# 《面向新型理财的强化学习技术研究》

## 引言

当前，随着生活条件的改善，个人存款增多，人们的理财意识逐渐增强。**理财方式有：定期存款、个人理财产品、基金或者股票**。定期存款理财方式最稳健，属于一种保本的理财方式。截止2021年，央行公布的定期存款的年利率为1.5%（整存整取1年）。个人理财产品系银行面向合格客户设计并销售的理财产品，属于一种低风险理财方式，具体收益由产品推出方决定，大致在2%-3%之间（时间一般为6个月）。基金，这里主要指证券投资基金。它种类繁多，根据投资对象可分为：债券基金、股票基金、货币基金和混合基金四种类型。它们分别具有不同的投资风险和投资收益。一般而言，投资风险和投资收益成反比关系。**股票**是股份公司为筹集资金而发行给股东作为持股凭证并借以取得股息和红利的一种有价证券，股票的投资风险为上述理财产品中最高，换言之其收益水平也最高。

然而，在全世界通货膨胀普遍存在的背景下，中国每年的通货膨胀率大约为2%~3%之间。换言之，如果我们的理财方式的收益低于通货膨胀率，这实际上是一种失败的理财方式。股票属于一种高风险高回报的理财方式。为获得更高回报，投资者需要承担更高的风险，这对于大众而言是无法接受的。理财产品的收益大致与通货膨胀相同，该种理财方式仅能帮我们做到对冲通货膨胀，而达不到财产增加的目的。基金是一种极其灵活的理财方式，投资者根据自身对风险承受程度选择不同类型的基金。一般的偏股型基金，其年收益往往在8%~12%之间，风险相对较高。偏债型基金年收益大约在5%~8%之间。货币基金年收益大约为3%~5%。尽管基金是一种适合大众的理财方式，但是绝大多数人由于缺乏金融知识，对于如何购买基金，何时补仓，何时清仓，如何分配持有的资金到不同基金，以获得高回报的同时承担更低风险往往难以做到理智，容易被短期的利益所迷惑。

强化学习是机器学习的范式和方法论之一。强化学习就是学习“做什么（即如何把当前的情景映射成动作）才能使得数值化的收益信号最大化”。强化学习有四个核心要素：策略、收益信号、价值函数以及（可选的）对环境建立的模型。策略定义了学习智能体在特定时间的行为方式。收益信号定义了强化学习问题中的目标；收益信号表明了在短时间内什么是好的，而价值函数则表示了从长远的角度看什么是好的。强化学习的第四个要素是对环境的建立。这是一种对环境的反应模式的模拟，或者更一般地说，它允许对外部环境的行为进行推断。

利用强化学习方法来找到一个最优投资策略以达到最大化累计收益的目的。