**T.C.**

**MANİSA CELAL BAYAR ÜNİVERSİTESİ**

**HASAN FERDİ TURGUTLU TEKNOLOJİ FAKÜLTESİ**

**YAZILIM MÜHENDİSLİĞİ BÖLÜMÜ**

**PROFESYONEL YAZILIM GELİŞTİRME**

**(LİSANS BİTİRME TEZİ)**

**SESLİ KOMUTLA BİLGİSAYAR KONTROLÜ**

**Muzaffer Enes Bulat**

**222802046**

**Danışman: Doç. Dr. Osman Altay**

****

**MANİSA – 2025**

**TAAHHÜTNAME**

Bu tezin Manisa Celal Bayar Üniversitesi Hasan Ferdi Turgutlu Teknoloji Fakültesi Yazılım Mühendisliği Bölümü’nde, akademik ve etik kurallara uygun olarak yazıldığını ve kullanılan tüm literatür bilgilerinin referans gösterilerek tezde yer aldığını beyan ederim.

Muzaffer Enes Bulat

# İÇİNDEKİLER

[İÇİNDEKİLER 3](#_Toc201689061)

[SİMGELER VE KISALTMALAR DİZİNİ 4](#_Toc201689062)

[TEŞEKKÜR 5](#_Toc201689063)

[ÖZET 6](#_Toc201689064)

[ABSTRACT 7](#_Toc201689065)

[1. GİRİŞ 8](#_Toc201689066)

[2. GENEL BİLGİLER 9](#_Toc201689067)

[2.1 Sesli Komut Sistemlerinin Evrimi 9](#_Toc201689068)

[2.2 Speech-to-Text ve Alternatif Yöntemler 9](#_Toc201689069)

[2.3 MFCC ile Ses Özellik Çıkarımı 9](#_Toc201689070)

[2.4 Makine Öğrenmesi ile Komut Sınıflandırma 10](#_Toc201689071)

[2.5 Gelişmiş Fonksiyonlar ve Doğal Diyalog Sistemleri 10](#_Toc201689072)

[3. MATERYAL VE YÖNTEM 11](#_Toc201689073)

[3.1 Yazılım Ortamı 11](#_Toc201689074)

[3.2 Veri Kümesi 11](#_Toc201689075)

[3.3 MFCC Öznitelik Çıkarımı 11](#_Toc201689076)

[3.4 Komut Sınıflandırma 12](#_Toc201689077)

[3.5 Ek Özellikler 12](#_Toc201689078)

[KAYNAKLAR 14](#_Toc201689079)

[ÖZGEÇMİŞ 15](#_Toc201689080)

# SİMGELER VE KISALTMALAR DİZİNİ

**AI**  Artificial Intelligence — Yapay Zeka  
**ML**  Machine Learning — Makine Öğrenmesi  
**MFCC**  Mel-Frequency Cepstral Coefficients — Mel Frekans Kepstral Katsayıları  
**RF**  Random Forest — Rastgele Orman Algoritması  
**STT**  Speech-to-Text — Sesi Metne Dönüştürme  
**CMD**  Command — Komut  
**GUI**  Graphical User Interface — Grafiksel Kullanıcı Arayüzü  
**UI**  User Interface — Kullanıcı Arayüzü  
**API**  Application Programming Interface — Uygulama Programlama Arayüzü  
**IDE**  Integrated Development Environment — Bütünleşik Geliştirme Ortamı  
**OS**  Operating System — İşletim Sistemi  
**CPU**  Central Processing Unit — Merkezi İşlem Birimi  
**WAV**  Waveform Audio File Format — WAV Ses Dosyası Biçimi  
**PCM**  Pulse-Code Modulation — Darbe Kodlu Modülasyon  
**QT**  Qt Framework (PyQt5’te kullanılan) — Qt Arayüz Çatısı  
**ACC**  Accuracy — Doğruluk Oranı  
**F1**  F1-Score — Harmonik Ortalama Skoru

# TEŞEKKÜR

Çalışmamın her aşamasında bilgi ve tecrübeleriyle yol gösteren, desteğini esirgemeyen değerli danışmanım Doç. Dr. Osman Altay’a; öğrenim hayatım boyunca maddi ve manevi olarak yanımda olan, sabır ve anlayışlarıyla her zaman destekleyen aileme; bilgi ve fikir paylaşımlarıyla katkıda bulunan tüm hocalarıma ve bu süreçte yanımda olan arkadaşlarıma en içten teşekkürlerimi sunarım.

**Muzaffer Enes Bulat**  
Manisa – 2025

# ÖZET

Bu çalışmada, kullanıcıların bilgisayarlarıyla yalnızca sesli komutlar aracılığıyla etkileşime geçebilmelerini sağlayan bir yapay zeka destekli sistem geliştirilmiştir. Sistem, geleneksel "speech-to-text" yerine doğrudan ses komutu tanıma yaklaşımını benimsemektedir. MFCC öznitelikleri çıkarılan ses verileri, Random Forest sınıflandırıcı ile analiz edilerek komut tahmini yapılmakta, tahmin sonucuna göre sistem eylemleri gerçekleştirilmektedir. Bu sistem, yazıya dökmeden doğrudan komut anlamı çıkarımı ile erişilebilirlik ve kullanım kolaylığı sağlamayı hedeflemektedir.

Anahtar Kelimeler: Sesli komut, yapay zeka, Random Forest, MFCC, PyQt5, ses tanıma

# ABSTRACT

In this study, an artificial intelligence-assisted system is developed to allow users to interact with their computers using only voice commands. The system adopts a direct command recognition approach instead of the conventional "speech-to-text" method. Extracted MFCC features from the audio input are analyzed using the Random Forest classifier, and system actions are executed based on prediction results. This system aims to improve accessibility and ease of use by deriving direct meaning from voice commands without transcription.

Keywords: Voice command, artificial intelligence, Random Forest, MFCC, PyQt5, speech recognition

# GİRİŞ

Günümüzde kullanıcı deneyimi, teknolojik sistemlerin başarısını belirleyen en kritik unsurlardan biri haline gelmiştir. Özellikle klavye ve fareye ihtiyaç duymadan, yalnızca sesli komutlarla bilgisayar ile etkileşime geçmek; hem ergonomi hem de erişilebilirlik açısından önemli avantajlar sunmaktadır. Bu bağlamda, sesli komut sistemleri yalnızca engelli bireyler için değil, tüm kullanıcılar için daha doğal bir iletişim aracı hâline gelmektedir.

Mevcut sistemlerde genellikle sesli komutlar yazıya çevrildikten sonra işlenmektedir. Ancak bu yöntem hem işlem süresini uzatmakta hem de doğal olmayan etkileşimler doğurmaktadır. Bu çalışmada amaç, doğrudan ses sinyali üzerinden anlam çıkararak komutları yorumlayabilen ve uygulayabilen bir sistem geliştirmektir. Geliştirilen sistem, herhangi bir yazıya dökme süreci olmadan sesli komutları tanımlayacak; MFCC (Mel Frekans Kepstrum Katsayıları) özellik çıkarımı ve Random Forest algoritması ile sınıflandırma gerçekleştirecektir.

Bu tez çalışmasının temel amacı, yazıya dökme adımına ihtiyaç duymadan, sesli komutları doğrudan işleyebilen yapay zekâ destekli bir bilgisayar kontrol sistemi geliştirmektir. Geliştirilen sistemde, ses verileri MFCC yöntemiyle özniteliklere dönüştürülmekte ve Random Forest algoritması ile sınıflandırılmaktadır. Kullanıcıdan alınan sesli komutlara karşılık olarak bilgisayar, PyQt5 tabanlı arayüz üzerinden doğrudan işlem gerçekleştirebilmekte, bu da sistemin hem kullanıcı dostu hem de hızlı tepki veren bir yapıda olmasını sağlamaktadır. Bu yönüyle çalışma, sesli komut sistemlerinin erişilebilirliğini artırmayı ve klasik etkileşim yöntemlerine alternatif bir çözüm sunmayı hedeflemektedir. Ayrıca sistemin doğal dil işleme (LLM), web araması, komut günlüğü, kullanıcı sesi doğrulama, güvenlik kontrolleri ve sohbet modu gibi gelişmiş yetenekleriyle yapay zekâ destekli bir asistan olarak genişletilebilirliği de hedeflenmiştir.

# 2. GENEL BİLGİLER

## 2.1 Sesli Komut Sistemlerinin Evrimi

Sesle kontrol sistemleri, özellikle erişilebilirlik ihtiyaçlarına yönelik olarak uzun süredir geliştirilmektedir. İlk örnekleri 1950’lerde Bell Labs tarafından geliştirilen sınırlı kelime tanıma sistemleridir. Günümüzde Amazon Alexa, Google Assistant, Apple Siri gibi gelişmiş sanal asistanlar bu teknolojinin geldiği noktayı göstermektedir.

## 2.2 Speech-to-Text ve Alternatif Yöntemler

Geleneksel sesli komut sistemleri genellikle 'speech-to-text' (sesi yazıya dökme) yöntemini temel alır. Ancak bu yöntemlerde metin dönüşüm hataları komutun yanlış algılanmasına neden olabilmektedir. Bu nedenle son yıllarda 'direct speech-to-command' yani doğrudan komut tanıma yaklaşımları ön plana çıkmıştır.

## 2.3 MFCC ile Ses Özellik Çıkarımı

Mel Frequency Cepstral Coefficients (MFCC), insan kulağının frekans algısını taklit eden ve konuşma tanımada yaygın kullanılan bir öznitelik çıkarım yöntemidir. MFCC, sesin spektral özelliklerini sadeleştirerek, sınıflandırma için uygun hale getirir.

## 2.4 Makine Öğrenmesi ile Komut Sınıflandırma

Ses verisinin sınıflandırılmasında kNN, SVM, Decision Tree ve Random Forest gibi makine öğrenmesi algoritmaları yaygın olarak kullanılmaktadır. Random Forest, yüksek doğruluğa sahip ve yorumlanabilir bir yapıya sahiptir. Karar ağaçlarından oluşan bu yöntem, farklı sınıflandırıcıların sonuçlarını birleştirerek tahmin gücünü artırır.

## 2.5 Gelişmiş Fonksiyonlar ve Doğal Diyalog Sistemleri

Bu proje yalnızca komut tanıma sistemi değildir; aynı zamanda aşağıdaki yetenekleri bünyesinde barındırır:

* Sesli Geri Bildirim: pyttsx3 TTS kütüphanesi ile yanıtlar üretir.
* Web Arama: Chrome üzerinden YouTube, Google gibi sitelere yönlendirme yapılabilir.
* Doğal Dil Diyaloğu: ChatGPT gibi LLM API’leri ile “2+2 kaçtır?” gibi sorulara yanıt üretilebilir.
* Sesli Not Tutma: “Hey bilgisayar, not al” komutuyla konuşmalar metne dökülebilir.
* Güvenlik Modu: Sistem, kullanıcı sesini tanıtarak yalnızca bu sesle çalışacak şekilde ayarlanabilir.
* Komut Günlüğü: Tüm komutlar zaman damgalı olarak bir .txt dosyasına kaydedilebilir.
* Kullanıcıya Hitap: Kamera yerine ses tanıma ile kullanıcının kimliği tanımlanarak ad ile hitap edilebilir.

# 3. MATERYAL VE YÖNTEM

## 3.1 Yazılım Ortamı

Python 3.10, librosa, scikit-learn, PyQt5, pyttsx3, OpenAI SDK gibi kütüphaneler kullanılmıştır. Proje Windows 11 ortamında geliştirilmiş ve test edilmiştir.

## 3.2 Veri Kümesi

9 farklı sesli komut 25 katılımcıdan toplanmıştır. Her komut için ortalama 8 örnek alınmıştır. Gürültülü varyantlarla birlikte toplam 2700 ses örneği oluşturulmuştur.

## 3.3 MFCC Öznitelik Çıkarımı

Ses örnekleri ön işleme tabi tutulmuş, MFCC, delta ve delta-delta öznitelikleri çıkarılarak tek bir vektöre dönüştürülmüştür.

## 3.4 Komut Sınıflandırma

Random Forest algoritması, ses vektörlerinin sınıflandırılmasında kullanılmıştır. GridSearchCV ile en uygun parametreler belirlenmiş, %92.4 doğruluk ve %91.8 F1 skoru elde edilmiştir.

## 3.5 Ek Özellikler

* Web üzerinden işlem: YouTube gibi platformlarda doğrudan arama yapılabilir.
* ChatGPT bağlantısı: Kullanıcıdan gelen anlamlı cümlelere doğal yanıtlar verilir.
* Komut günlüğü: Konuşmalar zaman damgalı olarak kayıt edilir.
* Sesli doğrulama: Sistemin sadece tanıtılan sesle çalışması sağlanır.
* Hitap sistemi: Kullanıcının adı tanınarak kişisel iletişim sağlanır.
* Güvenlik kontrolü: Kapama, silme gibi işlemler için ek onay istenir.

# KAYNAKLAR

1. Rabiner, L., & Juang, B. H. (1993). Fundamentals of Speech Recognition. Prentice Hall.

2. Davis, S., & Mermelstein, P. (1980)<https://doi.org/10.1109/TASSP.1980.1163420>

3. Breiman, L. (2001)<https://doi.org/10.1023/A:1010933404324>

4. Zhang, Y., & Glass, J. (2011)<https://doi.org/10.1109/ICASSP.2011.5947103>

5. Hinton, G. et al. (2012)<https://doi.org/10.1109/MSP.2012.2205597>

6. Python Software Foundation (2023)<https://docs.python.org/3/>

7. Pedregosa, F. et al. (2011)<http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html>

8. Muzaffer Enes (2025). Mairo Voice Assistant [GitHub repository] <https://github.com/muzafferenes/mairo-voice-assistant>

# ÖZGEÇMİŞ

Adı Soyadı : Muzaffer Enes Bulat

Doğum Yeri ve Yılı : Manisa, 2005

Medeni Hali : Bekar

Yabancı Dil : İngilizce

E‑posta : m.enesbulat@gmail.com  
  
Eğitim Durumu

Lise : Sami Evkuran Anadolu Lisesi, 2025