

機械学習勉強会

時系列解析

講師紹介

野田 真史 博士（工学）

分子科学研究所、筑波大学でポスドクを経た後、アカデミアに入社。

これまでの専門は第一原理計算を用いた物性予測。

SALMON(電子ダイナミクス計算プログラム)の講習会の講師を2度務める。

G検定所持。

RNN (Recurrent Neural Network)

$$h_j = \tanh(Wx_j + Uh_{j-1} + b)$$

h_j : 中間層からの出力

W : 入力層から中間層への重み

x_j : 入力

U : 中間層から中間層への重み

b : バイアス

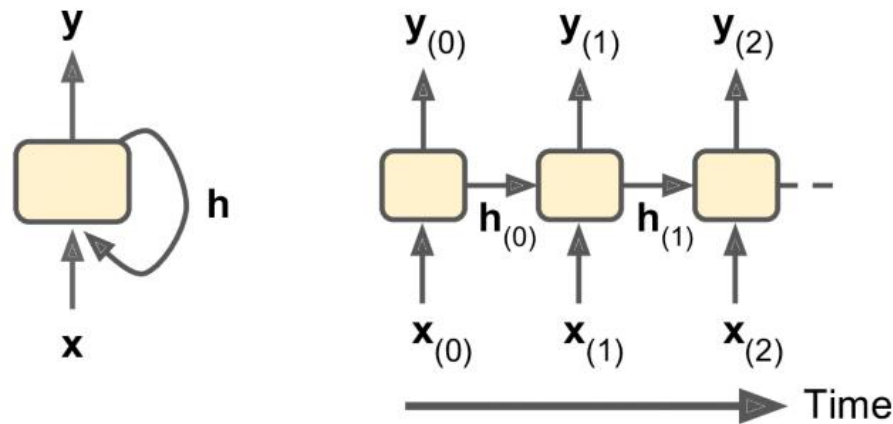


Fig. 15-3 (Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, Aurélien Géron)

勾配消失問題

$$\frac{\partial E_t}{\partial W} = \sum_{k=0}^T \frac{\partial E_t}{\partial y_t} \frac{\partial y_t}{\partial h_t} \left(\prod_{j=t+1}^T \frac{\partial h_j}{\partial h_{j-1}} \right) \frac{\partial h_t}{\partial W}$$

E_t : 誤差関数

$$h_j = \tanh(Wx_j + Uh_{j-1} + b)$$

$$\frac{\partial h_j}{\partial h_{j-1}} = U(1 - \tanh^2(Wx_j + Uh_{j-1} + b))$$

$$\left| \frac{\partial h_j}{\partial h_{j-1}} \right| < 1 \text{ のため } T \text{ が大きくなると誤差が消失する。}$$

LSTM (Long Short Term Memory)

$$\mathbf{i}_{(t)} = \sigma(\mathbf{W}_{xi}^T \mathbf{x}_{(t)} + \mathbf{W}_{hi}^T \mathbf{h}_{(t-1)} + \mathbf{b}_i)$$

$$\mathbf{f}_{(t)} = \sigma(\mathbf{W}_{xf}^T \mathbf{x}_{(t)} + \mathbf{W}_{hf}^T \mathbf{h}_{(t-1)} + \mathbf{b}_f)$$

$$\mathbf{o}_{(t)} = \sigma(\mathbf{W}_{xo}^T \mathbf{x}_{(t)} + \mathbf{W}_{ho}^T \mathbf{h}_{(t-1)} + \mathbf{b}_o)$$

$$\mathbf{g}_{(t)} = \tanh(\mathbf{W}_{xg}^T \mathbf{x}_{(t)} + \mathbf{W}_{hg}^T \mathbf{h}_{(t-1)} + \mathbf{b}_g)$$

$$\mathbf{c}_{(t)} = \mathbf{f}_{(t)} \otimes \mathbf{c}_{(t-1)} + \mathbf{i}_{(t)} \otimes \mathbf{g}_{(t)}$$

$$\mathbf{y}_{(t)} = \mathbf{h}_{(t)} = \mathbf{o}_{(t)} \otimes \tanh(\mathbf{c}_{(t)})$$

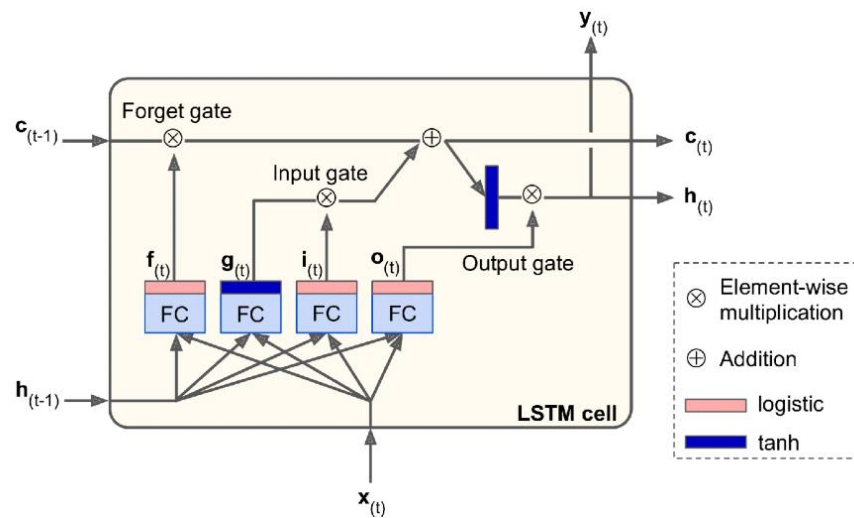


Fig. 15-9 (Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, Aurélien Géron)

忘却ゲートを設けることにより勾配消失問題を解消。