

Feintuning offener Sprachmodelle für die Branchenklassifikation von Unternehmen

Hintergrund zum Projektthema:

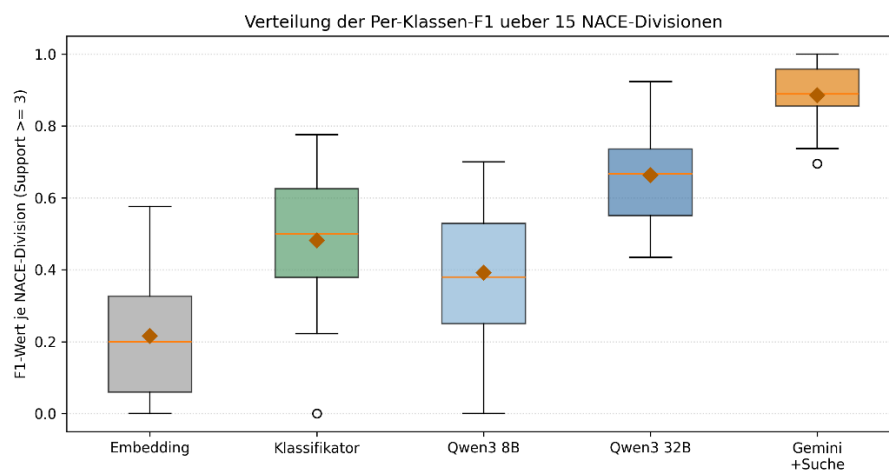
Große Sprachmodelle können Unternehmen erstaunlich gut Wirtschaftszweigen (NACE) zuordnen, doch die besten Ergebnisse liefern bislang proprietäre KI-Dienste. Offene, lokal betreibbare Modelle (z.B. Qwen3) erreichen am IKKU ohne Training bereits 65 bis 78 % Trefferquote. Erste eigene Untersuchungen zeigen, dass die Fähigkeit zur Nutzung von Web-Belegen erst mit der Modellgröße entsteht. Ob gezieltes, ressourcenschonendes Feintuning offener Modelle die Lücke zu den Closed-Source-KI Diensten schließt, ist die treibende Frage der Abschlussarbeit.

Aufgabenstellung des Projektes:

Pretrained Foundation Models (PFM's) sollen mit geeignetem Feintuning (z.B. LoRA) auf die Zuordnung von Firmenname plus Unternehmenstexten zu Wirtschaftszweigen feinjustiert werden. Der Vergleich erfolgt systematisch gegen die vorhandenen Referenzmethoden: Zero-Shot-Klassifikation, trainierte klassische Klassifikatoren und ein websuchengestütztes Referenzmodell, alle auf demselben Validierungssample. Untersucht werden der Einfluss von Modellgröße, Trainingsdatenmenge und Evidenzqualität sowie die Kalibrierung der Modellgenauigkeit.

Die Projektarbeit wird folgende Ziele bzw. Ergebnisse beinhalten:

- Literaturrecherche zu parameter-effizientem Feintuning (LoRA/QLoRA) und LLM-basierter Unternehmensklassifikation
- Aufbau der Trainingspipeline auf dem bwHPC-Cluster DACHS
- Feintuning-Experimente mit Ablation von Modellgröße und Trainingsdatenmenge
- Benchmark gegen Zero-Shot, klassische Klassifikatoren und Websuche-Referenz (Trefferquote, Makro-F1, Kalibrierung) auf dem eingefrorenen Validierungssample
- Einordnung gegenüber dem internationalen Stand (u. a. MONETA-Benchmark der Deutschen Bundesbank)
- Dokumentation der Ergebnisse, ggf. Beitrag zu einer wissenschaftlichen Veröffentlichung



Bearbeitung im:

WS 2026/27

Zielgruppe:

Master Studiengang

Projekt-Code:

27ws_CB_PFM

Betreuer (Professor): Prof. Constanze Bongs; Tel.: [+49 721 925-1847](tel:+497219251847); E-Mail: Constanze.Bongs@h-ka.de
Eventueller Zweitbetreuer: Philipp Kötter; Tel.: [+49 721 925-2083](tel:+497219252083); E-Mail: Philipp.koetter@h-ka.de