



Fragebogen der Fachschaft zu mündlichen Prüfungen im Informatikstudium

Dein Studiengang: Informatik (MA)

Barcode:

Prüfungsart:

- Wahlpflichtfach
- Vertiefungsfach
- Ergänzungsfach

Welches? Neuronale Netze

Prüfungsdatum: 12.02.2016

Prüfer/-in: Dr. Stüker

Beisitzer/-in: ?

Prüfungsfächer und Vorbereitung:

Veranstaltung	Dozent/-in	Jahr	regelmäßig besucht?
Neuronale Netze	Prof. Waibel / Dr. Kilgour	SS 2015	Ja

Note: 1,0

Prüfungsdauer: 40 Minuten

War diese Note angemessen? Ja

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

Die Fragen waren größtenteils klar. Teilweise war es mir nicht klar und in der Prüfung hatte ich das Gefühl, dass dies schlecht ist. Aber an der Note und der Begründung später sieht man ja, dass es kein Problem war.

Es wurden einige Standard-Fragen gestellt und wenn nicht die Standardantworten gekommen sind (insbesondere wenn etwas gefehlt hat) wurde nachgebohrt. Aber nie besonders tief. Sehr viele Fragen in die Breite.

→Rückseite bitte nicht vergessen!

⌚ Hat sich der **Besuch / Nichtbesuch** der Veranstaltung für dich gelohnt?

Ja. Wenn man wenig Erfahrung mit Neuronalen Netzen hat werden in der Vorlesung die Zusammenhänge klarer. Allerdings kann ich die Übung („Praktikum“) leider nicht empfehlen. Man sollte sich die Aufgaben sowie die Folien der Besprechung; die Besprechung an sich war in meinem Jahr leider nicht so gut. Aber mit Prof. Waibel kann man gut reden; eventuell wird das in kommenden Jahren besser.

⌚ Wie lange und wie hast du dich **alleine bzw. mit anderen vorbereitet**?

- viel Vorwissen
- Vorlesung zu 100% besucht
- ca. 10 Treffen à 2 Stunden mit einem Lernpartner
- ca. 2 Wochen mit ca. 8 Stunden pro Tag

⌚ Welche **Tips zur Vorbereitung** kannst du geben?
Folien lesen und verstehen, Protokolle durchgehen und meinen Blog lesen:
martin-thoma.com/neuronale-netze-vorlesung

⌚ Kannst du ihm/sie **weiterempfehlen**? Ja / Nein
Warum?

Die Prüfung war sehr routiniert. Ich glaube wenn man den Stoff gut beherrscht kann man gut eine 1.0 bekommen. Der Prüfer ist fair, weiß aber auch sehr genau was er will und ist fordernd. In anderen Fächern kann man sicherlich deutlich leichter eine (sehr) gute Note bekommen.

⌚ Fanden vor der Prüfung **Absprachen** zu Form oder Inhalt statt?

Nein, es gab keine Absprachen.

⌚ Kannst du Ratschläge für das **Verhalten in der Prüfung** geben?

Nein, ich kann keine Tipps geben die nicht offensichtlich sind: Sagt was ihr wisst, lernt den Stoff vorher gut und schaut euch Protokolle an.

Inhalte der Prüfung:

- **Welche Anwendungen von Neuronalen Netzen kennen Sie?**

⇒ Klassifikation, Funktionsapproximation, Prädiktion, Collaborative Filtering, Dimensionalitätsreduktion, Denoising / Rekonstruktion.

- **Welche Netztypen haben wir in der Vorlesung kennen gelernt?**

⇒ McCulloch-Pitts Perzepron, Rosenblatt Perzepron, Multi-Layer Perzepron (MLP), (Denoising) Auto-Encoder, Restricted Boltzmann Machines (RBMs), Hopfield-Netze, CNNs / TDNNs, Rekurrente Netze (LSTMs). Ich habe noch „Gated Recurrent Units“ (GRU) erwähnt; das war aber nicht Teil der Vorlesung.

- Sie bekommen von ihrem Boss einen Datensatz, der bereits Vorverarbeitet / gefiltert wurde. Sie müssen ein Klassifikationsproblem mit 3 Klassen lösen. Was machen sie als erstes?

⇒ **Split in Trainings- und Testset.** Gegebenenfalls noch Validation-Set und Development-Set.

- **Wie sieht ein Perzepron aus?**

⇒ Aufgezeichnet und erklärt: Inputs x_1, \dots, x_n , Bias 1, gewichte w_0, \dots, w_n . Summe davon ergibt **net**, darauf wird Aktivierungsfunktion φ angewendet. Bei McCulloch-Pitts / Rosenblatt die Heavy-Side step function

- Was macht der Bias? Wo kann man den sehen?

⇒ Trennebene aufgezeichnet. Bias ist der Abstand zur 0 wenn man das Lot von der Trennebene auf die 0 fällt. (Hier habe ich mich verhaspelt.)

- **Wie sieht ein Multi-Layer Perzepron aus?**

⇒ Aufgezeichnet (3 Layer - Input, Hidden, Ouptut. Bias nicht vergessen.)

- **Welche Aktivierungsfunktionen kennen sie?** Zeichnen sie auf.

⇒ Signum (smooth: tanh), Heavy-Side step function (smooth: sigmoid), ReLU (smooth: softplus), Leaky ReLU, ELU. Softmax und Maxout kann man nicht so einfach Zeichnen, weil sie auf der Ausgabe einer ganzen Schicht arbeiten.

Softmax ist für Klassifikationsprobleme gut, da Softmax eine Wahrscheinlichkeitsverteilung (WKV) ausgibt, wenn man One-Hot Encoding der Labels hat (Jede Klasse ein Neuron; label ist ein Vektor der eine 1 bei der richtigen Klasse hat und sonst 0.) Cross-entropy Fehlerfunktion erwähnt.

- Softmax gibt nur unter bestimmten Annahmen eine WKV aus. Welche?

⇒ (Wusste ich nicht): Normalverteilte Daten, sigmoid-Aktivierungsfunktion sind nötig, damit man eine A posteriori Verteilung annehmen kann.

- Was sind Bottleneck-features?

⇒ MLP, bei dem man überwacht lernt und einen kleinen Layer zwischendrin hat. (Hier habe ich anscheinend etwas falsch verstanden. Laut Sebastian versteht man unter Bottleneck-Features immer überwachtes lernen. Unüberwacht kann man zwar auch Bottlenecks haben und man lernt auch Features, aber es wird halt nicht so genannt.)

- Was sind Denoising Auto-Encoder?

⇒ Man hat Daten X . Diese werden als Labels verwendet. Dem Netz gibt man $\tilde{X} = X + \text{Noise}$. Das Netz lernt also den Noise herauszufiltern. Als Beispiel finde ich schlecht empfangene Fernsehbilder gut (ist mir spontan eingefallen :)) Man weiß wie der Noise aussieht und kann den vorher auf typische Fernsehbilder tun. Dann könnte man im Fernseher etwas haben, was diesen Noise entfernt und das unverrauschte Bild rekonstruiert.

- Nun haben sie ein MLP und folgende Fehlerkurve (Aufgezeichnet - x-Achse: Epoche, y-Achse: Fehler; Trainingsfehler und Validierungsfehler). Der Trainingsfehler geht runter, aber der Testfehler bleibt gleich. Was ist der mögliche Grund?

⇒ Overfitting. Das Netzwerk ist zu groß / für das Netzwerk gibt es zu wenig Daten.

- Nun machen sie das Netzwerk kleiner und bekommen folgende Kurven (Zeichnung; An beiden Fehlern tut sich nicht mehr wirklich was)

⇒ Nun ist das Netzwerk zu klein; es kann nicht mehr lernen.

- Wie kann man nun sinnvoll eine Topologie aufbauen?

⇒ Es gibt 3 Gruppen von Verfahren, aber nur zwei haben wir in der Vorlesung besprochen:

- **Generative Verfahren** (Meiosis, Cascade Correlation),
- **Pruning-Verfahren** (Weight Decay, Optimal Brain Damage, Optimal Brain Surgeon, Regularisierung (L_1 , L_2 , Dropout)),
- **Genetische Algorithmen** (NEAT, HyerNEAT).

- Beschreiben sie Meiosis.

⇒ Aufgezeichnet (schaut euch mein YouTube Video an sowie die Antwort in der StackExchange Frage. Ist in meinem Blog-Artikel <https://martin-thoma.com/neuronale-netze-vorlesung> verlinkt.)

- Wie funktioniert Meiosis?

⇒ Wir lernen Mean und Variance der Gewichte. Wenn ein Kriterium erfüllt ist (vgl. Folien / Meiosis-Paper), dann splitten wir den Knoten mit Mean μ in zwei Knoten. Es gilt $\mu = \frac{\mu_1 + \mu_2}{2}$, aber $\mu \neq \mu_1 \neq \mu_2$.

- Was heißt Meiosis auf Deutsch?

⇒ Zellteilung.

Hinweis: Das ist ein Gedächtnisprotokoll. Es gab noch ein paar andere Fragen, aber dies waren die Wichtigsten. Meine Antworten waren ausführlicher, aber die wichtigsten Punkte wurden hier genannt. Mit Denoising Auto-Encoder / Bottleneck-Features habe ich mich etwas verrannt.