# 1. 项目基本信息

|  |  |  |  |
| --- | --- | --- | --- |
| 项目名称 | 股票市场预测研究 | 当前阶段 | 对比实验 |
| 开始日期 | 2016.5.15 | 结束日期 | 2016.6.20 |
| 小组成员： | 黄姗、靳晓乐、马骁 | | |

# 2. 项目综述

通过分析股票技术关键指标数据，建立相应模型，实现股票第二天收盘价预测。

# 3. 项目执行情况

## 3.1 数据处理部分

### 3.1.1 原始数据来源：数据来自三个证券交易所

上证000001：从1990年12月19日到2016年6月01日

深证399001：从1991年04月03日到2016年6月01日

沪深399300：从2002年01月04日到2016年6月01日

原始数据结构如下（前十条）：



原始记录包括日期、股票代码、名称、收盘价、最高价、最低价、开盘价等。过滤后留下收盘价、最高价、最低价、开盘价和成交量。

### 3.1.2 股票技术关键指标数据的获取：

候选技术指标：

MACD WMS RSI OBV M30 M10 K指标 D指标 DIF DEA 收盘价

实验确定关键技术指标

## 3.2 神经网络拟合模型

### 3.2.1 神经网络配置

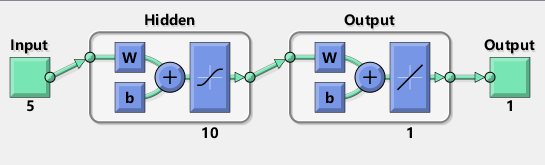
根据选择的技术指标的不同，有不同的输入神经元个数，从5个到11个不等，输出节点为1个，调整隐含层结点进行实验。

激活函数隐含层选用逻辑回归，输出层选择线性回归；

性能评价选用MSE以及准确率；

指定训练集为70%，验证集和测试集均为15%

由以上配置，得到神经网络如下：



如图所示是输入神经元个数为5，隐含层个数为10时的神经网络。

### 3.2.2实验结果

1、输入指标：

MACD WMS RSI OBV 收盘价 M30 M10 K指标 D指标 DIF DEA

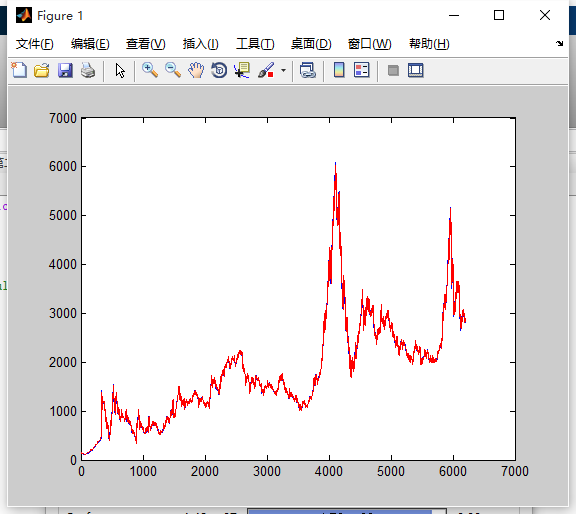
2、输出：

第二天收盘价

3、隐含层个数：

10

4、拟合结果：



0.01以内的误差准确率：0.495968

0.02以内的误差准确率：0.737258

0.05以内的误差准确率：0.930968

涨跌准确率：0.886485

1、输入指标：

MACD WMS RSI OBV 收盘价 M30 M10 K指标 D指标 DIF DEA

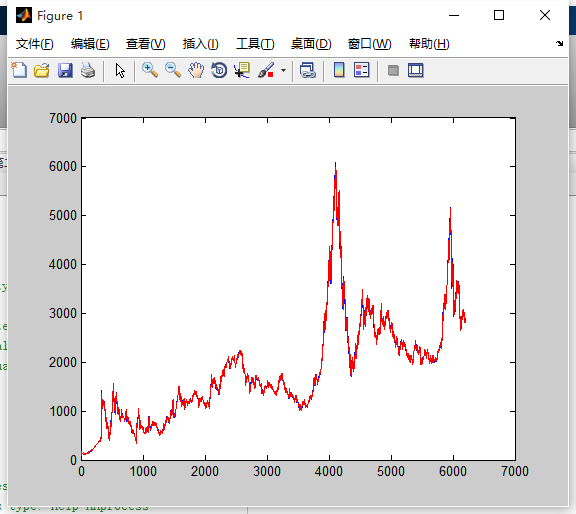
2、输出：

第二天收盘价

3、隐含层个数：

23

4、拟合结果：



0.01以内的误差准确率：0.499355

0.02以内的误差准确率：0.747903

0.05以内的误差准确率：0.948065

涨跌准确率：0.888422

1、输入指标：

MACD WMS RSI OBV 收盘价 M30 M10 K指标 D指标 DIF DEA

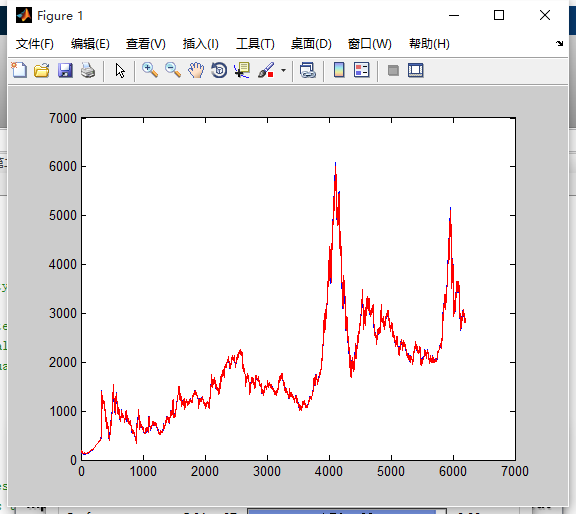
2、输出：

第二天收盘价

3、隐含层个数：

3

4、拟合结果：



0.01以内的误差准确率：0.525484

0.02以内的误差准确率：0.763226

0.05以内的误差准确率：0.938065

涨跌准确率：0.880995

1、输入指标：

MACD WMS RSI OBV 收盘价 M30 M10 K指标 D指标 DIF DEA

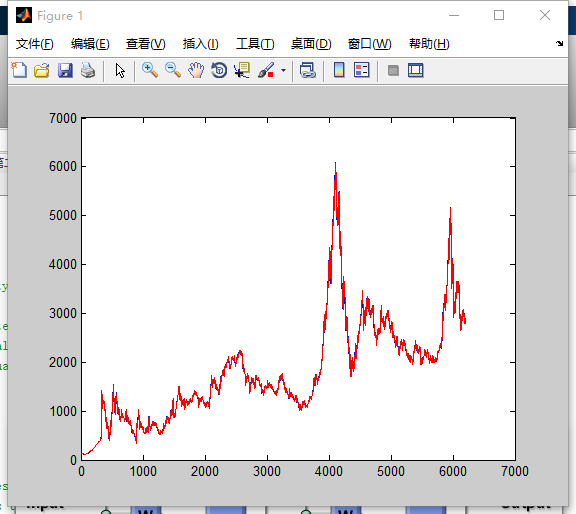
2、输出：

第二天收盘价

3、隐含层个数：

4

4、拟合结果：



0.01以内的误差准确率：0.533065

0.02以内的误差准确率：0.770806

0.05以内的误差准确率：0.958710

涨跌准确率：0.889714

1、输入指标：

MACD WMS RSI OBV 收盘价 M30 M10 K指标 D指标 DIF DEA

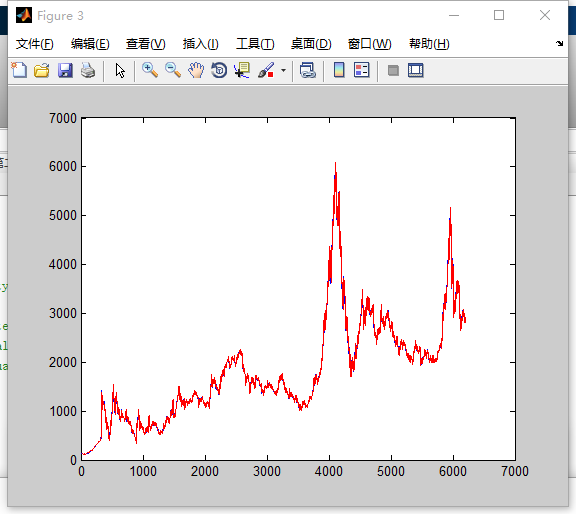
2、输出：

第二天收盘价

3、隐含层个数：

5

4、拟合结果：



0.01以内的误差准确率：0.517419

0.02以内的误差准确率：0.768710

0.05以内的误差准确率：0.955968

涨跌准确率：0.891975

1、输入指标：

MACD WMS RSI OBV 收盘价

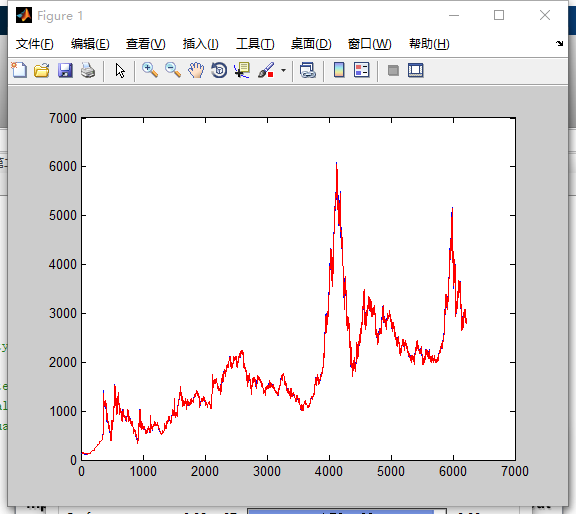
2、输出：

第二天收盘价

3、隐含层个数：

10

4、拟合结果：



0.01以内的误差准确率：0.512289

0.02以内的误差准确率：0.764980

0.05以内的误差准确率：0.945703

涨跌准确率：0.885815

## 3.3 神经网络时间序列模型Nonlinear Autoregressive (NAR)

### 3.3.1 神经网络配置

由于采用时间序列模型，故输出层节点个数为1，而选择滞后值d=5，即输入层节点个数为5，隐含层节点个数为10；

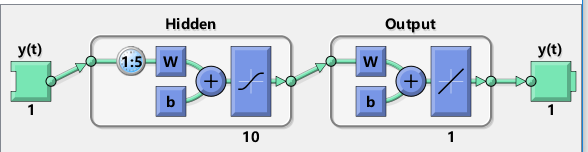
激活函数隐含层选用逻辑回归，输出层选择线性回归；

性能评价选用MSE；

训练方法选用Levenberg-marquardt，一种非线性最小二乘算法；

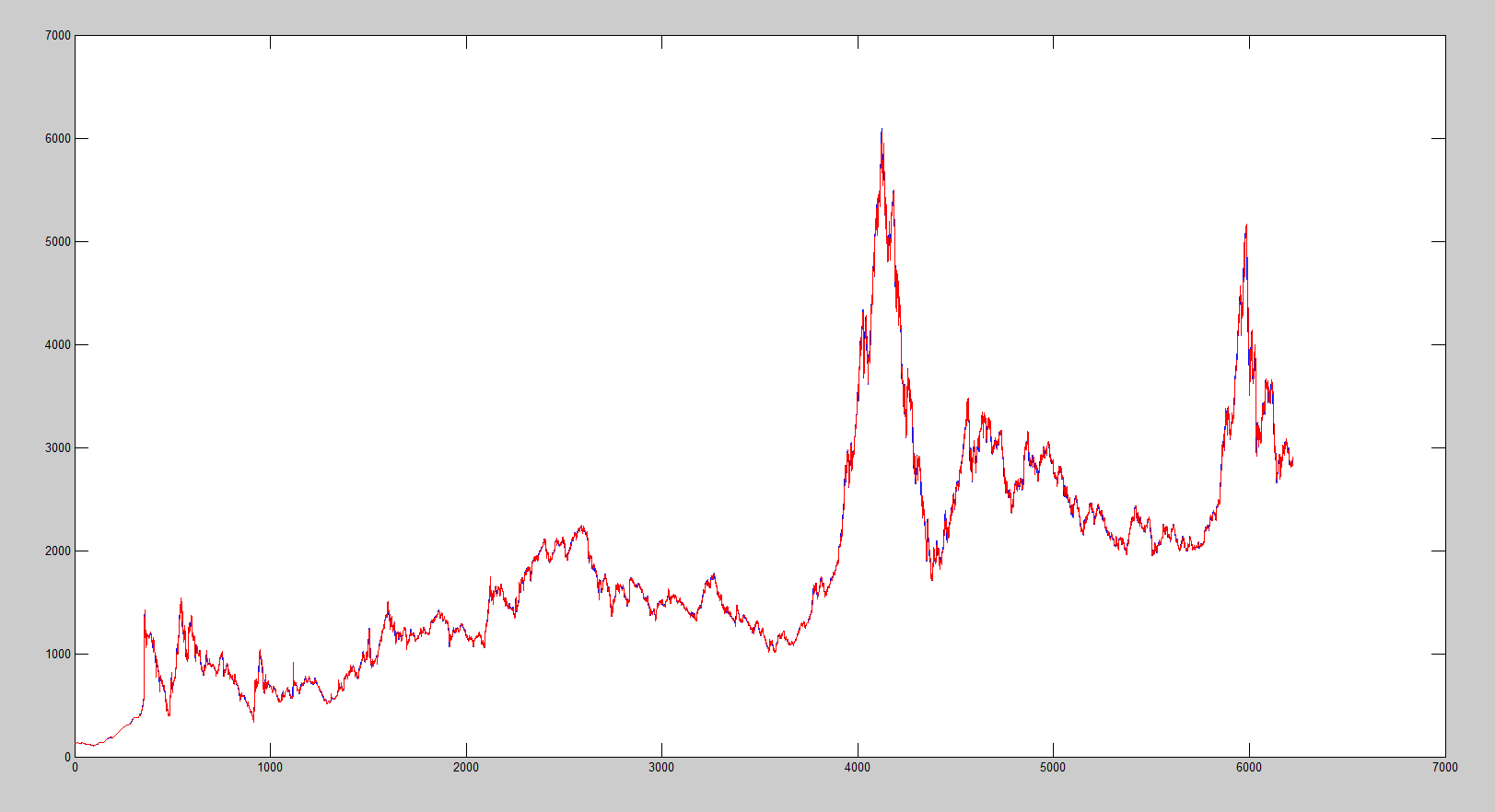
指定训练集为80%，验证集和测试集均为10%

由以上选择配置，得到神经网络如下：



### 3.3.2 数据拟合结果

（蓝色表示实际值，红色表示预测值）



涨跌幅在1%以内的为54.3%左右，涨跌幅在2%以内的为78.2%左右，涨跌幅在5%以内的为96.2%左右，有一定的参考价值。