

Sprache, Spracheingabe, Text & Übersetzung mit Google Cloud APIs

M.Sc. Onur Yilmaz

Angewandte Künstliche Intelligenz

Schriftliche Ausarbeitung - Cloud Computing
Fachhochschule Südwestfalen

Gutachter: Prof. Dr. Giefers

10. August 2023

Inhaltsverzeichnis

1	Grundlagen	3
1.1	Einführung in Cloud Computing	3
1.1.1	Was ist eine API und ihre Funktion?	3
1.1.2	Bedeutung von API-Keys	3
1.2	Erstellen und Aufrufen der API-Anfrage	4
1.2.1	Grundlagen des API-Aufrufs	4
1.2.2	Spezifikationen und Parameter	4
1.3	Interaktion mit der Google Cloud API	5
1.3.1	Umgang mit dem Google Cloud Dienst	5
1.3.2	Erstellen eines API-Keys in Google Cloud	5
2	Spracherkennung und -transkription	6
2.1	Schritte im Speech-to-Text Prozess	6
2.2	Google Cloud Speech API	6
3	Sprachübersetzung	7
3.1	Erkennung und Übersetzung von Texten	7
3.2	Cloud Translation API	7
4	Textanalyse	8
4.1	Klassifizierung von Text in Kategorien	8
4.2	Entitäten- und Sentimentanalyse	8
4.3	Natural Language API	8
5	Sprachsynthese	9
5.1	Erzeugung synthetischer Sprache	9
5.2	Natural Language API	9

Einleitung

In dieser Arbeit werden Technologien und Anwendungen aus dem Bereich der Sprach- und Textverarbeitung beleuchtet, die auf Google Cloud-Diensten aufbauen. Hierzu gehören die *Cloud Speech API*, die *Cloud Translation API*, die *Natural Language API* und die *Text-to-Speech API* [1].

Im Abschnitt über die *Grundlagen* wird eine Einführung in die Konzepte des Cloud Computings, der API und des Umgangs mit dem Google Cloud Dienst gegeben. Hierbei wird insbesondere auf die Themen API und API-Keys sowie Natural Language Processing (NLP) eingegangen.

Der Abschnitt *Spracherkennung und -transkription* fokussiert sich auf die *Cloud Speech API*, die die Transkription von Audio in Text ermöglicht, und die Methoden zur Messung und Verbesserung der Sprachgenauigkeit.

In der *Sprachübersetzung* wird die *Cloud Translation API* behandelt, die den Prozess der Übersetzung von Texten in verschiedene Sprachen ermöglicht.

Der Bereich *Textanalyse* befasst sich mit der *Natural Language API*, die Techniken zur Klassifizierung von Text in Kategorien und zur Analyse von Entitäten und Sentiments bietet.

Im Abschnitt *Sprachsynthese* wird die *Text-to-Speech API* vorgestellt, die die Erzeugung synthetischer Sprache ermöglicht.

Die Arbeit dient nicht nur als theoretischer Überblick, sondern bietet auch praktische Einblicke und Anleitungen zur Verwendung dieser Tools. Dabei werden unterschiedliche Schwierigkeitsgrade und Themenbereiche abgedeckt, um einen umfassenden Einblick in die Möglichkeiten der Sprach- und Textverarbeitung mit Google Cloud zu bieten.

1 Grundlagen

1.1 Einführung in Cloud Computing

Cloud Computing bezeichnet das Angebot und die Nutzung von IT Infrastrukturen und Anwendungen, die nicht lokal auf dem eigenen Computer, sondern in einem Netzwerk, meistens dem Internet, ausgeführt werden. Durch Cloud Computing können Unternehmen, aber auch Einzelpersonen Speicherplatz, Rechenleistung oder Softwareanwendungen als Service über das Internet nutzen.

Eine der Hauptanwendungen von Cloud Computing in der heutigen Zeit ist die Datenverarbeitung. Besonders hervorzuheben ist hier die Verwendung von Cloud-Plattformen für Natural Language Processing (NLP), da sie die Fähigkeit bieten, große Mengen an Textdaten effizient zu analysieren und zu verarbeiten. Nachfolgend wird speziell auf die Möglichkeiten und Dienste der Google Cloud Platform eingegangen.

1.1.1 Was ist eine API und ihre Funktion?

Wenn man über NLP in der Cloud spricht, ist eine der Hauptinteraktionen die Verwendung von APIs. Zum Beispiel bietet OpenAI, der Entwickler von ChatGPT, ein Cloud-basiertes API-Interface für sein Modell an.

Eine API (Application Programming Interface) dient als Schnittstelle, die es Entwicklern ermöglicht, bestimmte Funktionen eines Programms oder einer Plattform zu nutzen, ohne sich mit den internen Details auseinandersetzen zu müssen.

Im Kontext von NLP könnte eine API es ermöglichen, Textdaten zu senden und als Antwort eine Analyse oder Übersetzung dieses Textes zu erhalten.



1.1.2 Bedeutung von API-Keys

API-Keys sind ein wesentlicher Bestandteil vieler Webanwendungen und Cloud-Dienste. Sie dienen als Identifikator für den Benutzer oder Entwickler und stellen sicher, dass die Anfrage autorisiert ist. Durch den API-Key kann der

Serviceanbieter den Zugriff auf seine Ressourcen kontrollieren, den Verbrauch überwachen und gegebenenfalls Gebühren erheben.

1.2 Erstellen und Aufrufen der API-Anfrage

1.2.1 Grundlagen des API-Aufrufs

1.2.2 Spezifikationen und Parameter

1.3 Interaktion mit der Google Cloud API

1.3.1 Umgang mit dem Google Cloud Dienst

Die Google Cloud bietet eine Vielzahl von Diensten, darunter auch Machine Learning und NLP. Wie bei anderen Cloud-Anbietern auch, benötigen Nutzer einen API-Key, um auf diese Dienste zuzugreifen.

1.3.2 Erstellen eines API-Keys in Google Cloud

Um einen API-Key in der Google Cloud zu erstellen, muss man sich zuerst in der Google Cloud Console anmelden. Dann wählt man das gewünschte Projekt aus oder erstellt ein neues. Unter den „Credentials“-Optionen kann man dann einen neuen API-Key erstellen. Es ist wichtig, den generierten Schlüssel sicher aufzubewahren und nicht öffentlich zugänglich zu machen, da er den Zugriff auf Cloud-Ressourcen und möglicherweise auch Abrechnungsinformationen ermöglicht.

2 Spracherkennung und -transkription

2.1 Schritte im Speech-to-Text Prozess

2.2 Google Cloud Speech API

3 Sprachübersetzung

3.1 Erkennung und Übersetzung von Texten

3.2 Cloud Translation API

4 Textanalyse

4.1 Klassifizierung von Text in Kategorien

4.2 Entitäten- und Sentimentanalyse

4.3 Natural Language API

5 Sprachsynthese

5.1 Erzeugung synthetischer Sprache

5.2 Natural Language API

Literatur

- [1] Google Cloud. *Language, Speech, Text, & Translation with Google Cloud APIs*. Zugriffsdatum: 08. August 2023. 2023. URL: <https://www.cloudskillsboost.google/quests/179>.