

YİNELENEN SİNİR AĞLARI



01



YSA (RNN) Nedir

Yapay Sinir Ağları konu başlığından bahsedeceğiz.

02



LSTM

Long Short Term Memory konu başlığından bahsedeceğiz.

03



SRN

Basit Tekrarlayan Ağ konu başlığından bahsedeceğiz.

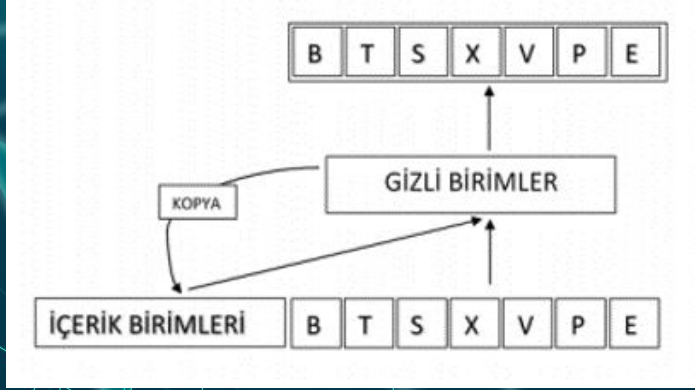
YSA'lar derin öğrenme ve insan beynindeki nöronların aktivitesini simüleden modellerin geliştirilmesinde kullanılır. Bağlamın bir sonucu tahmin etmek için kritik olduğu ve diğer yapay sinir ağlarının farklı olduğu kullanım durumlarında özellikle güçlüdürler, çünkü çıktıyı bildiren bir veri dizisini işlemek için geri besleme döngülerini kullanırlar.

YSA (RNN) Nedir?

RNN = Recurrent Neural Network = Yinelenen Sinir Ağları

Yinelenen Sinir Ağı, konuşma tanıma ve doğal dil işlemede (NLP) yaygın olarak kullanılan bir yapay sinir ağı türüdür . YSA'lar bir verinin sıralı özelliklerini tanımak ve bir sonraki olası senaryoyu tahmin etmek için kalıpları kullanmak üzere tasarlanmıştır.

NLP = Neuro-Linguistic Programming = Duyu-Dil Programlama

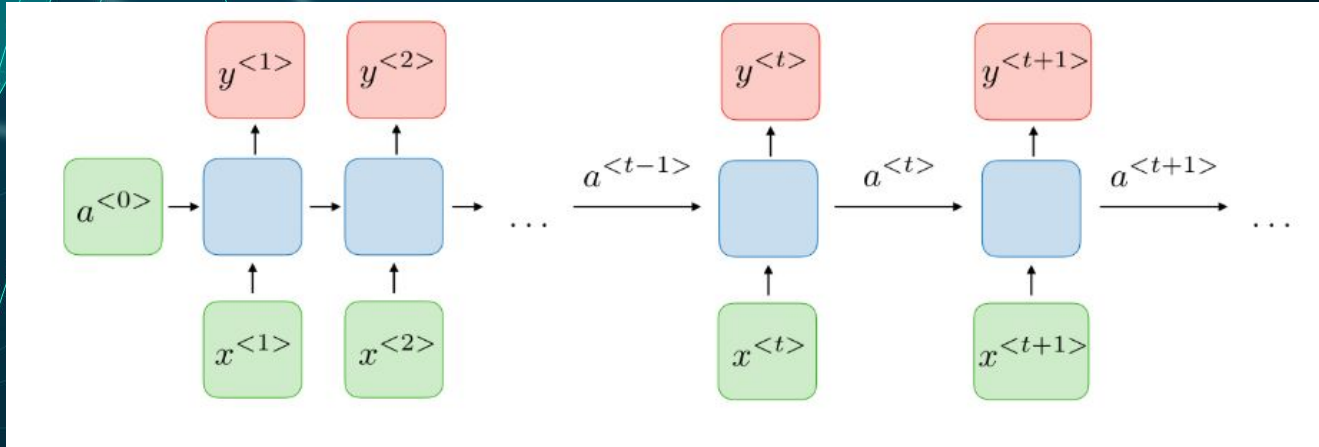


Basit Tekrarlayan Ağ (SRN), Jeff Elman tarafından tasarlanmıştır. Elman'ın cümle yapısı simülasyonunda kullanılan her bir kelime için gizli kalıpların üzerinde ortalama örüntü kümeleme sonucunda isim ve fiil kategorileri temiz şekilde ayrılmıştır.

Bunun yanında isimler arasında canlı-cansız ayrımları, hatta insan-hayvan, hayvanlar arasında avcı-yırtıcı gibi kümeler de ayrılmıştır

YSA İnşası

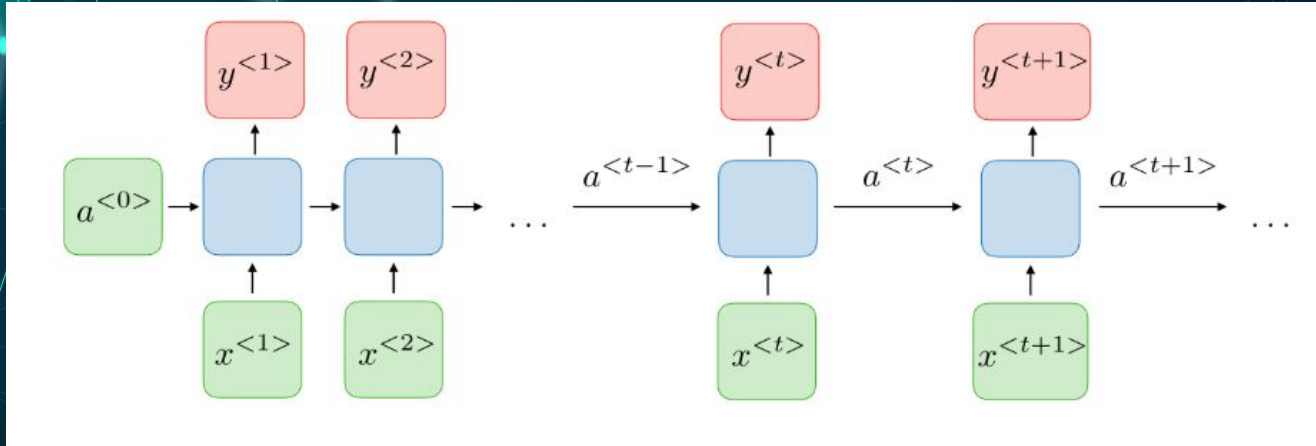
YSA'lar, gizli durumlara sahipken önceki çıktıların girdi olarak kullanılmasına izin veren bir sinir ağı sınıfıdır. Tipik olarak aşağıdaki gibidirler:



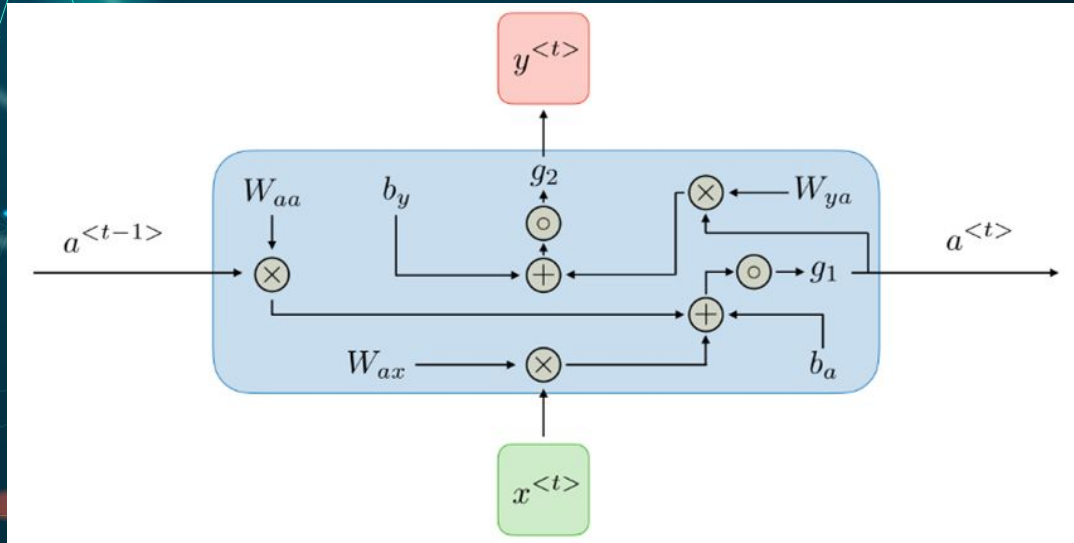
Her bir t zamanında, 'a' aktivasyonu ve 'y' çıktısı aşağıdaki gibi ifade edilir:

$$a^{<t>} = g_1(W_{aa} a^{<t-1>} + W_{ax} x^{<t>} + b_a)$$

$$y^{<t>} = g_2(W_{ya} a^{<t>} + b_y)$$

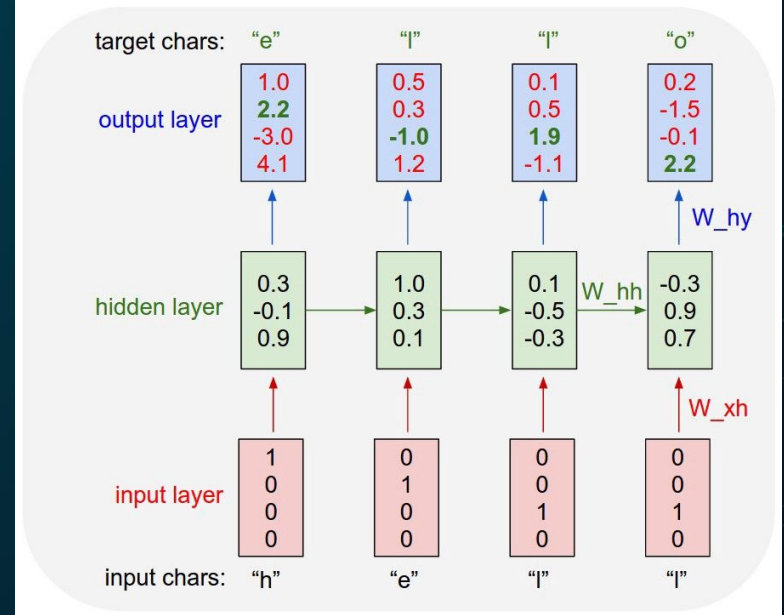


$W_{aa}, W_{ax}, W_{ya}, b_a$ ve b_y geici olarak paylaşılan kat sayılarıdır ve g_1 ve g_2 aktivasyon fonksiyonlarıdır.



YSA (RNN) Örnek

Yinelenen Sinir Ağları kullanılarak önce kullanılan kelimelerin gidişatına göre gerekli olan kelimeler tahmin edilebilmektedir. Dil modelleme yöntemi YSA ile gerçekleştirilebilmektedir. Eğitim verileri açık kod kullanan kütüphaneler yardımı ile veya manuel bir şekilde sinir ağına aktarılabilir. Ağ üzerinde yapılan bu eğitim çalışmaları sonrası yeni gelen kelimenin tahmin edilebilme olasılığı da artış göstermektedir.



Avantajları

- Herhangi bir uzunluktaki girdilerin işlenmesi imkanı
- Girdi büyüklüğüyle artmayan model boyutu
- Geçmiş bilgileri dikkate alarak hesaplama
- Zaman içinde paylaşılan ağırlıklar

Dezavantajları

- Yavaş hesaplama
- Uzun zaman önceki bilgiye erişme zorluğu
- Mevcut durum için gelecekteki herhangi bir girdinin tahmin edilememesi

Kayıp Fonksiyonu

Tekrarlayan bir sinir ağı olması durumunda, tüm zaman dilimlerindeki LL kayıp fonksiyonu, her zaman dilimindeki kaybı temel alınarak yandaki gibi tanımlanır.



$$\mathcal{L}(\hat{y}, y) = \sum_{t=1}^{T_y} \mathcal{L}(\hat{y}^{<t>}, y^{<t>})$$

Zamanda Geri Yayılım

Geriye yayılım zamanın her noktasında yapılır. T zaman diliminde, ağırlık matrisi W'ye göre L kaybının türevi aşağıdaki gibi ifade edilir.



$$\frac{\partial \mathcal{L}^{(T)}}{\partial W} = \sum_{t=1}^T \frac{\partial \mathcal{L}^{(T)}}{\partial W} \Big|_{(t)}$$

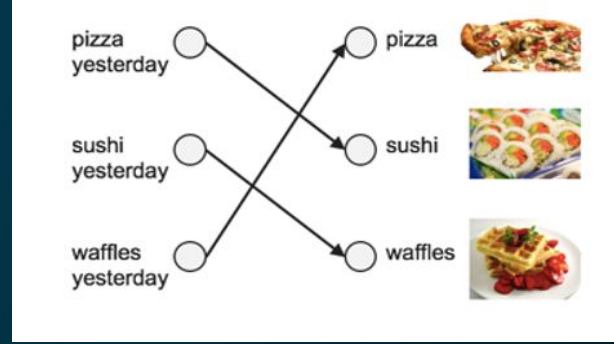
ÖRNEK

Bir aşçının 3 farklı yemeği sırasıyla yaptığını düşünelim.
1. gün pizza yaptıysa, 2. gün sushi ve 3. gün waffle yapıyor olsun. Bir sonraki gün ne yapacağını tahmin etmemiz gerekiyorsa öncelikle nasıl bir problemle karşı karşıya olduğumuz anlayarak, bu probleme uygun bir metot kullanmamız gerekir.

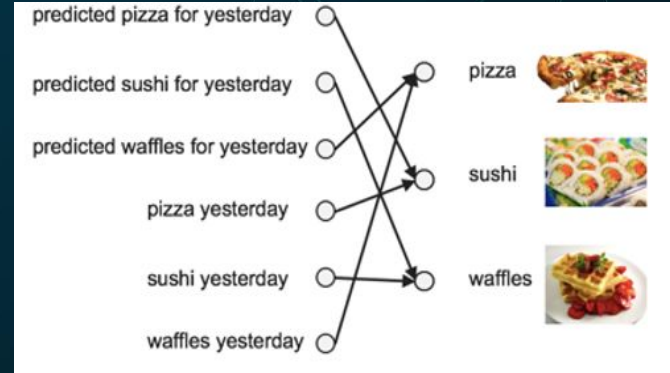
Burada önceki gün yapılan yemeğin sonraki günde etkisi olacağı için uygulanması gereken metot **YSA** metodudur.

Bu sayede elde var olan bilgilerle yeni bilgi olmasa dahi birkaç hafta sonraki günde yapılacak yemeği dahi **tahmin edilebilir** hale getirilebilir.

Aşçı yemek döngüsü



RNN Yapısıyla



YSA'lar ile Başka Ne Yapabilir?

YSA modelleri çoğunlukla kullanılan alanlar;

- **Dil Modelleme ve Metin Oluşturma**
- **Konuşma Tanıma**
- **Makine çevirisi**

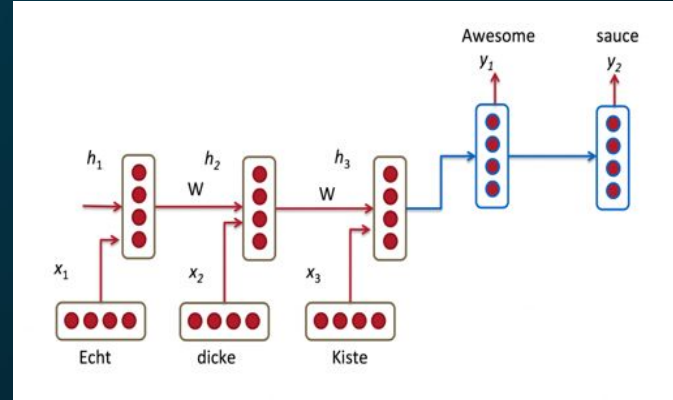
Bunlardan kısaca bahsedelim.

Konuřma Tanıma

Ses dalgasından gelen akustik sinyallerin bir giriş dizisi verildiğinde, fonetik bölümlerin bir dizisini olasılıkları ile birlikte tahmin edebiliriz. Uygulamayı tamamladığımız da alacağımız cevaplar basit tanımlamalar şeklinde olacaktır.

Makine Çevirisi

Makine Çevirisi, kaynak dilimizde bir dizi kelime girdiğimiz için dil modellemesine benzer. Hedef dilimizde bir kelime dizisi çıkarmak istiyoruz. Temel bir fark, sadece tüm **giriş** dizisini yakalamamız gerektiğidir, çünkü çevrilmiş cümlenin ilk sözcüğü, tam giriş dizisinden yakalanabilir.



YSA'ları Eđitmek

YSA eğitimi geleneksel bir sinir ağı eğitimine benzer. Ayrıca, **geri yayılım algoritmasını** kullanılıyor. Parametreler ağıdaki tüm zaman aşamaları tarafından paylaşıldığı için, her çıkıştaki gradyan sadece geçerli zaman adımının hesaplarına değil, aynı zamanda önceki zaman adımlarına da bağlıdır. Buna **Zamana Göre Geriye Yayılım** denir.

Zamana Göre Geri Yayılım (BPTT)

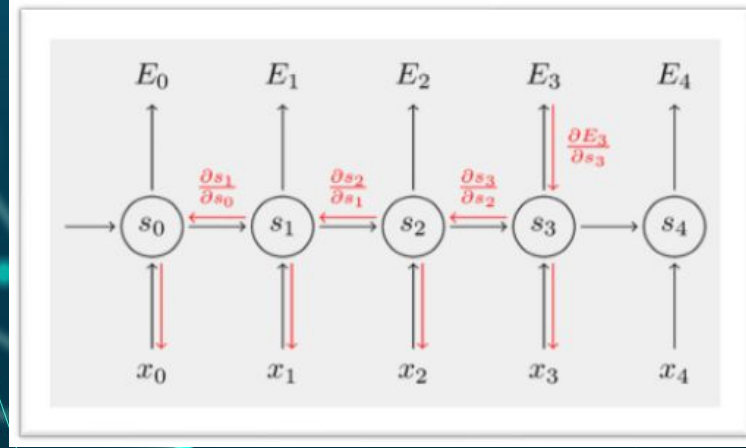
Recurrent ağların amacı sıralı inputları doğru bir şekilde sınıflandırmak denilebilir. Bu işlemleri yapabilmek için hatanın geri yayılımını ve gradient düşüşünü kullanırız. Backprop, feedforward ağlarda sonda çıkıtki hatayı geriye hatanın türevini ağırlıklara dağıtılarak yapılır. Bu türev kullanılarak öğrenme katsayısı, gradyan azalmasını düzenlenerek hatayı düşürecek şekilde ağırlıklar düzenlenir.

YSA için kullanılan yöntem ise BPTT diye bilinen zamana bağlı sıralı bir dizi hesaplamasının tümü için backprop uygulamasıdır. Yapay ağlar bir dizi fonksiyonu içiçe $f(h(g(x)))$ şeklinde kullanır. Buraya zamana bağlı değişken eklendiğinde türev işlemi zincir kuralı ile çözümlenebilir.

BPTT = Backpropagation Trough Time = Zamana Göre Geri Yayılım

Backprop = Geri Yayılım

Çalışma Mantığı



Yukardaki YSA'nın çalışma mantığı anlatılmıştır. Üstteki figürde ise sıralı 5 girdili bir recurrent ağ yapısı gösterilmektedir. 'E' burada oluşan hatayı ifade etmektedir. Örneğin, 'E₃' için backpropagation yaparken yaptığımız işlemde w ağırlığına göre türevi kullanılmaktadır. Bu türevi çözebilmek için zincir kuralı ile birkaç türevin çarpımını kullanılır.

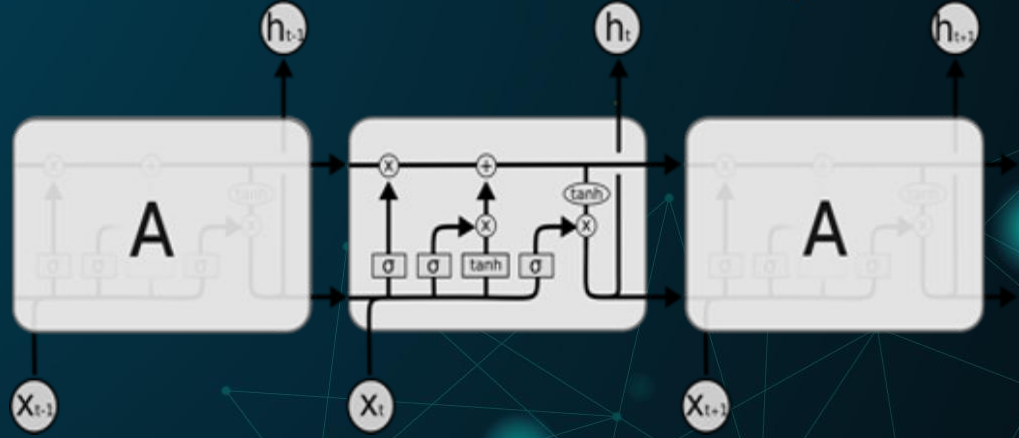
Long Short Term Memory (LSTM)

Sepp Hochreiter ve **Juergen Schmidhuber** 1997 yılında LSTM'i geliştirdiler. Daha sonra birçok kişinin katkısıyla düzenlenen ve popülerleşen LSTM şu anda geniş bir kullanım alanına sahiptir. LSTM backprop'ta farklı zaman ve katmanlardan gelen hata değerini korumaya yarıyor. Daha sabit bir hata değeri sağlayarak recurrent ağların öğrenme adımlarının devam edebilmesini sağlamaktadır.

LSTM = Long Short Term Memory = Uzun Kısa Süreli Bellek

LSTM yapısında tekrar eden modülün farkı tek bir neural network katmanı yerine, özel bir şekilde bağlı 4 katman bulunmasıdır. Bu katmanlara aynı zamanda kapı denmektedir. Bu bilgiler depolanabilir, hücreye yazılabilir, okunabilir.

Hücre neyi depolayacağını, ne zaman okumasına, yazmasına veya silmesine izin vereceğini kapılar sayesinde karar verir. Bu kapılarda bir ağ yapısı ve aktivasyon fonksiyonu bulunmaktadır. Aynı nöronlarda olduğu gibi gelen bilgiyi ağırlığına göre geçirir veya durdurur.



Soru 1



**Basit Tekrarlayan Ağ (SRN),
kim tarafından
tasarlanmıştır?**

- A) Sepp Hochreiter
- B) Juergen Schmidhuber
- C) Jeff Elman**
- D) David Rumelhart

Soru 2



LSTM yapısında tekrar eden modül özel bir şekilde bağlı 4 katman bulunur. Hangileri bu modülde bulunan bilgilerin özelliğidir?

- I- Depolanabilirlik
- II- Hücreye Yazılabilme
- III- Okunabilirlik
- IV- Filtreleme

- A) Yalnız I
- B) I - II - III**
- C) I - II - III - IV.
- D) Yalnız IV

Soru 3

Hangisi RNN kullanılan alanlara girmez ?

- A) Dil Modelleme ve Metin Oluřturma
- B) Konuřma Tanıma
- C) Makine Çevirisi
- D) Veri İletimi**

Soru 4

Kaynak dilimizin İngilizce olduđu bir modelleme işleminde giriş dizisinin “Awes” olduđu durumda beklenen tahmini çıktı ne olacaktır?

- A) Awful
- B) Allies
- C) Awaydays
- D) Awesome**

KAYNAKLAR

Araştırma

- <http://colah.github.io>
- <https://deeplearning4j.org>
- <https://ayearofai.com>
- <http://www.wildml.com>
- <http://curiously.com>
- <https://medium.com/@mliuzzolino/hello-rnn-55a9237b7112>
- <http://www.wildml.com>

Görsel

- <https://slidesgo.com/>



TEŞEKKÜRLER

Sormak istediğiniz soru var mı?

G171210014 Aslıhan ÇETİNER

G171210558 Muhammet ÖMER

G181210350 Elyase Tunahan SAĞLAM

G171210032 Esma Ülhüsna SİĞİRTMAÇ