest Wednesday before the tutorial

The deadline is before a tutorial a few weeks later

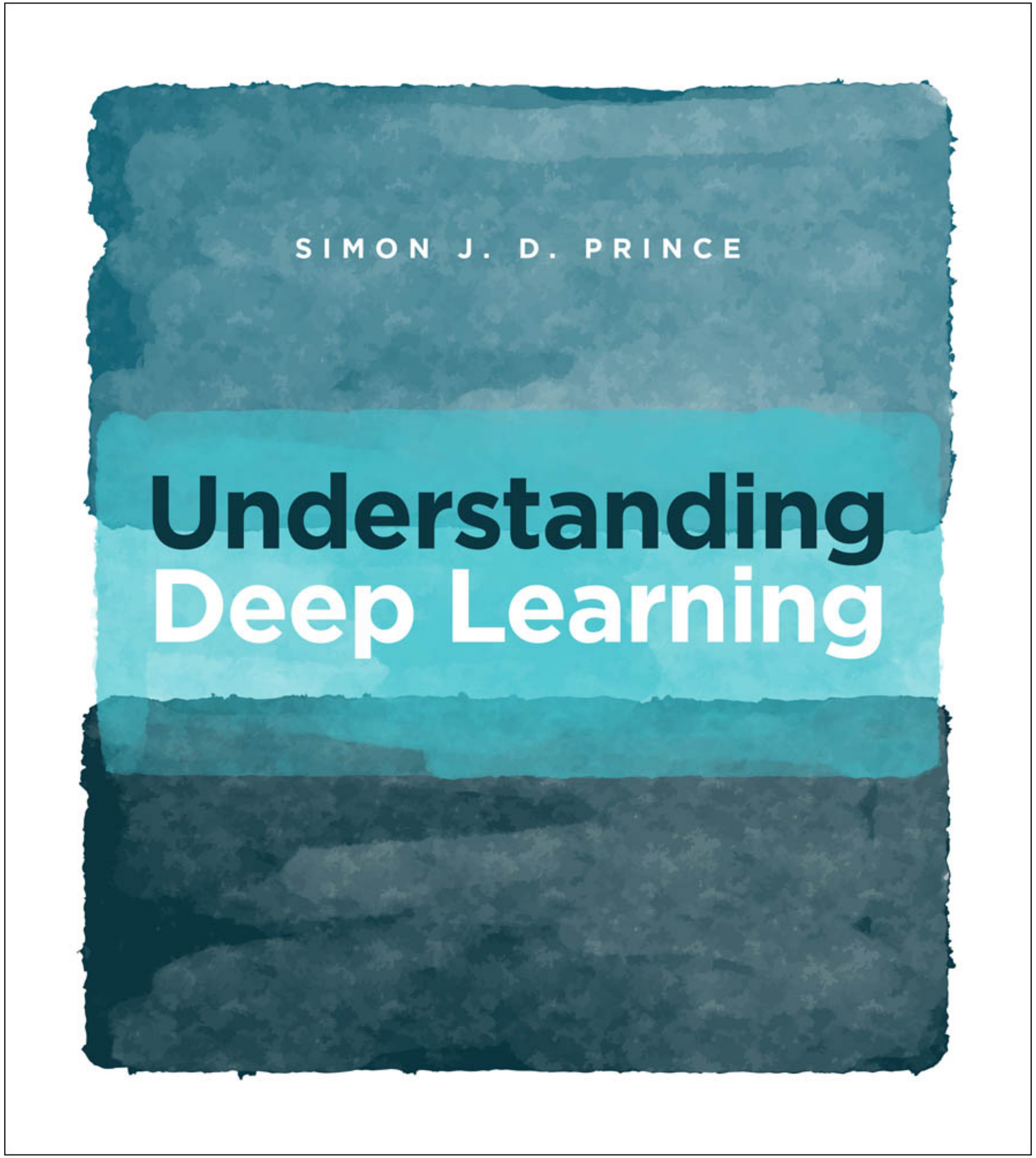
Assignment submissions through Github as in ICSS

4 assignments count 40% of the final grade of the course (10% each)

Project

Counts 60% of the grade

More information about the project will be delivered soon...


SIMON J. D. PRINCE
Understanding
Deep Learning

“Course Book”

Universität Konstanz

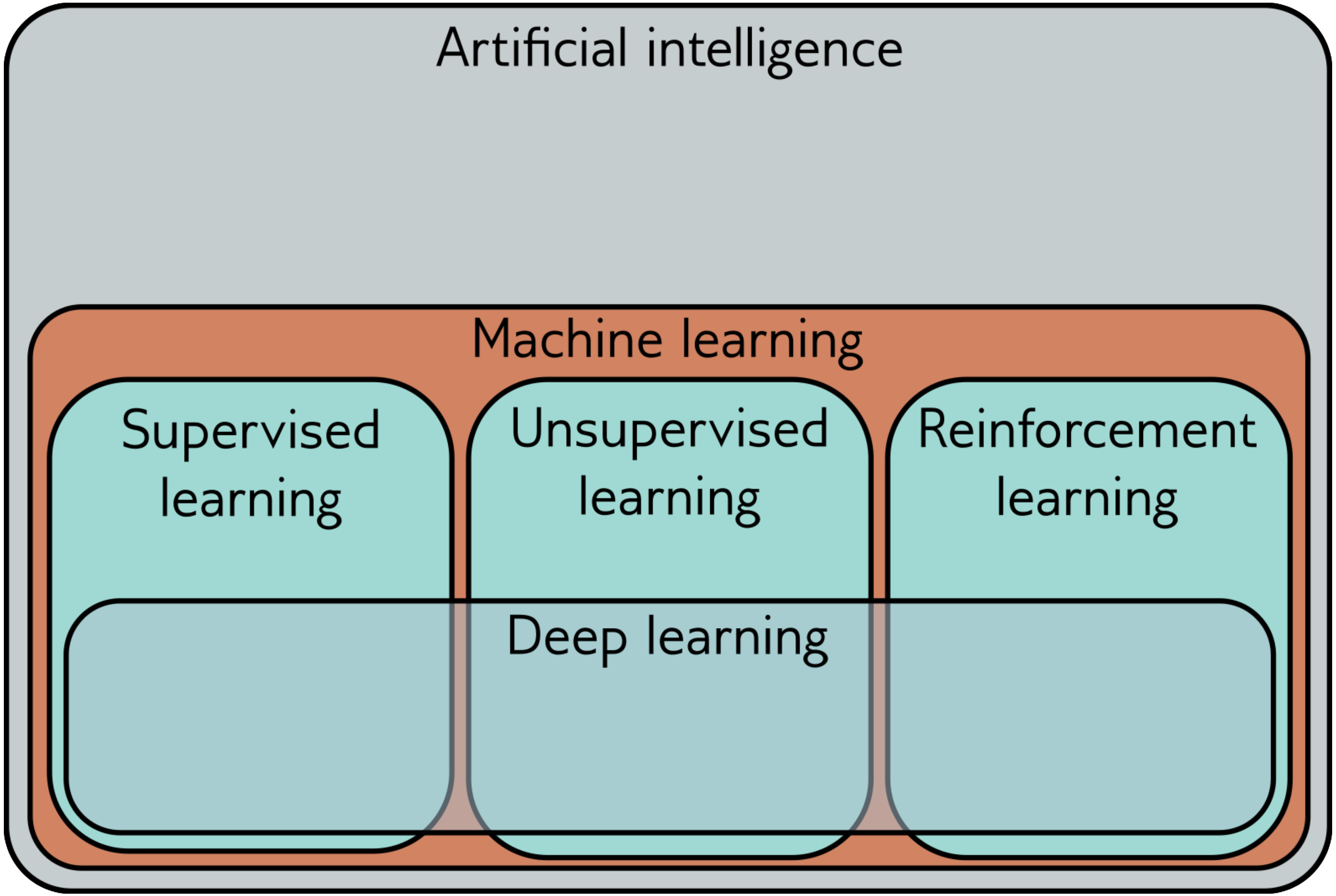
Prince, S. J. D. (2023). Understanding deep learning. The MIT Press.

Available in print or for free as a PDF:
https://udlbook.github.io/udlbook/

On the webpage you will find many additional materials

We will cover many topics from the book and it also helps you as additional materials to deepen your knowledge on specific aspects in-depth

In addition to the contents covered in the book, in this course we aim to keep the focus on applications in the social sciences


Artificial intelligence
Machine learning
Supervised learning
Unsupervised learning
Reinforcement learning
Deep learning

Supervised Learning

Basic workflow: Define a mapping from input to output

Learn this mapping from paired input/output data examples

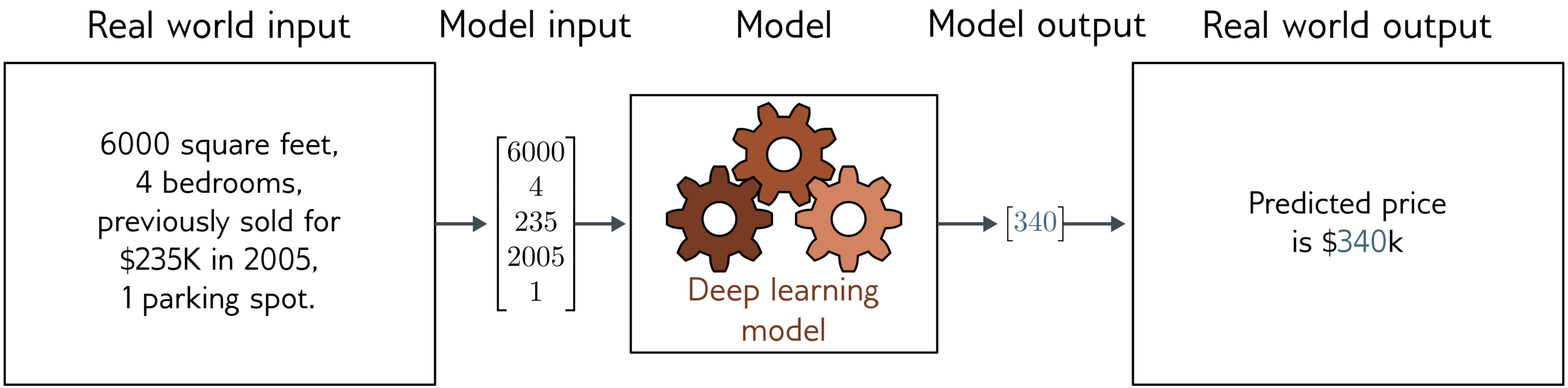
Often, the examples are from data sets of inputs that have been manually annotated by humans, i.e. the output are human-labeled *supervisory* signals

Often, the annotation is done by crowdworkers (if the task is not already outsourced to another model)

Veselovsky, V., Ribeiro, M. H., & West, R. (2023). Artificial Artificial Artificial Intelligence: Crowd Workers Widely Use Large Language Models for Text Production Tasks (arXiv:2306.07899). arXiv. http://arxiv.org/abs/2306.07899

Regression

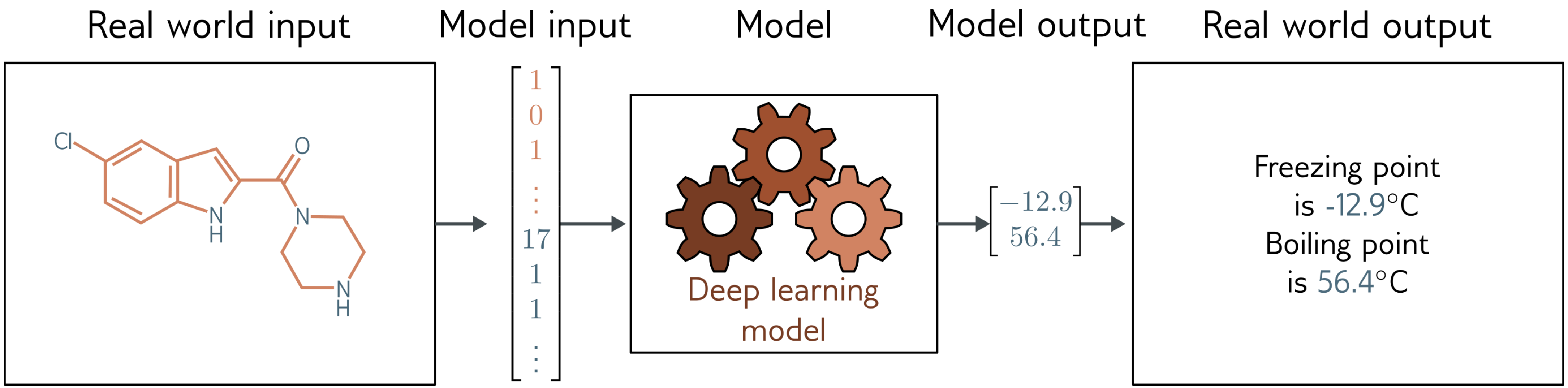
Universität Konstanz



Univariate regression problem (one output, real value)

Fully connected network

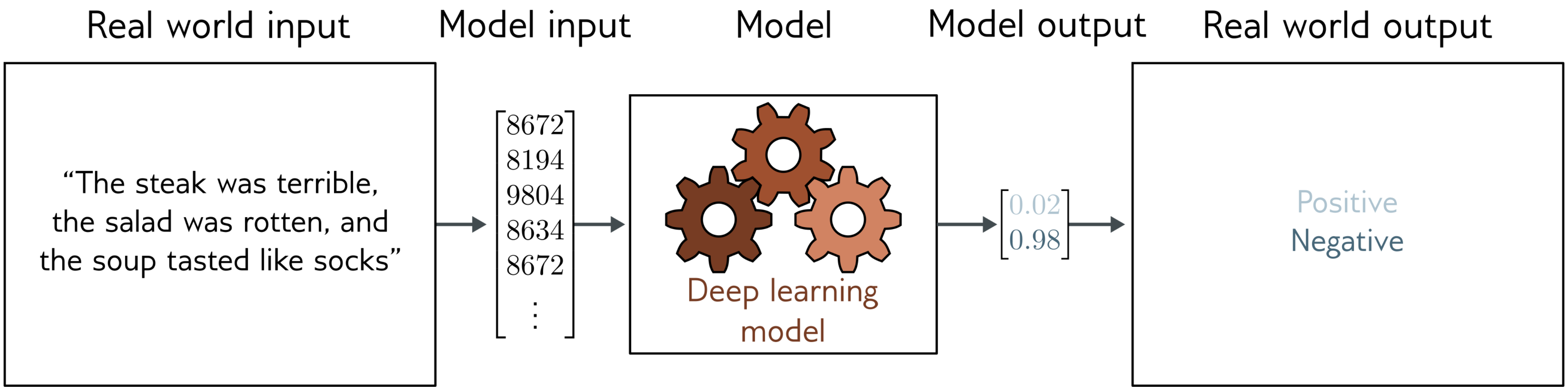
Graph Regression



Multivariate regression problem (>1 output, real value)

Graph neural network

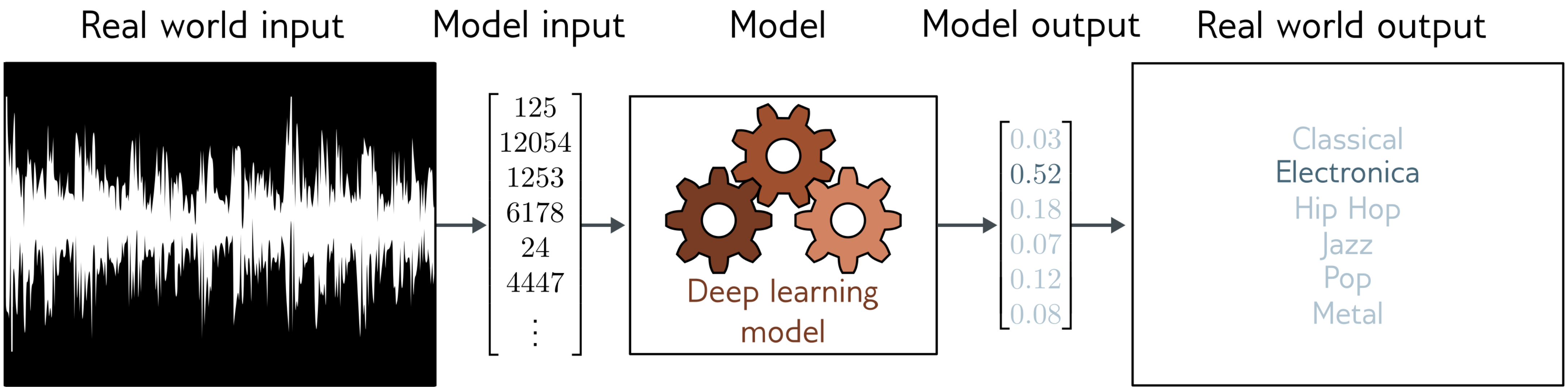
Text classification



Binary classification problem (two discrete classes)

Transformer network

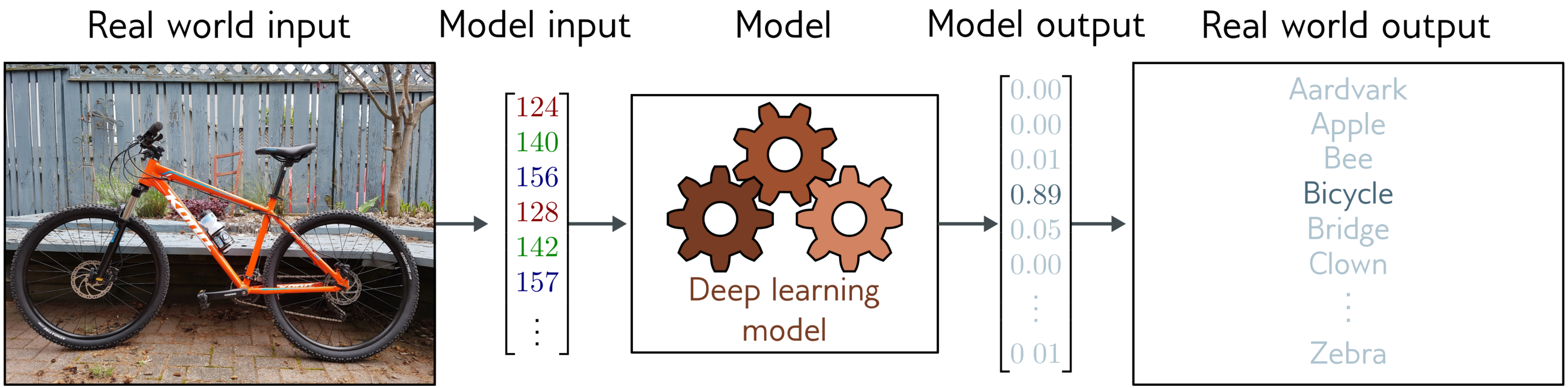
Music genre classification



Multiclass classification problem (discrete classes, >2 possible values)

Recurrent neural network (RNN)

Image classification

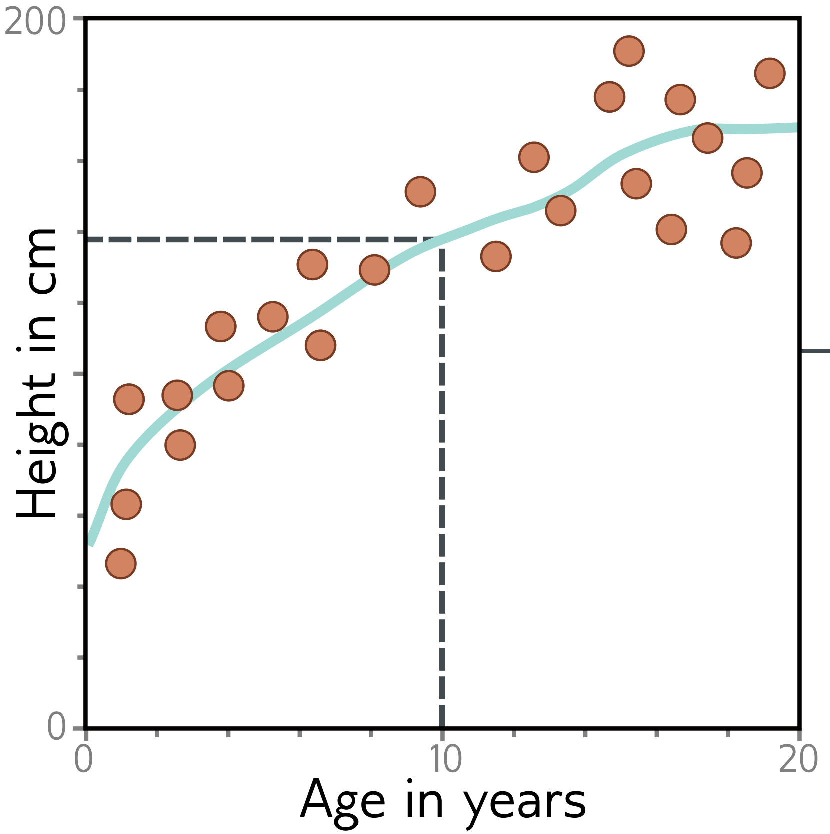


Multiclass classification problem (discrete classes, >2 possible classes)

Convolutional network

Age of child
is 10 years

[10]

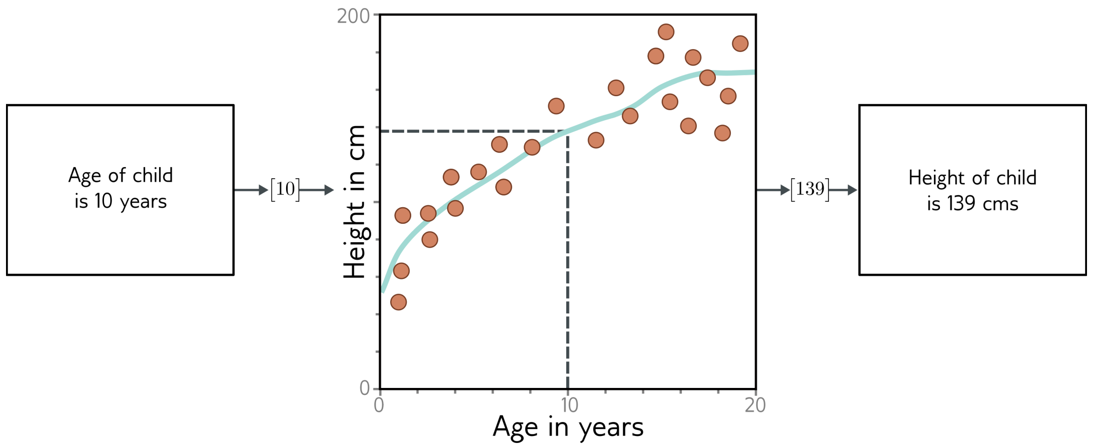


[139]

Height of child
is 139 cms

What is a supervised learning model?

Universität Konstanz



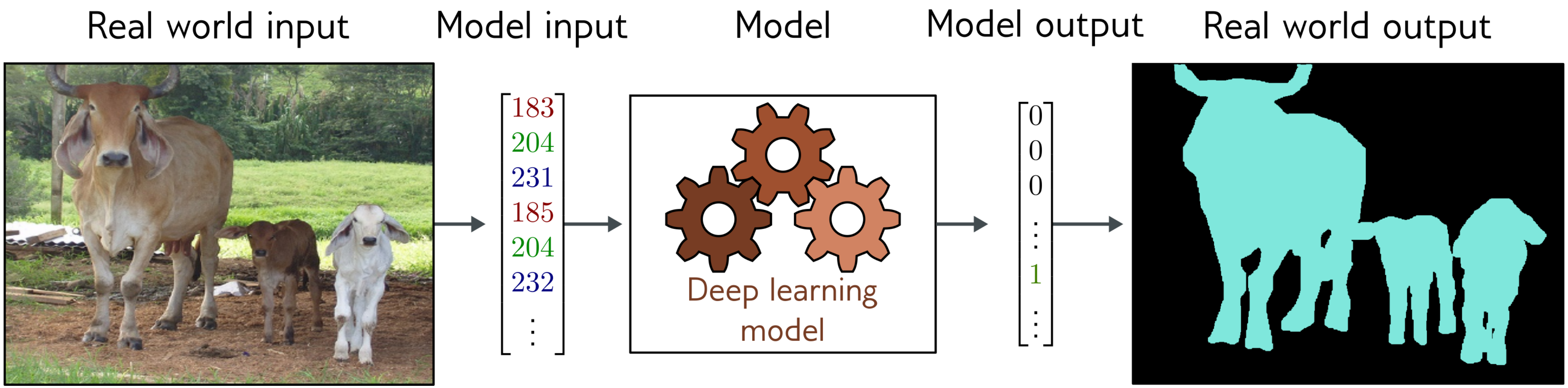
An equation relating input (age) to output (height)

Search through family of possible equations to find one that fits training data well

Deep neural networks are just a very flexible family of equations

Fitting deep neural networks = “Deep Learning”

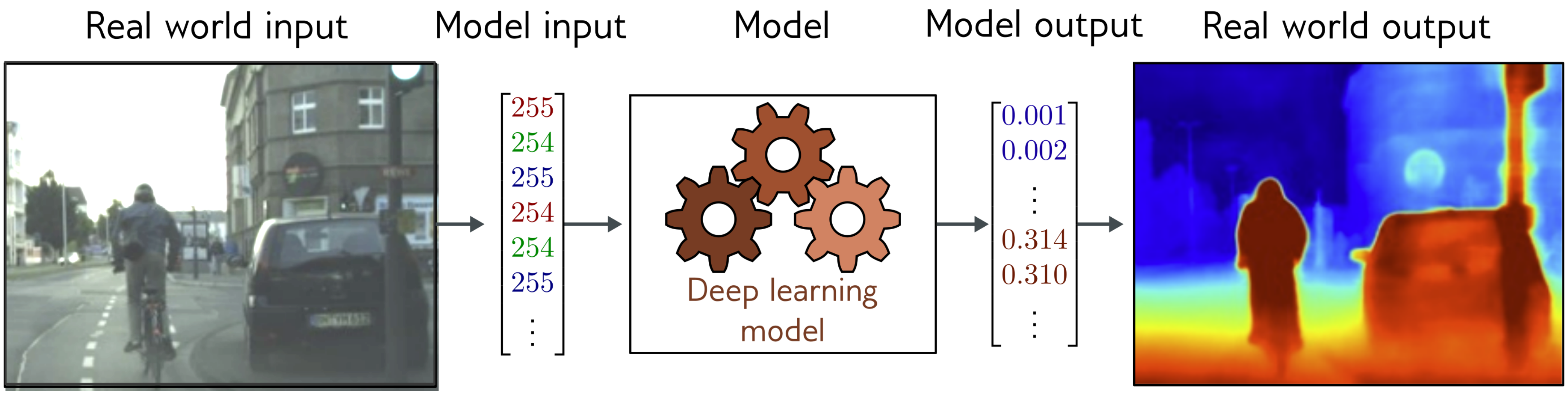
Image segmentation



Multivariate binary classification problem (many outputs, two discrete classes)

Convolutional encoder-decoder network

Depth estimation



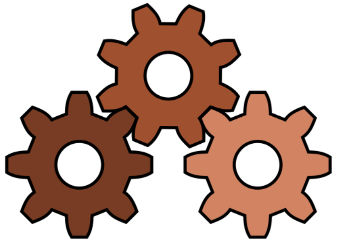
Multivariate regression problem (many outputs, continuous)

Convolutional encoder-decoder network

Some terms to remember

- Regression = continuous numbers as output
- Classification = discrete classes as output
- Two class (binary) and multiclass classification treated differently
- Multilabel = zero or more of x discrete classes
- Univariate = one output
- Multivariate = more than one output

Translation

Real world input	Model input	Model	Model output	Real world output
“Skill without imagination is craftsmanship and gives us many useful objects such as wickerwork picnic baskets. Imagination without skill gives us modern art.”	$\begin{bmatrix} 7800 \\ 9853 \\ 4520 \\ 4596 \\ 987 \\ \vdots \end{bmatrix}$	 Deep learning model	$\begin{bmatrix} 6003 \\ 3689 \\ 4432 \\ 6003 \\ 2149 \\ \vdots \end{bmatrix}$	“L’habileté sans l’imagination est de l’artisanat et nous donne de nombreux objets utiles tels que des paniers de pique-nique en osier. L’imagination sans habileté nous donne l’art moderne.”

Transcription

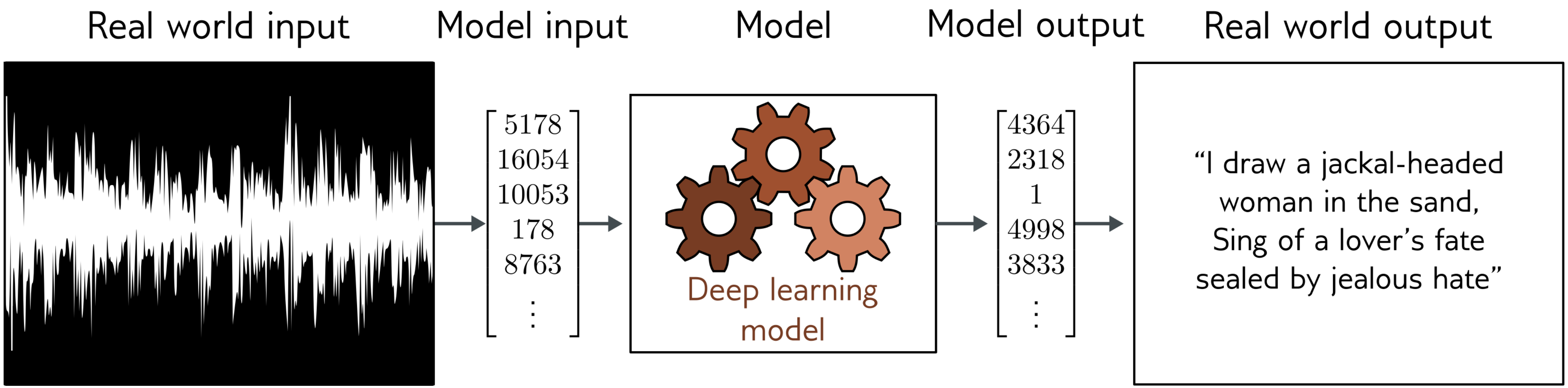
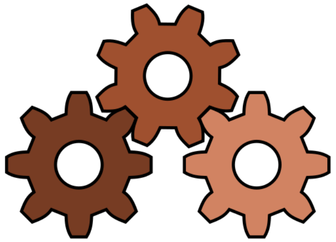


Image generation from text

Universität Konstanz

Real world input	Model input	Model	Model output	Real world output
"Teddy bears mixing sparkling chemicals as mad scientists, in a steampunk style."	$\begin{bmatrix} 8300 \\ 532 \\ 7676 \\ 7898 \\ 883 \\ \vdots \end{bmatrix}$	 Deep learning model	$\begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 1 \\ \vdots \end{bmatrix}$	

What do these examples have in common?

Universität Konstanz

Very complex relationship between input and output

Sometimes we may have many possible valid answers (think of translation for example)

But outputs (and sometimes inputs) obey rules

Can we learn the "grammar" of the data from unlabeled examples?

Can use an enormous amount of data to do this (as we don't need costly labels)

This has potential to make the supervised learning task earlier by having a lot of general knowledge of possible outputs (about grammatically correct sentences for example)

Unsupervised Learning

Learning about a dataset without labels

For example:

- Clustering
- Finding outliers
- Generating new examples
- **Filling in missing data**

In this course, we focus primarily on supervised approaches, but boundaries are sometimes a bit fuzzy as “self-supervision” shows

Self-supervised Learning

We can also create large amounts of "free" labeled data ourselves with two main approaches:

Generative self-supervised learning masks part of each data example and the task is to predict the masked part (this way we get a "label")

For example, take a corpus of unlabeled images, remove a part of each image and try to fill in ("inpaint") the missing part

Or we might take a large corpus of text (from the internet) and mask some words that we then try to predict

Or we might take cut off texts and try to predict the word that follows after the cut-off

Self-supervised Learning

Contrastive self-supervised learning uses pairs of examples that have a relationship and compares them to unrelated pairs.

With images, we could set up the task to decide if pairs of images are transformed versions of one another or if they are unconnected

Or, with text, we can determine if two sentences follow each other in the original document or not

We can also establish if two sentences are logically related

–> A lot of potential for creative approaches using and transforming *found* data

Landmarks in Deep Learning

Universität Konstanz

- 1958 Perceptron (Simple "neural" model)
- 1986 Backpropagation (Practical deep neural networks)
- 1989 Convolutional networks (Supervised learning)
- 2012 AlexNet Image classification (Supervised learning)
- 2014 Generative adversarial networks (Unsupervised learning)
- 2014 Deep Q-Learning - Atari games (Reinforcement learning)
- 2016 AlphaGo (Reinforcement learning)
- 2017 Machine translation (Supervised learning)
- 2019 Language models ((Un)supervised learning)
- 2022 Dall-E2 Image synthesis from text prompts ((Un)supervised learning)
- 2022 ChatGPT ((Un)supervised learning)
- 2023 GPT4 Multimodal model ((Un)supervised learning)

Applications of deep learning in the social sciences

Hugging Face | Search models, datasets, users... | Models | Datasets | Spaces | Posts | Docs | Solutions | Pricing

Tasks | Libraries | Datasets | Languages | Licenses | Other

Filter Tasks by name

Multimodal

Image-Text-to-Text | Visual Question Answering | Document Question Answering

Computer Vision

Depth Estimation | Image Classification | Object Detection | Image Segmentation | Text-to-Image | Image-to-Text | Image-to-Image | Image-to-Video | Unconditional Image Generation | Video Classification | Text-to-Video | Zero-Shot Image Classification | Mask Generation | Zero-Shot Object Detection | Text-to-3D | Image-to-3D | Image Feature Extraction

Natural Language Processing

Text Classification | Token Classification | Table Question Answering | Question Answering | Zero-Shot Classification | Translation | Summarization | Feature Extraction | Text Generation | Text2Text Generation | Fill-Mask | Sentence Similarity

Audio

Text-to-Speech | Text-to-Audio | Automatic Speech Recognition | Audio-to-Audio | Audio Classification | Voice Activity Detection

Tabular

Tabular Classification | Tabular Regression

Reinforcement Learning

Reinforcement Learning | Robotics

Other

Graph Machine Learning

Models 533,551 | Filter by name | new Full-text search | Sort: Trending

google/gemma-7b
Text Generation • Updated 5 days ago • 198k • 1.82k

ByteDance/SDXL-Lightning
Text-to-Image • Updated 1 day ago • 297k • 970

bigcode/starcoder2-15b
Text Generation • Updated 3 days ago • 4.15k • 296

playgroundai/playground-v2.5-1024px-aesthetic
Text-to-Image • Updated 5 days ago • 32.3k • 258

google/gemma-7b-it
Text Generation • Updated 10 days ago • 98.4k • 833

mistralai/Mixtral-8x7B-Instruct-v0.1
Text Generation • Updated 4 days ago • 1.04M • 3.15k

google/gemma-2b
Text Generation • Updated 11 days ago • 80.2k • 482

stabilityai/stable-cascade
Text-to-Image • Updated 14 days ago • 454k • 986

stabilityai/stable-video-diffusion-img2vid-xt
Image-to-Video • Updated 17 days ago • 182k • 1.91k

bigcode/starcoder2-7b
Text Generation • Updated 3 days ago • 1.89k • 86

meta-llama/Llama-2-7b-chat-hf
Text Generation • Updated Nov 13, 2023 • 1.18M • 2.92k

openai/whisper-large-v3
Automatic Speech Recognition • Updated 25 days ago • 884k • 1.83k

stabilityai/stable-diffusion-xl-base-1.0
Text-to-Image • Updated Oct 30, 2023 • 4.3M • 4.61k

meta-llama/Llama-2-7b
Text Generation • Updated Nov 13, 2023 • 3.56k

bigcode/starcoder2-3b
Text Generation • Updated 3 days ago • 2.57k • 68

mistralai/Mistral-7B-Instruct-v0.2
Text Generation • Updated 4 days ago • 927k • 1.02k

HuggingFaceH4/zephyr-7b-gemma-v0.1
Text Generation • Updated about 14 hours ago • 396 • 63

runwayml/stable-diffusion-v1-5
Text-to-Image • Updated Aug 23, 2023 • 3.96M • 10.4k

briaai/RMBG-1.4
Image-to-Image • Updated 20 days ago • 314 • 683

m-a-p/ChatMusician
Text Generation • Updated 2 days ago • 550 • 59

microsoft/phi-2
Text Generation • Updated 27 days ago • 432k • 2.88k

mistralai/Mistral-7B-v0.1
Text Generation • Updated Dec 11, 2023 • 1.17M • 2.88k

google/gemma-2b-it
Text Generation • Updated 4 days ago • 91.6k • 346

abacaj/phi-2-super
Text Generation • Updated 1 day ago • 543 • 50

stabilityai/sdxl-turbo
Text-to-Image • Updated Dec 7, 2023 • 559k • 1.79k

nerijs/pixelcascade128-v0.1
Text-to-Image • Updated 7 days ago • 225 • 47

togethercomputer/evo-1-131k-base
Text Generation • Updated 6 days ago • 3.48k • 46

h94/IP-Adapter-FaceID
Text-to-Image • Updated 28 days ago • 465k • 1.07k

BioMistral/BioMistral-7B
Text Generation • Updated 11 days ago • 7.48k • 286

yam-peleg/Experiment26-7B
Text Generation • Updated 5 days ago • 394 • 41

Previous 1 2 3 ... 17,786 Next

© Hugging Face | TOS | Privacy | About | Jobs | Models | Datasets | Spaces | Pricing | Docs

🤗 Hugging Face

The Hugging Face Model Hub can give you an idea about the vast number of possible application areas

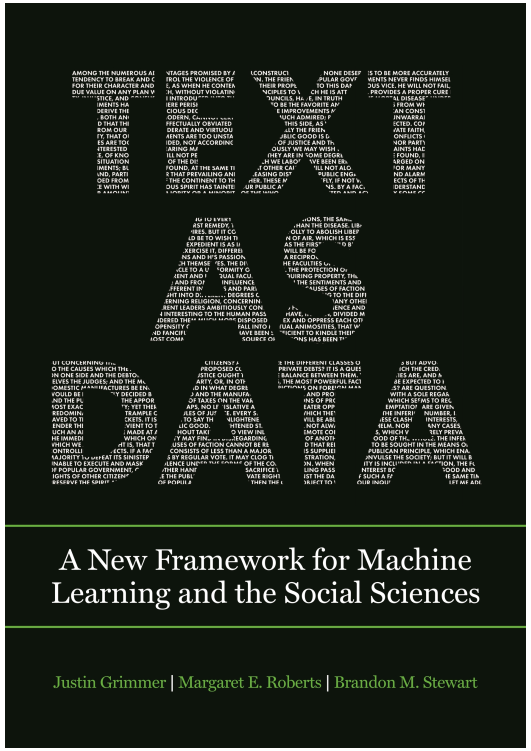
Also check out the Data Set Hub and Hugging Face Spaces

Spaces are often used for demos and to showcase interesting models and their applications

You can also rent dedicated hardware (billed by the minute, usually very cheap) to run spaces privately without queues

Examples from NLP

Endless research opportunities using “Text as Data”



Grimmer, J., Roberts, M. E., & Stewart, B. M. (2022). Text as data: A new framework for machine learning and the social sciences. Princeton University Press.

Text data can come from social media for example and be analysed for sentiment, emotions, arguments, stance, ...

Barbieri, F., Camacho-Collados, J., Neves, L., & Espinosa-Anke, L. (2020). TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification. arXiv:2010.12421 [Cs]. http://arxiv.org/abs/2010.12421

Synthetic Data

Universität Konstanz

"We propose and explore the possibility that language models can be studied as effective proxies for specific human subpopulations in social science research."

"Practical and research applications of artificial intelligence tools have sometimes been limited by problematic biases (such as racism or sexism), which are often treated as uniform properties of the models. We show that the "algorithmic bias" within one such tool—the GPT-3 language model—is instead both fine-grained and demographically correlated, meaning that proper conditioning will cause it to accurately emulate response distributions from a wide variety of human subgroups. We term this property algorithmic fidelity and explore its extent in GPT-3."

"We create 'silicon samples' by conditioning the model on thousands of sociodemographic backstories from real human participants in multiple large surveys conducted in the United States. We then compare the silicon and human samples to demonstrate that the information contained in GPT-3 goes far beyond surface similarity. It is nuanced, multifaceted, and reflects the complex interplay between ideas, attitudes, and sociocultural context that characterize human attitudes."

Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C., & Wingate, D. (2023). Out of One, Many: Using Language Models to Simulate Human Samples. Political Analysis, 31(3), 1–15.
https://doi.org/10.1017/pan.2023.2

Many other examples

Synthetic data: in-silico replication of experiments

In the vision domain: image classification, for example of satellite images to count attendance at events (cars) or migration flows

Whisper: Analysis of transcripts of videos (for example from Youtube) with NLP models

In the (near) future, more tools for analysing videos directly?

Promising advances in video generation for example

https://openai.com/research/whisper

https://github.com/openai/whisper

Deep learning is not always what you need

Letter | Published: 29 August 2018

Deep learning of aftershock patterns following large earthquakes

Phoebe M. R. DeVries, Fernanda Viégas, Martin Wattenberg & Brendan J. Meade

Nature **560**, 632–634 (2018) | Cite this article

Matters Arising to this article was published on 02 October 2019

DeVries, P. M. R., Viégas, F., Wattenberg, M., & Meade, B. J. (2018). Deep learning of aftershock patterns following large earthquakes. *Nature*, 560(7720), 632–634. https://doi.org/10.1038/s41586-018-0438-y

Matters Arising | Published: 02 October 2019

One neuron versus deep learning in aftershock prediction

Arnaud Mignan & Marco Broccardo

Nature **574**, E1–E3 (2019) | Cite this article

"[...] using two-parameter logistic regression (that is, one neuron) and obtain the same performance as that of the 13,451-parameter DNN."

"We further show that a logistic regression based on the measured distance and mainshock average slip (instead of derived stresses) performs better than the DNN."

Mignan, A., & Broccardo, M. (2019). One neuron versus deep learning in aftershock prediction. *Nature*, 574(7776), E1–E3. https://doi.org/10.1038/s41586-019-1582-8

Matters Arising | Published: 02 October 2019

Reply to: One neuron versus deep learning in aftershock prediction

Brendan J. Meade

Nature **574**, E4 (2019) | Cite this article

"Before commenting on the interesting philosophical issues raised by Mignan and Broccardo, I note that the authors were able to reproduce the results presented in our paper (available at https://github.com/phoebemrdevries/Learning-aftershock-location-patterns)."

"The perspective presented in our paper is that it was interesting to discover that a neural network learned a simple, non-exotic combination of stresses that provided considerably improved precision."

Hype vs. Underclaiming

Hype

The AI industry, as many other parts of the economy, depends heavily on the attention of all kinds of stakeholders to attract funding

Fanning the flames with musings about the close possibility of Artifical General Intelligence (AGI) is part of that game

We also see special emphasis on extreme dangers that are only very very remotely likely (if at all possible)

This can also, maybe counterintuitively, be seen as beneficial

Polemically: "Fund us because only we can protect you"

Inflated expecations can also backfire, see "AI Winter"

Underclaiming

Universität Konstanz

On the other hand, it's impossible to deny progress over the last years in areas such as NLP and also for other modalities than text such as image or video (analysis and generation)

Standardized **benchmark** test are indicators for the speed of progress (but still imperfect measures)

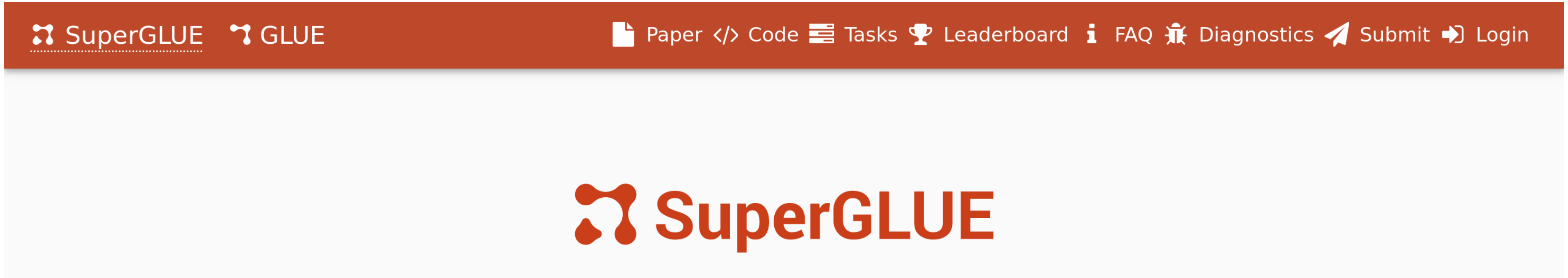
There is also a rewarding niche for experts that are by default downtalking every and each achievement

Often, these kind of experts have little to contribute besides that general criticism

There is a lot of questionable information around on "AI"

Benchmark Progress Example

https://gluebenchmark.com/



https://super.gluebenchmark.com/

Universität Konstanz

	Rank	Name	Model	URL	Score	BoolQ	CB	COPA	MultiRC	ReCoRD	RTE	WiC	WSC	AX-g	AX-b
	1	JDExplore d-team	Vega v2		91.3	90.5	98.6/99.2	99.4	88.2/62.4	94.4/93.9	96.0	77.4	98.6	100.0/50.0	-0.4
+	2	Liam Fedus	ST-MoE-32B		91.2	92.4	96.9/98.0	99.2	89.6/65.8	95.1/94.4	93.5	77.7	96.6	96.1/94.1	72.3
	3	Microsoft Alexander v-team	Turing NLR v5		90.9	92.0	95.9/97.6	98.2	88.4/63.0	96.4/95.9	94.1	77.1	97.3	93.3/95.5	67.8
	4	ERNIE Team - Baidu	ERNIE 3.0		90.6	91.0	98.6/99.2	97.4	88.6/63.2	94.7/94.2	92.6	77.4	97.3	92.7/94.7	68.6
	5	Yi Tay	PaLM 540B		90.4	91.9	94.4/96.0	99.0	88.7/63.6	94.2/93.3	94.1	77.4	95.9	95.5/90.4	72.9
+	6	Zirui Wang	T5 + UDG, Single Model (Google Brain)		90.4	91.4	95.8/97.6	98.0	88.3/63.0	94.2/93.5	93.0	77.9	96.6	92.7/91.9	69.1
+	7	DeBERTa Team - Microsoft	DeBERTa / TuringNLRv4		90.3	90.4	95.7/97.6	98.4	88.2/63.7	94.5/94.1	93.2	77.5	95.9	93.3/93.8	66.7
	8	SuperGLUE Human Baselines	SuperGLUE Human Baselines		89.8	89.0	95.8/98.9	100.0	81.8/51.9	91.7/91.3	93.6	80.0	100.0	99.3/99.7	76.6
+	9	T5 Team - Google	T5		89.3	91.2	93.9/96.8	94.8	88.1/63.3	94.1/93.4	92.5	76.9	93.8	92.7/91.9	65.6

The Dangers of Underclaiming: Reasons for Caution When Reporting How NLP Systems Fail

Samuel Bowman

Abstract

Researchers in NLP often frame and discuss research results in ways that serve to deemphasize the field's successes, often in response to the field's widespread hype. Though well-meaning, this has yielded many misleading or false claims about the limits of our best technology. This is a problem, and it may be more serious than it looks: It harms our credibility in ways that can make it harder to mitigate present-day harms, like those involving biased systems for content moderation or resume screening. It also limits our ability to prepare for the potentially enormous impacts of more distant future advances. This paper urges researchers to be careful about these claims and suggests some research directions and communication strategies that will make it easier to avoid or rebut them.

PDF

Cite

Search

Video

Universität Konstanz

Bowman, S. (2022). The Dangers of Underclaiming: Reasons for Caution When Reporting How NLP Systems Fail. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 7484–7499. https://doi.org/10.18653/v1/2022.acl-long.516