



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.

Barcode:



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

Dein Studiengang: Physik

L

- Prüfungsart:**  
 Wahlpflichtfach  
 Vertiefungsfach  
 Ergänzungsfach

Prüfungsdatum: 29.04.2021

Prüfer/-in: Pascal Friederich

Beisitzer/-in: Doktorand

Welches ? .....

**Prüfungsfächer und Vorbereitung:**

Veranstaltung	Dozent/-in	Jahr	regelmäßig besucht?
Maschinelles Lernen für die Naturwissenschaften	Pascal Friederich	SS20	Ja

**Prüfungsablauf:**

Prüfungsdauer: 30 Minuten

Note: 1,0

War diese Note angemessen?

**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?)

Sehr entspannte Atmosphäre, sehr wohlwollend. Prof. Friederich hat am Anfang erstmal 2 Minuten Smalltalk gemacht, um die Stimmung etwas aufzulockern.

Fragen waren verständlich, falls ich zunächst keine Antwort wusste, hat er mich zur Lösung geführt.

↳ Rückseite bitte nicht vergessen!

☞ Hat sich der Besuch / Nichtbesuch der Veranstaltung für dich gelohnt? ☞ Wie lange und wie hast Du Dich alleine bzw. mit anderen vorbereitet?

Vorlesung sollte man auf jeden Fall besuchen, da sehr aktuelle Anwendungsfälle und Beispiele besprochen werden. Diese sind auch prüfungsrelevant.

3 Wochen Vollzeit (alleine)

☞ Welche Tipps zur Vorbereitung kannst Du geben?

(Wichtige / Unwichtige Teile des Stoffes, gute Bücher / Skripten, Lernstil) Warum?

☞ Kannst Du ihn/sie weiterempfehlen?  Ja /  Nein

Formeln waren bei mir nicht relevant, es ging wirklich um die größeren Zusammenhänge. Fürs Verständnis unbedingt auch die besprochenen Paper anschauen (Introduction lesen und Schaubilder anschauen).

Ja, sehr faire, wohlwollende und entspannte Prüfung. Der Stoff ist sehr interessant und aktuell relevant - auch über Naturwissenschaften hinaus.

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

Keine

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

Wie immer: Nicht zu viel preisgeben, aber auch nicht zu wenig.

Inhalte der Prüfung: → Bitte auf die Rückseite und weitere Blätter!

- Schreibe bitte möglichst viele Fragen und Antworten auf.
- Wo wurde nach Herleitungen oder Beweisen gefragt oder anderweitig nachgehakt?
- Worauf wollte der Prüfer / die Prüferin hinaus?
- Welche Fragen gehörten nicht zum eigentlichen Stoff?

P: Prüfer, I: Ich

P: Für was für Aufgaben wird Maschinelles Lernen denn typischerweise, außerhalb der Naturwissenschaften, so eingesetzt?

I: Spracherkennung, Bildklassifikation, Übersetzung, Recommendation Systems

P: Wie kann man denn mit einem Neuronalen Netz übersetzen?

I: Sequence-To-Sequence-Modell, basierend auf RNNs. Encoder produziert context Vektor, Decoder erzeugt aus diesem wieder eine Sequenz.

P: Wie genau sieht denn ein solcher Encoder und Decoder aus?

I: Encoder: Aufbau mit hidden states grob erklärt. Hidden state zu  $t+1$  ergibt sich aus Input  $x$  zum Zeitpunkt  $t$  und hidden state zum Zeitpunkt  $t$ . Decoder ebenfalls grob beschrieben (Context Vektor ist Input in jedem Schritt).

P: D.h. um den nächsten hidden state zu erhalten, multipliziere ich eine Matrix mit  $x$  und eine mit  $h$  zum Zeitpunkt  $t$ . Können Sie sich vorstellen, wie man das in einer Operation zusammenfassen kann?

I: Bin nicht drauf gekommen, was er meinte. Er hat dann erklärt, dass man einfach einen gemeinsamen Vektor bestehend aus  $x$  und  $h$  (zum Zeitpunkt  $t$ ) untereinander basteln kann und dann eine große Matrix auf diesen draufmultipliziert.

P: Was ist beim naiven Sequence-To-Sequence Modell ein Problem?

I: Context Vektor hat feste Größe, nicht gut geeignet für Sequenzen mit beliebiger Länge.

P: Wie kann man das dann umgehen?

I: Attention Mechanismus erklärt. In jedem Zeitschritt enthält der Context Vektor bestimmte hidden states aus dem Enkodierschritt. Welche hidden states in welchem Zeitschritt zur Verfügung stehen, wird über Alignment Matrix bestimmt.

P: Wie sieht die Alignment Matrix beim Übersetzen typischerweise aus?

I: Hauptsächlich Diagonalmatrix, hängt aber von der Satzstellung in den beiden Sprachen ab.

P: Wie kann ich ein solches Modell jetzt für Moleküle verwenden?

I: Stelle Moleküle als SMILE Strings dar, Modell übersetzt Edukt in Produkte.

P: Wie funktioniert ein solcher SMILE String?

I: Erzeugung eines SMILE String grob erklärt.

P: Wie sieht hier jetzt die Alignment Matrix aus?

I: Vermutlich nicht mehr hauptsächlich Diagonalmatrix, da man ja an beliebiger Stelle im Molekül anfangen kann, den SMILE aufzubauen.

P: Wie funktioniert ein Entscheidungsbaum?

I: Eine Entscheidung auf jeder Ebene, am Ende Klassenzugehörigkeit oder Regressionswert.

P: Wie erzeugt man einen solchen Entscheidungsbaum?

I: Auf jeder Ebene den Schnitt auswählen, der die Unreinheit der child nodes minimiert (greedy Algorithmus).

P: Wie viele Schnitte muss ich beim Erzeugen eines Entscheidungsbaumes auf jeder Ebene prüfen?

I:  $k \cdot n$

P: Wie sieht die Zeit-Komplexität auf einer Ebene beim Bilden eines Entscheidungsbaumes aus?

I:  $n \cdot \log(n)$  fürs Sortieren, also insgesamt  $n \cdot \log(n) \cdot k$  für  $k$  Features

P: Was für eine Tiefe hat ein Entscheidungsbaum typischerweise?

I:  $\log(n)$ , also insgesamt Komplexität  $n \cdot k \cdot \log(n)^2$

P: Was für einen Nachteil hat ein einzelner Entscheidungsbaum?

I: Überfittet schnell, also hohe Varianz. Und hat Probleme mit korrelierten Daten.

P: Wie kann man das Überfitten umgehen?

I: Ensemble Methode, bspw. Random Forest. Habe dann Bagging und Random Split Selection erklärt.

P: Warum überfittet das jetzt nicht so stark?

I: Wenn die Bäume unabhängig sind, dann geht Varianz mit  $1/m$  runter.

P: Sind die Bäume eines Random Forest unabhängig?

I: Nicht komplett, da aus derselben Menge von Trainingsdaten mit Zurücklegen gezogen wird.

P: Wie funktioniert ein CNN?

I: Korrelation grob erklärt.

P: Wie kann ich das jetzt verwenden, um Bilder zu klassifizieren?

I: Convolution layer besteht aus Convolution, Pooling und Nichtlinearität. Davon verwendet man in der Regel mehrere hintereinander. Im Pooling layer wird die Größe der Channels typischerweise kleiner. Nach einigen solcher Convolution layer konkateniert man den Output zu einem Vektor. Diesen Vektor kann man noch durch ein normales MLP geben und erhält dann am Ende eine Klassenzugehörigkeit.

P: Wir haben in der Vorlesung auch über automatisierte Experimente gesprochen. Wie sind wir da vorgegangen?

I: Bayessche Optimierung mit Acquisition Function erklärt. Man nimmt das Maximum der Acquisition Function und macht dort ein neues Experiment. Wichtig war ihm hier, dass die in der VL definierte Acquisition Function (Erwartungswert von  $\max(0, f^-(x))$ ) nur eine von vielen möglichen Acquisition Functions ist. Welche man wählt,

hängt vom Problem ab.

P: Können Sie sich vorstellen, was bei der Bayesschen Optimierung ein Problem ist?

I: Nach etwas Hilfestellung: Ist iterativ, d.h. nicht direkt parallelisierbar.

P: Wie könnte man das dann trotzdem parallelisieren?

I: Nehme nicht nur das globale Maximum, sondern  $k$  Maxima der Acquisition Function. Diese kann man dann parallel im Experiment untersuchen und dann das Modell updaten.