

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

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 Was ist eine API?

1.1.1 Definition und Funktion einer API

1.1.2 Bedeutung von API-Keys

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 Einführung in Cloud Computing

1.3.2 Umgang mit dem Google Cloud Dienst

1.3.3 Erstellen eines API-Keys in Google Cloud

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>.