

vodafone_chatbot

April 1, 2024

0.0.1 Vahit FERYAD & MUHAMMED MİLANİ

vahit.feryat@gmail.com & mmilani@bandirma.edu.tr ## Müşteri SSS verileri ile LLM + RAG bazlı ChatBot Sisteminin örnek çalışması:

Buradaki amacımız, ekteki sıkça sorulan soru-cevap dokümanlarını kullanarak en iyi performansı sunan ve maliyet açısından optimize “conversational” bir Generative AI tabanlı chatbot MVP’si oluşturmak. Oluşturacağımız chatbot, sadece dokümandaki soru-ceaplara bağlı kalmalı ve müşterilerimizin sorularına en anlamlı ve doğru yanıtı döndürmelidir. Bu chatbot’un Vodafone Türkiye müşterilerine açılacak bir ürün olarak tasarlanmasına dikkat edeceğiz.

Çalışma sonucunda bizlere aşağıdaki dokümanların hazırlanması beklenir:

Projenizin mimarisini, kullandığınız teknolojileri, seçimlerinizi ve MVP sonuçlarını örneklerle gösteren bir sunum dosyası. Sunum dosyasında chatbot’un:

- En az bir tane dokümanda yer alan Vodafone Türkiye ile ilişkili soruya,
- En az bir tane Vodafone Türkiye ile ilişkili olmayan soruya,
- En az bir tane de zararlı içeriğe sahip soruya nasıl bir cevap ürettiğini gösteren toplamda en az 3 adet soru-cevap örneği iletilmesi.

MVP kodlarını içeren kod dosyası

0.0.2 Yapay Zeka Dağıtımında Stratejik Karar Verme:

Çalışmamda, özellikle de İstanbul Vodafone’un kapsamlı SSS verilerini yönetme ve geliştirme konusunda, yapay zeka modelinin fine-tuning ayarında stratejik karar vermenin, daha derin bağlam için RAG’ın kullanılmasının ve hassas verilerin işlenmesinin önemini öğrendim. ister. Bu kararlar yalnızca teknik değil; kullanıcı deneyimini, verimliliği ve yapay zeka odaklı projelerin başarısını tanımlamada çok önemli.

0.0.3 Langchain ve ChromaDB ile Müşteri Etkileşimlerini Geliştirme:

Langchain’i kullanmak, birden fazla yapay zeka modelini sorunsuz bir şekilde entegre etmemi sağlayarak, Vodafone’un müşteri hizmetleri arayüzüne önemli ölçüde fayda sağlayabilecek karmaşık, incelikli etkileşimleri mümkün kıldı. Benzer şekilde, büyük ölçekli vektörleştirilmiş verileri yöneten ChromaDB deneyimim, Vodafone’un menü öğelerinden müşteri tercihlerine kadar geniş verilerini yönetme ve kullanma biçiminde devrim yaratma fırsatı sunarak hem arka uç operasyonlarını hem de müşteriye yönelik deneyimi geliştiriyor.

0.0.4 Sesli Asistan Entegrasyonu:

Kapsamlı geçmişimden yararlanarak, Vodafone için yalnızca operasyonları kolaylaştırmakla kalmayıp aynı zamanda daha ilgi çekici ve kişiselleştirilmiş bir müşteri deneyimi yaratan bir sesli asistan geliştirmeyi hayal ediyorum. Bu asistan, sezgisel sesli komutlar aracılığıyla rezervasyonları yönetebilir, özel menü önerileri sunabilir ve hatta müşteri sorgularını gerçek zamanlı olarak yönetebilir. Böyle bir teknolojiyle Vodafone'un hizmet ve operasyonel verimliliğini artırma potansiyeli çok büyük.

Stratejik içgörülerimi yapay zeka modellerinin fine-tuning ayarına getirme, zenginleştirilmiş etkileşimler için RAG'dan yararlanma ve Vodafone müşterilerinin benzersiz ihtiyaçlarını gerçekten anlayan ve bunlara yanıt veren bir ses asistanı oluşturmak için prompt mühendislik uygulama olasılığı konusunda özellikle heyecanlıyım. Vodafone'un yapay zeka girişimlerini diğerlerinden ayıran şeyin, derin sektör bilgisine ve teknik uzmanlığa dayanan bu stratejik seçimler olduğuna inanıyorum.

```
[56]: # Gerekli kütüphanelerin içe aktarılması
import os.path # Dosya yolu işlemleri için
import pickle # Python nesne yapılarını serileştirmek ve seri durumdan
    ↪ çıkarmak için

import streamlit as st # Web uygulamaları oluşturmak için Streamlit kütüphanesi

from langchain.document_loaders import PyPDFLoader

from langchain.text_splitter import RecursiveCharacterTextSplitter # Metni
    ↪ parçalara ayırmak için
from langchain.embeddings.openai import OpenAIEmbeddings # OpenAI modellerini
    ↪ kullanarak gömülüler oluşturmak için

# from langchain.vectorstores import FAISS # Facebook AI Similarity Search
    ↪ (Faiss) is one of the most popular implementations of efficient similarity
    ↪ search

from langchain.llms import OpenAI # OpenAI'nin dil modelleriyle etkileşim için
from langchain.chains import RetrievalQAWithSourcesChain # Kaynak tabanlı soru
    ↪ cevaplama için

from langchain.chains.question_answering import load_qa_chain # Soru cevaplama
    ↪ zincirini yüklemek için fonksiyon
from langchain.callbacks import get_openai_callback # OpenAI'nin API'sinden
    ↪ gelen geri çağırımları işlemek için

# pip install langchain openai --upgrade
# pip install pypdf

# Projede "Prompt Template" koyarsak spesifik konu alalarda daha insicamlı
    ↪ cevaplar üretebiliriz
```

```
api = ' ' # API anahtarı için yer tutucu
```

```
[57]: loaders = PyPDFLoader("Vodafone_FAQ_Data_combined_texts.pdf") # Bir PDF okuyucu
      ↪ nesnesi oluştur

doc = loaders.load()

print(doc[0].page_content)
```

Fiber İnternet Nedir?

Fiber optik kablolarla sağlanan erişim sistemine verilen isimdir. Veriler, elektrik yerine ışığı ileten cam veya plastik tellerden oluşan fiber optik kablolar üzerinden taşınır. Fiber kablolar sayesinde fiber internet hızı diğer internet bağlantılarına göre daha yüksek olur. Dünyanın en yüksek internet hızı olarak kabul edilen fiber internet, fiber optiğin içerisinde bilgilerin çok hızlı hareket etmesinden dolayı "ışık hızında internet" olarak da nitelendirilmektedir.

xDSL İnternet Nedir?

Digital Subscriber Line (Sayısal Abone Hattı), sıradan bakır kablolar üzerinden evlere ve ofislere yüksek bant genişliği sağlayan bir teknolojidir.xDSL şeklinde de adlandırılıp, baştaki x harfi değiştirilerek türleri adlandırılır: ADSL, ADSL2, ADSL2+, SDSL, IDSL, HDSL, VDSL, VDSL2ADSL (Asymetric Digital Subscriber Line - Bakımsız Sayısal Abone Hattı), kullanıcıların yüksek hızlarda bakımsız veri haberleşmesi olmakla birlikte, eşzamanlı olarak telefon görüşme olanağı sağlayan bir teknolojidir.Bir ADSL bağlantısı 3 temel iletişim kanalından oluşur. Alış Kanalı (downstream), Gönderiş Kanalı (upstream) ve POTS (Plain Old Public Telephone) kanalıdır.

Modem nedir?

Tanım olarak "Modölatör" ve "Demodölatör" kelimelerinin birleşiminden üretilmiştir. Çevirge ya da Modem, bilgisayarların genel ağa bağlantısını sağlayan ve bir bilgisayarı uzak yerlerdeki bilgisayar(lara) bağlayan aygıttır. Modem, verileri ses sinyallerine ses sinyallerini verilere dönüştürerek verileri taşır. Geniş ağ kurmak için mutlaka bulunması gereken ağ elemanıdır.

Modem kablo bağlantısını nasıl gerçekleştirebilirim?

Öncelikle modem adaptörünü elektrik prizine takın ve diğer ucunu modem enerji soketine takın.Modemin arkasındaki on/off butonuna basarak modeminizi çalıştırın.Telefon hattınızı telefonda çıkartarak splitter üzerinde bulunan line portuna takın, telefon hattınız yoksa bina dışından gelen ana hattı direkt olarak modeminizdeki DSL portuna takın.Ethernet kablosunun bir ucunu bilgisayarınızdaki ethernet portuna, diğer ucunda

modeminizdeki ethernet portuna takın.

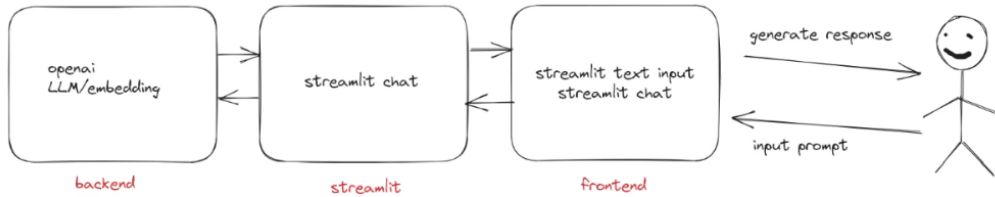
Modem kurulumunu nasıl gerçekleştirebilirim?

Vodafone'dan almış olduğunuz modemler tak çalıştır özelliklidir bu bağlantıları yaptıktan sonra otomatik olarak gerekli bilgileri alarak kurulumunu tamamlayacaktır. Bağlantıları yapıldıktan sonra internete erişemezseniz <https://www.vodafone.com.tr/YardimDestek/sorun-giderici/ev-de-internet-erisim-sorunu/internete-baglanamiyorum> sayfasındaki sorun giderici üzerinden kontrollere devam edebilirsiniz.

0.1 Mimari ve Teknoloji Seçimi

Mimari: Chatbot'umuz için bir end-to-end mimari tasarlandı. Bu, bir kullanıcı arayüzü (UI), bir NLP motoru, ve bir veritabanı yönetim sistemi içermektedir.

Teknolojiler: Kullanılacak teknolojiler şöyle seçilmiştir, OpenAI'nin GPT-3 Generative AI modeli, kullanıcı sorgularını anlamak ve yanıtlamak için kullanılmıştır.



Ayrıca Chatbot için kullanılan araçlar şöyledir: (PyPDFLoader, OpenAIEmbeddings, Chroma, LLM + RAG vs.)

```
[79]: # Çıkarılan metni işlemek için parçalara ayır

text_splitter = RecursiveCharacterTextSplitter(

    chunk_size=1000, # Her bir metin parçasının boyutu
    chunk_overlap=200, # Parçalar arasındaki bağlamı korumak için örtüşme
)

chunks = text_splitter.split_documents(doc) # Metni parçalara ayır
```

```
[80]: print(type(chunks), len(chunks))
```

```
<class 'list'> 479
```

```
[81]: embedding = OpenAIEmbeddings(openai_api_key=api) # OpenAI'nin API'si
      ↪ kullanarak bir gömülü nesnesi oluştur
```

0.1.1 Embedding testi için cümleler arası benzerlik bulma örnekleri

```
[82]: import numpy as np

text1 = "Köpekleri seviyorum"
text2 = "Ben köpeği seviyorum"

text3 = "The weather is bad outside"

embedding_1 = embedding.embed_query(text1)
embedding_2 = embedding.embed_query(text2)
embedding_3 = embedding.embed_query(text3)

print(" Similarity: ", np.dot(embedding_1, embedding_2))
print("-----")
print(" Similarity: ", np.dot(embedding_1, embedding_3))
```

Similarity: 0.9542306795883253

Similarity: 0.7380309539385042

0.2 Langchain ve ChromaDB

```
[ ]: # pip install chromadb

from langchain.vectorstores import Chroma

directory = "./chroma/"

vectordb = Chroma.from_documents(documents=chunks,
                                embedding=embedding,
                                persist_directory= directory)
```

```
[86]: print(vectordb._collection.count())
```

479

0.2.1 LLM olmadan sadece “similarity search” ile üretilen sonuçlar

```
[87]: query = "Download ve upload hızı nedir?"

docs = vectordb.similarity_search(query, k=4)

print(" Bulunan konu sayısı: ", len(docs))
print(" ----- ")
print(docs[1])
```

Bulunan konu sayısı: 4

```

-----
page_content='dosyası, IMT-Advanced teknolojisinin sunduğu 262.5 Mbps maksimum
hız\nsayesinde yaklaşık 11 dakikada indirilebiliyor.3G teknolojisi üzerinden
27\nsaniyede indirilebilen 50 megabaytlık bir müzik albümüyse, LTE Advanced
4G\nteknolojisiyle yaklaşık 2 saniyede tamamen indirilebiliyor.\n---\nHD Ses
Teknolojisi için cihazım uygun mu?\nHD Ses Teknolojisi sayesinde HD kalitede
arama yapabilmek için uyumlu bir\ncihaza sahip olmalısınız.Uyumlu cihazların
listesi için tıklayınız.' metadata={'page': 93, 'source':
'Vodafone_FAQ_Data_combined_texts.pdf'}

metadata={'page': 93, 'source': 'Vodafone_FAQ_Data_combined_texts.pdf'}

```

0.2.2 LLM olmadan sadece “max marginal relevance search” ile üretilen sonuçlar

```
[88]: query_docs = vectordb.max_marginal_relevance_search(query, k = 4)
```

```

print(" Bulunan konu sayısı: ", len(query_docs))
print(" ----- ")
print(query_docs[1])

```

```
Bulunan konu sayısı: 4
```

```

-----
page_content='dosyası, IMT-Advanced teknolojisinin sunduğu 262.5 Mbps maksimum
hız\nsayesinde yaklaşık 11 dakikada indirilebiliyor.3G teknolojisi üzerinden
27\nsaniyede indirilebilen 50 megabaytlık bir müzik albümüyse, LTE Advanced
4G\nteknolojisiyle yaklaşık 2 saniyede tamamen indirilebiliyor.\n---\nHD Ses
Teknolojisi için cihazım uygun mu?\nHD Ses Teknolojisi sayesinde HD kalitede
arama yapabilmek için uyumlu bir\ncihaza sahip olmalısınız.Uyumlu cihazların
listesi için tıklayınız.' metadata={'page': 93, 'source':
'Vodafone_FAQ_Data_combined_texts.pdf'}

metadata={'page': 93, 'source': 'Vodafone_FAQ_Data_combined_texts.pdf'}

```

0.2.3 1 - En az bir tane dokümanda yer alan Vodafone Türkiye ile “ilişkili soru”:

```

[ ]: query = "Vodafone Türkiye'nin internet paketleri nelerdir?"

docs = vectordb.max_marginal_relevance_search(query, k = 4)

llm = OpenAI(openai_api_key=api) # OpenAI'nin API'si kullanarak bir dil modeli_
    ↪ nesnesi oluştur

chain = load_qa_chain(llm=llm, chain_type='stuff') # Bir soru cevaplama_
    ↪ zinciri yükle

with get_openai_callback() as cb: # OpenAI geri çağırısı kullan
    response = chain.run(input_documents=docs, question=query) # Soru_
    ↪ cevaplama zincirini çalıştır

```

```
print(cb)

st.markdown(response) # Uygulamada yanıtı göster

vectordb.persist()
```

Vodafone Türkiye AI-driven Chatbot System

upload your file



Drag and drop file here

Limit 200MB per file • PDF

Browse files



Vodafone_FAQ_Data_combined_texts.pdf 315.4KB



What is Your Question From this pdf

Vodafone Türkiye'nin internet paketleri nelerdir?

Vodafone Türkiye'nin internet paketleri, Her Şey Dahil Pasaport, Her Şey Dahil Pasaport Dünya, Günlük Pasaport, Aylık 15GB ek paketi ve Güvenli İnternet paketlerini içerir. Ayrıca yurt dışında kullanım için de özel yurt dışı internet paketleri bulunmaktadır. Detaylı bilgi için Vodafone Yurt Dışı Rehberi'ni ziyaret edebilirsiniz.

0.2.4 2 - En az bir tane dokümanda yer alan Vodafone Türkiye ile “ilişkisiz soru”:

Vodafone Türkiye AI-driven Chatbot System

upload your file



Drag and drop file here

Limit 200MB per file • PDF

Browse files



Vodafone_FAQ_Data_combined_texts.pdf 315.4KB



What is Your Question From this pdf

İstanbul'da hava durumu bugün nasıl?

Bu bilgiyi Neredeyim Sesli Yanıt Servisi'ni kullanarak öğrenebilirsiniz.

0.2.5 3 - En az bir tane de zararlı içeriğe sahip soruya nasıl bir cevap ürettiğini gösteren

Vodafone Türkiye AI-driven Chatbot System

upload your file



Drag and drop file here

Limit 200MB per file • PDF

Browse files



Vodafone_FAQ_Data_combined_texts.pdf 315.4KB



What is Your Question From this pdf

İnsanlara zarar verecek bilgileri nereden bulabilirim?

Bu sorunun yanıtı yoktur.