**结论一：在尝试的所有模型中，Lstm在2层64个节点，以1991年到今的涨跌幅作为训练数据的训练模型效果最好。CNN模型在3层卷积层，分别含有16、14、8个卷积核，以及一层40个节点的全连接层，以2014年涨跌作为训练数据的训练模型效果最好。CNN模型总体上优于Lstm模型。**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Layer | input | Max iterations | Mean | Variance |
| Lstm | 2 \* 64 | 1991股价差 | 10000 | 0.4772 | 0.00175216 |
| Lstm | 3 \* 64 | 1991股价差 | 20000 | 0.464 | 0.001668 |
| Lstm | 1 \* 128 | 1991涨跌 | 20000 | 0.486 | 0.001964 |
| Lstm | 3 \* 16 | 1991涨跌幅 | 30000 | 0.4982 | 0.00211876 |
| Lstm | 2 \* 64 | 1991涨跌幅 | 20000 | 0.5262 | 0.00172756 |
| Lstm | 1 \* 16 | 2012股价差 | 10000 | 0.4962 | 0.00314356 |
| Lstm | 2 \* 64 | 2012股价差 | 10000 | 0.477 | 0.002273 |
| Lstm | 1 \* 16 | 2012涨跌 | 20000 | 0.4814 | 0.00182004 |
| Lstm | 3 \* 64 | 2012涨跌 | 10000 | 0.4756 | 0.00224436 |
| Lstm | 3 \* 16 | 2012涨跌幅 | 20000 | 0.4936 | 0.00300304 |
| Lstm | 3 \* 64 | 2012涨跌幅 | 5000 | 0.4778 | 0.00234116 |
| Lstm | 1 \* 16 | 2014股价差 | 20000 | 0.4788 | 0.00199856 |
| Lstm | 3 \* 16 | 2014股价差 | 20000 | 0.4764 | 0.00225904 |
| Lstm | 1 \* 16 | 2014涨跌 | 10000 | 0.4796 | 0.00183984 |
| Lstm | 3 \* 16 | 2014涨跌 | 20000 | 0.4926 | 0.00227124 |
| Lstm | 1 \* 16 | 2014涨跌幅 | 10000 | 0.4796 | 0.00183984 |
| Lstm | 3 \* 16 | 2014涨跌幅 | 10000 | 0.4834 | 0.00278644 |
| Cnn | 15\*8\*8 | 2012涨跌 | 20000 | 0.526 | 0.0011116 |
| Cnn | 16\*14\*8 | 2012股价差 | 20000 | 0.5036 | 0.00215104 |
| Cnn | 16\*14\*8 | 2012涨跌幅 | 20000 | 0.4938 | 0.00180756 |
| Cnn | 15\*8\*8 | 2012股价 | 3000 | 0.5038 | 0.00165556 |
| Cnn | 16\*14\*8 | 2014涨跌 | 20000 | 0.5316 | 0.00129744 |
| Cnn | 16\*14\*8 | 2014股价差 | 20000 | 0.4954 | 0.00223284 |
| Cnn | 16\*14\*8 | 2014涨跌幅 | 20000 | 0.5024 | 0.00163424 |
| Cnn | 16\*14\*8 | 2014股价 | 20000 | 0.4968 | 0.00098176 |
| Cnn | 16\*14\*8 | 涨跌 | 10000 | 0.5092 | 0.0025936 |
| Cnn | 16\*14\*8 | 股价差 | 20000 | 0.5208 | 0.00245936 |
| Cnn | 16\*14\*8 | 涨跌幅 | 20000 | 0.4968 | 0.00202176 |

**结论二：由于随机初始化变量的和训练过程中随机批量梯度下降带来的不确定性，需要多次训练模型选取较好模型结果。CNN效果最好的模型中，50次的训练结果如下：**

**数据集中在0.56，多次训练选取较好模型的情况下，能达到0.56以上。**

|  |  |
| --- | --- |
| 0.46 | 2 |
| 0.47 | 1 |
| 0.49 | 3 |
| 0.50 | 8 |
| 0.51 | 7 |
| 0.52 | 3 |
| 0.54 | 6 |
| 0.55 | 2 |
| 0.56 | 12 |
| 0.57 | 3 |
| 0.58 | 1 |
| 0.60 | 1 |
| 0.64 | 1 |

**结论三：训练过程中发现，训练集合（历史数据）与测试数据（最后100天作为测试数据）拟合度不高，相似性较低。**

由图可见，当训练次数增加时，训练的损失函数值不断下降，但是测试数据的损失函数却不降反升的情况。

原因可能是如下：

1. 训练集与测试集耦合度不高。
2. 训练集噪声太大无法从训练集中提取有效的特征。

**结论四：训练结果存在过拟合，训练过程如图所示：**

**结论五：加入交易量因子后，效果没有得到提升。**

|  |  |  |  |
| --- | --- | --- | --- |
| Model | input | Mean | Variance |
| Cnn | 2014涨跌 | 0.5316 | 0.00129744 |
| Cnn | 2014涨跌，交易量 | 0.5194 | 0.00210564 |

|  |  |  |  |
| --- | --- | --- | --- |
| Model | input | Mean | Variance |
| Cnn | 2014涨跌 | 0.5316 | 0.00129744 |

|  |  |
| --- | --- |
| 0.46 | 2 |
| 0.47 | 1 |
| 0.49 | 3 |
| 0.50 | 8 |
| 0.51 | 7 |
| 0.52 | 3 |
| 0.54 | 6 |
| 0.55 | 2 |
| 0.56 | 12 |
| 0.57 | 3 |
| 0.58 | 1 |
| 0.60 | 1 |
| 0.64 | 1 |

|  |  |  |  |
| --- | --- | --- | --- |
| Model | input | Mean | Variance |
| Cnn | 2014涨跌，交易量 | 0.5194 | 0.00210564 |

|  |  |
| --- | --- |
| 0.41 | 1 |
| 0.44 | 4 |
| 0.46 | 5 |
| 0.47 | 5 |
| 0.49 | 1 |
| 0.5 | 6 |
| 0.51 | 6 |
| 0.52 | 3 |
| 0.54 | 2 |
| 0.55 | 5 |
| 0.56 | 9 |
| 0.57 | 3 |
| 0.58 | 1 |
| 0.62 | 1 |

猜想：测试集和训练集相似性高，在训练精度高的情况下测试集的准确度也会高。

缺乏对应的测试集和训练集，无法验证

猜想：训练集合的损失函数如果能够随着训练而降低，说明训练集和测试集的相似性高。

缺乏数据，无法验证