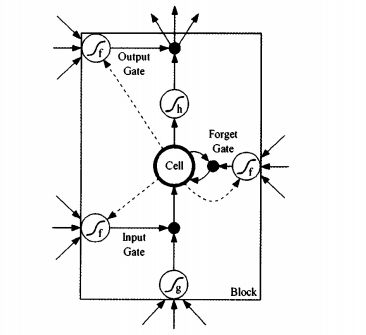
## 1.背景

在复杂的股票市场环境中，神经网络算法在股票预测中已经得到了广泛使用，这是由于其自身具有较好的学习性能和高度的模拟能力，相对于传统的经济计量学方法，神经网络在金融时间序列预测方面更具优势。  
近年来，国内外学者对于在股票市场的神经网络预测问题做了很多的研究工作。Shapiro…将神经网络、遗传算法和粗糙集组合成集成算法对股票市场价格趋势进行综合预测，但是文中没有作对比验证，而且模型中没有考虑到金融时间序列的依赖关系，预测结果并不客观；Ozbayoglu等旧1通过对比人工神经网络和贝叶斯方法在金融市场的预测性能，发现这两种算法均有效，但是人工神经网络的预测效果更佳；Bildirici等"1将BP神经网络与条件异方差模型相结合，对1987年到2008年的伊斯坦布尔市场的股票数据做训练及预测，实证表明，这种结合模型的预测精度更加可靠，但是面对海量数据，此模型提取特征比较困难；Hammda等H1采用多层BP神经网络对约旦股票市场的指数价格的趋势做预测，研究发现多层BP神经网络具有预测精度高、泛化能力强的优点，但文中没有解决BP神经网络容易陷入局部最小的问题；，得出该组合模型预测精度较高。考虑到股票市场的数据繁杂及其具有长期依赖性等问题以上模型在时间序列预测上还存在一些问题，而递归神经网络RNN(Recurrent Neural Networks)可以改善这些问题。孙瑞奇¨4 J将BP神经网络、RNN和LSTM模型分别对美股和上证指数做预测，实验得出LSTM模型的预测精度最高；金雪军等¨纠采用深度LSTM模型研究美国的扩张性货币政策对中国膨胀的影响。综上可知，利用LSTM模型对于股票价格预测的研究极少。  
本文针对RNN存在的梯度问题和神经网络的过拟合问题，提出LSTM模型对股票指数进行预测。通过修改LSTM模型的目标函数，从而提高了LSTM模型对金融时间序列预测性能。LSTM模型误差低，对股票指数有更佳的预测效果。

## 2.模型构建

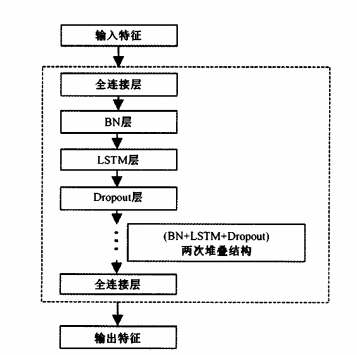
## 2.1LSTM神经元结构

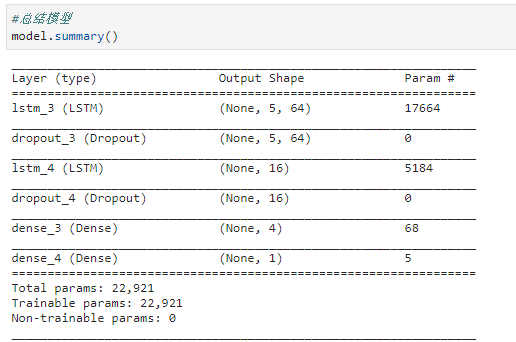
LSTM神经网络最早由Hochreiter和Schmidhuber(1997)提出，由于能更好地发现长期依赖关系而被广泛用于处理序列信息，如语音识别、机器翻译等。LSTM每个神经元的结构见图l(Grave，2012)013]，其内部包括一个记忆储存(Cell)和三个门控(Gates)设置，Cell记录神经元状态，输人门(Input Gate)和输出门(Output Gate)用来接收、输出参数和修正参数，遗忘门(Forget Gate)用来控制上一单元状态的被遗忘程度。



## 2.2 LSTM神经网络的构建

深层神经网络存在两种可能影响模型训练效果的问题，一是梯度消失导致神经网络难以收敛，二是过拟合导致测试集失效。研究表明，Batch．Normalization(BN)能有效解决梯度消失问题，Dropout技术通过阻止神经元共适应能够缓解过拟合问题。因此，本文在设计LSTM神经网络时特别添加了Dropout层以优化神经网络结构。具体来说，本文LSTM神经网络的主体结构包括2层LSTM神经层和一层全连接层，第一个LSTM层包括64个节点，第二个LSTM层16个节点，其后加人Dropout层并将失活概率设为0．2。本文构建的深层LSTM神经网络计算图结构见图2，虚线方框内表示神经网络结构。





## 2.3训练方法和损失函数

本文采用Mini—Batch方法训练LSTM网络。本文目标是预测股指未来收盘价，故选取均方误差(Mean Square Error，MsE)预测损失。优化器方面，本文采用Adam优化器(Adaptive MomentEstimation，适应性矩估计)①进行优化训练。Adam优化器由Kingma和Sa(2015)…1提出，是目前最常用的算法(Sabour等2017‘151；Wang等2018‘怕1)。与其他自适应学习率算法相比，Adam算法收敛速度更快、学习效果更为有效。本文基于Python语言环境，并以PyTorch作为深度学习框架进行训练及预测。最后求训练集和测试集的MAE平均绝对误差/MSE最小均方误差/MAPE百分比误差绝对值的平均值。





MAPE:范围[0,+∞)，MAPE 为0%表示完美模型，MAPE 大于 100 %则表示劣质模型。MAPE 的值越小，说明预测模型拥有更好的精确度.



## 3.实验

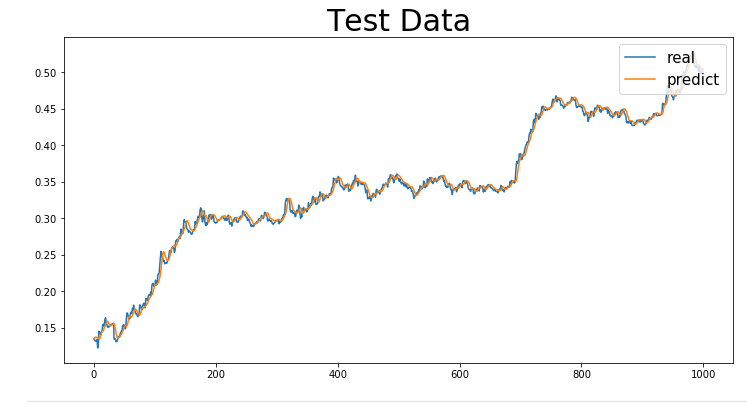
本文针对6322组数据，即包括股票的开盘价格，收盘价，最低价和最高价四种数据。将收盘价作为输出变量，其他5个指标作为模型的输入变量。

在模型拟合数据之前，使用Python3.7中的sklearn函数库对数据集做转换处理。首先，为了使时间序列数据趋于稳定，对数据做滞后一次差分处理；然后将预测收盘价转化为有监督学习问题，也就是将数据组合成输入和输出模式，将上次时间步长的观测值用作预测当前时间步长观测值的输入；最后对原始数据进行归一化处理，将其转换到[0，1]之间。

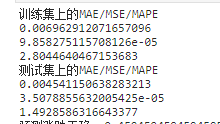
将处理后的数据按顺序分成训练集和测试集，其中前5422个数据做训练集，后1000个数据做测试集，训练模型50次epoch，在数据测试完成之后，将预测值做反转换处理，以便计算预测性能指标。

64个的LSTM——遗忘层——16个LSTM——遗忘层——4层全连接层





MAE平均绝对误差/MSE最小均方误差/MAPE百分比误差绝对值的平均值



## 4.总结

经过这次实验，我深刻学习了LSTM在股票预测中的运用，了解了LSTM层，dropout层，浅层神经网络LSTM模型具有较强的泛化能力，其预测效果也更加优于传统的BP神经网络和RNN。

LSTM的主要优点概括如下：

1)因为LSTM模型具有独特的网络结构，所以它可以更好地学习过去的股票数据，并且能够找出时间序列之间的关系，还能利用选择性记忆的功能，对股票价格的固有规律进行更进一步挖掘，从而具有较好的预测效果。

2)通过利用正则化参数的先验信息，对LSTM模型的目标函数做出修改，使其在对股指数据的应用中具有更好的泛化能力。

当面对复杂多变的股票市场，对股指预测的要求将会更加便捷、更加精确。而在股指预测中应用人工智能算法，将是解决此问题的有效途径。当然LSTM还可以和很多优化算法结合，实现优化效率的提高。