

# 机器学习纳米学位

---

## 算式识别 金阳 Udacity

2018年12月16日

## I. 问题的定义

---

### 项目概述

使用深度学习识别一张图片中的算式。从数据图片观察，图片中有噪点，且每个符号有一定角度的旋转，应该是验证码。这个数据集和mnist手写体识别挺像的，最大的区别是算是识别图片中，有多个字符。

### 问题陈述

这是一个图片识别问题，所以需要用到卷积神经网络（CNN），并且需要对图片数据做一些预处理。算式图片中出现的长度是不定长的，需要用到递归神经网络（RNN）得到计算结果。我决定使用卷积神经网络提取出特征之后，输入到递归神经网络中，识别出其中的算式。

### 评价指标

正确率=识别正确的算式数量/算式的总数

当算式图片识别出来的每个字符都正确时，该算式为识别正确。

## 数据以及观察结果

---

此数据集包含10万张图片，每张图里面都有一个算式。

实例图片：



实例图片中的字符:(0+0)+9=9

- 每个算式可能包含  $+$ 、 $-$ 、 $*$  三种运算符，可能包含一对括号，可能包含0-9中的几个数字，以及每个算式包含一个等号。所以一共出现的字符总数是16种。
- 每个字符都可能旋转。
- 图片大小统一是300\*64。
- 图片字体是各种颜色的，背景也是各种颜色的，但是背景都是浅色（接近白色）
- 图片中有一些噪点。

## 预期的解决方案

---

- plan A 图片预处理后，从算式图中切分出每个字符的图片，然后使用每个字符的图片做识别比较。
- plan B 图片预处理后，直接将算式图输入模型，然后结合CNN和RNN，直接输出。

## 数据预处理

---

图片先转成灰度图，然后除以255方便模型计算。把label转换成one-hot的形式。

## 项目设计

---

1. 数据预处理。
2. 设计一个结合CNN和RNN的模型。
3. 训练模型直到达到，大于等于99%正确率。