

Abstract

Smart personal assistants like Siri or Alexa use machine learning based text analysis to understand commands and questions. Underlying technologies require a huge amount of training data and their production often involves human annotators doing repetitive and monotonous work. The purpose of the present study is to identify possibilities to support the text annotation task with automated assistance. Since these annotation assistance systems cannot be perfectly accurate the influence of reliability is analyzed, too. We developed a fully functional, autonomous assistance system, specialized on named entity recognition to prove the feasibility of such an assistance. It creates suggestions for annotators based on what it has learned from previous annotations. For the study we simulated such a system. Subjects ($N = 66$) accomplished annotation tasks with and without the system in a counterbalanced order. Three levels of reliability (10%, 50%, or 90% correct suggestions) of the system were benchmarked. Dimensions measured were correct annotations (hits), misses of annotations, and time per correct annotation. A 2×3 mixed design (assistance present / not present, within and the three levels of reliability, between) was used. An assistance system providing a reliability of 50% or 90% improves accuracy, average time per correct annotation, as well as reducing misses of the human annotations significantly. Supporting the task of text annotation by using an automated assistance will improve humans' performance if the suggestions of the system are mature.

Zusammenfassung

Persönliche Assistenzsysteme wie Siri oder Alexa nutzen Textanalysen, um Eingaben und Fragen zu verstehen. Die darunterliegende, auf maschinellem Lernen beruhende Technologie benötigt große Trainingsdatensätze, deren Erstellung oft mit monotonen und repetitiven Aufgaben für menschliche Annotatoren einhergeht. Ziel der vorliegenden Studie ist es, Möglichkeiten zu identifizieren, um die Textannotationen mit einem automatischen Assistenzsystem zu unterstützen. Da solche Systeme nicht unfehlbar sein können, wird untersucht, welchen Einfluss die Verlässlichkeit des Assistenzsystems hat. Wir entwickelten ein vollständig funktionierendes, auf Eigennamenerkennung spezialisiertes, autonomes Assistenzsystem, um die grundsätzliche Umsetzbarkeit einer solchen Assistenz zu beweisen. Das System lernt aus vorhergehenden Annotationen und generiert Vorannotationen in neuen Texten. Diese dienen als Vorschläge für die Annotatoren. Für die Studie wurde ein solches System simuliert. Die Probanden ($N = 66$) annotierten Texte abwechselnd mit und ohne Assistenzsystem. Dabei gab es drei Verlässlichkeitsstufen der Assistenz: Es lieferte 10%, 50%, oder 90% korrekte Vorschläge. Als abhängige Variablen wurden korrekte Annotationen, fehlende Annotationen und Zeit pro korrekter Annotation erhoben. Die Studie war in einem 2×3 gemischten Design (mit / ohne Assistenzsystem und die drei Stufen der Verlässlichkeit) aufgebaut. Wir zeigen, dass ein Assistenzsystem mit 50% oder 90% richtigen Vorschlägen bei den Probanden die Rate der richtigen Annotationen steigert, die Zeit pro korrekter Annotation senkt und die Fehleranzahl vermindert. Die menschliche Leistung bei Textannotationen kann durch ein automatisches Assistenzsystem verbessert werden — wenn es eine hohe Treffsicherheit in seinen Vorschlägen hat.