

Sprache, Spracheingabe, Text & Übersetzung mit Google Cloud APIs

M.Sc. Onur Yilmaz

Angewandte Künstliche Intelligenz

Schriftliche Ausarbeitung - Cloud Computing

Betreuer: Prof. Dr. Giefers

Fachhochschule Südwestfalen 4. August 2023

Inhaltsverzeichnis

1	Grundlagen	3
1.1	Was ist eine API	3
1.2	Erstellen eines API-Schlüssels	3
1.3	Erstellen und Aufrufen der API-Anfrage	3
2	Spracherkennung und -transkription	4
2.1	Transkription von Sprache zu Text	4
2.2	Cloud Speech API	4
2.3	Messung und Verbesserung der Sprachgenauigkeit	4
3	Sprachübersetzung	5
3.1	Erkennung und Übersetzung von Texten	5
4	Textanalyse	6
4.1	Klassifizierung von Text in Kategorien	6
4.2	Entitäten- und Sentimentanalyse	6
5	Sprachsynthese	7
5.1	Erzeugung synthetischer Sprache	7

Einleitung

Im Rahmen dieser Arbeit werden verschiedene Technologien und Anwendungen im Bereich Sprache und Textverarbeitung vorgestellt, die auf den Diensten von Google Cloud basieren, einschließlich der *Cloud Speech API*, der *Cloud Translation API*, der *Natural Language API* und der *Text-to-Speech API* [1].

Im Abschnitt über die *Grundlagen* wird eine Einführung in die Konzepte der API, des Cloud Computing und die Relevanz von Sprachtechnologien in der heutigen Zeit gegeben.

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

Eine API (Application Programming Interface) ist eine Schnittstelle, die es verschiedenen Softwareanwendungen ermöglicht, miteinander zu kommunizieren. Es handelt sich im Wesentlichen um eine Reihe von Regeln und Protokollen, die von den Entwicklern befolgt werden müssen, um auf die Funktionen eines Softwareprodukts zuzugreifen.

1.2 Erstellen eines API-Schlüssels

1.3 Erstellen und Aufrufen der API-Anfrage

2 Spracherkennung und -transkription

Die Spracherkennung und -transkription, auch als *Automatic Speech Recognition* (ASR) bekannt, ist ein wichtiger Bereich der künstlichen Intelligenz und der Signalverarbeitung. In diesem Abschnitt werden wir uns damit beschäftigen, wie gesprochene Sprache in geschriebenen Text umgewandelt werden kann. Dieser Prozess wird als Transkription bezeichnet.

2.1 Transkription von Sprache zu Text

Die Transkription von Sprache zu Text ist das Verfahren, bei dem Audiosignale, die menschliche Sprache enthalten, analysiert werden, um den entsprechenden Textinhalt zu erzeugen. Dies kann in vielen Anwendungen nützlich sein, von automatischen Untertiteln für Videos bis hin zur sprachbasierten Interaktion mit virtuellen Assistenten.

2.2 Cloud Speech API

2.3 Messung und Verbesserung der Sprachgenauigkeit

3 Sprachübersetzung

3.1 Erkennung und Übersetzung von Texten

4 Textanalyse

4.1 Klassifizierung von Text in Kategorien

4.2 Entitäten- und Sentimentanalyse

5 Sprachsynthese

5.1 Erzeugung synthetischer Sprache

Literatur

- [1] Google LLC. Language, speech, text, & translation with google cloud apis, 2023.