SPP 2298: Theoretical Foundations of Deep Learning

Gitta Kutyniok

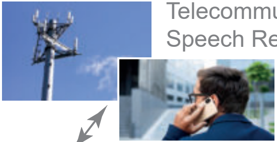
(Ludwig-Maximilians-Universität München)

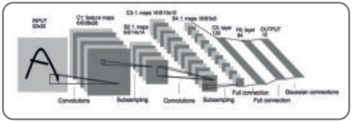
Virtual Information Meeting
October 18, 2023

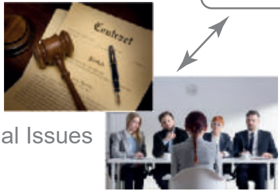
The Dawn of Artificial Intelligence in Public Life

Self-Driving Cars

Telecommunication/
Speech Recognition







Legal Issues

Health Care

Artificial Intelligence = Alchemy?

AAAS | Science

AI researchers allege that machine learning is alchemy

By Matthew Hutson | May. 3, 2018 , 11:15 AM

Ali Rahimi, a researcher in artificial intelligence (AI) at Google in San Francisco, California, took a swipe at his field last December—and received a 40-second ovation for it. Speaking at an AI conference, Rahimi charged that machine learning algorithms, in which computers learn through trial and error, have become a form of "alchemy." Researchers, he said, do not know why some algorithms work and others don't, nor do they have rigorous criteria for choosing one AI architecture over another. Now, in a paper presented on 30 April at the International Conference on Learning Representations in Vancouver, Canada, Rahimi and his collaborators document examples of what they see as the alchemy problem and offer prescriptions for bolstering AI's rigor.

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Problem with Reliability

Problems with Safety

Example:
Accidents involving robots

Problems with Security

Example:
Risks in self-driving cars

Problems with Privacy

Example:
Privacy violations of health data

Problems with Responsibility

Example:
Black-box and biased decisions

Problem with Reliability

Problems with Safety

Example:
Accidents involving robots

Problems with Security

Example:
Risks in self-driving cars

Problems with Privacy

Example:
Privacy violations of health data

Problems with Responsibility

Example:
Black-box and biased decisions

Current major problem worldwide:
Lack of reliability of AI technology!

Strong Requirements for Reliability

International Position on Reliable AI:

- AI Act of the European Union
- G7 Hiroshima AI Process

Strong Requirements for Reliability

International Position on Reliable AI:

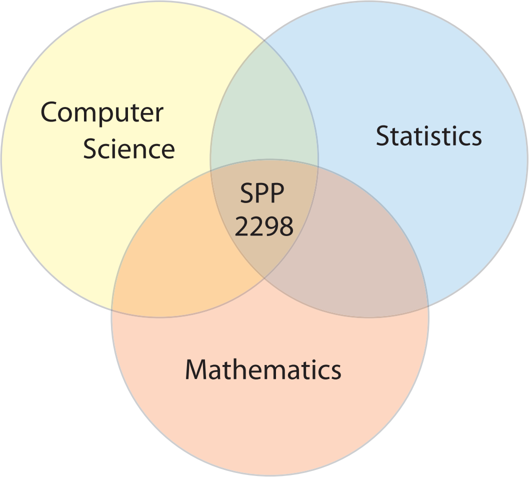
- AI Act of the European Union
- G7 Hiroshima AI Process

Major Challenge:

Derive a profound theoretical understanding!

SPP 2298: Theoretical Foundations of Deep Learning

Key Research Areas:



Key Parameters of SPP 2298

Important Dates:

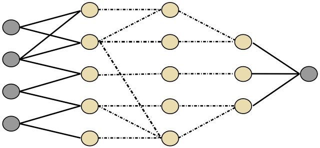
- December 1, 2023: Deadline for Submission
- Mai 2+3, 2024: Review
- Summer/Fall 2024: Start of Projects

Team:

- Martin Burger (DESY): Mathematics
- Matthias Hein (U Tübingen): Computer Science
- *Gitta Kutyniok (LMU Munich)*: Mathematics
- Sebastian Pokutta (ZIB): Mathematics
- Ingo Steinwart (U Stuttgart): Statistics

What are the Key Goals of this SPP?

Definition of a Deep Neural Network



Definition:
Assume the following notions:

- $d \in \mathbb{N}$: Dimension of input layer.
- L: Number of layers.
- $\rho : \mathbb{R} \to \mathbb{R}$: (Non-linear) function called *activation function*.
- $T_\ell : \mathbb{R}^{N_{\ell-1}} \to \mathbb{R}^{N_\ell}$, $\ell = 1, \dots, L$, where $T_\ell x = W^{(\ell)} x + b^{(\ell)}$

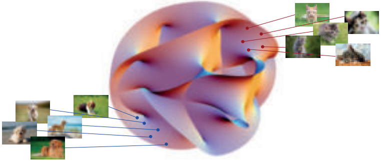
Then $\Phi : \mathbb{R}^d \to \mathbb{R}^{N_L}$ given by

$$\Phi(x) = T_L \rho(T_{L-1} \rho(\dots \rho(T_1(x)))), \quad x \in \mathbb{R}^d,$$

is called *(deep) neural network (DNN)*.

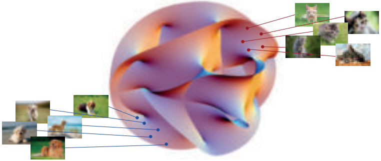
High-Level Set Up:

- Samples $(x_i, f(x_i))_{i=1}^m$ of a function such as $f : \mathcal{M} \to \{1, 2, \ldots, K\}$.
  $\rightsquigarrow$ *Training- and test data set.*

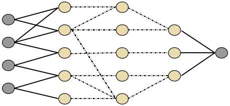


Training of Deep Neural Networks

High-Level Set Up:



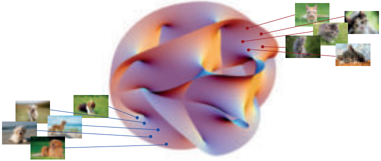
- Samples $(x_i, f(x_i))_{i=1}^m$ of a function such as $f : \mathcal{M} \to \{1, 2, \ldots, K\}$.
  $\rightsquigarrow$ *Training- and test data set.*



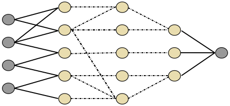
- Select an architecture of a deep neural network, i.e., a choice of d, L, $(N_\ell)_{\ell=1}^L$, and ρ.
  Sometimes selected entries of the matrices $(W^{(\ell)})_{\ell=1}^L$, i.e., weights, are set to zero at this point.

Training of Deep Neural Networks

High-Level Set Up:



- Samples $(x_i, f(x_i))_{i=1}^m$ of a function such as $f : \mathcal{M} \to \{1, 2, \ldots, K\}$.
  $\rightsquigarrow$ *Training- and test data set.*



- Select an architecture of a deep neural network, i.e., a choice of d, L, $(N_\ell)_{\ell=1}^L$, and ρ.
  Sometimes selected entries of the matrices $(W^{(\ell)})_{\ell=1}^L$, i.e., weights, are set to zero at this point.
- Learn the affine-linear functions $(T_\ell)_{\ell=1}^L = (W^{(\ell)} \cdot + b^{(\ell)})_{\ell=1}^L$ by

$$\min_{(W^{(\ell)}, b^{(\ell)})_\ell} \sum_{i=1}^m \mathcal{L}(\Phi_{(W^{(\ell)}, b^{(\ell)})_\ell}(x_i), f(x_i))$$

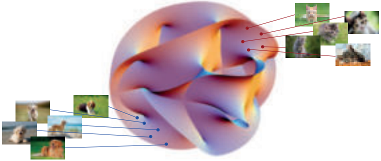
  yielding the network $\Phi_{(W^{(\ell)}, b^{(\ell)})_\ell} : \mathbb{R}^d \to \mathbb{R}^{N_L}$,

$$\Phi_{(W^{(\ell)}, b^{(\ell)})_\ell}(x) = T_L \rho(T_{L-1} \rho(\ldots \rho(T_1(x)))).$$

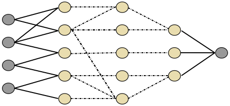
  This is often done by stochastic gradient descent.

Training of Deep Neural Networks

High-Level Set Up:



- Samples $(x_i, f(x_i))_{i=1}^m$ of a function such as $f : \mathcal{M} \to \{1, 2, \ldots, K\}$.
  $\rightsquigarrow$ *Training- and test data set.*



- Select an architecture of a deep neural network, i.e., a choice of d, L, $(N_\ell)_{\ell=1}^L$, and ρ.
  Sometimes selected entries of the matrices $(W^{(\ell)})_{\ell=1}^L$, i.e., weights, are set to zero at this point.
- Learn the affine-linear functions $(T_\ell)_{\ell=1}^L = (W^{(\ell)} \cdot + b^{(\ell)})_{\ell=1}^L$ by

$$\min_{(W^{(\ell)}, b^{(\ell)})_\ell} \sum_{i=1}^m \mathcal{L}(\Phi_{(W^{(\ell)}, b^{(\ell)})_\ell}(x_i), f(x_i))$$

  yielding the network $\Phi_{(W^{(\ell)}, b^{(\ell)})_\ell} : \mathbb{R}^d \to \mathbb{R}^{N_L}$,

$$\Phi_{(W^{(\ell)}, b^{(\ell)})_\ell}(x) = T_L \rho(T_{L-1} \rho(\ldots \rho(T_1(x)))).$$

  This is often done by stochastic gradient descent.

Goal: $\Phi_{(W^{(\ell)}, b^{(\ell)})_\ell}(x_i) \approx f(x_i)$ for the test data!

Main Research Directions, I

- **Expressivity:**
  - Which *aspects of a neural network architecture* affect the performance of deep learning?

$\rightsquigarrow$ *Applied Harmonic Analysis, Approximation Theory, ...*

Main Research Directions, I

- **Expressivity:**
  - Which *aspects of a neural network architecture* affect the performance of deep learning?

$\rightsquigarrow$ *Applied Harmonic Analysis, Approximation Theory, ...*

- **Learning:**
  - Why does *stochastic gradient descent* converge to good local minima despite the non-convexity of the problem?

$\rightsquigarrow$ *Algebraic/Differential Geometry, Optimal Control, Optimization, ...*

Main Research Directions, I

- **Expressivity:**
  - Which *aspects of a neural network architecture* affect the performance of deep learning?

  $\rightsquigarrow$ *Applied Harmonic Analysis, Approximation Theory, ...*
- **Learning:**
  - Why does *stochastic gradient descent* converge to good local minima despite the non-convexity of the problem?

  $\rightsquigarrow$ *Algebraic/Differential Geometry, Optimal Control, Optimization, ...*
- **Generalization:**
  - Can we derive overall *success guarantees* (on the test data set)?

  $\rightsquigarrow$ *Learning Theory, Probability Theory, Statistics, ...*

Main Research Directions, I

- **Expressivity:**
  - Which *aspects of a neural network architecture* affect the performance of deep learning?

  ⇝ *Applied Harmonic Analysis, Approximation Theory, ...*
- **Learning:**
  - Why does *stochastic gradient descent* converge to good local minima despite the non-convexity of the problem?

  ⇝ *Algebraic/Differential Geometry, Optimal Control, Optimization, ...*
- **Generalization:**
  - Can we derive overall *success guarantees* (on the test data set)?

  ⇝ *Learning Theory, Probability Theory, Statistics, ...*
- **Safety, Robustness, Interpretability, and Fairness:**
  - How can *adversarial attacks* be prevented?
  - How does a trained deep neural network *reach a certain decision*?
  - How can *fair decisions* be ensured?

  ⇝ *Information Theory, Uncertainty Quantification, ...*

Main Research Directions, II

- **Inverse Problems:**
  - How do we *optimally combine* AI-based with model-based approaches?
  - Is artificial intelligence capable of *replacing highly specialized numerical algorithms* in natural sciences?

$\rightsquigarrow$ *Imaging Science, Inverse Problems, Microlocal Analysis, ...*

Main Research Directions, II

- **Inverse Problems:**
  - How do we *optimally combine* AI-based with model-based approaches?
  - Is artificial intelligence capable of *replacing highly specialized numerical algorithms* in natural sciences?

$\leadsto$ *Imaging Science, Inverse Problems, Microlocal Analysis, ...*

- **Partial Differential Equations:**
  - Why do AI-based approaches perform well in *very high-dimensional environments*?

$\leadsto$ *Numerical Mathematics, Partial Differential Equations, ...*

Summary

Key Research Areas:

- *The statistical point of view:*
  - Regarding neural network training as a *statistical learning problem*.
  - Studying *expressivity*, *learning*, *optimization*, and *generalization*.

Summary

Key Research Areas:

- *The statistical point of view:*
  - Regarding neural network training as a *statistical learning problem*.
  - Studying *expressivity*, *learning*, *optimization*, and *generalization*.
- *The applications point of view:*
  - Focusing on *safety*, *robustness*, *interpretability*, and *fairness*.

Summary

Key Research Areas:

- *The statistical point of view:*
  - Regarding neural network training as a *statistical learning problem*.
  - Studying *expressivity*, *learning*, *optimization*, and *generalization*.
- *The applications point of view:*
  - Focusing on *safety*, *robustness*, *interpretability*, and *fairness*.
- *The mathematical methodologies point of view:*
  - developing and theoretically analyzing novel deep learning-based approaches to solve
    - *inverse problems* and
    - *partial differential equations*.

Summary

Five Key Interconnections:

- *Computational Efficiency.*
  - How to improve optimization, reduction of overparametrization,...?

Summary

Five Key Interconnections:

- *Computational Efficiency.*
  - How to improve optimization, reduction of overparametrization,...?
- *Deep Learning with Expert/Physical Knowledge.*
  - How to optimally combine physics with deep learning?

Summary

Five Key Interconnections:

- *Computational Efficiency.*
  - How to improve optimization, reduction of overparametrization,...?
- *Deep Learning with Expert/Physical Knowledge.*
  - How to optimally combine physics with deep learning?
- *Identification of Limitations of Deep Neural Networks.*
  - Critical assessment for which tasks deep learning is beneficial.

Summary

Five Key Interconnections:

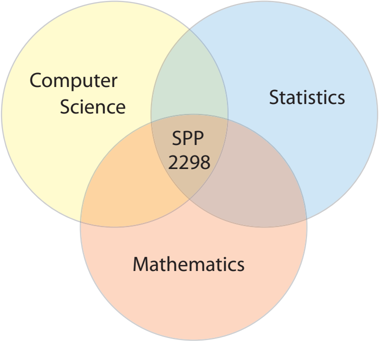
- *Computational Efficiency.*
  - How to improve optimization, reduction of overparametrization,...?
- *Deep Learning with Expert/Physical Knowledge.*
  - How to optimally combine physics with deep learning?
- *Identification of Limitations of Deep Neural Networks.*
  - Critical assessment for which tasks deep learning is beneficial.
- *Curse of Dimensionality.*
  - Under which conditions can the curse of dimensionality be overcome by deep neural networks?

Summary

Five Key Interconnections:

- *Computational Efficiency.*
  - How to improve optimization, reduction of overparametrization,...?
- *Deep Learning with Expert/Physical Knowledge.*
  - How to optimally combine physics with deep learning?
- *Identification of Limitations of Deep Neural Networks.*
  - Critical assessment for which tasks deep learning is beneficial.
- *Curse of Dimensionality.*
  - Under which conditions can the curse of dimensionality be overcome by deep neural networks?
- *Uncertainty Quantification.*
  - What is the uncertainty of outcome of a deep learning algorithm?

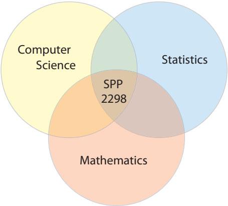
Interdisciplinarity



The research questions to be addressed within this Priority Programme are of a *truly interdisciplinary nature* and can only be solved by a *joint effort of computer science, mathematics, and statistics!*

Conclusions

SPP 2298: Theoretical Foundations of Deep Learning



Important Dates:

- December 1, 2023: Deadline for Submission
- Mai 2+3, 2024: Review
- Summer/Fall 2024: Start of Projects

Webpage:

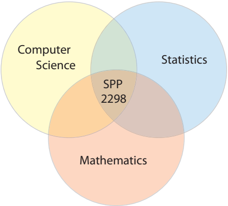
- `https://www.foundationsofdl.de`

Conclusions

SPP 2298: Theoretical Foundations of Deep Learning

Important Dates:

- December 1, 2023: Deadline for Submission
- Mai 2+3, 2024: Review
- Summer/Fall 2024: Start of Projects



Webpage:

- `https://www.foundationsofdl.de`

THANK YOU!