



Ausschreibung Mono-Informatik Bachelorarbeit

„Chat Bots – Unterschiede der Absichtserkennung bei kleinen vs. großen Trainingsdatensmengen“

Beschreibung:

Mithilfe von NLP- und Machine Learning Methoden soll in zwei Domänen für frei per Chat geschriebene und thematisch vorklassifizierte Kundenanfragen zwischen „allgemeiner Frage“ und dem Wunsch, eine bestimmte Aktion auszulösen, unterschieden werden. Zum Beispiel soll beim Thema die allgemeine Frage „Welche Unterlagen benötige ich, um meine Kreditkarte zu sperren?“ vom Wunsch der Sperrung „Ich möchte meine Kreditkarte sperren.“ unterschieden werden. Dabei sollen die Herausforderungen der Unterscheidung herausgearbeitet und die Lösungsansätze in Abhängigkeit der Menge an Trainingsdaten priorisiert und beispielhaft implementiert werden.

Die Arbeit erfolgt in Zusammenarbeit mit dem KI-Startup Aaron.ai und ist in folgende Phasen/ Aufgaben aufgeteilt

- Anforderungsanalyse. Literaturreview zu Intent Detection Methoden; Auswahl Kriterien für Priorisierung
- Entwicklung. Preprocessing Trainingsdaten; Priorisierung geeigneter Methoden; Implementierung ausgewählter Methoden auf beispielhaften Domänen mit unterschiedlicher Trainingsdatensmenge
- Evaluierung. Statistische Auswertung von Präzision der Methoden auf den unterschiedlichen Datensätzen im Vergleich; Fazit und Ausblick

Anforderungen: gute natural language processing und machine learning Kenntnisse; solide Python Kenntnisse

Kontaktperson: N.-T. Le, R. 3.406, nguyen-thinh.le@hu-berlin.de