

GİRİŞ

Sohbet botu, kullanıcı ile genellikle metin, bazı durumlarda ise konuşma yoluyla diyalog kurarak bilgi veren veya bir işlemi gerçekleştiren bir yazılımdır. Sohbet botu terimi ilk olarak “chatter bot”[24] olarak Michael L. Mauldin tarafından 1994 yılında kullanılmıştır.

Sohbet botları, AI sistemlerinin tipik bir örneğidir ve akıllı İnsan-Bilgisayar Etkileşiminin (HCI) en temel ve yaygın örneklerinden biridir [5]. Sohbet botları, çeşitli alanlarda hedef kullanıcılarla verimli bir şekilde diyalog sağlamak amacıyla anlık mesajlaşma frameworklerini kullanan yazılımlardır. Sohbet botları, herhangi bir çevrimiçi olma durumu veya son görülme bilgileri olmadığı için insan kullanıcılardan tamamen farklıdır [6]. Bu sohbet botlarından bazıları gelişmiş Doğal Dil İşleme Sistemlerini (NLPS) kullanır fakat bir kısmı da sadece giriş metnindeki anahtar kelimeleri tarayıp onlara odaklanarak cevap üretir. Sohbet botları halihazırda gelişim sürecinde olan bir teknoloji olduğu için diyalog kalitelerinde büyük farklılıklar gözlemlenmektedir [7]. İnsanlar gibi sohbet botları da kendi bilgilerini ortama dayalı olarak, anlamak için kullanabilirler. Böyle bir sistem, bilgisayarların hesaplama gücü avantajına sahip olarak, bazı görevlerde bir insandan bile daha başarılı olabilir [8].

Sohbet botları, müşterilere hızlı ve doğal bir etkileşim imkânı sunmalarını sağlaması ve aynı zamanda verimliliği arttırmaları ile şirketlerin de ilgisini çekmektedir. Günümüz müşterisi alınan hizmeti kesintisiz ve hızlı bir biçimde almak istemektedir. Sohbet botları bunu sağlama vaadi ile şirketler tarafından denenmeye ve kullanılmaya başlanmıştır. Şirketlerinin gösterdiği ilginin bir diğer sebebi de Sohbet botlarının şirket verimliliğini artırma vaadidir. Günümüzde birçok sektörde daralan kar marjları ve yeni teknolojilerin sunduğu imkânlar şirketleri verimlilik arayışına itmiştir [12].

Sohbet botu tasarlariken duyulan kaygılardan birisi kullanıcılar tarafından bota duyulacak güvenin beklenenden az olmasıdır. Kullanıcının güven derecesini etkileyen faktörlere baktığımızda ise kullanıcının yaşadığı deneyimler, tasarımcıyla ilgili bilgiler ve güvenlik sorunlarından bahsedebiliriz. Bu güveni arttırmak istersek de botumuzun arayüzü, insansı isim belirleme, botumuza katacağımız kişilik özellikleri, insan dilini işleme becerisi gibi konulara göz atmak gerekir. Bunların yanında sohbet botlarına duygu katarak bilinçli bir sohbet robotu tasarlamaya yönelim de mevcuttur [9].

Sıklıkla kullanılan iki tür sohbet botu bulunmaktadır. İlk sohbet botları, konuyu anlamaya çalışan, aynı konu hakkında her soru için her zaman aynı cevabı veren yapıdaki sohbet botlarıdır. Genel olarak, bu tür sohbet botları cevap için sınıflandırma algoritmalarını kullanmaktadır. Bu tür bot uygulamalarında, N adet konu ve N adet cevabın listesi bulunmaktadır. Uygulamada, sınıflandırılan konunun olasılığı az ise kullanıcıdan daha özel ifadeler kullanarak soruyu tekrar etmeleri istenir. İkinci türdeki sohbet botları ise daha gelişmiş, güncel teknolojileri kullanır ve aynı zamanda daha karmaşıktır. Bu botlardaki cevaplar, genellikle RNN (Tekrarlayan Sinir Ağı) kullanılarak oluşturulmaktadır. Bu sohbet botları daha kişiselleştirilmiş ve daha uygun cevaplar verebilmektedir. Özel bir RNN türü olan Uzun Kısa Süreli Bellek (LSTM) kullanıcının sorularını daha iyi anlayıp mümkün olan en iyi cevabı üretmektedir. Geleneksel seq2seq sohbet bot modelleri, çıktı cümlelerinin duyarlılığını düşürmeden sadece girdi dizilerinde şartlandırılmış en yüksek olasılıklara sahip cümleleri bulmaya çalışmaktadır [2, 3].

İlgili bilgiler ile ihtiyaç duyulan cevap arasındaki farkın az olduğu durumlarda, RNN'ler geçmişteki bilgileri kullanmayı öğrenebilirler. Tahmin edilmesi gereken kelime sayısı çoğaldıkça, RNN'ler bilgileri birleştirmeyi öğrenemezler. LSTM'ler ise bu soruna bir çözüm olarak bilginin uzun süre hatırlanması için tasarlanmıştır [2]. Çalışmada LSTM ve seq2seq TensorFlow Kütüphanesi kullanılarak bir sohbet botu tasarımı yapılmıştır.

1. GENEL BİLGİLER

Günümüzde teknolojinin gelişimi ile beraber yavaş yavaş yapay zeka kavramını hayatımızın her alanında görmeye başladık. Yapay zekaların artması ile birlikte şirketlerin iş gücünü arttırabilmek, maliyeti azaltabilmek ve zamandan tasarruf edebilmek gibi temel ihtiyaçlarından bazıları kolayca halledilebilmekte. Yapay zeka, görevleri yerine getirmek için insan zekasını taklit eden ve topladığı bilgilerle kendisini seviyeli olarak ilerletebilen sistemler veya makineler anlamına gelir. Yapay zekanın amacı insanın yerini almak değil insanın yeteneklerini geliştirmek ve katkıda bulunmaktadır. Yapay zekanın birçok kullanım alanı vardır. Örneğin;

- Sohbet robotları
- Akıllı asistanlar
- Öneri robotları vb.

Yukarıda belirtilen ve daha birçok farklı alanda karşımıza çıkan yapay zeka ticaret içinde oldukça değerli bir varlıktır.



Şekil.1 Yapay Zeka Dalları [1]

1.1. Sohbet Botu Teknolojisi

Sohbet botları yaklaşık 2016'dan bu yana yoğun şekilde konuşulmaktadır. Ancak yapay zeka konusundaki bilimsel araştırmalar 1950'lere kadar gitmektedir. 1950 yılında Alan Turing tarafından Mind dergisinde bir seminer raporu yayınlanmış, yazıda bugün yapay zekanın ilk örneği olarak kabul edilen Turing testinden söz edilmiştir.[2] Turing testi bir bilgisayar programının, insan ve yazılım ile iletişime geçmesini ve hangisinin insan, hangisinin yazılım olduğunu ayırt etmesine dayanan bir testtir. Eğer denenler, insan ve yazılımı birbirinden ayıramazsa yazılım testi geçmiş sayılmaktadır.

İlk sohbet botu olarak ise 1966 yılında bir MIT profesörü olan Joseph Weizenbaum tarafından geliştirilen "ELIZA" olarak kabul görmektedir. ELIZA bir psikolog olarak yaratılarak insani diyaloglar kurması amaçlanmıştır. Fakat makine öğrenmesi ve doğal dil işleme teknolojileri henüz yeterince ilerlemediği için, o dönemde sadece kelime eşleştirerek yanıt oluşturabilmiştir. Chatbot terimi ilk olarak "chatter bot"[3] olarak Michael L. Mauldin tarafından 1994 yılında kullanılmıştır.

Sohbet botları dijital ortamda kullanıcıya bilgi vermek, bir görevi yerine getirmek, çeşitli konularda işlem yapabilmek gibi kullanıcıya yardım etmek amacıyla ile çalışan yapay zekalardır. Amaçları doğrultusunda zaman ve enerji tasarrufu sağlamaktadırlar.

Sohbet botlarını kullanım amaçlarına göre sınıflara ayırmak mümkündür. Örneğin hangi platformda yayınlanacağı, hangi amaç doğrultusunda geliştirildiği veya hangi teknoloji ile geliştirildiği konusunda sınıflandırabiliriz.

1.1.1 Kullanılan Sohbet Bot Türleri

Kelime bazlı sohbet botlar: Bu tür sohbet botlarda, bir yapay zeka teknolojisinden söz etmek mümkün değil. Kullanıcı mevcut menü ve butonlardan seçimler yaparak veya bazı ifadeler yazarak ilerliyor. Ön yüzde menü ve butonlar, arka tarafta da daha çok IVR sistemlerinde gördüğümüz karar ağaçları kullanılıyor. Sistem, yazılan ifadelerdeki bazı kelimeleri tanıyıp, uygun bir cevap ile eşleştirmeye çalışıyor. Bu tür sohbet botlar, kısıtlı sayıda ve türde soruyu yanıtlayabiliyor. Ancak kelime bazlı sohbet botlar, yazılan ifadeyi bir bütün olarak anlamlandırmada yetersiz kalıyor. Kelimelerin yanlış yazımında kullanıcıyı doğru yönlendirebilmek için olası tüm yanlış yazımların sisteme tanıtılması gerekiyor. Çok basit soru-cevap sohbet botlarında kullanılabilirler ancak karmaşık durumlarda çalışmazlar. Bir örnekle açıklık getirmek istenirse, yazılımda İngilizce-Türkçe dönüşümü tanımlanmış olduğunu varsayalım. Kelime bazlı bir sohbet bota kullanıcı “Hi Türkçede ne demek?” sorusunu sorduğunda, sohbet bot “hi” kelimesini tanıyarak Türkçe karşılığını verir, ifadenin gerçekten ne sorduğunu anlayamaz.

İfade bazlı sohbet botlar: Bu tür sohbet botlarda bir NLP ve makine öğrenmesi teknolojisinden söz etmek mümkün. Sohbet bot, kullanıcının yazdığı ifadeyi bir bütün olarak tanıyor ve anlamlandırıyor. Ancak karşılıklı konuşma devam ettiğinde, diyaloga bağlı kalamıyor. Dil dönüşümleri ile ilgili örnek soru üzerinden devam edecek olursak, bu tür bir sohbet bot “Hi Türkçede ne demek” sorunu bir ifade olarak anlıyor, ancak ardından sorulan “Peki, Almancada ne demek?” sorusunu anlamada yetersiz kalıyor. Çünkü bu noktada diyalogu başından itibaren izleme ve bağlama uygun şekilde sürdürebilme yetkinliği devreye giriyor.

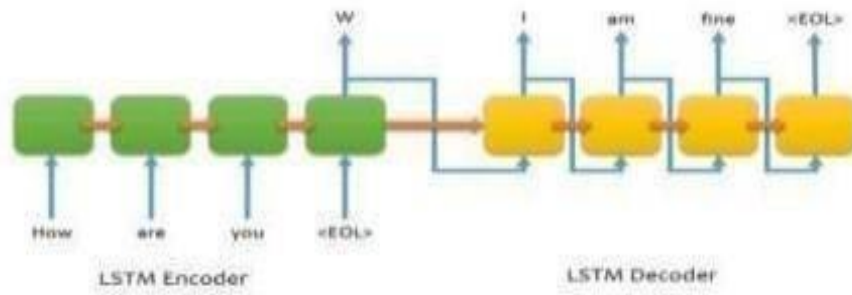
3. Diyalog bazlı sohbet botlar: Bu tür sohbet botlar da yapay zeka tabanlı sohbet botlar ve NLP (Doğal dil işleme) ve makine öğrenmesi / derin öğrenme teknolojilerini kullanarak geliştiriliyorlar. Diyalog bazlı sohbet botlar kelime veya ifade tabanlı bir anlayışla sınırlı kalmıyor, yazılan ifadeyi anlamlandırabildikleri gibi, konuşmanın akışını takip edip, bu akışına göre cevaplar da üretebiliyor. Aynı örneği bir de diyalog bazlı sohbet bot için verirsek, “Hi Türkçede ne demek?” sorusunu ifade olarak anlayıp cevap verir, ardından sorulan “Peki Almancada ne demek?” sorusunda neden söz edildiğini, nasıl bir dönüştürme yapacağını anlar ve doğru sonucu çıkartır.

Bu 3 türden hangisi daha iyi diye sorulursa öncelikle kullanım amacına bakmak gerekmektedir. Örneğin bir inşaat şirketinde etkileşim alanı sınırlı olacağı için ifade bazlı bir sohbet bot yeterli olabilir. Sohbet bot türüne karar vermek için kullanıcı için yaratacağı değere bakmak da önemli. Eğer kullanıcı için yaratılacak değer, kelime bazlı veya ifade bazlı bir sohbet bot ile sağlanıyorsa, diyalog bazlı bir sohbet bot geliştirmek anlamlı olmayabilir. Bir diğer parametre, hedef kullanıcıların tercihleri. Hedef kitleyi çok iyi tanımak ve onların beklentilerine uygun bir sohbet bot geliştirmek önemli. Bazı kullanıcı tipleri menülerle ilerlemek isteyebilir, bazı gruplar bir diyalogu serbest şekilde sürdürmeyi tercih edebilir. Bu 3 etmene dikkat ederek sohbet bot türünü belirlemek en uygun çözüm.

Sonuç olarak bu 3 amaca en iyi şekilde hizmet verecek sohbet botu bulmak için önce kendi stratejiniz içinde sohbet botlarının yerinin belirlenmesi, sohbet bottan nasıl bir soruna çözüm getireceğinin tanımlanması ve buna göre nasıl bir sohbet bot istendiğinin ortaya konması gerekiyor. Ayrıca botun gelişmişlik seviyesi de önemlidir. Botun çok daha efektif kullanılması için nasıl bir seviye istendiği de belirlenmelidir.

2. LİTERATÜR

Derin öğrenme, yapay zekâ ve makine öğrenmesi alanlarında çok önemli bir konu haline gelmiştir. Son yıllarda, farklı derin öğrenme yöntemlerini önermekte olan çalışmaların ve mevcut yöntemleri değişik problemler üzerinde uygulayan çalışmaların sayıları hızla artmaktadır. Derin Öğrenme Yöntemleri; Derin Sinir Ağları (Deep Neural Networks), Derin Oto-Kodlayıcılar (Deep Autoencoders), Derin İnanç Ağları (Deep Belief Networks), Derin Boltzmann Makinesi (Deep Boltzmann Machine), Yinelenen Sinir Ağları (Recurrent Neural Networks), Evrişimsel Sinir Ağları (Convolutional Neural Networks)ndan oluşmaktadır. Otomatik soru cevaplama sistemi için Evrişimsel Sinir Ağları (Convolutional Neural Networks) ve LSTM (Long Short-Term Memory) kullanılmaktadır. Evrişimsel sinir ağları, çok boyutlu girdiler için ve özellikle iki boyutlu görsel veriler için önerilmiş bir derin öğrenme yöntemidir. Otomatik Soru Cevaplama Sistemi; doğal dilde ifade edilen sorunun cevabını doğal dildeki büyük veri kümeleri içerisinde arayıp bularak yine doğal dilde otomatik olarak cevaplandırmayı planlayan sistemdir [1].



Şekil.2 LSTM Mimarisi [1]

Chatbot, işitsel veya metinsel olarak kullanıcı ile sohbet edebilen bir yazılımdır. Gelişmiş chatbotlar, ilgili konularda gerekli cevapları verebilmektedir. Chatbotta yapay zekâ kullanılması verimliliği yükseltmektedir. Sıradan Sıraya Modeli (Sequence to Sequence Model): RNN mimarisine dayanmaktadır. İki adet RNN den oluşmaktadır. Bu iki RNN kodlayıcı ve kod çözücüdür. Kodlayıcı girdiyi kontrol etmekte olup girişte bir dizi alıp her adımda bir sembol işlemektedir. Bundaki amaç gereksiz bilgileri kaybedip her adımda önemli bilgileri kodlayan bir sembol dizisini sabit boyutlu bir özellik vektörüne dönüştürmektir. • Kelime Gömme (Word Embedding): Düşük boyutlu bir vektör uzayda kelimelerin yoğun temsilini öğrenmek için uygulanan tekniktir. • LSTM Modeli (Long Short-Term Memory): Basit RNN' nin giriş ve çıkış kapılarına ek olarak, özel unutkan kapıları olan özel bir RNN çeşididir. Bu kapsamda, yapılan çalışmada LSTM (Uzun Kısa Süreli Bellek) ve seq2seq modeli ile telegram bot uygulaması geliştirilmiştir. Çalışmada LSTM kullanılması bir sonraki konuşma eylemini tahmin etmek için konuşma geçmişine geri dönülebilmesini sağlamıştır. [26]

Akıllı sohbet botlarının insan hayatına adapte oluşu yaygınlaştıkça, insanların bu botlardan beklentileri değişebilmektedir. Görev gerçekleştirme amacı güden sohbet robotlarına, bu görevlerin üstüne günlük diyalog ve arkadaşlık gibi görevler de eklenebilmektedir. Kullanıcı ihtiyaçlarını karşılamak ve kullanıcılarla insan-insan benzeri bağ kurmak için, sohbet robotlarının daha insancıl ve gelişmiş dil özelliklerini içermesi beklenmektedir. Bu dil özelliklerinden birisi de metafor kullanımıdır. Bu konuda yapılan bir çalışmada, konuya duyarlı ve konu hakkında metafor üretebilen sohbet robotu tasarlamak hedeflenmiştir. Yine bu çalışmada insan gözlemcilere, üretilen metaforların yeniliğini ve uygunluğunu sorgulatılmıştır. Bu çalışmadaki deneyler, kurulan sistemin sohbet robotumuzla iletişim kurma konusunda kullanıcıların ilgisini etkin bir şekilde uyandırdığını ve bunun sonucunda önemli ölçüde daha uzun insan-chatbot sohbetleri sağladığını gösteriyor [15].

Canlı sohbet arayüzleri aracılığıyla müşterilerle iletişim kurmak, e-ticaret şirketlerinin müşteri hizmeti sağlamak için başvurduğu yöntemlerden biridir. Günümüzde, arka planda müşteri hizmet çalışanının olduğu sohbet hizmeti araçlarının yerini, genellikle yapay zekaya dayalı doğal dil aracılığıyla insan kullanıcılarla iletişim kurmak için tasarlanmış yazılımlar olan sohbet botu araçları almaktadır. Maliyet, zaman ve iş gücünden tasarruf sağlayan bu çözüm, yapay zeka tabanlı sohbet robotlarının hızla yayılmasını sağlamış olsa da yine de müşteri beklentilerini karşılamada başarısız olabildiklerini ve müşterilerin sohbet robotu tarafından yapılan taleplere uymaya daha az meyilli olabildiklerini gözlemlenebilmektedir. Bağlılık-tutarlılık teorisinden yardımıyla, sözlü antropomorfik tasarım ipuçlarının ve foot-in tekniğinin kullanıcı isteği uyumluluğunu nasıl etkilediğini çevrimiçi bir deney aracılığıyla incelenmiştir.

Alınan sonuçlara göre

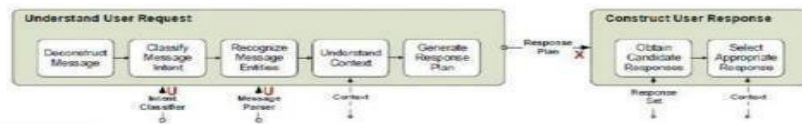
, hem antropomorfizmin hem de tutarlı kalma ihtiyacının, müşterilerin bir sohbet robotunun hizmet geri bildirim talebine uyma olasılığını kayda değer ölçüde artırdığını göstermiştir. [4]

Derin Öğrenme, Makine öğrenimini orijinal ve ana hedeflerinden birine, yani Yapay zekâya yaklaştırmak amacıyla tanıtılan bir Makine Öğrenimi ve araştırma alanıdır. Otomatik öğrenme algoritmasından bahsederek, bunlar doğrusal olma eğilimindedir, DL (Derin Öğrenme) algoritmaları karmaşıklığı ve soyutlamayı artıracak şekilde yapılandırılmıştır. Derinlemesine öğrenmek için ilk kelimesi kedi olan 5 yaşında bir çocuk düşünüm. Çocuk, nesneleri göstererek ve kedi kelimesini söyleyerek kedinin ne olduğunu öğrenmeye devam eder. Anne, "Evet, o bir kedi" veya "Hayır, o bir kedi değil" der. Çocuk nesneleri işaret etmeye devam ettikçe, tüm kedilerin özelliklerinin ve özelliklerinin daha çok farkına varır. [3]

Yapılan bir çalışmada, lokanta ortamında müşterilerden gelen soruları okuyan, her bir kelimeyi analiz eden ve ilgili anahtar kelimelere göre cevabı oluşturmaya çalışan bir sohbet robotu tasarlamak amaçlanmıştır [14].

Yapılan bu çalışmaya göre sohbet robotu, siparişleri almaya ve restoran hakkında daha fazla bilgi vermeye hazır garsonu temsil etmektedir. Kullanıcı; Menüden bir veya daha fazla yemek siparişi verebilir, restoranın günlük açılış ve kapanış saatleri hakkında bilgi alabilir, restoranın sipariş alma ile ilgili bilgilerini alabilir, siparişlerini içeren tam listeyi alabilir. Bu çalışmadan yola çıkarak da sohbet robotlarının kullanım alanlarının ne denli geniş olduğu sonucuna varabiliriz.[14]

Bir sohbet-bot geliştirme çalışmasında genellikle son aşama, botu istenilen ortama entegre etmektir. Bu aşama temel olarak, botu entegre ettiğiniz ortamda kullanıcıların botla iletişim kurmaya başlamasına izin vermenizdir. [13]. Botunuzu oluştururken halihazırda hedef ortamınızın belli olması gerekmektedir, çünkü farklı ortamlar farklı tipte diyalogları gerçekleştirebilir ve farklı özellikleri destekleyebilir. Örneğin, Microsoft Teams, Uyarlanabilir Kartlar gibi farklı içerikleri işleyebilirken Twilio metin mesajları için kullanılır. Bu verilen örnekte, metin mesajları Uyarlanabilir Kartlar'ı görüntüleyemediğinden, Twilio ortamında Uyarlanabilir Kartları kullanamazsınız. Bu nedenle, sohbet-botu'nuzun ilerleyen güncellemelerinde büyük değişiklik taleplerinden kaçınmak için geliştirme aşamasında mutlaka hedef kanalları düşünmeniz gerekir. [13]



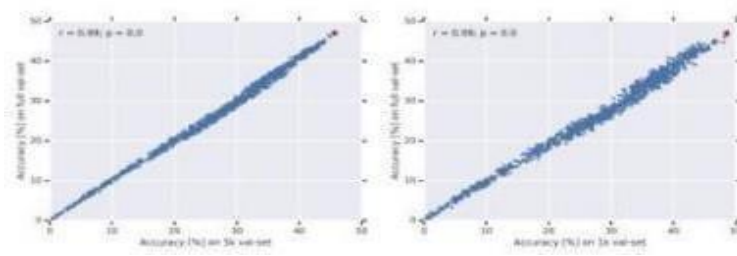
Şekil.3 Alma Tabanlı Sohbet Robotları için Süreç Mimarisi Modeli [11]

Bu çalışmada görüntü sınıflandırıcıların yarı denetimli öğrenme problemi ele alınmıştır. Ana hedef, yarı denetimli öğrenme alanının, hızla gelişen kendi kendini denetleyen görsel temsili öğrenme alanından faydalanılabileceğidir. Bu iki yaklaşımı birleştirerek, kendi kendine denetimli, yarı denetimli öğrenme çerçevesi önerilmiştir ve bunu iki yeni yarı denetimli görüntü sınıflandırma yöntemini türetmek için kullanılır. [16]

NBA ve EPL (İngiltere Premier Ligi) dahil olmak üzere büyük spor ligleri, oyun bilgileri sunmak ve taraftarlarının içeriklere ulaşmalarını kolaylaştırmak için yenilikçi bir çıkış noktası olarak sohbet botlarını benimsiyor. Bununla birlikte, mevcut spor sohbet robotları, istenen istatistiksel verileri karşılamakta genellikle yetersiz olan skor ve maçın öne çıkan videoları sağlayabilmektedir. Yapılan bir çalışmada, spor hayranlarının daha fazla oyun istatistik verilerini keşfetmeleri için etkileşimli bir sohbet robotu olan GameBot adında sohbet robotu tasarlanmıştır. [22]

Yinelemeli bir geliştirme süreciyle bir sohbet robotu oluşturmak, sohbet robotu geliştiricisi için belirli zorluklar doğurur. Geliştiriciden beklenen şey, kısa bir döngü sonunda botun hizmete hazır olan ilk sürümü üretmesidir. Bahsedilen döngünün her yinelenmesi, sohbet robotunun kapasitesini aşamalı olarak ilerletmeli ve bir öncelik listesine dayalı olarak genel kullanıcı öykülerinin bir alt kümesini uygulamalıdır. Bu bağlamda, ticari sohbet robotu oluşturma platformları, geliştiricinin bu kullanıcı hikayelerini belirli bir biçimde eşleyebilmesi koşuluyla, sohbet robotu geliştiricisine birçok avantaj sağlar. Bunu yapmak için, sohbet robotunun yanıtlaması beklenen her sorgu için geliştiricinin kullanıcının isteklerini değerlendirmesi gerekir. Amaca bağlı olarak, sorgunun amaca hizmet edecek işlenmesi gerekir, bu da bazı iş mantığının yürütülmesini gerektirebilir. Ek olarak, sorgunun işlenmesi, kullanıcının sohbet robotu ile konuşmanın bir parçası olarak sağlaması gereken belirli veri öğelerini gerektirebilir. Böylece sohbet botu, karşılaşılabileceği bir "bağlam" ve bağlam gözlemlendiğinde gerçekleşmesi gereken "tepki" sağlanarak tanımlanır. [23]

Modern bilgisayarlı görü sistemleri, görüntü tanıma, nesne algılama, anlamsal görüntü bölütleme vb. gibi çeşitli zorlu bilgisayarlı görü kısıtlarında olağanüstü performans sergilendiği görülmektedir. Beraberinde getirdiği sonuçlar, büyük miktarda zaman alan ve elde edilmesi pahalı olan açıklamalı veriler. Birçok gerçek dünya bilgisayarlı görü uygulaması, standart kıyaslama veri kümelerinde olmayan görsel kategorilerle veya görsel kategorilerin veya görünümünün zaman içinde değişebileceği dinamik nitelikteki uygulamalar ilgili olduğudur. Fakat, tüm bu senaryolar için büyük etiketli veri kümeleri oluşturmak pratik olarak mümkün değildir. Dolayısıyla, yalnızca az miktarda etiketlenmiş örnekten yararlanarak yeni kavramları tanımayı başarılı bir şekilde öğrenebilen bir öğrenme yaklaşımı tasarlamak önemli bir araştırma zorluğu olarak tanımlanmış, insanların sadece birkaç örnek gördükten sonra yeni kavramları çabucak anlamaları ve bu amaca prensipte ulaşabileceğinden bahsedilir [16].



Şekil.4: Doğrulama setindeki doğrulama puanı arasındaki korelasyon [16]

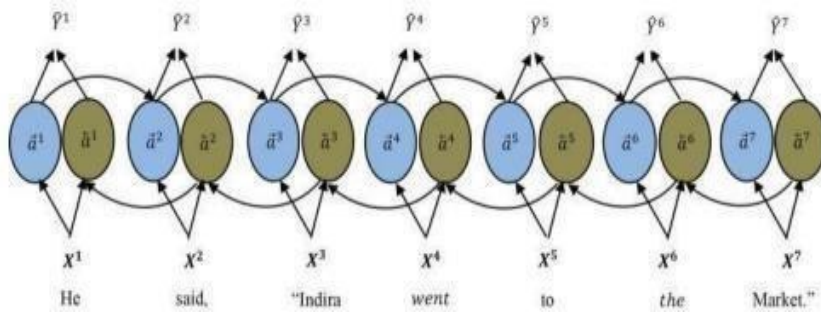
Akademik danışmanlık, öğrencilerin akademik hayatında önemli rol oynar; ancak, zaman alıcı ve sürdürülmesi zordur. Yapılan bir çalışmada, hızlı, ölçeklenebilir ve uygun maliyetli bir sohbet robotu kullanarak etkili çözüm üretme amaçlanmıştır. Bu çalışmada tasarlanan bot, Microsoft Teams ile entegre edilebilen bir bottur. Natural Language Processing'den (NLP) gelişmiş semantik analiz tekniklerini uygular ve öğrencinin sorgularının bağlamını analiz eder ve buna göre yanıt verir [21].

Başlıca değinilen konular:

- Kendi kendini denetleyen öğrenmedeki son gelişmelerden yararlanan doğal görüntülerle yan denetimli öğrenme için yeni bir teknikler ailesi önerilir.
- Kendi kendini denetleyen yarı denetimli tekniklerin, etiketlenmemiş veriler olmadan eğitilmiş dikkatle ayarlanmış temel çizgilerden daha iyi performans gösterdiğini ve önceden önerilen yarı denetimli öğrenme teknikleriyle rekabet edebilir performans ettiği gözlemlenmiştir.

Bir gün içerisinde çok fazla miktarda veri işlenmekte ve veri alışverişi yapılmakta ve veri tabanı uzun süredir verileri düzenlemek ve işlemek için kullanılan ve aktif olarak üzerine çalışılan bir araştırma konusudur. Veritabanı konusu birçok bilgisayar sisteminde büyük role sahiptir. Konuyla ilgili kişi ya da kişilerce her zaman yapılan işlemleri kolaylaştırma ve harcanan vakti azaltma amacıyla tercih sebebidir. Bir veritabanıyla doğrudan yolla etkileşim kurmak için “doğal dil” i kullanmak, güzel ve kullanıcıya kolaylık sağlayan bir çözümdür. İnsan ve bilgisayar arasındaki bu tarz bir iletişimi sağlayabilmek için bilgisayarın kendinden ne istendiğini anlaması ve istenilen doğrultusunda doğru cevap veriyor olabilmeliyiz. Doğal dil işleme (NLP), Doğal Dil anlama veya Doğal Dil oluşturma veya hatta ikisini birlikte gerektiren birçok uygulamaya sahiptir. Örnek olarak, metni bir insan dilinden diğerine kendiliğinden çevirmeye odaklanan Makine Çevirisi. NLIDB sistemler fikri, veritabanını sorgulamak için veritabanı dili yerine Doğal Dil sorgularını kullanan veritabanı sorgu yönteminden gelir. Kullanıcı, çok basit bir şekilde veritabanıyla etkileşim kurmak için Doğal Dil’i kullanabilir hatta kullanıcının bununla ilgili herhangi bir eğitime sahip olup olmaması sorun olmayacaktır. Ayrıca birden çok veritabanı tablosu içeren sorgularda Doğal Dil’i kullanmak daha kolaydır [17].

Chatbot için derin öğrenme hesaplamasındaki modellemeyi ve performansı gösterir. Neural Machine Translation için Tensorflow yazılım kitaplığının kullanımı Modelleme için data elde etmek en önemli işlemlerden biridir ve onu işleyebilmek oldukça zordur. Dikkat katmanlarını içeren Çift Yönlü Tekrarlayan Sinir Ağları (BRNN) kullanılır, böylece çok sayıda zarf içeren giriş cümlesi daha uygun bir şekilde cevaplanabilir. İnsanoğlu gerekli cevapları alabilmek için sorduğu sorularda çok detay verir. Verilen bu detaylar içerisinde chatbotların en doğru cevapları verebilmesi için BRNN ve Attention modelinin birleşimini mükemmel yapan sebep budur. BRNN yapısı Şekil.2 te gösterilmektedir. Sinyaller ileri ve geri olarak 2 tiptedir. $A \rightarrow$ ileri yinelenen bileşendir ve $a \leftarrow$ geriye doğru yinelenen bileşendir [2]. Birçok aktivasyon fonksiyonu türü vardır, ancak BRNN ile en çok sigmoid ve lojistik aktivasyon fonksiyonları kullanılır. Örneğin, Şekil.2'te gösterilen ağda, birkaç ifadeden oluşan bir girdi verilmiştir. Açıklama 1- He said, "Indira went to the market." ve açıklama 2- He said, "Indira Gandhi was a great Prime Minister." Bu durumda, deyim 2 ağa eklendiğinde, "Indira" kelimesinin eklendiği adım 3'te, Y^3 hücre tahmini veya çıkış sinyali kontrol edilir ve çift yönlü olması nedeniyle, ileri bilgi 1. hücreden akar ve 2. veya 3. hücrenin yanı sıra, 9. hücreden 3. hücreye doğru geriye bilgi akışı ile 2. ifadedeki 'Indira'nın Başbakan olduğunu ve kendisine giden 'Indira' olmadığını tahmin etmesine yardımcı olur [2].



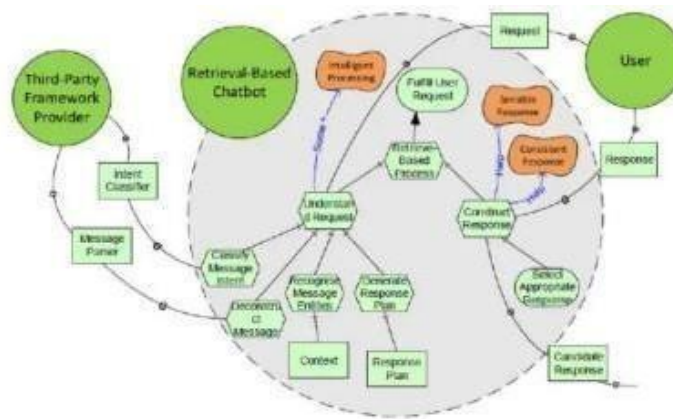
Şekil.5 İleri ve geri bilgi akışını tanımlamak için gizli katman ile BRNN'ye kaynak girişi gösterimi [2]

Bilgi iletişimindeki dönüşüm yirminci yüzyılın son yirmi yılında gerçekleştiğinde medya endüstrisinde de önemli değişiklikler meydana geldi. Gazetecilik yöntemleri multimedya ve etkileşimli özelliklerle zenginleştirilmiştir, ancak iletişimin temel medyası (metin, fotoğraflar ve videolar) sabit kalır. Son 10 yılda sohbet robotlarının kullanıma sunulması, doğal dil (Natural Language) biçiminde bile gerçekleşebilen etkileşimli sohbetleri kullanarak gazeteciliği dönüştürmek için yeni imkanlar sağlamıştır. [18]

Bilgi portalları genellikle göçmenlerin ev sahibi ülkeye uyum sağlama süreçlerini desteklemek için oluşturulur. Bununla birlikte bilgi arama süreci, doküman listelerinden istenen bilgileri almanın uzun sürebileceği için göçmenler için yorucu ve karıştırıcı olabilir. Sohbet robotlarının kullanımı kolay, doğal ve sezgiseldir ve bu nedenle bilgi aramayı destekleyebilir. Sohbet botu geliştirmede ortak tasarım katılımcıları olarak göçmenleri ve diğer paydaşları dahil eden ve güçlendiren araştırma eksikliği olduğu düşünülmektedir ve göçmenlerin sosyal entegrasyonlarını destekleyen bir sohbet robotu tasarlamayı hedefleyen çalışma mevcuttur. Bu çalışmada ortak tasarım yaklaşımı kullanarak, göçmenler çevrimiçi anketler, empati araştırmaları, anketler ve ortak tasarım atölyeleri gibi etkinlikler yürütülmüştür. [19]

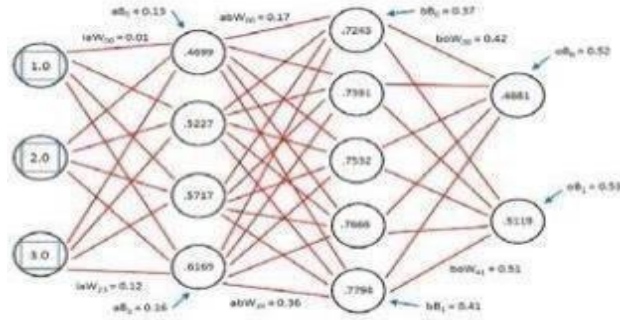
Yapılan bir çalışmada Hindistan'da ülkenin tamamına etki eden covid-19 sorunun üstesinden gelinmesine yardımcı olmayı amaçlayan sohbet robotu geliştirilmiştir. Hastalar, hastanelerdeki yatak gibi kritik kaynakları elde etmek için hastanelerin dışında uzun kuyruklarda beklemekte fakat zaman zaman yeterli yatak olmaması nedeniyle elleri boş yere geri dönebilmektedirler. Aynı çalışmada önerilen sohbet botu, kullanıcıların seyahat etmelerine gerek kalmadan tüm hastanelerdeki mevcut yatakları ve gerekli tıbbi yardımın bulunup bulunmadığını sorgulayabilmeleri hedeflenmiştir. Çalışmada bahsedilen deneysel sonuçlara göre önerilen sohbet botunun, yataklar ve oksijen tüplerinin mevcudiyeti hakkında hastane bilgileriyle ilgili tüm kullanıcıların sorularına başarıyla yanıt verdiğini gösteriyor. [20]

Bu araştırmanın amacı, kendisine verilen önermeye dayalı olarak kişiye tavsiye vererek yanıt veren bir hizmet olarak bir danışman Chatbot oluşturmaktır. Xen, kullanıcılarla yaptığı konuşmalara dayanarak kullanıcının ne iletmeye çalıştığını öğrenir ve anlar. Üretken sohbet robotları, veriler üzerinde eğitim alır ve verileri anlamak için kurallar oluşturarak içerik sınıflandırması veya içerik sınıflandırması temelinde çalışır ve böylece tüm senaryolar için çalışabilir. Bağlam sınıflandırmasında hangisinin ne anlama geldiğini ve ne anlama gelebileceğini anlamak için birçok bilginin analiz edilmesi gerekir. Bağlam, muğlaklık, eksik ve yanlış verilerle birlikte birçok anlam taşıyabilir. İnsanlar tarafından sohbet etmek veya sorunlarını anlamak ya da sadece bir düşünce akışına ihtiyaç duyduklarında konuşmak için kullanılabilir. Önerilen sohbet robotları, kullanıcıyla etkileşim kurarak yeterli veriyi topladıktan sonra, K-Means kullanarak kendi kendine öğrenmeyi kullandığı Alım, AIML (Yapay Zeka İşaretleme Dili) ve Üretken sohbet botudur. [10] Üretken sohbet robotu, yanıtlarını iyileştirebilmesi ve kullandığı yanıtları genişletebilmesi için bağlam sınıflandırmasını kullanarak kendini eğitir. Önerilen sohbet robotu Xen, verildiği bağlama göre uygun yanıtlar verir ve uygun şekilde yanıt verir ve tasvir edilen duyguyu iyi anlar. Yaptığı konuşmalar üzerinde eğitim alır ve bir dahaki sefere benzer bir bağlamda daha uygun yanıtlar verir ve ayrıca daha fazla duruma yanıt verebilir. [10]



Şekil.6 i* Erişim Tabanlı Sohbet Robotları için Stratejik Mantık Modeli [11]

Derin Öğrenme, Makine öğrenimini orijinal ve ana hedeflerinden birine, yani Yapay zekâya yaklaştırmak amacıyla tanımlanan bir Makine Öğrenimi ve araştırma alanıdır. Otomatik öğrenme algoritmasından bahsederseniz, bunlar doğrusal olma eğilimindedir, DL (Derin Öğrenme) algoritmaları karmaşıklığı ve soyutlamayı artıracak şekilde yapılandırılmıştır. Derinlemesine öğrenmek için ilk kelimesi kedi olan 5 yaşında bir çocuk düşünün. Çocuk, nesneleri göstererek ve kedi kelimesini söyleyerek kedinin ne olduğunu öğrenmeye devam eder. Anne, "Evet, o bir kedi" veya "Hayır, o bir kedi değil" der. Çocuk nesneleri işaret etmeye devam ettikçe, tüm kedilerin özelliklerinin ve özelliklerinin daha çok farkına varır. Çocuğun ne yaptığını bile bilmeden ne yaptığı çok karmaşık bir soyutlama düzeyine açıklık getirir ve her bir soyutlama düzeyinin, ölçekte bir önceki katmanda edindiği bilgi yardımıyla bir ölçek oluşturur. Derin öğrenme kullanan sistemde çalışan yazılımlar da aynı süreçte ilerler. Her algoritma girdisine doğrusal olmayan bir yaklaşım uygulayacak şekilde öğrendiklerini verilen çıktıdan istatistiksel bir model oluşturmak için kullanır. Çıktı kabul edilebilir bir doğruluk noktasına ulaşana kadar tekrar devam eder. [3]



Şekil.7 Sinir Ağı [3]

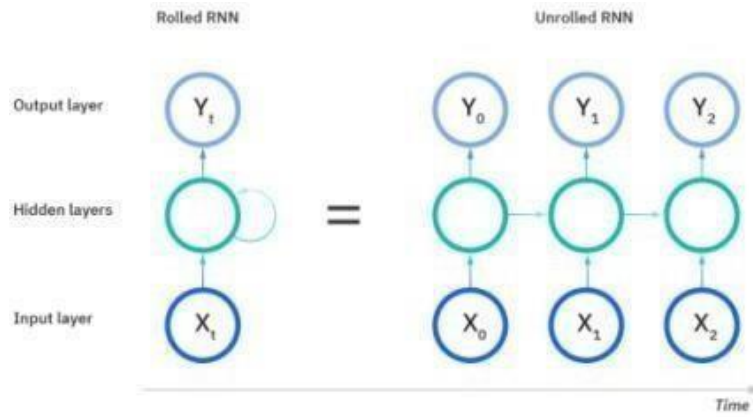
3. EVRİŞİMLİ SİNİR AĞLARI

3.1 Yinelemeli Sinir Ağı (Recurrent Neural Network) Rnn Encoder-Decoder

Yinelemeli Sinir Ağı (RNN), sıralı verileri veya zaman serisi verilerini kullanan bir tür yapay sinir ağıdır. Bu derin öğrenme algoritmaları, dil çevirisi, doğal dil işleme (NLP), konuşma tanıma ve resim yazısı gibi sıralı veya zamansal sorunlar için yaygın olarak kullanılır; Siri, sesli arama ve Google çeviri gibi popüler uygulamalara dahil edilirler. İleri beslemeli ve Evrişimli Sinir Ağları (CNN'ler) gibi, Yinelemeli Sinir Ağları da öğrenmek için eğitim verilerini kullanır. Mevcut girdi çıktıyı etkilemek için önceki girdilerden bilgi aldıkları için “hafızaları” ile ayırt edilirler. Geleneksel derin sinir ağları, girdilerin ve çıktılarının birbirinden bağımsız olduğunu varsayarken, yinelemeli sinir ağlarının çıktısı, dizi içindeki önceki öğelere bağlıdır. Gelecekteki olaylar, belirli bir dizinin çıktısının belirlenmesinde de yardımcı olurken, tek yönlü tekrarlayan sinir ağları, bu olayları tahminlerinde açıklayamaz. Bir Yinelemeli Sinir Ağı (RNN), bir sembol dizisini sabit uzunlukta bir vektör temsiline kodlar ve diğeri, gösterimi başka bir sembol dizisine çözer. Önerilen modelin kodlayıcısı ve kod çözücüsü, bir kaynak dizisi verilen bir hedef dizinin koşullu olasılığını en üst düzeye çıkarmak için ortaklaşa eğitilir. İstatistiksel bir makine çevirisi sisteminin performansının, mevcut log-lineer modelde ek bir özellik olarak RNN EncoderDecoder tarafından hesaplanan tümce çiftlerinin koşullu olasılıkları kullanılarak ampirik olarak iyileştirildiği bulunmuştur. Niteliksel olarak, önerilen modelin dilsel ifadelerin anlamsal ve sözdizimsel olarak anlamlı bir temsili öğrendiği gözlemlenir.

RNN'lerin açıklanmasında bize yardımcı olması için, ingilizcede bulunan “havanın altında hissetmek” deyiminin anlamlı olması için, o belirli sırayla ifade edilmesi gerekir. Sonuç olarak, yinelemeli ağların deyimdeki her kelimenin konumunu hesaba

katması gerekir ve bu bilgiyi dizideki bir sonraki kelimeyi tahmin etmek için kullanırlar. Aşağıda verilmiş olan görselde, RNN'nin “yuvarlanmış” görseli, tüm sinir ağını veya daha doğrusu "havanın altında hissetmek” gibi öngörülen tüm ifadeyi temsil ediyor. “Kaydırılmamış” görsel, sinir ağına tek tek katmanlarını veya zaman adımlarını temsil eder. Her katman, “hava durumu” gibi o ifadedeki tek bir kelimeyle eşlenir. “Hissetme” ve “altında” gibi önceki girdiler, “the” dizisindeki çıktıyı tahmin etmek için üçüncü zaman adımında gizli bir durum olarak temsil edilecektir.[25]



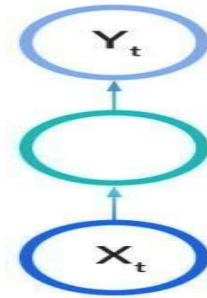
Şekil.8 ,RNN'nin “yuvarlanmış” görseli [25]

Tekrarlayan ağların bir diğer ayırt edici özelliği, ağına her katmanında parametreleri paylaşmalarıdır. İleri beslemeli ağlar her düğümde farklı ağırlıklara sahipken, tekrarlayan sinir ağları ağına her katmanında aynı ağırlık parametresini paylaşır. Bununla birlikte, bu ağırlıklar, pekiştirmeli öğrenmeyi kolaylaştırmak için geri yayılım ve gradyan iniş süreçlerinde hala ayarlanır. Tekrarlayan sinir ağları, dizi verilerine özgü olduğu için geleneksel geri yayılımdan biraz farklı olan gradyanları belirlemek için zaman içinde geri yayılım (BPTT) algoritmasından yararlanır. BPTT'nin ilkeleri, modelin çıktı katmanından girdi katmanına hataları hesaplayarak kendini eğittiği geleneksel geri yayılımla aynıdır.

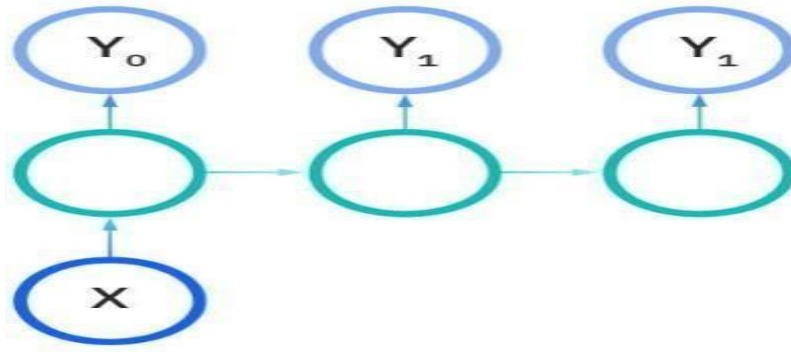
Bu hesaplamalar, modelin parametrelerini uygun şekilde ayarlamamızı ve uydurmamızı sağlar. BPTT geleneksel yaklaşımdan farklıdır, çünkü BPTT her zaman adımında hataları toplar, ileri beslemeli ağlar ise her katmanda parametreleri paylaşmadıklarından hataları toplamaya ihtiyaç duymazlar. Bu süreç boyunca, RNN'ler, patlayan gradyanlar ve kaybolan gradyanlar olarak bilinen iki problemle karşılaşma eğilimindedir. Bu sorunlar, hata eğrisi boyunca kayıp fonksiyonunun eğimi olan gradyanın boyutu ile tanımlanır. Gradyan çok küçük olduğunda, küçülmeye devam eder, ağırlık parametreleri önemsiz hale gelene kadar güncellenir. Bu gerçekleştiğinde, algoritma artık öğrenmiyor. Patlayan gradyanlar, gradyan çok büyük olduğunda meydana gelir ve kararsız bir model oluşturur. Bu durumda, model ağırlıkları çok büyüyecek ve sonunda NaN olarak temsil edilecektir. Bu sorunlara bir çözüm, sinir ağı içindeki gizli katmanların sayısını azaltarak RNN modelindeki karmaşıklığın bir kısmını ortadan kaldırmaktır.

3.2 Tekrarlayan Sinir Ağları Türleri

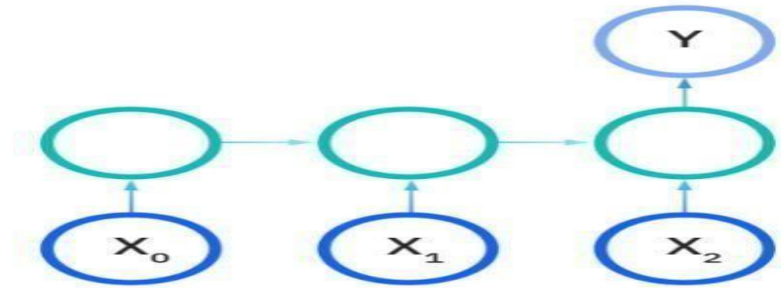
İleri beslemeli ağlar, bir girdiyi bir çıktıya eşler ve yukarıdaki diyagramlarda tekrarlayan sinir ağlarını bu şekilde görselleştirirken, aslında bu kısıtlamaya sahip değildir. Bunun yerine, giriş ve çıkışlarının uzunluğu değişebilir ve müzik oluşturma, duygu sınıflandırması ve makine çevirisi gibi farklı kullanım durumları için farklı RNN türleri kullanılır. Farklı RNN türleri genellikle aşağıdaki diyagramlar kullanılarak ifade edilir:



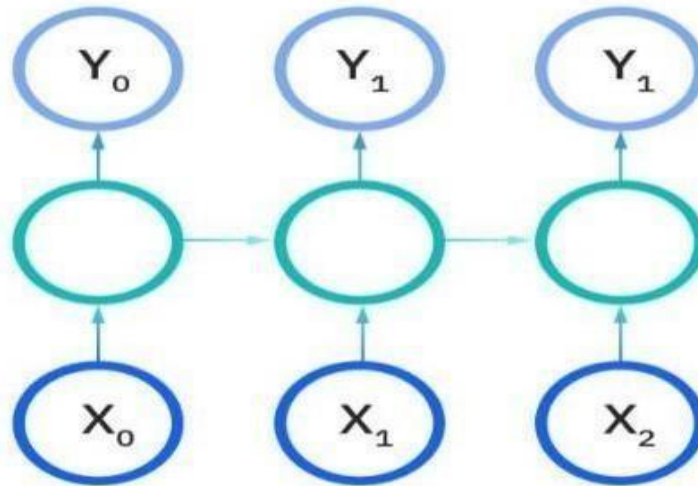
Şekil.9 Bir Girdinin Bir Çıktıya Eşlenmesi [26]



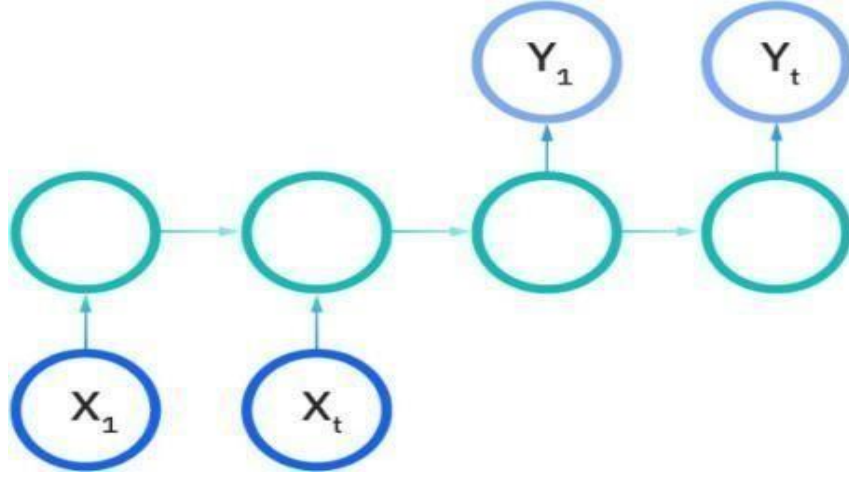
Şekil.10 Bir Girdinin Birden Çok Çıktıya Eşlenmesi [26]



Şekil.11 Birden Fazla Girdinin Bir Çıktıya Eşlenmesi [26]



Şekil.12 Birden Fazla Girdinin Birden Fazla Çıktıyla Eşlenmesi [26]

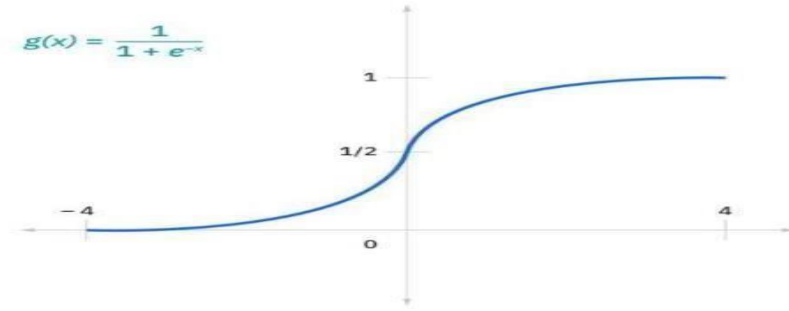


Şekil.13 Birden Fazla Girdinin Birden Fazla Çıktıyla Eşlenmesi
[26]

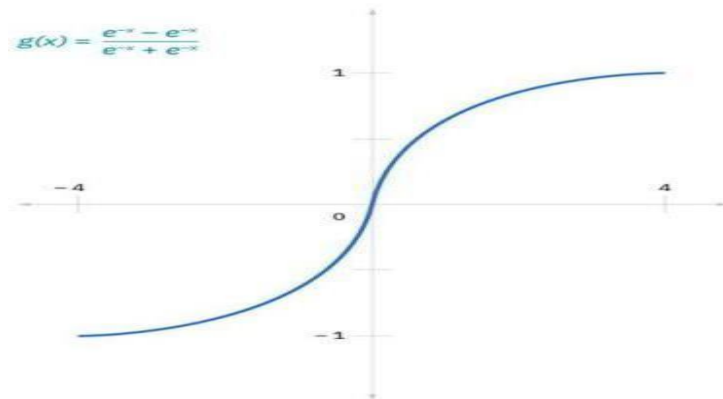
Tekrarlayan bir sinir ağı (RNN), bir katman içindeki ağırlıklı bağlantıları içeren bir sinir ağıları sınıfıdır (beslemeyi yalnızca sonraki katmanlara bağlayan geleneksel ileri beslemeli ağlarla karşılaştırıldığında). RNN'ler döngüler içerdiğinden, yeni girdiyi işlerken bilgi depolayabilirler. Bu bellek, onları önceki girdilerin dikkate alınması gereken (zaman serisi verileri gibi) işleme görevleri için ideal kılar. Bu nedenle mevcut derin öğrenme ağları RNN'lere dayanmaktadır. Bu öğretici, RNN'lerin arkasındaki fikirleri araştırır ve seri veri tahmini için sıfırdan bir tanesini uygular. Sinir ağıları, yüksek düzeyde bağlantılı işlem elemanları (nöronlar) ağına dayalı olarak bir girdiyi bir çıktıya eşleyen hesaplama yapılarıdır.

3.3 Ortak Aktivasyon Fonksiyonları

Sinir Ağları hakkındaki Öğren makalesinde tartışıldığı gibi, bir aktivasyon işlevi, bir nöronun etkinleştirilip etkinleştirilmeyeceğini belirler. Doğrusal olmayan fonksiyonlar tipik olarak belirli bir nöronun çıktısını 0 ile 1 veya -1 ile 1 arasında bir değere dönüştürür. En sık kullanılan fonksiyonlardan bazıları şu şekilde tanımlanır:



Şekil.14 Sigmoid: $g(x) = 1/(1 + e^{-x})$ formülüyle temsil edilir. [26]



Şekil.15 Tanh: $g(x) = (e^x - e^{-x})/(e^x + e^{-x})$ formülüyle temsil edilir. [26]

Basitçe söylemek gerekirse, ortak aktivasyon fonksiyonları ağıın verilerdeki karmaşık kalıpları öğrenmesine yardımcı olmak için bir yapay sinir ağına eklenen bir işlevdir. Beynimizde bulunan nöron temelli bir modelle karşılaştırıldığında, aktivasyon işlevi sonunda bir sonraki nörona neyin ateşleneceğine karar verir. Bir ortak aktivasyon fonksiyonunun ANN'de (Yapay Sinir Ağları) tam olarak budur. Bir önceki hücreden gelen çıkış sinyalinin alır ve onu bir sonraki hücreye girdi olarak alınabilecek bir forma dönüştürür.

Şekil.14'de görmüş olduğumuz Sigmoid fonksiyonu Yapay Sinir Ağları'nda en sık kullanılan aktivasyon fonksiyonudur. Sigmoid fonksiyonunun bu kadar sık kullanılmasının ana nedeni (0 ile 1) arasında olmasıdır. Bu nedenle özellikle çıktı olarak olasılığı tahmin etmemiz gereken modeller için kullanılır. Herhangi bir şeyin olasılığı sadece 0 ile 1 aralığında olduğu için sigmoid doğru seçimdir.

Şekil.15'de görmüş olduğumuz Tanh fonksiyonu Yapay Sinir Ağları'nda kullanılan bir başka alternatif aktivasyon fonksiyonudur. Hiperbolik tanjant aktivasyon fonksiyonu aynı zamanda basitçe Tanh (ayrıca "tanh" ve "TanH") fonksiyonu olarak da adlandırılır. Sigmoid aktivasyon fonksiyonuna çok benzer ve hatta aynı S şekline sahiptir. Fonksiyon, herhangi bir gerçek değeri girdi olarak alır ve -1 ile 1 aralığındaki değerleri çıkarır. Girdi ne kadar büyükse (daha pozitif), çıktı değeri 1.0'a o kadar yakın olur, oysa girdi ne kadar küçükse (daha negatif), o kadar yakın olur. çıktı -1.0 olacaktır.

3.4 Varyant Rnn Mimarileri

Çift yönlü tekrarlayan sinir ağları (BRNN): Bunlar, RNN'lerin değişken bir ağ mimarisidir. Tek yönlü RNN'ler yalnızca mevcut durum hakkında tahminlerde bulunmak için önceki girdilerden alınabilirken, çift yönlü RNN'ler doğruluğunu artırmak için gelecekteki verileri çeker. Bu makalenin başlarındaki "havanın altında hissetmek" örneğine geri dönersek, model dizideki son kelimenin "hava" olduğunu bilseydi, bu ifadedeki ikinci kelimenin "under" olduğunu daha iyi tahmin edebilir.

Uzun kısa süreli bellek (LSTM): Bu, Sepp Hochreiter ve Juergen Schmidhuber tarafından kaybolan gradyan sorununa bir çözüm olarak tanıtılan popüler bir RNN mimarisidir. Makalelerinde (PDF, 388 KB) (bağlantı IBM dışındadır), uzun vadeli bağımlılıklar sorununu çözmek için çalışırlar. Yani, mevcut tahmini etkileyen önceki durum yakın geçmişte değilse, RNN modeli mevcut durumu doğru bir şekilde tahmin edemeyebilir. Örnek olarak, aşağıdaki italik kelimeleri tahmin etmek istediğimizi varsayalım, "Alice fındıklara alerjisi var. Fıstık ezmesi yiyemez." Bir fındık alerjisi bağlamı, yenemeyen yiyeceklerin fındık içerdiğini tahmin etmemize yardımcı olabilir. Bununla birlikte, bu bağlam birkaç cümle önce olsaydı, RNN'nin bilgiyi bağlamasını zorlaştırır, hatta imkansız hale getirirdi. Bunu düzeltmek için, LSTM'ler, sinir ağının gizli katmanlarında üç kapıya sahip "hücrelere" sahiptir: bir giriş kapısı, bir çıkış kapısı ve bir unutma kapısı. Bu kapılar, ağdaki çıktıyı tahmin etmek için gerekli olan bilgi akışını kontrol eder. Örneğin, "she" gibi cinsiyet zamirleri önceki cümlelerde birden çok kez tekrarlandıysa, bunu hücre durumundan hariç tutabilirsiniz. Kapılı yinelenen birimler (GRU'lar): Bu RNN varyantı, RNN modellerinin kısa süreli bellek sorununu da çözmeye çalıştığı için LSTM'lere benzer. Bir "hücre durumu" düzenleme bilgisi kullanmak yerine, gizli durumları kullanır ve üç kapı yerine iki tane vardır: bir sıfırlama kapısı ve bir güncelleme kapısı. LSTM'lerdeki kapılara benzer şekilde, sıfırlama ve güncelleme kapıları, ne kadar ve hangi bilgilerin tutulacağını kontrol eder.

3.5 Tekrarlayan Sinir Ağları Ve Ibm Cloud

IBM, on yıllardır, IBM Watson'ın gelişimi ve evrimi ile vurgulanan yapay zeka teknolojilerinin ve sinir ağlarının geliştirilmesinde öncü olmuştur. Watson, artık yapay zekanın benimsenmesi ve uygulanmasına yönelik kanıtlanmış katmanlı bir yaklaşım kullanarak gelişmiş doğal dil işleme ve derin öğrenme tekniklerini sistemlerine uygulamak isteyen kuruluşlar için güvenilir bir çözümdür. IBM Watson Machine Learning gibi IBM ürünleri, yinelenen sinir ağlarında yaygın olarak kullanılan TensorFlow, Keras ve PyTorch gibi popüler Python kitaplıklarını da destekler. Kuruluşunuz, IBM Watson Studio ve Watson Machine Learning gibi araçları kullanarak, modellerinizi herhangi bir bulutta devreye alıp çalıştırırken açık kaynaklı yapay zeka projelerinizi sorunsuz bir şekilde üretime getirebilir.

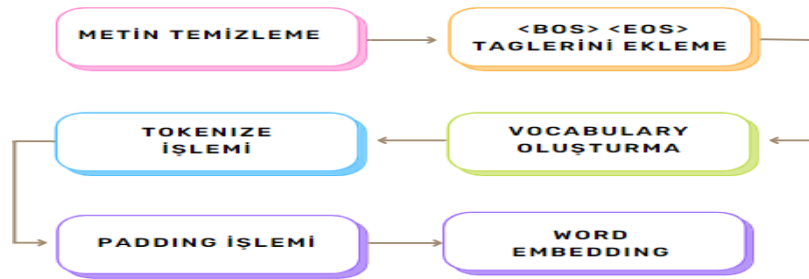
4. YÖNTEM

4.1 Kullanılan Materyaller ve Aşılan Zorluklar

Bu bölümde LSTM katmanı içeren seq2seq modeli uygulayarak sohbet bot uygulaması tasarlanmıştır. Uygulamayı geliştirme aşamasında ise yapay zeka uygulamaları için en uygun dillerden Python kullanılmıştır. Ayrıca, ihtiyaç duyulan kütüphaneler ve frameworkler kullanılarak özellikle, ücretsiz GPU desteği sağlaması sebebi ile, Google'ın hizmeti olan Colab kullanılmıştır. Uygulama geliştirilirken performans düşüklükleri, tutarsız veya anlamsız çıktı gibi sorunlarla karşılaşmıştır. Karşılaşılan bu sorunlar sonucunda, konfigürasyon işlemleri tekrar düzenlenip gerekli adımlar atılarak yeniden eğitim sağlanmış ve başarı oranı tatmin edici seviyelere getirilmiştir.

4.2 Hazırlık Aşaması

Çalışmaya başlamadan önce sohbet botumuza soracağımız soruları ve bu sorulara karşılık beklediğimiz cevapları düşündük. Bunu düşünme sebebimiz konumuz hakkında çok fazla veri seti mevcut olması ve botumuzdan almayı beklediğimiz cevaplar hem veri setine hem soracağımız sorulara bağlıydı. Veri setimizi seçtikten sonra izlediğimiz ön işleme adımları aşağıdaki gibidir.



Şekil.16 Hazırlık Aşaması

4.3 Metin Temizleme

Bu aşama için kullanılan fonksiyonlar genel olarak birbirleriyle benzerdir. Bizim kullandığımız fonksiyonun kod parçası şeklindeki gibidir.

```
def temizle_txt(txt):  
    txt = txt.lower()  
    txt = re.sub(r"i'm", "i am", txt)  
    txt = re.sub(r"he's", "he is", txt)  
    txt = re.sub(r"she's", "she is", txt)  
    txt = re.sub(r"it's", "it is", txt)  
    txt = re.sub(r"that's", "that is", txt)  
    txt = re.sub(r"what's", "that is", txt)  
    txt = re.sub(r"where's", "where is", txt)  
    txt = re.sub(r"how's", "how is", txt)  
    txt = re.sub(r"\ll", " will", txt)  
    txt = re.sub(r"\ve", " have", txt)  
    txt = re.sub(r"\re", " are", txt)  
    txt = re.sub(r"\d", " would", txt)  
    txt = re.sub(r"\re", " are", txt)  
    txt = re.sub(r"won't", "will not", txt)  
    txt = re.sub(r"can't", "cannot", txt)  
    txt = re.sub(r"n't", " not", txt)  
    txt = re.sub(r"n'", "ng", txt)  
    txt = re.sub(r"'bout", "about", txt)  
    txt = re.sub(r"'til", "until", txt)  
    txt = re.sub(r"[-()\"#/@;:<>{}`+=~|.!?,]", "", txt)  
    return txt
```

Şekil.17 Metin Temizleme Kodu

```
["Drink up, Charley. We're ahead of you.",  
 'Did you change your hair?',  
 'I believe I have found a faster way.']
```

Şekil.18 Fonksiyon çalışmadan önceki data

```
['drink up charley we are ahead of you',  
 'did you change your hair',  
 'i believe i have found a faster way']
```

Şekil.19 Fonksiyon çalıştıktan sonra data

4.4 Etiket Ekleme

Ön işleme aşamalarımızdan birisi de decoder, encoder inputlarına etiketler eklemektir. <PAD> etiketini eğitim aşamasında ağa göndereceğimiz batchlerin eşit boyutta olmasını sağlamaktadır. <EOS> etiketimiz cümlelerin bitiş noktasını belirtir. <SOS> etiketimiz cümlelerin başlangıcını işaretler. Şekil.20'deki döngüde temizle_soru datamızı önce satır satır sonrasında kelime kelime dolaşarak vocab datasında bulunmayan kelimeleri tespit edip yerine <OUT> etiketini ekledik.

```
151 for cumle in temizle_soru:
152     lst = []
153     for word in cumle.split():
154         if word not in vocab:
155             lst.append(vocab['<OUT>'])
156         else:
157             lst.append(vocab[word])
158
```

Şekil.20 Etiket ekleme

```
for i in range(len(temizle_cevap)):
    temizle_cevap[i] = '<SOS> ' + temizle_cevap[i] + ' <EOS>'
```

Şekil.21 Cevap datamızın başlangıcını ve sonunu işaretleyen kod parçası

4.5 Vocabulary Oluşturma

Bu aşamada datamızın sözlüğünü oluşturuyoruz diyebiliriz. Sözlüğümüzde kelimelere atanmış key değerleri olacaktır. İlk önemli adımımız şekildeki fonksiyon ile her kelimenin kaç kez kullanıldığı belirlemektir. Sonrasında belirlediğimiz thresh değerinden fazla kez tekrarlanan kelimeleri sözlüğümüze ekleyerek sözlüğümüzü oluşturmuş olduk. Thresh değerimiz uygulamamızın tasarlama süresince değişiklikler görmüştür.

```
word2count = {}

for line in clean_ques:
    for word in line.split():
        if word not in word2count:
            word2count[word] = 1
        else:
            word2count[word] += 1
for line in clean_ans:
    for word in line.split():
        if word not in word2count:
            word2count[word] = 1
        else:
            word2count[word] += 1
```

Şekil.22 Vocabulary oluşturma kod parçası

4.6 Piksel Ekleme

Bu aşamada decoder ve encoder inputlarının eşit uzunlukta olmasını sağlamayı amaçlıyoruz. Birçok algoritma input datasının son halinde eşit uzunluk beklemektedir. Belirlediğimiz uzunluk 13'tür.

4.7 Seq2seq Modeli Oluşturma

4.7.1 tf.keras.Model

Uygulamamızda kullandığımız katmanları gruplayan, eğitimimiz için kullanışlı özellikler sağlayan bir modüldür. Sequential modülü yerine Model kullanmamızın sebebi eğitim ve tahmin aşamalarında daha fazla kontrol sahibi olmamız.

4.7.2 Katmanlar

4.7.2.1 Input

Encoderımıza yollayacağımız datayı girdiler haline getirdik. İlk parametremiz bu girdilerin boyutudur. Kodumuzun önceki kısımlarında pad_sequence oluştururken de 13 değerini belirlemiştik burada da aynı değeri kullanıyoruz.

```
5 enc_inp = Input(shape=(13, ))  
6 dec_inp = Input(shape=(13, ))
```

Şekil.23 Girdi Oluşturma kod parçası

```
179 encoder_inp = pad_sequences(encoder_inp, 13, padding='post', truncating='post')  
180 decoder_inp = pad_sequences(decoder_inp, 13, padding='post', truncating='post')
```

Şekil.24 pad_sequences oluşturma kod parçası

4.7.2.2 Kelime Gömme İşlemi

Bu katmanımız ile daha önce oluşturduğumuzdan bahsettiğimiz (bkz. 3.2.3) vocab datamızı sıkıştırıyoruz. Birinci parametremiz olan input_dim değerini vocab datamızın büyüklüğü olan VOCAB_SIZE olarak belirliyoruz. VOCAB_SIZE değerini 1 arttırma sebebimiz ekstra bir etiketi eklememizdir. Output_dim parametresini 50 olarak belirledik. Bu parametre, sıkıştırma sonucu oluşmasını istediğimiz boyutu temsil etmektedir. Input_length parametresinin 13 olduğunu daha önce pad_sequence bölümünde belirlemiştik. Son olarak trainable parametresini de true olarak belirledik.

4.7.2.3 LSTM

Tf.keras.layer'dan import ettiğimiz LSTM modülünde encoder ve decoder için 3'er parametre kullandık. Bu parametrelerden ilki LSTM yapımızın vermesini istediğimiz boyut sayısını belirtir. Biz 300 olarak belirledik. Her adımdan hidden state ve sequence çıktısı almak için şekildeki gibi return_sequences ve return_state parametrelerini true olarak belirledik.

```
17 enc_lstm = LSTM(400, return_sequences=True, return_state=True)
24 dec_lstm = LSTM(400, return_sequences=True, return_state=True)
```

Şekil.25 dec_lstm ve enc_lstm oluşturma kod parçası

4.7.2.4 Dense

Dense katmanımız, katmanlar arası geçişler sağlamaktadır. Şekil.26’de görüldüğü üzere iki adet parametre ile kullanım sağladık. Activation = “softmax” kullanımının sebebi modelimizin olasılık odaklı olmasıdır. Yine Şekil.26’de görüldüğü gibi decoder outputunu dense layerda parametre olarak kullandık.

```
28 dense = Dense(VOCAB_SIZE, activation='softmax')
29
30 dense_op = dense(dec_op)
31
```

Şekil.26 Dense katmanı

4.8 Model

4.8.1 Decoder Model

LSTM için 300 hücre kullandık. Bu yüzden 300 adet h ve c çıktımızın oluşması beklenmektedir. Şekil.27 kod satırı 10 ve 11’de görüldüğü üzere decoder_state_input_h ve decoder_state_input_c için Input katmanı kullandık. Decoderimizin inputu için gerekli state’ler oluştuktan sonra decoder_states_inputs değişkenimiz ile decoder inputumuzun son halini oluşturmuş olduk. Oluşturduğumuz inputları kullanarak elde ettiğimiz outputlar Şekil.27 satır 13’de görüldüğü üzere state_h, state_c ve decoder_outputs değişkenlerine atanmıştır. Daha sonra decoder outputundan elde ettiğimiz stateler ile Şekil.27 satır 13’deki görüldüğü gibi state listesine ekledik. Son olarak satır 15’de görüldüğü gibi önceki satırlarda oluşturduğumuz değişkenleri kullanarak decoder modelimizi oluşturduk.

```

decoder_state_input_h = Input(shape=(400,))
decoder_state_input_c = Input(shape=(400,))
decoder_states_inputs = [decoder_state_input_h, decoder_state_input_c]
decoder_outputs, state_h, state_c = dec_lstm(dec_embed, initial_state=decoder_states_inputs)
decoder_states = [state_h, state_c]
dec_model = Model([dec_inp, decoder_states_inputs],
                  [decoder_outputs] + decoder_states)

```

Şekil.27 Decoder Model

4.8.2 Encoder Model

Kullandığımız LSTM decoderde olduğu gibi 300 hücrelidir. Önce Şekil.28’da görüldüğü gibi Input katmanımız ile input verimizi oluşturduk. Bu input verimizi kullanarak LSTM’den çıktı elde edebilmek için embedding katmanı ile verimizi işlenebilir hale getirdik. Hazır LSTM fonksiyonu ile oluşturduğumuz LSTM’e embedding katmanında hazırladığımız input verisini yollayarak encoder output, h state ve c state elde ettik. Bu elde ettiğimiz stateler bize decoderda lazım olacaktır. Bu yüzden state list oluşturup bu stateleri listemize koyduk.

```

enc_embed = embed(enc_inp)
enc_lstm = LSTM(400, return_sequences=True, return_state=True, dropout=0.05)
enc_op, h, c = enc_lstm(enc_embed)
enc_states = [h, c]

```

Şekil.28 Encoder Model

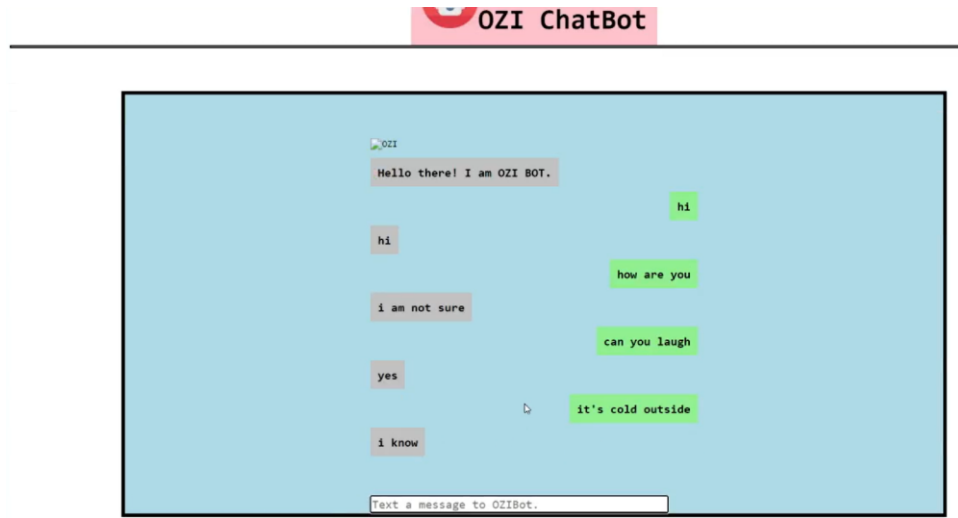
4.9 Model Test Aşaması

Bu aşamada botumuza bir cümle yazıyoruz. Yazdığımız bu cümle ilk olarak clean_text fonksiyonuyla ön işlemeye giriş yapıyor. Bu aşamadan sonra girdimiz dizi haline getiriliyor. Bir sonraki aşama için key değerlerini tutmaya yarayacak txt[] dizisi oluşturuluyor. Dizi haline getirdiğimiz cümle girdimiz döngü içinde kelimelere ayrılarak vocab datamızda var olup olmadığı kontrol ediliyor. Var ise hangi key değerine sahip olduğu ayrı bir lst[] listesinde tutuluyor.

Bu list döngünün içindeki geçici bir listtir. Döngünün farklı bir katmanında lst'den txt'ye veri aktarımı söz konusudur. Döngünün ardından elde ettiğimiz txt listini pad_sequence'te kullanıyoruz. Elde ettiğimiz yeni verimizi txt değişkenine atayıp “enc_model” modelimizde parametre olarak kullanıyoruz. Daha sonra 1 adet satır ve sütundan, 0 değerlerinden oluşan numpy dizisi tanımlıyoruz. Bu dizinin 0. indislerine <SOS> etiketini koyuyoruz. Daha sonra bir döngü içerisinde decoder modelimize encoder predictlerimizi ve boş numpy dizimizi parametre olarak atayarak çalıştırıyoruz. Oluşan decoder outputumuzu modelimizde tanımlı dense katmanına yolluyoruz. Dense katmanındaki softmax özelliği sayesinde sonuç olarak olasılıklar elde ediyoruz. Kullandığımız arg_max hazır fonksiyonu ile en yüksek olasılıklı datanın indeksini elde ediyoruz. Elde ettiğimiz indis ile sözlükte tahmin ettiğimiz kelimeyi sample_word değişkenine atıyoruz. Daha sonra if else yapısı ile <EOS> etiketi kontrolü yapıyoruz. Etiket varsa veya tahmin ettiğimiz sonuç cümlesinin kelime sayısı 13'ten büyük olursa stop_condition ile döngümüzden çıkıyoruz. Etiket yoksa döngümüz devam ediyor.

5. SONUÇ

Sohbet botu çalışmasında, LSTM ve sıradan seq2seq modeli ile sohbet bot uygulaması geliştirilmiştir. Çalışmada, veri kümesi olarak Movie Dialog Corpus veri kümesi kullanılmıştır. Çalışma ile, Movie Dialog Corpus veri kümesinden alınan verilerle kullanıcının sorduğu sorulara uygun cevaplar veren bir sohbet botu tasarımı yapılmıştır. Çalışma, categorical_crossentropy kayıp fonksiyonu ile 40 adımlık dönem için eğitilmiştir. Kayıp oranı son adımdan sonra 0.5621 değerlerine düşürülmüştür. Elde edilen doğruluk oranı ise yüzde 85 dir. Elde edilen verilere göre eğitim sayısı arttıkça doğruluk oranı da artmaktadır. TPU destekli bilgisayarlarda daha kapsamlı veri kümesi kullanıldığında kapsamı geniş cevaplar üretilebileceği ve yüksek doğrulukta sonuçlar elde edilebileceği düşünülmektedir.



6. KAYNAKLAR

- [1] Textminingonline, "Getting Started with Word2Vec and GloVe", Available: <https://textminingonline.com/getting-started-withword2vec-and-glove>, [Accessed: Sept. 12, 2019].
- [2] An intelligent Chatbot using deep learning with Bidirectional RNN and attention
modelManyuDhyani-RajivKumar-<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7283081/>
- [3] Designing and Developing a Chatbot Using Machine Learning
<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8746972>
- [4] AI-based chatbots in customer service and their effects on user compliance-Martin AdamMichael Wessel -Alexander Benlian:
- [5] Adamopoulou E., Moussiades L. (2020) An Overview of Chatbot Technology. In: Maglogiannis I., Iliadis L., Pimenidis E. (eds) Artificial Intelligence Applications and Innovations. AIAI 2020. IFIP Advances in Information and Communication Technology, vol 584. Springer, Cham. https://doi.org/10.1007/978-3-030-49186-4_31
- [6] A. M. Rahman, A. A. Mamun and A. Islam, "Programming challenges of chatbot: Current and future prospective," 2017 IEEE Region 10 Humanitarian Technology Conference (R10- HTC), 2017, pp. 75-78, doi: 10.1109/R10- HTC.2017.8288910.
- [7] N. Ali, M. Hindi and R. V. Yampolskiy, "Evaluation of authorship attribution software on a Chat bot corpus," 2011 XXIII International Symposium on Information, Communication and Automation Technologies, 2011, pp. 1-6, doi: 10.1109/ICAT.2011.6102123.
- [8] D. Carlander-Reuterfelt, Á. Carrera, C. A. Iglesias, Ó. Araque, J. F. Sánchez Rada and S. Muñoz, "JAICOB: A Data Science Chatbot," in IEEE Access, vol. 8, pp. 180672-180680, 2020, doi: 10.1109/ACCESS.2020.3024795.
- [9] Eleni Adamopoulou, Lefteris Moussiades, Chatbots: History, technology, and applications, Machine Learning with Applications, Volume 2, 2020, 100006, ISSN 2666- 8270, <https://doi.org/10.1016/j.mlwa.2020.100006>
- [10] COUNSELLOR CHATBOT
-Rishabh Chawlahttps://www.academia.edu/36353256/COUNSELLOR_CHATBOT
- [11] Babar, Zia, Alexei Lapouchnian, and Eric Yu. "Chatbot Design-Reasoning about design options using i* andprocess architecture.
- [12] ChatBot <https://tr.wikipedia.org/wiki/Chatbot>

- [13] Connecting a Chatbot with Channels-Stephan Bisserhttps://www.researchgate.net/publication/349451342_Connecting_a_Chatbot_with_Channels
- [14] Artificial Intelligent ChatBot for Food Related Question- Alessandra Amato-Giovanni Cozzolino-Antonino Ferraro
<https://www.springerprofessional.de/en/artificial-intelligentchatbot-for-food-related-question/18464288>
- [15] “Love Is as Complex as Math”: Metaphor Generation System for Social ChatbotDanning Zheng-Ruihua Song-Tianran Hu-Hao Fu-Jin Zhou-<https://arxiv.org/abs/2001.00733>
- [16] Kendi Kendine Denetimli Yarı Denetimli Öğrenme-Xiaohua Zhai-Avital OliverAlexander Kolesnikov-Lucas Beyer-Google Research Brain Team
- [17] Embedding a chatbot in a news article: design and implementation
https://www.academia.edu/64107792/An_Intelligent_Database_System_Using_Natural_Language_Processing
- [18] Andreas Veglis and Theodora A. Maniou. 2019. Embedding a chatbot in a news article: design and implementation. In *Proceedings of the 23rd Pan-Hellenic Conference on Informatics (PCI '19)*. Association for Computing Machinery, New York, NY, USA, 169– 172. DOI:<https://doi.org/10.1145/3368640.3368664>
- [19] Zhifa Chen, Yichen Lu, Mika P. Nieminen, and Andrés Lucero. 2020. Creating a Chatbot for and with Migrants: Chatbot Personality Drives Co-Design Activities. *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. Association for Computing Machinery, New York, NY, USA, 219–230. DOI:<https://doi.org/10.1145/3357236.3395495>
- [20] Farah Shaik, Arooba Khalid Bashir, and Heba Mahmoud Ismail. 2021. Covisstance Chatbot. In *The 7th Annual International Conference on Arab Women in Computing in Conjunction with the 2nd Forum of Women in Research (ArabWIC 2021)*. Association for Computing Machinery, New York, NY, USA, Article 16, 1–4. DOI:<https://doi.org/10.1145/3485557.3485574>
- [21] Heba Ismail, Nada Hussein, Rawan Elabyad, and Salma Said. 2021. A Serverless Academic Adviser Chatbot. In *The 7th Annual International Conference on Arab Women in Computing in Conjunction with the 2nd Forum of Women in Research (ArabWIC 2021)*. Association for Computing Machinery, New York, NY, USA, Article 27, 1–5. DOI:<https://doi.org/10.1145/3485557.3485587>
- [22] Qiyu Zhi and Ronald Metoyer. 2020. GameBot: A Visualization-augmented Chatbot for Sports Game. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–7. DOI:<https://doi.org/10.1145/3334480.3382794>

- [23] Saurabh Srivastava, Sumit Kalra, and T.V. Prabhakar. 2020. Contextual Reactive Pattern on Chatbot building Platforms. In *Proceedings of the European Conference on Pattern Languages of Programs 2020 (EuroPLoP '20)*. Association for Computing Machinery, New York, NY, USA, Article 6, 1–8.
DOI:<https://doi.org/10.1145/3424771.3424815>
- [24] Facebook launches Messenger platform with chatbots<https://techcrunch.com/2016/04/12/agents-on-messenger/>
- [25] NLP vs. NLU: from Understanding a Language to Its Processing
<https://www.kdnuggets.com/2019/07/nlp-vs-nlu-understanding-language-processing.html>
- [26] What are recurrent neural networks?-<https://www.ibm.com/cloud/learn/recurrent-neural-networks>
- [27] RNN NEDİR NASILÇALIŞIR?-<https://medium.com/deep-learning-turkiye/rnn-nedir-nasil-çalışır-9e5d572689e1>
- [28] Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network
<https://www.sciencedirect.com/science/article/abs/pii/S0167278919305974>
- [29] Improving RNN Transducer Modeling for End-to-End Speech Recognition <https://ieeexplore.ieee.org/abstract/document/9003906>