



Fancy Flashcards Chatbot

von Tamara Bucher, Lea Kleemann und Maria Eichenlaub

Gliederung

- Ziele
- Datenquellen
- Funktionsweise des Chatbots
- Verwendete Modelle
- Evaluierung der Modelle
- Herausforderungen
- Weitere Implementierungsmöglichkeiten
- Live Präsentation

Ziele

- Entwicklung eines Chatbots, der Studenten beim Lernen hilft
- User interagiert mit Chatbot
 - User stellt Fragen
 - Lernen: User beantwortet fragen

**TELEGRAM
BOT WITH
PYTHON**



Datenquellen

- Idee
 - Fancy Flashcards <https://github.com/fancy-flashcard/ffc>
 - Webanwendung zum Lernen von Karteikarten
- Daten
 - Fancy Flashcards Decks <https://github.com/fancy-flashcard/deck-collection/tree/main/wirtschaftsinformatik>
 - Decks enthalten Fragen-Antwort-Paare
 - Decks in deutscher Sprache



Decks

```
1  {
2    "meta": {
3      "author": "Florian Klinke && FFC Quick Check",
4      "version": "1.0"
5    },
6    "decks": {
7      "d0": {
8        "meta": {
9          "deck_name": "Finanzbuchhaltung",
10         "next_card_id": 41,
11         "description": "Das Deck beinhaltet Grundlagen für das Fach Finanzbuchhaltung."
12       },
13       "cards": {
14         "0": {
15           "q": "Was sind die Aufgaben des Rechnungswesens",
16           "a": "Dokumentation, Information & Rechenschaftslegung, Kontrolle und Disposition."
17         },
18         "1": {
19           "q": "Was sind die Bereiche des Rechnungswesens",
20           "a": "Buchführung, Statistik, Kosten-/Leistungsrechnung und Planung"
21         },

```

Funktionsweise des Chatbots

- Fragen an Chatbot stellen
 - Chatbot sucht ähnlichste Frage
 - Beantwortet Frage
- Chatbot stellt Frage
 - Deck wird gewählt
 - Chatbot stellt zufällige Frage aus diesem Deck
 - Evaluert die vom User gegebene Antwort anhand der Musterantwort
 - Gibt Rückmeldung
- Timer zum Lernen starten

Verwendete Modelle

Chatbot stellt Fragen an User und User antwortet

- Sentence Transformer
 - Erstellung von Sentence Embeddings
 - Berechnung der Ähnlichkeit zwischen der Musterlösung und der eingegebenen Antwort des Users

Verwendete Modelle

User stellt Frage an Chatbot

- tf-idf
 - Erstellung von Dokumentenvektoren der Fragen
 - Berechnung der Ähnlichkeiten
- Sentence Transformer
 - Berechnung der Ähnlichkeiten zwischen den Fragen
- Naïve Bayes
 - Erkennung ob gesendete Nachricht eine Frage ist
 - Trainieren eines Klassifikators mit den Decks

Evaluierung der Modelle

- Keine “klassische” Evaluierung, da keine Labels vorhanden
- Rechenzeit
- Performance aus User-Sicht
- Verlässlichkeit
- Wie deutlich ist der Unterschied der Kosinus-Ähnlichkeit bei tatsächlich ähnlichen Fragen und sehr unterschiedlichen Fragen?

Herausforderungen

- Evaluierung der Modelle
 - Synonyme
 - Wechsel zwischen Deutschen und Englischen Begriffen
 - Hauptsächlich Verwendung von deutscher Sprache
 - tf-idf: tatsächliche ähnliche Fragen und sehr unterschiedliche Fragen haben ähnliche hohe Kosinus-Ähnlichkeiten
-
- Message Handler Python Telegram Bot

Weitere Implementierungsmöglichkeiten

- Automatische Erkennung, ob User Message eine Frage ist
- Statistik der Lernsession
- Rechtschreibkorrektur

Live Präsentation Chatbot



Vielen Dank für die
Aufmerksamkeit



Quellen

- <https://morioh.com/p/b78b4b571e33>
- <https://avatars.githubusercontent.com/u/65548303?s=200&v=4>