

# Hauptseminar

## Thema: Mitteln von Gewichten Neuronaler Netzwerke als Alternative zu Ensembles

Aus dem Forschungsbereich des Ensemble Learning ist bekannt, dass sich die Leistungsfähigkeit tiefer Neuronaler Netzwerke steigern lässt indem die Entscheidungen mehrerer Klassifikatoren gemittelt werden. Dies hat in der realen Anwendung jedoch den Nachteil, dass die Inferenzzeit mit der Anzahl der Klassifikatoren im Ensemble zunimmt. Aufgrund dieser Eigenschaft können in der Praxis oft nur kleine Ensembles eingesetzt werden. Eine Alternative zum Mitteln der Entscheidungen ist das Mitteln der Gewichte der einzelnen Klassifikatoren, was zu ähnlich guten Ergebnissen führen kann, aber nicht mit den Nachteilen einer gesteigerten Inferenzzeit einhergeht, weil weiterhin nur ein Neuronales Netzwerk angewendet werden muss. Das Mitteln von Gewichten mehrerer tiefer Neuronaler Netzwerke ist jedoch nur unter bestimmten Voraussetzungen möglich. Im Rahmen dieses Hauptseminars sollen die Grundvoraussetzungen aufgearbeitet und die Vorteile des Mitteln von Gewichten bezüglich verschiedener Aspekte (Leistungsfähigkeit, Generalisierung, usw.) aufgezeigt werden.

### Aufgabenstellung:

- Aufbereitung des State of the Art zum Thema ausgehend von der gegebenen Literatur [1-7]
- Aufbereitung der Grundvoraussetzungen für die Anwendbarkeit des Mitteln von Gewichten, sowie Beantwortung der Fragen warum es unter diesen Voraussetzungen funktioniert [4-7] und welche Vorteile es bringt (verschiedene Aspekte)
- Übersicht zu aktuellen Verfahren [1-3] mit Aufbereitung der Grundideen
- Abschätzung, wo sich diese Verfahren für die Robotik einsetzen lassen
- Vortrag im Rahmen des Hauptseminars

### Geeignet für:

Bachelor- / Masterstudiengänge

### Themengebiet / Schwerpunkte:

Deep Learning

### Erforderliche Vorkenntnisse:

Guter Abschluss der Vorlesung „Neuroinformatik und Maschinelles Lernen“ und Erfahrungen im Bereich Deep Learning oder erfolgreicher Abschluss der Vorlesung „Deep Learning for Computer Vision“

### Zu verwendende Literatur:

- [1] Izmailov et al.: [Averaging Weights Leads to Wider Optima and Better Generalization](#). UAI, 2018.
- [2] Wortsman et al.: [Model soups: averaging weights of multiple fine-tuned models improves accuracy without increasing inference time](#). ICML, 2022.
- [3] Ilharco et al.: [Patching open-vocabulary models by interpolating weights](#). arXiv, 2022.
- [4] Garipov et al.: [Loss Surfaces, Mode Connectivity, and Fast Ensembling of DNNs](#). NeurIPS, 2018.
- [5] Athiwaratkun et al.: [There Are Many Consistent Explanations of Unlabeled Data: Why You Should Average](#). ICLR, 2019.
- [6] Neyshabur et al.: [What is being transferred in transfer learning?](#) NeurIPS, 2020.
- [7] [Erläuterungen zu Stochastic Weight Averaging auf PyTorch-Webseite](#)

- Elektronische Literaturdatenbank des FG NI&KR mit Recherchemöglichkeiten
- Elektronische Konferenzproceedings-Datenbank des FG NI&KR
- IEEE Recherchesystem [www.ieeeexplore.ieee.org](http://www.ieeeexplore.ieee.org) (nur aus dem Uni-Netz bzw. via VPN)
- Google Scholar [scholar.google.com](http://scholar.google.com)
- Suche nach ähnlichen Publikationen [connectedpapers.com](http://connectedpapers.com), [arxiv-sanity-lite.com](http://arxiv-sanity-lite.com)
- Proceedings der relevanten Konferenzen (NeurIPS, ICML, ICLR, IJCNN, WCCI, ICANN, CVPR, ICCV, ECCV, BMVC, AVSS, ICPR, ICIP, ...)

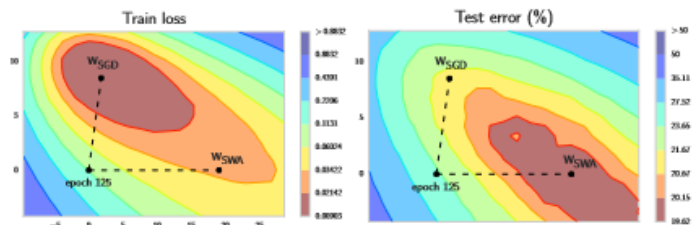
Betreuer:

Dr. Markus Eisenbach ([Markus.Eisenbach@tu-ilmenau.de](mailto:Markus.Eisenbach@tu-ilmenau.de))

Betr. Hochschullehrer: Prof. Dr. H.-M. Groß

Bearbeiter:

Omar Ashraf Mohamed Mohamed Hassan



Das Fehlergebirge für Training und Test ist oft leicht unterschiedlich. Neuronale Netzwerke mit gemittelten Gewichten treffen das Minimum auf den unbekanntem Testdaten oft besser.

Bildquelle: [1]