

# Deep Convolution Networks für Graphen

Forschungsprojekt Datenanalyse (Master AM)

Research Project (Master AP)

Verfahren des maschinellen Lernens konnten mittlerweile in einigen verschiedenen Anwendungsbereichen erfolgreich eingesetzt werden. Dazu gehören zum Beispiel, *Recommender* Systeme, Teilprobleme für das autonome Fahren und vieles mehr. Hier, kommen mittlerweile oft tiefe neuronale Netzwerke zum Einsatz. Insbesondere zur Bilderkennung und auch anderen Klassifizierungsproblemen konnten sogenannte Deep Convolutional Networks erfolgreich entwickelt und in Anwendung gebracht werden. Diese werden bisher jedoch meist nur auf Daten in Form von 2d Tensoren erfolgreich eingesetzt. Im Gegensatz dazu gibt es jedoch auch eine größere Klasse von Problemen, in der die Eingangsobjekte zunächst nur als graphenähnliche Strukturen vorliegen. Dazu zählen zum Beispiel prediktive Modelle für Mehrkörpersysteme, wie zum Beispiel die Vorhersage von Eigenschaften von Molekülen oder Materialien, wie sie innerhalb von Ansätzen des daten-getriebenen virtuellen Molekül- oder Materialdesigns zum Einsatz kommen.

**Ziel dieses Forschungsprojektes** ist es Deep Convolutional Network Verfahren für Graphen anzupassen und insbesondere für Problemstellungen aus der Physik und Chemie anzuwenden. Hierzu sollen spezifische Problem angepasste Netze mit Frameworks wie TensorFlow oder pyTorch praktisch umgesetzt und validiert werden. Des Weiteren soll in diesem Zusammenhang neben dem Grundkonzept und die wichtigsten Techniken des Deep Learnings auf weitere aktuelle Entwicklungen wie *Autoencoder* und *Generative Modelle* zum Beispiel zum Design von Molekülen oder Materialien eingegangen werden.

**Prüfungsleistung:** Präsentation des Projekts

**Inhaltliche Voraussetzungen:** Kenntnisse in Analysis, Linearer Algebra, Numerik und Programmierung im Umfang der Module der Bachelorstudiengänge. Vorkenntnisse in Python (oder zumindest Matlab) sind wünschenswert.

Der **Dozent**, Dr. Jan Hamaekers, Fraunhofer Institute for Algorithms and Scientific Computing SCAI, steht unter [jan.hamaekers@scai.fraunhofer.de](mailto:jan.hamaekers@scai.fraunhofer.de) für Rückfragen gerne zur Verfügung.

## Literatur:

Duvenaud, D. K., Maclaurin, D., Iparraguirre, J., Bombarell, R., Hirzel, T., Aspuru-Guzik, A., & Adams, R. P. (2015). Convolutional networks on graphs for learning molecular fingerprints. In *Advances in neural information processing systems* (pp. 2224-2232).

Gómez-Bombarelli, R., Wei, J. N., Duvenaud, D., Hernández-Lobato, J. M., Sánchez-Lengeling, B., Sheberla, D., ... & Aspuru-Guzik, A. (2018). Automatic chemical design using a data-driven continuous representation of molecules. *ACS central science*, 4(2), 268-276.

Barker, J., Bulin, J., Hamaekers, J., & Mathias, S. (2017). LC-GAP: Localized Coulomb descriptors for the Gaussian approximation potential. In *Scientific Computing and Algorithms in Industrial Simulations* (pp. 25-42). Springer, Cham.