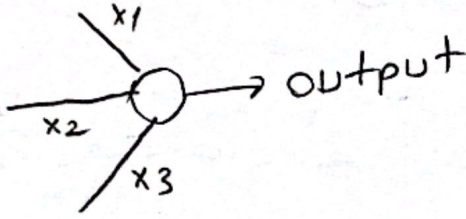


Neural Networks

Kaan İNCE
13.01.2023

Classification, regressiondan datayi sonuç verir. (Görüntü için)

Perceptron \rightarrow NN'in en temel yapı taşıdır. İnsan beyninden esinlenilmiştir.



$$\text{output} = \begin{cases} 0 & \text{if } \sum_j w_j x_j \leq \text{threshold} \\ 1 & \text{if } \sum_j w_j x_j > \text{threshold} \end{cases} \rightarrow \text{Önemli bir algoritmadır.}$$

Thresholder, bias'e evrilir! Bias bir ön yetkidir. Sistemi bias ile data eğitimi ya da eğimsiz yapıyoruz. İki önemli şey ile ilgileniyoruz. Doğru ağırlık ve bias değerlerini bulmak. Nasıl buluyoruz?

- Deneme / Yanılma vardır. Train olması \rightarrow Geçmiş dataya ihtiyaç gerektir. Bunu data sağlar.

Okula gidip gitmemek?

Akşam dışarı çıkmak çıkmamak?

AND veya OR kapısını tek nöronla yapabiliriz.

NAND kapısına ikiden fazla nöron gerektir.

Master algoritma \rightarrow her şey öğrenir.

Sigmoid function \rightarrow aktivasyon fonksiyonudur. Kaan İNCE
13.01.2023

Ağın çıkışına uygulanır. Daha sonra bu,

nihai çıkış ya da bir başka katmanın girişi olacaktır.

Perceptron, tek yapılı sinir ağ modelidir. (6)

<u>Giriş</u>	<u>Ara</u>	<u>Çıkış</u>
Featurelar	hidden layerlar	Output

NN geliştiren, matematiksel mantığın aktığı katmandır.

Datanın miktarı önemlidir. Az datada weightler güncellenemez. Görüntü datası için geçitli olması iyidir. Geçitli ve birbirinden farklı

garbage-in } ne verersen
garbage-out } onu alırsın. :)

İyi dataya iyi sonuç verir. (NN) Kompleksli yapısından kaynaklıdır.

No free lunch

Az data ile linear regresyonda iyi sonuçlar alabilecekken neural network ile zorlanmaya gerek yoktur. Çünkü data azdır.

Cost function

learning rate az olunca yavaş öğrenicek
Data fazla dataya ihtiyaç duyacak. (6)

Nasıl öğrenir?

Cost bulundu. Adım adım güncellenecek. (Weightlar.)
Geriden gelinip tüm sistem güncellenir.

Convolutional Neural Networks

İşlem maliyetini düşük boyuta getirirler.
Deep learning'e giden bir yoldur.

Multi-layer olan mimariyi convolutional nn'e
dönüştüren işlemler nelerdir? Niye böyle bir
ihtiyaç doğdu?

Bazı optimizasyonlar yapılıyor. Data sıkıştırılıyor
gibi düşünülebilir. İşlem sayısını azaltıyoruz.
Bu sebeple deep learning'e giden yol diyoruz.

Convolutional Garpım kullanılır. Bu Garpım optimizasyon-
nu getirir. (Formülü koymadım. Rotaat ulaşılabilir.)

5×5 , 3×3 'lük vb. filtreler ile garpılır. Bir
sonraki layer bu şekilde oluşur.

Sıkıştırma işlemi uygulanır. Tekrar bir
formül oluşmuş olur. (Hocanın notlarında mevcut.)

Nasıl parametre sayısı azalır?

Kaan İNCE
14.01.2023

Filtrelerde paylaşımlı ağırlıklar mevcuttur.
5x5'lik filtrenin bağlanması için 25 ağırlığı vardır.

$$(25+1) \rightarrow \text{bias}$$

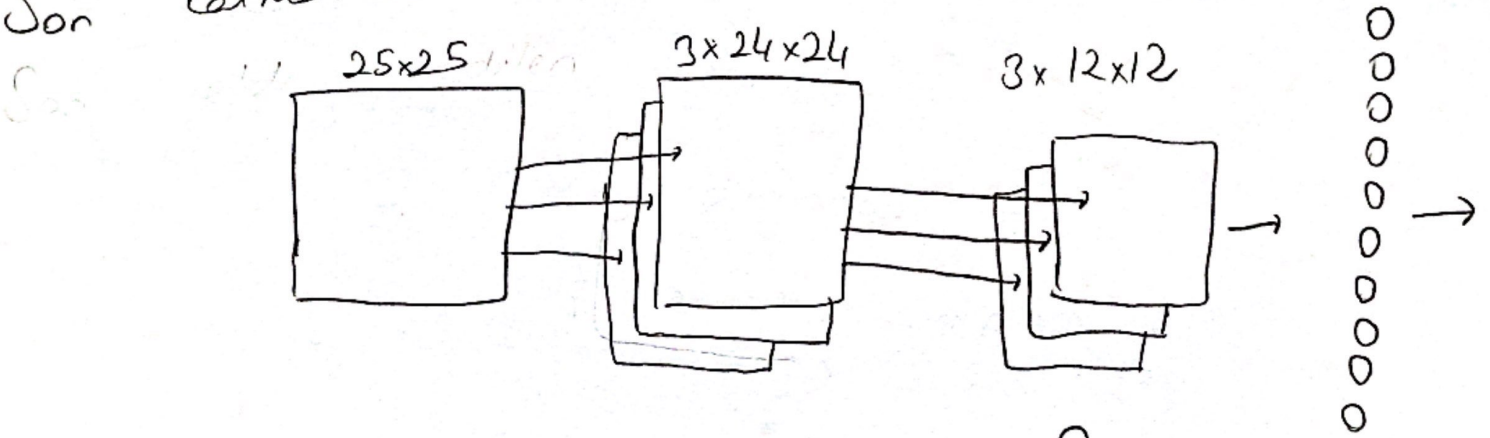
20 feature maps the convolutional layer.
 $20 \times 26 = 520$ parameters defining.

Pooling layers

- max pooling unit kaybetmez ise
- Ortalama pooling

parametre sayısını azaltırız.
boyut çok büyük olacak.

Son katmanda artık Dense layera dönüştürülür.



NN'i train etmek neden zordur?

kedi mi köpek mi? tahmini

First layer → Farklı bilgi içerir. CNN yaparak daha derine indik.

The vanishing gradient problem (Türevin bitmesi problemi)

Layer ekledik yükseldi ve sonraki layer için düştü. Overfittige girerek ezberleniş oluyor. Ön önemi öğrenme farklı hiperparametre arayarak olur. Ama maliyetli. ⑧

Layer etledikçe etledik. Fakat öğrenme hızı yavaşlar. Ara katmanlara gittikçe türevin değeri küçülür. Türev kaybolur. Hatanın değeri yok oluyor.

multi-label / multi-class problemler için

Bilgi	Softmax	→	değilim	→	toplam 1'e eşit
	Sigmoid	→	Overlapse	→	toplam 1'e eşit değil

learning rate, cost function, backpropagation değeriyle bilir. İlerde bunlara değineceğiz.

Son katman daha rahat öğrenir. (Backpropagation da) Alt katmanlara indikçe yavaşlama gerçekleşir.

Bazen hiç öğrenemiyor türev gidiyor. Başta hiperparametreler ayarabilir veya başka teknikler uygulatabilir. Bu teknikler regülasyon dersinde daha ayrıntıya ulaşacaktır.

The Exploding Gradient Problem

Türev değeri, 0'a değil tam tersine gidiyor. Uçlarda oluruz. Ya çok iyi ya çok kötü. Sonuç değeri vs. verilebilir.

Q1 = Given an image of size 27×27 and a filter of size 5×5 with a stride of 2 and no padding. What is the output size?

$$(27 - 5 + 1) / 2 + 1 = 12 \text{ dir.}$$

$$\text{Formula} = (27 + 2 * \underset{\substack{\downarrow \\ \text{padding}}}{p} - 5 + 1) / 2 + 1$$