




Turistler İçin, Engelli Bireylere Yönelik Ekler de İçeren, Görüntü Altyazılama Destekli Bilgilendirme ve Öneri Sistemi

Image Caption Generation Supported Information and Recommender System for Tourists, Including Supplements for Individuals With Disabilities

Muhammed Salih Tatar^{*1} , Rabia Kök² , Aybars Uğur³ 

¹Bilgisayar Mühendisliği Bölümü, Ege Üniversitesi, İzmir, Türkiye

²Bilgi İşlem Daire Başkanlığı, Bursa Teknik Üniversitesi, Bursa, Türkiye

³Bilgisayar Mühendisliği Bölümü, Ege Üniversitesi, İzmir, Türkiye

(msalihtatar@outlook.com.tr, rabia.kok@btu.edu.tr, aybars.ugur@ege.edu.tr)

Received:Aug.24,2023

Accepted:Aug.26,2023

Published:Oct.18,2023

Özetçe— Turizm, bireylerin farklı kültürlerle etkileşimi için en önemli araçlardan biridir. Turizm sektörüne önemli bir teknolojik ve sosyal yenilik getirecek, turistler için engelli bireylere yönelik ekler de içeren görüntü altyazılama destekli bilgilendirme ve öneri sistemi işlevlerinden oluşan bir web uygulaması geliştirilmiştir. Bu uygulamada, GPS sistemiyle elde edilen konum bilgisi kullanılarak yakın çevredeki gözde mekanların, tarihi ve turistik bölgelerin turistlere daha efektif bir şekilde ulaştırılması amacıyla İzmir'deki gözde mekanları kapsayan bir veri seti oluşturulmuş ve istatistiksel bir yöntem olan Apriori Algoritması kullanılmıştır. Görme engelli bireylerin şehri keşfetmelerine yardımcı olmak amacıyla kullanıcıdan alınan girdi görüntüsü için transfer öğrenme modellerinden olan VGG16 ve LSTM modelleri ile görüntü altyazısı üretme ve nesne tanıma işlevlerini gerçekleştiren bir Kolaylaştırıcı Modül tasarlanmıştır. Bu modül sayesinde görme engelli bireylerin, şehirdeki sokakların ve nesnelerin anlık görüntülerini metinsel ve işitsel olarak tasvir etmeleri ayrıca daha önceden sisteme girdikleri anahtar kelimelerin (yangın, metro, heykel vs.) görüntüde olma durumunu belirleyebilmeleri sağlanmıştır. MS-COCO, Flicker8k, Flicker30k ve Tourism48 veri setlerindeki görüntülerden oluşturulan veri seti; %80 eğitim, %20 test verisi olarak ayrılmıştır. Yapılan testlerde başarı değerleri, BLEU-1 için %55,41 ve BLEU-2 için ise %30,15 olarak elde edilmiştir.

Anahtar Kelimeler : Turizm, Derin Öğrenme, Nesne Tanıma, Görüntü Altyazısı Üretme, Öneri Sistemi

Abstract— Tourism is one of the most important tools for individuals to interact with different cultures. A web application has been developed that includes information and recommendation system functions supported by image captioning, which will bring significant technological and social innovation to the tourism sector and includes additions for disabled individuals. In this application, a data set covering popular places in Izmir, including nearby popular, historical, and touristic places was created using location information obtained with the GPS system. The Apriori Algorithm, which is a statistical method, was used to deliver popular places more effectively to tourists. A Facilitator Module was designed that performs image captioning and object recognition functions using VGG16 and LSTM models, which are transfer learning models for user input images obtained from visually impaired individuals to help them explore the city. With this module, visually impaired individuals can describe the instant images of streets and objects in the city in textual and auditory form and can also determine whether previously entered keywords (fire, metro, statue, etc.) are present in the image. The data set created from images in MS-COCO, Flicker8k, Flicker30k and Tourism48 data sets was divided into 80% training and 20% test data. Success values were obtained as %55.41 for BLEU-1 and %30.15 for BLEU-2 in the tests conducted.

Keywords : Tourism, Deep Learning, Object Recognition, Image Captioning, Recommender System

1. Giriş

Turizm sektörü, değişik şehir, bölge, ülke ve coğrafyaların tanınması, farklı kültürlerin kaynaşması, mental olarak vücudun dinlenmesi ve ayrıca ülkemizin gayri safi milli hasılası için de önemli bir yere sahiptir. Küresel amaçlarda da belirtildiği gibi, yüksek katma değerli sektörlerde farklı teknolojik çözümler üretme ve var olan

teknolojilerin yenilenmesi ile ekonomik verimliliğin daha üst düzeylere taşınması gerekmekte ve ayrıca kültürel ürünlerin kullanımını artıran istikrarlı turizmin teşviği için düzenlemeler yapılmalı ve uygulanmalıdır (Küresel Amaçlar-8, 2015). Böylece turizm sektörünün yeni teknolojilere uyum sağlamasıyla daha geniş kitlelere ulaşılması sağlanabilir.

Turizm endüstrisinde çeşitli amaçlarla yapay zeka teknolojilerinin kullanımı gerçekleştirilmektedir. Yapay zekâ destekli robotların turizm sektöründe kullanılmasıyla müşterilere hizmet sunulması, müşteri verileri üzerinde analiz yapılması ve öneriler getirilmesi gibi amaçlarla kullanılabilceği belirtilmiş ve turizm işletmeleri için çözümler sıralanmıştır. (Ercan, 2020).

Bir görselin bağlamını tanıma, bilgisayarlı görü ve doğal dil işleme kullanılarak görsele açıklama ekleme sürecidir. Görüntü üzerindeki nesnelere, özelliklerini, gerçekleştirilen eylemleri, nesnelere arasındaki etkileşimi vb. metinsel olarak ifade etmesi beklenmektedir. Görüntüye yazı eklenmesi temel olarak iki aşamada gerçekleştirilir: Görüntü Tabanlı Model (Kodlayıcı) ve Dil Tabanlı Model (Kod Çözücü) (Shaikh, 2018). Görüntü tabanlı modelde nesne ve işlevlerinin belirlenmesi amacıyla CNN (Evrişimli Sinir Ağları) modelleri, dil tabanlı modelde de görüntü tabanlı modellerin çıktılarında üretilen metinsel ifadelerin oluşturulması amacıyla RNN (Tekrarlı Sinir Ağları) modelleri kullanılmaktadır. CNN modellerinden olan ve transfer öğrenme yöntemi olarak kullanılan VGG16, 2014 yılında Visual Geometry Group tarafından geliştirilmiş ve ILSVRC yarışmasında büyük bir başarı elde etmiş, 13 adet evrişim katmanı ve 3 adet tam bağlantılı katman olmak üzere 16 adet katmandan oluşan bir derin öğrenme modelidir (Zhang, 2015). RNN modellerinden olan ve görüntülere tasvir üretilmesi amaçlı kullanılan, özellikle de zamansal verilerin işlenmesi gereken problemlerde etkili olan LSTM (Uzun Kısa Süreli Bellek) modeli, birer adet giriş, çıkış ve unutma kapısı mekanizmaları sayesinde bilgiyi koruyabilen ve güncelleyebilen özel bir yapıya sahiptir (Liu, 2019).

Öneri sistemleri, son zamanlarda popülerliğini gittikçe artıran konulardan biridir. E-ticaret sitelerinden müzik endüstrisine kadar birçok alanda kullanılan öneri sistemleri ile müşterilere sevindikleri ürünlerle kişiselleştirilmiş ve belki de hiçbir zaman keşfedemeyecekleri ürünler ve hizmetler sunulur (Çilingir, 2019; Dereli, 2020). Öneri sisteminde, birliktelik kuralları yöntemlerinden olan ve genellikle market sepeti analizi uygulamalarında kullanılan Apriori algoritması, bilgileri bir önceki adımdan elde eder. Temelde kullanıcı tercihlerinin tahmin edilmesine dayalı ve veriler için hesaplanan destek, güven ve ilginçlik değerlerini kullanır. (Sertçelik, 2023)

Bu çalışma, özellikle GPS sistemiyle beraber Derin Öğrenme yöntemlerinden Transfer Öğrenme ile Görüntü Altyazısı Üretme fonksiyonunun kullanılması, derin öğrenme modellerinin MS-COCO, Flickr8k, Flickr30k ve Tourism48 veri setlerinin harmanlanarak eğitim ve test verisi üzerinde eğitilmesi ve üretilen altyazıların konuşmaya dönüştürülmesi sayesinde engelli bireylere de önemli ölçüde destek olacak bir sistem içermesi yönü ile yenilikçidir. Çok sayıda derin öğrenme yönteminin bir arada kullanılmasıyla özellikle engelli bireyler için turistik faaliyetlerin daha da kolaylaştırılması yönüyle özgündür. Eşitsizliklerin sosyal alanlarda da azaltılması ve fırsat eşitliğini sağlaması amacıyla bu çalışmanın hedef kitleleri bütün insanlardır. Küresel amaçlara göre bireylerin statülerinden bağımsız olarak, sosyal, ekonomik ve politik olarak kapsanması, ayrımcı yönergelerin ortadan kaldırılıp uygun politikaların desteklenerek herkese eşit fırsatlar sunulması gerekmektedir (Küresel Amaçlar-10, 2015).

İkinci bölümde, literatürdeki öneri sistemi ve görüntü altyazılama alanındaki çalışmalara yer verilmiştir. Üçüncü bölümde, çalışmada kullanılan yöntem adımları detaylı olarak açıklanmıştır. Dördüncü bölümde, çalışma kapsamında gerçekleştirilen deneysel sonuçlar gösterilmiştir. Beşinci bölümde ise sonuç ve değerlendirmeye yer verilmektedir.

2. Literatür Araştırması

Turizm alanında yayınlanmış makale, bildiri, kitap vb. çalışmaların yanında çeşitli internet siteleri gibi harici kaynaklar analiz edilmiştir. Bu analiz sonucunda, turizm sektöründe yapılan çalışmalarda yapay zekâ ve derin öğrenme kullanımının nispeten kısıtlı olduğu tespit edilmiştir. Yapay zekâ yöntemlerinden Öneri Sistemi ve Görüntü Altyazılama alanlarında literatürde bulunan çalışmalar incelenmiştir.

2.1. Literatürde Öneri Sistemleri:

Bayrak et al. (2022) yaptıkları çalışmada öneri sistemi geliştirmede kullanılan Apriori ve TF-IDF, doğal dil işleme tabanlı Word2Vec ve FastText modelleriyle karşılaştıran bir çalışma gerçekleştirmişlerdir. Bu çalışma, Kaggle üzerinden erişilebilen, Grocery Products Purchase Data veri kümesi kullanılarak gerçekleştirilmiştir. 5 adet ürün için yapılan inceleme sonucunda, satın alma işleminde ürünlerin birliktelik ilişkilerini önemsemeyen TF-IDF algoritması, %13,9 başarı oranı ile en başarısız yöntem olmuştur. Anlamsal modellerden FastText, %32,6 ile Word2vec'ten daha başarılıdır. Apriori algoritması ise %24,4 başarı oranı ile TF-IDF algoritmasından daha başarılı, anlamsal modellerden daha başarısızdır.

R. Burke (2022) tarafından en uygun öneriyi gerçekleştirmek için içeriğe, iş birliğine ve bilgiye dayalı bu teknikler birleştirilerek EntreeC hibrit öneri sistemi geliştirilmiştir. Bu sistemde, UCI KDD arşivinde bulunan Entree veri seti üzerinde yapılan deneysel çalışmalar sonucunda, bilgiye dayalı anlamsal derecelendirmelerin sayısal derecelendirmelere kıyasla daha doğru tahminler ürettiği tespit edilmiştir.

2.2. Literatürde Görüntü Altyazılama:

M. Kuyu vd. (2018) tarafından Türkçe dilinde görüntü altyazılama probleminin zorluklarını aşmak için alt sözcük öğeleri kullanılarak görüntü altyazılama yaklaşımı Türkçe dilinin kuralları temel alınarak geliştirilmiştir. Türkçe eğitim kümesinde yer alan sözcükler öncelikle n-gram istatistikleri ve BPE algoritması ile alt sözcüklere ayrıştırılmış ve alt sözcüklere dayalı temsili sözlükler oluşturulmuştur. Alt sözcük tabanlı modelin, sözcük tabanlı modele göre BLEU-2, BLEU-3, BLEU-4, CIDEr ve WMD ölçüm metriklerinde daha başarılı olduğu belirlenmiştir. MS-COCO veri kümesindeki 500 adet görüntü için doğrulama yapıldıktan sonra başarı düzeyi en yüksek sonuçları MS-COCO ile eğitilen ve TasvirEt veri kümesi ile ince ayar yapılan alt sözcük tabanlı model vermiştir. Bu modelin başarımı BLEU-1 metriğine göre 0,265; BLEU-2 metriğine göre 0,163 olarak ölçülmüştür. Ancak Flicker30k, MS-COCO veri kümesinden daha küçük olduğu için elde edilen sonuçlar daha az başarılıdır.

Wang et al. (2016) tarafından görüntü altyazılama için Bi-LSTMA olarak adlandırılan AlexNet ve iki ayrı LSTM ağı üzerine kurulu olan, değişken uzunlukta ve yeni cümleler oluşturmadaki zorlukların üstesinden gelebilecek bir model önerilmiştir. AlexNet diğer modellere göre daha zayıf bir model olmasına karşın, rekabet edebilecek düzeyde bir performans sergilemiştir. Bi-LSTMA ve DeepVSV BLEU-1 ve BLEU-2 metriklerine göre karşılaştırıldığında, Bi-LSTMA modeli biraz daha iyi bir performans göstermiştir. Bi-LSTMA ve DeepVSV modellerinin BLEU-1 ve BLEU-2 metriklerine göre başarı değerleri sırasıyla, Flicker8k veri seti için 61,9/43,3-57,9/38,3; Flicker30k veri seti için 58,2/39,3-57,3/36,9; MS-COCO veri seti için 63,4/44,7-62,5/45,0 olarak ölçülmüştür. Elde edilen başarı sonuçlarına göre, büyük veri kümelerindeki eğitimler için derin LSTM mimarilerinin daha uygun olduğu ve böylece ezberleme sorununun önüne geçildiği tespit edilmiştir.

3. Turistler İçin Engelli Bireylere Yönelik Ekler de İçeren Görüntü Altyazılama Destekli Bilgilendirme ve Öneri Sistemi

3.1. Veri setleri:

MS-COCO: Lin (2014) tarafından Nesne tanıma, segmentasyon, key-point algılama ve altyazılama için kullanılmak üzere oluşturulmuş veri setidir. 328K adet görüntüden oluşur. Altyazılama için her bir görüntü için 5 farklı metin yer almaktadır. Her bir amaca yönelik kategori ve görüntü sayısı farklıdır.

Flicker8k: Hodosh (2013) tarafından cümle tabanlı görüntü altyazılama problemine yönelik olarak oluşturulan 8 bin görüntüye sahip bir veri setidir ve her görüntü için 5 adet metin içermektedir.

Flicker30k: Plummer (2015) tarafından cümle tabanlı görüntü altyazılama problemine yönelik olarak oluşturulmuştur. Anlamsal bir bağlam etrafında şekillendirilmiş cümlelerden oluşan 30 bin civarı görüntüye ve 158 bin altyazı metnine sahip bir veri setidir. Her görüntü için beş adet altyazı metni bulunmaktadır.

Tourism48: Bounab (2020) tarafından oluşturulan veri seti, kullanıcı yorumlarına dayalı görüntü altyazısı oluşturmak için tasarlanmıştır. Her ülkenin 10 adet galeri ile temsil edildiği 48 adet ülkeyi içermektedir. Her galeride yalnızca bir resim bulunur. Her görüntü 2 adet JSON dosyasıyla birlikte gelir.

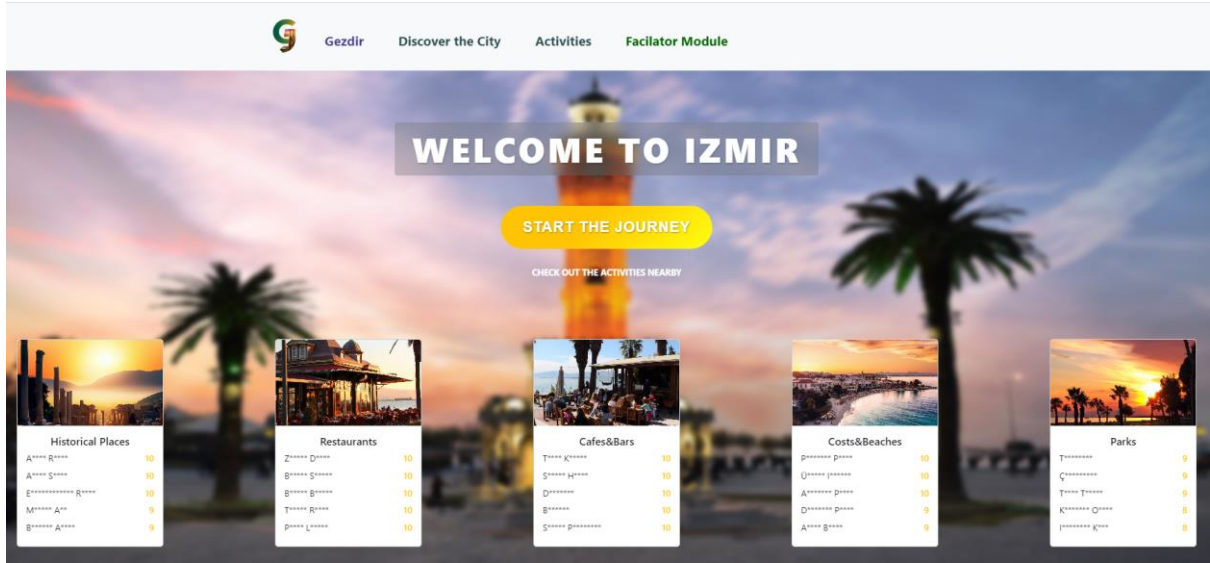
İzmir'deki Gözde Mekanlar Veri Seti: Turizm sektöründe veri seti bulmak oldukça zordur. Her şirketin kendine özgü veri setleri bulursa da dış kullanıcılara kapalı durumdadır. Bu yüzden çalışmada kullanılmak üzere üniversite öğrencilerinin katkılarıyla öneri sistemi için bir veri seti hazırlanmıştır. 63 adet öğrenciye İzmir'de en sevdikleri restoran, tarihi mekan, kafe, plaj ve park sorulmuştur. Her öğrenci dilediği kadar cevap vermiştir. Yapılan anket sonucunda 100 adet gözde mekan kümeler halinde belirlenmiştir. Öğrencilerden herhangi bir kişisel bilgi talep edilmemiştir. Rumuz kullanılarak mekan önerilerinin ankete girilmesi sağlanmış ve öğrenciler bu verilerin nerede ve nasıl kullanılacağı konusunda bilgilendirilmiştir. Şekil 3.1'de, hazırlanan veri setinin bir örneği gösterilmektedir.

Rumuz	Restoranlar/Lokantalar vb.	Kafeler vb.	Tarihi ve Turistik Mekanlar/Ören Yerleri/Müzeler vb.	Plajlar	Parkları/Piknik Alanları vb.	Belirtmek İstedığınız Diğer Yerler
n****	Z***** D****	S***** H****	A**** Ö**** Y****	Ç***** P***** P****	Ç*****	
nl****	B***** S*****	R**** C*****	T***** A*****	D***** P****	İ**** D**** Y**** P****	
nl****	Ç***** D****	D*****	A**** S**** M*****		K*****	K***** S*****
ba****	B***** B*****	B*****	E****	Ü***** P****	B*****	
ba****	H***** S*****	B*****	E***** M****	D***** P****	A**** V***** P****	
ba****	B*** T*****	S***** H****		Ç***** A***** P****		
ba****	T***** R****	B*****			G***** İ*****	
Nu****	Ö*** U****	o*****	a*****			
Nu****	S***** H****	k*****	s*** k*****			

Şekil 3.1 İzmir'deki Gözde Mekanlar Veri Setinin 9 Satırı

3.2. Önerilen ve Geliştirilen Sistem:

Bu çalışmada, turizm sektöründe görüntü altyazılama ve öneri sistemi kullanılarak turistlere özel bir gezi rehberi oluşturulmuştur. Buldukları şehri keşfetmek isteyen turistlere yönelik kendi lokasyonlarına ait hava durumu bilgisi, şehir içi ulaşım bilgisi ve şehirdeki etkinliklerin sunulduğu bir web uygulaması geliştirilmiştir. Bu uygulamada kullanıcı anasayfa ile karşılaşmaktadır. Ana sayfanın üst kısmında navbar, üst sağ kısmında ise kullanıcının konumuna göre getirilen hava durumu bilgisi bulunmaktadır. Kullanıcı, ana sayfanın orta kısmında bulunan "Start The Journey" butonu ile "Discover the City", "Check out the Activities Nearby" butonu ile "Activities" sayfasına yönlendirmektedir. Ana sayfanın alt tarafında en yüksek skora sahip 5'er adet Tarihi Mekan, Restoran, Kafe, Sahil ve Park yer almaktadır. Her lokasyonun skor bilgisi sistem yöneticisi tarafından girilmektedir. Mekanlara ait skor bilgilerinin öneri sistemi ile bağlantısı bulunmamaktadır. Şekil 3.2'de, web uygulamasının anasayfa ekranı gösterilmektedir.



Şekil 3.2 Anasayfa Görünümü

Turistlere yönelik, şehirdeki en gözde restoranları, konaklama alanlarını, tarihi ve turistik mekanları (ören yerleri, müzeler vb.) gösteren bir bilgilendirme ve öneri sistemi geliştirilmiştir. Bu sistemin akış şeması Şekil 3.3'te gösterilmektedir.



Şekil 3.3 Akış Şeması

Öneri sistemlerinin turizm alanına uyarlanması ile kullanıcıya ürün yerine mekan önerisi yapılmıştır. Bu sayede kullanıcı, daha önceden birlikte tercih edilen en gözde mekanları keşfedebilmektedir. Çalışmadaki öneri sistemi bölümünde yeni bir veri seti üretilmiş ve bu veri seti kullanılarak Apriori algoritmasının performansı gözlemlenmiştir. Discover the City sayfasında Google Maps haritası ve kullanıcının konumuna göre harita üzerinde işaretlenmiş mekanlar kullanıcıyı karşılamaktadır. Sayfanın sağ tarafında harita üzerinde işaretlenmiş mekanların isim, mekan tipi ve skor bilgileri listelenmiştir. Harita ve mekan listesinin üzerinde mekan tipi ve Apriori algoritması ile oluşturulan önerilerin filtreleri bulunmaktadır. Kullanıcının kendi konumu kırmızı marker ile harita üzerinde gösterilmektedir. Haritanın merkezi, kullanıcının konumuna göre ayarlanmıştır. Apriori Algoritması ile önerilen mekanlar turuncu, diğer mekanlar yeşil renk marker ile harita üzerinde işaretlenmiştir. Activities sayfasında ise kullanıcıya şehirdeki etkinlikler sunulmaktadır. Sayfada görülen resimlere tıklayarak kullanıcı bilet almak için bilet satışının yapıldığı dış bir bağlantıya yönlendirilmektedir. Şekil 3.4'te, Discover the City; Şekil 3.5'te, Activities sayfasının görünümü verilmiştir.

Rating	Category	Score	Suggested
Z***** D****	Restaurant	10	Yes
B***** S*****	Restaurant	10	No
T**** K*****	Cafe / Bar	10	Yes
B***** B*****	Restaurant	10	Yes
H***** S*****	Restaurant	9	No
B*** T*****	Restaurant	8	No

Şekil 3.4 Discover the City Sayfası



Gezdir

Discover the City

Activities

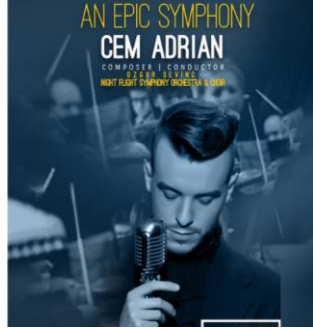
Facilator Module

26.8°

Clear

Activities Nearby

< Previous



Next >

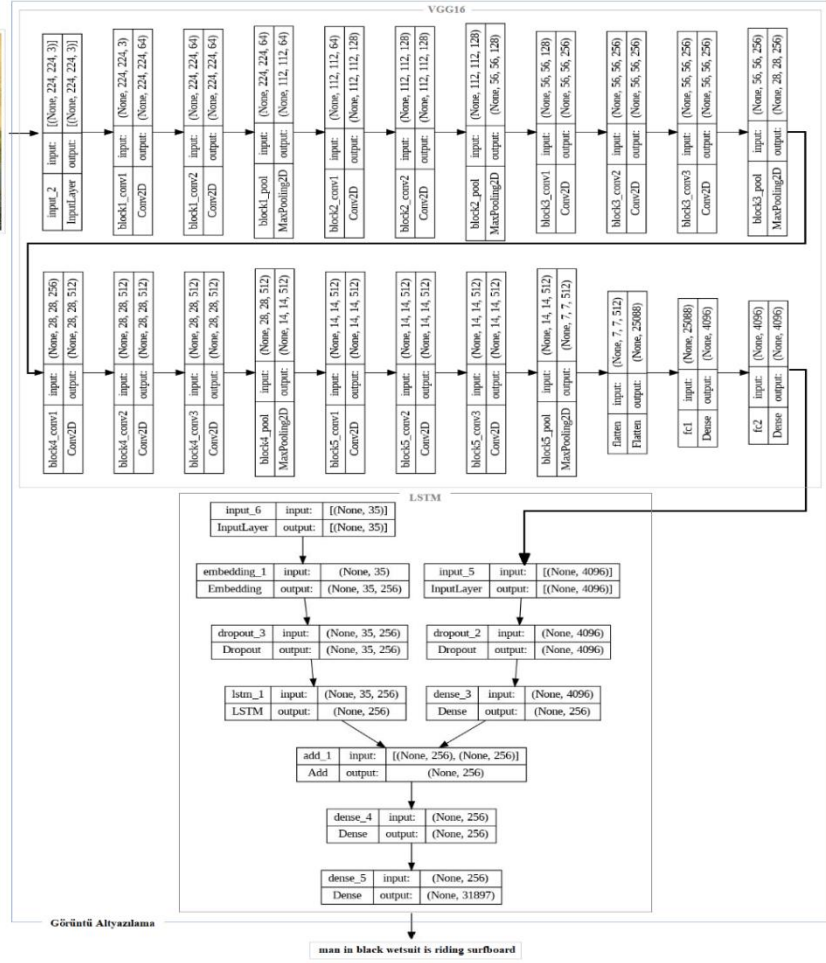
Şekil 3.5 Activities Sayfası

Ayrıca engelli bireyleri de belli ölçüde kapsayacak şekilde şehre ait görüntüleri metinsel ve işitsel olarak tasvir edecek bir Kolaylaştırıcı Modül tasarlanmıştır. Engelli bireylerin şehirdeki tehlikelere karşı önceden uyarılabilmeleri, şehrin keşfinin kolaylaştırılması ve keşfin daha güvenli bir hale getirilmesi amacıyla kullanıcının önceden sisteme yüklediği nesne ve olguların (ateş, metro, araba, vs.), sisteme yüklenen görüntüde olma durumunun belirlenmesi sağlanmıştır. Bir anlamda engelsiz turizm denilen kavrama bir adım daha yaklaşılmaya çalışılmıştır. Çalışmadaki görüntü altyazılama yönteminin kodlayıcı modülünde VGG16, kod çözücü modülünde ise LSTM modeli kullanılmıştır (Şekil 3.6). Bu modellerin performansları Flicker8k, Flicker30k, MS-COCO ve Tourism48 veri setleri üzerinde incelenmiştir. Facilator Module sayfasında kullanıcının şehre ait görüntüleri metinsel ve işitsel olarak tasvir edebilmesi için geliştirilen Kolaylaştırıcı Modül bulunmaktadır. Sayfaya girildiğinde kullanıcıyı görüntü yükleyebileceği ve anahtar kelime girebileceği bir alan karşılamaktadır. Web uygulaması React ile geliştirildiği için platform bağımsızdır. Bu nedenle engelli bireyler de engel durumlarına özel olarak üretilen cihazlar üzerinden uygulamayı rahatlıkla kullanabilirler. Kullanıcı, tasvirini istediği veya nesne tespiti yapmak istediği görüntüyü yüklediği anda Görüntü Altyazılama API çalışır ve görüntü altyazılama ile oluşturulan görüntü tasviri sayfanın altındaki alana yazılır. Kullanıcı görüntüyü yükledikten sonra “Describe the Image” butonu ile görüntünün tasvirini sesli olarak alabilmektedir. Kullanıcı görüntüyü yükledikten sonra “Check Keywords” butonu ile anahtar kelimelerin yüklenen görüntüde bulunma durumunu kontrol edebilmektedir. Şekil 3.7’de Facilator Module sayfasının görünümü verilmiştir.

4. Deneysel Çalışmalar

4.1. Görüntü Altyazılama:

Bu çalışmadaki görüntü altyazılama yönteminde kodlayıcı olarak VGG16, kod çözücü olarak ise LSTM modeli kullanılmıştır. Görüntü altyazılama modelini eğitmek için MS-COCO, Flicker8k, Flicker30k ve Tourism48 veri setlerinden rastgele görüntüler alınarak hibrit bir veri seti oluşturulmuştur. Bu hibrit veri setinin %80’i eğitim, %20’si test veri seti olarak kullanılmıştır. Görüntü altyazılama çalışması; Tensorflow, Keras, Numpy, os, Pickle gibi kütüphaneler kullanılarak Anaconda3 ortamında Jupyter Notebook editöründe gerçekleştirilmiştir. Görüntü altyazılama uygulaması için gerçekleştirilen VGG16 kodlayıcı ve LSTM kod çözücü modellerinin hiper parametre değerleri Tablo 4.1’de gösterilmektedir.



Şekil 3.6 Görüntü Altyazılama Yöntemi Blok Diyagramı

Gezdir Discover the City Activities Facilator Module 25.9° Clear

Change / Remove Image

Add New Keyword Add Keyword

Fire X Volcano X Traffic Light X

Check Keywords

group of people are standing around fire Describe the Image

Şekil 3.7 Facilator Module Sayfası

Tablo 4.1 VGG16 Kodlayıcı ve LSTM Kod Çözücü Modellerinin Hiper Parametre Değerleri

Kodlayıcı Model	Input-Shape	4096
	Input-Droupout	0,3
	Input-Dense-Layers	256
	Input-Dense-Activation Function	Relu
Kod Çözücü Model	Input-Shape	35 (maksimum altyazı uzunluğu)
	Embedding-VocabSize	31897
	Embedding-Layers	256
	Embedding-Mask_zero	true
	Droupout	0,3
	LSTM	256

Görüntü altyazılama için oluşturulan model 20 epokta 5 kez eğitilmiştir. BatchSize değeri 32 olarak belirlenmiştir. Şekil 4.1'de, Görüntü Altyazılama Modeli kullanılarak elde edilen örnek çıktılar gösterilmiştir.

```
Detected Objects: [(('n03388043', 'fountain', 0.14823267), ('n02951358', 'canoe', 0.14814322), ('n03873416', 'paddle', 0.09827725))]  
-----Predicted-----  
startseq man in black wetsuit is riding surfboard endseq
```



Şekil 4.1 Görüntü Altyazılama Örneği

Görüntü Altyazılama modellerinin başarısını test etmek için BLEU-1 ve BLEU-2 başarı metrikleri kullanılmıştır. Yapılan 5 eğitimin ortalaması alınarak test ölçüm sonuçları elde edilmiştir. Tablo 4.2'de görüntü altyazılama modellerinin test verisi üzerindeki BLEU-1 ve BLEU-2 başarı metriklerine ait ölçüm sonuçları gösterilmektedir.

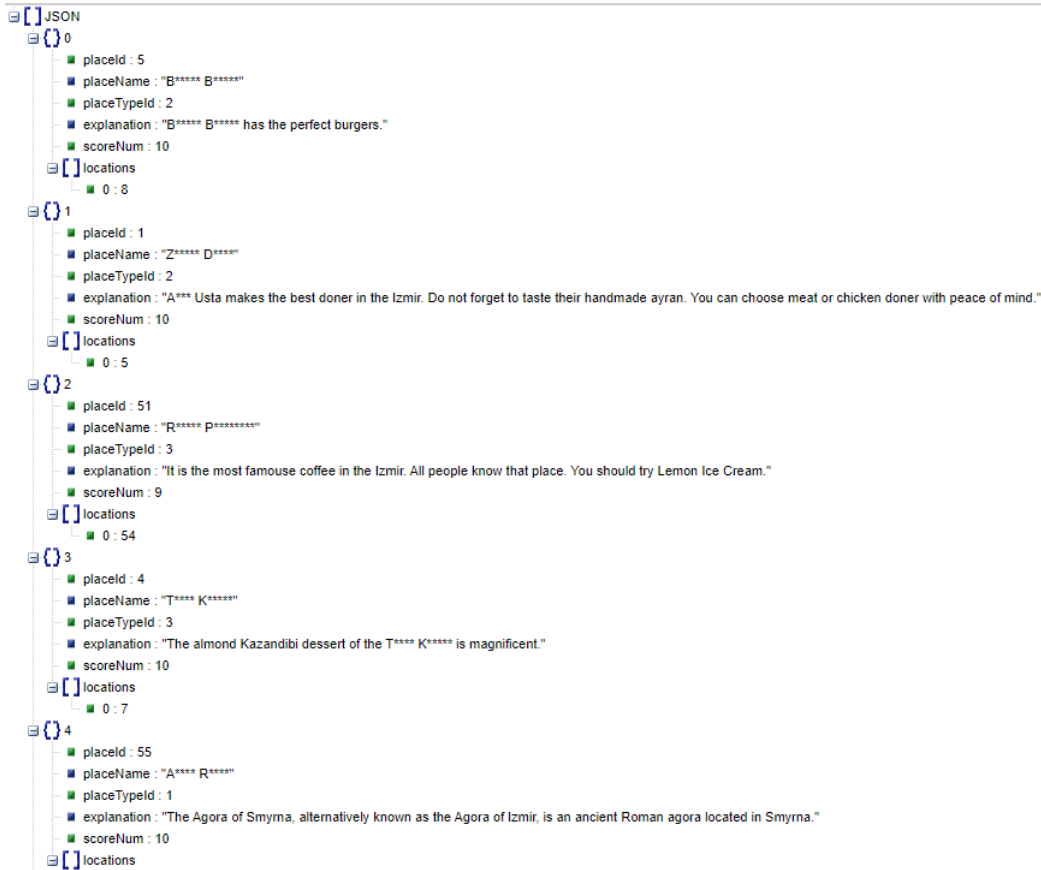
Tablo 4.2 Görüntü Altyazılama Modellerinin Test Veri Seti Üzerindeki Başarı Ölçüm Sonuçları

	BLEU-1 (%)	BLEU-2 (%)
Deney-1	55,94	30,39
Deney-2	54,87	29,74
Deney-3	55,32	29,92
Deney-4	55,86	29,82
Deney-5	55,05	30,88
Ortalama	55,41	30,15

4.2. Öneri Sistemi:

Bu çalışmada, öneri sisteminin gerçekleştirilmesi için Apriori algoritması kullanılmıştır. Algoritmada en küçük destek değeri 0,1, en küçük güven değeri ise 0,5 olarak alınmıştır. Apriori algoritması ile elde edilen 15 adet öneri kümesindeki verilerden en fazla destek değerine sahip olan 10 adet mekan kullanıcılara önerilmiştir. Apriori algoritmasında öncelikle kullanıcıların önerdiği bütün öneri kümeleri alınmış ve mekan tiplerine göre (tarihi mekan, restoran, kafe, park, plaj) ayrılmıştır. Ardından belirlenen en küçük destek değeri kullanılarak Apriori algoritması ile sık önerilen mekan kümeleri bulunmuştur. Sık önerilen mekan kümeleri, bütün öneri kümeleri ve en küçük güven değeri kullanılarak ilişkilendirme kuralları oluşturulmuştur. İlişkilendirme kuralları güven değerine göre yapılmıştır. Her ilişkilendirme kuralında öncül, sonuç, destek ve güven bilgileri bulunmaktadır. Öncül, bir kurallı veri kümesindeki öğelerin hangi sıklıkla bir arada görüldüğünü ifade eder. Apriori algoritmasında sonuç, ilişki kuralın sonucunu oluşturan öğe kümesini belirtir. İlişki kuralı, bir veri kümesine ait bir öğe kümesinin alt kümelerinin de aynı veri kümesinde sıklıkla görüldüğünü gösterir. Geliştirilen öneri sisteminde öncül ve sonuç ile önerilen mekan kümelerinin, tekrarlamayacak şekilde, birleştirilerek kullanıcıya tek bir öneri kümesi olarak sunulması sağlanmıştır. Bu şekilde oluşturulan 15 adet öneri kümesinin içinden ise en sık tekrarlayan 10 adet mekanın kullanıcıya sunulması sağlanmıştır. Öneri kümeleri belirlenirken sonuç kümesi boş olmayan, öncül kümesine göre gruplanmış, güven değeri en yüksek gruplar alınmıştır.

Şekil 4.2’de, bir kısmı listelenen öneri kümelerinin gruplanarak oluşturulan son hali görülmektedir. Bu noktada birbirleriyle en çok tercih edilen mekanlar kümeler halinde gösterilmiştir. İlk sırada gelen B**** B*****, R***** P***** ve Z***** D**** mekanları beraber tercih edilen, güven değeri en yüksek öneri kümesidir. İkinci sırada ise Z***** D**** ve T*** K***** K*** gelmektedir.



Şekil 4.2 Kullanıcıya Önerilen Örnek Mekanlar

Bu çalışmada gerçekleştirilen öneri sistemi ilk test edildiğinde öneriler, kümeler şeklinde yapılmaktaydı. Apriori algoritması tam da bunu gerektirmektedir. Yani en çok tercih edilen öğeler kullanıcılara kümeler halinde sunulmaktadır. Ancak birden fazla öneri kümesinin kullanıcıya sunulmasının kullanışlı olmayacağı düşünüldüğünden özel bir öneri algoritması geliştirilmiştir. Bu algoritma daha önce de belirtildiği gibi öneri

kümelere içerisinde, öncül ve sonuç kümeleri birleştirildikten sonra elde edilen, birbirleriyle sık kullanılan, güven değeri en yüksek 10 adet mekanın kullanıcıya iletilmesini sağlamaktadır. Bu yeniliğin sonucunda; kullanıcıya birden fazla öneri kümesi göstermekten daha kullanışlı bir yapı ortaya çıktığı görülmüş, daha anlaşılır ve uygun öneriler sunulduğu tecrübe edilmiştir.

5. Sonuç

Bu çalışmada, kullanıcılardan alınan görüntüler üzerinden metinsel tasvir üretilmiş ve bu tasvir engelli bireylere turistik faaliyetlerde yardımcı olması amacıyla seslendirilmiştir. Ayrıca kullanıcının lokasyon bilgisi kullanılarak turistik amaçlı mekan önerisinde bulunulmuştur.

VGG16 ve LSTM modellerinin, test veri seti üzerinde yapılan testlerde BLEU-1 ve BLEU-2 başarı ölçüm metriklerine göre görüntü altyazılama başarısı sırasıyla %55,41 ve %30,15 olarak elde edilmiştir. Bu sonuçlar, model eğitiminin 5 kez tekrarlanıp her tekrarda elde edilen test sonuçlarının ortalaması alınarak hesaplanmıştır. Görüntü altyazılama için kullanılan modelin eğitimi ve testi için MS-COCO, Flicker8k, Flicker30k ve Tourism48 veri setlerinden görüntüler birleştirilerek hibrit bir veri seti oluşturulmuştur. Bu veri seti, %80 eğitim; %20 test veri seti olarak kullanılmıştır.

Hazırladığımız veri seti üzerinde öneri sistemi yöntemlerinden Apriori algoritması kullanılarak, kullanıcının lokasyon bilgisine göre mekan önerisi yapılmaktadır.

Deney sayılarının tekrarı artırılarak ve farklı transfer öğrenme modelleri kullanılarak görüntü altyazılamadaki model başarısının artırılacağı düşünülmektedir.

Mekan önerisine ilave olarak kullanıcıların önerilerden yaptıkları seçimlerin gözde mekanlar veri setine eklenmesiyle kullanıcılara üretilecek önerilerde daha elverişli tahminler yapılabileceği düşünülmektedir.

Teşekkür

Bu çalışma, TÜBİTAK 2209-A programı desteği ile yürütülmüştür. Proje adı: (Turistler İçin Engelli Bireylere Yönelik Ekler de İçeren Görüntü Altyazılama Destekli Bilgilendirme ve Öneri Sistemi), proje numarası: (1919B012222734)

Kaynaklar

- Bayrak, A. T., Öner, S. C., Gencer, M., Cerit, O. S., Oymagil, A., & Dalva, D. (2022, May). Using word embedding methods for product recommendation. In *2022 30th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.
- Bounab, Y., Oussalah, M., & Ferdenache, A. (2020, November). Reconciling image captioning and user's comments for urban tourism. In *2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA)* (pp. 1-6). IEEE.
- Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12, 331-370.
- Çilingir İ., medium.com, <https://medium.com/@iremcilingir/%C3%B6neri-sistemleri-recommendation-systems-28a3f341c0a9>, (Erişim Tarihi: 04.06.2023)
- Dereli S. M., Veri Bilimi Okulu, <https://www.veribilimiokulu.com/oneri-sistemleri-101/>, (Erişim Tarihi: 04.06.2023)
- Ercan, F. (2020). Turizm pazarlamasında yapay zekâ teknolojilerinin kullanımı ve uygulama örnekleri. *Ankara Hacı Bayram Veli Üniversitesi Turizm Fakültesi Dergisi*, 23(2), 394-410.
- Hodosh, M., Young, P., & Hockenmaier, J. (2013). Framing image description as a ranking task: Data, models and evaluation metrics. *Journal of Artificial Intelligence Research*, 47, 853-899.
- Kuyu, M., Erdem, A., & Erdem, E. (2018). Altsözcük Ögeleri ile Türkçe Görüntü Altyazılama Image Captioning in Turkish with Subword Units.
- Küresel Amaçlar, Eşitsizliklerin Azaltılması – Madde:10.2, Madde:10.3, Madde:10.4, <https://www.kureselamaclar.org/amaclar/esitsizliklerin-azaltilmasi/>, (Erişim Tarihi: 05.06.2023)
- Küresel Amaçlar, İnsana Yakışır İş ve Ekonomik Büyüme – Madde:8.2, Madde:8.9, <https://www.kureselamaclar.org/amaclar/esitsizliklerin-azaltilmasi/>, (Erişim Tarihi: 05.06.2023)

- Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13* (pp. 740-755). Springer International Publishing.
- Liu, X., Xu, Q., & Wang, N. (2019). A survey on deep neural network-based image captioning. *The Visual Computer*, 35(3), 445-470.
- Plummer, B. A., Wang, L., Cervantes, C. M., Caicedo, J. C., Hockenmaier, J., & Lazebnik, S. (2015). Flickr30k entities: Collecting region-to-phrase correspondences for richer image-to-sentence models. In *Proceedings of the IEEE international conference on computer vision* (pp. 2641-2649).
- Sertçelik, Ş., & Önder, E. (2023). Yönetim Bilişim Sistemleri Kapsamında Akademik Araştırma Alanlarının İncelenmesi: Apriori Algoritması ile Bir Analiz. *Gümüşhane Üniversitesi Sosyal Bilimler Dergisi*, 14(2), 680-690.
- Shaikh, F. (2018). Automatic image captioning using deep learning (CNN and LSTM) in PyTorch, Analytics vidhya.
- Wang, C., Yang, H., Bartz, C., & Meinel, C. (2016, October). Image captioning with deep bidirectional LSTMs. In *Proceedings of the 24th ACM international conference on Multimedia* (pp. 988-997).
- Zhang, X., Zou, J., He, K., & Sun, J. (2015). Accelerating very deep convolutional networks for classification and detection. *IEEE transactions on pattern analysis and machine intelligence*, 38(10), 1943-1955.