



Fragebogen der Fachschaft zu  
**mündlichen Prüfungen**  
im Informatikstudium

Dieser Fragebogen gibt den KommilitonInnen, die nach dir die Prüfung ablegen wollen, einen Einblick in Ablauf und Inhalt der Prüfung. Das erleichtert die Vorbereitung.

Bitte verwende zum Ausfüllen einen schwarzen Stift. Das erleichtert das Einscannen.

Dein Studiengang: Physik

**Prüfungsart:**

- Wahlpflichtfach
- Vertiefungsfach
- Ergänzungsfach

Welches ? .....

Barcode:



L

Prüfungsdatum: 07.09.2020

Prüfer/-in: Pascal Friederich

Beisitzer/-in: Matthias Schniewind

**Prüfungsfächer und Vorbereitung:**

Veranstaltung	Dozent/-in	Jahr	regelmäßig besucht?
Maschinelles Lernen für die Naturwissenschaften	Friederich	2020	Sofern man 2020 halt von besucht sprechen kann ja.

**Prüfungsablauf:**

Prüfungsdauer: 30 Minuten

Note: 1,0

War diese Note angemessen? Ja, lief größtenteils supi.

**Wie war der Prüfungsstil des Prüfers / der Prüferin?**

(Prüfungsatmosphäre, (un)klare Fragestellungen, Frage nach Einzelheiten oder eher größeren Zusammenhängen, kamen häufiger Zwischenfragen oder ließ er/sie Dich erzählen, wurde Dir weitergeholfen, wurde in Wissenslücken gebohrt?)

Prüfer hat viel öfters nachgefragt und unterbrochen, jedoch auch manchmal Sachen gesagt, die mir weitergeholfen haben. Fragen waren größtenteils recht standardmäßig Neuronale Netze mit dann jeweils den Anwendungen, die in der VL angesprochen wurden. Insgesamt sehr angenehm, jedoch kannte ich ihn auch schon wegen HiWi.

☞ Hat sich der Besuch / Nichtbesuch der Veranstaltung für dich gelohnt? Jop

☞ Wie lange und wie hast Du Dich alleine bzw. mit anderen vorbereitet? Alleine ca eine Woche, jedoch direkt nachdem ich Neuronale Netze geschrieben hatte und da entsprechend schon Ahnung hatte.

☞ Welche Tipps zur Vorbereitung kannst Du geben? (Wichtige / Unwichtige Teile des Stoffes, gute Bücher / Skripten, Lernstil)

Das Deep Learning Buch ist super, die Website towardsdatascience auch.

☞ Kannst Du ihm/sie weiterempfehlen?  Ja /  Nein Warum?

Nett, hab ne gute Note bekommen.

☞ Fanden vor der Prüfung Absprachen zu Form oder Inhalt statt? Wurden sie eingehalten? Nope.

☞ Kannst Du Ratschläge für das Verhalten in der Prüfung geben?

Wenn man grad keine Ahnung hat gut drauf hören was er sagt, er fasst dann Sachen nochmal ein bisschen zusammen um dich drauf zu bringen.

Inhalt:

S: Student; P: Prüfer

P: Was macht man denn mit ML so, nicht nur in den Naturwissenschaften?

S: Klassifizieren, Regression, Sampling, Übersetzen, Transkription.

P: Wie funktioniert denn ein CNN?

S: Filter, alles erklärt, sehr viel weniger Gewichte als Feed Forward NN

P: Also der Output ist auch wieder ein Bild.

S: Joa, hat halt auch Tiefe durch verschiedene Filter.

P: Wie ändert sich denn die Größe bei dem Bild?

S: Kann kleiner werden, gleich bleiben, größer werden, je nach Filtergröße, Stride und Padding.

Außerdem kann man es auch verkleinern durch Pooling.

P: Was für Anwendungsmöglichkeiten gibt es denn für ein CNN in den Naturwissenschaften?

S: Brustkrebserkennung und auch wenn man von einer DFT Rechnung die 3D Daten von der Elektronendichte von einem quantenmechanischen System weiß, kann man darauf auch ein CNN anwenden.

P: War das so ein ganz normales Standard CNN, das da verwendet wurde?

S: Es war ein Unet. Hab dann ein allgemeines Unet wieder mit nem Bild als Beispiel erklärt.

P: Diese Idee, dass man den Input wieder zum Output dazu macht und damit irgendwelche Transformationen überspringt, wo finden wir diese Idee noch?

S: ResNet, Residual Connections, hab ich erklärt wie es funktioniert.

P: Genug der CNNs, wie funktioniert denn ein Entscheidungsbaum.

S: Datensatz wird aufgesplittet mit vielen Entscheidungen hintereinander, etc.

P: Wie funktioniert denn diese Entscheidung genau, wo gesplittet wird?

S: Probiere alle möglichen Schnitte, finde Schnitte mit größter Reinheit nach irgendeinem Reinheitskriterium.

P: Wie viele verschiedene mögliche Schnitte gibt es denn?

S:  $k$  mal  $n$ , da  $k$  Feature Dimensionen und  $n$  Datenpunkte.

P: Und was ist die Komplexität von so einem gesamten Baum?

S:  $k n \log(n)^2$ .

P: Und woher kommt das?

S: Hab bisschen gestruggelt, mit etwas Hilfe:  $k$  mal  $n$  kommt aus dieser Entscheidung. Dann ein  $\log(n)$  kommt weil wir diese ganzen Schnitte noch nach dem besten Schnitt sortieren müssen, Sortieralgorithmen haben oft  $O(n \log(n))$ . Das zweite  $\log(n)$  kommt weil die Tiefe des Baumes im Schnitt  $\log(n)$  ist.

P: Coolio, dann machen wir mal Reinforcement Learning.

Innerer spitzer Aufschrei meinerseits.

P: Was gibt es denn da so für Grundbegriffe?

S: Agent in Umgebung, macht actions, befindet sich in states, handelt mit Policy. Nach kurzer Hilfe: Oh, und es gibt rewards für jede Aktion.

Danach ging es noch irgendwie um die Bellman Gleichung und die value function. Hatte allgemein nicht so viel Ahnung, aber hab versucht anhand dessen was der Prüfer manchmal eingeworfen hat mir schnell was Schlaues zu überlegen das ich sagen könnte. Ansonsten nicken und lächeln. Hab am Ende auf die Q Table übergeleitet und grob erklärt.

S: Q table ist eine Tabelle mit (Anzahl States) mal (Anzahl Actions) Einträgen. Die Q table enthält den Q Wert für jede Aktion in jeder Tabelle, der berechnet wird durch die rewards der nächsten Aktion und die erwarteten rewards in der Zukunft, gewichtet mit einem abfallenden Faktor.

P: Was gibt es denn für Anwendungen von Reinforcement Learning in den Naturwissenschaften? So mit Molekülen?

S: Gehe aus von einem Molekül, gib dem System bestimmte chemische Regeln (die Actions) wie es das Molekül verändern kann. Am Ende schaue welche Eigenschaften das finale Molekül hat (z.B. mit DFT) und gib ihm Reward entsprechend ob das gut oder nicht gut ist.

P: Und geht das so mit einer Q table?

Hab ich ja gesagt, weil ich dachte in der einen vorgestellten Publikation wurde das so gemacht. Hab mir schon gedacht, dass er am Ende darauf hinauswollte, dass man die Table durch ein NN ersetzt (Q Learning) aber er wollte mir nicht glauben, dass in der einen Publikation in der Vorlesung wirklich eine Table benutzt wurde. Bin mir grad auch nicht mehr sicher und zu faul nachzugucken. P hat dann noch gesagt, dass man dann halt keine Probleme mit der Größe der Tabelle hat aufgrund der hohen Zahl an möglichen States.

P: Wie kann man denn Moleküle repräsentieren?

S: SMILES String, einfach aber nicht gut weil keine natürliche Darstellung eines Moleküls. Fingerprint, entweder durch einfache binäre Kodierung einer chemischen Gruppe in einem Vektor oder durch das hashen, wodurch wir auch wieder einen Vektor bekommen. Hab mich kurz gefasst weil nur noch drei Minuten und ich wollte noch MPNNs erklären.

S: Und heutzutage benutzt man Message Passing Neural Networks. Hab diese dann auch erklärt. Der Sinn ist einfach der, von einer komischen Darstellung (dem Graphen des Moleküls) zu einer schönen Darstellung (dem Output Vektor des MPNNs) zu gelangen. Dieser Output Vektor repräsentiert das Molekül dann, und dann könnte man zum Beispiel noch ein fully connected layer dahinter schalten, das aus diesem Output Vektor dann irgendeine Klassifizierung macht oder so.