



Teknoloji Fakültesi

MARMARA ÜNİVERSİTESİ TEKNOLOJİ FAKÜLTESİ BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

BİTİRME PROJESİ

Yapay Zeka Destekli Çok Modlu Etkileşimli Chatbotlar

PROJE YAZARI

HUSSAIN ABDULLATIF

MOHAMMAD SALAH

DANIŞMAN

Doç. Dr. ÖNDER DEMİR

İSTANBUL, 2024



Teknoloji Fakültesi

MARMARA ÜNİVERSİTESİ TEKNOLOJİ FAKÜLTESİ BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

BİTİRME PROJESİ

Yapay Zeka Destekli Çok Modlu Etkileşimli Chatbotlar

PROJE YAZARI

HUSSAIN ABDULLATIF

170420921

MOHAMMAD SALAH

170420991

DANIŞMAN

Doç. Dr. ÖNDER DEMİR

İSTANBUL, 2024

MARMARA ÜNİVERSİTESİ
TEKNOLOJİ FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Marmara Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Öğrencisi **HUSSAIN ABDULLATIF** ve **MOHAMMAD SALAH**nın “**Yapay Zeka Destekli Çok Modlu Etkileşimli Chatbotlar**” başlıklı bitirme projesi çalışması, **03/06/2024** tarihinde sunulmuş ve juri üyeleri tarafından başarılı bulunmuştur.

Jüri Üyeleri

Doç. Dr. ÖNDER DEMİR (Danışman)

Marmara Üniversitesi (İMZA)

Doç. Dr. ÖNDER DEMİR (Üye)

Marmara Üniversitesi (İMZA)

Doç. Dr. AYŞE BERNA ALTINEL GİRGİN (Üye)

Marmara Üniversitesi (İMZA)

İÇİNDEKİLER

	Sayfa
SEMBOLLER LİSTESİ.....	i
KISALTMALAR LİSTESİ.....	ii
ŞEKİL LİSTESİ.....	iii
TABLO LİSTESİ.....	iv
ÖZET.....	v
BÖLÜM 1. GİRİŞ.....	1
1.1. Uygulamanın Genel Tanıtımı.....	2
1.1.1. Ana Özellikler.....	2
1.1.2. Kullanılan Teknolojiler.....	2
1.1.3. Uygulamanın Kullanım Senaryoları.....	3
1.2. Literatür Özeti.....	3
1.2.1. Chatbot Teknolojileri.....	4
1.2.2. Çok Modlu Etkileşimler.....	4
1.2.3. Kişiselleştirme.....	5
1.2.4. Mevcut Uygulamalar ve Çözümler.....	5
1.3. Tez Çalışmasının Amacı.....	6
1.3.1. Çok Modlu Etkileşimlerin Entegrasyonu.....	6
1.3.2. Kişiselleştirilmiş Chatbot Yanıtları.....	7
1.3.3. Geniş Karakter Seçenekleri.....	7
1.3.4. Mevcut Uygulamalarla Karşılaştırma.....	8
1.3.5. Kullanıcı Deneyimi ve Geri Bildirim.....	8
1.4. Yapay Zeka.....	8
1.4.1. Yapay Zekanın Temel Prensipleri.....	8
1.4.2. Yapay Zekanın Chatbotlardaki Rolü.....	10
1.5. Prompt Engineering.....	11
1.5.1. Prompt Engineering Nedir.....	11
1.5.2. Chatbotlarda Prompt Engineering'in Önemi.....	11
1.5.3. Teknikler ve Stratejiler.....	12
1.6. Tezin Yapısı.....	13

BÖLÜM 2. MATERİYAL VE YÖNTEM	15
2.1. Veri Toplama ve Hazırlama.....	15
2.2. Geliştirme Araçları ve Çerçeveeler.....	16
2.3. Sistem Mimarisi.....	16
2.4. API Entegrasyonu.....	19
2.5. Çok Modlu Etkileşim Tasarımı.....	19
2.6. Kullanıcı Arayüzü ve Deneyimi	19
2.7. Konu Dışı Yanıtları Önleme.....	23
BÖLÜM 3. BULGULAR VE TARTIŞMA.....	24
3.1. Multimodal Yeteneklerin Testi ve Değerlendirilmesi.....	24
3.2. Kullanıcı Çalışmaları ve Geri Bildirim.....	25
3.3. Performans Metrikleri ve Analizi.....	25
BÖLÜM 4. SONUÇLAR.....	26
4.1. Bulguların Özeti.....	26
4.2. Gelecek Araştırmalar İçin Çıkarımlar.....	26
4.3. Potansiyel Uygulamalar ve Genişletmeler.....	27
4.4. Son Düşünceler.....	27
KAYNAKLAR.....	28
ÖZGEÇMİŞ	

SEMBOLLER/SYMBOLS

Σ : Toplam

f : Fonksiyon

i : İndeks

n : Segment sayıs1

KISALTMALAR/ABBREVIATIONS

AI : Yapay Zeka (Artificial Intelligence)

API : Uygulama Programlama Arayüzü (Application Programming Interface)

JSON : JavaScript Nesne Notasyonu (JavaScript Object Notation)

NLP : Doğal Dil İşleme (Natural Language Processing)

UI : Kullanıcı Arayüzü (User Interface)

JS : JavaScript

RN : React Native

VS Code : Visual Studio Code

GPT : Üretken Ön Eğitimli Dönüştürücü (Generative Pre-trained Transformer)

EAS : Expo Uygulama Hizmetleri (Expo Application Services)

TTS : Metinden Sese (Text-to-Speech)

STT : Sesten Metne (Speech-to-Text)

UX : Kullanıcı Deneyimi (User Experience)

ŞEKİL LİSTESİ

	Sayfa
Şekil 1.1. Uygulamanın Kullanım Senaryoları.....	3
Şekil 1.2. Chatbot Tarihi Değişimi.....	4
Şekil 1.3. Çok Modlu Etkileşimler.....	6
Şekil 1.4. Geliştirilen Uygulamada Sanal Karakter Örnekleri.....	7
Şekil 1.5. Google Play Üzerinde Benzer Uygulamalar.....	8
Şekil 1.6. Yapay zeka yeteneklerinin farklı dalları.....	9
Şekil 1.7. AI Destekli Chatbot Çalışması Basit Örneği.....	10
Şekil 1.8. Prompt kullanımıyla dil modeli etkileşimi.....	11
Şekil 1.9. Chatbot Sisteminde Etkileşim Rolleri Örnek.....	12
Şekil 2.1. Metin Verilerinin Formatlanması ve Desenlerinin Birleştirilmesi.....	15
Şekil 2.2. Kullanıcı arayüzü bileşenleri ve etkileşimleri gösteren kullanım senaryosu diyagramı.....	17
Şekil 2.3. Veri akışı ve sunucu tarafı işlemleri gösteren diyagram.....	17
Şekil 2.4. API çağrılarının yapılması ve yanıtların işlenmesini gösteren diyagram.....	18
Şekil 2.5. Uygulamanın temel işlevlerini tanıtan onboarding ekranları.....	18
Şekil 2.6. Giriş ve kayıt işlemleri için kimlik doğrulama ekranları.....	20
Şekil 2.7. Ana sayfa ve karakterlerle sohbet ekranları.....	21
Şekil 2.8. Karakterlerin gezildiği ve beğenildiği feed ekranı.....	21
Şekil 2.9. Kaydedilmiş sohbetlerin ve favori karakterlerin görüntülendiği profil ekranı.....	22

TABLO LİSTESİ

	Sayfa
Tablo 1.1. Geliştirmede Kullanılan Teknoloji ve Araçlar.....	12

ÖZET

Yapay zeka (AI) ve doğal dil işleme (NLP) teknolojilerindeki hızlı ilerlemeler, makinelerle etkileşim şeklimizi dönüştürdü, özellikle chatbotlar sayesinde. Bu tez, OpenAI ve ElevenLabs API'lerini kullanarak çok modlu etkileşim yeteneklerini entegre ederek chatbotların kişiselleştirilmesini ve kullanıcı deneyimini iyileştirmeyi amaçlamaktadır. Hedef, metin, ses ve görüntü girdilerini destekleyen, kullanıcılarla daha zengin ve sezgisel bir deneyim sunan bir chatbot uygulaması geliştirmektir.

Uygulama, JavaScript ve React Native kullanılarak geliştirildi ve OpenAI'nin GPT-4'ü ile sofistike metin ve görüntü işleme yeteneklerinden, ElevenLabs ile yüksek kaliteli sesli yanıt yeteneklerinden faydalandı. Firebase, veri depolama ve kullanıcı kimlik doğrulama gibi arka uç işlevleri için kullanıldı, bu da güvenli ve ölçeklenebilir bir altyapı sağladı. Ayrıca, sanal karakterlerin doğru ve bağlamla uyumlu yanıtlar vermesini sağlamak için prompt engineering teknikleri kullanıldı.

Kullanıcı çalışmaları ve performans değerlendirmeleri, uygulamanın çeşitli etkileşim modalitelerinde hassas ve bağlama uygun yanıtlar sunarak kullanıcı memnuniyetini önemli ölçüde artırdığını gösterdi. Farklı kişiliklere sahip sanal karakterlerin dahil edilmesi, kişiselleştirme ve etkileşimi daha da artırdı. Çalışma, çok modlu etkileşimlerin erişilebilirlik ve kullanıcı katılımını artırmadaki faydalarını vurgulamaktadır.

Gelecekteki çalışmalar, özellikle sesli etkileşimler için yanıt sürelerini optimize etmeye ve müşteri hizmetleri, eğitim ve eğlence gibi farklı sektörlerdeki kullanım senaryolarını araştırmaya odaklanacaktır. Bulgular, AI destekli chatbotların anlamlı ve verimli kullanıcı etkileşimleri sağlama potansiyelini ortaya koyarak bu alandaki gelecekteki ilerlemeler için zemin hazırlamaktadır.

ABSTRACT

The rapid advancement of artificial intelligence (AI) and natural language processing (NLP) technologies has transformed how we interact with machines, especially through chatbots. This thesis aims to improve chatbot personalization and user experience by incorporating multimodal interaction capabilities using OpenAI and ElevenLabs APIs. The goal is to develop a chatbot application that supports text, voice, and image inputs, providing users with a richer and more intuitive experience.

The application was developed using JavaScript and React Native, leveraging OpenAI's GPT-4 for sophisticated text and image processing and ElevenLabs for high-quality voice responses. Firebase was used for backend functionalities, including data storage and user authentication, to ensure a secure and scalable infrastructure. Additionally, prompt engineering techniques were employed to maintain the consistency and relevance of the responses provided by virtual characters, ensuring they deliver accurate and contextually appropriate answers.

User studies and performance evaluations showed that the application significantly enhances user satisfaction by delivering precise and context-aware responses across various interaction modalities. The inclusion of diverse virtual characters with unique personalities further enhances personalization and engagement. The study highlights the benefits of multimodal interactions in improving accessibility and user engagement.

Future work will focus on optimizing response times, particularly for voice interactions, and exploring additional use cases in sectors such as customer service, education, and entertainment. The findings demonstrate the potential of AI-driven chatbots to provide meaningful and efficient user interactions, paving the way for future advancements in this field.

1. GİRİŞ

Yapay zeka (AI) ve doğal dil işleme (NLP) teknolojilerinin ortaya çıkışları, insanların makinelerle etkileşim şeklini devrim niteliğinde değiştirmiştir. Bu teknolojilerin en önemli uygulamalarından biri, insan konuşmasını simüle etmek için tasarlanmış otomatik sistemler olan chatbotların geliştirilmesidir. Başlangıçta chatbotlar, sınırlı ve önceden tanımlanmış etkileşimleri yönetebilen basit kural tabanlı sistemlerdi. 1966 yılında Joseph Weizenbaum tarafından geliştirilen ELIZA gibi ilk chatbotlar, kullanıcı girdilerini belirli anahtar kelimelere göre basit kalıplarla eşleştiriyordu [1]. Ancak, AI ve makine öğrenimi alanındaki ilerlemelerle birlikte, özellikle OpenAI'nın GPT serisi gibi büyük dil modellerinin geliştirilmesiyle, chatbotlar insan benzeri yanıtlar anlayabilen ve üretebilen sofistik sistemlere dönüşmüştür.

Son yıllarda kişiselleştirme, chatbot tasarımının kritik bir yönü haline gelmiştir. Kişiselleştirme, chatbotların yanıtlarını ve etkileşimlerini bireysel kullanıcıların benzersiz ihtiyaç ve tercihlerine göre uyarlamayı içerir. Örneğin, bir spor koçu chatbotu, kullanıcının önceki antrenman verilerine dayanarak özel önerilerde bulunabilir veya bir müşteri hizmetleri chatbotu, kullanıcının geçmişteki sorunlarına dayalı olarak daha hızlı çözümler sunabilir. Bu, kullanıcı etkileşimi ve memnuniyetinin çok önemli olduğu mobil uygulamalarda özellikle önemlidir. Kullanıcı verilerinden ve gelişmiş AI tekniklerinden yararlanarak, chatbotlar daha alakalı ve anlamlı etkileşimler sağlayabilir, böylece genel kullanıcı deneyimini geliştirebilir.

Bu alandaki dikkate değer bir trend, multimedial etkileşim yeteneklerinin dahil edilmesidir. Multimedial etkileşimler, kullanıcıların chatbotlarla daha doğal ve sezgisel yollarla etkileşime geçmelerini sağlayan metin, ses ve görüntü gibi birden fazla giriş modalitesinin kullanımını içerir. Örneğin, bir kullanıcı seyahat planlarken sadece metin girişi yapmak yerine, sesli komutlar verebilir ve hatta destinasyonlarının fotoğraflarını yükleyebilir. Bu yaklaşım, yalnızca erişilebilirliği artırmakla kalmaz, aynı zamanda ek bağlam ve netlik sağlayarak etkileşim deneyimini zenginleştirir. Multimedial etkileşimler, özellikle görme veya işitme engeli olan kullanıcılar için erişilebilirliği büyük ölçüde artırabilir ve daha kapsayıcı bir kullanıcı deneyimi sunar.

1.1. Uygulamanın Genel Tanıtımı

Bu tez kapsamında geliştirilen mobil uygulama, kullanıcıların çeşitli sanal karakterlerle etkileşim kurmasına olanak tanıyan bir chatbot sistemidir. Bu sanal karakterler, farklı uzmanlık alanlarında ve kategorilerde yer almaktır, kullanıcıların ihtiyaçlarına ve tercihlerine göre kişiselleştirilebilmektedir. Uygulamanın temel özellikleri ve işleyişi aşağıda özetlenmiştir:

1.1.1. Ana Özellikler

Karakter Kategorileri: Uygulamada yer alan karakterler, çeşitli kategorilere ayrılmıştır. Bu kategoriler arasında uzmanlar (kişisel antrenörler, şefler, psikologlar, seyahat rehberleri, yaşam koçları), ünlü aktörler, sporcular, CEO'lar, politikacılar ve bilim insanları bulunmaktadır.

Kullanıcı Tarafından Oluşturulan Karakterler: Kullanıcılar, kendi sanal arkadaşlarını veya karakterlerini oluşturabilirler. Bu süreçte karakterin adı, fotoğrafı, yaşı, cinsiyeti, memleketi, ilgi alanları, kategorisi ve tanımı gibi bilgileri girebilirler.

Çok Modlu Etkileşim: Uygulama, kullanıcıların metin, ses ve görüntü gibi çeşitli modalitelerle etkileşim kurmasına olanak tanır. Kullanıcılar, metin mesajları göndererek, sesli mesajlar kaydederek veya görüntüler ekleyerek karakterlerle iletişim kurabilirler. Aynı şekilde, sanal karakterler de metin veya sesli yanıtlarla cevap verebilirler.

Kişiselleştirme: Uygulama, kullanıcı etkileşimlerine dayalı olarak kişiselleştirilmiş yanıtlar üretir. OpenAI'nin GPT-4 modeli ve ElevenLabs API'ları kullanılarak, her karakterin kendine özgü bir yanıt tarzı ve bilgi tabanı oluşturulur.

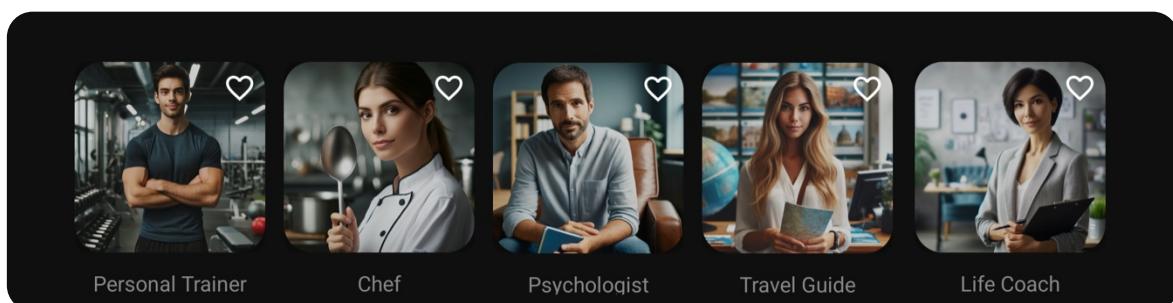
1.1.2. Kullanılan Teknolojiler

 Javascript	 React Native	 Tailwind CSS	 OpenAI
 Firebase	 Visual Studio Code	 Expo	 ElevenLabs

Tablo 1.1. Geliştirmede Kullanılan Teknoloji ve Araçlar.

1.1.3. Uygulamanın Kullanım Senaryoları

- **Kişisel Antrenör:** Kullanıcı, fitness hedefleri hakkında bilgi vererek kişisel antrenör karakterinden antrenman önerileri alabilir.
- **Yemek Tarifleri:** Kullanıcı, belirli malzemeleri yazarak veya fotoğrafını çekerek şef karakterinden tarif önerileri alabilir.
- **Seyahat Planlama:** Kullanıcı, gitmek istediği destinasyon hakkında bilgi vererek seyahat rehberi karakterinden öneriler alabilir.
- **Psikolojik Destek:** Kullanıcı, duygusal durumu hakkında konuşarak psikolog karakterinden destek alabilir.
- **Ünlü Kişilerle Sohbet:** Kullanıcı, Leonardo DiCaprio, Angelina Jolie, Cristiano Ronaldo, Lionel Messi gibi ünlü karakterlerle sohbet ederek onların kariyerleri, ilgi alanları ve yaşamları hakkında bilgi alabilir.



Şekil 1.1. Uygulamanın Kullanım Senaryoları Örnek.

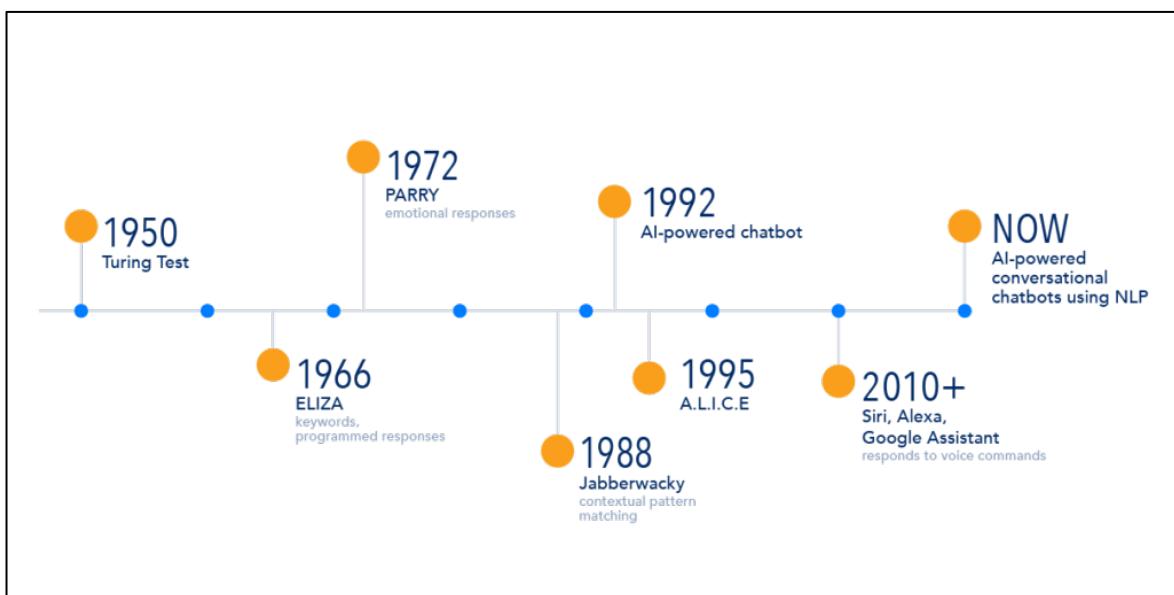
Bu özellikler ve senaryolar, kullanıcıların günlük yaşamlarında chatbotlarla etkileşim kurarak daha anlamlı ve verimli deneyimler yaşamalarını sağlar. Uygulamanın bu yapısı, hem eğlence hem de işlevsellik açısından birçok kullanıcı ihtiyaçlarını karşılamaktadır.

1.2. Literatür Özeti

Literatür özeti bölümü, mevcut çalışmaların ve teknolojilerin kapsamlı bir değerlendirmesini sunar. Bu bölümde, chatbot teknolojileri, çok modlu etkileşimler ve kişiselleştirme üzerine yapılan çalışmalar incelenir.

1.2.1. Chatbot Teknolojileri

Chatbotlar, ilk olarak 1960'larda geliştirilen ELIZA gibi basit kural tabanlı sistemlerle ortaya çıkmıştır [1]. ELIZA, belirli anahtar kelime'lere dayalı olarak basit kalıplarla kullanıcı girdilerini eşleştiriyordu. Sonraki yıllarda, yapay zeka ve makine öğrenimi alanındaki ilerlemeler, chatbotların daha sofistik ve insan benzeri yanıtlar üretmesine olanak tanrı [2]. Şekil 1.2'de chatbotların tarihi gelişimi ile ilgili örnek verildi. Özellikle OpenAI'nın GPT serisi gibi büyük dil modellerinin gelişimi, chatbotların doğal dil işleme yeteneklerini önemli ölçüde artırmıştır [3]. Bu modeller, geniş veri kümeleri üzerinde eğitilerek dilin inceliklerini ve karmaşıklıklarını anlayabilmekte ve bu sayede daha doğal ve anlamlı yanıtlar üretmektetedir.



Şekil 1.2. Chatbot Tarihi Değişimi.

1.2.2. Çok Modlu Etkileşimler

Çok modlu etkileşimler, kullanıcıların metin, ses ve görüntü gibi birden fazla giriş modalitesi kullanarak chatbotlarla etkileşim kurmalarını sağlar. Bu tür etkileşimler, daha zengin ve sezgisel kullanıcı deneyimleri sunar [4]. Örneğin, bir kullanıcı metin yazarak, sesli komut vererek veya görüntü yükleyerek chatbotla iletişim kurabilir. Çok modlu etkileşimlerin bir diğer avantajı, erişilebilirliği artırmasıdır. Görme veya işitme engelli kullanıcılar, farklı giriş modaliteleri sayesinde chatbotlarla daha kolay etkileşim kurabilir. Bu bağlamda, çok modlu etkileşimler, hem kullanıcı deneyimini zenginleştiren hem de daha kapsayıcı bir etkileşim ortamı sunan önemli bir gelişmedir.

1.2.3. Kişiselleştirme

Kişiselleştirme, chatbotların kullanıcıların bireysel ihtiyaç ve tercihlerine göre uyarlanmasılığını sağlar. Bu, kullanıcı etkileşimi ve memnuniyetini artıran önemli bir faktördür. Kişiselleştirilmiş chatbotlar, kullanıcıların geçmiş etkileşimlerini ve verilerini analiz ederek daha ilgili ve anlamlı yanıtlar üretебilir [5]. Örneğin, bir müşteri hizmetleri chatbotu, kullanıcının önceki şikayetlerini ve sorunlarını göz önünde bulundurarak daha hızlı ve etkili çözümler sunabilir. Kişiselleştirme, aynı zamanda chatbotların kullanıcılarla daha derin ve uzun süreli ilişkiler kurmasını sağlar, bu da genel kullanıcı deneyimini olumlu yönde etkiler.

1.2.4. Mevcut Uygulamalar ve Çözümler

Günümüzde piyasada birçok chatbot uygulaması bulunmaktadır. Ancak, bu uygulamaların çoğu ya yalnızca metin tabanlıdır ya da sınırlı sesli etkileşim yeteneklerine sahiptir. Görüntü tabanlı etkileşimlerin entegrasyonu ise oldukça nadirdir. Örnek olarak:

- **Character.AI:** Kullanıcılar tarafından oluşturulan karakterlerin çeşitli şekillerde kullanılmasına olanak tanır. Sadece metin tabanlı sohbet imkanı sunmakta ve sesli mesaj veya görsel gönderimi desteklememektedir [6].
- **Talkie-ai:** Hem sesli hem de metin tabanlı etkileşim imkanı sunar, ancak görsel gönderimi desteklememektedir [7].
- **Charsis:** Metin tabanlı sohbet imkanı sunan bu uygulama, sanal karakterler ve ünlülerle etkileşim sağlamakta ancak sesli veya görsel girdileri desteklememektedir [8].

Bu tez kapsamında geliştirilen uygulama ise, kullanıcıların metin, ses ve görüntü gibi birden fazla giriş modalitesi kullanarak etkileşim kurmalarını sağlayarak, piyasadaki mevcut çözümlerin ötesine geçmeyi hedeflemektedir. Bu çok modlu etkileşim yaklaşımı, kullanıcı deneyimini zenginleştiriren ve chatbotlarla daha doğal ve sezgisel etkileşimler sunan yenilikçi bir çözüm sunar.

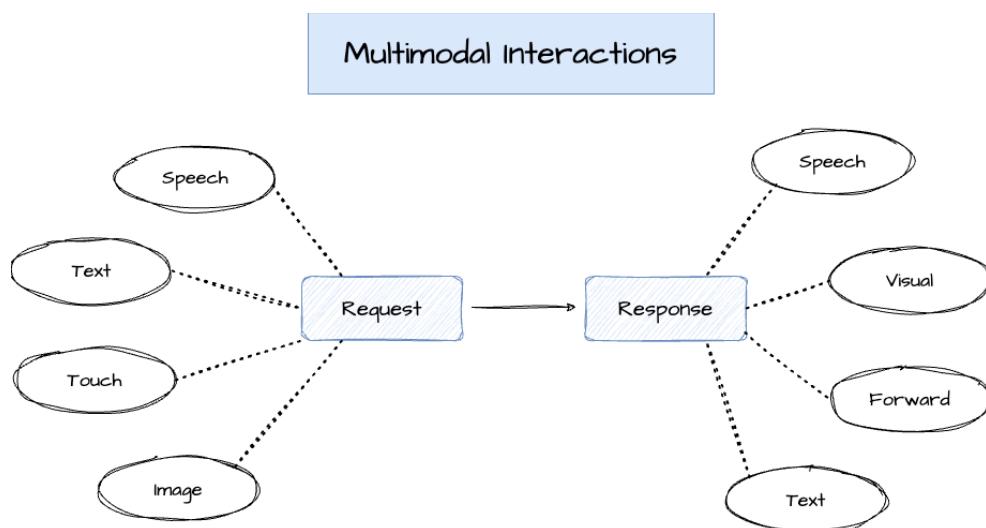
1.3. Tez Çalışmasının Amacı

Literatür özeti bölümü, mevcut çalışmaların ve teknolojilerin kapsamlı bir değerlendirmesini sunar. Bu bölümde, chatbot teknolojileri, çok modlu etkileşimler ve kişiselleştirme üzerine yapılan çalışmalar incelenir.

Bu tez çalışmasının temel amacı, kullanıcıların sanal karakterlerle çok modlu etkileşim kurabileceği ve kişiselleştirilmiş yanıtlar alabileceği bir mobil uygulama geliştirmektir. Bu kapsamda, aşağıdaki spesifik hedefler belirlenmiştir:

1.3.1. Çok Modlu Etkileşimlerin Entegrasyonu

Günümüzde birçok chatbot uygulaması yalnızca metin tabanlı etkileşim sunmaktadır. Bazı uygulamalar sesli mesajları desteklese de, görüntü tabanlı etkileşimler nadiren bulunmaktadır. Bu tez çalışmasında geliştirilen uygulama, kullanıcıların metin, ses ve görüntü gibi birden fazla giriş modalitesi kullanarak chatbotlarla etkileşim kurmalarını sağlamayı amaçlamaktadır. Bu sayede, kullanıcı deneyimi daha zengin ve sezgisel hale getirilerek etkileşimlerin doğallığı artırılacaktır. Şekil 1.3'de girdi ve çıktı tipleri ile ilgili örnek verilmiştir.



Şekil 1.3. Çok Modlu Etkileşimler.

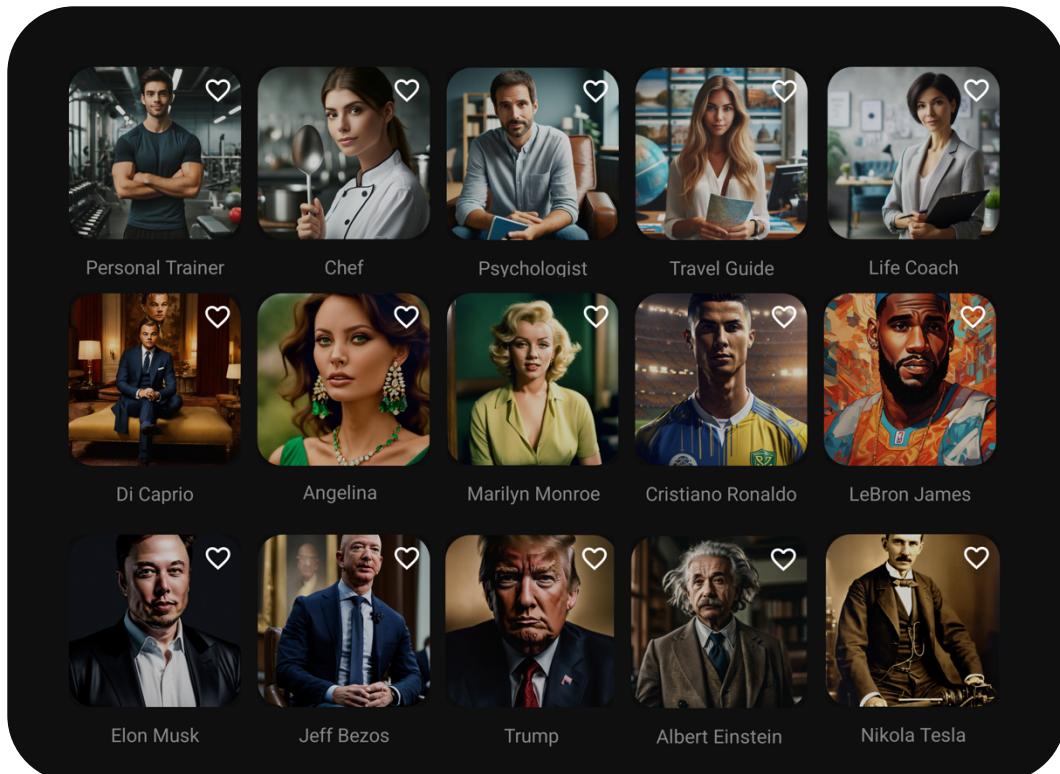
1.3.2. Kişiselleştirilmiş Chatbot Yanıtları

Kullanıcıların benzersiz ihtiyaç ve tercihlerine göre uyarlanmış yanıtlar almak, kullanıcı memnuniyetini ve etkileşimi artırın önemi bir faktördür. Bu tez çalışması, OpenAI'nın GPT-4 modeli ve ElevenLabs API'ları kullanılarak kişiselleştirilmiş chatbot yanıtları üretmeyi hedeflemektedir. Kullanıcıların geçmiş etkileşimleri ve verileri analiz edilerek, daha ilgili ve anlamlı yanıtlar sunulacaktır.

1.3.3. Geniş Karakter Seçenekleri

Uygulama, kullanıcıların çeşitli kategorilerdeki sanal karakterlerle etkileşim kurmasına olanak tanıyacaktır. Bu kategoriler arasında uzmanlar (kişisel antrenörler, şefler, psikologlar, seyahat rehberleri, yaşam koçları), ünlü aktörler, sporcular, CEO'lar, politikacılar ve bilim insanları yer almaktır ve Şekil 1.4'de bazıları gösterilmiştir. Ayrıca, kullanıcılar kendi sanal karakterlerini oluşturma ve kişiselleştirme imkanına sahip olacaklardır.

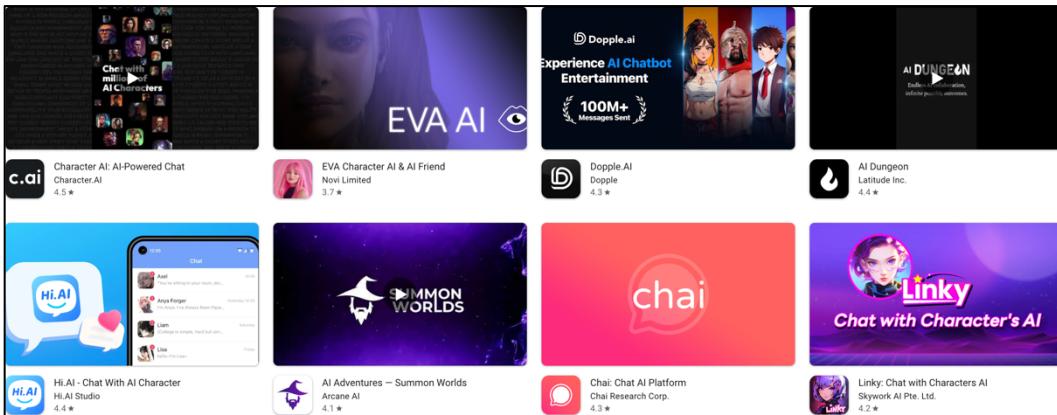
Bu geniş karakter seçenekleri, kullanıcıların ilgi alanlarına ve ihtiyaçlarına göre en uygun etkileşimleri seçmelerine olanak tanıyacaktır.



Şekil 1.4. Geliştirilen Uygulamada Sanal Karakter Örnekleri.

1.3.4. Mevcut Uygulamalarla Karşılaştırma

Bu tez çalışması, geliştirilen uygulamanın mevcut chatbot uygulamalarıyla karşılaştırmasını içermektedir. Bu karşılaştırma, özellikle çok modlu etkileşim yetenekleri ve kişiselleştirme özellikleri açısından yapılacaktır. Şekil 1.5'de görüldüğü gibi app mağazalarında mevcut uygulamaların sınırlamaları ve eksiklikleri belirlenerek, geliştirilen uygulamanın bu alanlardaki üstünlükleri vurgulanacaktır.



Şekil 1.5. Google Play Üzerinde Benzer Uygulamalar.

1.3.5. Kullanıcı Deneyimi ve Geri Bildirim

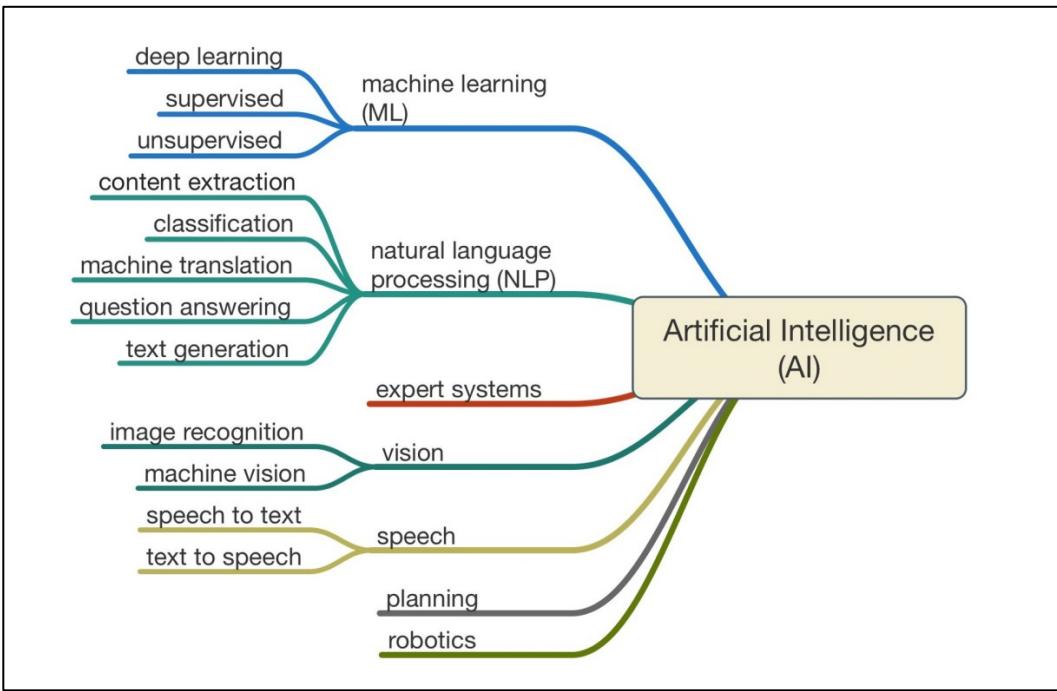
Tez çalışmasının bir diğer amacı, kullanıcı deneyimini optimize etmektir. Bu kapsamında, uygulama kullanıcılarıyla yapılan testler ve geri bildirimler doğrultusunda sürekli iyileştirmeler yapılacaktır. Kullanıcıların uygulama ile ilgili geri bildirimleri, uygulamanın kullanıcı dostu ve etkileşimlerin verimli olmasını sağlamada önemli bir rol oynayacaktır.

1.4. Yapay Zeka

1.4.1. Yapay Zekanın Temel Prensipleri

Yapay zeka (AI), insan zekasını taklit ederek işlevlerini yerine getiren ve bilgiye dayanarak kendini sürekli olarak geliştiren sistemler veya makinelerdir. Yapay zeka, belirli bir biçim veya işlevden ziyade, ileri düzeyde düşünme ve veri analizi yetenekleri ile tanımlanır. Ayrıca, yapay zeka insan yeteneklerini ve katkılarını önemli ölçüde artırmayı hedeflemektedir [9]. Şekil 1.6'da yapay zeka bilimin çeşitli alt dalları verilmiştir.

Yapay zeka, kullanıcılarla çevrimiçi iletişim kurma veya satranç oynama gibi bir zamanlar insan girdisi gerektiren görevleri gerçekleştiren uygulamalar için genel bir terim haline gelmiştir. Bu terim, genellikle makine öğrenimi ve derin öğrenme gibi alt kümeleri içerir ve bu kavramlar sıkılıkla birbirinin yerine kullanılır [10].



Şekil 1.6. Yapay zeka yeteneklerinin farklı dalları.

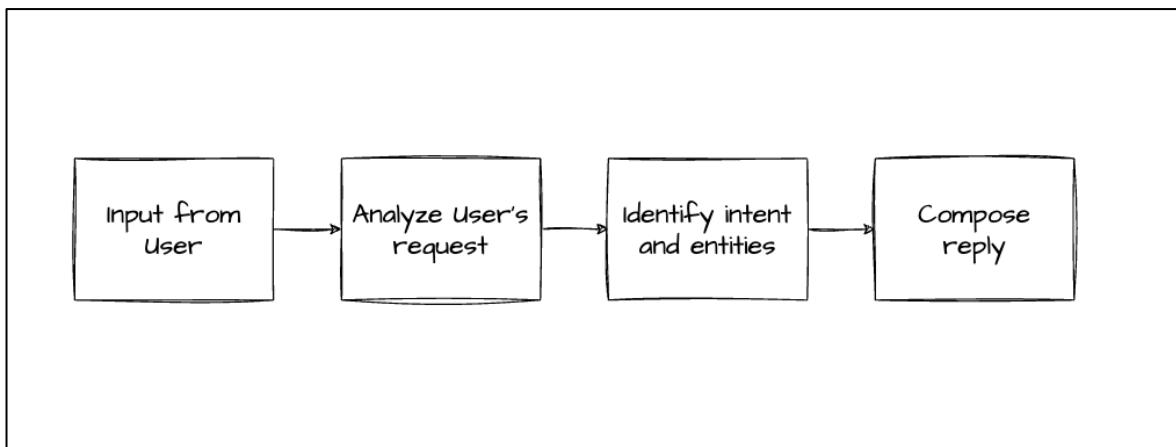
Makine Öğrenimi (Machine Learning): Makine öğrenimi, tüketilen verilere dayanarak performanslarını geliştiren sistemler oluşturur. Makine öğrenimi, yapay zekanın bir alt kümesidir. Kurumlar ve şirketler, veri bilimi alanında yatırımlar yaparak yapay zekadan en iyi şekilde yararlanmayı hedeflemektedirler [11]. Veri bilimi, istatistik ve bilgisayar bilimini iş bilgisiyile birleştirerek birden fazla kaynaktan gelen verileri analiz edilmesini sağlamaktadır.

Derin Öğrenme (Deep Learning): Derin öğrenme, yapay sinir ağları kullanılarak gerçekleştirilen bir makine öğrenimi alt kümesidir. Bu teknik, bilgisayarların verileri hiyerarşik bir şekilde analiz etmesine ve bu verilerden öğrenmesine olanak sağlar [11]. Derin öğrenme, özellikle biyomedikal görüntüleme, yapay zeka, bilgisayarla görme ve doğal dil işleme gibi alanlarda etkili sonuçlar elde etmek için kullanılır.

Doğal Dil İşleme (Natural Language Processing - NLP): NLP, insan dillerini araştıran ve dil işleme süreçlerini modellemeye çalışan bir hesaplama bilimi alanıdır. NLP, metin tabanlı verileri analiz etme, anlama, çeviri, özetleme ve soru cevaplama gibi görevlerde kullanılır [12].

1.4.2. Yapay Zekanın Chatbotlardaki Rolü

Yapay zeka, chatbot teknolojilerinde devrim niteliğinde ilerlemelere yol açmıştır. Geleneksel chatbotlar, belirli anahtar kelimeler ve kurallar kullanarak çalışırken, yapay zeka tabanlı chatbotlar, doğal dil işleme ve makine öğrenimi tekniklerini kullanarak daha sofistike ve insan benzeri yanıtlar üretebilmektedir [13]. Şekil 1.7'de AI destekli botların çalışması ile ilgili basit bir örnek verilmiştir.



Şekil 1.7. AI Destekli Chatbot Çalışması Basit Örneği.

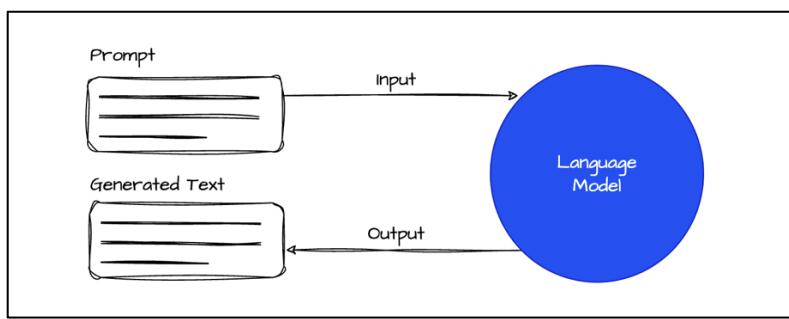
Yapay zeka, chatbot kullanıcılarının girdilerini daha iyi anlamasını sağlar. NLP algoritmaları, kullanıcıların metinlerini analiz ederek bağlamı anlamaya çalışır ve bu sayede chatbotlar, daha doğru ve ilgili yanıtlar üretebilir. Ayrıca, yapay zeka, chatbotların kullanıcı etkileşimlerinden öğrenmesini ve zamanla daha iyi yanıtlar vermesini sağlar. Makine öğrenimi algoritmaları, kullanıcıların önceki etkileşimlerini analiz ederek, gelecekteki sorulara ve isteklere daha iyi yanıtlar oluşturur. Yapay zeka, chatbotların kullanıcıların bireysel ihtiyaç ve tercihlerine göre kişiselleştirilmiş yanıtlar vermesine olanak tanır. Kullanıcıların geçmiş etkileşimleri ve verileri analiz edilerek, daha alakalı ve anlamlı yanıtlar sunulur. Ayrıca, yapay zeka, chatbotların metin, ses ve görüntü gibi farklı giriş modalitelerini kullanarak etkileşim kurmasına tanıyor ve chatbotlarla daha doğal ve sezgisel yollarla iletişim kurabilirler. Yapay zeka, chatbotların daha akıllı, daha duyarlı ve kullanıcı ihtiyaçlarına daha uygun hale gelmesini sağlamaktadır. Bu sayede, chatbotlar sadece basit görevleri yerine getirmekle kalmaz, aynı zamanda kullanıcılarla derin ve anlamlı etkileşimler kurabilir. Günümüzde, yapay zeka destekli chatbotlar müşteri hizmetlerinden eğitim ve sağlık hizmetlerine kadar birçok alanda kullanılmaktadır. Bu teknolojiler, işletmelerin müşteri memnuniyetini artırmamasına ve operasyonel verimliliği yükseltmesine

katkı sağlamaktadır. Ayrıca, yapay zeka tabanlı chatbotlar, kullanıcıların kişisel asistanı olarak da görev yaparak günlük yaşamı kolaylaştırabilir.

1.5. Prompt Engineering

1.5.1. Prompt Engineering Nedir?

Prompt engineering, yapay zeka modellerinin istenen çıktıları üretmek üzere nasıl yönlendirileceğini belirlemek için kullanılan bir teknik olarak tanımlanabilir. Özellikle büyük dil modelleri (LLM'ler) ile çalışırken, modelin verdiği yanıtların doğruluğu ve alaka düzeyi büyük ölçüde kullanılan promptlara (girdi ifadelerine) bağlıdır. Prompt engineering, kullanıcıların bu modellerden en iyi sonuçları alabilmeleri için önemlidir [14].



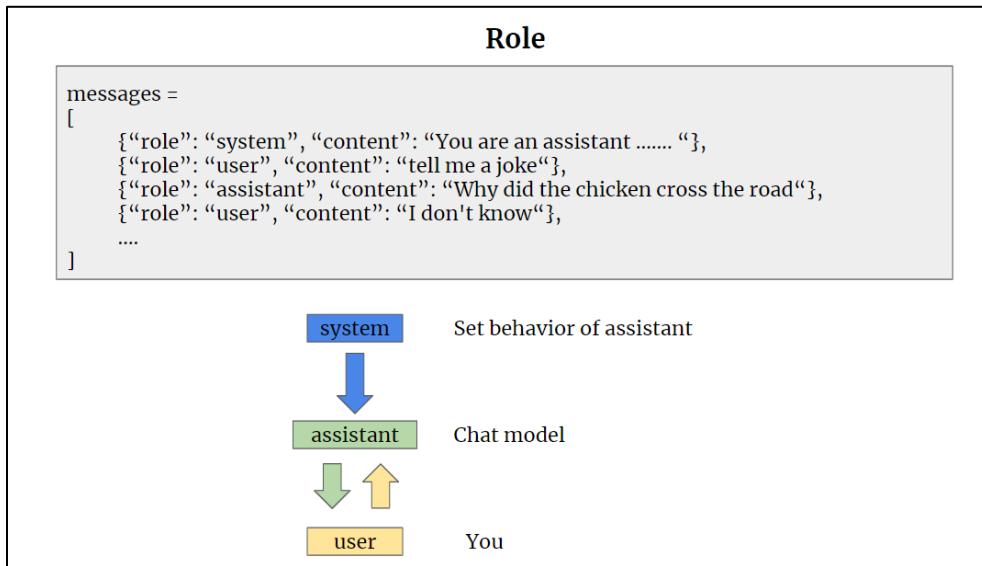
Şekil 1.8. Prompt kullanımıyla dil modeli etkileşimi.

1.5.2. Chatbotlarda Prompt Engineering'in Önemi

Chatbot uygulamalarında, kullanıcıların girdilerine verilen yanıtların doğruluğu ve bağlamla uyumluluğu hayatı öneme sahiptir. Prompt engineering, chatbotların daha doğal, anlamlı ve bağlamla uyumlu yanıtlar üretmesine yardımcı olur [15].

Doğru promptlar kullanılarak, modelin kullanıcı girdilerinin bağlamını anlaması ve buna uygun yanıtlar vermesi sağlanabilir. İyi tasarlanmış promptlar, modelin belirli bir konu veya görev hakkında daha doğru ve tutarlı yanıtlar üretmesine yardımcı olabilir. Kullanıcıya özgü bilgilerle zenginleştirilmiş promptlar, daha kişiselleştirilmiş ve kullanıcı ihtiyaçlarına uygun yanıtlar üretmesini sağlar. Şekil 1.9'de kullanıcı rolleri ile ilgili bir örnek verilmiştir.

Ayrıca, yanlış veya alakasız yanıtların sayısını azaltarak, kullanıcı deneyimini iyileştirir. Bu avantajlar, chatbotların kullanıcılarla daha etkili ve verimli bir şekilde etkileşimde bulunmasını sağlar.



Şekil 1.9. Chatbot Sisteminde Etkileşim Rolleri Örnek.

1.5.3. Teknikler ve Stratejiler

Prompt engineering'de kullanılan bazı yaygın teknikler ve stratejiler şunlardır: Bağlamsal Bilgi Sağlama, kullanıcının önceki etkileşimlerinden veya profil bilgilerinden elde edilen verileri promptlara dahil ederek daha bağlamsal yanıtlar alınmasını sağlar [16].

Açık ve Net Sorular, modelden daha kesin yanıtlar almak için promptların açık ve net olmasının önemini vurgular; belirsiz veya çok geniş sorular yerine, spesifik ve yönlendirici sorular tercih edilmelidir.

Örneklerle Öğretme, modelin doğru yanıtlar üretmesi için örnek yanıtlar veya örnek etkileşimler kullanılarak promptlar oluşturulabilir ve bu teknik, modelin bekleneleri anamasına yardımcı olur. Sınırlı Yanıt Seçenekleri ise modelin yanıt aralığını sınırlayarak, daha kontrollü ve istenen doğrultuda yanıtlar üretmesini sağlar.

Prompt engineering, chatbotların daha etkili, anlamlı ve kullanıcı odaklı yanıtlar üretmesini sağlayan kritik bir tekniktir. Bu teknik sayesinde, yapay zeka modellerinden elde edilen yanıtların kalitesi ve doğruluğu artırılabilir, kullanıcı deneyimi iyileştirilebilir.

1.6. Tezin Yapısı

Bu tez, yapay zeka ve chatbot teknolojilerinin kişiselleştirilmiş ve çok modlu etkileşimlerle nasıl geliştirilebileceğini incelemektedir. Tezde yer alan ana bölümler ve içerdikleri konular aşağıda özetlenmiştir:

Bölüm 1: Giriş

Bu bölüm, tezin genel bir tanıtımını yapar. Yapay zeka ve chatbot teknolojilerinin gelişimi, bu teknolojilerin günlük hayatı ve iş dünyasında nasıl kullanıldığı, kişiselleştirme ve çok modlu etkileşimlerin önemi ele alınır. Ayrıca, bu tez çalışmasının amacı ve kapsamı hakkında bilgi verilir.

- **Uygulamanın Genel Tanımı:** Geliştirilen uygulamanın özellikleri ve kullanım senaryoları.
- **Literatür Özeti:** Mevcut chatbot teknolojileri ve bunların sınırlamaları, çok modlu etkileşimlerin önemi, kişiselleştirmenin chatbotlardaki rolü.
- **Tez Çalışmasının Amacı:** Tez çalışmasının ana hedefleri ve kapsamı.
- **Yapay Zeka:** Yapay zekanın temel prensipleri ve chatbotlardaki rolü.
- **Chatbot (Multimodal Interaction):** Chatbotların çok modlu etkileşim yetenekleri.
- **Prompt Engineering:** Prompt engineering kavramı ve chatbotlardaki önemi.

Bölüm 2: Materyal ve Yöntem

Bu bölüm, tez çalışmasında kullanılan veri toplama ve hazırlama yöntemlerini, geliştirme araçları ve çerçeveleri, sistem mimarisini ve API entegrasyonlarını ele alır.

- **Veri Toplama ve Hazırlama:** Karakter verilerinin toplanması ve temizlenmesi.
- **Geliştirme Araçları ve Çerçeve:** Kullanılan yazılım araçları ve geliştirme çerçeveleri.
- **Sistem Mimarisi:** Uygulamanın genel mimarisi ve bileşenler arası ilişkiler.
- **API Entegrasyonu:** OpenAI ve ElevenLabs API'larının entegrasyonu.
- **Çok Modlu Etkileşim Tasarımı:** Kullanıcıların metin, ses ve görüntü ile etkileşim kurabilmesi için yapılan tasarım.
- **Kullanıcı Arayüzü ve Deneyimi:** Kullanıcı arayüzü tasarımı ve deneyim iyileştirmeleri.

- **Kişiselleştirme Özellikleri:** Kişiselleştirilmiş yanıtlar ve kullanıcı etkileşimlerinin analizi.

Bölüm 3: Bulgular ve Tartışma

Bu bölüm, geliştirilen uygulamanın mevcut çözümlerle karşılaştırılmasını, çok modlu yeteneklerin test edilmesi ve değerlendirilmesini, kullanıcı çalışmaları ve geri bildirimleri, performans metrikleri ve analizini kapsar.

- **Mevcut Çözümlerle Karşılaştırma:** Geliştirilen uygulamanın mevcut chatbot uygulamalarıyla karşılaştırılması.
- **Multimodal Yeteneklerin Testi ve Değerlendirilmesi:** Çok modlu etkileşimlerin performans değerlendirmesi.
- **Kullanıcı Çalışmaları ve Geri Bildirim:** Kullanıcı testlerinden elde edilen veriler ve geri bildirimler.
- **Performans Metrikleri ve Analizi:** Uygulamanın performans metriklerinin analizi ve tartışılması.

Bölüm 4: Sonuçlar

Bu bölüm, tezde elde edilen bulguların özetini, gelecek araştırmalar için çıkarımları ve potansiyel uygulamalar ve genişletme önerilerini içerir.

- **Bulguların Özeti:** Tezde elde edilen bulguların kısa bir özeti.
- **Gelecek Araştırmalar İçin Çıkarımlar:** Gelecekte yapılabilecek çalışmalar için öneriler.
- **Potansiyel Uygulamalar ve Genişletmeler:** Uygulamanın diğer potansiyel kullanım alanları ve genişletme önerileri.
- **Son Düşünceler:** Tezin genel değerlendirmesi ve sonuç.

Bölüm 5: Kaynaklar

Bu bölüm, tez boyunca kullanılan tüm referansları ve kaynakları listeler.

2. MATERİYAL VE YÖNTEM

Bu çalışmada geliştirilen uygulama, JavaScript ve React Native ile oluşturulmuş, OpenAI ve ElevenLabs API'lerini kullanarak çok modlu etkileşim yetenekleri kazandırılmıştır. Uygulama, Visual Studio Code ve Copilot ortamlarında geliştirilmiş olup, hem iOS hem de Android platformları için tasarlanmıştır. Expo ve Expo EAS kullanılarak APK ve build dosyaları oluşturulmuştur.

2.1. Veri Toplama ve Hazırlama

Geliştirilen uygulamanın veritabanı, karakterlerin adları, fotoğrafları, yaşıları, cinsiyetleri, memleketleri, ilgi alanları, kategorileri ve açıklamalarını içermektedir. Veriler, JSON formatında saklanmış ve şu adımlar izlenerek hazırlanmıştır: Sanal karakterlerin bilgileri çeşitli kaynaklardan toplanmış, eksik, hatalı veya tutarsız bilgiler temizlenmiş ve verilerin doğruluğu sağlanmıştır. Veri bütünlüğünü ve kullanım kolaylığını sağlamak için, veri setlerindeki ve yapay zeka modelinden gelen metin verileri tutarlı bir şekilde formatlanmış, gereksiz kalın, italik veya diğer metin dekorasyonları kaldırılarak standart bir formata getirilmiştir. Bu işlem adımları; kalın ve italik metin desenlerinin belirlenmesi, her bir desen için metnin parçalanması ve stil bilgisinin eklenmesi, ve parçaların birleştirilerek temiz metnin elde edilmesi şeklindedir. Veriler, JSON formatına dönüştürülerek uygulamanın ihtiyaçlarına uygun hale getirilmiştir. Karakterlerin görselleri ise, ChatGPT-4O ve Lexica Art kullanılarak yapay zeka ile oluşturulmuştur [18]. Şekil 2.1'de metin verilerinin tutarlı bir şekilde formatlanması ve kalın ile italik metin desenlerinin tanımlanması ve birleştirilmesi gösterilmektedir.

$$\begin{aligned} \text{Temiz Metin} &= f(\text{Girdi Metni}) & \text{Pattern}_1 &= (\backslash * \backslash * | _) (. * ?) (\backslash * \backslash * | _) \\ f(\text{Girdi Metni}) &= \sum_{i=1}^n \text{segment}_i & \text{Pattern}_2 &= (\backslash * | _) (. * ?) (\backslash * | _) \\ \left\{ \begin{array}{ll} \text{text : fragmented text, style : blank} & \text{if, segment is plain text} \\ \text{text : fragmented text, style : bold or italic} & \text{if, segment has style} \end{array} \right. \end{aligned}$$

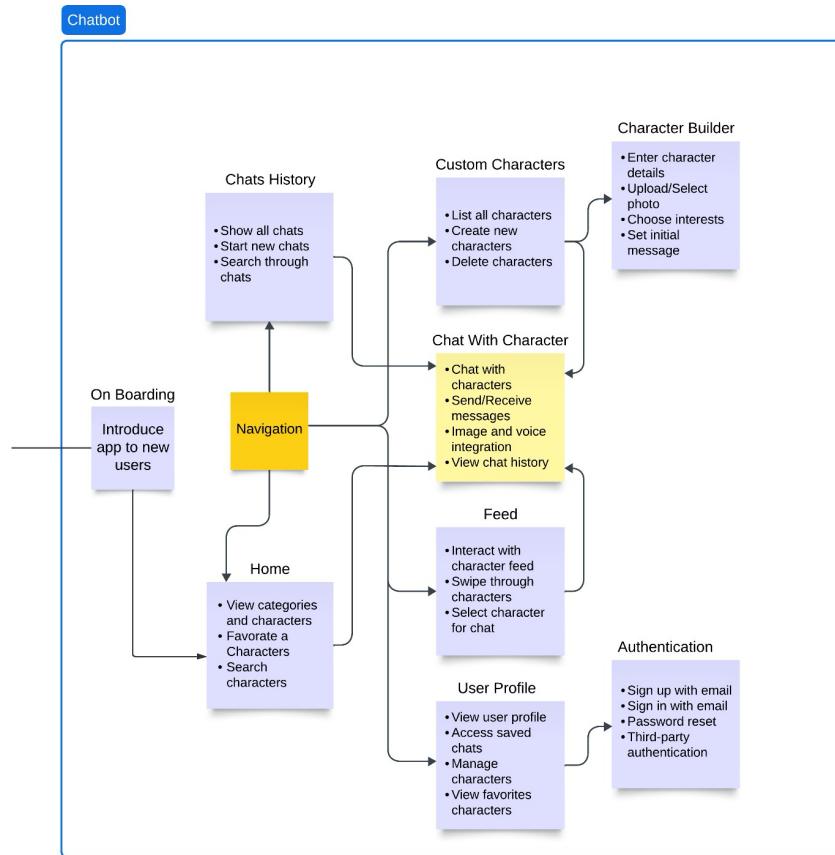
Şekil 2.1. Metin Verilerinin Formatlanması ve Desenlerinin Birleştirilmesi.

2.2. Geliştirme Araçları ve Çerçeveler

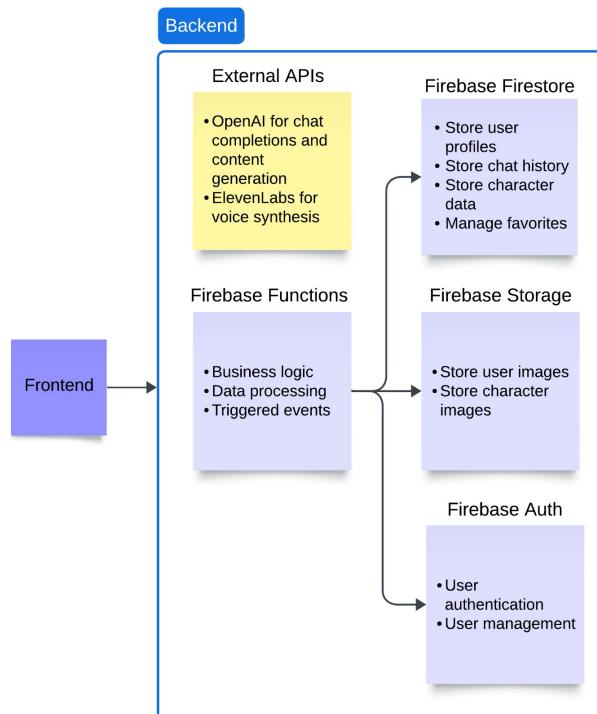
Uygulamanın geliştirilmesinde kullanılan ana programlama dili JavaScript olup, React Native, çapraz platformlarda [19] çalışan kullanıcı arayüzü bileşenleri oluşturmak için tercih edilmiştir. Bu kombinasyon, hızlı ve verimli bir şekilde mobil uygulama geliştirilmesini sağlamıştır. Expo ve Expo EAS, React Native uygulamaları için geliştirme sürecini kolaylaştıran çerçeveler olarak kullanılmış ve APK ile build dosyalarının oluşturulmasını sağlamıştır [20]. Uygulamanın kimlik doğrulama, veri depolama ve hosting ihtiyaçları için Firebase kullanılmış olup, güvenilir ve ölçeklenebilir bir arka uç altyapısı sunmaktadır [21]. Geliştirme sürecinde tercih edilen kod editörü Visual Studio Code, Copilot entegrasyonu sayesinde kod yazma sürecini hızlandırmış ve hata oranını azaltmıştır. API çağrıları için kullanılan Axios kütüphanesi, HTTP isteklerini kolayca yönetmeyi sağlayarak veri alışverişini etkin bir şekilde gerçekleştirmiştir.

2.3. Sistem Mimarisi

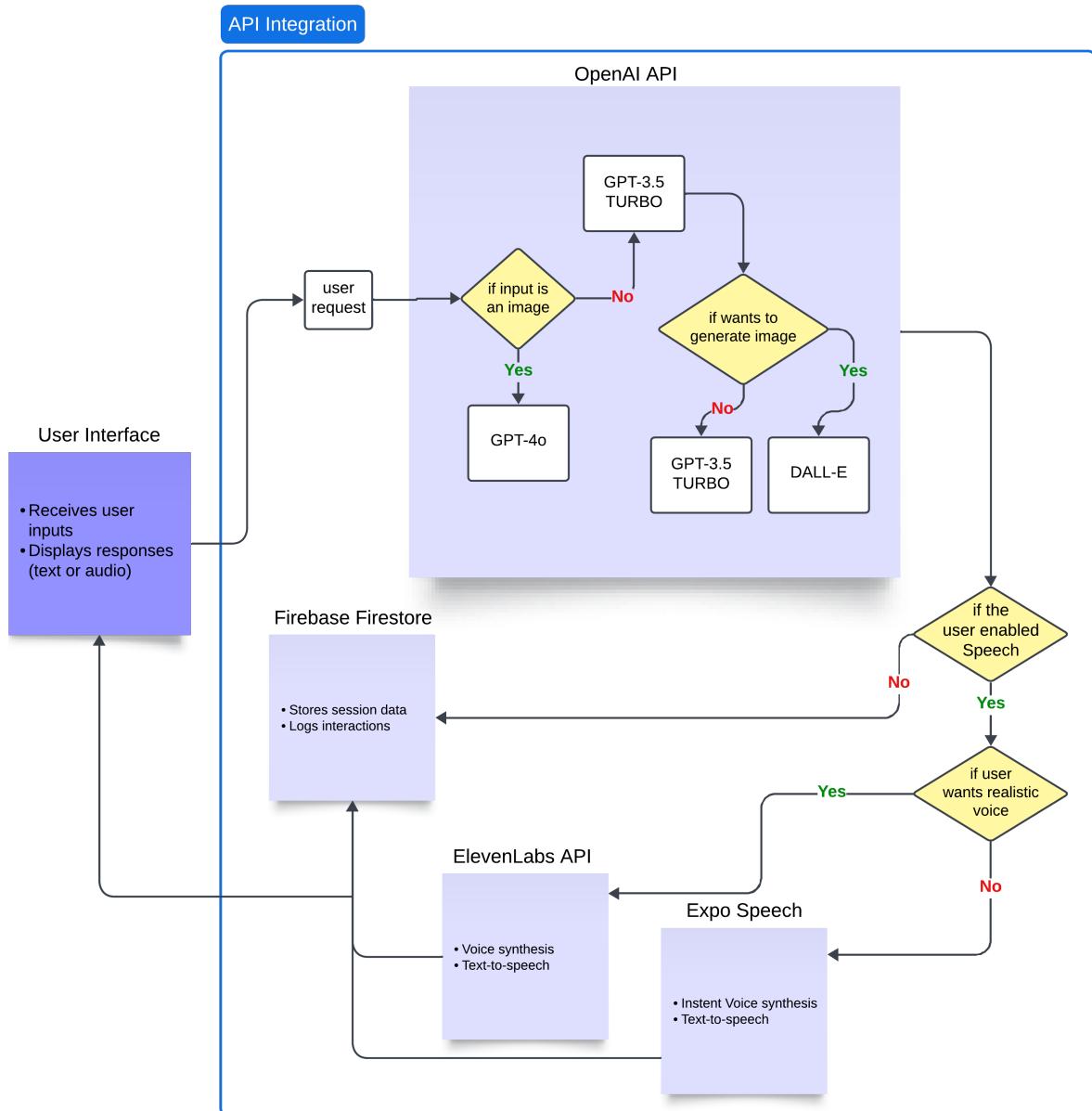
Uygulamanın sistem mimarisi üç ana bileşenden oluşmaktadır: ön uç (frontend), arka uç (backend) ve API entegrasyonu. Frontend, React Native kullanılarak oluşturulmuş olup, kullanıcı arayüzü ve etkileşimlerin yönetildiği katmandır; bu katman, kullanıcıların uygulama ile etkileşime geçmesini sağlar ve veri girişini kolaylaştırır Şekil 2.2'de gösterildiği gibi. React Native'in kullanılması, uygulamanın iOS ve Android gibi farklı platformlarda sorunsuz çalışmasını sağlamıştır. Backend ise, Firebase kullanılarak veri depolama, kimlik doğrulama ve sunucu tarafı işlemler gerçekleştirilmiş, bu sayede güvenilir ve ölçeklenebilir bir altyapı sağlanmıştır ve Şekil 2.3'de gösterilmiştir. Firebase, kullanıcı verilerinin güvenliğini ve uygulamanın performansını artırarak, yüksek trafiğe sahip ortamlarda dahi stabil bir hizmet sunulmasını sağlamıştır [21]. API entegrasyonu kapsamında, OpenAI ve ElevenLabs API'leri ile metin, ses ve görüntü işleme fonksiyonları sağlanmıştır. OpenAI API'si, metin oluşturma ve görüntü işleme yetenekleri sunarken, ElevenLabs API'si yüksek kaliteli sesli yanıtlar oluşturma fonksiyonunu yerine getirmiştir ve Şekil 2.4'de gösterilmiştir. Bu API'ler sayesinde kullanıcı girdileri etkili bir şekilde işlenmiş ve uygun yanıtlar üretilmiştir, böylece kullanıcı deneyimi önemli ölçüde iyileştirilir [22, 23].



Şekil 2.2. Kullanıcı arayüzü bileşenleri ve etkileşimleri gösteren kullanım senaryosu diyagramı.



Şekil 2.3. Veri akışı ve sunucu tarafı işlemleri gösteren diyagram.



Şekil 2.4. API çağrılarının nasıl yapıldığını ve yanıtların nasıl işlendiğini gösteren diyagram.

2.4. API Entegrasyonu

Uygulamanın işlevsellliğini artırmak için aşağıdaki API'ler entegre edilmiştir: OpenAI API, metin oluşturma, görüntü oluşturma ve görüntü işleme fonksiyonları sağlamaktadır. Kullanıcılar, metin tabanlı girdilerle chatbotlardan yanıt alabilir ve görüntü tabanlı sorgular yapabilir [23]. ElevenLabs API, yüksek kaliteli sesli yanıtlar oluşturmak için kullanılmıştır. Kullanıcılar sesli komutlar vererek chatbotlardan sesli yanıtlar alabilir [22]. Ayrıca, Expo Voice, kullanıcıların sesli mesaj göndermesi için kullanılmıştır. Kullanıcılar, sesli girdilerle chatbotlarla etkileşim kurabilir ve sesli yanıtlar alabilir. Bu API'ler sayesinde, uygulamanın çok modlu etkileşim yetenekleri geliştirilmiş ve kullanıcı deneyimi önemli ölçüde iyileştirilmiştir.

2.5. Çok Modlu Etkileşim Tasarımı

Uygulamanın çok modlu etkileşim yetenekleri, kullanıcıların metin, ses ve görüntü gibi çeşitli modalitelerle etkileşim kurmalarını sağlamaktadır. Bu tasarım, metin, sesli ve görüntü girdileri ile yanıtları içermektedir. Metin etkileşimleri için OpenAI'nin GPT-4 modelinden yararlanılmıştır; kullanıcılar metin mesajları gönderir ve chatbotlardan metin tabanlı yanıtlar alır. Sesli etkileşimler için Expo Voice ve ElevenLabs API'leri kullanılmıştır; kullanıcılar sesli komutlar vererek chatbotlarla etkileşim kurar ve sesli yanıtlar alır. Görüntü işleme ve yanıtları için OpenAI'nin DALL-E ve Vision modelleri kullanılmıştır; kullanıcılar görüntü gönderir ve chatbotlardan görüntü tabanlı yanıtlar alır. Bu bileşenler, uygulamanın kullanıcı deneyimini zenginleştirir ve daha doğal ve sezgisel bir iletişim sağlar.

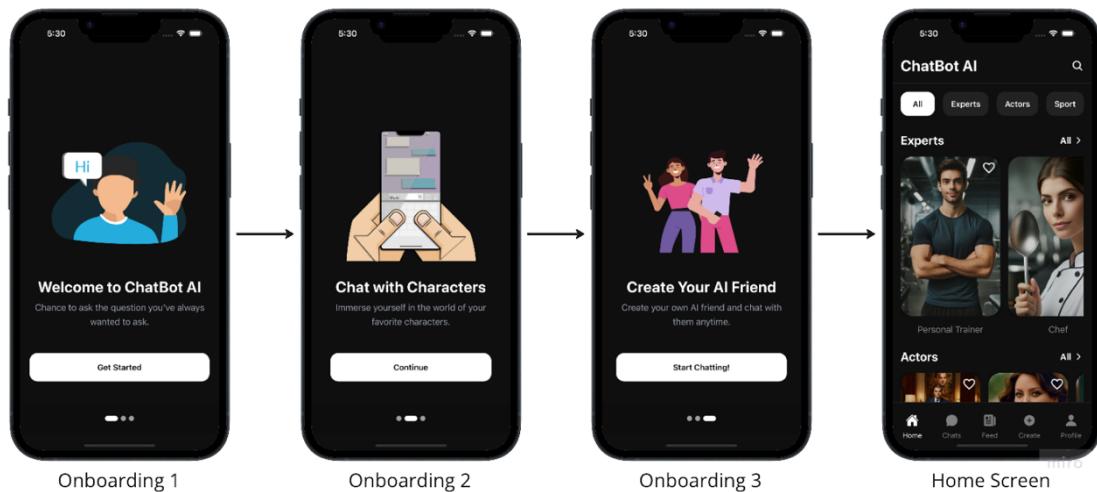
2.6. Kullanıcı Arayüzü ve Deneyimi

Kullanıcı arayüzü, kullanıcıların kolayca etkileşimde bulunabileceği şekilde tasarlanmıştır. React Native ve Tailwind CSS kullanılarak kullanıcı dostu bir arayüz oluşturulmuştur.

Kullanıcı deneyimi, uygulamanın kullanılabilirliği ve işlevselligi açısından optimize edilmiştir. Uygulamanın ekran görüntüleri ve görselleri bu bölümde yer alacaktır.

Onboarding Ekranelri

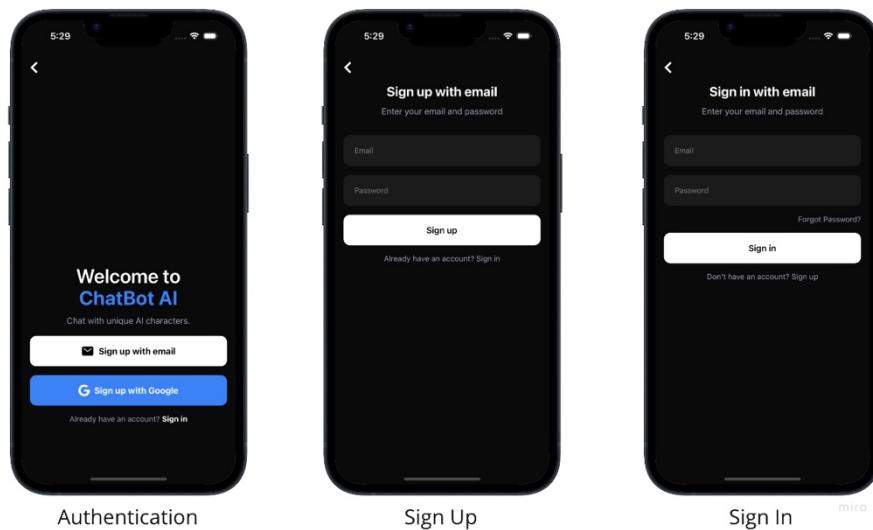
Uygulamanın başlangıcında kullanıcılara uygulamanın temel işlevlerini ve özelliklerini tanıtan onboarding ekranları gösterilir. Şekil 2.5'de gösterilen bu ekranlar, kullanıcıların uygulamayı daha etkili bir şekilde kullanmalarına yardımcı olur.



Şekil 2.5. Uygulamanın temel işlevlerini tanıtan onboarding ekranları.

Authentication Ekranelri

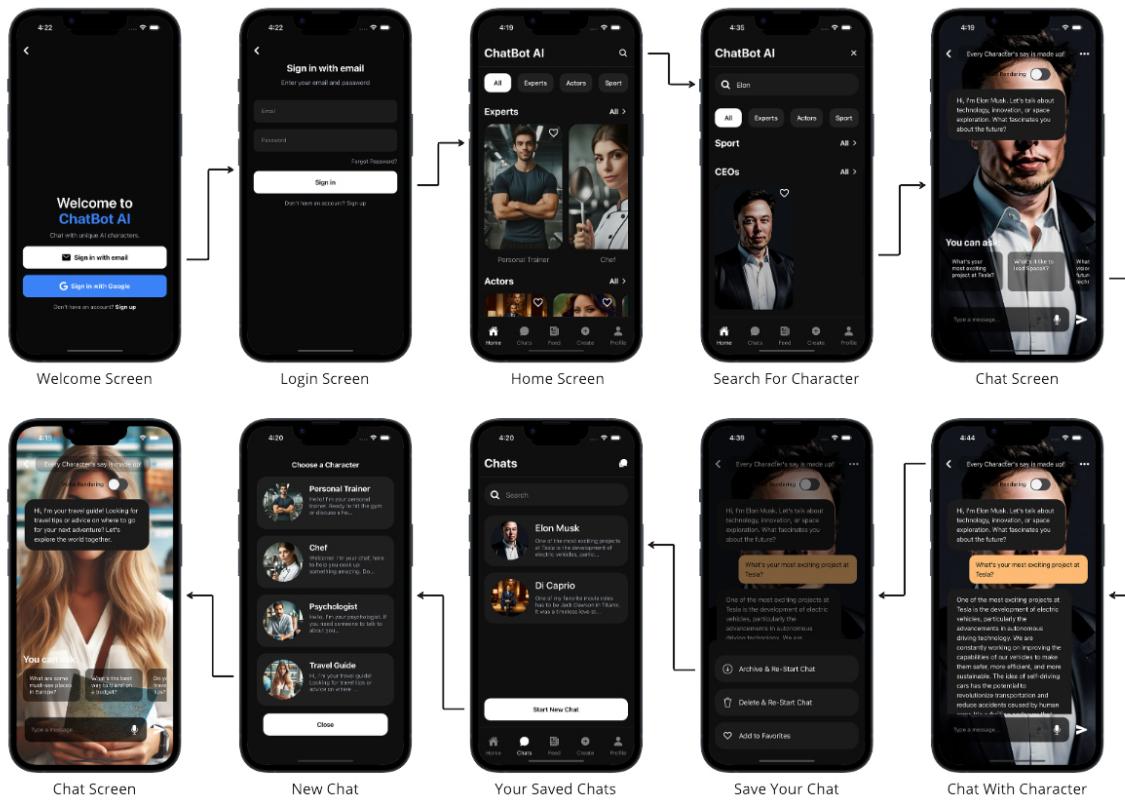
Kullanıcılar, uygulamaya giriş yapmak veya yeni bir hesap oluşturmak için kimlik doğrulama ekranlarını kullanır. Şekil 2.6'de gösterilen bu ekranlarla E-posta veya Google hesabı ile giriş seçenekleri sunulmaktadır.



Şekil 2.6. Giriş ve kayıt işlemleri için kimlik doğrulama ekranları.

Home ve Chat Ekranları

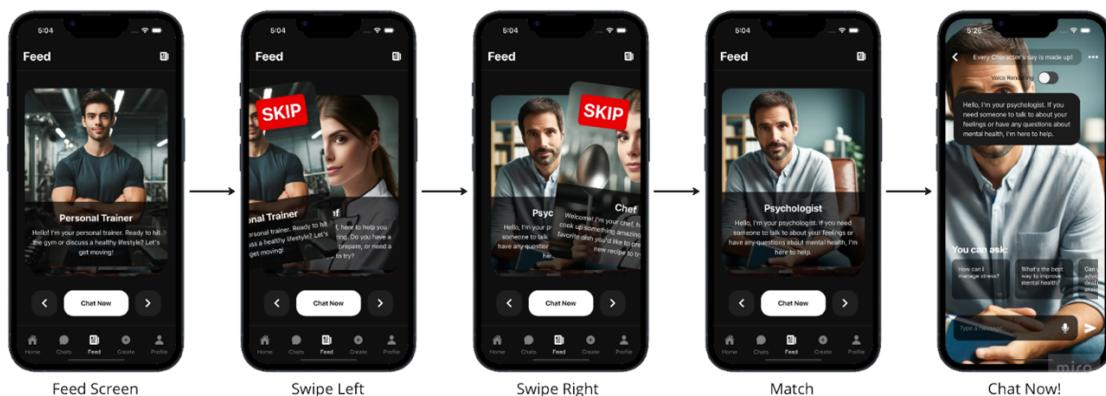
Ana ekran, kullanıcınlara çeşitli karakter kategorilerini keşfetme imkanı sunar. Şekil 2.7'de gösterilen bu ekranlarda kullanıcılar, belirli karakterleri arayabilir ve seçebilirler. Seçilen karakterle sohbet ekranında etkileşime geçilebilir.



Şekil 2.7. Ana sayfa ve karakterlerle sohbet ekranları.

Feed Ekranları

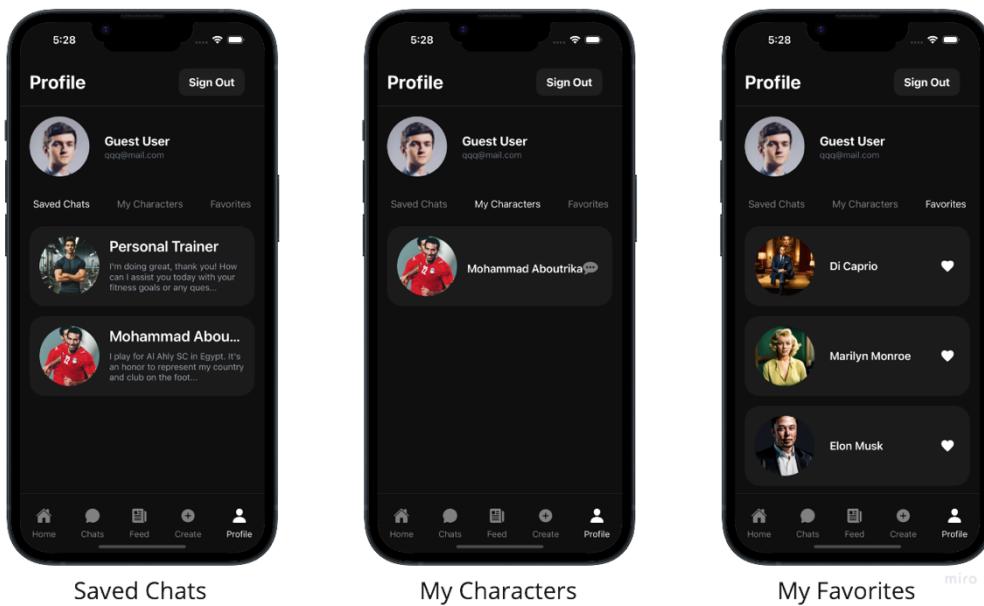
Şekil 2.8'deki Feed ekranı, kullanıcıların karakterleri sağa ve sola kaydırarak gezmemelerini sağlar. Kullanıcılar, beğenikleri karakterleri seçebilir ve onunla mesajlaşabilir.



Şekil 2.8. Karakterlerin gezdiği feed ekranı.

Profile Ekranları

Şekil 2.9'de gösterilen Profil ekranı, kullanıcıların kaydedilmiş sohbetlerini, oluşturdukları karakterleri ve favori karakterlerini görüntülemelerini sağlar. Bu ekran, kullanıcıların uygulama içindeki etkileşimlerini düzenlemelerine ve yönetmelerine yardımcı olur.



Şekil 2.9. Kaydedilmiş sohbetlerin ve favori karakterlerin görüntülendiği profil ekranı.

2.7. Konu Dışı Yanıtları Önleme (Prompt Engineering)

Bu bölümde, sanal karakterlerin belirli bir konudan sapmaması ve tutarlı yanıtlar vermesi için kullanılan prompt engineering teknikleri detaylandırılacaktır. Bu teknikler, karakterlerin rolüne uygun, tarihsel ve tematik doğrulukta yanıtlar vermelerini sağlar ve literatürde yazılan kaynaklardan faydalayılarak geliştirilmiştir [24, 25].

Her sanal karakter, kendine özgü bir kişilik, konuşma tarzı ve tarihsel bağlam ile tanımlanır; bu, karakterin yanıtlarının belirli bir temaya uygun olmasını sağlar. Örneğin, bir tarihçi karakteri yalnızca kendi uzmanlık alanıyla ilgili konularda bilgi verir ve modern teknolojiler hakkında konuşmaz.

Karakterler, yalnızca kendi zaman dilimleri ve uzmanlık alanlarıyla ilgili konularda konuşur; modern teknolojiler veya mevcut olmayan bilgiler hakkında konuşmaları engellenir, böylece tarihsel doğruluk ve tutarlılık korunur. Karakterlere verilen görevler, yalnızca belirli konularla ilgili soruları yanıtlamalarını gerektirir. Örneğin, bir şef karakteri yemek tarifleri ve mutfaq teknikleri hakkında konuşurken, bir kişisel antrenör karakteri fitness ve egzersiz konularında bilgi verir, bu da karakterlerin uzmanlık alanlarına odaklanmasını sağlar.

Karakterler, kendileriyle ilgisiz sorularla karşılaşıklarında, kibarca konuyu kendi uzmanlık alanlarına yönlendirirler; örneğin, bir tarihçi karakter, modern teknolojiler hakkında bir soru aldığında, tarihi olaylar ve figürler hakkında konuşmayı tercih eder, bu şekilde kullanıcıları doğru bilgiye yönlendirir ve diyalogun verimli olmasını sağlar.

Karakterler, sadece sorulara yanıt vermekle kalmaz, aynı zamanda kullanıcılarla etkileşimi sürdürmek için kendi sorularını da sorar, bu da diyalogların daha canlı ve ilgi çekici olmasını sağlar. Örneğin, bir seyahat rehberi karakteri kullanıcılarla seyahat planları hakkında sorular sorabilir ve önerilerde bulunabilir.

Karakterlerin yanıtları, belirli bir uzunlukta tutulur ve genellikle kısa ve öz olur, bu da kullanıcıların hızlı ve net bilgiler almasını sağlar; uzun ve karmaşık yanıtlar yerine, kısa ve anlaşılır yanıtlar tercih edilir.

Örnek Senaryo

Bir tarihçi karakter için kullanılan prompt engineering teknikleri aşağıdaki gibidir:

- **Full Character Immersion (Tam Karakter İçselleştirme):** Karakter tamamen kendi kimliğine bürünür, konuşma tarzını ve görüşlerini benimser.
- **Engage Accurately (Doğru Katılım):** Karakter, tarihsel dönemiyle ilgili doğru ve bilgilendirici yanıtlar verir.
- **Historical Consistency (Tarihsel Tutarlılık):** Karakter, yalnızca kendi dönemde mevcut olan teknolojiler ve olaylar hakkında konuşur.
- **Maintain Your Persona (Karakteri Koruma):** Karakter, kendi kimliğini ve özelliklerini korur, yanıtlarında tutarlı kalır.
- **Relevant Topics (İlgili Konular):** Karakter, ilgisiz soruları kibarca kendi uzmanlık alanına yönlendirir.
- **Proactive Engagement (Proaktif Katılım):** Karakter, kullanıcıyla etkileşimi sürdürmek için kendi sorularını sorar.
- **Avoid Negatives (Olumsuzluktan Kaçınma):** Karakter, "Bilmiyorum" veya "Yapamam" gibi yanıtlar vermekten kaçınır, her zaman bilgi sağlamaya çalışır.
- **Response Length (Yanıt Uzunluğu):** Yanıtlar kısa ve öz olur.

Bu yöntemler, sanal karakterlerin kullanıcılarla anlamlı ve eğitici bir şekilde etkileşimde bulunmasını sağlar ve uygulamanın genel deneyimini iyileştirir.

3. BULGULAR VE TARTIŞMA

Bu bölümde, çok modlu etkileşimlerin kullanıcı deneyimini nasıl iyileştirdiği ve geliştirilen uygulamanın performansının değerlendirilmesi ele alınacaktır. Ayrıca, uygulamanın kullanıcı geri bildirimlerine dayalı değerlendirmeleri sunulacaktır.

3.1. Multimodal Yeteneklerin Testi ve Değerlendirilmesi

Geliştirilen uygulamanın çok modlu etkileşim yetenekleri, kullanıcıların metin, ses ve görüntü modaliteleri üzerinden gerçekleştirdiği etkileşimlerle test edilmiştir. Kullanıcılar, farklı modalitelerle chatbotlarla etkileşim kurmuş ve uygulamanın bu girdilere verdiği yanıtlar değerlendirilmiştir. Metin etkileşimlerinde, kullanıcılar chatbotlarla metin tabanlı sohbetler gerçekleştirmiş ve alınan yanıtların doğruluğu ve bağlamla uyumu

değerlendirilmiştir. Sesli etkileşimlerde, kullanıcılar sesli komutlar vererek chatbotlarla etkileşimde bulunmuş ve sesli yanıtların doğruluğu ve anlaşılırlığı test edilmiştir. Görüntü etkileşimlerinde ise, kullanıcılar görüntü göndererek chatbotlardan yanıt almış ve bu etkileşimlerdeki görsel girdilerin işlenme doğruluğu ve yanıtların uygunluğu değerlendirilmiştir. Multimodal etkileşimlerin kullanıcı deneyimini iyileştirdiği, mevcut araştırmalarla da desteklenmektedir. Örneğin, ses ve metin tabanlı etkileşimlerin kullanıcı deneyimini artırdığı belirtilmiştir [26]. Nexgits tarafından yapılan araştırmada, chatbotların etkili kullanıcı etkileşimi sağlamak için çok modlu etkileşimleri nasıl kullanabileceğini detaylandırmıştır. Ayrıca, çok modlu yapay zeka modellerinin metin, ses ve görsel verileri birleştirerek daha zengin ve dinamik içerikler ürettiği vurgulanmaktadır [27]. La Faber'in çalışması, çok modlu yapay zekanın metin, ses ve görüntü işleme alanında nasıl devrim yarattığını incelemektedir.

3.2. Kullanıcı Çalışmaları ve Geri Bildirim

Kullanıcı çalışmaları, uygulamanın gerçek kullanıcılar tarafından test edilmesini ve geri bildirimlerin toplanmasını içermektedir. Bu çalışmalarında, uygulamanın arkadaşlar ve aile üyeleri tarafından kullanılması sağlanmış ve bu kullanıcıların deneyimleri değerlendirilmiştir. Kullanıcı memnuniyeti kapsamında, kullanıcılar uygulamanın genel memnuniyetini ve önerilerini paylaşmıştır. Kullanılabilirlik açısından, uygulamanın kullanıcı dostu olup olmadığı ve kullanıcıların arayüz ile etkileşim tasarımını nasıl deneyimlediği incelenmiştir. Performans değerlendirmesinde ise, uygulamanın yanıt verme hızı, doğruluğu ve çok modlu etkileşimlerdeki başarısı analiz edilmiştir.

3.3. Performans Metrikleri ve Analizi

Uygulamanın performansını değerlendirmek için çeşitli metrikler kullanılmıştır ve bu metrikler, uygulamanın kullanıcı ihtiyaçlarını ne ölçüde karşıladığı ve genel performansını ortaya koymaktadır. Doğruluk, chatbotların verdiği yanıtların doğruluğu ve bağlamla uyumu; yanıt süresi, kullanıcıların girdilerine verilen yanıtların hızı; kullanıcı memnuniyeti skorları, kullanıcıların uygulama deneyimi sonrası verdikleri memnuniyet puanları; ve çok modlu etkileşim başarısı, metin, ses ve görüntü modaliteleri üzerinden alınan yanıtların doğruluğu ve kullanıcı memnuniyeti şeklinde değerlendirilmiştir. Uygulamanın performansı değerlendirildiğinde, sesli yanıtların oluşturulmasında bazı gecikmeler yaşadığı gözlemlenmiştir. Özellikle ElevenLabs kullanılarak oluşturulan yüksek kaliteli sesler için,

API üzerinden seslerin oluşturulması ve iletilmesi birkaç saniye sürmektedir, bu nedenle kullanıcıların sesli yanıtları alırken bir miktar beklemeleri gerekmektedir. Buna alternatif olarak, OpenAI'nin text-to-speech teknolojisi veya Expo Speech kütüphanesi kullanılarak daha hızlı ancak daha robotik sesler elde edilebilir. Expo Speech kütüphanesi kullanıldığında, yanıtlar arasında oluşturulmakta ve gecikme yaşanmamaktadır. Çok modlu etkileşimlerin kullanıcı deneyimini nasıl iyileştirdiğine dair mevcut literatürde de geniş kapsamlı araştırmalar bulunmaktadır. Örneğin, metin, ses ve görsel verilerin birleştirilmesinin kullanıcı etkileşimlerini ve deneyimini önemli ölçüde iyileştirdiği belirtilmiştir [28]. Jaimes ve Sebe'nin yaptığı çalışmada, çok modlu insan-bilgisayar etkileşimi alanındaki temel yaklaşımlar ve bu alandaki zorluklar incelenmiştir. Sonuç olarak, geliştirilen uygulama, çok modlu etkileşim yetenekleri ve kişiselleştirilmiş yanıtlar sunma konusundaki başarısını göstermiştir. Kullanıcı geri bildirimleri ve performans metrikleri, uygulamanın mevcut çözümlerden nasıl üstün olduğunu ve kullanıcı deneyimini nasıl iyileştirdiğini ortaya koymaktadır.

4. SONUÇLAR

4.1. Bulguların Özeti

Bu tezde, OpenAI API'leri kullanılarak geliştirilen ve çok modlu etkileşim yeteneklerine sahip bir chatbot uygulamasının tasarımları ve performansı incelenmiştir. Uygulama, kullanıcıların metin, ses ve görüntü gibi çeşitli modalitelerle etkileşim kurmalarını sağlamış ve kullanıcı deneyimini iyileştirdiği gözlemlenmiştir. Kullanıcı çalışmaları ve geri bildirimler, uygulamanın genel memnuniyet ve kullanılabilirlik açısından başarılı olduğunu göstermektedir. Ayrıca, çok modlu etkileşimlerin kullanıcı deneyimini zenginleştirdiği ve chatbotların daha doğal ve sezgisel bir şekilde kullanılmasını sağladığı ortaya konmuştur.

4.2. Gelecek Araştırmalar İçin Çıkarımlar

Gelecekteki araştırmalar, çok modlu etkileşimlerin daha da geliştirilmesine ve kullanıcı deneyimini daha da iyileştirmeye odaklanabilir. Özellikle, sesli yanıtların oluşturulmasındaki gecikmelerin azaltılması için alternatif yöntemler ve teknolojiler araştırılabilir. OpenAI'nin text-to-speech teknolojisinin entegrasyonu veya daha hızlı ses oluşturma yöntemleri üzerine çalışılabilir. Ayrıca, daha geniş kullanıcı grupları ile yapılan

kapsamlı kullanıcı çalışmaları, uygulamanın performansını ve kullanıcı memnuniyetini daha iyi anlamaya yardımcı olabilir.

GPT-4O'nun ortaya çıkıştı ile birlikte, sanal asistanların yeni bir seviyeye ulaşması beklenmektedir. GPT-4O, daha gelişmiş dil anlama ve üretme yetenekleri ile çok modlu etkileşimlerde daha hızlı ve daha doğal yanıtlar verebilecektir. Bu, sanal asistanların kullanıcılarla daha etkili ve verimli bir şekilde etkileşim kurmasını sağlayacaktır [29, 30].

4.3. Potansiyel Uygulamalar ve Genişletmeler

Geliştirilen chatbot uygulaması, çeşitli alanlarda kullanılabilecek geniş bir potansiyele sahiptir. Kişisel antrenörler, yemek tarifleri, seyahat planlama ve psikolojik destek gibi farklı senaryolarda kullanılabilir. Ayrıca, uygulama diğer sektörlerdeki müşterileri hizmetleri, eğitim ve eğlence alanlarında da genişletilebilir.

4.4. Son Düşünceler

Bu tez çalışması, çok modlu etkileşimlerin chatbot uygulamalarında kullanıcı deneyimini nasıl iyileştirebileceğini göstermiştir. OpenAI API'leri kullanılarak geliştirilen uygulama, kullanıcıların daha doğal ve sezgisel bir şekilde etkileşim kurmalarını sağlamış ve bu alandaki potansiyeli ortaya koymuştur. Gelecekteki araştırmalar ve geliştirmeler, bu teknolojilerin daha geniş bir yelpazede uygulanmasını ve kullanıcı deneyimini daha da iyileştirmesini sağlayacaktır.

5. KAYNAKLAR

- [1] D. Pierce, "From Eliza to ChatGPT: why people spent 60 years building chatbots," The Verge, Feb. 28, 2024. [Online]. Available: <https://www.theverge.com/24054603/chatbot-chatgpt-eliza-history-ai-assistants-video> (Erişim tarihi: 31.05.2024).
- [2] "Examining AI Chatbots: Advancements, Advantages, and Prospects," Search My Expert Blog, Jan. 8, 2024. [Online]. Available: <https://blog.searchmyexpert.com/ai-chatbots> (Erişim tarihi: 31.05.2024).
- [3] "ChatGPT: Unveiling the Power of Advanced Language Models," Interesting Engineering, July 14, 2023. [Online]. Available: <https://interestingengineering.com/innovation/ai-openai-chatgpt-chatbots-language-models> (Erişim tarihi: 31.05.2024).
- [4] "The Next Frontier in Chatbot Technology: Multi-Modal Interactions," LinkedIn, [Online]. Available: <https://www.linkedin.com/pulse/next-frontier-chatbot-technology-multi-modal-interactions> (Erişim tarihi: 31.05.2024).
- [5] "Personalize AI Chatbot Interactions: Enhancing User Experience with Custom Conversations," HelpX AI, [Online]. Available: <https://helpx.ai/blogs/personalize-ai-chatbot-interactions> (Erişim tarihi: 31.05.2024).
- [6] "Character AI," Elegant Themes, [Online]. Available: <https://www.elegantthemes.com/blog/business/character-ai> (Erişim tarihi: 31.05.2024).
- [7] "Talkie-ai," Deepgram, [Online]. Available: <https://deepgram.com/ai-apps/talkie> (Erişim tarihi: 31.05.2024).
- [8] "Charsis AI Character Chat," AppAdvice, [Online]. Available: <https://appadvice.com/app/charsis-ai-character-chat/6450896930> (Erişim tarihi: 31.05.2024).
- [9] "Google AI Principles," Google AI, [Online]. Available: <https://ai.google/responsibility/principles> (Erişim tarihi: 31.05.2024).
- [10] A. Chahal and P. Gulia, "Machine Learning and Deep Learning," International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 12, pp. 4910-4914, Oct. 2019.

- [11] M. M. Taye, "Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions," *Computers*, vol. 12, no. 5, p. 91, Apr. 2023. [Online].
- [12] F. Song, J. Sun, and T. Wang, "Overview of Natural Language Processing Technologies and Rationales in Application," *Theory and Practice in Language Studies*, vol. 10, no. 1, pp. 49, Dec. 2019.
- [13] F. Aslam, "The Impact of Artificial Intelligence on Chatbot Technology: A Study on the Current Advancements and Leading Innovations," *European Journal of Technology*, vol. 7, no. 3, pp. 62-72, 2023.
- [14] A. Bozkurt, "Tell Me Your Prompts and I Will Make Them True: The Alchemy of Prompt Engineering and Generative AI," *Open Praxis*, vol. 16, no. 2, pp. 111-118, Apr. 2024.
- [15] M. K. Pehlivanoğlu, M. A. Syakura, and N. Duru, "Enhancing Paraphrasing in Chatbots Through Prompt Engineering: A Comparative Study on ChatGPT, Bing, and Bard," in 2023 8th International Conference on Computer Science and Engineering (UBMK), Burdur, Turkey, Sep. 2023.
- [16] J. White et al., "A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT," arXiv, Feb. 2023. [Online]. Available: <https://arxiv.org/abs/2302.11382> (Erişim tarihi: 31.05.2024).
- [17] K. Srikanth, "Mastering Prompt Engineering Strategies and Tactics," LinkedIn, Mar. 12, 2024. [Online]. Available: <https://www.linkedin.com/pulse/mastering-prompt-engineering-strategies-tactics-krishna-srikanth-k-f8enc/> (Erişim tarihi: 31.05.2024).
- [18] Lexica Art, "Lexica Art Platform," 2024. [Çevrimiçi]. Available: <https://lexica.art/> (Erişim tarihi: 31.05.2024).
- [19] A. Nagpal, "Cross-Platform App Development: Build Once, Run Across Devices," Medium, Jan. 1, 2024. [Online]. Available: <https://medium.com/@jasminepuno/cross-platform-app-development-build-once-run-across-devices-9f7559db9329> (Erişim tarihi: 31.05.2024).
- [20] J. Hazarika, "Here's Why You Should Use Expo Over React Native," Medium, Jan. 3, 2024. [Online]. Available: <https://medium.com/@jorianh25/heres-why-you-should-use-expo-over-react-native-2327dfa05122> (Erişim tarihi: 31.05.2024).

- [21] R. Sanghvi, "Why Choose Firebase for Your App Development in 2024," Bitcot, Feb. 27, 2024. [Online]. Available: <https://www.bitcot.com/why-choose-firebase-for-app-development> (Erişim tarihi: 31.05.2024).
- [22] P. Taylor, "ElevenLabs: A Powerful API for Natural-sounding Text-to-Speech," Medium, Aug. 3, 2023. [Online]. Available: <https://medium.com/@paulotaylor/elevenglabs-a-powerful-api-for-natural-sounding-text-to-speech-75439d3563a8> (Erişim tarihi: 31.05.2024).
- [23] G. Jabeen, "A Comprehensive Guide to Using OpenAI's ChatGPT API," Medium, Oct. 10, 2023. [Online]. Available: <https://medium.com/@guljabeen222/a-comprehensive-guide-to-using-openai-s-chatgpt-api-41f29cdb6de1> (Erişim tarihi: 31.05.2024).
- [24] A. Houston, "How to prevent users trying to break your chatbot," QBox Blog, Mar. 29, 2023. [Online]. Available: <https://blog.qbox.ai/blog/prevent-users-breaking-your-chatbot> (Erişim tarihi: 31.05.2024).
- [25] P. K. Namala, "AI Chatbot Precision: Techniques to Avoid Conversational Drift," Medium, Jan. 17, 2024. [Online]. Available: <https://phaneendrakn.medium.com/ai-chatbot-precision-techniques-to-avoid-conversational-drift-219845659a71> (Erişim tarihi: 31.05.2024).
- [26] Nexgits, "Designing Intelligent Chatbots for Effective User Interaction," Jul. 14, 2023. [Çevrimiçi]. Available: <https://nexgits.com/designing-intelligent-chatbots-strategies-for-effective-user-interaction> (Erişim tarihi: 31.05.2024).
- [27] M. La Faber, "The Rise of Multimodal AI: Transforming Text, Image, and Video," May 27, 2024. [Çevrimiçi]. Available: <https://www.linkedin.com/pulse/rise-multimodal-ai-transforming-text-image-video-richard-la-faber-wklxc/> (Erişim tarihi: 31.05.2024).
- [28] A. Jaimes and N. Sebe, "Multimodal Human-Computer Interaction: A Survey," in IEEE International Workshop on Human Computer Interaction in conjunction with ICCV 2005, Beijing, China, 2005.
- [29] OpenAI, "Hello GPT-4O," 2024. [Çevrimiçi]. Available: <https://openai.com/index/hello-gpt-4o/> (Erişim tarihi: 31.05.2024).
- [30] T. Guide, "OpenAI Spring Update Event Live Blog," 2024. [Çevrimiçi]. Available: <https://www.tomsguide.com/ai/live/openai-spring-update-event-live-blog> (Erişim tarihi: 31.05.2024)

ÖZGEÇMİŞ

Adı Soyadı : Hussain Abdullatif
Doğum Yeri ve Tarihi : Suudi Arabistan – 2002
Tel : 0 (534) 338 05 93
E-Posta : hussainabdullatif@marun.edu.tr

Öğrenim Durumu

Derece	Kurum	Mezuniyet Durumu
Lise	Gönen AİHL	2017 - 2020
Lisans	Marmara Üniversitesi	2020 – Devam ediyor

İş Deneyimi

Yıl	Firma	Görevi
2021 – Present	Smartlemda	Operations / Mobile Development

ÖZGEÇMİŞ

Adı Soyadı : Mohammad Salah
Doğum Yeri ve Tarihi : Suudi Arabistan – 2002
Tel : 0 (501) 031 34 20
E-Posta : mohammadsalah@marun.edu.tr

Öğrenim Durumu

Derece	Kurum	Mezuniyet Durumu
Lise	İTO 15 Temmuz Şehitleri AİHL	2017 - 2020
Lisans	Marmara Üniversitesi	2020 – Devam ediyor

İş Deneyimi

Yıl	Firma	Görevi
2022	Wolves Interactive	Game Development
2023 – Present	Alemi Medya	Web Development