# 一．背景：

受益于过去数十年的经济增长,中国的私营及公共部门积累了数量惊人的巨大财富。经过调查表明，中国已成为亚太地区排名第二的财富管理市场。 相应地，中国的私人和企业持续为财富的保值和增长寻求良好的投资机会。贯穿中国大陆、香港和澳门市场，为数众多的国内国际私人银行、财富管理公司、投资基金和资产管理公司建立和发展起 来。在海外直接投资不断增长和私募基金、风险投资基金实现跨越式发展的浪潮中，中国的投资管理行业在快速增长下也面临着日益复杂的市场环境和全球竞争。近年来，创业公司面对更加激烈的市场环境，如何从创业公司的墓碑中，总结获取投资风险建议，对于创业者来说成为了一个重要的问题

# 二．介绍：

本文使用了近年来倒闭破产的2000家初创公司列表的历史数据，挖掘数据表中的96组特征与参数，其中包括人力资源，资金流等多个纬度下的数据特征，利用机器学习中的BP神经网络算法。训练了一个公司当前数据是否导致破产的的二分类预测模型，算法模型能根据一个公司数据的多个纬度，判断初创公司是否有破产风险，给出风险评估，训练的模型，达到了97%的准确度。并基于历史数据。从几个关键变量出发，分析这些变量对破产的影响，给出建议

# 三．风险预测：

能够准确预测一个公司是否有破产的危机，可以对于一个企业来说是至关重要的，近年来倒闭破产的2000家初创公司列表的历史数据，其中包含了96组数据特征，与一组是否破产的标签

将数据划分为训练集与测试集，对数据表中的数据进行归一化处理，能够消除不同数据项的数值范围干扰，使用机器学习中的多层感知机BP神经网络作为算法模型，将归一化的训练数据传入模型进行训练，根据96个数据特征，最终输出是否会破产。训练完成后，使用测试数据在训练好的模型上进行测试，算法在测试集上达到了97%的准确度，说明预测模型能有效根据公司当前的数据进行风险判断

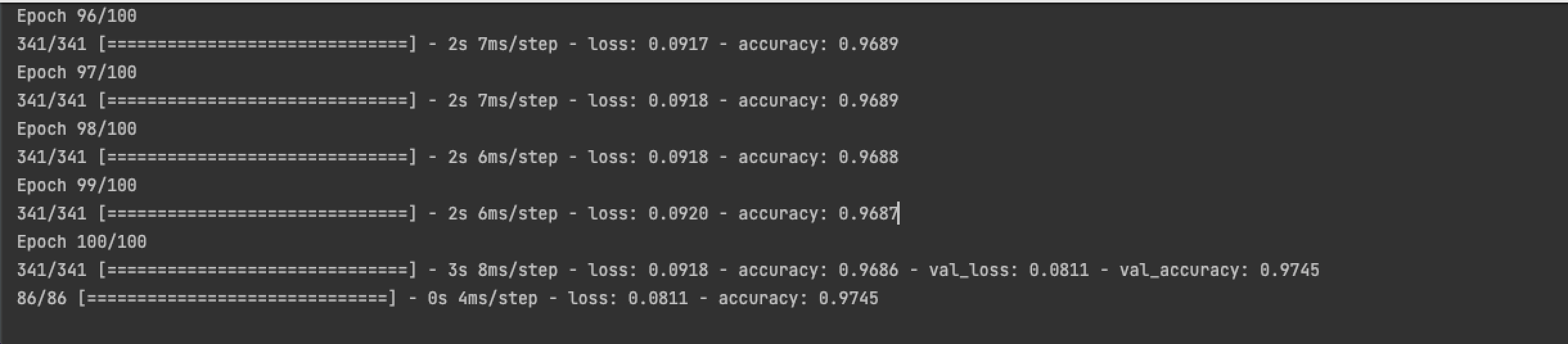


图1. 预测模型精度

# 数据分析

在得到预测模型后，修改对模型的输入，发现当改变几个关键变量后，会决定性的改变最终是否会破产。以下将数据分析角度与破产的关系，并给出调控建议

## 4.1 花销投入

在数据中列表中两个主要的花销是研究发展花销和运营花销，是连续对分类的问题，可利用Matplotlib生成2D散点图，其中红色代表破产公司，蓝色代表未破产的公司，从散点图中可看出，是否破产与运营花销强相关，破产公司主要集中在运营花销较少的区域，而于0.4-0.8运营花销率/0.2-0.6研究发展花销区域内的公司，破产概率较低

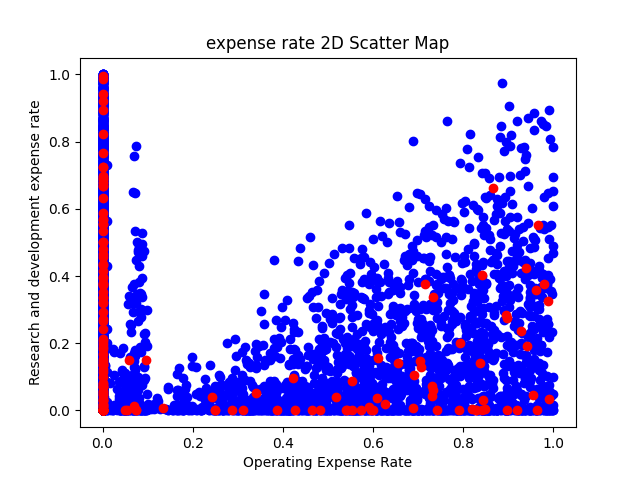


图2. 研究发展花销与运营花销散点图

## 4.2 周转率

周转率主要分为流动资产周转率与营运资金周转率，从散点图可看出，流动资产周转率更高时，通过对该指标的对比分析，可以促进企业加强内部管理，充分利用流动资产，如调动暂时闲置的货币资金用于短期的投资创造收益等。还可以促进企业采取措施扩大销售，提高流动资产的综合使用率。

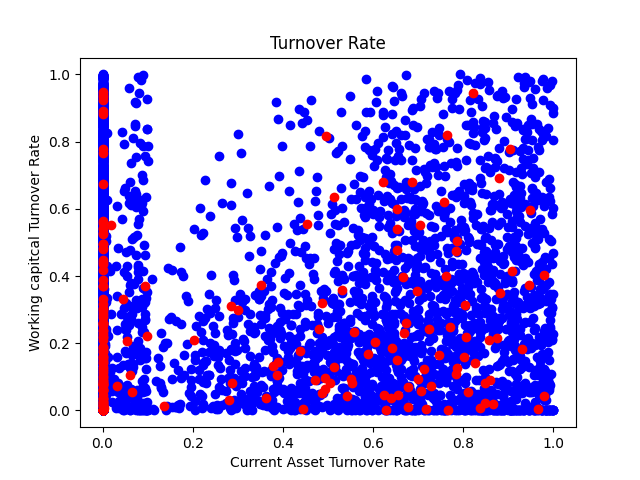


图3. 流动资产周转率与营运资金周转率散点图

## 4.3 ROA

以下给出了一个可视化ROA指标，当花销需要资产收益率是业界应用最为广泛的衡量银行盈利能力的指标之一，该指标越高，表明企业资产利用效果越好，说明企业在增加收入和节约资金使用等方面取得了良好的效果

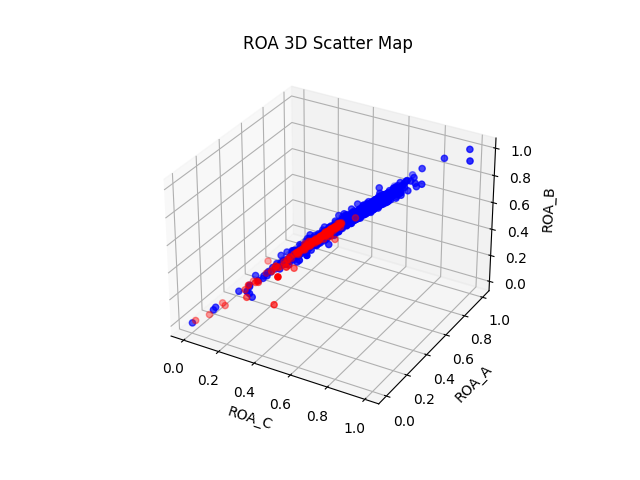


图4. ROA散点图

# 五. 结论

从花销角度，研究与开发花销投入产出比过低，投入率应该适中，不应过高，更多投入营运花销，这样对于公司稳定是有帮助的，另外，从资金流转率来看，需要提升资金流转速度，能够更快的在短期内创造更多收益，从财务的角度来看也能提升公司的ROA指标，当ROA指标稳定时，公司就不会有破产风险