

## 神经网络系列一

### 基于卷积神经网络(CNN)的期货价格预测

#### 内容摘要

自 1955 年达特茅斯会议的计划书《A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence》发表，人工智能走过了 60 余年，伴随着神经网络在学术研究和商业应用多次起伏，自 2006 年辛顿《A Fast Learning Algorithm for Deep Belief Nets》的突破性文章发表，作者在此文章中介绍了训练多层神经网络的方法，自此神经网络在安防，交通，在线购物，信息检索，无人机，机器人等领域的发展一日千里，从 ImageNet 竞赛到 AlphaGo 战胜李世石，深度学习的应用热潮再也没熄灭过。

在此，我们将深度神经网络引进金融资产价格预测，将金融时间序列转换成图像，以涨跌作为标识，通过卷积神经网络建模训练，分析价格预测的可行性和应用前景。

投资咨询业务资格：

证监许可【2012】669 号

研究咨询部  
量化组

刘宾  
0755-83212741  
liubin@citicsf.com  
从业资格号：F0231268  
投资咨询号：Z0000038

王建伟 CFA FRM  
021-60812992  
wangjianwei@citicsf.com  
从业资格号：F3014595  
投资咨询号：Z0013229

#### 联系人

邹天舒  
021-60812993  
zoutianshu@citicsf.com  
从业资格号：F3027249

王炳瑜  
021-60812989  
wangbingyu@citicsf.com  
从业资格号：F3018918

肖璋瑜  
0755-82723054  
xiaozhangyu@citicsf.com  
从业资格号 F3034888

陈舜尧  
0755-82723054  
chenshun Yao@citicsf.com  
从业资格号：F3029712

## 目 录

内容摘要 .....	1
一、 人工神经网络 .....	3
二、 卷积神经网络 .....	5
三、 针对时间序列的卷积神经网络分类 .....	9
免责声明 .....	12

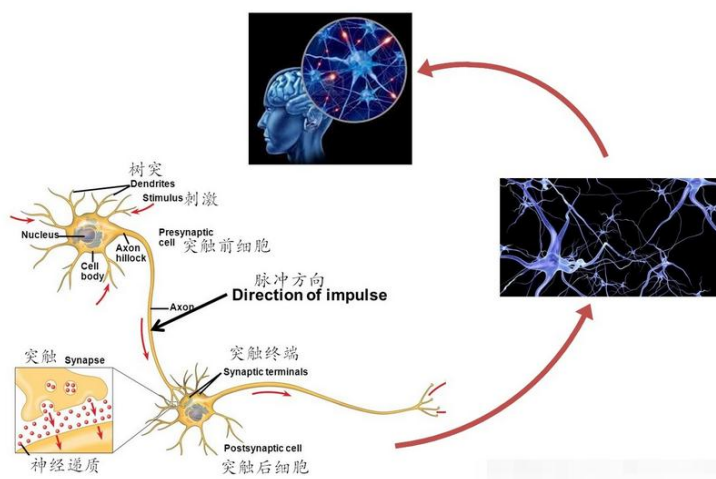
## 图表目录

图 1 神经网络结构 .....	3
图 2 感知器结构 .....	4
图 3 局部感受野 .....	6
图 4 参数共享 .....	7
图 5 LeNet-5 .....	8
图 6 LeNet-5 链接图 .....	8
图 7 多元时间序列卷积网络结构 1 .....	9
图 8 多元时间序列卷积网络结构 2 .....	10
图 9 多元时间序列卷积网络训练结果 .....	10

## 一、人工神经网络

1903 年，西班牙科学家卡哈尔通过改良的高尔基染色法首次观察到了神经网络的细微结构，并在 1904 年发表著作确立了神经元学说，提出了神经活动的基本原则。在一个神经元细胞周围会有多个呈放射状的突起结构，称为树突。另外每个神经元会有一个细长的突起结构，称为轴突。轴突末端的分支称为轴突末梢，神经元从树突接受信号，当某一时刻累计的信号达到阈值，则通过轴突释放信号。当成千上万个神经元通过轴突链接其他神经元的树突时，便形成了复杂的神经网络。

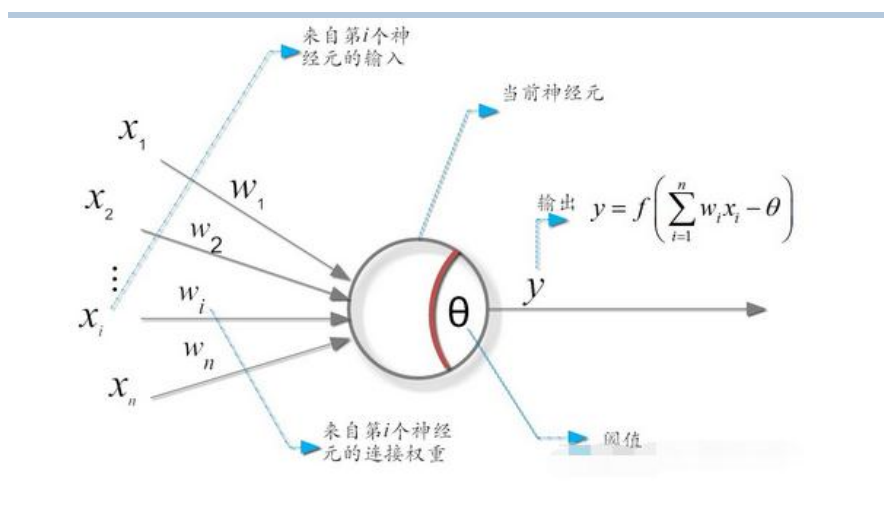
图 1：神经网络结构



数据来源：公开资料 中信期货研究部

1943 年，美国神经生理学家沃伦·麦克洛奇和逻辑学家沃尔特·匹茨一起提出了通过简单的计算模型来模拟神经元，对人类的神经系统进行了仿生建模，我们将此模型用两位作者的首字母作为简称，成为 M-P 模型。

图 2：感知器结构



数据来源：公开资料 中信期货研究部

在 M-P 模型中，所有的输入和输出都为 0 或者 1，用来模拟最基本的二进制信号，每个输入信号的权重保持一致，激活函数采用阶跃函数。在此启发下，1956 年，美国心理学家罗森布拉特提出了著名的感知机(Perceptron)，改进了信号权重，并且输入信号不限于二值信号，罗森布拉特当时通过光传感器阵列模拟了视网膜并识别了一些简单的字母。自此，人工神经网络开启了应用的大门。

然而，在当时，计算机硬件发展处于初级时期，二层感知器已经带来了难以完成的计算成本，而单层感知器却不能解决基的异或逻辑问题，由此进入了长期的研究寒冬。直到 1982 年，加州理工的霍普菲尔德提出了反馈型神经网络，成功的解决了一些识别和约束优化问题，使学术研究再次拉回到神经网络。1986 年，辛顿和鲁梅哈特合作发表了论文《Learning Representations by Back-Propagating Errors》提出了著名的反向传播算法，BP 神经网络使多层神经网络进入了实用阶段。1989 年，美国应用数学家塞班克证明了神经网络可以被看作一个通用逼近函数，一个隐含层可以逼近任意连续函数，两个隐含层的网络可以逼近任意函数，即神经网络的拟合能力是可以接近无限强，任意负责的多分类决策边界都可以被逼近。

随着实用化，多层神经网络的缺陷一一暴露。首先，随着神经网络层数的增加，BP 算法的梯度计算出现较不稳定的现象，也即是越远离输出层的参数越难以被训练，要么变化缓慢，要么变化过于剧烈，被称为梯度消失和梯度爆炸问题。其次，随着神经网络层数的增加，参数将以几何级扩张，导致神经网络

很容易出现过拟合现象。最后，神经网络的调参也是较明显的问题，使用者一般不关注细节，偏向于输入输出，隐层数，单元数的调整，导致调参无迹可寻。

直到 2006 年辛顿发表了一篇《A Fast Learning Algorithm for Deep Belief Nets》介绍训练多层神经网络的办法，将神经网络的研究引向了深度神经网络。此文章中，作者介绍了深度信念网络，用受限玻尔兹曼机(RBM)的结构得到生成网络。可以理解为，通过学习大量无标注数据，在隐层中提取数据的特征，通过此类特征来重现数据的分布。这样的过程叫做预训练，得到了网络的初始化参数，后续的工作只需要适当的微调。

至此，神经网络在近 10 年的发展历程，在计算机视觉中得到了极大的发展，另外在语音识别也得到了较大的进展。安防方面，人脸识别和指纹识别已经是我们耳熟能详的应用，苹果 X 手机将此两类应用带到了每个人的日常应用。交通方面，车辆识别，行人识别，无人驾驶已经逐渐走入我们的生活。在线购物，定向推介和定向广告越来越成为销售的趋势。

## 二、卷积神经网络

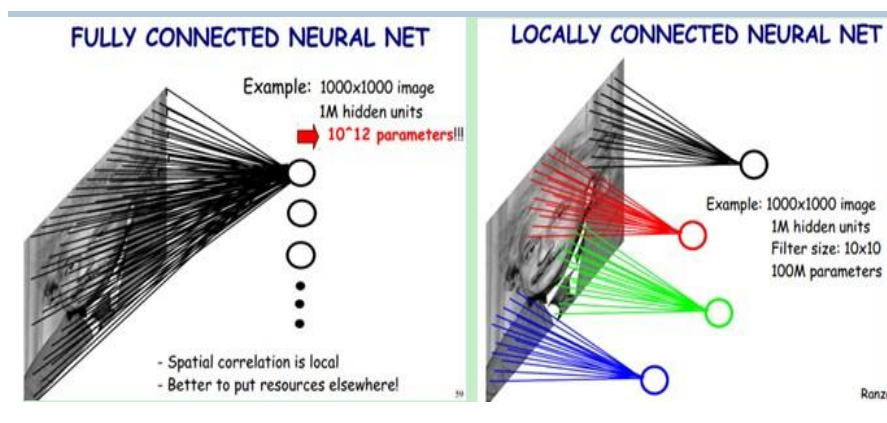
早在 1989 年，LeCun 就发明了卷积神经网络，被广泛应用于美国的银行系统，用来识别支票上的手写数字。卷积神经网络在图像的应用上得到了极大的成功，其最初的启发也是从动物的视觉神经系统获得。动物的眼睛是一个成像系统，图像通过瞳孔，晶状体在视网膜上成像，视网膜上布满了光感受细胞，把光刺激转换成神经冲动，神经冲动再经过视觉通路传递近大脑的初级视觉皮层。在视觉皮层中，对图像的处理是分层处理，每一层处理特定的信息。

卷积神经网络与简单多层神经网络并不是相同的含义，其主要区别在于卷积神经网络包含了一个由卷积层、激活层和下采样层构成的特征提取结构。在卷积神经网络的卷积层中，一个神经元只与部分邻层神经元链接。在卷积神经网络中，通常包含若干个特征图，每个特征图的神经元共享权重参数，共享的权值便是卷积核。卷积核一般随机的形式初始化，也可以通过预训练进行初始化。卷积核的应用，给多层神经网络带来了较大的提升，减少了网络各层的连接，降低了过拟合的风险。下采样也叫池化(Pooling)，可以看做一种特定的卷积过程，通过卷积和下采样使卷积神经网络的复杂度大大降低。

## 1. 局部感受野

卷积神经网络通过局部感受野来降低参数数量。一般动物对外界的感知是从局部再到全面。在图像中，一般相近的像素相关性较高，距离远的像素往往相关性较低。因而，每个神经元往往不需要对全局图像进行感知，单个神经元只通过局部感知，而深层多个神经元的综合便能得到对浅层的综合信息。

图 3：局部感受野



数据来源：公开资料 中信期货研究部

上面的图像例子中，对于 1000X1000 像素的图片，如果有一百万个隐层神经元，将会有 $10^{12}$ 个参数，而如果设置一个 10X10 的卷积核，即每个神经元只有 100 范围内的感受野，最终的参数为 $10^8$ 个。

## 2. 参数共享

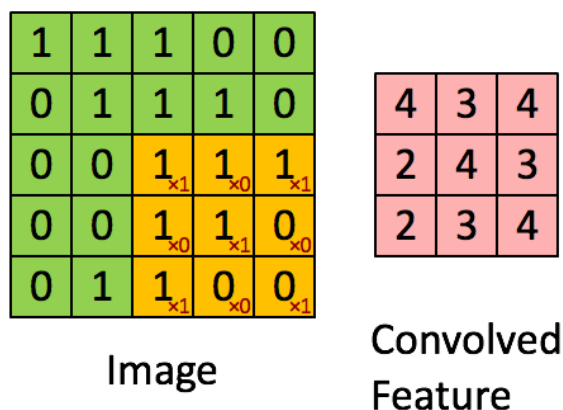
在以上的例子中， $10^8$ 个参数依然是一个很庞大的工程。如果所有的隐层神经元所对应的 10X10 的参数都是相等的，那此时的参数就变成了 100 个。此时，模型的复杂度则降低到了卷积核大小的程度。

参数共享的实际意义是如何呢？我们可以理解为卷积操作是对图像的特征提取。隐含的原理是：从小样本区域学习的特征作为探测器，去探查整个图片，将图像中符合条件的部分筛选出来。

当我们通过多个卷积核来对图像进行特征提取时，变实际上得到了多个局部信息，再下一层中，再将此类局部信息综合。下图展示了参数共享的计算。



图 4：参数共享



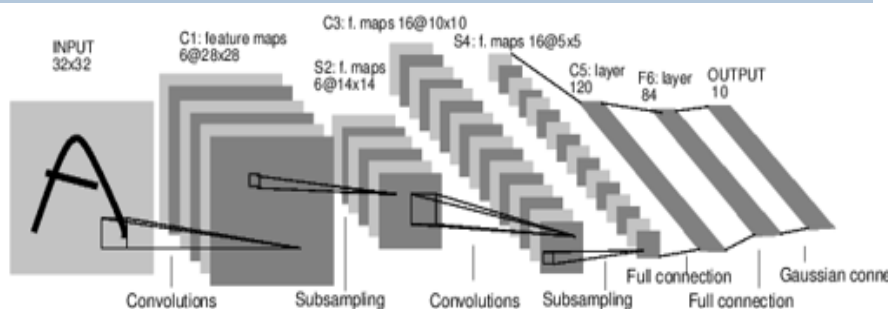
数据来源：公开资料 中信期货研究部

### 3. 下采样

理论上讲，通过提取的特征，我们便可以分类的操作。但此时已然面临着计算量的挑战。针对刚刚 1000X1000 的图像，假设采用的是 10X10 的卷积核，卷积操作后的单个特征图将是(1000-10+1)X(1000-10+1)的矩阵，依然是一个庞大的矩阵，若提取的特征数是 100，那数量级将再次提升。因此，通过下采样来计算一个图像特征中某些区域里的平均值或者最大值，用平均值或者最大值来代替某一区域，这样，参数的数量又得到了极大的缩减。

介绍完卷积神经网络的大部分特点，可以参考一下最初 LeNet-5 来理解整个过程。LeNet-5 包括输入层在内，总共 8 层。两个卷积层、两个下采样层以及两个全连接层。

图 5：LeNet-5



数据来源：公开资料 中信期货研究部

输入层为 32x32 的黑白图像，C1 层为卷积层，卷积核大小为 5x5，共有 6 个卷积核，相当于 6 个特征提取器。S2 层为下采样层，由 C1 层每组特征中 2x2 领域均值采样，因此 S2 层大小为 14x14。

后续 C3 和 S4 操作类似，不同的是，由于 S2 层特征图数量较少，C3 层加入一个连接表，用来指定 C3 层对 S2 层的映射关系。

图 6：LeNet-5 连接图

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	X				X	X	X			X	X	X	X		X	X
1	X	X				X	X	X			X	X	X	X		X
2	X	X	X				X	X	X			X		X	X	X
3			X	X	X		X	X	X	X			X		X	X
4				X	X	X		X	X	X	X		X	X		X
5					X	X	X		X	X	X	X		X	X	X

数据来源：公开资料 中信期货研究部

最后，通过 C5 再次卷积和 F6 的全连接网络形成输出。其中输出层由 10 个欧式径向基函数组成。



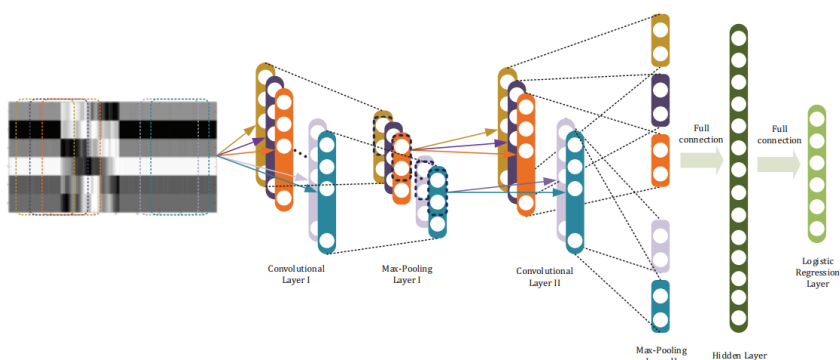
### 三、针对时间序列的卷积神经网络分类

以上，了解了卷积神经网络在图像识别中强大应用，那在金融时间序列上的应用如何展开呢？一类应用考虑到金融时间序列的时间性，将循环神经网络引入到金融时间序列预测，类似于语义识别，寻找最近数据中的特征，从而达到分类效果。鉴于卷积神经网络的强大特征提取能力，我们处理方式借用了图像识别的处理，将金融时间序列表征为图像。

图像在卷积神经网络中的表示实际上是一个二维/三维矩阵，对于黑白图像而言是二维矩阵，对于彩色图像而言，一个多通道的三维矩阵。那将金融时间序列表征为  $N \times T$  的矩阵， $N$  表示为多元时间序列的维度， $T$  表示时间序列的长度， $N \times T$  的矩阵即可表示为样本内图像。对于分类问题，我们需要对样本进行标识，在简单模型中，我们可以用时间序列的样本外涨跌进行标识。例如， $T+1 \sim T+5$  的涨跌幅大于 0 标为 A，否则为 B。因此将涨跌预测问题，转换到了卷积神经网络的分类问题。

然而对于卷积神经网络而言，网络的建筑设计是至关重要的一环。我们应用多元时间序列来对未来一周涨跌进行预测。卷积核的大小选择为  $N \times M$ ， $N$  是上面设定的多元时间序列的维度， $M$  是一个待定的参数。我们可以将样本内数据看作一个数据带，卷积核看作特征提取环，一般的图像处理中，特征图依然是一个二维矩阵，在我们的模型中，将形成一个一维的特征链。

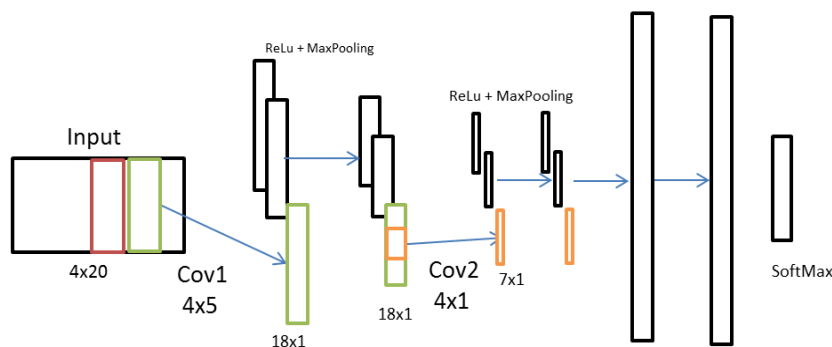
图 7：多元时间序列卷积网络结构 1



数据来源：公开资料 中信期货研究部

我们第一步选择 20 日的开高低收数据来作为输入数据，未来 5 日涨跌进行分类表示。网络结构借用 LeNet-5 的网络设计方式，采用两次卷积-激活-池化后再链接两层全连接网络。

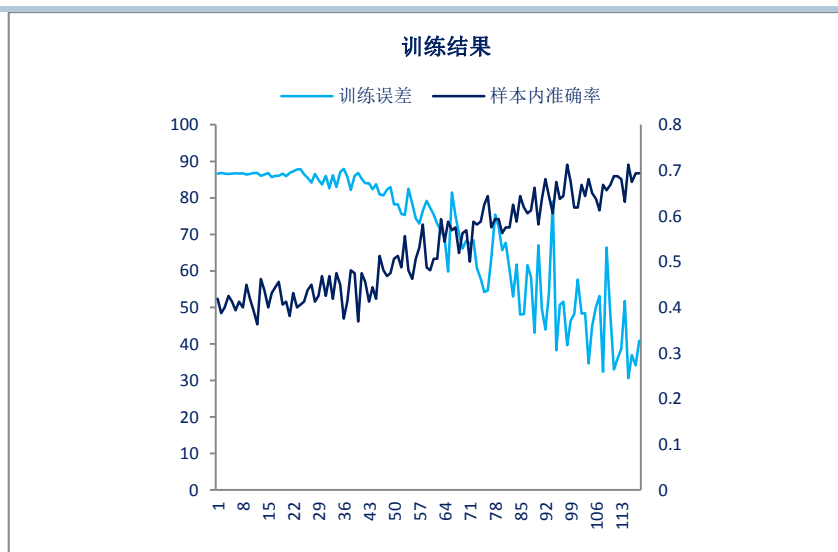
图 8：多元时间序列卷积网络结构 2



数据来源：中信期货研究部

对 33 个期货合约品种的拼接主力合约数据进行切分成 10000 个样本，并随机分成 8000 个样本内数据供网络训练以及 2000 个数据供检测有效性。

图 9：多元时间序列卷积网络训练结果



数据来源：中信期货研究部

图表中迭代次数 1 代表 100 次迭代，样本内准确率随着迭代次数的增加而增加，训练误差也在逐步降低。样本外检测准确率为 55.9%。

后续工作将从多方面提升模型的有效性。一方面，可以考虑加入更加能反应后期涨跌的数据；另一方面，修改网络架构，调整一部分参数；再者，可以考虑将模型应用到策略层面，设置相应的止损止盈。

## 免责声明

除非另有说明，本报告的著作权属中信期货有限公司。未经中信期货有限公司书面授权，任何人不得更改或以任何方式发送、复制或传播此报告的全部或部分材料、内容。除非另有说明，此报告中使用的所有商标、服务标记及标记均为中信期货有限公司的商标、服务标记及标记。中信期货有限公司不会故意或有针对性的将此报告提供给对研究报告传播有任何限制或有可能导致中信期货有限公司违法的任何国家、地区或其它法律管辖区域。

此报告所载的全部内容仅作参考之用。此报告的内容不构成对任何人的投资建议，且中信期货有限公司不因接收人收到此报告而视其为客户。

中信期货有限公司认为此报告所载资料的来源和观点的出处客观可靠，但中信期货有限公司不担保其准确性或完整性。中信期货有限公司不对因使用此报告及所载材料而造成的损失承担任何责任。此报告不应取代个人的独立判断。中信期货有限公司可提供与本报告所载资料不一致或有不同结论的报告。本报告和上述报告仅反映编写人的不同设想、见解及分析方法。本报告所载的观点并不代表中信期货有限公司或任何其附属或联营公司的立场。

此报告中所指的投资及服务可能不适合阁下，我们建议阁下如有任何疑问应咨询独立投资顾问。此报告不构成投资、法律、会计或税务建议，且不担保任何投资及策略适合阁下。此报告并不构成给予阁下的私人咨询建议。

中信期货有限公司2018版权所有并保留一切权利。

## 深圳总部

地址：深圳市福田区中心三路8号卓越时代广场（二期）北座13层1301-1305、14层

邮编：518048

电话：400-990-8826

传真：(0755)83241191

网址：<http://www.citicsf.com>