

Örüntü Tanıma

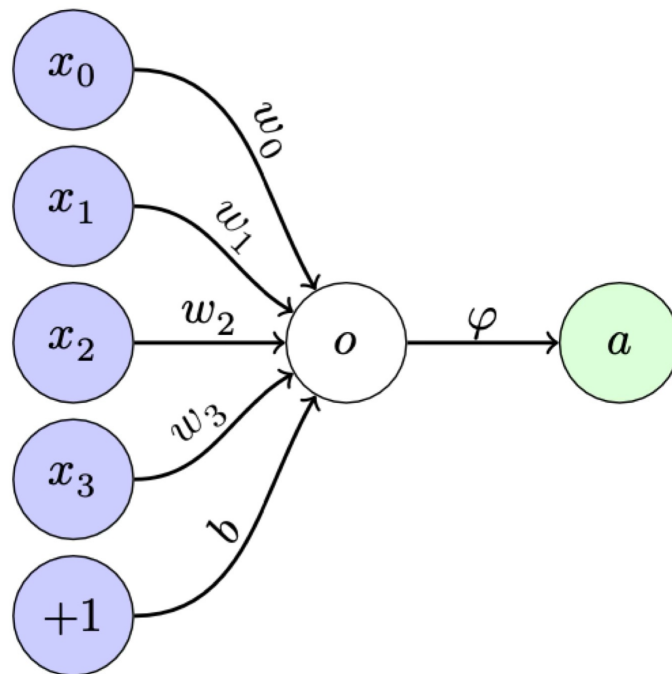
Emir Öztürk

Oğuz Kırat

Sinir Ağları ve Deep Learning

- Sinir ağıları
- Derin öğrenme
- Derin öğrenme modelleri

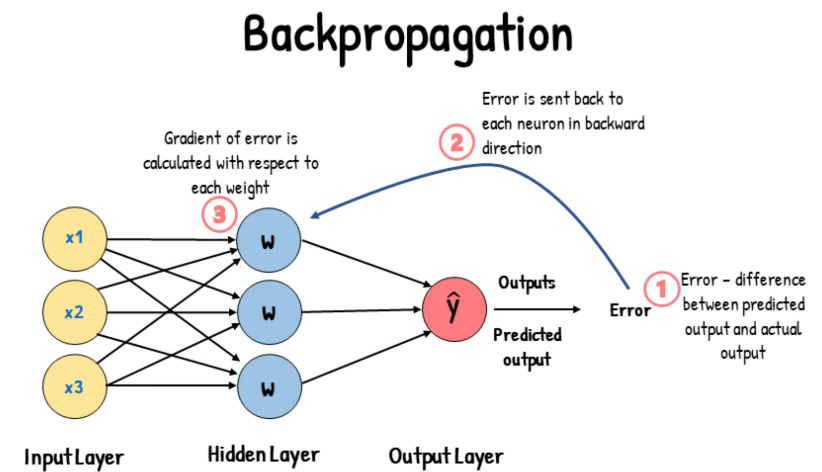
Perceptron



$$a = \varphi(\underbrace{\mathbf{w}\mathbf{x}}_o + b),$$

Backpropagation

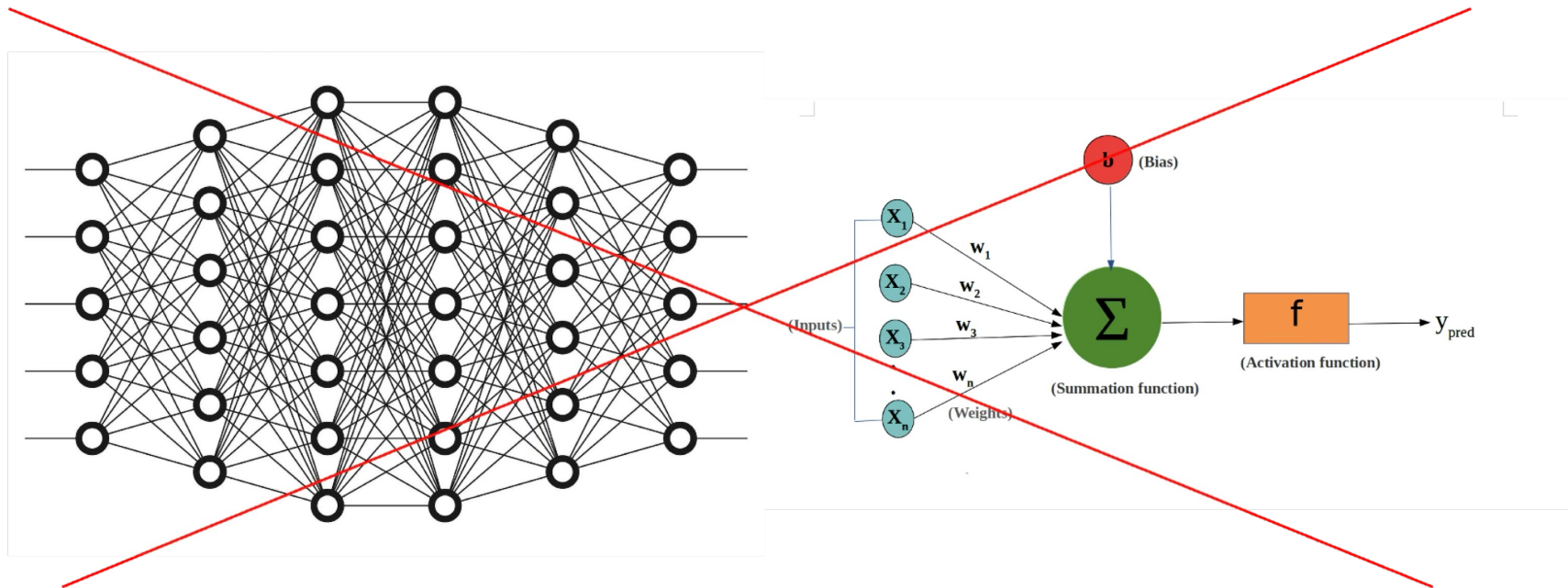
- Geri besleme
- Kaybın girdi olarak kullanımı
- Ağırlık güncellemesi



Deep Learning

- Özellik seçiminde daha az işlenmiş veri
- Yapının detayının tespitine gerek olmaması
- İhtiyaç duyulan temel yapı
 - Veri
 - Çıktı
 - Başarı hesabı - Kayıp
 - Modelin yapısı

Deep Learning



Deep Learning

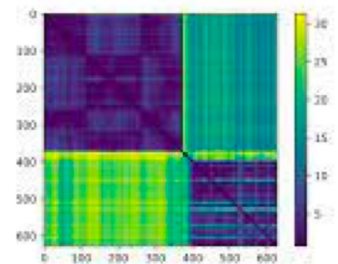
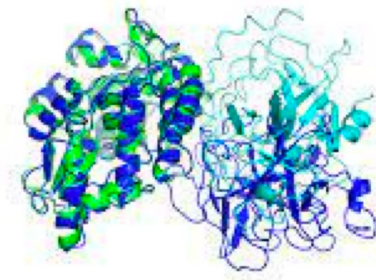
- Deep learning'in popülerleşme sebebi
 - GPU'lar - TPU'lar
 - Verinin artması
 - Belirli benchmark ve geliştirme toollarının ortaya çıkması

Deep Learning

- Modellerin geliştirilmesi hala sınırlı bir uzayda
- Çok büyük uzaya sahip sorunların belirli sadeleştirmelere ihtiyaç duyması
- Örneğin
 - Satranç
 - Go

Uzay Küçültme

- Belirli problemler deep learning için hala fazla karmaşık
- Boyutun bağımsız küçültülebilmesi
- 3D - 2D haritalama
- AlphaFold



TM-score = 67.4, DockQ = 0.04, N_{res} = 628, PDB ID = 6QF7

Bilinen Algoritmalar



Convolutional Neural Network (CNN)

AlexNet, VGGNet, ResNet, InceptionNet, DenseNet, EfficientNet, YOLO, Faster R-CNN, Mask R-CNN



Recurrent Neural Network (RNN)

Vanilla RNN, LSTM, GRU, Bidirectional RNN, Sequence-to-Sequence (Seq2Seq)



Generative Modeller

Generative Adversarial Network (GAN), Variational Autoencoder (VAE), Deep Boltzmann Machine (DBM), Restricted Boltzmann Machine (RBM)



Deep Belief Network (DBN)

Deep Boltzmann Machine (DBM), Stacked Autoencoder



Transformer'lar

Transformer, BERT, GPT, T5, RoBERTa, XLNet



Autoencoder

Variational Autoencoder (VAE), Denoising Autoencoder (DAE), Sparse Autoencoder, Contractive Autoencoder

Veri eldesi

Ham veri

Crawling

Veriseti kaynakları

- <https://huggingface.co/datasets>
- <https://www.kaggle.com/datasets/>
- <https://research.google/resources/datasets/>
- <https://www.datasetlist.com/>
- <https://datasetsearch.research.google.com/>
- <https://dumps.wikimedia.org/>
- <https://paperswithcode.com/datasets>

Önişleme

Veri Temizleme

- Eksik Verilerle İlgili İşlemler
- Aykırı Değerlerin Tespiti ve Kaldırılması
- Yinelenen Kayıtların Silinmesi

Veri Dönüşümü

- Normalizasyon
- Standartlaştırma
- Log Dönüşümü

Veri Kodlama

- Etiket Kodlama (Label Encoding)
- One-Hot Encoding
- Embedding

Özellik Mühendisliği

- Özellik Çıkartma
- Özellik Seçimi
- Boyut İndirgeme

Veri Artırma (Görüntü ve Metin Verisi İçin)

- Görüntü Artırma
- Metin Artırma

Önişleme



Veri Bölme

Eğitim ve Test Verisi
Bölme
Çapraz Doğrulama



Dengesiz Veriyle Baş Etme

Yeniden Örnekleme
Yöntemleri

- Aşırı Örnekleme
- Azaltma
- Sentetik Veri



Metin İşleme (NLP Görevleri İçin)

Tokenizasyon
Stopword Temizleme
Kök Bulma /
Lemmatizasyon
Vektörleştirme

Eğitim Aşaması

- Sıfırdan Eğitim
 - Kaynak problemi
 - Zaman
 - Veri ihtiyacı
 - Başarı oranı
- Finetuning
 - PEFT (Parameter efficient Fine Tuning)
 - LoRA (Low-Rank Adaptation)
- One shot learning
- Zero shot learning