



› RESEARCH DAY 2024

Towards Quantum Graph Neural Networks for Molecular Property Prediction

Präsentation Master Thesis | Tobias Roth / WV / Wirtschaftsinformatik | WiSe 23/24

AGENDA



Exposé / Einleitung

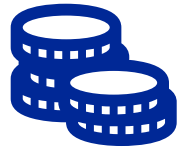
Paper 1

Paper 2

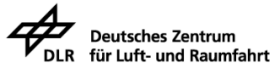
Konklusion

PROJEKTVORSTELLUNG: AQUAS

Gefördert vom
Bundesministerium für
Wirtschaft und Klimaschutz



Innovationen in der
Wasserstoffforschung durch
Quantensimulation



universität
uulm

Präzisere sowie effizientere Berechnung
und Simulation für spezifische Prozesse
und Materialien im Rahmen
wasserstoffbasierter Energie



Aufteilung von
Materialsimulationsprobleme in
Teilprobleme, welche mithilfe
klassischer und Quantencomputer
gelöst werden



<https://quantumsimulations.de/aquas>

https://www.digitale-technologien.de/DT/Navigation/DE/ProgrammeProjekte/AktuelleTechnologieprogramme/Quanten_Computing/Projekte/AQUAS/aquas.html

EINLEITUNG

PROBLEMSTELLUNG

- Bemühungen zur Reduzierung der Umweltbelastung geraten aufgrund des stetig voranschreitenden Klimawandels rücken in den Fokus der Öffentlichkeit (Amin et al. 2022)
- Suche nach effizienten Speicherlösungen für die Deckung des wachsenden globalen Energiebedarfs (Kilkis et al. 2019)
→ Umwandlung von Strom in Wasserstoff sowie Verbrennungsprozesse spielen eine wichtige Rolle (Amin et al. 2022)
- Erforschung neuer Materialien durch den Einsatz von Methoden im Bereich des maschinellen Lernens (Chen et al. 2023)
 - Bsp: Berechnung katalytischer Eigenschaften mithilfe **graphenbasierter Neuronaler Netze** (Tran et al. 2022)
- Prozess der Vorhersage von **Potentialflächen** und anderer relevanter Eigenschaften läuft auf molekularer bzw. atomarer Ebene (Tran et al. 2022)
- Erste Ansätze in der Literatur zur Realisierung zur Realisierung von GNNs auf Quantencomputer (Verdon et al. 2019, Beer et al. 2021, Ai et al. 2023)



Untersuchung, inwiefern Quantencomputer bzw. quanten-graphenbasierte Neuronale Netze für die Vorhersage katalytischer Eigenschaften im Rahmen der Wasserstoffproduktion genutzt werden können



Globale Forschungsfrage:

„Wie können klassische und quanten-graphenbasierte Neuronale Netze für die Vorhersage molekularer Eigenschaften genutzt werden?“

- Q1: Wie funktionieren GNNs und Quanten-GNNs?**
- Q2: Wie müssen molekulare Daten verarbeitet werden, damit sie mit einem Quantencomputer verwendet werden können?**
- Q3: Eignen sich Quanten-GNNs für die Vorhersage molekularer Eigenschaften im Kontext der Elektrokatalyse?**
- Q4: Wie ist die Leistung im Vergleich zu klassischen GNNs?**

EINLEITUNG

ZIELSETZUNG



Ziel: Implementierung eines klassischen GNNs und eines Quanten-GNNs

Auf Basis der vorhandenen Literatur einen Prototyp konzeptionieren, implementieren und evaluieren

Paper #1:

- Literaturanalyse zu GNNs, Darstellung des State-of-the-Arts
- Identifikation einer geeigneten Architektur
- Implementierung und Evaluierung eines klassischen GNNs

Paper #2:

- Analyse der vorhandenen Literatur (Foundation Paper) zu Quanten-GNNs
- Anforderungserhebungen für das zu implementierende Quanten-GNN
- Implementierung und Evaluierung eines Quanten-GNNs
- Einbindung und Veröffentlichung in das sQULearn-Python-Framework des Fraunhofer IPA



EINLEITUNG

METHODIK

Design-Science-Research nach Peffers et al. (2007)

Problemstellung & Motivation

Auseinandersetzung mit dem Thema

Zielsetzung definieren

Definition der Forschungsfragen und Artefakte in Kooperation mit dem Fraunhofer IPA

Design & Entwicklung

Literaturanalyse, Analyse der Foundation Paper → Implementierung der GNNs nach CRISP-DM (Wirth und Hipp 200)

Demonstration

Durchführung der Vorhersage von Potentialflächen mithilfe verschiedener Datensätze

Evaluation

Evaluierung, Vergleich der implementierten GNNs

Kommunikation

Mögliche Veröffentlichung der Paper und Einbindung des Codes in sQULearn Python-Bibliothek

› PAPER 1: ARCHITECTURES OF GRAPH NEURAL NETWORKS

2 Architectures of Graph Neural Networks

Abstract

As deep learning algorithms increasingly address graph-structured data, Graph Neural Networks (GNNs) play a more and more important role in various fields, including computational chemistry and the prediction of molecular properties. However, currently in the literature different architectures and approaches of GNNs exist. Therefore, this paper analyzes the current state-of-art by exploring architectures and approaches of GNNs in the literature. By creating an overview, it will show that convolutional GNNs are commonly used and particularly reliable when working with molecular properties. With the help of structured process CRISP-DM, later a convolutional GNN will be developed that predicts the potential energy surface of molecules with the QM9 dataset. The evaluation shows high accuracy, therefore the GNN was successfully developed.

Keywords: graph neural networks, architectures, literature review, potential energy surface prediction

2.1 Introduction

While deep learning algorithms effectively capture hidden patterns of Euclidean data, there is an increasing number of application areas where the data is represented in the form of graphs or structures similar to graphs [12]. Due to the expressive capabilities of graphs, researches on analyzing these kind of structures with machine learning have been receiving more and more attention. This is evident in the areas of social science, such as social networks, natural science like physical systems, knowledge graphs and other research areas [13]. Deep learning based methods that operate in a graph domains are called Graph Neural Networks (GNNs) [14].

In computational chemistry, molecules are modeled as graphs enabling various experiments [12]. This includes the prediction of potential energy surfaces for molecules made of hydrogen compounds or other materials [15]. In the current literature, different approaches or architectures for GNNs exist [12]. This raises the question of which approaches

of hydrogen compounds or other materials [15]. In the current literature, different approaches or architectures for GNNs exist [12]. This raises the question of which approaches

PAPER #1

PROBLEMSTELLUNG

- Daten werden vermehrt in Graphen oder graphenähnlichen Strukturen dargestellt (Wu et al. 2021)
- Mehr Aufmerksamkeit in der Forschung zur Analyse von Graphen mit maschinellem Lernen
 - z.B. Sozialwissenschaften (soziale Netzwerke), Naturwissenschaften (**physikalische Systeme**), etc. (Zhou et al. 2020)
- **GNNs**: Deep Learning Methoden, die Graphenstrukturen verarbeiten können (Velickovic 2023)
- Computergestützte Chemie: Moleküle werden als Graphen modelliert, verschiedene Experimente möglich (Wu et al. 2021)
 - Vorhersage von Potentialflächen und anderer molekularer Eigenschaften in Wasserstoffverbindungen oder anderer Materialien (Liu et al. 2023)



Literatur bietet verschiedene Ansätze bzw. Architekturen für GNNs – was eignet sich dabei besonders gut für die Vorhersage molekularer Eigenschaften?

“

Forschungsfrage:

„Wie können graphenbasierte Neuronale Netze für die Vorhersage von Moleküleigenschaften entwickelt werden?“

- Q1: Welche Ansätze existieren für GNNs in der Literatur?**
- Q2: Welche Architektur eignet sich am besten für die Vorhersage von Moleküleigenschaften?**

PAPER #1

ZIELSETZUNG



Ziel: Implementierung eines klassischen GNNs und eines Quanten-GNNs

Auf Basis der vorhandenen Literatur einen Prototyp konzeptionieren, implementieren und evaluieren

Paper #1:

- Literaturanalyse zu GNNs, Darstellung des State-of-the-Arts
- Identifikation einer geeigneten Architektur
- Implementierung und Evaluierung eines klassischen GNNs

Paper #2:

- Analyse der vorhandenen Literatur (Foundation Paper) zu Quanten-GNNs
- Anforderungserhebungen für das zu implementierende Quanten-GNN
- Implementierung und Evaluierung eines Quanten-GNNs
- Einbindung und Veröffentlichung in das sQULearn-Python-Framework des Fraunhofer IPA



PAPER #1

GRUNDLAGEN – GRAPHENBASIERTE NEURONALE NETZE

- Framework für das Lernen von Abhängigkeiten und Interaktionen zwischen Datenpunkten → klassische neuronale Netze würden daran scheitern (Velickovic 2023)
- Im Kontext von GNNs: Graph besteht aus Nodes (Knoten, Datenpunkte) und Edges (Kanten, Beziehung zwischen diesen Datenpunkten) (Wu et al. 2021)
- Gewichtete Summierung der Merkmale benachbarter Nodes bzw. Zusammenfassung der Informationen der benachbarten Nodes (Velickovic 2023, Liu et al. 2023)
- Schlüsselaspekt: Message-Passing Mechanismus (Gasteiger, Groß und Günnemann 2022):
 - Definiert, wie die Informationen zwischen den Nodes ausgetauscht werden
 - Schrittweise Erlernung der Node Representation über verschiedene Schichten des GNN
- **Ziel von GNNs:** Erlernen der Node Representations unter lokalen und globalen Grapheninformationen (Wu et al. 2021)
- Essentiell für Anwendungen, bei welchen **Beziehung zwischen Datenpunkten** eine wichtige Rolle spielen (Zhou et al. 2020)

PAPER #1

GRUNDLAGEN - POTENTIALFLÄCHEN

In computergestützter Chemie und Materialwissenschaft:

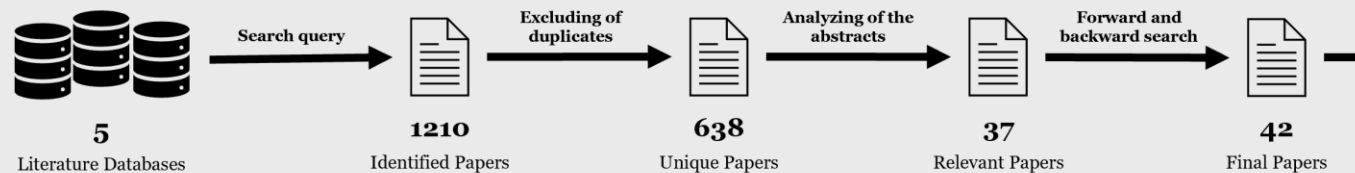
- Beschreibung der Energie eines Moleküls in Abhängigkeit verschiedener Parameter (Liu et al. 2023)
 - Hier: Zustandsenergie des Moleküls durch die Position der Atome im Raum
 - Vorhersage gibt etwa Aufschluss über die Stabilität und weiteren Eigenschaften von Molekülen und respektiven Materialien
- Zeigt, inwiefern Materialien genutzt werden können, um effiziente katalytische Prozesse im Bereich der Wasserstoffproduktion zu ermöglichen (Chen et al. 2023)

PAPER #1

FORSCHUNGSPROZESS

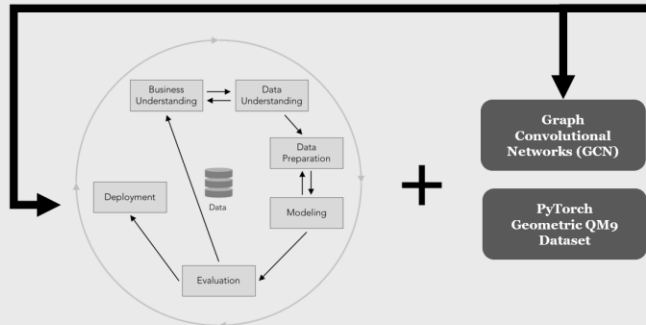
1 Literature Search

(Fettke 2006, vom Brocke et al. 2009, Webster and Watson 2002)



3 Prototypical Implementation of the GNN

(according to Crisp-DM, Wirth and Hipp 2000)

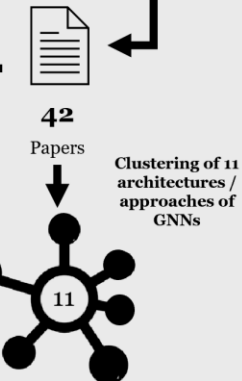


2 Literature Analysis

(Mayring 2010, Cruzes and Dyba 2011)

Architectures and approaches of Graph Neural Networks

Molecular property related (4)	In General (8)	
Graph Convolutional Networks (GCN)	Graph Convolutional Networks (GCN)	GraphSAGE (Sample and Aggregated)
Message Passing Neural Networks (MPNN)	Recurrent Graph Neural Networks (RecGNN)	Spatial-temporal Graph Neural Networks (STGNN)
DeepChem	Graph Attention Networks (GAT)	Graph Isomorphism Network (GIN)
AttentiveFP	Graph Symmetric Neural Network (GraphSNN)	Graph Autoencoder (GAE)



- Systematische Literaturanalyse in 5 Datenbanken: IEEE, ScienceDirect, SpringerLink, Emerald Insight, Google Scholar
- Suchterme: [("Graph Neural Networks") AND (("architectures") OR ("approaches"))] and [("Graph Neural Networks") AND ("*molecular*")]
- Clustering von 11 Ansätzen / Architekturen in zwei Kategorien
- Basierend auf den Ergebnissen Entwicklung eines GNNs nach CRISP-DM

PAPER #1

ERGEBNISSE - LITERATURANALYSE

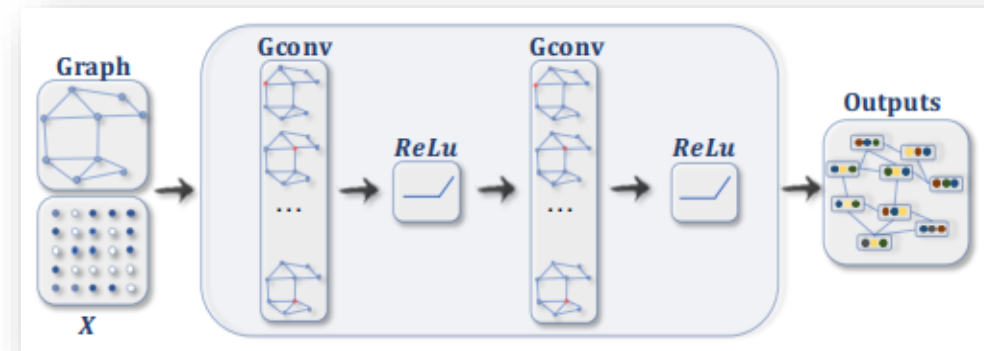
Architectures and approaches of Graph Neural Networks		
Molecular property related (4)	In General (8)	
Graph Convolutional Networks (GCN)	Graph Convolutional Networks (GCN)	GraphSAGE (Sample and Aggregated)
Message Passing Neural Networks (MPNN)	Recurrent Graph Neural Networks (RecGNN)	Spatial-temporal Graph Neural Networks (STGNN)
DeepChem	Graph Attention Networks (GAN)	Graph Isomorphism Network (GIN)
AttentiveFP	Graph Symmetric Neural Network (GraphSNN)	Graph Autoencoder (GAE)

- Literaturanalyse liefert zwei Kategorien: "Molecular property related" und "In General"
- Erste Kategorie zeigt Architekturen und Ansätze der computergestützten Chemie bzw. der Molekülforschung
- Zweite Kategorie zeigt Architekturen für andere Anwendungsfelder

PAPER #1

ERGEBNISSE – GRAPH CONVOLUTIONAL NETWORKS

- Graph Convolutional Neural Networks (GCNs) als eine der ersten und am meisten verwendeten Architekturen von GNNs (Kipf und Welling 2016)
- Kernaspekte (Zhang et al. 2019):
 - Aggregation von Informationen aus benachbarten Nodes in den Convolutional Layer
 - **Gewichtung** dieser Aggregationen und Hervorheben wichtiger Beziehungen → Iterative Informationsausbreitung durch den Graphen
- Gewichtungparameter werden während des Trainings erlernt und angepasst → Anpassung des GCNs an die Charakteristiken eines Datensatzes (Wu et al. 2021)



Beispielhafte GCN-Architektur
für die Node Classification (Wu et al. 2021)

PAPER #1

ERGEBNISSE – ENTWICKLUNG

- Entwicklung eines GCNs basierend auf den Ergebnissen der Literaturrecherche
- Strukturierter Entwicklungsprozess nach CRISP-DM
- QM9 Datensatz für Benchmarks in der computergestützten Chemie als Trainings- und Testdatensatz
- Ziel: Vorhersage der Potentialfläche von Molekülen des genutzten Datensatzes
- Business Understanding: Verständnis für das Ziel der Forschung und den Anforderungen → **siehe Einleitung / Exposé**



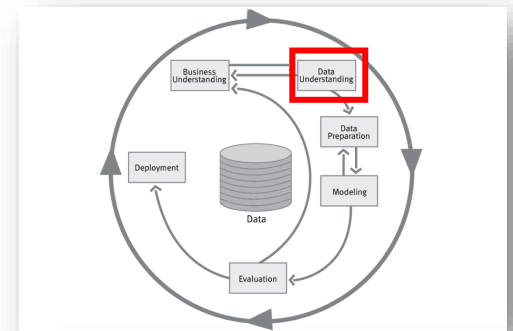
CRISP-DM Prozess nach Wirth und Hipp 2000

PAPER #1

ERGEBNISSE – ENTWICKLUNG: DATA UNDERSTANDING

Ziel: Schaffung des Verständnisses für den gegebenen Datensatz

- Nutzung des QM9 Datensatzes von PyTorch (PyTorch Geometric 2023)
- Beinhaltet ca. 130.000 Moleküle mit insgesamt 19 Regression Targets (Zielvariablen)
- Insgesamt fünf verschiedene Atomtypen, u.a. Wasserstoff, Sauerstoff und Kohlenstoff sowie Moleküle (z.B. Wasser, H₂O)
- Transformation der Moleküle zur Nutzung notwendig:
 - Atome bilden die Knoten des Graphen
 - Kanten repräsentieren die chemischen Verbindungen der Atome (Position im Raum)
- Zielvariable: Potentielle Energiefläche, interne Energie bei 0K



Target	Property	Description	Unit
0	μ	Dipole moment	D
1	α	Isotropic polarizability	a_0^3
2	ϵ_{HOMO}	Highest occupied molecular orbital energy	eV
3	ϵ_{LUMO}	Lowest unoccupied molecular orbital energy	eV
4	$\Delta\epsilon$	Gap between ϵ_{HOMO} and ϵ_{LUMO}	eV
5	$\langle R^2 \rangle$	Electronic spatial extent	a_0^2
6	ZPVE	Zero point vibrational energy	eV
7	U_0	Internal energy at 0K	eV
8	U	Internal energy at 298.15K	eV
9	H	Enthalpy at 298.15K	eV
10	G	Free energy at 298.15K	eV
11	c_v	Heat capacity at 298.15K	$\frac{\text{cal}}{\text{mol K}}$

Regression Targets des QM9
Datensatzes von PyTorch (Auszug)

PAPER #1

ERGEBNISSE – ENTWICKLUNG: DATA PREPARATION

Ziel: Vorbereitung der Daten zum Training des GCNs

- Extraktion der relevanten Informationen (Atome, xyz-Koordinaten der Atome, Zielvariable Interne Energie bei 0K)
- Normalisierung der Zielvariable mithilfe des Mittelwertes und der Standardabweichung
- Umwandlung der Koordinaten in die inversen paarweisen Abstände
 - Einbeziehung der räumlichen Beziehungen wie Translation (geradliniger Abstand) und Rotation (Drehung) zwischen den Atomen des Moleküls im Raum
- Aufteilen in Trainings- und Testdatensätze



```
# normalize labels using mean and standard deviation
def normalize_labels(labels):
    return (labels - train_labelm) / train_labelstd

# retransform predictions into original size
def transform_predictions(labels):
    return labels * train_labelstd + train_labelm
```

```
def x2e(x):
    """convert xyz coordinates to inverse pairwise distance"""
    r2 = np.sum((x - x[:, np.newaxis, :]) ** 2, axis=-1)
    e = np.where(r2 != 0, 1 / r2, 0.0)
    e = torch.from_numpy(e).double()
    return e
```

Entwicklungsauszüge

PAPER #1

ERGEBNISSE – ENTWICKLUNG: MODELING

Ziel: Umsetzung des GCN

- Implementierung der Architektur des GCN mithilfe des PyTorch Python Bibliothek
- GCN besteht aus zwei Convolutional Layer mit den zuvor präparierten Nodes, Edges und Features sowie den Gewichtungen als Inputparameter
- LeakyRelu als Aktivierungsfunktion (Erlernen komplexerer, nichtlinearer Zusammenhänge)



```
def gnn_layer(nodes, edges, features, we, wb, wv, wu):  
  
    leaky_relu = torch.nn.LeakyReLU(0.01)  
    new_nodes_tensor = torch.repeat_interleave(nodes[None, :], nodes.size()[0], dim=0)  
    input_tensor = wb + torch.matmul(new_nodes_tensor, we) * torch.transpose(edges[:,None],1,2)  
    ek = leaky_relu(input_tensor)  
  
    # sum over neighbors to get N x features  
    ebar = torch.sum(ek, 1)  
  
    # dense layer for new nodes to get N x features  
    new_nodes = leaky_relu(torch.matmul(ebar, wv) + nodes)  
  
    # sum over nodes to get shape features  
    global_node_features = torch.sum(new_nodes,0)  
  
    # dense layer for new features  
    new_features = leaky_relu(torch.matmul(global_node_features, wu)) + features  
  
    # just return features for ease of use  
    return new_nodes, edges, new_features
```

Entwicklungsauszug

PAPER #1

ERGEBNISSE – ENTWICKLUNG: EVALUATION

Ziel: Evaluation des entwickelten Modells

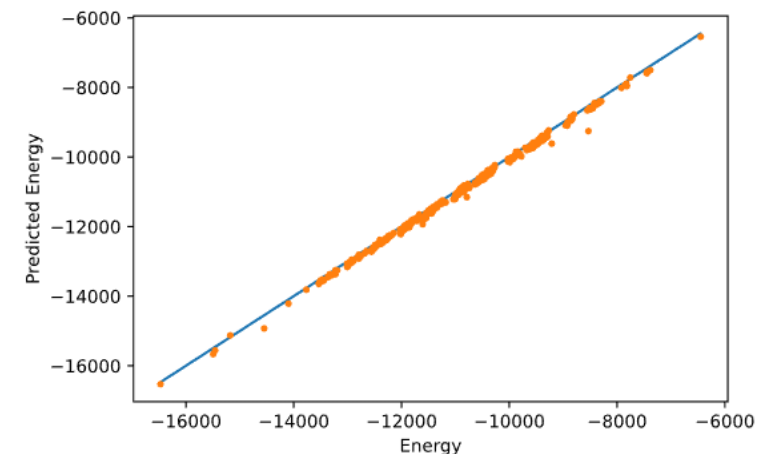
- Optimierung des Modells während des Trainings mithilfe des Adam Optimizers
 - Fähigkeit des Modells zur Generalisierung und genauen Vorhersagen auf anderen Daten neben den Trainingsdaten wird verbessert
- Bewertung der Genauigkeit des Modells mithilfe des Testdatensatzes
- Messung mithilfe der gängigen Regressionsmetriken



Performance Metric	Computed Value
Mean Squared Error (MSE)	97.5177
Root Mean Squared Error (RMSE)	9.3987
R^2 -Score (R^2)	0.9919

Tabelle der Regressionsmetriken des entwickelten Modells

Training time: 8 min 37.5s



Plot der tatsächlichen und vorhergesagten Energie

PAPER #1

KONKLUSION

Zusammenfassung	Kritische Reflektion	Ausblick
<ul style="list-style-type: none">• Insgesamt 11 verschiedene Architekturen bzw. Ansätze für GNNs im Rahmen der Literaturanalyse identifiziert• Aufbauend darauf Entwicklung eines GCNs → Beantwortung der Forschungsfrage• Genauigkeit der Vorhersage ermöglicht die Nutzung anderer Datensätze mithilfe des Modells, bspw. in Form von anderen Molekülen	<ul style="list-style-type: none">• Gewählte Forschungsmethodik hat sich als passend für die Zielsetzung herausgestellt• Weitere Ansätze bzw. Architekturen hätten untersucht bzw. implementiert werden können → möglicher Vergleich verschiedene Modelle durch die Regressionsmetriken• Keine Nutzung des Modells mit anderen Datensätzen, um Generalisierung zu testen	<ul style="list-style-type: none">• Erkenntnisse dieser Arbeit als Ausgangssituation für weitere Forschung in diesem Bereich• Vorhersage von Moleküleigenschaften wie die Potentialfläche verläuft auf molekularer bzw. atomarer Ebene → quantenchemische Eigenschaften• Schneller Fortschritt von Quantencomputer wirft Frage auf, ob diese für die Vorhersage von quantenchemischen Eigenschaften mittels ML-Methoden geeignet sind

› PAPER 2: QUANTUM GRAPH NEURAL NETWORKS FOR MOLECULAR PROPERTY PREDICTION

3 Quantum Graph Neural Networks for Molecular Property Prediction

Abstract

Hydrogen research, particularly in the fields of energy storage and catalysis, hinges on a deep understanding of molecular interactions and properties. The surface area of molecules is a crucial parameter influencing hydrogen adsorption and catalytic efficiency. Traditional computational methods for predicting molecular surface areas, while effective, face challenges in terms of computational complexity and accuracy when dealing with large and complex molecular structures. The accurate prediction of molecular surface area is paramount in optimizing hydrogen storage and catalysis processes. Conventional computational models, primarily based on classical algorithms, struggle with the vast complexity and quantum nature of molecular systems. This gap necessitates a novel approach that can handle the quantum mechanical properties inherent in molecules more naturally and efficiently.

Keywords: *quantum machine learning, quantum graph neural networks, hydrogen research, potential energy surface prediction*

3.1 Introduction

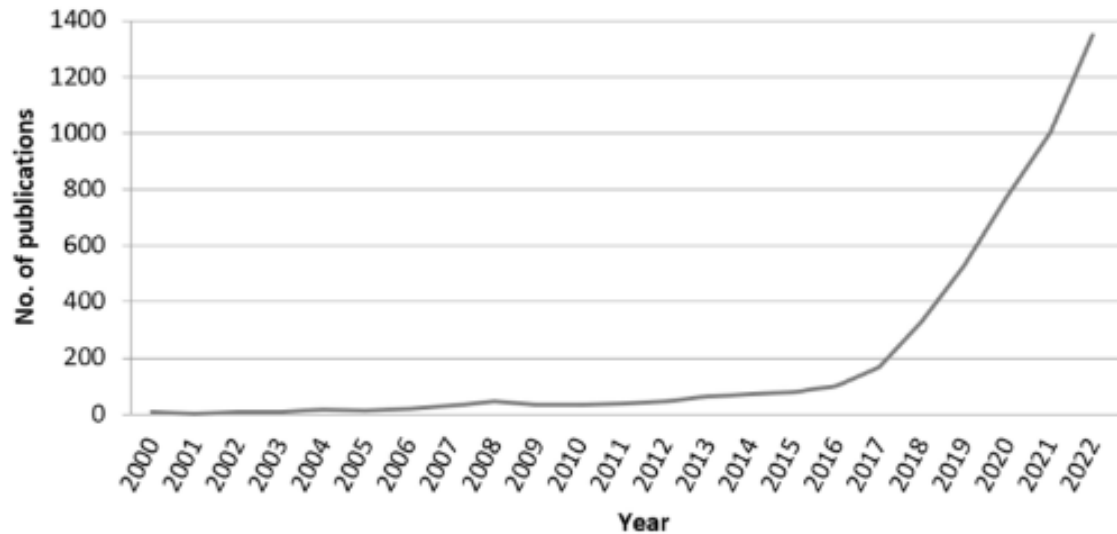
While there is growing research on the application of quantum computing in various domains, including chemistry and materials science, the specific application of QGCNNs for predicting molecular surface area in hydrogen research remains underexplored. This presents an opportunity to bridge the gap between quantum computing and molecular science, particularly in the context of hydrogen storage and catalysis. Recent advancements in quantum computing offer promising avenues for addressing complex problems in molecular sciences. Quantum Graph Convolutional Neural Networks (QGCNNs) have emerged as a potential tool for processing graph-structured data, which is naturally suited for molecular representations (Bronstein et al., 2017). The application of quantum computing in neural networks has shown significant improvements in handling high-dimensional data and complex interactions (Biamonte et al., 2017). Moreover, the representation of molecules as graphs aligns well with the capabilities of GNNs in capturing the intricate

molecules as graphs aligns well with the capabilities of GNNs in capturing the intricate

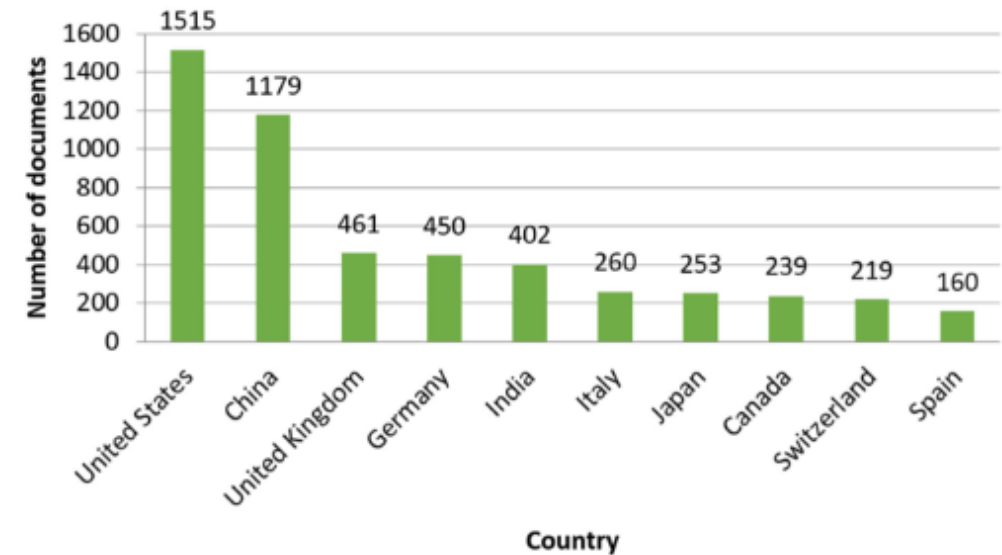
PAPER #2

PROBLEMSTELLUNG

Documents by year



Documents by the 10 first countries



Quelle: Quantum Machine Learning — An Overview (Tychola et al. 2023)



Forschungsfrage:

„Wie können quanten-graphenbasierte Neuronale Netze für die Vorhersage von Moleküleigenschaften entwickelt werden?“

- Q1: Wie müssen molekulare Daten verarbeitet werden, damit sie mit einem Quantencomputer verwendet werden können?**
- Q2: Eignen sich Quanten-GNNs für die Vorhersage molekularer Eigenschaften im Kontext der Elektrokatalyse?**
- Q3: Wie ist die Leistung im Vergleich zu klassischen GNNs?**

PAPER #2

ZIELSETZUNG



Ziel: Implementierung eines klassischen GNNs und eines Quanten-GNNs

Auf Basis der vorhandenen Literatur einen Prototyp konzeptionieren, implementieren und evaluieren

Paper #1:

- Literaturanalyse zu GNNs, Darstellung des State-of-the-Arts
- Identifikation einer geeigneten Architektur
- Implementierung und Evaluierung eines klassischen GNNs

Paper #2:

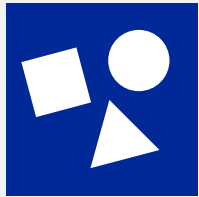
- Analyse der vorhandenen Literatur (Foundation Paper) zu Quanten-GNNs
- Anforderungserhebungen für das zu implementierende Quanten-GNN
- Implementierung und Evaluierung eines Quanten-GNNs
- Einbindung und Veröffentlichung in das sQULearn-Python-Framework des Fraunhofer IPA



PAPER #2

GRUNDLAGEN – QUANTUM MACHINE LEARNING

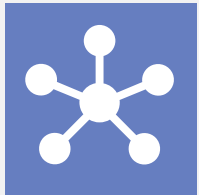
→ Kombination aus Quantencomputing und klassischem Machine Learning



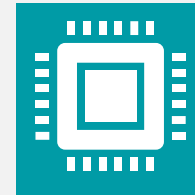
Quantum Bits, Gates und Circuits (Zheng et al. 2021)
Nutzung von Qbits (equiv. klassischer Bits), Quantum Gates (Operationen auf den Qubits) und Quantum Circuits (Sequenz von Qbits und Quantum Gates, Ausführung von Berechnungen)



Quantum Data Encoding (Schuld & Petruccione 2018)
Enkodierung klassischer Daten in einen Quantum State unter der Nutzung von Feature Maps
Effiziente Repräsentation von Daten durch die Nutzung des hochdimensionalen Quantum States



Entanglement und Superposition (Dunjko & Briegel 2018)
Essentielle Eigenschaften des Quantencomputings, effiziente Optimierung und Mustererkennung durch die Nutzung der quantenmechanischen Eigenschaften und Berechnungen auch in großen Rechenräumen



Quantum Processing (Biamonte et al. 2017)
Verarbeitung der enkodierten Daten mit Quantenalgorithmen wie bspw. Quantum Support Vector Machines oder Quantum Neural Networks



Ziel: Quantum Advantage, Überlegenheit von Quantenalgorithmen zu herkömmlichen Algorithmen in bestimmten ML-Aufgaben

PAPER #2

GRUNDLAGEN – QUANTUM GRAPH NEURAL NETWORKS

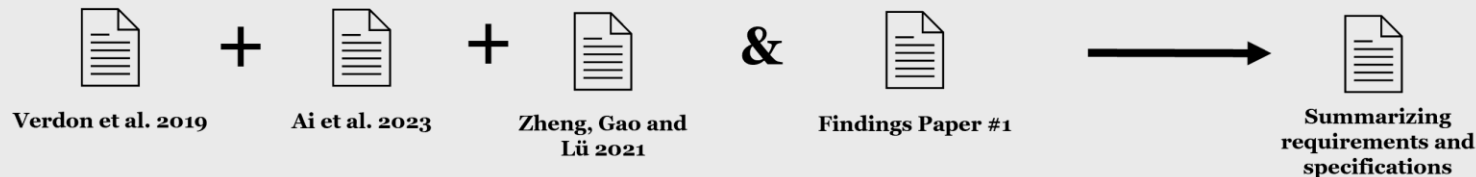
- Integration von GNNs und Prinzipien des Quantencomputings bzw. Quantum Machine Learning
- Relativ neue Forschungsrichtung: Bisher nur vereinzelte Veröffentlichungen zu QGNNs
- Encoding von klassischen Daten (Nodes, Edges) in einen Quantumstate (Zheng, Goa und Lü 2021):
 - Mapping von Features des Graphen (bspw. Koordinaten der Atome) durch Methoden wie Amplitude Encoding oder der Nutzung von Feature Maps
- Nutzung von Quantum Gates und Circuits (bzw. eines spezifischen **Ansatzes**) zum Erlernen der Strukturen des Graphen während des Trainings (Zheng, Goa und Lü 2021)
- Umsetzung klassischer GNN-Strukturen (siehe: GCN) als Quanten-Algorithmen (Convolutional, Pooling Layer) (Zheng, Goa und Lü 2021)
- Anwendungsspezifische Architektur: Anzahl der Qbits, Wahl des Datenenkodings, Wahl des Ansatzes, Wahl des Optimizers, etc. (Ai et al. 2023)

PAPER #2

FORSCHUNGSPROZESS

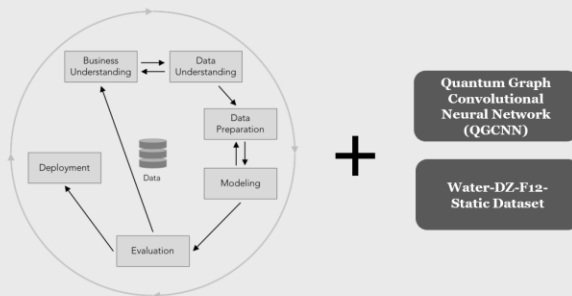
1 Foundation Paper Analysis

Phase: Design & Development (Peffer et al. 2007)



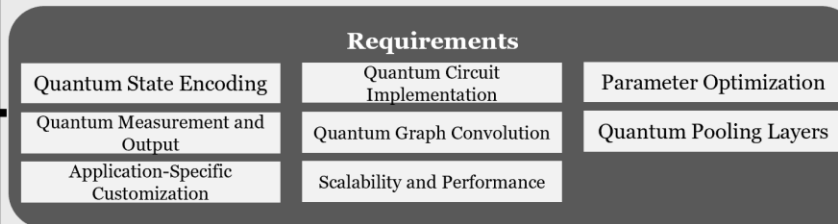
3 Prototypical Implementation of the QGCN

Phase: Design & Development | Demonstration | Evaluation (Peffer et al. 2007), according to CRISP-DM (Wirth and Hipp 200)



2 Requirements Analysis

Phase: Design & Development (Peffer et al. 2007)



- Methodik DSR nach Peffer et al. 2007
- Untersuchung von drei relevanten Foundation Paper
- Zusammenfassung der Erkenntnisse aus den Foundation Paper und aus Paper#1
- Identifikation von Anforderungen zur Entwicklung von QGNN
- Entwicklung des QGNN nach CRISP-DM basierend auf den identifizierten Anforderungen

PAPER #2

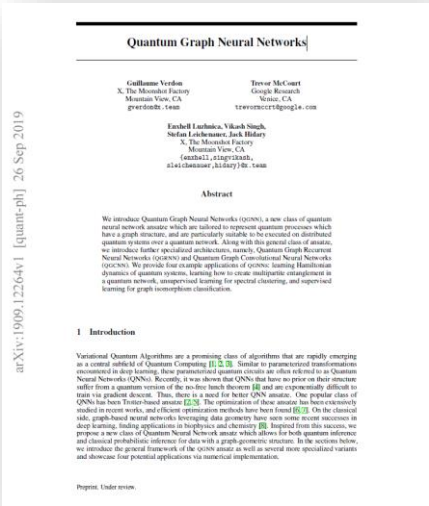
FOUNDATION PAPER

Paper

Quantum Graph Neural Networks

Verdon et al. 2019

- Erster Ansatz zur Realisierung von QGNNs in der Literatur
- Repräsentation / Berechnung von Graphenstrukturen mit Quantencomputing
- Vorstellung von QGCN- und QGRN-Architekturen, realisiert durch parametrisierte Quantum Circuits
- Demonstration anhand vier verschiedener Anwendungsfelder, darunter z.B. unüberwachtes Lernen für spektrale Clusterbildung oder Klassifikationsaufgaben



Paper

Towards Quantum Graph Neural Networks: An Ego-Graph Learning Approach

Ai et al. 2023

- Einführung eines hybriden quantum-klassischen ML-Algorithmus für Graphdaten
- Dekompositionsstrategie: große Graphen in kleinere aufteilen, um auch große Datenmengen verarbeiten zu können
- Architektur: Quantum Circuits mit eher weniger Qubits
- Algorithmus liefert gute bis sehr gute Performance für graphenbasierte Aufgaben
- Interessanter Ansatz für große Datensätzen mit vielen Dimensionen



PAPER #2

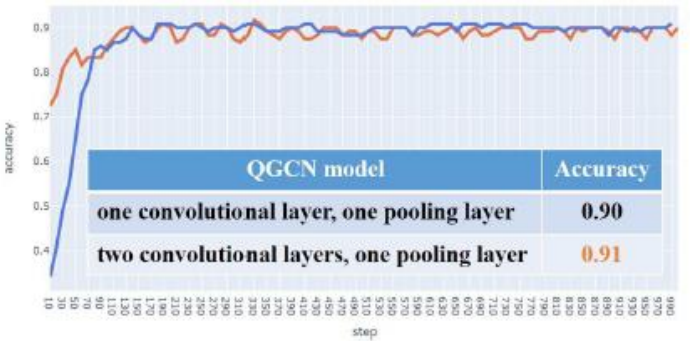
FOUNDATION PAPER

Paper

Quantum Graph Convolutional Neural Networks
Zheng, Gao und Lü 2021

Ziel: Entwicklung eines Quantum Graph Convolutional Neural Network (QGCN)

- Strukturierte Vorgehensweise zur Entwicklung von QGCNs
- Nutzung von Amplitude Encoding zur Umwandlung der Graphdaten in Quantum States
- Einsatz von parametrisierten Quantum Circuits zur Verarbeitung der Graphdaten
- Nutzung von Convolutional und Pooling Layer zum Training der Daten, optimiert für Graphdaten
- Test des Modells mithilfe des MNIST Datensatzes (Klassifikation)
 - Erfolgreiche Tests bzw. Vorhersagen mit Trainings- und Testdatensatz



Trainingsgenauigkeit mit dem MNIST Datensatz

PAPER #2

ANFORDERUNGSANALYSE – DATENVORBEREITUNG

Quantum State Encoding:

- Enkodierung der Graphdaten in Quantum States (Features (Koordinaten), Labels (Energie))
- Testing verschiedener Enkodierungsstrategien (Erhöhung der Genauigkeit bzw. Effizienz)

Quantum Circuit Implementation:

- Design und Implementierung von Quantum Circuits für die verschiedenen Layer des QGCNs
- Umsetzung von Graph Convolutional Operations mithilfe der Circuits, Berechnungen der Merkmale des Graphen
- Testing verschiedener Circuits zur Effizienz bzw. Skalierbarkeit des QGCNs

Parameter Optimization:

- Optimierung der Parameter der Quantum Circuits (bspw. Anzahl Parameter der Circuits \geq Anzahl der Features im Datensatz)
- Nutzung von Optimierungsalgorithmen (bspw. VQE) → Erhöhung der Genauigkeit des Modells

Quantum Measurement and Output:

- Einsatz von Messprotokollen (sog. Observables), um die verarbeiteten Daten aufzunehmen
- Messungsprozess muss dabei mit dem gewünschten Output (bspw. Vorhergesagte Energie) einhergehen

Requirements		
Quantum State Encoding	Quantum Circuit Implementation	Parameter Optimization
Quantum Measurement and Output	Quantum Graph Convolution	Quantum Pooling Layers
Application-Specific Customization	Scalability and Performance	

PAPER #2

ANFORDERUNGSANALYSE – ARCHITEKTUR GQCN

Quantum Graph Convolution:

- Umsetzung von Circuits, die den Convolutional Layern in klassischen GNNs ähneln
- effektive Extraktion und Berechnung der Weights der Features bzw. Edges und der Nodes

Quantum Pooling Layers:

- Einführung des Pooling-Mechanismus (Reduzierung der Graphengröße unter möglichst niedrigem bzw. keinem Informationsverlust)
- Abstraktion / Bewertung der Features hinsichtlich ihres Einflusses auf das Regressionsziel

Application-Specific Customization:

- Anpassung des QGCN an den speziellen Anwendungsbereich (bspw. Anzahl der Qbits, Struktur der Circuits, Parameter in den Circuits, Anzahl der Convolutional / Pooling Layer, etc.)
- Experimentelles Testen und Validieren notwendig, um eine optimale Struktur zu ermitteln

Scalability and Performance:

- Optimierung des Models für Recheneffizienz und Ressourcennutzung (letzter Schritt!)
- Testen von Datensätzen unterschiedlicher Größe
- Kontinuierliche Weiterentwicklung

Requirements		
Quantum State Encoding	Quantum Circuit Implementation	Parameter Optimization
Quantum Measurement and Output	Quantum Graph Convolution	Quantum Pooling Layers
Application-Specific Customization	Scalability and Performance	



Entwicklungseinblick

Quantum Graph Convolutional
Neural Networks



› KONKLUSION

KONKLUSION

ZUSAMMENFASSUNG

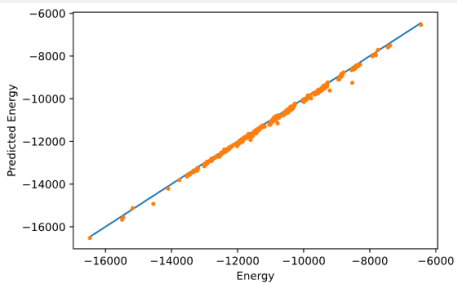
Architectures and approaches of Graph Neural Networks		
Molecular property related (4)	In General (8)	
Graph Convolutional Networks (GCN)	Graph Convolutional Networks (GCN)	GraphSAGE (Sample and Aggregated)
Message Passing Neural Networks (MPNN)	Recurrent Graph Neural Networks (RecGNN)	Spatial-temporal Graph Neural Networks (STGNN)
DeepChem	Graph Attention Networks (GAN)	Graph Isomorphism Network (GIN)
AttentiveFP	Graph Symmetric Neural Network (GraphSNN)	Graph Autoencoder (GAE)

Implementierung
Umsetzung eines GCNs basierend auf den Ergebnissen der Literaturanalyse und nach CRISP-DM

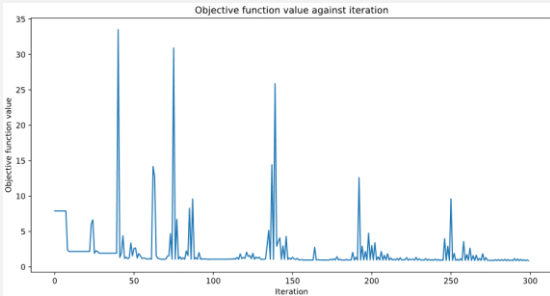
Requirements		
Quantum State Encoding	Quantum Circuit Implementation	Parameter Optimization
Quantum Measurement and Output	Quantum Graph Convolution	Quantum Pooling Layers
Application-Specific Customization	Scalability and Performance	

Implementierung
Umsetzung des QGCN mithilfe nach den Anforderungen und der Qiskit Python Bibliothek unter der Nutzung von CRISP-DM

Literaturanalyse
11 mögliche Architekturen für die Implementierung eines GNNs identifiziert



Anforderungsanalyse
Analyse dreier Foundation Paper zur Identifikation notwendiger Anforderung an das zu entwickelnde QGCN



KONKLUSION

KRITISCHE WÜRDIGUNG



Ausgewählte Forschungsmethodik hat sich als geeignet herausgestellt



Erweiterung der Suchterme und Literaturdatenbanken, zur Identifizierung weitere möglicher Ansätze und Architekturen für GNNs



Limitation auf eine ausgewählte Architektur nach der Literaturanalyse



Bisher noch kein Performance Vergleich der Modelle möglich

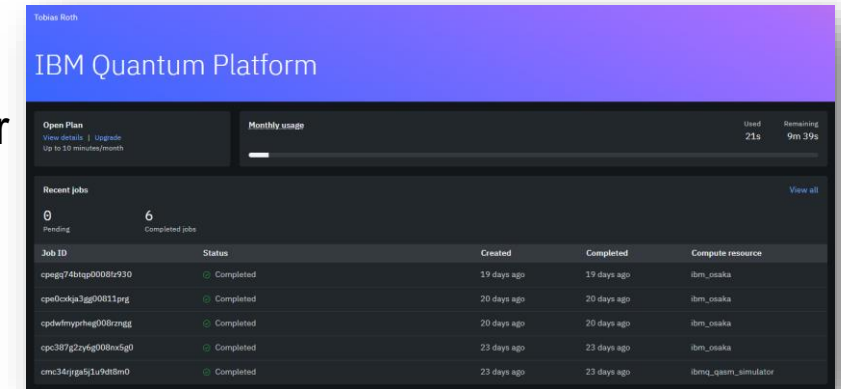
KONKLUSION

AUSBLICK

- Vollständige Umsetzung des QGNN
- Mögliches Testing der Ergebnisse auf einem echten Quantencomputer (IBM Quantum Experience)
- Vergleich der Performance der Modelle nach Beendigung der Entwicklung (Geschwindigkeit, Regressionsmetriken, etc.)

→ Beantwortung aller Forschungsfragen

- Fraunhofer **sQUlearn Python-Bilbiothek**: Vereinfachung der Nutzung und Implementierung verschiedener QML – Algorithmen
 - Ziel: Veröffentlichung der Ergebnisse in der sQUlearn Bilbiothek und als Paper
- Mögliche Nutzung / Durchführung von Experimenten mithilfe des entwickelten QGCN mit verschiedensten Datensätzen



<https://quantum.ibm.com/>



<https://sqlearn.github.io/index.html>

LITERATUR

- Ai, Xing, Zhihong Zhang, Luzhe Sun, Junchi Yan, and Edwin Hancock. 2023. "Decompositional Quantum Graph Neural Network." arXiv. <http://arxiv.org/abs/2201.05158>.
- Altae-Tran, Han, Bharath Ramsundar, Aneesh S. Pappu, and Vijay Pande. 2017. "Low Data Drug Discovery with One-Shot Learning." *ACS Central Science* 3 (4): 283–93. <https://doi.org/10.1021/acscentsci.6b00367>.
- Amin, Muhammad, Hamad Hussain Shah, Anaiz Gul Fareed, Wasim Ullah Khan, Eunhyea Chung, Adeel Zia, Zia Ur Rahman Farooqi, and Chaehyeon Lee. 2022. "Hydrogen Production through Renewable and Non-Renewable Energy Processes and Their Impact on Climate Change." *International Journal of Hydrogen Energy* 47 (77): 33112–34. <https://doi.org/10.1016/j.ijhydene.2022.07.172>.
- Beer, Kerstin, Megha Khosla, Julius Köhler, and Tobias J. Osborne. 2021. "Quantum Machine Learning of Graph-Structured Data." arXiv. <http://arxiv.org/abs/2103.10837>.
- Biamonte, Jacob, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. 2017. "Quantum Machine Learning." *Nature* 549 (7671): 195–202. <https://doi.org/10.1038/nature23474>.
- Brocke, Jan vom, Alexander Simons, Bjoern Niehaves, and Kai Reimer. n.d. "RECONSTRUCTING THE GIANT: ON THE IMPORTANCE OF RIGOUR IN DOCUMENTING THE LITERATURE SEARCH PROCESS."
- Chen, Zhijie, Sining Yun, Lan Wu, Jiaqi Zhang, Xingdong Shi, Wei Wei, Yiwen Liu, Renji Zheng, Ning Han, and Bing-Jie Ni. 2023. "Waste-Derived Catalysts for Water Electrolysis: Circular Economy-Driven Sustainable Green Hydrogen Energy." *Nano-Micro Letters* 15 (1): 4. <https://doi.org/10.1007/s40820-022-00974-7>.
- Cruzes, Daniela, and Tore Dybå. 2011. *Recommended Steps for Thematic Synthesis in Software Engineering. International Symposium on Empirical Software Engineering and Measurement*. <https://doi.org/10.1109/ESEM.2011.36>.
- "DeepChem." n.d. Accessed November 27, 2023. <https://deepchem.io/about/>.
- Dunjko, Vedran, and Hans J. Briegel. 2017. "Machine Learning & Artificial Intelligence in the Quantum Domain." arXiv. <http://arxiv.org/abs/1709.02779>.
- Fettke, Peter. 2006. "State-of-the-Art des State-of-the-Art: Eine Untersuchung der Forschungsmethode „Review“ innerhalb der Wirtschaftsinformatik." *WIRTSCHAFTSINFORMATIK* 48 (4): 257. <https://doi.org/10.1007/s11576-006-0057-3>.
- Gasteiger, Johannes, Janek Groß, and Stephan Günnemann. 2022. "Directional Message Passing for Molecular Graphs." arXiv. <http://arxiv.org/abs/2003.03123>.
- Gilmer, Justin, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. 2017. "Neural Message Passing for Quantum Chemistry." In *Proceedings of the 34th International Conference on Machine Learning*, 1263–72. PMLR. <https://proceedings.mlr.press/v70/gilmer17a.html>.
- Jiang, Dejun, Zhenxing Wu, Chang-Yu Hsieh, Guangyong Chen, Ben Liao, Zhe Wang, Chao Shen, Dongsheng Cao, Jian Wu, and Tingjun Hou. 2021. "Could Graph Neural Networks Learn Better Molecular Representation for Drug Discovery? A Comparison Study of Descriptor-Based and Graph-Based Models." *Journal of Cheminformatics* 13 (1): 12. <https://doi.org/10.1186/s13321-020-00479-8>.
- Kipf, Thomas N., and Max Welling. 2017. "Semi-Supervised Classification with Graph Convolutional Networks." arXiv. <http://arxiv.org/abs/1609.02907>.
- Kılış, Şiir, Goran Krajačić, Neven Duić, Luca Montorsi, Qiuwang Wang, Marc A. Rosen, and Moh'd Ahmad Al-Nimr. 2019. "Research Frontiers in Sustainable Development of Energy, Water and Environment Systems in a Time of Climate Crisis." *Energy Conversion and Management* 199 (November): 111938. <https://doi.org/10.1016/j.enconman.2019.111938>.
- Liu, Jin-Cheng, Hai Xiao, Xiao-Kun Zhao, Nan-Nan Zhang, Yuan Liu, Deng-Hui Xing, Xiaohu Yu, Han-Shi Hu, and Jun Li. 2023. "Computational Prediction of Graphdiyne-Supported Three-Atom Single-Cluster Catalysts." *CCS Chemistry* 5 (1): 152–63. <https://doi.org/10.31635/ccschem.022.202201796>.

LITERATUR

Mayring, Philipp. 2010 "Qualitative Content Analysis."

Peppers, Ken, Tuure Tuunanen, Marcus A. Rothenberger, and Samir Chatterjee. 2007. "A Design Science Research Methodology for Information Systems Research." *Journal of Management Information Systems* 24 (3): 45–77. <https://doi.org/10.2753/MIS0742-1222240302>.

"PyTorch." n.d. PyTorch. Accessed November 28, 2023. <https://pytorch.org/>.

Schuld, M., I. Sinayskiy, and F. Petruccione. 2015. "An Introduction to Quantum Machine Learning." *Contemporary Physics* 56 (2): 172–85. <https://doi.org/10.1080/00107514.2014.964942>.

"Torch_geometric.Datasets.QM9 — Pytorch_geometric Documentation." n.d. Accessed November 22, 2023. https://pytorch-geometric.readthedocs.io/en/latest/generated/torch_geometric.datasets.QM9.html.

Tran, Richard, Janice Lan, Muhammed Shuaibi, Brandon M. Wood, Siddharth Goyal, Abhishek Das, Javier Heras-Domingo, et al. 2023. "The Open Catalyst 2022 (OC22) Dataset and Challenges for Oxide Electrocatalysts." *ACS Catalysis* 13 (5): 3066–84. <https://doi.org/10.1021/acscatal.2c05426>.

Tychola, Kyriaki A., Theofanis Kalampokas, and George A. Papakostas. 2023. "Quantum Machine Learning—An Overview." *Electronics* 12 (11): 2379. <https://doi.org/10.3390/electronics12112379>.

Veličković, Petar. 2023. "Everything Is Connected: Graph Neural Networks." *Current Opinion in Structural Biology* 79 (April): 102538. <https://doi.org/10.1016/j.sbi.2023.102538>.

Verdon, Guillaume, Trevor McCourt, Enxhell Luzhnica, Vikash Singh, Stefan Leichenauer, and Jack Hiday. 2019a. "Quantum Graph Neural Networks." arXiv. <http://arxiv.org/abs/1909.12264>.

Webster, Jane, and Richard T. Watson. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review." *MIS Quarterly* 26 (2): xiii–xxiii.

Wirth, Rüdiger, and Jochen Hipp. n.d. "CRISP-DM: Towards a Standard Process Model for Data Mining."

"Wu et al. - 2021 - A Comprehensive Survey on Graph Neural Networks.Pdf." n.d. Accessed November 21, 2023. <https://ieeexplore.ieee.org/ielaam/5962385/9312808/9046288-aam.pdf>.

Wu, Zonghan, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. 2021. "A Comprehensive Survey on Graph Neural Networks." *IEEE Transactions on Neural Networks and Learning Systems* 32 (1): 4–24. <https://doi.org/10.1109/TNNLS.2020.2978386>.

Xiong, Zhaoping, Dingyan Wang, Xiaohong Liu, Zhong Feisheng, Xiaozhe Wan, Xutong Li, Zhaojun Li, et al. 2019. "Pushing the Boundaries of Molecular Representation for Drug Discovery with Graph Attention Mechanism." *Journal of Medicinal Chemistry* 63 (August). <https://doi.org/10.1021/acs.jmedchem.9b00959>.

Xu, Keyulu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. 2019. "How Powerful Are Graph Neural Networks?" arXiv. <http://arxiv.org/abs/1810.00826>.

Yan, Dafeng, Chalachew Mebrahtu, Shuangyin Wang, and Regina Palkovits. 2023. "Innovative Electrochemical Strategies for Hydrogen Production: From Electricity Input to Electricity Output." *Angewandte Chemie International Edition* 62 (16): e202214333. <https://doi.org/10.1002/anie.202214333>.

Zhang, Si, Hanghang Tong, Jiejun Xu, and Ross Maciejewski. 2019. "Graph Convolutional Networks: A Comprehensive Review." *Computational Social Networks* 6 (1): 11. <https://doi.org/10.1186/s40649-019-0069-y>.

Zheng, Jin, Qing Gao, and Yanxuan Lv. 2021. "Quantum Graph Convolutional Neural Networks." arXiv. <http://arxiv.org/abs/2107.03257>.

Zhou, Jie, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. 2020. "Graph Neural Networks: A Review of Methods and Applications." *AI Open* 1: 57–81. <https://doi.org/10.1016/j.aiopen.2021.01.001>.

VIELEN DANK!

Bei Fragen kontaktieren Sie bitte:



Tobias Roth

Fakultät WV | Wirtschaftsinformatik
toroth@hs-heilbronn.de