

---

# 3. 神经网络入门

---

WU Xiaokun 吴晓堃

xkun.wu [at] gmail

2021/03/29

# 本章内容

神经网络剖析。IMDB电影评论分类：二分类问题。路透社新闻分类：多分类问题。  
波士顿预测房价：标量回归问题。实践：三类基本问题。

**重点：**层、神经网络拓扑、监督学习的三类基本使用场景；

**难点：**监督学习中三类基本问题的Keras基础实现。

# 学习目标

- 理解神经网络的基本构成单元（层）及其拓扑结构；
- 理解监督学习的三类基本使用场景：二分类问题、多分类问题和标量回归问题；
- 掌握监督学习中三类基本问题的Keras基础实现。

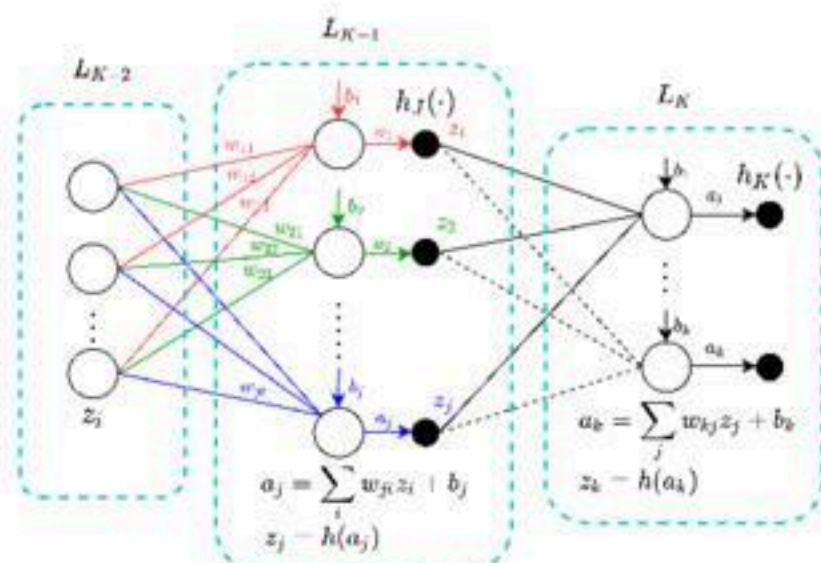
# 概念

## 层

层是数据处理单元，将输入张量转换为输出张量。

## 模型

层构成的网络（有向无环图）；网络的拓扑结构定义了一个假设空间。例如：密集连接网络假设输入特征中没有特定结构。



$$\mathbf{y} = L_K \circ L_{K-1} \circ L_{K-2} \cdots L_1(\mathbf{x})$$

# 监督学习的三大基本问题

## 二分类

- 概率输出: sigmoid 激活;
- 损失函数: 二元交叉熵 (binary\_crossentropy) 。

## 多分类

- 概率输出: softmax 激活;
- 损失函数: 分类交叉熵 (categorical\_crossentropy) 。

## 回归

- 标量输出: 无激活;
- 损失函数: 均方误差 (MSE, mean squared error) 。

# 二分类问题

```
from keras import models
from keras import layers

model = models.Sequential()
model.add(layers.Dense(32, activation='relu', input_shape=
    (num_input_features,)))
model.add(layers.Dense(32, activation='relu'))
model.add(layers.Dense(1, activation='sigmoid'))
```

# 单标签多分类问题

```
from keras import models
from keras import layers

model = models.Sequential()
model.add(layers.Dense(32, activation='relu', input_shape=
    (num_input_features,)))
model.add(layers.Dense(32, activation='relu'))
model.add(layers.Dense(num_classes, activation='softmax'))
```

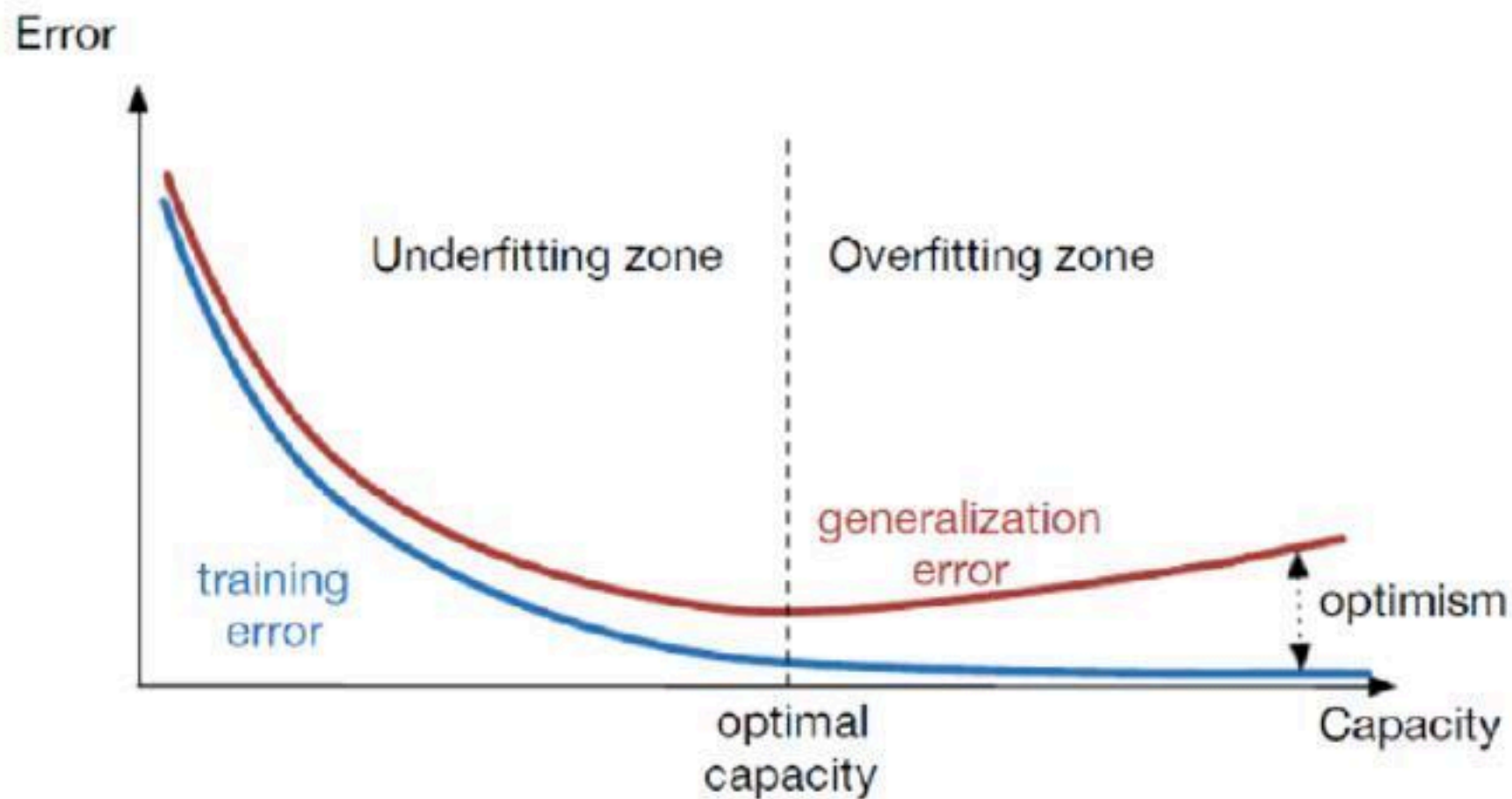


# 多标签多分类问题

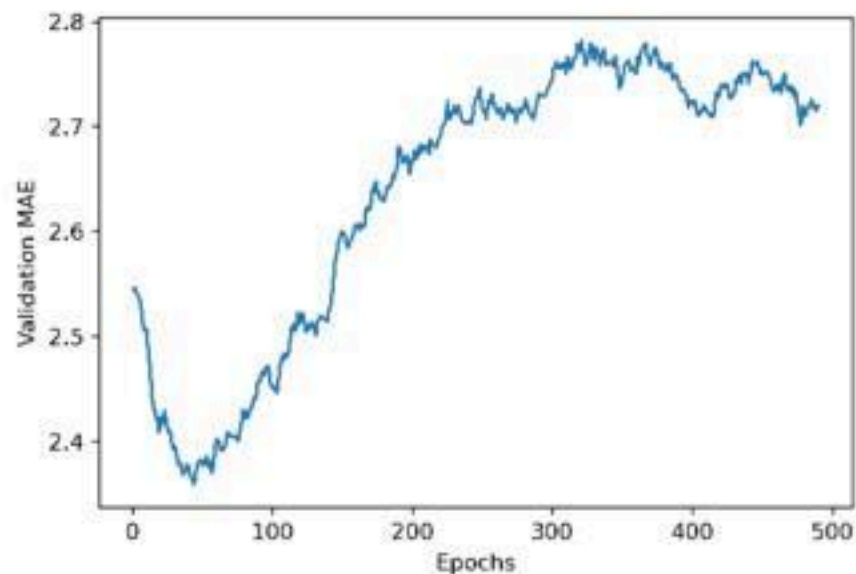
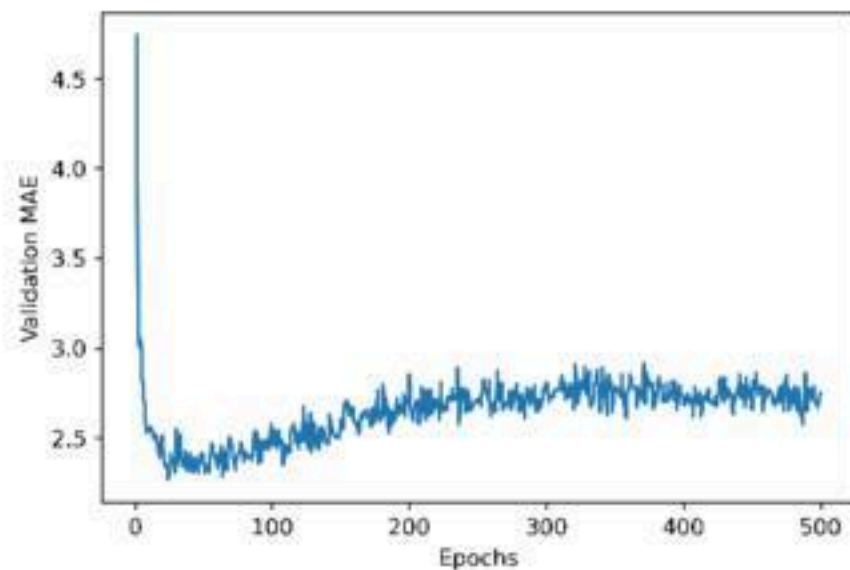
```
from keras import models
from keras import layers

model = models.Sequential()
model.add(layers.Dense(32, activation='relu', input_shape=
    (num_input_features,)))
model.add(layers.Dense(32, activation='relu'))
model.add(layers.Dense(num_classes, activation='sigmoid'))
```

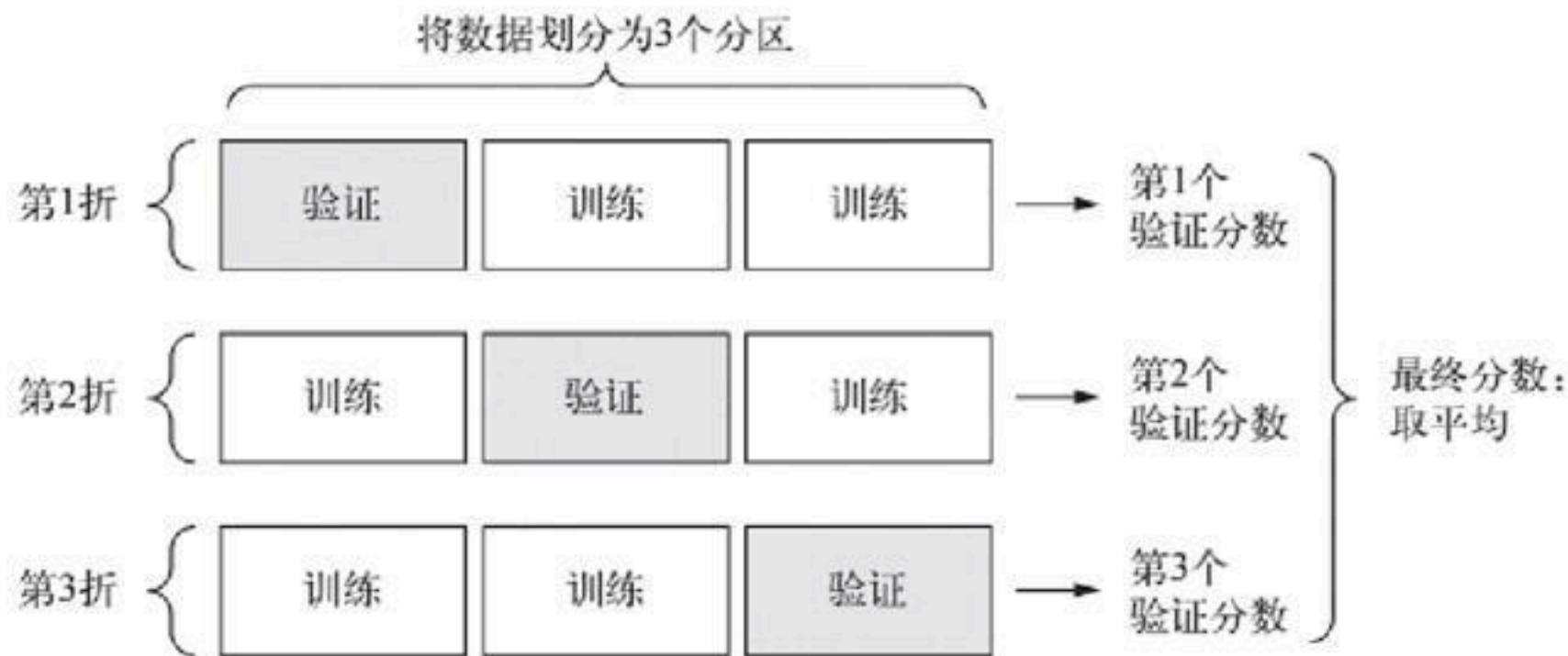
# 过拟合与欠拟合：理想曲线



# 过拟合与欠拟合：实际曲线



# K-fold 交叉验证



# 回归问题

```
from keras import models
from keras import layers

model = models.Sequential()
model.add(layers.Dense(32, activation='relu', input_shape=
    (num_input_features,)))
model.add(layers.Dense(32, activation='relu'))
model.add(layers.Dense(num_values))
```