

Self-Instruct

SOTA2 - Group C

Anton Kesy, Étienne Muser, Katharina Schindler, Lukas Fehrenbacher, Nico Ruschmann

Offenburg University of Applied Sciences

WS 2024/2025



Instruction Tuning (IT)

Instruction Tuning (IT) ist eine Methode, um große Sprachmodelle (LLMs) darauf zu trainieren, menschliche Anweisungen zu verstehen und zu befolgen.

Ziel: LLMs sollen verschiedene Aufgaben ausführen können, indem sie einfach neuen Anweisungen folgen, ohne dass eine zusätzliche Feinabstimmung erforderlich ist.

Vorgehensweise: Anstelle des Trainings auf rohen Textdaten verwendet IT Datensätze, die aus Paaren von Anweisungen und gewünschten Ausgaben bestehen.

- Ermöglicht schnelle und effiziente Anpassung an neue Aufgaben.
- Reduziert den Bedarf an großen, aufgabenspezifischen Datensätzen.
- Näher an menschlicher Intelligenz im Hinblick auf das Erlernen neuer Aufgaben.

Dual-Track Scaling ist die gleichzeitige Skalierung von Modellgröße und Aufgabenanzahl beim Instruction Tuning

Skalierung der Modellgröße: - Größere Modelle profitieren tendenziell stärker von IT als kleinere Modelle. - Große LLMs ohne IT können kleinere, auf vielen Aufgaben abgestimmte Modelle übertreffen.

Skalierung der Aufgabenanzahl: - Modelle gleicher Größe, die auf einer größeren Anzahl von Aufgaben trainiert werden, erzielen in der Regel eine bessere Leistung. - Die Leistungsverbesserung durch die Skalierung der Aufgabenanzahl ist jedoch instabil, insbesondere bei kleinen Modellen.

Die Vorteile der Skalierung der Modellgröße und der Aufgabenanzahl sind eng miteinander verbunden. Dual-Track Scaling ist entscheidend für ein effektives Instruction Tuning.

Empirische Analyse: Instruction Tuning

Aufgabenspezifische Anweisungen

- **Datensatz:** FLAN-T5
- **Umfang:** 1.836 Aufgaben
- **Anweisungen:** 15 Millionen
- **Fokus:** Spezifische NLP-Aufgaben

Chat-Anweisungen

- **Datensatz:** ShareGPT
- **Umfang:** 63.000 Anweisungen
- **Quelle:** Reale Mensch-ChatGPT-Konversationen
- **Charakteristik:** Dialogorientiert

Synthetische Anweisungen

- **Datensatz:** Self-Instruct-52K
- **Umfang:** 82.439 Anweisungen
- **Generierung:** Automatisch durch LLM
- **Besonderheit:** Künstlich erzeugte Trainingsdaten

Leistungsvergleich

- Aufgabenspezifisch: Optimal für Frage-Antwort
- Chat-Anweisungen: Beste dialogische Interaktion
- Synthetische Daten: Höchstes Verbesserungspotenzial

Schlüsselergebnis

Kombinierte Datensätze zeigen die beste Modellperformance

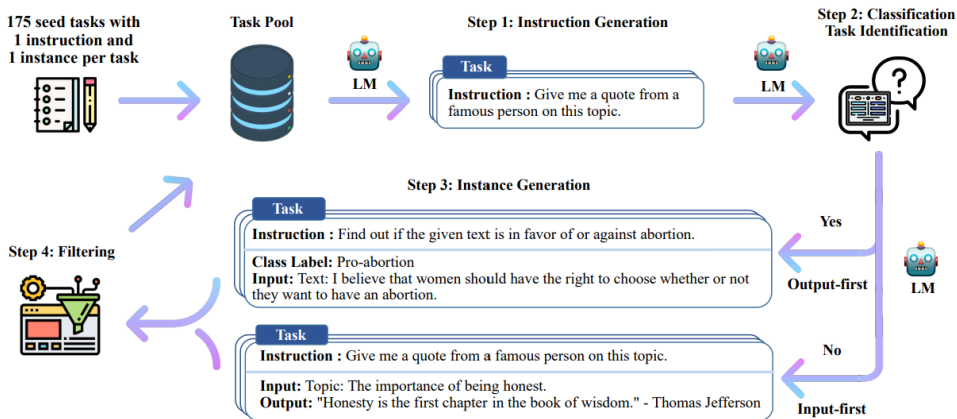
Experimentelle Erkenntnisse

Models	Dataset Mixtures	Instruction Numbers	Lexical Diversity	Chat	QA	
				AlpacaFarm	MMLU	BBH3k
LLaMA (7B)	① FLAN-T5	80,000	48.48	23.77	38.58	32.79
	② ShareGPT	63,184	77.31	81.30	38.11	27.71
	③ Self-Instruct-52K	82,439	25.92	/*	37.52	29.81
	② + ③	145,623	48.22	71.36	41.26	28.36
	① + ② + ③	225,623	48.28	70.00	43.69	29.69
	③ Self-Instruct-52K	82,439	25.92	/*	37.52	29.81
	w/ complexity	70,000	70.43	76.96	39.73	33.25
	w/ diversity	70,000	75.59	81.55	38.01	30.03
	w/ difficulty	70,000	73.48	79.15	32.55	31.25
	w/ scaling	220,000	57.78	51.13	33.81	26.63
LLaMA (13B)	① FLAN-T5	80,000	48.48	22.12	34.12	34.05
	② ShareGPT	63,184	77.31	77.13	47.49	33.82
	③ Self-Instruct-52K	82,439	25.92	/*	36.73	25.43
	② + ③	145,623	48.22	72.85	41.16	29.49
	① + ② + ③	225,623	48.28	69.49	43.50	31.16
	③ Self-Instruct-52K	82,439	25.92	/*	36.73	25.43
	w/ complexity	70,000	70.43	77.94	46.89	35.75
	w/ diversity	70,000	75.59	78.92	44.97	36.40

Was ist Self-Instruct?

- Verwendung zur Anleitung von LLM um neue Anweisungen und Instanzen zu generieren
- Iterativer Prozess, der bis zu Stoppkriterien durchgeführt wird
- Vorbereitung: Generieren von manuell geschriebenen Aufgaben (Seed-Tasks)

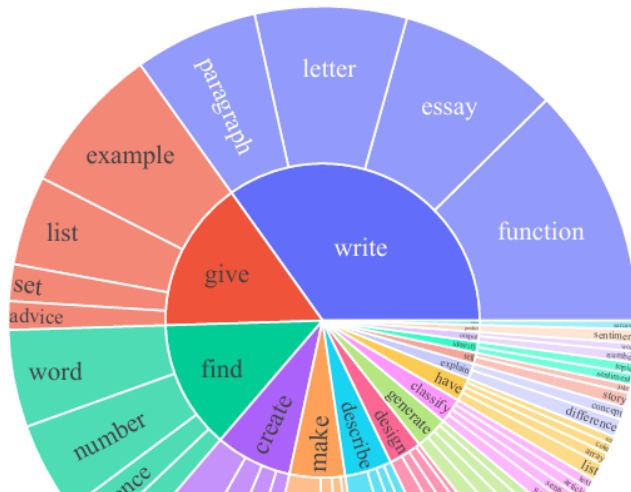
Der Self-Instruct Prozess



- GPT3 generiert instructions
- GPT3 wird mit den instructions fine getuned
- Fine tuning mittels OpenAI API
- Fine tuning hyperparameter:
 - default
 - prompt loss weight: 0
 - Epochen: 2

Ergebnisse - Diversität der generierten instructions

- Top 20 root Verben + deren top 4 Substantive:
- Auch einige neue instructions



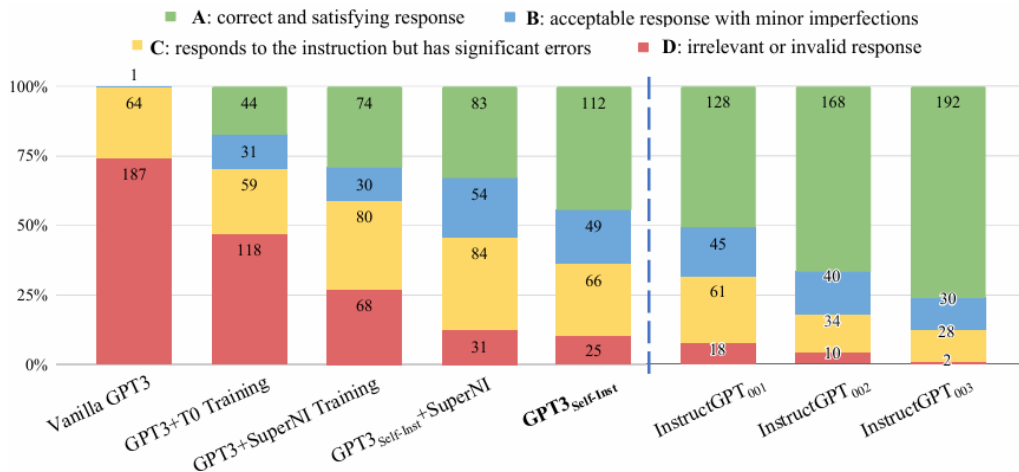
Ergebnisse - Qualität der generierten instructions

- Stichprobenartige Überprüfung durch Experten

Quality Review Question	Yes %
Does the instruction describe a valid task?	92%
Is the input appropriate for the instruction?	79%
Is the output a correct and acceptable response to the instruction and input?	58%
All fields are valid	54%

Qualität

Ergebnisse - Qualität des fine-getuneten Modells



Ergebnisse

- **Geringerer Bedarf an menschlichen Annotationen**
 - reduziert die Abhängigkeit von menschlich geschriebenen Anweisungen (teuer und zeitaufwendig)
- **Verbesserte Generalisierung**
 - verbesserte Fähigkeit Anweisungen für neue und unbekannte Aufgaben zu befolgen
- **Erstellung vielfältiger Datensätze**
 - ermöglicht die Generierung vielfältiger Anweisungsdaten, die über typische NLP-Aufgaben hinausgehen

- **Abhängigkeit von großen Modellen**
 - hängt von den induktiven Verzerrungen ab (aus großen Sprachmodellen extrahiert)
 - Ansatz für Benutzer mit begrenzten Rechenressourcen möglicherweise unzugänglich
- **Verstärkung von Modellverzerrungen**
 - iterative Prozess könnte unbeabsichtigt problematische soziale Verzerrungen verstärken
- **Tail-Phänomene**
 - größten Verbesserungen könnten auf häufig verwendete Aufgaben oder Anweisungen beschränkt sein
 - seltenen und kreativen Anweisungen möglicherweise weniger effektiv