

Derin öğrenme - LSTM (uzun-kısa süreli belek) Sinir Ağı

HÜSEYİN CANIK-21360859064

Bilgisayar Mühendisliği/3 - Seminer Dersi

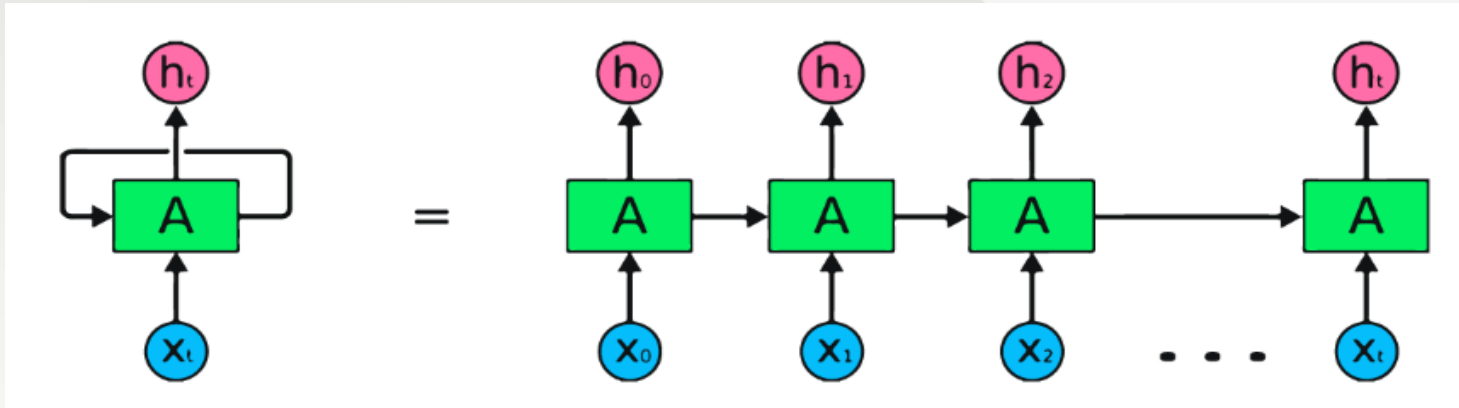
2023-2024 BAHAR DÖNEMİ

İÇERİK

- 1-LSTM Nedir ?
- 2-LSTM Model Mimarisi
 - > 2.1-Hücre Kapıları
 - > 2.2-Sigmoid Fonksiyon Tanıtımı
 - > 2.3-Unutma Kapısı
 - > 2.4-Girdi Kapısı
 - > 2.5-Güncelleme Kapısı
 - > 2.6-Çıkış Kapısı
- 3-Kullanım Aşaması
 - > 3.1-Kütüphane Yükleme
 - > 3.2-Model Oluşturma
 - > 3.3-Optimizasyon Metrikleri
- Örnek Kullanım Alanları
- Kaynaklar
- Sorular
- Teşekkürler

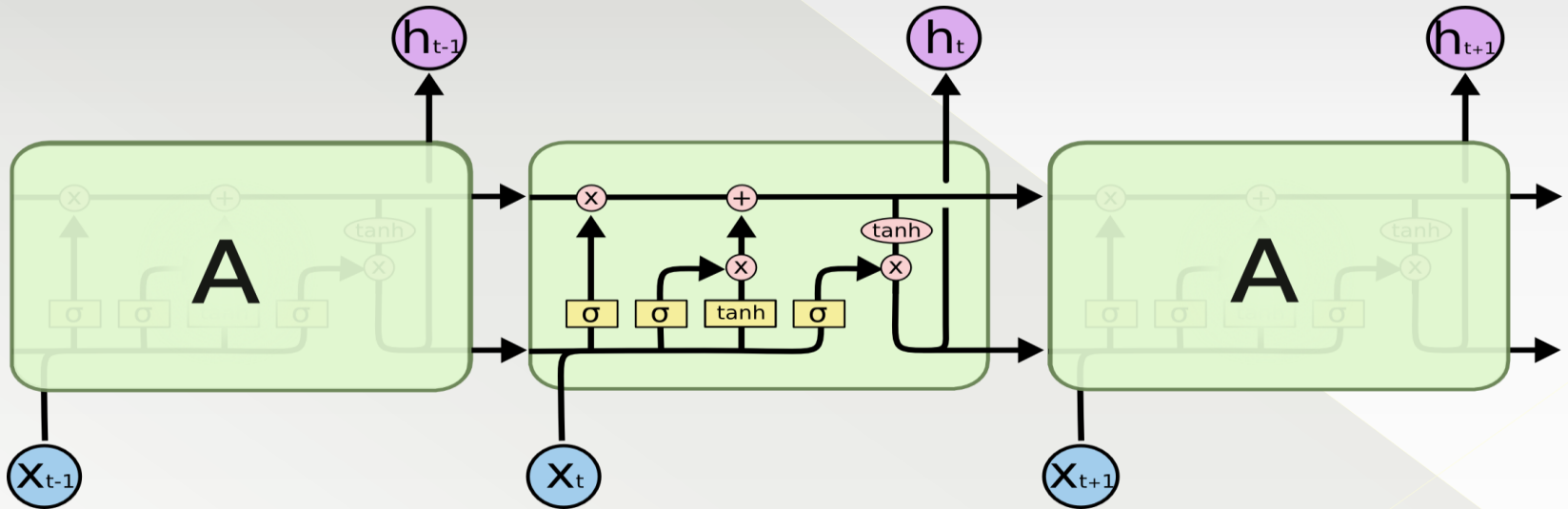
1-LSTM Nedir ?

- Dizi şeklinde sıralanmış yapay sinir ağı olarak çalışır.
- Zaman serilerinde , metin üretme ve metin sınıflandırmada
- Ses , görüntü ve video sinyali analizi gibi alanlarda kullanılır.
- LSTM'in amacı, girilen veriler arasındaki zamanla değişen karmaşık ilişkileri öğrenmek ve bu ilişkileri kullanarak gelecekteki çıktıları tahmin etmektir.



2-LSTM model mimarisi

- Üç hücreden oluşan örnek bir LSTM ağı.



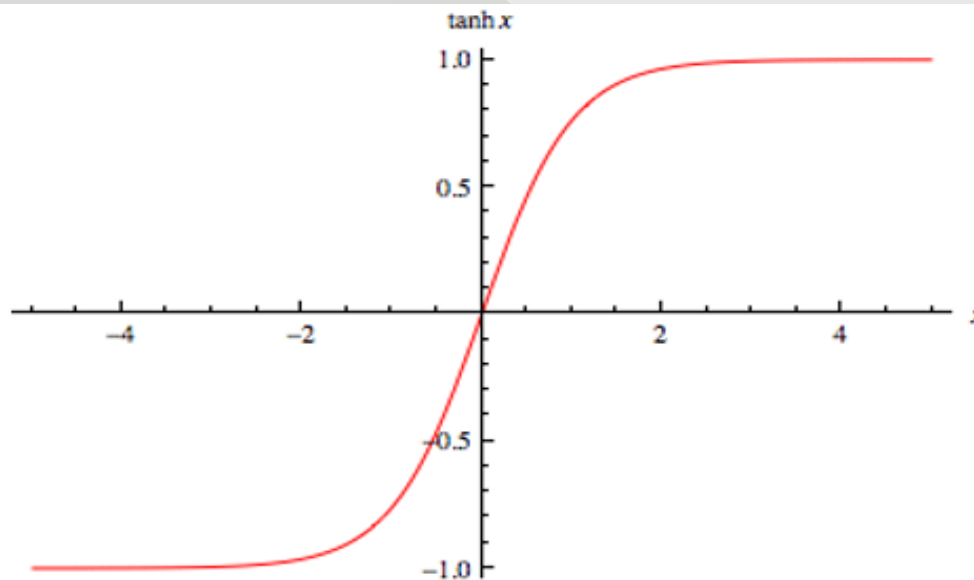
2.1-Hücre Kapıları

- ❖ Her hücre 4 temel kapı içerir.
- ❖ İç içe geçmiş bu kapıların görevi hatırlanacak veya unutulacak bilgileri belirlemektir.



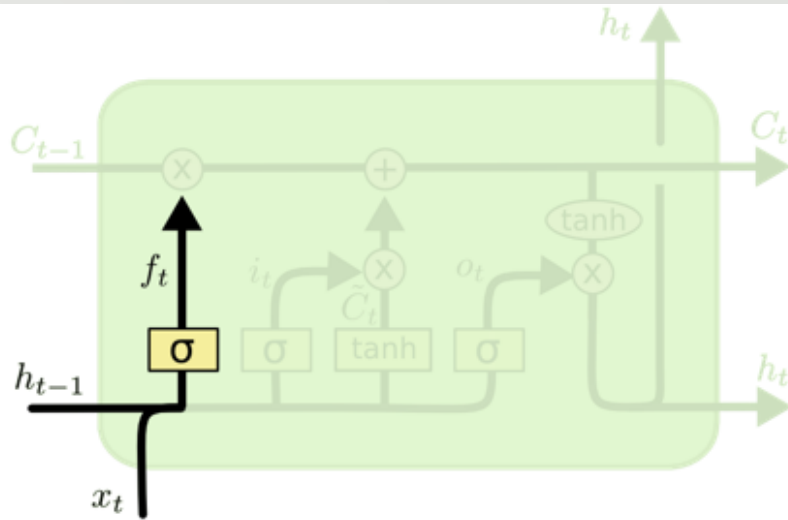
2.2-Sigmoid Fonksiyonu Tanıtımı

- Hiperbolik tanjant(\tanh) fonksiyonu -1 ile 1 arasında çıktı üreten doğrusal olmayan bir fonksiyondur.
- **Sigmoid fonksiyonu** hiperbolik tanjant fonksiyonunu 0 ile 1 arasında çıkış fonksiyonuna dönüştüren ve doğrusal olmayan fonksiyon olarak kullanılır.



2.3-Unutma Kapısı

- ✓ Yeni bilgi(x_t), geçmiş bilgi(h_{t-1}) ve bayes değeriyle oluşturulan değer sigmoid fonksiyonunda 0 ila 1 arasında bir değere dönüşür
- ✓ f_t fonksiyonu sonucuna göre, kapının 0'a yakın bir çıktı verdiği bilgiyi unuttur 1'e yakın verdiği değeri hatırlar .



Rastgele oluşturulan ağırlık değeridir. Bilgi ağırlıklandırılır.

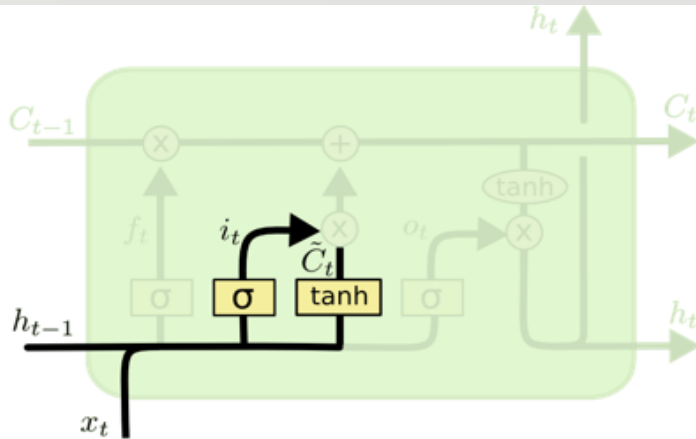
$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f)$$

Sigmoid fonksiyonu

Bayes değeri

2.4-Girdi Kapısı

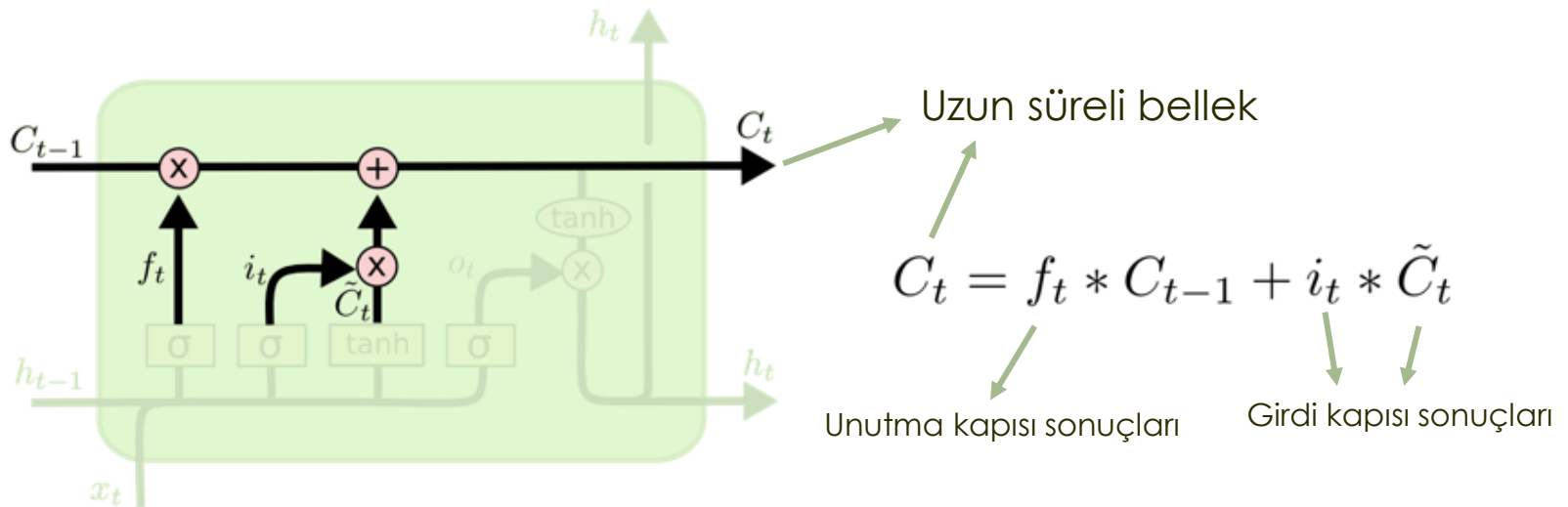
- Yeni bilgi ve geçmiş bilgi ağırlıklar eklenerek sigmoid ve hiperbolik tanjant fonksiyonundan geçirilir.
- Gelen bilginin hücre durumuna eklenip eklenmeyeceğini belirler amaç hücre durumunu güncellemek ve uzun süreli belleğe veri aktarmaktır.



$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

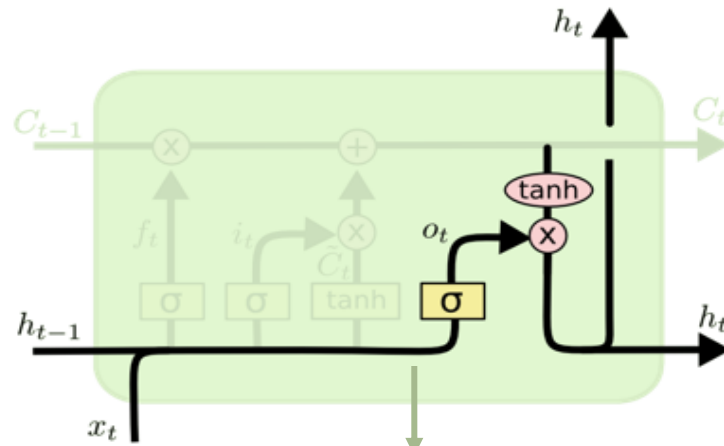
2.5-Güncelleme Kapısı

- Unutma ve girdi kapısından elde edilen sonuçlar uzun süreli bellek kısmına aktarılır ve bellek güncellenir.



2.6-Çıkış Kapısı

- Çıkış kapısı hücrenin son kapısı olup kendinden önceki bilgileri işleyerek mevcut durumları sonlandırır.
- Bu kapıdan sonra üretilen bilgi çıktı olarak tutulur ve bir sonraki hücreye aktarılır.



Girdiyi ve eski bilgiyi taşıyan kısa süreli bellek

$$o_t = \sigma (W_o [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh (C_t)$$

Yeni bilgi hafızaya alınıyor ve bir sonraki hücreye aktarılıyor

3.1-Kütüphane Yükleme

```
In [ ]: # Veriyi okuma ve işleme adımımda kullanılacak olan kütüphaneler
import numpy as np
import pandas as pd
import tensorflow as tf
import matplotlib.pyplot as plt

# Model değerlendirme ve veriyi scale edebilmemiz için kullanılacak olan kütüphaneler
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error

# Model için kullanılacak olan kütüphaneler
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, LSTM, Dropout
from tensorflow.keras.callbacks import ModelCheckpoint, EarlyStopping
```

- LSTM modeli oluşturmak için gerekli adımlar

3.2-Model oluşturma

```
In [ ]: #yapay sinir ağı modeli oluşturuldu
model = Sequential()

#Modele LSTM hücreleri eklendi
model.add(LSTM(units=50,
               activation='sigmoid',
               input_shape=(X_train.shape[1], lookback) ))

#Modelin veriyi ezberlememesi için rastgele sönümleme oranı verildi
model.add(Dropout(0.2))

#çıkış katmanı eklendi
model.add(Dense(1))
```

❑ Örnek bir model şeması

3.3-Modelin optimizasyon Metrikleri

```
In [ ]: # Optimizasyon ve Değerlendirme Metrikleri Ayarlanması
model.compile(loss='mean_squared_error', optimizer='adam')

# !!! EarlyStopping işlemi ve loss fonksiyonu
callbacks = [EarlyStopping(monitor='val_loss', patience=3, verbose=1, mode="min"),
ModelCheckpoint(filepath='mymodel.h5', monitor='val_loss', mode='min',
save_best_only=True, save_weights_only=False, verbose=1) ]
```

❖ Loss

❖ EarlyStopping

❖ Patience

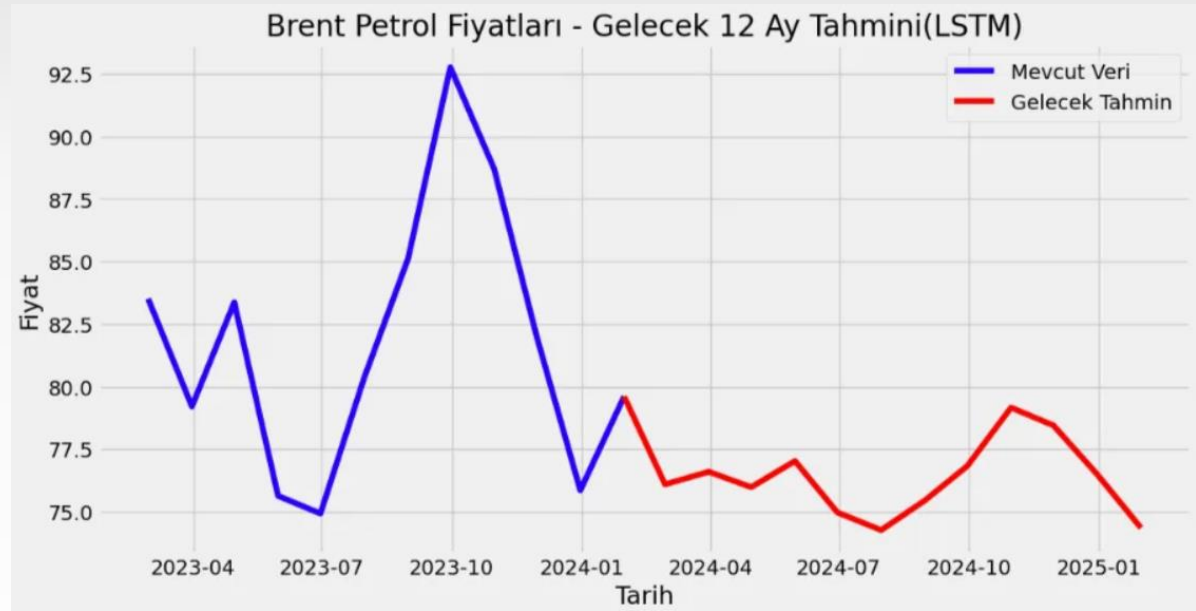
4.1-Örnek kullanım alanı

Eğitim verisi

Tahmin Verisi

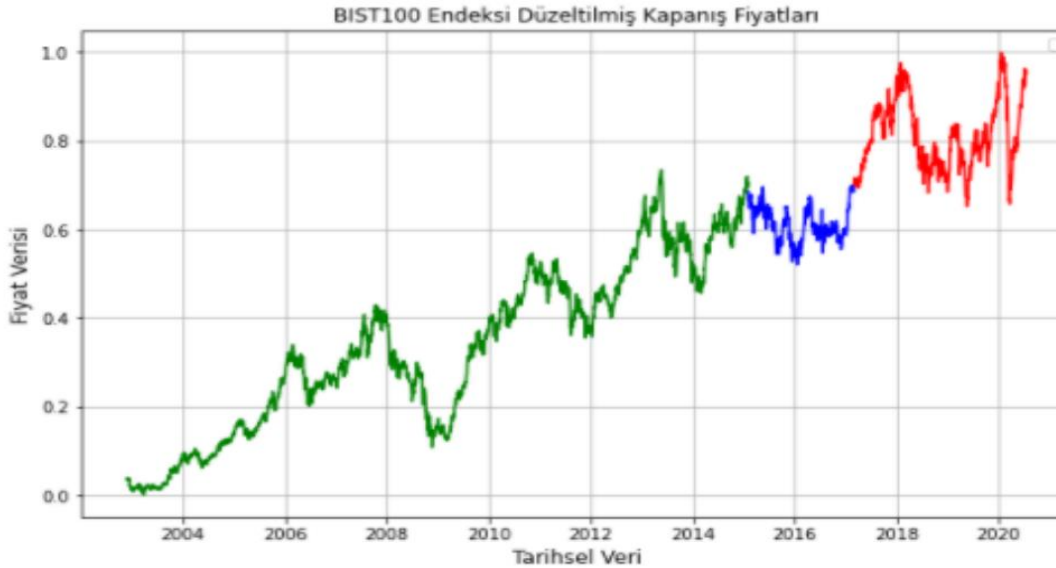
Tarih	USD/TL	ParaArzi	Motorin	Perakende Satış	Sanayi	Kredi	M3	Borç	TÜFE
2010-02-28	0.574311	-0.468328	-0.470323	0.935668	2.358713	2.473323	-0.317515	1.774400	1.447693
2010-03-31	0.644001	1.850070	2.622296	1.971218	1.929881	2.894349	1.831928	1.064192	0.583272
2010-04-30	0.478728	1.038067	0.713406	0.918462	1.021775	3.033220	0.969354	0.366513	0.596780
2010-05-31	0.646630	0.240025	-0.856462	0.610769	2.763430	3.592510	0.095471	0.729264	-0.358182
2010-06-30	0.922501	1.537408	-0.175370	-0.094972	0.563442	3.358199	1.609402	-0.769974	-0.561672
...
2023-07-31	11.475241	15.146856	33.393665	2.986884	-0.403339	5.675617	15.439600	9.689321	9.488824
2023-08-31	10.505114	5.943431	26.763908	-4.603827	-0.795784	1.850455	5.968276	2.618269	9.086793
2023-09-30	5.071342	4.764088	4.843457	-0.605829	-0.083836	2.080815	4.845313	4.384930	4.753114
2023-10-31	1.709911	2.694244	-0.714650	1.976717	-0.388520	2.752662	2.613859	3.653039	3.433981
2023-11-30	2.001541	3.356285	-3.136247	0.000000	0.000000	2.547674	3.373976	0.000000	3.281097

166 rows × 9 columns



4.2-Kur tahmini

- Eğitim için verilen verinin grafiği



- Gerçek değerler ile tahmin edilen değerlerin karşılaştırılması

Kaynaklar

- ◉ **Raymond J. Mooney**-University of Texas at Austin/pdf
- ◉ Turkcell/Geleceği yazarlar/LSTM
- ◉ <https://www.cs.sjsu.edu/faculty/pollett/masters/Semesters/Fall19/parnika/LSTM.pdf>

SORULARINIZ...

DİNLEDİĞİNİZ İÇİN TEŞEKKÜR EDERİM



Sorular

- 1-Hiperbolik tanjant fonksiyonunun çıktılarını 0 ila 1 arasında çıktılarına dönüştüren aktivasyon fonksiyonu nedir?
- Relu
- Sigmoid
- MSE

Sorular

- 2-LSTM hücresi kaç kapıdan oluşur?
- 3
- 4
- 5

Sorular

- ③ 3-Modelin gerçek değerler ile tahmin edilen değerlerin arasındaki hata oranını gösteren fonksiyon nedir ?
- ③ -Sigmoid
- ③ -Loss
- ③ -Hiperbolik tanjant

Sorular

- ◎ 4-Modelin öğrenememesi durumunda erken bitirilmesi için gerekli kütüphane nedir?
- ◎ -Sequential
- ◎ -EarlyStopping
- ◎ -Dropout