

Parkbank oder Bankkonto?

Wenn Computer lernen uns zu verstehen

Autorin: Morna Florack, Account Executive, Financial Services/Travel & Tourismus

Virtuelle Assistenten haben in vielen Lebensbereichen Einzug gehalten, spielen Musik ab, führen uns durch Webseiten oder machen das Licht an. In diesem Kontext verzeihen wir so manche Stilblüte bzw. sie bringt uns eher zum Schmunzeln.

Wenn es allerdings um Geldanlagen geht, ist das richtige Verständnis des Gesagten oder Getippten eminent wichtig, um das benötigte Vertrauen der Kunden und Mitarbeiter aufzubauen.

Um Sparern die Geldanlage zu erleichtern und diese kosteneffizient per Software anzubieten, haben sogenannte Robo-Advisor Einzug gehalten. Diese benötigen Angaben zum Risiko-Typ und Anlageziel, um die Geldanlagen durchführen zu können. Damit man sie dialogorientiert oder auch als telefonischen Self-Service sinnvoll einsetzen kann, ist es hilfreich die Hintergründe zu beleuchten, wie diese Systeme funktionieren und wie sie den Nutzer verstehen.

Sprache verstehen und richtig umsetzen

Sprachverstehen ist die Kernkomponente für erfolgreiche virtuelle Assistenten, da sie es ermöglicht das Anliegen des Nutzers beziehungsweise die Bedeutung eines Satzes richtig zu verstehen. Dies gilt sowohl für die Nutzung im Rahmen von Sprachdialogen, die z.B. mit dem telefonischen Kundenservice geführt werden als auch für Chatbots, die auf der Webseite, über SMS oder allgemein in den digitalen Kanälen genutzt werden. Der Begriff *natürliche Sprache* bezieht sich dabei auf Sprache, die Menschen in der Kommunikation untereinander verwenden und wenn Computer diese „versteht“, spricht man von *Natural Language Understanding* (NLU).

Das weitere Feld der Verarbeitung, *Natural Language Processing* (NLP) oder in Deutschland auch als *Computerlinguistik* bekannt, umfasst NLU und erweitert es z.B. mit Spracherkennung, Sprachsynthese, mit syntaktischen Analysen oder auch automatischer Übersetzung.

BERT – Ein wichtiger Entwicklungssprung

Einen Schub erhielt die Entwicklung von NLU durch Google im Jahr 2018 durch die Veröffentlichung ihres neuen Suchmaschinenalgorithmus mit dem Ziel komplexe Suchanfragen zu verstehen. BERT¹, steht für *Bidirectional Encoder Representations from Transformers* und kann durch den bidirektionalen Ansatz den

¹ published in October 2018 as BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, the paper was authored by Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova.

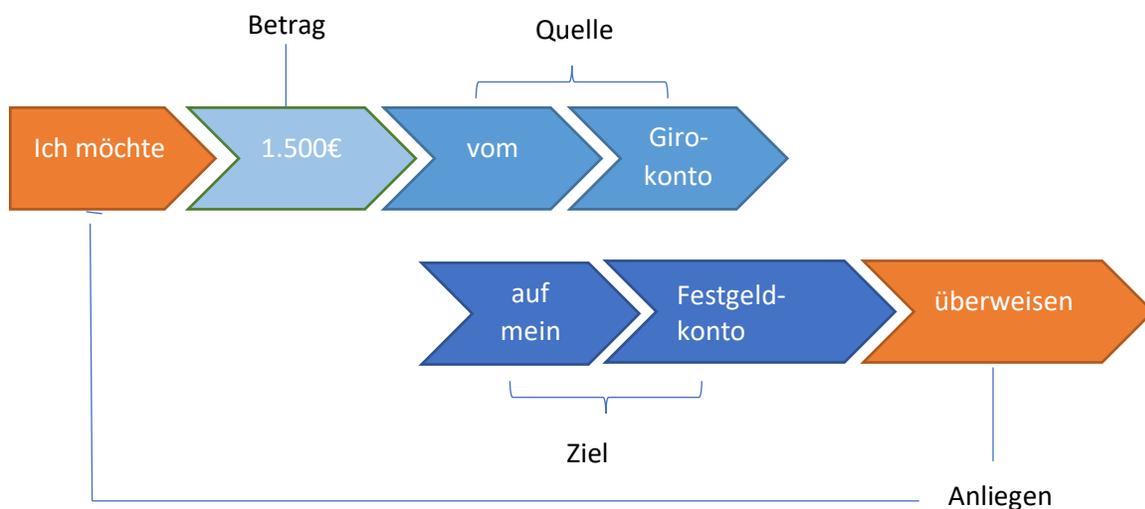
Kontext mehrerer Worte im Satz verstehen, und zwar unabhängig von der Aufeinanderfolge dieser Worte („... eröffne bei der **Bank** ein **Konto**“ versus „im **Park** wurde eine neue **Bank**...“). BERT ist vortrainiert auf dem Text-Korpus von Wikipedia und hat so ein Verständnis der jeweiligen Sprache, das in einigen Bereichen menschliche Vergleichswerte übertrifft².

Seit der Einführung von BERT hat es zahlreiche Weiterentwicklungen und Spezialisierungen gegeben, die die Leistungsfähigkeit gesteigert haben. So konnten in der letzten Zeit wegweisende Verbesserungen im Bereich NLU erzielt werden, um die Nutzer korrekt zu verstehen und dialogorientierte Anwendungen zu entwickeln bei denen die Anwender nicht in der Vielfalt ihrer Ausdrucksweisen limitiert sind. Neben dem Ziel der Genauigkeitssteigerung wurde auch an dem Ziel des ressourcenschonenden Hardwareeinsatzes gearbeitet (z.B. ALBERT A Lite BERT).

Mehr als Verstehen des Anliegens

NLU-Vergleichstests zielen auf die verschiedenen Einsatzbereiche des Sprachverstehens ab. Als Teil der General Language Understanding Evaluation (GLUE³) gibt es zum Beispiel Test-Sets zum „Question Answering“ (richtige Antworten auf Basis von Wikipedia) oder „Textual Entailment“ (die richtigen Hypothesen zu einem Text) und weitere. Diese zeigen eindrucksvoll die Leistungsfähigkeit von NLU, lassen aber nicht immer den direkten Rückschluss zu, dass das System, welches hier am besten abschneidet, auch das geeignetste für den jeweiligen kommerziellen Einsatzzweck ist.

Im kommerziellen Bereich legen viele Vergleichstests ihren Schwerpunkt auf das korrekt erkannte Anliegen. Dies ist vordergründig auch richtig, da dem Anliegen in der Konversation eine hohe Bedeutung zukommt und sichergestellt werden muss, dass der Nutzer des Systems sich „richtig verstanden“ fühlt. Dennoch fehlen dabei wichtige Aspekte, auf die ich im Folgenden näher eingehen werde:



Um kommerziell erfolgreiche, effiziente Dialoge führen zu können, benötigt man weitere Fähigkeiten aus dem NLP Bereich, wie die sogenannte „Entity-Recognition“. Unter Entity-Recognition oder auch Named-

² <https://rajpurkar.github.io/SQuAD-explorer/>

³ General Language Understanding Evaluation: <https://gluebenchmark.com/>

Entity-Recognition, versteht man in diesem Zusammenhang die Extraktion und Zuordnung von unstrukturiertem Text teilen zu vordefinierten Kategorien wie z.B. Namen, Beträgen oder Daten. **Hierbei sollen relevante Inhalte der Äußerung ggf. einer oder mehrerer Entitäts erkannt und zugeordnet werden können**, um z.B. Sätze wie: „ich möchte 1.500 € vom Girokonto auf mein Festgeldkonto überweisen“ zu verarbeiten.

Weitere praktische relevante Fähigkeiten, die einen effizienten und menschlichen Dialog zwischen virtuellem Assistenten und Nutzer ermöglichen, sind:

Verneinung: Für den Mensch ist der Einsatz von Verneinung natürlich, doch für die Maschine ist es eine große Herausforderung, da es hierbei besonders auf die Stellung im Satz ankommt. Die Antworten innerhalb eines Überweisungsdialogs: „Spar **nicht** Giro“ oder „**Nein**, 1.600 €“ stellen Systeme vor die Herausforderung richtig zuzuordnen, worauf sich „Nein“ oder „nicht“ bezieht und zu erkennen, ob die ursprüngliche Aufgabe mit einer korrigierten Quelle bzw. einem korrigierten Wert durchgeführt werden soll.

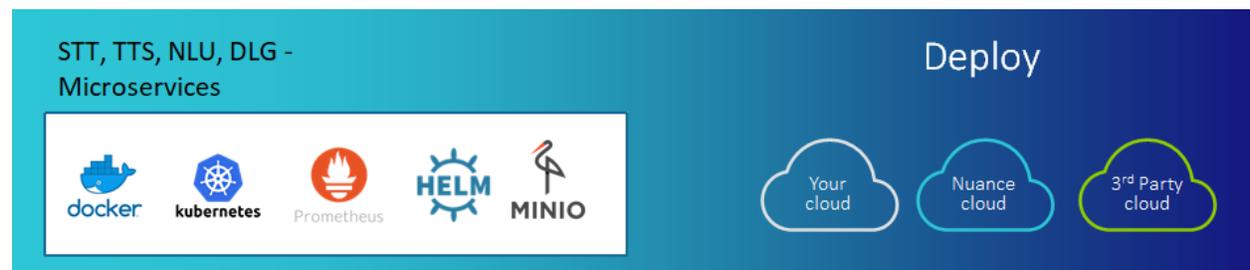
Zur Verneinung zählen auch Aufgabenstellungen wie: „Zeige mir alle Abbuchungen von XXX **außer** die monatlichen Abos“ oder „investiere im Energiesektor **außer** in Kernkraft“. An diesen Beispielen sehen Sie, dass mit Nutzung der neuen NLU-Möglichkeiten spürbar natürlichere Dialoge mit automatischen Systemen möglich sind und diese Zeit für den Kunden einsparen. Ohne diese erweiterten NLU-Fähigkeiten waren zuvor zwei oder mehr Dialogschritte notwendig, um denselben Inhalt auszutauschen (1. „In welchem Bereich möchten Sie investieren?“ „Energiesektor“ 2. „Gibt es Ausnahmen hierzu?“ „Ja, Kernenergie“). Mehr Dialogschritte in veralteten Systemen erhöhen nicht nur die aufgewendete Zeit, sondern auch das Risiko eines Dialogabbruchs und somit je nach Kontext unzufriedenere Kunden und ggf. auch verpasste Verkaufsmöglichkeiten.

Referenz: In Folgesätzen nutzen Menschen gerne Verweise, wie „Ich bezahle mit der *anderen* Karte“ oder „Liefere Sie es *dorthin*“. Hier ist es Aufgabe der NLU auch diese Verweise zu erkennen, so dass der virtuelle Assistent sie dann im Kontext richtig verstehen und bearbeiten kann. Damit werden für den Nutzer nicht notwendige Wiederholungen vermieden und wiederum ein kürzerer, natürlicherer Dialog ermöglicht.

Nächste Schritte in die Cloud

Neue Möglichkeiten mit NLU erfordern auch mehr Rechenleistung. Hier ist es wichtig auf ein System zu setzen, welches durch einen Cloud-native Ansatz, die volle Flexibilität im Einsatz ermöglicht.

So kann der Übergang vom Bestandssystem in die zukünftige Infrastruktur als SaaS Angebot nahtlos gestaltet werden ohne zusätzliches Investitionsvolumen in Hardware zu binden.



Die einzelnen Komponenten eines Gesamtsystems (Speech-to-Text, Text-to-Speech, NLU und die eigentliche Dialog-Engine) werden bei dieser Architektur als sogenannte Microservices bereitgestellt. Um diese zu nutzen wird ein Rahmenwerk um die Systeme Docker⁴ und Kubernetes⁵ entsprechend der Bedürfnisse zusammengestellt und betrieben. Dieser Einsatz kann dann innerhalb der jeweiligen Unternehmens-Infrastruktur („Your Cloud“) in einer öffentlichen Cloud oder als Service durch Anbieter wie Nuance bereitgestellt werden.

Nutzen Sie die Erfahrung anderer, um Vertrauen aufzubauen

Künstliche Intelligenz „lebt“ von den Trainingsdaten. Je relevanter die Trainingsdaten sind, desto besser auch die zu erwartenden Ergebnisse. Systeme, die bereits mit branchenspezifischen Daten vortrainiert sind, helfen so Trainingsaufwand für die einzelne Applikation zu verringern und ermöglichen dadurch geringere Erstellungskosten und eine schnellere Markteinführung.

Anbieter mit relevanter Erfahrung aus Kundenkontakten sind hier klar im Vorteil, da z.B. geschriebene Sprache, wie sie in Wikipedia-Artikeln jedem zur Verfügung steht, sich deutlich von gesprochener Sprache unterscheidet und daher weniger gut zum NLU-Training geeignet ist. Deshalb ist es wichtig auf einen Anbieter zu setzen, der Erfahrung mitbringt und seine Produkte sowohl mit anwendungs-relevanten Daten trainieren konnte als auch bei Bedarf sicherstellen kann, dass die eigenen Kundendaten nur für das jeweilige Unternehmen eingesetzt werden und nicht vom NLU-Anbieter anderweitig kommerziell genutzt werden. Dies ist für das Unternehmen notwendig, um sicherzustellen, dass DS-GVO-konform Daten nur zu den Zwecken eingesetzt werden, zu denen der Nutzer auch zugestimmt hat bzw. die von der Verordnung abgedeckt sind.

Ausblick: Conversational Banking

Die Grundfähigkeit den Kunden zu verstehen eröffnet einen weiteren Einsatzbereich virtueller Assistenten über den eigentlichen Kundenservice hinaus – das Conversational Banking. Bisher war die Automatisierung einfacher Aufgaben wie Kontostands-Ansagen und Durchführung von Überweisungen im Fokus. Mit Conversational Banking können den Kunden weitere Serviceleistungen angeboten werden, z.B. interaktive Analysen des Ausgabeverhaltens oder Hinweise bei Überschreiten einer Sparquote auf Anlagemöglichkeiten oder auch der oben erwähnten dialogorientierte Robo-Advisor.

Solche Serviceleistungen werden von Kunden als hilfreich empfunden und bieten dadurch sowohl weitere Chancen im individuellen Kundenkontakt als auch eine stärkere Bindung des Kunden an das Institut. Und machen dadurch schlussendlich den Unterschied in der Kundenzufriedenheit und -bindung.

⁴ <https://www.docker.com/>

⁵ <https://kubernetes.io/>