



Natural Language Processing Approach to Time Management for Meetings and Lectures: From Audio Recordings to Summary Listening

Asude Kosova^a , Soydan Serttaş^{*b}

ABSTRACT

In today's business and education world, meetings are an important tool for managing projects and business processes, but they also take up a significant amount of employees' time. To solve this problem, it can be considered as a great advantage to transcribe the audio recordings of the meetings and then summarise them. Converting summaries into audio files allows employees to master the meetings by listening only to the important points and thus time savings can be achieved. This method can both increase work efficiency and help employees maintain their work-life balance. In this study, the methods of converting audio files into text, summarising the text, and speaking the summarised text using today's technologies are investigated. Firstly, the performance of Sonix, an artificial intelligence-supported speech recognition technology, is evaluated. Then, the transcription times of audio files with five different sizes of the Whisper model presented by OpenAI, namely tiny, base, small, medium, and large, were measured and compared with the original texts. Smodin, Resoomer, Summarizing Tool, Rephrase, ChatGPT 3.5, T5 artificial intelligence models, Transformers and NLTK natural language processing libraries were used to summarise the texts. The success of these software was evaluated with Rouge-1, Rouge-2, Rouge-L and Meteor performance metrics. As a result of the evaluation, the most successful methods for English and Turkish meetings were determined.

^a Kütahya Dumlupınar University, Engineering Faculty, Dept. of Computer Engineering 43100 - Kütahya, Türkiye
ORCID: 0009-0006-7946-4124

^{b,*} Kütahya Dumlupınar University, Engineering Faculty, Dept. of Computer Engineering 43100 - Kütahya, Türkiye
ORCID: 0000-0001-8887-8675

*Corresponding author.
e-mail: soydan.serttas@dpu.edu.tr

Keywords: Natural language processing, text summarisation, text-to-audio, audio-to-text

Anahtar Kelimeler: Doğal dil işleme, metin özetleme, metni sese çevirme, sesi metne çevirme

Submitted: 01.06.2025
Revised: 29.06.2025
Accepted: 30.06.2025

doi:10.30855/ais.2025.08.01.04

Toplantılar ve Dersler için Zaman Yönetimine Doğal Dil İşleme Yaklaşımı: Ses Kayıtlarından Özet Dinlemeye

ÖZ

Günümüz iş ve eğitim dünyasında toplantılar, projelerin ve iş süreçlerinin yönetimi için önemli bir araç olsa da, çalışanların zamanını önemli ölçüde almaktadır. Bu sorunu çözmek için toplantıların ses kayıtlarının alınarak metne dönüştürülmesi ve ardından bu metinlerin özetlenmesi büyük bir avantaj olarak düşünülebilir. Özetlerin ses dosyalarına dönüştürülmesi, çalışanların sadece önemli noktaları dinleyerek toplantılara hakim olmasını sağlar ve böylece zaman tasarrufu elde edilebilir. Bu yöntem, hem iş verimliliğini artırabilir hem de çalışanların iş-yaşam dengesini korumalarına yardımcı olabilir. Bu çalışmada günümüz teknolojilerinden yararlanarak ses dosyalarını metne dönüştürme, metin özetleme ve özetlenmiş metni seslendirme yöntemleri incelenmiştir. İki farklı konuşma tanıma aracından ilk olarak, yapay zeka destekli konuşma tanıma teknolojisi olan Sonix'in performansı değerlendirilmiştir. Ardından, OpenAI tarafından sunulan Whisper modelinin beş farklı boyutu olan tiny, base, small, medium ve large ile ses dosyalarının metne dönüştürülme süreleri ölçülmüş ve orijinal metinlerle karşılaştırılmıştır. Metinlerin özetlenmesi için ise Smodin, Resoomer, Summarizing Tool, Rephrase, ChatGPT 3.5, T5 yapay zeka modelleri, Transformers ve NLTK doğal dil işleme kütüphaneleri kullanılmıştır. Bu yazılımların başarısı Rouge-1, Rouge-2, Rouge-L ve Meteor performans metrikleriyle değerlendirilmiştir. Değerlendirme sonucunda İngilizce ve Türkçe toplantılar için en başarılı bulunan yöntemler belirlenmiştir.

1. Giriş (Introduction)

Zaman yönetimi, günümüz iş dünyasında verimliliği artırmanın önemli yollarından biridir. Teknolojinin ilerlemesiyle, bireylerin zamanını etkili bir şekilde yönetmesi daha da önemli hale gelmiştir. Bu noktada yapay zeka devreye girerek iş süreçlerini optimize etme ve kolaylaştırma becerisine sahiptir. Yapay zeka tabanlı yardımcı araçlar, çalışanların toplantı planlaması, e-posta yönetimi ve günlük görevlerin otomasyonu gibi işlemleri daha hızlı ve verimli bir şekilde yapmalarına olanak tanır. Bu, çalışanların enerjilerini daha stratejik ve yaratıcı görevlere yönlendirmelerini sağlayarak iş yeri verimliliğini artırır [1,2].

Eğitim dünyasında da zaman yönetimi, öğrencilerin başarılarını şekillendiren önemli bir faktördür. Yapay zeka, öğrencilere zaman yönetimi konusunda kişiselleştirilmiş çözümler sunarak eğitimde devrim yaratmaya başlamıştır. Yapay zeka tabanlı öğrenme platformları, öğrencilerin bireysel öğrenme hızlarına göre özelleştirilmiş ders programları sunarak, öğrencilerin ders çalışma sürelerini ve yöntemlerini optimize etmektedir. Ayrıca, yapay zeka, öğrencilerin akademik performansını izleyerek ve analiz ederek, hangi alanlarda daha fazla çalışma gerektirdiğini belirlemekte ve öğrencilerin zamanlarını daha etkili bir şekilde kullanmalarına yardımcı olmaktadır [3,4].

Günümüzde, iş toplantıları çoğalmış, bilgiye erişim için eğitim ve öğrenme kaynakları çeşitlenmiş ve artmış durumdadır. Bununla birlikte, bir konuya ayrıntılı olarak vakit ayırmak ve bir eğitimi tamamlamak için geleneksel derslerin, kursların, eğitimlerin yanı sıra kayıt altına alınmış konuşmalar da oldukça yaygın hale gelmiştir. Bu kayıtlar, bir bilgiye ulaşmak veya bir eğitimi tamamlamak için zamanı etkili bir şekilde kullanmayı sağlar. Bu çalışmanın amacı, ilk aşamada toplantı ya da ders anlatılarını ses dosyasına çevirmek, sonrasında ses kayıt dosyasını metne çeviren modelleri karşılaştırarak doğru sonuca en yakın değeri veren modele ulaşmaktır. Ardından metnin anlam bütünlüğünü bozmadan özetleme yapan modelleri karşılaştırarak en doğru sonucu veren modele ulaşmayı sağlamak ve son olarak bu özeti seslendirmektir. Bu nedenle, farklı yapay zeka modellerinin performanslarının değerlendirilmesi ve karşılaştırılması büyük bir önem arz etmektedir.

2. Önceki Çalışmalar (Previous Works)

Nagdwani ve Jain, çalışmalarında [5], konuşmadan metne dönüşüm (Speech to Text, STT) için Gizli Markov Modeli (Hidden Markov Model, HMM) ve Derin Sinir Ağı (Deep Neural Network, DNN) kombinasyonunun en uygun teknik olduğunu belirtmişlerdir. Metinden konuşmaya dönüşüm (Text to Speech, TTS) için ise HMM modelini kullanarak en yüksek doğruluğa ulaşmışlardır. Bu sistem, farklı dillerde uygulanabilir olup Python'da Google Speech Recognition API, pyttsx3 ve gTTS modülleri ile gerçekleştirilebilmektedir.

Jain ve arkadaşları, çalışmalarında [6], kullanıcıdan alınan sesli girdiyi metne dönüştürüp duygu analizi yapan bir sistem geliştirmiştir. Model, konuşmayı tokenize (metni cümlelere, kelimelere veya karakterlere ayırmak) ederek kutupluluğunu ve doğasını belirlemektedir. Çalışma canlı sesli girdilerle başarılı bir şekilde çalışmakta, ancak aynı anda yalnızca bir sesli girdiyi işleyebilmektedir. Çalışmada geliştirilen sistemin performansına bakıldığında 0.8 kesinlik (precision) değeri, 0.2 duyarlılık (recall) değeri ve 0.32 F-skor sonuçları elde edilmiştir.

Madahana ve arkadaşları, çalışmalarında [7], COVID-19 pandemisi sırasında iletişimin öneminin artmasıyla, işitme engelliler için gerçek zamanlı konuşma-metin işaret dili çevirisi sağlayan yapay zeka uygulamalarını incelemişlerdir. İnceleme, yapay zeka ve makine öğrenimi teknolojilerinin işitme engelliler için kullanımının yeterince araştırılmadığını ve Afrika kıtasında bu teknolojilerin klinik kullanımında geri kaldığını ortaya koymuştur. Güney Afrika dillerinde konuşma-metin-işaret dili çevirisi yapabilen yapay zeka tabanlı çözümler geliştirilmesi hedeflenmektedir.

Saraswathi ve arkadaşları yaptıkları çalışmada [8], kullanıcılara mp4 veya wav dosyalarını yükleyebilme seçeneği sunan, çeşitli konuşma tanıyıcıları arasından seçim yapılmasına izin veren ve dosya boyutu ve uzunluğu gibi daha fazla ayrıntının gösteren bir yazılım önermişlerdir. Çalışmada, kullanıcı deneyiminin geliştirilmesi için PyQt ve Tkinter Designer gibi araçlarla grafik kullanıcı arayüzleri oluşturulabileceği belirtilmiştir. Performans sorunlarının iş parçacıkları kullanılarak çözülebileceği de ifade edilmiştir.

Basystiuk ve arkadaşları çalışmalarında [9], Harry Bot adlı bir sistem geliştirme üzerine odaklanmıştır. Sesin otomatik olarak metne dönüştürülmesini sağlayan bu sistem, yapay zeka ve özellikle RNN (Recurrent Neural Network - Yinelemeli Sinir Ağı) gibi ardışık öğrenme tekniklerini kullanmaktadır. Kullanılan Keras ve PyTorch kütüphaneleri temelli modeller, öğrenme süresi 4150 milisaniye ile 5800 milisaniye arasında, eğitim döngüleri ise 400 ile 650 arasında değişmektedir. Kaybın katsayısı Keras için 0.0027 ile 0.0021 arasındadır. Kullanılan Keras Kütüphanesi temelli modeller, mevcut veri setleriyle daha verimli sonuçlar vermiştir.

Sharma ve Sardana çalışmasında [10], bir konuşma tanıma sistemi geliştirmiş ve performansını artırmak için Kalman filtresi kullanmışlardır. Kendi veritabanlarını oluşturdukları ve TIDIGIT veri setiyle karşılaştırdıkları gerçek zamanlı testlerde, sistemleri farklı gürültü koşullarında değişen bir performans göstermiştir. Ev ortamında fan gürültüsü ile %100 doğruluk sağlanırken, laboratuvar ortamında %80 başarı elde edilmiştir. Korelasyon değerleri, 20dB SNR (Signal to Noise Ratio- Sinyal-Gürültü Oranı) seviyelerinde yüksek olup %97.8'e kadar çıkarken, 5dB SNR seviyelerinde ise %77.2'ye düşmüştür. Bu değerlendirmeler, sistemin yüksek SNR seviyelerinde ve düşük gürültü koşullarında yüksek performans gösterdiğini, ancak düşük SNR ve karmaşık gürültü koşullarında performansın düştüğünü ortaya koymaktadır. Kalman filtresi arka plan gürültüsünü etkili bir şekilde gidermiş, ancak gürültüyü filtrelemenin zaman aldığı ve sürekli konuşma alımında metni göstermenin uzun sürdüğünü belirtmiştir.

Elakkiya ve arkadaşlarının çalışmasında [11], hızla gelişen bilgisayar bilimi alanlarından biri olan konuşma metin sentezinin önemli bir role sahip olduğu belirtilmiştir. Gizli Markov Model yaklaşımının kullanımının, STT operasyonlarının etkinliğini artırırken, Derin Sinir Ağı tabanlı Python'da Google'ın Konuşma Tanıma API paketiyle daha kaliteli metin ve konuşma üretimi sağladığı ifade edilmiştir. Mevcut yönteme göre noktalama işaretlerinin dikkate alınması önemlidir. Programları, Amerikan İngilizcesi için optimize edilmiş bir konuşma-metin motoruna bağlanmıştır. DNN-HMM modeli, diğer algoritmalarla kıyasla en yüksek doğruluk oranlarına ulaşmıştır. Eğitim doğruluğu %92.4, test doğruluğu %93.6 ve doğrulama doğruluğu %95.2 olarak ölçülmüştür. Çalışma gelecekte Nijerya dili için yerel olarak optimize edilmiş motorların geliştirilmesini amaçlamaktadır.

Widyassari ve arkadaşları çalışmalarında [12], SLR (Systematic Literature Review - Sistemantik Literatür Taraması) yöntemiyle metin özetleme alanındaki son araştırmaları sunmuşlardır. Makine öğrenimi, istatistikler ve bulanık mantık gibi çeşitli teknikler kullanılarak özetleme performansı geliştirilmeye çalışılmıştır. MMI (Maximal Marginal Importance - Maksimal Marjinal Önem), Parçacık Sürü Optimizasyonu (Particle Swarm Optimization, PSO) ve bulanık mantık yöntemlerinin hibrit olarak kullanıldığı bir çalışmada, duyarlılık değeri 0.40463 ve f-skör değeri 0.42291 olarak elde edilmiştir. Ancak, bu yöntem, diğer özetleme sistemleri karşısında kesinlik açısından daha düşük bir performans göstermiştir.

Thange ve arkadaşları çalışmalarında [13], metin özetleme alanında BERT, sinir ağları ve Pegasus gibi modellerin önemli rolleri olduğunu vurgulamış ve çeşitli metin özetleme yaklaşımlarının farklı alanlardaki uygulamalarını incelemiştir. Bu modellerin, metin özetleme görevlerini geliştirmede önemli bir rol oynadıkları ve doğal dil işleme alanında önemli bir bileşen olarak kalacaklarını belirtmişlerdir.

Yadav ve arkadaşları çalışmalarında [14], metin özetlemenin temel tekniklerini ve bu tekniklerin performansını incelemektedir. Çalışma, hem cümle seçerek özetleme (extractive) hem de yorumlayarak özetleme (abstractive) yöntemlerini ele alır ve her iki yöntemin avantajları ile sınırlamaları arasında bir karşılaştırma sunar. Hibrit özetleme yöntemleri, bu iki yaklaşımın birleşimini sunarak her iki yöntemin de güçlü yönlerini birleştirir. SUMMARIST sistemi, insan özetlerinden %30 daha yüksek performans sergilerken, BERT tabanlı özetleme yaklaşımları Rouge-1 ve Rouge-2 değerlerinde sırasıyla 43.16 ve 20.32'lik başarılar elde etmiştir.

Çalışır Zenci çalışmasında [15], öğretmen adaylarının metin özetleme stratejilerini incelemiştir. Çalışmada yüksek, orta ve düşük düzeyde tercih edilen stratejiler belirlenmiştir. Yüksek tercih edilen stratejiler arasında özgün metni tamamen okuma, uygun bağlaç kullanımı ve özgün metnin yapısını dikkate alarak özet metni oluşturma yer alırken, orta düzeyde tercih edilen stratejilerde özet metnin ilk cümlesinde özgün metnin ana fikrini kullanma ve özgün metinde geçen örnekleri özet metinde belirtmeme görülmüştür. Düşük düzeyde tercih edilen stratejiler arasında ise özgün metnin yazarına

gönderme yapma yer almaktadır. Cinsiyetin ve bölümün özetleme stratejileri üzerinde belirgin bir etkisinin olmadığı belirlenmiştir.

Kaynar ve arkadaşları çalışmalarında [16], Türkçe metinler için özellik tabanlı bir doküman özetleme sistemi önermişlerdir. Sistem, her bir dokümana ait cümlelerin ayrıştırılması ve 8 farklı özniteliğin çıkartılmasıyla oluşturulmuştur. Bu cümleler, Genetik Algoritma (GA - Genetic Algorithm) kullanılarak 80 veriyle eğitilmiş ve her bir özniteliğin optimum ağırlığı bulunmuştur. GA'nın elde ettiği optimum ağırlıklar, 40 test verisi üzerinde denenmiş olup Rouge-1 (0.828) ve Rouge-2 (0.744) skorları elde edilmiştir. Ayrıca, sistem ağırlıklandırması kullanmadan sadece öznitelik skorlarının toplanarak yapılan bir test de gerçekleştirilmiş ve bu işlem sonucunda elde edilen ağırlıksız öznitelikler, yine başarılı skorlar (Rouge-1: 0.725, Rouge-2: 0.626) elde etmiştir. Tüm bu sonuçlar incelendiğinde, Genetik Algoritma ile belirlenen ağırlıkların sistem başarısını önemli ölçüde artırdığı görülmüştür.

Ranjan'ın çalışmasında [17], metni dört farklı şekilde ('mutlu', 'üzgün', 'kızgın' ve 'nötr') sınıflandırabilen bir model geliştirilmiştir. Bu model, birleşik veri seti üzerinde %86.77 doğruluk oranına sahiptir. Kelime gömme ve RNN kullanılarak metindeki ilişkiler korunmuş ve duygusal analiz yapılmıştır. Sinir ağı, cümlelerdeki kelimeler arasındaki ilişkileri anlayarak metindeki duyguları belirlemeye yardımcı olmuştur. Önerilen yöntem, metni duygu temelli konuşmaya dönüştürmek için kullanılabilir ve özellikle hikâye anlatma uygulamaları ve sesli kitaplar için faydalı olabilir. Bu çalışmanın, özellikle görsel olarak engelli bireyler için metinden duyguları analiz etmede önemli bir araç olabileceği belirtilmiştir.

Chowdhury ve Hussan çalışmasında [18], farklı TTS teknolojilerinin avantajlarını, sınırlamalarını ve evrimini incelemişlerdir. Derin öğrenmedeki ilerlemelerin TTS sistemlerinin kalitesini artırdığını belirten çalışma, sentezlenmiş konuşmada duygusal ifade eksikliği gibi bazı zorlukların hala ele alınması gerektiğini vurgulamıştır.

Tamboli ve arkadaşları tarafından yapılan çalışmada [19] geliştirilen bir metin belgesi okuma yazılımı olan TTS projesi özetlenmektedir. Windows tabanlı bu uygulama, metin dosyalarını okur ve kullanıcıya telaffuz eder. Özellikle görsel olarak engelli bireyler veya sürekli ekrana bakmak istemeyenler için tasarlanmıştır. Önerilen tasarım, bilgisayarın sadece metin belgelerini okuyarak hızlı yanıtlar vermesini sağlar. TTS modeli, ön uç ve arka uç olmak üzere iki bölümden oluşur: Ön uç, sembollerin ve kısaltmaların yazılı sözlere dönüştürülmesini sağlar, arka uç ise bu sembolik dönüşümü ses olarak çıktıya dönüştürür.

Shastri ve Vishwakarma çalışmalarında [20], İngilizceden Hintçe'ye metin dönüşümünü ve ardından metinden konuşmaya geçişi araştırmışlardır. Çalışma, görüntülerden metin tanıma, metin karakter tanıma, geometrik filtreleme, yazım denetimi ve optik karakter tanıma gibi bir dizi teknik ve yöntemi içermektedir. Naive Bayes ve Karar Ağaçları algoritmaları yazım denetimi amacıyla kullanılmıştır. Bu yöntemlerle elde edilen sonuçlar, görüntülerden metin çıkarma işleminde %100 doğruluk oranı sağlandığını ve yüksek performans elde edildiğini göstermektedir. Önerilen yöntem TTS teknolojisini kullanarak metni Hintçe'ye dönüştürmekte ve bu alanda önemli bir gelişme sağlamaktadır.

3. Materyal ve Metot (Material and Method)

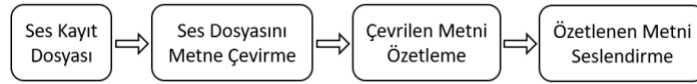
Bu çalışmada Avrupa'ya matbaayı tanıtan Alman Johannes Gutenberg hakkındaki 193 kelimelik bir İngilizce metin [21] ve Ziya Gökalp'in 152 kelimelik "İlim" adlı Türkçe metni [22] kullanılmıştır. Kullanılan ses dosyaları, bu metinlerin 'Google Translate' seslendirmesi ile kaydedilmiş halidir. İngilizce metnin ses kayıt süresi 1 dakika 27 saniye iken Türkçe metnin ses kayıt süresi 1 dakika 43 saniyedir.

Ses dosyasını metne dönüştürme işlemi için iki farklı araç incelenmiştir. İlk incelememiz yapay zeka destekli konuşma tanıma teknolojisi olan "Sonix"tir. Bir STT platformu olan Sonix 50'den fazla dili destekler ve API'si aracılığıyla ses veya video dosyalarını hızlıca metne dönüştürür. Bunun yanı sıra kanal bazlı konuşmacı tespiti yapabilir ve transkript üzerinde otomatik düzeltilmiş zaman damgaları içeren altyazı üretimi, konuşmacı ayrımı gibi gelişmiş özellikler sunar [23]. 30 dakikalık ücretsiz deneme sürümü mevcuttur. Sonix ile kaydettiğimiz İngilizce ve Türkçe ses dosyalarının metne çevrilme süreleri ölçülmüştür. Ardından oluşturulan metin ile orijinal metnin Rouge-1, Rouge-2, Rouge-L ve Meteor değerlendirme metriklerine göre karşılaştırılması yapılmıştır.

İkinci incelememiz Whisper modelidir. Whisper, OpenAI'nin sunduğu genel amaçlı bir konuşma tanıma modelidir. Çeşitli seslerden oluşan geniş bir veri kümesi üzerinde eğitilmiş olan Whisper'ın beş farklı model boyutu bulunmaktadır ve bu modeller içerdikleri parametreler, mevcut donanım hızları gibi faktörlere bağlı olarak değişir. Bu projede Whisper'ın beş farklı model boyutu olan tiny, base, small, medium ve large ile hem İngilizce hem Türkçe ses dosyalarını ne kadar sürede metne dönüştürdükleri incelenmiş ve orijinal metin ile Rouge-1, Rouge-2, Rouge-L ve Meteor değerlendirme metriklerine göre karşılaştırılması yapılmıştır.

Metinlerin özetlenmesi için 8 farklı model incelenmiştir. Bu modeller ile metin özetleme yaparken farklı yüzdelik değerler denenmiş ve orijinal özet ile Rouge-1, Rouge-2, Rouge-L ve Meteor değerlendirme metriklerine göre karşılaştırılması yapılmıştır. İncelenen yapay zeka destekli araçlar ve doğal dil işleme kütüphaneleri; Smodin/Extractive3, Smodin/Extractive4, Smodin/Abstractive3, Smodin/Abstractive4, Resoomer %30- %50- %70, Summarizing Tool %35- %50- %70, Rephrase %25- %50- %75, ChatGPT3.5 %30- %50- %70, NLTK, T5 Small- Base- Large ve HuggingFace Transformers 'dır. Bu özetleme araçlarındaki yüzdelikler "metnin ne kadar kısaltılacağını" ya da "ne kadar dönüştürüleceğini" ifade etmektedir. Örneğin %30 özetleme orijinal metnin yaklaşık %30'u kadar uzunlukta bir özet oluşturur. Yani daha yoğun ve kısa bir özet verir.

Metinlerin seslendirmesi için de çeşitli yapay zeka araçları incelenmiş ve kullanıcıya uygun seçenekler sunulmuştur. Şekil 1'de ses kayıtlarından özet dinlemeye çalışmasının akış diyagramı gösterilmiştir.



Şekil 1. Çalışmanın akış diyagramı
(Flow diagram of the study)

3.1. Sesten metne dönüşüm aşamasında incelenen modeller (Models analyzed in the voice to text transformation phase)

Whisper: Whisper, genel amaçlı bir konuşma tanıma modelidir. Çeşitli seslerden oluşan geniş bir veri kümesi üzerinde eğitilmiştir ve aynı zamanda çok dilli konuşma tanıma, konuşma çevirisi ve dil tanımlama işlemlerini gerçekleştirebilen çok görevli bir modeldir. Parametre ve hızı bağlı beş farklı boyutu vardır: tiny, base, small, medium ve large. Aşağıdaki Tablo 1'de modellerin boyutları, parametreleri, gerekli hafıza alanı ve hızları belirtilmiştir.

Tablo 1. Whisper modeli içeriği (Whisper model content)

Whisper Model Çeşidi	Parametreler	RAM	Hız
<i>Tiny</i>	39 milyon	1GB	-32x
<i>Base</i>	74 milyon	1GB	-16x
<i>Small</i>	244 milyon	2GB	-6x
<i>Medium</i>	769 milyon	5GB	-2x
<i>Large</i>	1550 milyon	10GB	-1x

Sonix: Sonix, ses ve video içeriğini kopyalamak için kullanılan ve sestem metne dönüşüm yapan bir araçtır. 30 dakikalık ücretsiz deneme sürümü mevcuttur. Çeşitli dil seçenekleri de bulunmaktadır.

3.2. Metin özetleme aşamasında incelenen modeller (Models analyzed in text summarization phase)

Metin özetlemek için incelenen modeller aşağıda listelenmiştir:

- Smodin: Smodin, yazma ve araştırmanın tüm yönlerine yardımcı olmak için çeşitli araçlar sağlayan, metin özetleme de yapabilen yapay zeka destekli bir yazma asistanıdır.
- Resoomer (%30, %50, %70): Resoomer, metin özetleme işlemi yapmamızı sağlayan çevrim içi bir araçtır. %30, %50, %70 oranlarıyla özetleme seçeneği sunar.
- Summarizing Tool (%35, %50, %70): Summarizing Tool, yapay zeka desteğiyle metinleri anlamını bozmadan özetleyen ücretsiz bir çevrimiçi araçtır. Özet uzunluğu ayarlanabilir ve birden fazla dili destekler. %35- %50- %70 oranlarıyla özetleme seçeneği sunar.
- Rephrase (%25, %50, %75): Rephrase, verilen metnin orijinal anlamını koruyarak farklı kelimelerle yeniden ifade eden yapay zeka destekli bir araçtır.
- ChatGPT3.5 (%30, %50, %70): OpenAI tarafından geliştirilen, yapay zekâya dayalı bir sohbet

robotudur.

- NLTK (Natural Language Toolkit): Python programlama dilinde yazılmış sembolik ve istatistiksel doğal dil işleme için bir dizi kütüphane ve programdır.
- T5 (small, base, large): T5, denetlenmiş ve denetlenmemiş görevlerin çok görevli bir karışımı üzerinde önceden eğitilmiş ve her görevin metinden metne biçimine dönüştürüldüğü bir kodlayıcı-kod çözücü modelidir.
- Transformers: Transformers, önceden eğitilmiş modelleri kolayca indirip eğitmek için API'ler ve araçlar sağlar. Önceden eğitilmiş modellerin kullanılması, bilgi işlem maliyetlerini ve karbon ayak izini azaltabilir ve bir modeli sıfırdan eğitmek için gereken zamandan ve kaynaklardan tasarruf edilmesini sağlar.

3.3. Metin seslendirme (Text to speech)

Metni seslendirmek için kullanılabilecek yapay zeka modelleri ve kütüphaneleri aşağıda listelenmiştir.

- Murf AI: Reklamlar, açıklayıcı videolar, YouTube videoları, podcastler ve benzeri medya ürünleri için yüksek kaliteli ses üretmek amacıyla yapay zeka kullanan gelişmiş bir ses oluşturma aracıdır.
- Speech Gen: Çeşitli dil seçenekleri ile yazılan metinleri seslendiren web tabanlı, yapay zeka destekli araçtır.
- Descript: Yazdığımız yazılarla, videoları ve podcastleri düzenlememize yarayan ve metin seslendirme yapma gibi çeşitli yapay zeka düzenleme araçları sunan araçtır.
- Genny Lovo: Metinden seslendirme yapma seçeneği ile birlikte sesi klonlamak ve klonlanmış sesle oynamamızı sağlamak gibi birçok özelliği mevcut olan yapay zeka platformudur.
- Audiosonic: Metinsel içerikleri konuşmaya dönüştüren, yapay zeka destekli ücretli bir platformdur.
- gTTS kütüphanesi: gTTS (Google Text-to-Speech) bir python kütüphanesidir ve Google Translate'in konuşma sentezleyici komut satır arayüzüdür.
- Pyttsx3 kütüphanesi: Pyttsx3, konuşma sentezlemek için kullanılan Python kütüphanesidir. Diğer kütüphanelerden farklı olarak çevrimdışı olarak çalışır.

3.4. Performans değerlendirme metrikleri (Performance evaluation metrics)

Performans değerlendirme metrikleri, sesten metne dönüşüm ve metin özetleme aşamalarında kullanılan modellerin doğruluğunu ve etkinliğini ölçmek için kullanılmaktadır. Rouge ve Meteor, metin özetleme aşamasında, oluşturulan özetlerin referans özetlerle ne kadar uyumlu olduğunu değerlendirirken; kesinlik, duyarlılık ve F1 skoru, sesten metne dönüşüm aşamasında, modelin tahminlerinin doğruluğunu ve eksiksizliğini ölçmek amacıyla kullanılmaktadır.

- ROUGE (Recall-Oriented Understudy for Gisting Evaluation): Otomatik olarak oluşturulan özetlerin referans özetlerle ne kadar uyumlu olduğunu değerlendirmek için kullanılan bir metrik sistemidir. Rouge-1, bireysel kelimelerin referans metinle eşleşme oranını ölçerken, Rouge-2, iki kelimedenden oluşan grupların eşleşme frekansını değerlendirir. Rouge-L ise iki metin arasındaki en uzun ortak kelime dizisini esas alarak ölçüm yapar. Bu metrikler, özetlerin içerik uyumunu ve doğruluğunu belirlemek için yaygın olarak kullanılır.
- Meteor (Metric for Evaluation of Translation with Explicit Ordering): Makine çevirisi kalitesini değerlendirmek için kullanılan bu metrik, unigram eşleştirme kavramına dayanır ve kelime kökleri, eşanlamlılar ve esnek sıralama gibi daha geniş eşleştirme özelliklerini dikkate alarak insan referans çevirileriyle karşılaştırmalar yapar.
- Kesinlik (Precision (p)): Modelin pozitif olarak tahmin ettiği değerlerin ne kadarının gerçekten doğru olduğunu ölçer.
- Duyarlılık (Recall (r)): Aslında pozitif olan tüm değerlerin ne kadarını modelin doğru şekilde pozitif olarak tahmin ettiğini ölçer.
- F-Skor: Kesinlik ve duyarlılık değerlerinin harmonik ortalamasıdır.

4. Bulgular ve Tartışma (Findings and Discussion)

Modeller test edilirken Avrupa'ya matbaayı tanıtan Alman Johannes Gutenberg hakkındaki İngilizce bir

metin ve Ziya Gökalp'in "İlim" adlı Türkçe metni kullanılmıştır. Kullanılan ses dosyaları bu metinlerin Google Translate tarafından seslendirilmiş biçimleridir. İngilizce metnin ses kayıt dosyası 1 dakika 27 saniye, Türkçe metnin ses kayıt dosyası 1 dakika 43 saniyedir. Sekiz farklı model ile özetleme yapılmış ve bu modellerin farklı yüzdeliklerde oluşturdukları özetler denenmiştir.

4.1. Ses dosyasından metne dönüşüm (Audio-to-text conversion)

Bu çalışmada kullandığımız İngilizce metnin kelime sayısı 193, Türkçe metnin kelime sayısı 152'dir. Çalışmanın bu kısmında orijinal metinlerin Google Translate ile seslendirilmiş ses dosyaları kullanılarak Sonix, Whisper-Tiny, Whisper-Base, Whisper-Small, Whisper-Medium, Whisper-Large modelleri ile metne çevrilmiştir. Çevrilen metinlerde oluşan kelime sayıları orijinal metinlerle karşılaştırılmıştır.

4.1.1. İngilizce ses dosyasından metne dönüşüm için kullanılan modellerin değerlendirilmesi (Evaluation of models for English audio-to-text conversion)

Tablo 2'de Sonix, Whisper-Tiny, Whisper-Base, Whisper Small, Whisper-Medium, Whisper-Large modellerinin, İngilizce metnin ses dosyasından çevirdikleri metinlerin kelime sayıları verilmiştir. Orijinal İngilizce metnin kelime sayısı 193'tür. Tablo 2'de görüldüğü üzere kullandığımız İngilizce metnin kelime sayısına ulaşan model Whisper Medium'dur.

Tablo 2. Modeller ve çevirdikleri kelime sayıları (Models and the number of words they converted)

Yapay Zeka / Model	Kelime sayısı
<i>Sonix</i>	198
<i>Whisper-Tiny</i>	195
<i>Whisper-Base</i>	195
<i>Whisper- Small</i>	196
<i>Whisper- Medium</i>	193
<i>Whisper- Large</i>	219

Tablo 3'te İngilizce ses dosyasını metne çeviren modellerin oluşturduğu metinler ile orijinal metnin kıyaslanması sonucu oluşan Rouge-1 değerleri verilmiştir. Tablo 3'te görüldüğü üzere modellerin ses dosyasından ürettiği metinlerin orijinal metin ile en yüksek eşleşme gösteren modeller; 0.992 Rouge-1 değeri ile Sonix, Whisper-Base, Whisper-Small, Whisper-Medium modelleridir.

Tablo 3. İngilizce ses dosyasını metne dönüştüren modellerin Rouge-1 değerleri (Rouge-1 values of models that convert English audio-to-text)

Yapay Zeka /Model	Duyarlılık değeri (r)	Kesinlik değeri (p)	Rouge-1 (F-Skor)
<i>Sonix</i>	0.994	0.989	0.992
<i>Whisper-Tiny</i>	0.989	0.984	0.987
<i>Whisper-Base</i>	0.994	0.989	0.992
<i>Whisper- Small</i>	0.994	0.989	0.992
<i>Whisper- Medium</i>	0.994	0.989	0.992
<i>Whisper- Large</i>	0.984	0.873	0.926

Tablo 4'te İngilizce ses dosyasını metne çeviren modellerin oluşturduğu metinler ile orijinal metnin kıyaslanması sonucu oluşan Rouge-2 değerleri verilmiştir. Tablo 4'te görüldüğü üzere modellerin ses dosyasından ürettiği metinlerin orijinal metin ile en yüksek eşleşme gösteren modeller; 0.982 Rouge-2 değeri ile Sonix, Whisper-Base, Whisper-Small, Whisper-Medium modelleridir.

Tablo 4. İngilizce ses dosyasını metne dönüştüren modellerin Rouge-2 değerleri (Rouge-2 values of models that convert English audio-to-text)

Yapay Zeka /Model	Duyarlılık değeri (r)	Kesinlik değeri (p)	Rouge-2 (F-Skor)
<i>Sonix</i>	0.984	0.979	0.982
<i>Whisper-Tiny</i>	0.974	0.969	0.972
<i>Whisper-Base</i>	0.984	0.979	0.982
<i>Whisper- Small</i>	0.984	0.979	0.982
<i>Whisper- Medium</i>	0.984	0.979	0.982
<i>Whisper- Large</i>	0.974	0.864	0.916

Tablo 5'te İngilizce ses dosyasını metne çeviren modellerin oluşturduğu metinler ile orijinal metnin kıyaslanması sonucu oluşan Rouge-L değerleri verilmiştir. Tablo 5'te görüldüğü üzere modellerin ses dosyasından ürettiği metinlerin orijinal metin ile en yüksek eşleşme gösteren modeller; 0.992 Rouge-L değeri ile Sonix, Whisper-Base, Whisper-Small, Whisper-Medium modelleridir.

Tablo 5. İngilizce ses dosyasını metne dönüştüren modellerin Rouge-L değerleri (Rouge-L values of models that convert English audio-to-text)

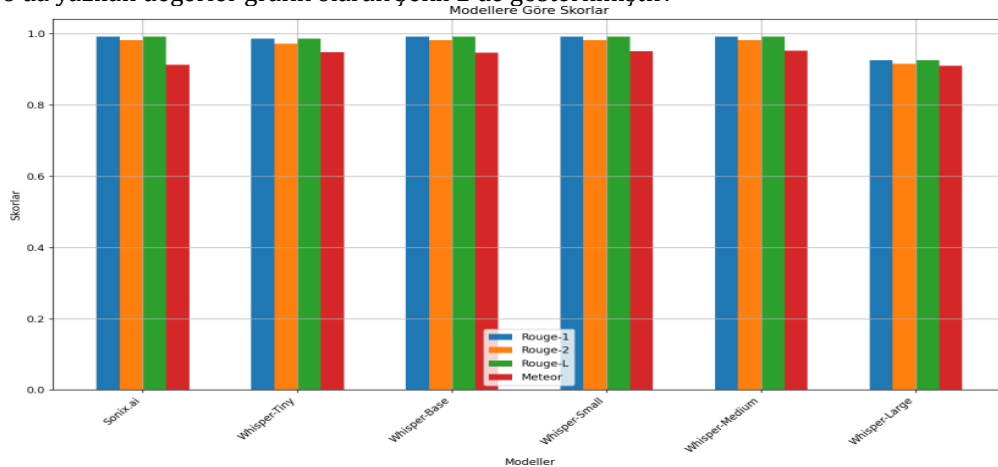
Yapay Zeka /Model	Duyarlılık değeri (r)	Keskinlik değeri (p)	Rouge-L (F-Skor)
Sonix	0.994	0.989	0.992
Whisper-Tiny	0.989	0.984	0.987
Whisper-Base	0.994	0.989	0.992
Whisper- Small	0.994	0.989	0.992
Whisper- Medium	0.994	0.989	0.992
Whisper- Large	0.984	0.873	0.926

Tablo 6’da İngilizce ses dosyasını metne çeviren modellerin oluşturduğu metinler ile orijinal metnin kıyaslanması sonucu oluşan Rouge-1, Rouge-2, Rouge-L ve METEOR değerleri verilmiştir. Tablo 6’da görüldüğü üzere modellerin Rouge-1, Rouge-2, Rouge-L ve METEOR sonuçları karşılaştırıldığında en yüksek eşleşme gösteren modeller; 0.992 Rouge-1, Rouge-2, Rouge-L değerleri ile Sonix, Whisper-Base, Whisper-Small, Whisper-Medium modelleridir.

Tablo 6. İngilizce ses dosyasını metne dönüştüren modellerin Rouge-1, Rouge-2, Rouge-L ve Meteor değerleri karşılaştırması (Comparison of Rouge-1, Rouge-2, Rouge-L and Meteor values of the models that convert English audio-to-text)

Yapay Zeka /Model	Rouge-1	Rouge-2	Rouge-L	METEOR
Sonix	0.992	0.982	0.992	0.913
Whisper-Tiny	0.987	0.972	0.987	0.949
Whisper-Base	0.992	0.982	0.992	0.947
Whisper- Small	0.992	0.982	0.992	0.951
Whisper- Medium	0.992	0.982	0.992	0.953
Whisper- Large	0.926	0.916	0.926	0.911

Tablo 6’da yazılan değerler grafik olarak Şekil 2’de gösterilmiştir.

Şekil 2. Modellere göre alınan skorlar grafiği
(Graph of scores according to models)

4.1.2. Türkçe ses dosyasından metne dönüşüm için kullanılan modellerin değerlendirilmesi (Evaluation of models for Turkish audio-to-text conversion)

Tablo 7’de Sonix, Whisper-Tiny, Whisper-Base, Whisper-Small, Whisper-Medium, Whisper-Large modellerinin, Türkçe metnin ses dosyasından çevirdikleri metinlerin kelime sayıları verilmiştir. Orijinal Türkçe metnin kelime sayısı 152’dir. Tablo 7’de görüldüğü üzere kullandığımız metnin kelime sayısına ulaşan modeller Sonix ve Whisper-Small olmuştur.

Tablo 7. Modeller ve çevirdikleri kelime sayıları (Models and the number of words they converted)

Yapay Zeka / Model	Kelime sayısı
Sonix	152
Whisper-Tiny	157
Whisper-Base	153
Whisper- Small	152
Whisper- Medium	151
Whisper- Large	149

Tablo 8’de Türkçe ses dosyasını metne çeviren modellerin oluşturduğu metinler ile orijinal metnin kıyaslanması sonucu oluşan Rouge-1 değerleri verilmiştir. Tablo 8’de görüldüğü üzere modellerin ses

dosyasından ürettiği metinlerin orijinal metin ile en yüksek eşleşme gösteren model, 0.968 Rouge-1 değeri ile Whisper-Medium modelidir.

Tablo 8. Türkçe ses dosyasını metne dönüştüren modellerin Rouge-1 değerleri (Rouge-1 values of models that convert Turkish audio-to-text)

Yapay Zeka /Model	Duyarlılık değeri (r)	Kesinlik değeri (p)	Rouge-1 (F-Skor)
<i>Sonix</i>	0.946	0.954	0.950
<i>Whisper-Tiny</i>	0.838	0.827	0.832
<i>Whisper-Base</i>	0.883	0.875	0.879
<i>Whisper- Small</i>	0.955	0.963	0.959
<i>Whisper- Medium</i>	0.964	0.972	0.968
<i>Whisper- Large</i>	0.950	0.972	0.961

Tablo 9'da Türkçe ses dosyasını metne çeviren modellerin oluşturduğu metinler ile orijinal metnin kıyaslanması sonucu oluşan Rouge-2 değerleri verilmiştir. Tablo 9'da görüldüğü üzere modellerin ses dosyasından ürettiği metinlerin orijinal metin ile en yüksek eşleşme gösteren model, 0.938 Rouge-2 değeri ile Whisper-Large modelidir.

Tablo 9. Türkçe ses dosyasını metne dönüştüren modellerin Rouge-2 değerleri (Rouge-2 values of models that convert Turkish audio-to-text)

Yapay Zeka /Model	Duyarlılık değeri (r)	Kesinlik değeri (p)	Rouge-2 (F-Skor)
<i>Sonix</i>	0.909	0.918	0.914
<i>Whisper-Tiny</i>	0.693	0.684	0.689
<i>Whisper-Base</i>	0.783	0.776	0.780
<i>Whisper- Small</i>	0.927	0.936	0.932
<i>Whisper- Medium</i>	0.936	0.945	0.941
<i>Whisper- Large</i>	0.927	0.949	0.938

Tablo 10'da Türkçe ses dosyasını metne çeviren modellerin oluşturduğu metinler ile orijinal metnin kıyaslanması sonucu oluşan Rouge-L değerleri verilmiştir. Tablo 10'da görüldüğü üzere modellerin ses dosyasından ürettiği metinlerin orijinal metin ile en yüksek eşleşme gösteren model, 0.968 Rouge-L değeri ile Whisper-Medium modelidir.

Tablo 10. Türkçe ses dosyasını metne dönüştüren modellerin Rouge-L değerleri (Rouge-L values of models that convert Turkish audio-to-text)

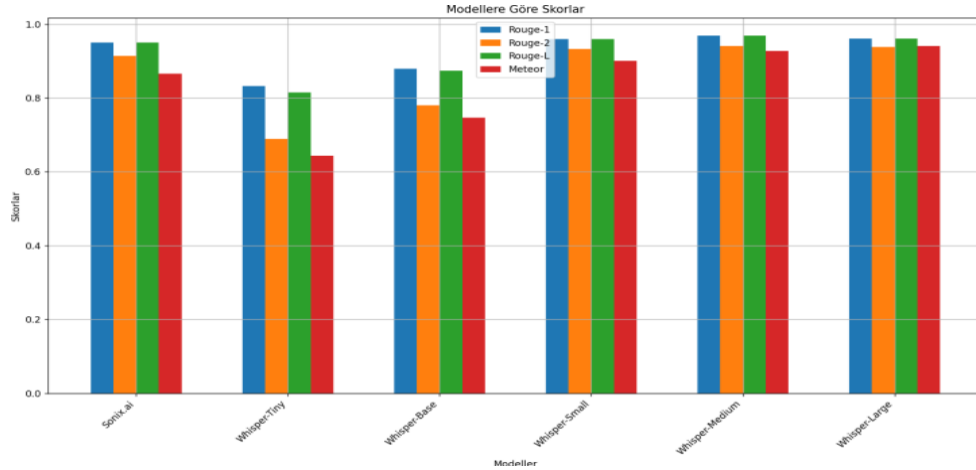
Yapay Zeka /Model	Duyarlılık değeri (r)	Kesinlik değeri (p)	Rouge-L (F-Skor)
<i>Sonix</i>	0.946	0.954	0.950
<i>Whisper-Tiny</i>	0.820	0.809	0.815
<i>Whisper-Base</i>	0.878	0.871	0.874
<i>Whisper- Small</i>	0.955	0.963	0.959
<i>Whisper- Medium</i>	0.964	0.972	0.968
<i>Whisper- Large</i>	0.950	0.972	0.961

Tablo 11'de Türkçe ses dosyasını metne çeviren modellerin oluşturduğu metinler ile orijinal metnin kıyaslanması sonucu oluşan Rouge-1, Rouge-2, Rouge-L ve Meteor değerleri verilmiştir. Tablo 11'de görüldüğü üzere modellerin Rouge-1, Rouge-2, Rouge-L ve Meteor sonuçları karşılaştırıldığında en yüksek eşleşme gösteren model, 0.968 Rouge-1, Rouge-L değerleri ile Whisper-Medium modelidir.

Tablo 11. Türkçe ses dosyasını metne dönüştüren modellerin Rouge-1, Rouge-2, Rouge-L ve Meteor değerleri karşılaştırması (Comparison of Rouge-1, Rouge-2, Rouge-L and Meteor values of the models that convert Turkish audio-to-text)

Yapay Zeka /Model	Rouge-1	Rouge-2	Rouge-L	METEOR
<i>Sonix</i>	0.950	0.914	0.950	0.866
<i>Whisper-Tiny</i>	0.832	0.689	0.815	0.643
<i>Whisper-Base</i>	0.879	0.780	0.874	0.747
<i>Whisper- Small</i>	0.959	0.932	0.959	0.900
<i>Whisper- Medium</i>	0.968	0.941	0.968	0.927
<i>Whisper- Large</i>	0.961	0.938	0.961	0.941

Tablo 11'de yazılan değerler grafik olarak Şekil 3'te gösterilmiştir.



Şekil 3. Modellere göre alınan skorlar grafiği
(Graph of scores according to models)

4.1.3. İngilizce ve Türkçe ses dosyasından metne dönüşümü için kullanılan modellerin değerlendirilmesi (Evaluation of models for English and Turkish audio-to-text transcription)

Tablo 12’de ses dosyasını metne çevirmede kullandığımız modellerin hem İngilizce hem Türkçe metinlerin ses kayıtlarını ne kadar sürede metne çevirdikleri ve performans değerlendirme metrik sonuçları gösterilmektedir. Tablo 12’de görüldüğü üzere hem Türkçe hem İngilizce metinlerde Whisper modelinde boyut arttıkça ses dosyasının metne dönüşüm işlemi süresi de artmıştır.

Whisper-Tiny en kısa sürede çıktı veren model olmasına rağmen en düşük doğruluk oranını (0.987) veren modeldir. Whisper-Base, Whisper-Small, Whisper-Medium modelleri boyut arttıkça metne dönüşüm işlemi süresi artmaktadır fakat Whisper-Tiny’den daha yüksek doğruluk oranı vermektedirler. Whisper-Large ise hem metne dönüşüm işlemi süresi en uzun (4 dakika 28 saniye) hem de Whisper-Base, Whisper-Small, Whisper-Medium modellerinden daha düşük doğruluk oranına (0.926) sahiptir.

Süre ve performansı birlikte değerlendirdiğimizde Sonix diğer modellere göre daha uygundur. Performans değerlendirme metriklerinden ise en yüksek sonuçları veren Rouge-1 ve Rouge-2’dir.

Tablo 12. Türkçe ve İngilizce metinler için modellerin ses dosyasını metne dönüştürme süreleri ve en yüksek doğruluk değerleri (Models’ audio-to-text conversion times for Turkish and English texts, and highest accuracy values)

Yapay Zeka /Model	İngilizce		Türkçe	
	Süre (sn)	En Yüksek Değer	Süre (sn)	En Yüksek Değer
Sonix	25	0.987 Rouge -1, Rouge -L	21	0.832 Rouge -1
Whisper-Tiny	22	0.992 Rouge -1, Rouge -L	30	0.879 Rouge -1
Whisper-Base	55	0.992 Rouge -1, Rouge -L	70	0.959 Rouge -1, Rouge -L
Whisper- Small	153	0.992 Rouge -1, Rouge -L	207	0.968 Rouge -1, Rouge -L
Whisper- Medium	268	0.926 Rouge -1, Rouge -L	337	0.961 Rouge -1, Rouge -L
Whisper- Large	78	0.992 Rouge-1, Rouge-L	90	0.950 Rouge -1, Rouge -L

4.2. Metin Özetleme (Text summarization)

Tablo 13’te İngilizce metninden özetleme yaptırdığımız modellerin oluşturduğu özetlerin kelime sayısı bulunmaktadır. Smodin yapay zeka destekli platformunun sunduğu 3-4 cümlelik seçenekleri ile cümle seçerek özetleme (extractive) ve yorumlayarak özetleme (abstractive) sonuçları yer almaktadır. Resoomer, Summarizing Tool, Rephrase, ChatGPT3.5 modellerinin kendi içlerinde farklı yüzdeliklerdeki özetlemeleri; NLTK, HuggingFace Transformers kütüphaneleri; T5 modelinin small, base, large boyutlarıyla yapılan özetlemeleri yer almaktadır. Orijinal özetin kelime sayısı 61’dir.

Tablo 13 incelediğinde; cümle seçerek özetleme (extractive) yöntemi ile yapılan özetlemenin daha

düşük cümleli özeti, ideal özet kelime sayısına daha yakın olduğu görülmektedir. Modeller arasındaki yüzdelerle baktığımızda, Resoomer, Summarizing Tool ve ChatGPT3.5'un %70'lik yaklaşımının, Rephrase modelindeki %25'lik yaklaşıma eş değere yakın aralıkta olduğu görülmektedir. İdeal özet kelime sayısına en yakın kelime sayısına sahip modellerin Smodin/Extractive-3, Resoomer %30 ve T5/Base olduğu görülmektedir.

Tablo 13. Modellerin özetlediği İngilizce metinlerin kelime sayıları (Word counts of English texts summarized by models)

Yapay Zeka/ Model	Kelime Sayısı
<i>Smodin/ Extractive-3</i>	65
<i>Smodin/ Extractive-4</i>	98
<i>Smodin/ Abstractive-3</i>	68
<i>Smodin/ Abstractive-4</i>	94
<i>Resoomer %30</i>	62
<i>Resoomer %50</i>	88
<i>Resoomer %70</i>	141
<i>Summarizing Tool %35</i>	75
<i>Summarizing Tool %50</i>	118
<i>Summarizing Tool %75</i>	140
<i>Rephrase %25</i>	155
<i>Rephrase %50</i>	108
<i>Rephrase %75</i>	56
<i>ChatGPT3.5 %30</i>	114
<i>ChatGPT3.5 %50</i>	100
<i>ChatGPT3.5 %70</i>	79
<i>NLTK /3cümlelik özet</i>	54
<i>T5 / Small</i>	53
<i>T5 / Base</i>	61
<i>T5 / Large</i>	38
<i>Transformers</i>	54

Tablo 14'te İngilizce orijinal özet ile modellerin oluşturdukları özetlerin kıyaslanması sonucu oluşan Rouge-1 değerleri verilmiştir. Modellerin ürettiği özetler ve metinlerin ideal özetleri arasında en yüksek benzerlik oranına sahip olan model 0.527 Rouge-1 değeri ile Resoomer %30'luk modelidir.

Tablo 14. İngilizce metin özetlemek için kullanılan modellerin Rouge-1 değerleri (Rouge-1 values of the models used to summarize English text)

Yapay Zeka /Model	Duyarlılık değeri (r)	Kesinlik değeri (p)	Rouge-1 (F-Skor)
<i>Smodin/ Extractive-3</i>	0.460	0.453	0.456
<i>Smodin/ Extractive-4</i>	0.587	0.373	0.456
<i>Smodin/ Abstractive-3</i>	0.523	0.471	0.496
<i>Smodin/ Abstractive-4</i>	0.619	0.397	0.484
<i>Resoomer %30</i>	0.539	0.515	0.527
<i>Resoomer %50</i>	0.603	0.413	0.490
<i>Resoomer %70</i>	0.730	0.317	0.442
<i>Summarizing Tool %35</i>	0.507	0.410	0.453
<i>Summarizing Tool %50</i>	0.730	0.377	0.497
<i>Summarizing Tool %75</i>	0.761	0.333	0.463
<i>Rephrase %25</i>	0.777	0.312	0.445
<i>Rephrase %50</i>	0.619	0.351	0.448
<i>Rephrase %75</i>	0.460	0.5	0.479
<i>ChatGPT3.5 %30</i>	0.666	0.655	0.464
<i>ChatGPT3.5 %50</i>	0.571	0.346	0.431
<i>ChatGPT3.5 %70</i>	0.476	0.365	0.413
<i>NLTK /3cümlelik özet</i>	0.412	0.472	0.440
<i>T5 / Small</i>	0.269	0.320	0.293
<i>T5 / Base</i>	0.444	0.444	0.444
<i>T5 / Large</i>	0.301	0.5	0.376
<i>Transformers</i>	0.476	0.588	0.526

Tablo 15'te İngilizce orijinal özet ile modellerin oluşturdukları özetlerin kıyaslanması sonucu oluşan Rouge-2 değerleri verilmiştir. Modellerin ürettiği özetler ve metinlerin ideal özetleri arasında en yüksek benzerlik oranına sahip olan model 0.211 Rouge-2 değeri ile Rephrase %25'lik modelidir.

Tablo 15. İngilizce metin özetlemek için kullanılan modellerin Rouge-2 değerleri (Rouge-2 values of the models used to summarize English text)

Yapay Zeka /Model	Duyarlılık değeri (r)	Kesinlik değeri (p)	Rouge-2 (F-Skor)
<i>Smodin/ Extractive-3</i>	0.145	0.142	0.144
<i>Smodin/ Extractive-4</i>	0.177	0.112	0.137
<i>Smodin/ Abstractive-3</i>	0.145	0.130	0.137

<i>Smodin/ Abstractive-4</i>	0.193	0.123	0.150
<i>Resoomer %30</i>	0.112	0.107	0.110
<i>Resoomer %50</i>	0.177	0.120	0.143
<i>Resoomer %70</i>	0.306	0.131	0.184
<i>Summarizing Tool %35</i>	0.161	0.129	0.143
<i>Summarizing Tool %50</i>	0.274	0.140	0.185
<i>Summarizing Tool %75</i>	0.322	0.139	0.195
<i>Rephrase %25</i>	0.370	0.147	0.211
<i>Rephrase %50</i>	0.193	0.109	0.139
<i>Rephrase %75</i>	0.048	0.052	0.050
<i>ChatGPT3.5 %30</i>	0.193	0.102	0.134
<i>ChatGPT3.5 %50</i>	0.161	0.097	0.121
<i>ChatGPT3.5 %70</i>	0.080	0.061	0.069
<i>NLTK /3cümlelik özet</i>	0.129	0.148	0.137
<i>T5 / Small</i>	0.064	0.076	0.070
<i>T5 / Base</i>	0.096	0.096	0.096
<i>T5 / Large</i>	0.096	0.162	0.121
<i>Transformers</i>	0.177	0.22	0.196

Tablo 16’da İngilizce orijinal özet ile modellerin oluşturdukları özetlerin kıyaslanması sonucu oluşan Rouge-L değerleri verilmiştir. Modellerin ürettiği özetler ve metinlerin ideal özetleri arasında en yüksek benzerlik oranına sahip olan model 0.348 Rouge-L değeri ile Resoomer %50’lik modelidir.

Tablo 16. İngilizce metin özetlemek için kullanılan modellerin Rouge-L değerleri (Rouge-L values of the models used to summarize English text)

Yapay Zeka /Model	Duyarlılık değeri (r)	Kesinlik değeri (p)	Rouge-L (F-Skor)
<i>Smodin/ Extractive-3</i>	0.317	0.312	0.314
<i>Smodin/ Extractive-4</i>	0.365	0.232	0.283
<i>Smodin/ Abstractive-3</i>	0.365	0.328	0.345
<i>Smodin/ Abstractive-4</i>	0.365	0.234	0.285
<i>Resoomer %30</i>	0.333	0.318	0.325
<i>Resoomer %50</i>	0.428	0.293	0.348
<i>Resoomer %70</i>	0.507	0.220	0.307
<i>Summarizing Tool %35</i>	0.285	0.230	0.255
<i>Summarizing Tool %50</i>	0.412	0.213	0.281
<i>Summarizing Tool %75</i>	0.460	0.201	0.280
<i>Rephrase %25</i>	0.492	0.197	0.281
<i>Rephrase %50</i>	0.349	0.198	0.252
<i>Rephrase %75</i>	0.269	0.293	0.280
<i>ChatGPT3.5 %30</i>	0.412	0.220	0.287
<i>ChatGPT3.5 %50</i>	0.365	0.221	0.275
<i>ChatGPT3.5 %70</i>	0.285	0.219	0.248
<i>NLTK /3cümlelik özet</i>	0.238	0.272	0.254
<i>T5 / Small</i>	0.190	0.226	0.206
<i>T5 / Base</i>	0.285	0.285	0.285
<i>T5 / Large</i>	0.206	0.342	0.257
<i>Transformers</i>	0.253	0.313	0.280

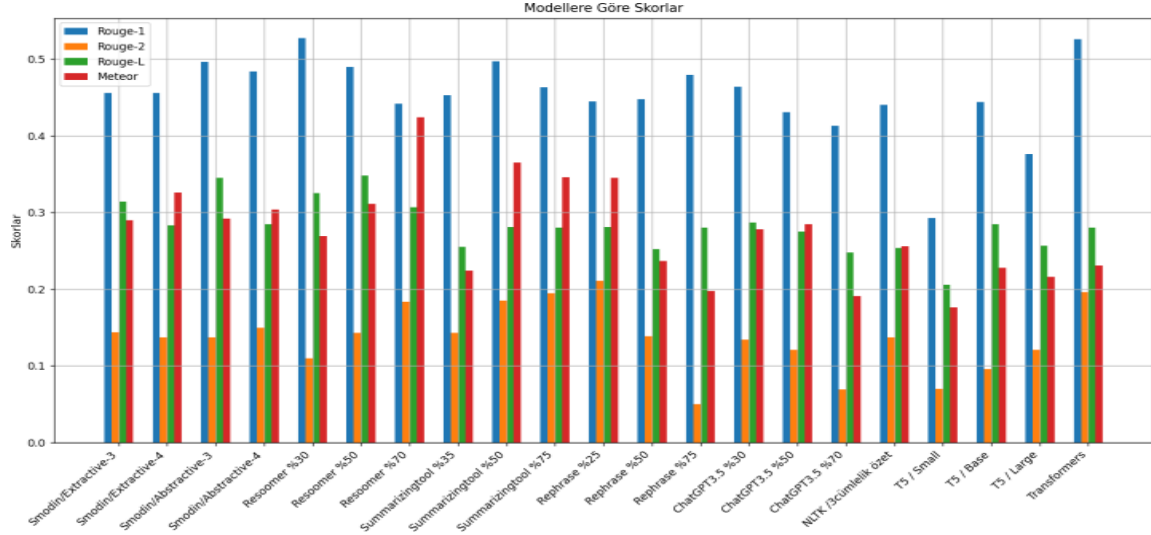
Tablo 17’de İngilizce orijinal özet ile modellerin oluşturdukları özetlerin kıyaslanması sonucu oluşan Rouge-1, Rouge-2, Rouge-L ve Meteor değerleri verilmiştir. Tablo 17’de görüldüğü üzere Rouge-1, Rouge-2, Rouge-L ve Meteor değerleri karşılaştırıldığında en yüksek eşleşmelere sahip olan modeller 0.527 Rouge-1 değeri ile Resoomer %30 ve 0.526 Rouge-1 değeri ile Transformers modelleridir.

Tablo 17. İngilizce metin özetlemek için kullanılan modellerin Rouge-1, Rouge-2, Rouge-L ve Meteor değerleri karşılaştırması (Comparison of Rouge-1, Rouge-2, Rouge-L and Meteor values of the models used for English text summarization)

Yapay Zeka /Model	Rouge-1	Rouge-2	Rouge-L	METEOR
<i>Smodin/ Extractive-3</i>	0.456	0.144	0.314	0.29
<i>Smodin/ Extractive-4</i>	0.456	0.137	0.283	0.326
<i>Smodin/ Abstractive-3</i>	0.496	0.137	0.345	0.292
<i>Smodin/ Abstractive-4</i>	0.484	0.150	0.285	0.304
<i>Resoomer %30</i>	0.527	0.110	0.325	0.269
<i>Resoomer %50</i>	0.490	0.143	0.348	0.311
<i>Resoomer %70</i>	0.442	0.184	0.307	0.424
<i>Summarizing Tool %35</i>	0.453	0.143	0.255	0.224
<i>Summarizing Tool %50</i>	0.497	0.185	0.281	0.365
<i>Summarizing Tool %75</i>	0.463	0.195	0.280	0.346
<i>Rephrase %25</i>	0.445	0.211	0.281	0.345
<i>Rephrase %50</i>	0.448	0.139	0.252	0.237
<i>Rephrase %75</i>	0.479	0.050	0.280	0.198
<i>ChatGPT3.5 %30</i>	0.464	0.134	0.287	0.278
<i>ChatGPT3.5 %50</i>	0.431	0.121	0.275	0.285

ChatGPT3.5 %70	0.413	0.069	0.248	0.191
NLTK /3cümlelik özet	0.440	0.137	0.254	0.256
T5 / Small	0.293	0.070	0.206	0.176
T5 / Base	0.444	0.096	0.285	0.228
T5 / Large	0.376	0.121	0.257	0.216
Transformers	0.526	0.196	0.280	0.231

Tablo 17’de yazılan değerler grafik olarak Şekil 4’te gösterilmiştir.



Şekil 4. Modellere göre alınan skorlar grafiği
(Graph of scores according to models)

Tablo 18’de Türkçe metninden özetleme yaptırdığımız modellerin oluşturduğu özetlerin kelime sayısı bulunmaktadır. Smodin yapay zeka destekli platformunun sunduğu 3-4 cümlelik seçenekleri ile cümle seçerek özetleme (extractive) ve yorumlayarak özetleme (abstractive) sonuçları yer almaktadır.

Resoomer, Summarizing Tool, Rephrase, ChatGPT3.5 modellerinin kendi içlerinde farklı yüzdeliklerdeki özetlemeleri; NLTK, HuggingFace Transformers kütüphaneleri; T5 modelinin 3 farklı boyut modeli (small, base, large) ile özetlemeleri yer almaktadır. Orijinal özetin kelime sayısı 33’tür.

Tablo 18 incelediğinde; ideal özetin kelime sayısına en yakın kelime sayısına sahip modellerin Resoomer %30 ve Summarizing Tool %35 olduğu görülmektedir. Modeller arasındaki yüzdeliklere baktığımızda, Resoomer, Summarizing Tool ve ChatGPT3.5’un %70’lik yaklaşımının, Rephrase modelindeki %25’lik yaklaşıma eş değere yakın aralıkta olduğu görülmektedir.

Tablo 18. Modellerin özetlediği Türkçe metinlerin kelime sayıları (Word counts of Turkish texts summarized by models)

Yapay Zeka/ Model	Kelime Sayısı
Smodin/ Extractive-3	26
Smodin/ Extractive-4	70
Smodin/ Abstractive-3	56
Smodin/ Abstractive-4	63
Resoomer %30	35
Resoomer %50	72
Resoomer %70	115
Summarizing Tool %35	34
Summarizing Tool %50	59
Summarizing Tool %75	93
Rephrase %25	109
Rephrase %50	93
Rephrase %75	56
ChatGPT3.5 %30	83
ChatGPT3.5 %50	80
ChatGPT3.5 %70	97
NLTK /3cümlelik özet	23
T5 / Small	22
T5 / Base	24
T5 / Large	11
Transformers	21

Tablo 19’da Türkçe orijinal özet ile modellerin oluşturdukları özetlerin kıyaslanması sonucu oluşan Rouge-1 değerleri verilmiştir. Modellerin ürettiği özetler ve metinlerin ideal özetleri arasında en yüksek benzerlik oranına sahip olan model 0.444 Rouge-1 değeri ile T5-Base modelidir.

Tablo 19. Türkçe metin özetlemek için kullanılan modellerin Rouge-1 değerleri (Rouge-1 values of the models used to summarize Turkish text)

Yapay Zeka /Model	Duyarlılık değeri (r)	Kesinlik değeri (p)	Rouge-1 (F-Skor)
<i>Smodin/ Extractive-3</i>	0.235	0.315	0.269
<i>Smodin/ Extractive-4</i>	0.372	0.177	0.240
<i>Smodin/ Abstractive-3</i>	0.549	0.345	0.424
<i>Smodin/ Abstractive-4</i>	0.509	0.279	0.361
<i>Resoomer %30</i>	0.274	0.233	0.252
<i>Resoomer %50</i>	0.294	0.140	0.189
<i>Resoomer %70</i>	0.431	0.128	0.198
<i>Summarizing Tool %35</i>	0.333	0.354	0.343
<i>Summarizing Tool %50</i>	0.392	0.243	0.300
<i>Summarizing Tool %75</i>	0.490	0.196	0.280
<i>Rephrase %25</i>	0.431	0.138	0.209
<i>Rephrase %50</i>	0.372	0.143	0.207
<i>Rephrase %75</i>	0.372	0.220	0.277
<i>ChatGPT3.5 %30</i>	0.529	0.216	0.306
<i>ChatGPT3.5 %50</i>	0.529	0.226	0.317
<i>ChatGPT3.5 %70</i>	0.529	0.184	0.274
<i>NLTK /3cümlelik özet</i>	0.156	0.242	0.190
<i>T5 / Small</i>	0.098	0.185	0.128
<i>T5 / Base</i>	0.444	0.444	0.444
<i>T5 / Large</i>	0.176	0.473	0.257
<i>Transformers</i>	0.196	0.370	0.256

Tablo 20’de Türkçe orijinal özet ile modellerin oluşturdukları özetlerin kıyaslanması sonucu oluşan Ruge-2 değerleri verilmiştir. Modellerin ürettiği özetler ve metinlerin ideal özetleri arasında en yüksek benzerlik oranına sahip olan modeller 0.153 Rouge-2 değerleri ile Smodin/Abstractive-3 ve Smodin/Abstractive-4 modelleridir.

Tablo 20. Türkçe metin özetlemek için kullanılan modellerin Rouge-2 değerleri (Rouge-2 values of the models used to summarize Turkish text)

Yapay Zeka /Model	Duyarlılık değeri (r)	Kesinlik değeri (p)	Rouge-2 (F-Skor)
<i>Smodin/ Extractive-3</i>	0.02	0.027	0.022
<i>Smodin/ Extractive-4</i>	0.04	0.018	0.025
<i>Smodin/ Abstractive-3</i>	0.2	0.125	0.153
<i>Smodin/ Abstractive-4</i>	0.08	0.043	0.153
<i>Resoomer %30</i>	0.04	0.033	0.036
<i>Resoomer %50</i>	0.04	0.018	0.025
<i>Resoomer %70</i>	0.06	0.017	0.027
<i>Summarizing Tool %35</i>	0.02	0.021	0.020
<i>Summarizing Tool %50</i>	0.08	0.049	0.061
<i>Summarizing Tool %75</i>	0.1	0.039	0.056
<i>Rephrase %25</i>	0.08	0.025	0.038
<i>Rephrase %50</i>	0.04	0.015	0.022
<i>Rephrase %75</i>	0.04	0.023	0.029
<i>ChatGPT3.5 %30</i>	0.14	0.056	0.08
<i>ChatGPT3.5 %50</i>	0.14	0.059	0.083
<i>ChatGPT3.5 %70</i>	0.14	0.048	0.071
<i>NLTK /3cümlelik özet</i>	0.02	0.031	0.024
<i>T5 / Small</i>	0.0	0.0	0.0
<i>T5 / Base</i>	0.02	0.033	0.025
<i>T5 / Large</i>	0.02	0.055	0.029
<i>Transformers</i>	0.02	0.038	0.026

Tablo 21’de Türkçe orijinal özet ile modellerin oluşturdukları özetlerin kıyaslanması sonucu oluşan Rouge-L değerleri verilmiştir. Modellerin ürettiği özetler ve metinlerin ideal özetleri arasında en yüksek benzerlik oranına sahip olan modeli 0.222 Rouge-L değeri ile Summarizing Tool %35’lik modelidir.

Tablo 21. Türkçe metin özetlemek için kullanılan modellerin Rouge-L değerleri (Rouge-L values of the models used to summarize Turkish text)

Yapay Zeka /Model	Duyarlılık değeri (r)	Kesinlik değeri (p)	Rouge-L (F-Skor)
<i>Smodin/ Extractive-3</i>	0.156	0.210	0.179
<i>Smodin/ Extractive-4</i>	0.235	0.112	0.151

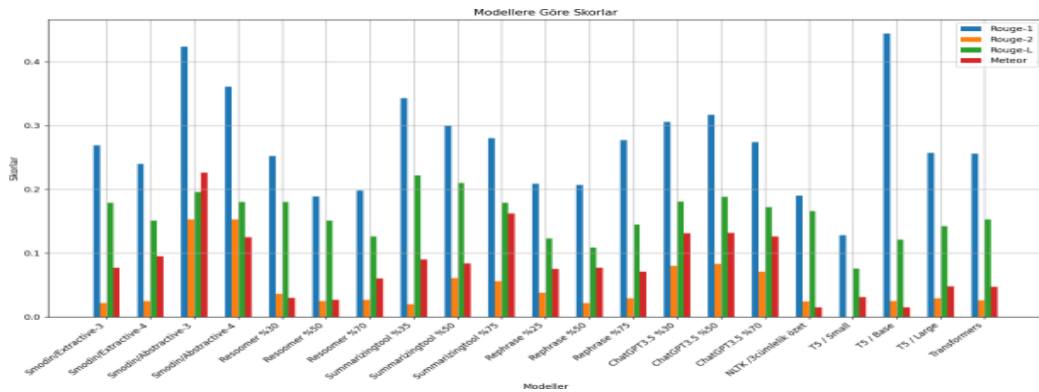
<i>Smodin/ Abstractive-3</i>	0.254	0.160	0.196
<i>Smodin/ Abstractive-4</i>	0.254	0.139	0.180
<i>Resoomer %30</i>	0.196	0.166	0.180
<i>Resoomer %50</i>	0.235	0.112	0.151
<i>Resoomer %70</i>	0.274	0.081	0.126
<i>Summarizing Tool %35</i>	0.215	0.229	0.222
<i>Summarizing Tool %50</i>	0.274	0.170	0.210
<i>Summarizing Tool %75</i>	0.313	0.125	0.179
<i>Rephrase %25</i>	0.254	0.081	0.123
<i>Rephrase %50</i>	0.196	0.075	0.109
<i>Rephrase %75</i>	0.196	0.116	0.145
<i>ChatGPT3.5 %30</i>	0.313	0.128	0.181
<i>ChatGPT3.5 %50</i>	0.313	0.134	0.188
<i>ChatGPT3.5 %70</i>	0.333	0.116	0.172
<i>NLTK /3cümlelik özet</i>	0.137	0.212	0.166
<i>T5 / Small</i>	0.058	0.111	0.076
<i>T5 / Base</i>	0.098	0.161	0.121
<i>T5 / Large</i>	0.098	0.263	0.142
<i>Transformers</i>	0.117	0.222	0.153

Tablo 22’de Türkçe orijinal özet ile modellerin oluşturdukları özetlerin kıyaslanması sonucu oluşan Rouge-1, Rouge-2, Rouge-L ve Meteor değerleri verilmiştir. Tablo 22’de görüldüğü üzere Rouge-1, Rouge-2, Rouge-L ve Meteor değerleri karşılaştırıldığında en yüksek benzerlik oranına sahip olan model 0.444 Rouge-1 değeri ile T5-Base modelidir.

Tablo 22. Türkçe metin özetlemek için kullanılan modellerin Rouge-1, Rouge-2, Rouge-L ve Meteor değerlerinin karşılaştırmaları (Comparison of Rouge-1, Rouge-2, Rouge-L and Meteor values of the models used for summarizing Turkish text)

Yapay Zeka /Model	Rouge-1	Rouge-2	Rouge-L	METEOR
<i>Smodin/ Extractive-3</i>	0.269	0.022	0.179	0.077
<i>Smodin/ Extractive-4</i>	0.240	0.025	0.151	0.095
<i>Smodin/ Abstractive-3</i>	0.424	0.153	0.196	0.226
<i>Smodin/ Abstractive-4</i>	0.361	0.153	0.180	0.125
<i>Resoomer %30</i>	0.252	0.036	0.180	0.030
<i>Resoomer %50</i>	0.189	0.025	0.151	0.027
<i>Resoomer %70</i>	0.198	0.027	0.126	0.060
<i>Summarizing Tool %35</i>	0.343	0.020	0.222	0.090
<i>Summarizing Tool %50</i>	0.300	0.061	0.210	0.084
<i>Summarizing Tool %75</i>	0.280	0.056	0.179	0.162
<i>Rephrase %25</i>	0.209	0.038	0.123	0.075
<i>Rephrase %50</i>	0.207	0.022	0.109	0.077
<i>Rephrase %75</i>	0.277	0.029	0.145	0.071
<i>ChatGPT3.5 %30</i>	0.306	0.08	0.181	0.131
<i>ChatGPT3.5 %50</i>	0.317	0.083	0.188	0.132
<i>ChatGPT3.5 %70</i>	0.274	0.071	0.172	0.126
<i>NLTK /3cümlelik özet</i>	0.190	0.024	0.166	0.015
<i>T5 / Small</i>	0.128	0.0	0.076	0.031
<i>T5 / Base</i>	0.444	0.025	0.121	0.015
<i>T5 / Large</i>	0.257	0.029	0.142	0.048
<i>Transformers</i>	0.256	0.026	0.153	0.047

Tablo 22’de yazılan değerler grafik olarak Şekil 5’te gösterilmiştir.



Şekil 5. Modellere göre alınan skorlar grafiği
(Graph of scores according to models)

5. Sonuç ve Öneriler (Conclusions and Recommendations)

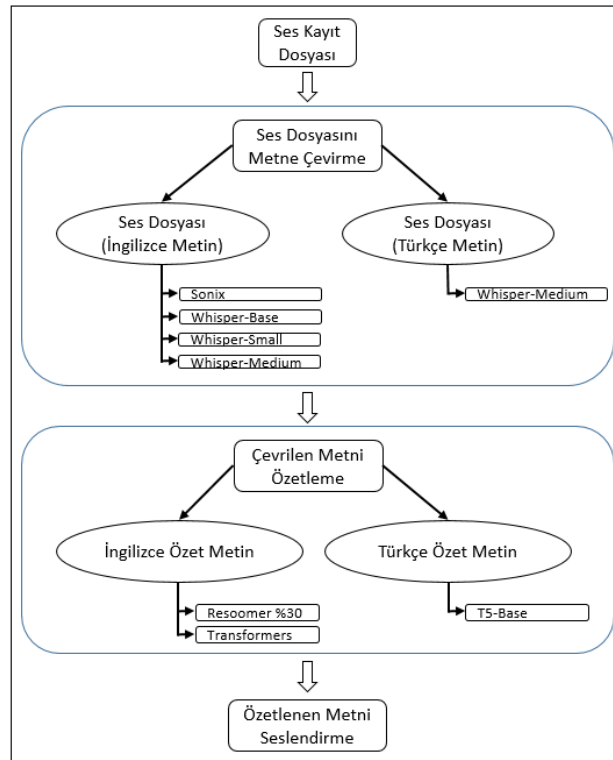
Bu çalışmada, günümüzde bilgiye erişim ve eğitim kaynaklarının çeşitlenmesiyle ortaya çıkan ses kayıtlarının metne dönüştürülmesi ve ardından özetlenerek seslendirilmesi süreci incelenmiştir. Önerilen bu yöntemin kullanımıyla, zamanın daha etkili bir şekilde kullanılması ve bilgiye erişim kolaylaştırılmak hedeflenmektedir.

Bu çalışmada sırasıyla ses dosyasının metne dönüştürülmesi, metnin özetlenmesi ve özetlenen metnin seslendirilmesi aşamalarında hem İngilizce metin hem de Türkçe metin kullanılarak kapsamlı bir araştırma yapılmıştır. Çıkan sonuçlar Rouge-1, Rouge-2, Rouge-L ve meteor performans değerlendirme metrikleri ile değerlendirilmiştir.

Ses dosyasını metne çevirme aşamasında İngilizce metin için en yüksek eşleşme gösteren modeller; 0.992 Rouge-1, Rouge-2, Rouge-L değerleri ile Sonix, Whisper-Base, Whisper-Small ve Whisper-Medium modelleri olmuştur. Türkçe metin için ise en yüksek eşleşme gösteren model, 0.968 Rouge-1, Rouge-L değerleri ile Whisper-Medium modeli olmuştur.

Metin özetleme aşamasında İngilizce metin için en yüksek benzerlik oranına sahip olan iki model, 0.527 Rouge-1 değeriyle Resoomer (%30'luk özet) modeli ve 0.526 Rouge-1 değeriyle Transformers modelidir. Türkçe metin için ise en yüksek benzerlik oranına sahip olan model 0.444 Rouge-1 değeri ile T5-Base modeli olmuştur.

Şekil 6'da, bu çalışma sonucunda en yüksek doğruluk oranını bulan modellerden oluşan yapı gösterilmiştir.



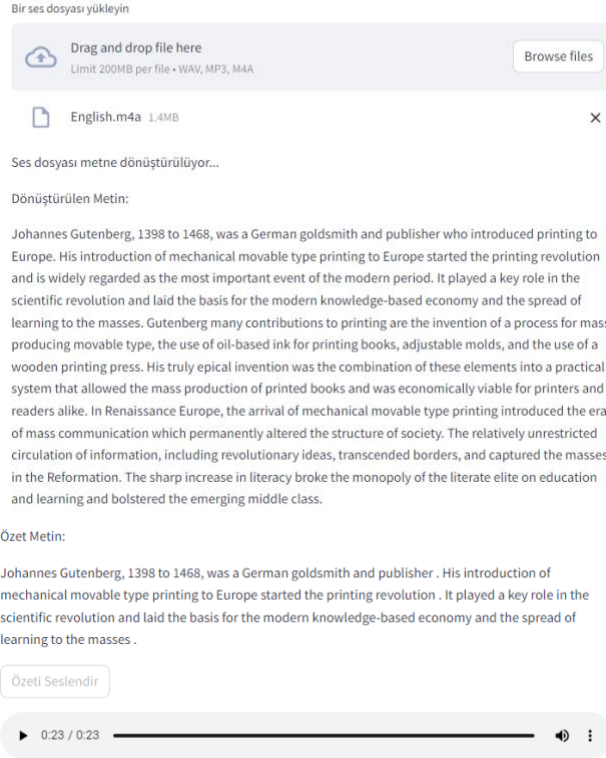
Şekil 6. En iyi sonuçları veren modellerden oluşan yapı
(Structure consisting of models that give the best results)

Bu sonuçlardan faydalanılarak çıkan en yüksek değerler kullanılarak seçtiğimiz ses dosyasını önce metne çeviren ardından özetleyip seslendirme işlemi yapan bir arayüze sahip web sitesi de oluşturulmuştur. Web sayfası görüntüleri Şekil 7'de gösterilmektedir.

Mevcut çalışmamızın sonrasında metin özetleme aşamasında daha doğru ve anlamlı özetler üretebilen modeller geliştirilebilir. Ayrıca, seslendirme aşamasında daha doğal ve anlaşılır bir ses üreten modellerin araştırılması ve geliştirilmesi gerekmektedir.

Bununla birlikte, farklı dil ve aksanlarda daha etkili sonuçlar elde etmek için modellerin çok dilli ve çeşitli veri kümeleriyle eğitilmesi gerekmektedir. Ayrıca, daha geniş kullanım senaryolarına uygun olarak, ses dosyalarının metne dönüştürülmesi ve özetlenmesi yöntemlerinin pratik uygulamaları geliştirilebilir. Bu tür gelişmeler, bilgiye erişimi ve iletişimi kolaylaştırmak için ses teknolojilerinin daha etkin bir şekilde kullanılmasını sağlayacaktır.

Ses Dosyasının Metne Dönüştürülüp Özetlenmesi ve Seslendirilmesi



Şekil 7. Web sayfası görüntüleri
(Web page views)

Çıkar Çatışması Beyanı (Conflict of Interest Statement)

Yazarlar tarafından herhangi bir çıkar çatışması bildirilmemiştir.

Kaynaklar (References)

- [1] J. Manyika, M. Chui, M. Miremadi, J. Bughin, K. George, P. Willmott, M. Dewhurst, A Future That Works: Automation, Employment, and Productivity, McKinsey & Company, New York, USA, 2017.
- [2] T. H. Davenport & R. Ronanki, "Artificial Intelligence for the Real World", Harvard Business Review, 96(1), 108-116, 2018.
- [3] W. Holmes, M. Bialik, & C. Fadel, Artificial Intelligence in Education: Promises and Implications for Teaching and Learning, Center for Curriculum Redesign, Boston, MA, USA, 2019.
- [4] R. Luckin, W. Holmes, M. Griffiths, & L. B. Forcier, Intelligence Unleashed: An Argument for AI in Education, Pearson Education, London, 2016.
- [5] S. Nagdewani & A. Jain, "A Review on Methods for Speech-to-Text and Text-to-Speech Conversion", International Research Journal of Engineering and Technology (IRJET), 7(5), 4459-4464, 2020.
- [6] N. Jain, M. Goyal, A. Gupta, V. Kumar, "Speech-to-Text Conversion and Sentiment Analysis on Speaker Specific Data", International Research Journal of Modernization in Engineering, Technology and Science, 3(6), 3050-3057, 2021.
- [7] M. C. Madahana, K. Khoza-Shangase, N. Moroe, D. Mayombo, O. Nyandoro, J. Ekoru, "A Proposed Artificial Intelligence-Based Real-Time Speech-to-Text to Sign Language Translator for South African Official Languages for the COVID-19 Era and Beyond: In

Pursuit of Solutions for the Hearing Impaired”, South African Journal of Communication Disorders. 69(2), 1-11, 2022.

[8] M. Saraswathi, V. V. S. Ronit, S. Pranav, “Implementation of Video and Audio to Text Converter”, International Journal of Research Publication and Reviews, 4(5), 1204-1208, 2023.

[9] O. Basystiuk, N. Shakhovska, V. Bilynska, O. Syvokon, O. Shamuratov, & V. Kuchkovskiy, “The Developing of the System for Automatic Audio to Text Conversion”, Symposium on Information Technologies & Applied Sciences, Bratislava, Slovak Republic, March 5-6, 2021.

[10] N. Sharma, & S. Sardana, “A Real-Time Speech to Text Conversion System Using Bidirectional Kalman Filter in MATLAB”, International Conference on Advances in Computing, Communications and Informatics, Jaipur, India, September 21-24, 2016.

[11] A. Elakkiya, k. J. Surya, K. Venkatesh, & S. Aakash, “Implementation of Speech to Text Conversion Using Hidden Markov Model”, 6th International Conference on Electronics, Communication and Aerospace Technology, Coimbatore, India, December 1-3, 2022.

[12] A. P. Widyassari, S. Rustad, G. F. Shidik, E. Noersongko, A. Syukur, A. Affandy, D. I. M. Setiadi, “Review of Automatic Text Summarization Techniques & Methods”, Journal of King Saud University – Computer and Information Sciences, 34, 1029-1046, 2022.

[13] S. Thange, J. Dange, V. Karjule, J. Sase, & N. Kokate, “A Survey on Text Summarization Techniques”, International Journal of Scientific and Research Publications, 13(11), 528-535, 2023.

[14] D. Yadav, J. Desai, A. K. Yadav, “Automatic Text Summarization Methods: A Comprehensive Review”, ArXiv. <https://arxiv.org/abs/2204.01849>, 2022.

[15] S. Çalışır Zenci, “Öğretmen Adaylarının Metin Özetleme Stratejilerini Kullanım Tercihleri”, Dil ve Edebiyat Araştırmaları (DEA), 21, 341-359, 2020.

[16] O. Kaynar, Y. E. Işık, Y. Görmez, F. Demirkoparan, “Otomatik Metin Özetleme İçin Genetik Algoritma Tabanlı Cümle Çıkarımı”, Yönetim Bilişim Sistemleri Dergisi, 3(2), 62-75, 2017.

[17] N. M. Ranjan, “Text to Speech Conversion Based on Emotion Using Recurrent Neural Network”, International Journal for Research in Applied Science & Engineering Technology (IJRASET), 11(8), 970-978, 2023.

[18] M. J. Chowdhury & A. Hussan, “A review-based study on different Text-to-Speech technologies”, ArXiv. <https://arxiv.org/abs/2312.11563>, 2023.

[19] S. Tamboli, P. Raut, L. Satgaonkar, A. Atram, S. Kawane, V. K. Barbudhe, “A Review Paper on Text-to-Speech Converter”, International Journal of Research Publication and Reviews, 3(5), 3807-3810, 2022.

[20] S. Shastri & S. Vishwakarma, “An Efficient Approach for Text-to-Speech Conversion Using Machine Learning and Image Processing Technique”, International Journal of Engineering and Manufacturing, 13(4), 44-49, 2023.

[21] <http://explainwell.org/index.php/table-of-contents-synthesize-text/examples-of-summaries/>. [Accessed: Sept. 12, 2024].

[22] <https://www.ege-edebiyat.org/docs/437.pdf>. [Accessed: Dec. 23, 2024].

[23] <https://sonix.ai/docs/api>. [Accessed: May. 12, 2024].

This is an open access article under the CC-BY license

