

Leitfaden Abschlussarbeiten - Professur Computerlinguistik

Das folgende Dokument versucht, Antworten auf oft gestellte Fragen zum Ablauf von Abschlussarbeiten an der Forschungsprofessur "Computerlinguistik" zu geben.

Stand: März 2024

Themenfindung & Themenvergabe

- Die Arbeit wird an der Forschungsprofessur "Computerlinguistik" durchgeführt, d.h. die Themen sollten etwas mit der automatischen Verarbeitung von Sprache zu tun haben, wobei "Sprache" sehr weit definiert ist (neben digitalen geschriebenen Texten sind auch Handschriften, Gebärdensprache, gesprochene Sprache, Text-Bild Kombinationen wie Memes, etc. gemeint)
- Einen Eindruck der Bandbreite (und Ausgangspunkt für mögliche Themen) verschaffen unsere Publikationen:
<https://www.fernuni-hagen.de/computerlinguistik/publikationen.shtml>
Wir gehen bei Kontaktaufnahme davon aus, dass Sie sich hier umgeschaut haben.
- Einen weiteren Eindruck vermittelt die Liste der abgeschlossenen Abschlussarbeiten
https://www.fernuni-hagen.de/forschung/schwerpunkte/catalpa/forschung/forschungsprofessuren/Forschungsprofessur_Computerlinguistik.shtml (Lehre)
- Bitte füllen Sie vor einer Kontaktaufnahme zur Themenfindung das Thesis Topic Form aus https://www.fernuni-hagen.de/computerlinguistik/docs/thesis_topic_formular_fuh.pdf
- Themenfindung
 - Gute Themen/Forschungsfragen sind schwer zu definieren. Wir beraten und unterstützen Sie bei der Themenfindung.
 - Wir erwarten aber, dass Sie sich bereits etwas eingelesen haben (s.o.)
 - Gute Themen zeichnen sich durch eine Forschungsfrage mit folgenden Eigenschaften aus:
 - als offene Frage formuliert:
 - Nicht nur Themenbereich, z.B. nicht "Large language models"
 - Keine Ja/Nein Fragen, z.B. nicht "Kann ein large language model eingesetzt werden um Task X zu bearbeiten?"
 - abgegrenzt
 - Zu generisch "Wie können large-language Models im Bereich X unterstützen?"
 - Zu spezifisch "Untersuchung der optimalen learning rate bei fine-tuning von open-instruct-stanford-alpaca-7b auf Daten der Buggy Software GbR"
 - spannend

- Sie finden es so spannend, dass Sie Monate damit verbringen wollen
- Antwort ist nicht bereits bekannt oder trivial zu erzeugen.

Anmeldung der Arbeit

- Zur Anmeldung der Arbeit füllen Sie bitte das folgende Formular aus (https://www.fernuni-hagen.de/mi/studium/pdf/anlage_aa.pdf) und schicken es per Mail an torsten.zesch@fernuni-hagen.de
- Titelvorschlag bitte von Ihnen und mit uns kurz abstimmen
- Um einen Zweitbetreuer kümmern wir uns, es sei denn, Sie möchten explizit jemanden auswählen oder ausschließen.
- Die Bearbeitungszeit beginnt mit der Anmeldung.
- Die Bearbeitungszeit richtet sich nach Ihrer Prüfungsordnung. Normalerweise 3 Monate für Bachelor, 6 Monate für Master.

Ausarbeitung

- Sprache
 - Arbeiten können in Englisch oder Deutsch verfasst werden. Wir empfehlen Englisch, da 99% der relevanten Literatur auch in Englisch verfasst ist und die Übersetzung des Fachvokabulars oft Probleme bereitet. Sie können die Arbeit aber auch ohne Abzüge bei der Bewertung auf Deutsch schreiben.
- Seitenzahl
 - Es gibt keine Vorgaben. Die Länge der Ausarbeitung richtet sich vollständig nach den Erfordernissen des Themas. Lieber kürzer schreiben und auf das Wesentliche konzentrieren. "Füllmaterial" hilft weder Ihnen noch uns. Wer bis hierhin gelesen hat: Bachelorarbeiten haben oft dreißig bis fünfzig Seiten und Masterarbeiten fünfzig bis siebzig.
- Template
 - Wir stellen unter <https://www.overleaf.com/read/twjqrqhcypwc> ein Overleaf Template für die Ausarbeitung zur Verfügung. Wir empfehlen LaTeX und Overleaf ist eine komfortable Möglichkeit, Ihnen ist aber vollkommen freigestellt, mit welcher Software Sie die Ausarbeitung schreiben. Das Endergebnis sollte aber in den wesentlichen Punkten (Zitierstil, Schriftgröße, etc.) dem Template entsprechen.
- Struktur
 - Die im Template zu findende Struktur passt höchstwahrscheinlich nicht auf Ihre spezifische Aufgabenstellung. Überlegen Sie genau, welche Inhalte die Arbeit enthalten soll. Stimmen Sie diese Outline (d.h. Überschriften und ein paar Stichpunkte über die Inhalte) mit der Betreuung ab.
- Zeitplanung
 - Fangen Sie möglichst früh damit an, die Ausarbeitung zu schreiben. Setzen Sie sich Schreibziele und überprüfen Sie deren Einhaltung.

- Verwendung von Hilfsmitteln
 - chatGPT
- Plagiate
 - Wenn Sie Inhalte aus anderen Veröffentlichungen/Webseiten/etc. übernehmen, müssen Sie das entsprechend angeben, normalerweise mittels einer passenden Referenz.

Code

- Der Code sollte lauffähig und dokumentiert sein
 - Es sollte auch jemand ohne Vorkenntnisse des Themas in der Lage sein, die grundlegenden Ergebnisse der Arbeit zu reproduzieren.

Vorgehen während der Bearbeitungszeit

- Eigenständigkeit
 - Arbeiten Sie möglichst eigenständig, aber scheuen Sie sich nicht, Fragen zu stellen. Schlechte Frage: "Was soll ich als nächstes machen?" Gute Frage(n): "Ich habe mir das angeschaut und denke ich sollte als nächstes X tun. Allerdings verstehe ich hier nicht, wie ich Problem X lösen soll. In der Literatur wird es als ungelöstes Problem beschrieben. Ich denke ich sollte stattdessen Y tun, bin mir aber unsicher, ob es dann das Thema der Arbeit sprengt."
- Literatur
 - Das Finden relevanter Literatur ist ein wichtiger Teil Ihrer eigenständigen wissenschaftlichen Leistung. Neben einer reinen Internetsuche gibt es auch Portale wie Google Scholar (<https://scholar.google.com/>) oder Semantic Scholar (<https://www.semanticscholar.org/me/research>). Für computerlinguistische / sprachtechnologische Themen auch (<https://aclanthology.org/>).
 - Sobald Sie wichtige Referenzen identifiziert haben, empfiehlt sich auch die Vorwärtssuche (welche neueren Publikationen referenzieren die gefundenen Kernpublikationen), sowie die Rückwärtssuche (welche wichtigen Papiere werden von bereits gefundenen Papieren zitiert).
- Spiralprinzip
 - Arbeiten Sie möglichst nach dem Spiralprinzip, d.h. einmal das Gesamtproblem mit möglichst einfachen Lösungen abarbeiten und dann verfeinern. Beispiel: auf einem reduzierten Datensatz (der dann schneller verarbeitbar ist) eine zufällige Baseline implementieren und mit der gewählten Metrik evaluieren. Dabei überprüfen, ob die Ergebnisse plausibel sind. Dann verfeinern, indem komplexere Methoden implementiert werden. Parallel an der Ausarbeitung schreiben.

Betreuung während der Bearbeitungszeit

- Prinzip
 - es ist Ihre Arbeit, wir beraten nur
- Status Updates
 - Üblicherweise vereinbaren wir regelmäßige Treffen
 - Bachelorarbeiten jede Woche
 - Masterarbeiten alle zwei Wochen
 - In den Treffen gilt: bereiten Sie eine Agenda vor und stellen Sie Ihre wichtigsten Fragen
- Folgendes Feedback bieten wir in jedem Fall an
 - Machbarkeit des Forschungsplanes prüfen
 - Initiale Outline der Ausarbeitung kommentieren
 - Probekapitel lesen
 - Feedback zu initialem Draft der Kolloquiumsfolien

Kolloquium

- Zeitpunkt
 - Das Kolloquium findet in der Regel nach Abgabe der Arbeit statt, damit über die finalen Ergebnisse berichtet werden kann, d.h. üblicherweise nach Ende der Bearbeitungszeit. Es kann jedoch auf Wunsch auch vorgezogen werden.
- Form
 - 15min Vortrag + 15min Diskussion
 - Die Zeit ist bewusst knapp. Wählen Sie aus, was die Kernergebnisse sind.
 - Es gibt keine formalen Vorgaben für das Design der Folien
 - So wenig Text wie möglich
- Vortragsstil
 - Versuchen Sie, frei zu sprechen. Vorlesen ermüdet das Publikum
 - Die meisten Hinweise zu Vorträgen in Seminaren gelten auch für Abschlussarbeiten

<https://docs.google.com/document/d/1PS0uZCsCJs2ehfhyxt3lg23VWuIAjjUWhOhXNj7mKM/edit?usp=sharing>

Bewertungskriterien

Die Bewertung erfolgt nach dem folgenden strukturierten Schema.

Die Gewichtung der Kriterien richtet sich auch nach der inhaltlichen Ausrichtung der Arbeit und dem Schwierigkeitsgrad der Aufgabenstellung.

Kolloquium

Kriterium	Prüfpunkte (Beispiele)							
Umfang und Struktur	- klare Gliederung - sinnvolle Auswahl an Inhalten - Zeitmanagement	<i>unzureichend</i>	<input type="checkbox"/>	<i>sehr gut</i>				
Folien	- Beispiele - Informationsdichte / Anschaulichkeit - Graphiken / Diagramme / Tabellen	<i>unzureichend</i>	<input type="checkbox"/>	<i>sehr gut</i>				
Präsentation	- sprachliche Gestaltung (frei sprechen vs. ablesen) - Blickkontakt	<i>unzureichend</i>	<input type="checkbox"/>	<i>sehr gut</i>				
Diskussion	- angemessene Reaktion auf Nachfragen	<i>unzureichend</i>	<input type="checkbox"/>	<i>sehr gut</i>				
Ergänzender Kommentar:								

Schriftliche Ausarbeitung

Kriterium	Prüfpunkte (Beispiele)	Bewertung						
Umfang und Strukturierung	- klare Gliederung - angemessener Umfang, Teile sinnvoll gewichtet - Vollständigkeit der Darstellung	<i>unzureichend</i>	<input type="checkbox"/>	<i>sehr gut</i>				
Sprachliche Darstellung	- Verständlichkeit, schlüssige Argumentation - Präzision und Korrektheit der Darstellung - angemessener Sprachstil - korrekte Verwendung der Fachsprache	<i>unzureichend</i>	<input type="checkbox"/>	<i>sehr gut</i>				
Nicht-sprachl. Inhalte	- verständliche aussagekräftige Abbildungen und Tabellen, Formeln, Dokumentation von Artefakten (Code...)	<i>unzureichend</i>	<input type="checkbox"/>	<i>sehr gut</i>				
Formale Güte	- korrekte Rechtschreibung und Grammatik - Layout und Erscheinungsbild	<i>unzureichend</i>	<input type="checkbox"/>	<i>sehr gut</i>				
Referenzen	- regelgerechte Referenzierung der Quellen - vollständiges und korrektes Literaturverzeichnis - keine Anhaltspunkte für Plagiat (andernfalls begründ.)	<i>unzureichend</i>	<input type="checkbox"/>	<i>sehr gut</i>				
Ergänzender Kommentar:								

Fachliches Ergebnis

Kriterium	Prüfpunkte (Beispiele)	Bewertung
Ergebnisgüte	<ul style="list-style-type: none"> - Qualität und Vollständigkeit der Aufgabenerfüllung - Güte der Implementierung - Güte der Untersuchung 	<i>gering</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <i>hoch</i>
Wissenschaftliche Fundierung	<ul style="list-style-type: none"> - Ausarbeitung der theoretischen Grundlagen - Abdeckung der relevanten und aktuellen Literatur - Fokussierung des Themas - Ableitung der Forschungsfragen 	<i>gering</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <i>hoch</i>
Originalität	<ul style="list-style-type: none"> - Neuigkeitsgrad der Lösung - Fortschritt gegenüber Stand der Forschung 	<i>gering</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <i>hoch</i>
Methodik	<ul style="list-style-type: none"> - Sinnvolle Werkzeugverwendung - Angemessene Untersuchungsmethodik u. Datenauswertung 	<i>nicht angemessen</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <i>angemessen</i>
Potenzial	<ul style="list-style-type: none"> - Umsetzbarkeit, praktische Anwendbarkeit - Weiterverwendungsmöglichkeit - Herausarbeitung weiterführender Forschungsfragen 	<i>gering</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <i>hoch</i>
Ergänzender Kommentar:		

Vorgehensweise bei der Bearbeitung (nur bei Erstgutachten)

Kriterium	Prüfpunkte (Beispiele)	Bewertung
gezeigte Fachkompetenz	<ul style="list-style-type: none"> - Kenntnis relevanter Methoden und Techniken - Fähigkeit zur Problemlösung 	<i>gering</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <i>hoch</i>
Selbständigkeit	<ul style="list-style-type: none"> - Selbständige Recherche - Eigeninitiative bei der Bearbeitung - Eigenständige Lösungsvorschläge - zielführende Kommunikation mit Betreuer 	<i>gering</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <i>hoch</i>
Systematik des Vorgehens	<ul style="list-style-type: none"> - Planung der Arbeitsschritte, Zeitmanagement - Sorgfalt und Verlässlichkeit bei der Bearbeitung 	<i>gering</i> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <i>hoch</i>
Ergänzender Kommentar:		