智能证券投资学课程总结报告

(2020 秋)

学号: 20S051009

姓名: 陶靖枞

学院: 计算机学院

专业: 计算机科学与技术

日期: 2020-11-04

目录

应用篇总结	3
2.1 算法内容	
2.2 研究设想	7
总结,感悟,关于课程或平台的建议	7
	2.2 研究设想

1. 应用篇总结

1.1 课程内容总结

本课程一共八个章节. 第一章介绍了什么是智能证券投资以及如何评价投资 效果,理性投资自动化和自动投资理性化是本课程主要教授的两个方向。 二章则介绍了证券投资相关的基础知识,主要包括证券相关知识(股票、债券和 现金等价物及其相关的一些概念与指标)以及最基本的证券买卖操作。 章的内容是海天 4S 方法中的"势",即宏观判势。要对于当前在熊市还是牛市 还是震荡有一个基本的判断。在这一章中首先介绍了股市债市以及现金等价物等 相关市场的趋势有哪些类型及相应的特征,其次详细介绍了如何进行股市判势及 每种趋势下可以采取的不同策略,最后介绍了如何对自己的投资行为和判势能力 进行量化评价。 第四章的内容是海天 4S 方法中的"施",即具体实施。具体 而言是如何选择购买股票, 讲述了一些对于股票表现的评价指标. 包括股票的 个体表现以及股票所在行业的群体表现,告诉我们如何择行择股 第五章的内 容是以史为鉴。主要讲述了多种评价指标. 我们需要一个指标来对自己过去一段 时间的投资行为及相应的投资效果进行评价。最基本的是收益率,在本章一开始 介绍了适用于不同情景与场合下的收益率计算方法,之后则给出了一种比较不同 投资策略的量化指标,最后给出了一种归因方法.有了这些知识,我们就能够对于 自己过去的投资表现有一个更直观的评价, 来调整自己的投资策略

第六章的内容是海天 4S 方法中的"师",即悟道出师。积累了一定的投资经验以后,投资者应当能够给出自己对于市场的独特见解,并总结自己的投资经验与投资偏好,从而真正了解自己。 第七章与第八章是投资自动化的相关介绍与总结实践。给出了一个自动投资智能体的实现思路,将海天 4S 方法的四个阶段分为感知量化、任务执行、评测诊断和学习进化四个功能。我们可以通过指定投资策略的方式让投资机器人来完成投资过程。此外,老师还为我们补充了遗传算法在智能证券投资领域的应用方面的内容。

1.2 比赛内容总结

1.2.1 盘感训练比赛总结

盘感训练在给定的一段交易日内,通过对股票 K 线形态的主观分析,来决定股票的买卖操作。直观来看,根据股票的形态走势进行交易的例子在现实中有很多,最典型的比如"追涨杀跌"(跟随涨股买入,跟随跌股卖出)或者"高抛低吸"(股价上涨卖出,股价跌落买入)或者无操作(无法判断股价走势).

盘感训练能提高投资者对股票走势的熟悉程度,增加投资者对股票未来走势判断的准确性,提高投资者的看盘能力。通过测试排名,投资者可以更好地总结经验教训,提高投资水平。

1.2.2 模拟投资比赛总结

模拟投资比赛是在课程配套的海知平台上进行的,整个模拟投资比赛使用的是平台提供的虚拟资金,而除此之外全部与真实的股市情况一致。实践出真知,通过参与这场模拟投资比赛,可以在实践中运用课程讲授的相关知识。

图 1.1 是我的模拟投资收益状况

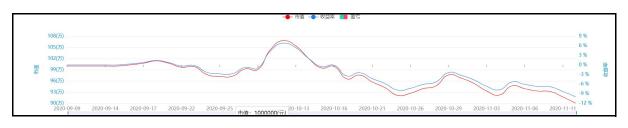


图 1.1

1.2.3 自动投资比赛总结

自动投资比赛同样是在海知模拟平台上进行,与模拟投资比赛的唯一区别是,自动投资比赛中是由自动投资智能体机器人代替人来进行投资决策,通过模拟平台提供的策略定制功能,只需将自己的投资策略告诉机器人就可以让它代替我们进行投资操作,也即实现了理性投资自动化。要实现自动投资,制定自己的投资策略很关键。在进行自动投资实验的过程中经过三次迭代实验,最后得到了相对来说最优的投资策略如表格 1-1 所示:

表 1-1

项目	值
选股范围	沪深 300 成分股
股债金比例	股票 100%
调仓周期	5 天
调仓比例	50%
单股持仓最大天数	4 天
买入规则	涨幅>5%
卖出规则	跌幅>2%
风险控制	单股止盈 10%; 单股止损 4%

制定好策略后,将该策略在之前模拟投资的同一时段进行运行,回测得到的结果如图表 1-3 所示,总收益达到了近 24 万元。

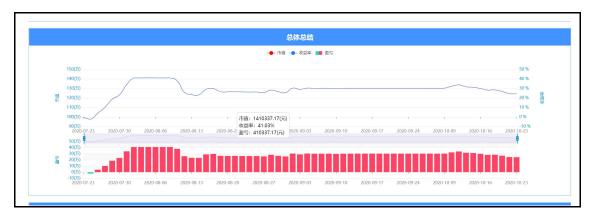


图 1-3

可以看到自动投资机器人在相同时间段内的表现比我自己投资收益要高出许多,但也可能是由于这个策略仅仅是在选定的时间范围内表现较好,我在自动投资比赛中开启该策略,现在比赛结束时投资收益为-5000,如图 1-4 所示。

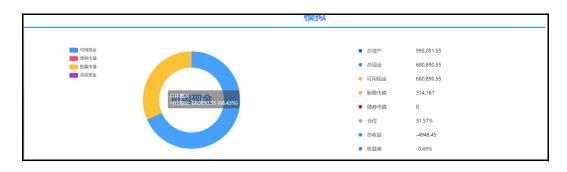


图 1-4

1.3 投资经历与平台实训总结

投资经历: 9月18日,中国长城一路上涨,带动个人市值的提高,同时卓翼科技也有不小的上涨。9月24日,所有股票均下跌,择时能力较差,没有及时抛出有下跌趋势的股票,导致亏损严重。9月25日,抛出大量股票,亏损近2万元,卓翼科技亏损近万元,是这一个月投资中最大的亏损。10月9日,一直持有的宁德时代终于回暖,适时抛出,获取近3万元利润,这是最成功的一次战斗。直到10月14日,股票一直保持稳步上涨阶段。10.25日,持有的股票开始全线下跌,由于没有及时调仓,导致损失惨重,到10.30日已经亏损10%.

平台实训总结:

实时数据模块提供了沪深市场所有股票与债券的详细数据,同时还提供了多因子选股、智能优选等结合了计算机相关技术的实用功能,方便我们在投资时选择股票。在最初进行投资时,由于没有相关的证券投资经验,在选择股票时遇到了许

多不便,有了实时数据模块,学完证券投资基础知识以后就可以利用它来筛选符 合自己要求的股票。

评测诊断模块则是用来辅助投资者对自身的投资行为进行反思的功能。对应海天4S 方法中的"以史为鉴"。如果由自己进行人工评价,既费时又费力,评测诊断模块提供了许多实用功能如总体分析、个体分析、业绩归因和用户画像,这些功能使用了课程中讲授的评测诊断量化方法,通过计算机自动针对过去的投资历史进行数据分析,以打分、图表等形式给投资者提供过往投资行为的分析,帮助投资者不断改进自己的投资行为。在我过去两个月的投资过程中,多次用到这个模块,对应着我在比赛中的几次失误,相应的择股择行等的评分也降低,业绩归因的功能直接帮我找到了我所不足的部分,可以有针对性地改进。用户画像的功能给我的标签非常准确,我比较偏好于短期投资,并且经常性的满仓买入股票,"短线投资"和"激进型"两个标签描述的非常贴切。

伯乐相马模块也是一种辅助选股的工具。初听这个功能的名字感觉有些摸不着头脑,这其实是对这个模块功能的一种比喻,这个模块提供了股票表现动画回放的功能,选定股票及时间区间后可以相互比较它们之间的表现,如同伯乐选千里马一般。选股时可以根据目标股票所属的类别来筛选,不仅可以看到类别内股票的表现与排名,还能看到类别的整体表现,这有助于预测某板块或某些股票的未来表现。我在投资比赛的过程中就经常使用它来挑选打算投资的板块。

2. 算法篇总结

2.1 算法内容

在算法篇当中,主要讲了三个算法:回归分析、动态时间规整和遗传算法。回归分析是一种预测建模技术的方法,研究因变量(目标)和自变量(预测器)之前的关系。这一技术被用在预测、时间序列模型和寻找变量之间因果关系。自回归模型尤其适用于对序列化的数据进行建模,而实际当中的金融时序数据如股票价格变动数据就恰好符合这一特征,因此在智能证券投资方面自回归模型较为常用,其的目标就是基于历史数据学习一个模型,使得这个模型能够预测某只股票在未来一段时期内的价格。

动态时间规整算法则是基于动态规划的思想,在时间序列中,需要比较相似性的两段时间序列的长度可能并不相等,在语音识别领域表现为不同人的语速不同。而且同一个单词内的不同音素的发音速度也不同,比如有的人会把"A"这个音拖得很长,或者把"i"发的很短。另外,不同时间序列可能仅仅存在时间轴上的位移,亦即在还原位移的情况下,两个时间序列是一致的。在这些复杂情况下,使用传统的欧几里得距离无法有效地求的两个时间序列之间的距离(或者相似性)。在智能证券投资领域,这一算法被用来衡量两只股票在某些时间范围内走势的相似程度,即将表示股票的价格变动的时间序列作为该算法的输入。

遗传算法是进化算法的一种。应用到智能证券投资领域来说,可以将一些智能投资机器人作为初始种群,将其编码表示,通过借鉴现实生物世界中的一些现象如

遗传、突变、自然选择以及杂交等来对种群中的个体模拟相应的操作,适应度的判断既可以用机器人在实际投资中的收益表现情况、亦可以根据自己的投资偏好制定相应的规则作为标准,以此不断迭代,选出最优的个体。

2.2 研究设想

均线指标实际上是移动平均线指标的简称。由于该指标是反映价格运行趋势的重要指标,其运行趋势一旦形成,将在一段时间内继续保持,趋势运行所形成的高点或低点又分别具有阻挡或支撑作用,因此均线指标所在的点位往往是十分重要的支撑或阻力位,这就为我们提供了买进或卖出的有利时机,均线系统的价值也正在于此。如图 2-1 所示,图中蓝色、黄色和紫色线即为均线。



图 2-1

由此我们可以根据这些均线走势与交点的特征来衡量股票之间的相似程度,以此找到相似的股票和对应时间段,并以此来预测股票后续的价格走势,也可以根据均线相对位置关系和趋势来辅助买入或卖出的决策,这里可以借鉴股市常用的均线策略来进行操作.

3. 总结, 感悟, 关于课程或平台的建议

作为一个专业为 CS 的工科生, 我一直对于股票金融领域想有一些了解,但是一直没有实际的行动. 刚好发现有这么一门智能证券投资课程,于是就选修了这门课.

学了这门课以后我认为对我最大的帮助则是在老师讲解以及模拟投资比赛中掌握了很多证券投资的知识。让我明白我确实不适合投资,还是老老实实写代码吧.

关于课程的建议主要有一点,就是每次课程之后的实验报告看起来有很多问,但实际上问的东西都差不多,让人写很多重复的文字实在是很困难.