**投资项目**

COMS W4995机器学习在金融中的应用

最终报告

2023年7月

吴彤，tw2906

卡塔鲁埃

[1. 范围 3](#_Toc139248719)

[1.1. 语境的定义 3](#_Toc139248720)

[1.2. 投资者的定义 3](#_Toc139248721)

[2. 投资、回报和风险目标 4](#_Toc139248722)

[2.1. 总体投资目标 4](#_Toc139248723)

[2.2. 回报、分配和风险要求 4](#_Toc139248724)

[2.3. 投资者的风险容忍度 4](#_Toc139248725)

[2.4. 相关制约因素 5](#_Toc139248726)

[2.5. 其他相关考虑因素 5](#_Toc139248727)

[3. 办法 6](#_Toc139248728)

[3.1. 机器学习算法 6](#_Toc139248729)

[3.2. 平均值-方差优化 7](#_Toc139248730)

[3.3. 指数模型 9](#_Toc139248731)

[3.4. 资本资产定价模型 11](#_Toc139248732)

[3.5. 套利理论和多因素模型 12](#_Toc139248733)

[3.6. 股权评估模型 13](#_Toc139248734)

[3.7. 布莱克-利特曼模型 16](#_Toc139248735)

[3.8. 算法交易 18](#_Toc139248736)

[4. 股票选择 21](#_Toc139248737)

[5. 投资组合评估 23](#_Toc139248738)

[6. 投资组合表现 24](#_Toc139248739)

[6.1. 投资组合和基准统计数据 24](#_Toc139248740)

[6.2. 业绩归属 27](#_Toc139248741)

[6.3. 因素风格分析--Fama French三因素模型 28](#_Toc139248742)

[7. 分析和结论 31](#_Toc139248743)

[7.1. 超出期望值 超出期望值 31](#_Toc139248744)

[7.2. 表现不佳的预期 31](#_Toc139248745)

[7.3. 主要学习内容 32](#_Toc139248746)

[7.4. 未来的影响 32](#_Toc139248747)

[8. 参考文献 34](#_Toc139248748)

# 范围

在动态的投资世界中，彻底了解投资计划的执行背景和投资计划的委托人（投资者）是至关重要的。这种详细的理解是财务规划和投资组合管理的骨干。

## 语境定义

在这种情况下，背景涉及100万美元，这是客户自由资产的一部分，将被投资于股票以获得最佳回报。鉴于这一金额的巨大性质，我们对这一投资的方法需要有条不紊、精确，并完全符合客户的风险状况。我们的目标是创建一个多元化的投资组合，在有效管理风险的同时优化收益，我们将利用各种金融模型，加上机器学习算法优化，来实现这一目标。

考虑到我们将在不到一个月的时间内管理这些资产，投资范围是相当短期的。鉴于我们的投资组合选择必须平衡潜在的短期市场波动，同时以最佳回报为目标，因此，短的时间框架将成为决策过程的因素。然而，我们必须在长期框架内优先考虑投资组合的表现。投资组合的长期回报决定了客户是否能继续与我们签约。

## 投资者的定义

我们的客户是一位年轻的企业家，拥有一家国际上成功的公司。财富的来源是通过该公司的几条产品线，这些产品在全球范围内销售，并占有相当大的市场份额。这种商业上的成功使客户积累了大量的自由资产，而客户正是将其中的一部分--100万美元--委托给我们进行投资。

尽管客户的商业头脑在了解股市动态方面可能会有优势，但我们必须记住，投资者的主要专业领域是创业，而不一定是投资管理，这构成了他们信任我们有效管理其资产的基础。

因此，我们致力于创建一个尊重客户的创业精神、风险偏好和短期投资时间表的策略，同时深深扎根于既定的金融模式，并通过机器学习算法的力量来丰富投资组合的内容。我们将在这一旅程中注重透明度、沟通和实现风险与回报之间的最佳平衡。

# 投资、回报和风险目标

## 总体投资目标

我们这项投资的主要目标是在不使客户的投资组合面临过度风险的情况下，获得尽可能高的回报。我们的目标是建立一个主要由股票组成的多元化投资组合，考虑到其在较短时期内获得较高回报的潜力。所选择的策略将平衡投资组合增长与风险管理的需要，与客户的风险承受能力保持一致。 ("个人投资者投资政策声明的要素" n.d.)

## 回报、分配和风险要求

鉴于投资期限相对较短，不到一个月，我们的策略必须适应潜在的短期市场波动。虽然我们的目标是实现最佳回报，但我们也需要考虑市场波动的可能性，这可能会在短期内影响投资组合的价值。

这项投资的回报要求将基于投资时的市场条件。然而，应该注意的是，高的短期回报往往伴随着高的风险，我们的目的是通过多样化和战略性的资产配置来减轻风险。

至于风险，我们将使用投资组合回报的标准偏差作为我们的主要风险指标，并承认较高的投资组合标准偏差通常意味着更显著的风险。然而，我们的目标是通过仔细选择股票和充分的多样化来尽量减少这种风险。投资组合的风险水平将通过仔细选择股票并将每个股票的风险限制在总资产的5%至10%之间来控制。这也将确保充分的多样化来减轻风险。

## 投资者的风险容忍度

作为一个年轻的企业家，在通过商业投资承担风险方面有着良好的记录，我们的客户被认为具有适度的高风险承受能力。然而，鉴于投资金额巨大，我们的目标是平衡高回报的潜力和谨慎的风险方法。 ("个人投资者投资政策声明的要素" n.d.).我们的风险管理策略将涉及资产多元化，选择具有稳定收益的公司，并有可能投资于防御性股票以对冲市场的衰退。

## 相关制约因素

我们的客户已经提供了具体的要求，为我们的投资战略建立了框架。我们致力于遵守这些限制条件，同时在投资领域进行导航，以确保获得最佳结果。客户提供的主要限制条件如下：

**股票选择：**我们的客户规定，投资组合必须由10至20只股票组成。这些股票应该是在美国交易所专门交易的，我们可以灵活地以长仓或短仓的方式持有这些股票。

**仓位规模：**为了防止对任何单一资产的过度暴露并促进投资组合的多样化，客户规定每个头寸，无论是多头还是空头，应占总资产的5%至10%。这一规则的例外情况是指数模型或算法交易，这些头寸规模限制在这里不适用。

**现金持有量：**为了确保投资组合得到充分投资，客户规定现金持有量不应超过投资组合总价值的20%。这一要求指导我们保持积极的投资姿态。

**数据要求：**客户强调，投资组合的选择和优化必须基于过去五年所选股票的历史数据。这意味着我们的策略必须以数据为导向，利用历史市场模式为我们的投资决策提供依据。

**模型和算法：**为了生成一个优化的投资组合，客户要求从预定的清单中纳入至少三种机器学习算法。这些算法将被应用于各种金融模型，包括但不限于平均方差优化、指数模型、资本资产定价模型、套利定价理论和多因素模型、股票估值模型、Black Litterman模型和算法交易。从这些金融模型和机器学习算法中，我们将选择具有最高夏普比率的投资组合，应用于投资组合配置。

## 其他相关考虑因素

其他可能影响我们投资决策的因素包括当前的经济状况和市场条件。宏观经济和微观经济指标在为我们的投资战略提供信息方面都将是至关重要的。

此外，客户的个人价值观也可能影响我们的投资决策。例如，如果客户对社会责任投资感兴趣，我们会在选择股票时考虑ESG（环境、社会和治理）因素。

在本报告的下一部分，我们将深入研究我们将采用的具体金融模型和机器学习算法，以帮助我们在确定的风险容忍度、约束条件和其他相关考虑因素下实现既定目标。

# 办法

## 机器学习算法

在这个项目中，我们使用各种机器学习算法，包括随机森林、决策树、支持向量机（SVM）和神经网络来预测选定股票的潜在价格和指数变化。这些预测后来被用来分析高级模型的信心值，如Black Litterman。

**随机森林：**这种算法的优势在于它能够处理具有高维度的大数据集。它通过创建随机的特征子集并使用这些子集构建较小的树来减少过拟合。(Breiman 2001).然而，一个主要的缺点是它的复杂性，导致训练时间延长。

**决策树**：决策树算法很容易理解和可视化。它隐含地进行了变量筛选或特征选择(Bekena 2017).相反，决策树很容易过度拟合，尤其是在处理连续数字变量时。

**支持向量机（SVM）：支持向量**机在高维空间中特别有效，最适合于数据中具有明显分离边缘的复杂领域的问题。然而，它的效率随着数据噪音的增加而下降 (Luca 2020).

**神经网络：**神经网络是强大的算法，可以对复杂的模式和预测问题进行建模。在处理高维、非线性数据时，它们是有益的。然而，它们需要大量的数据和大量的计算资源。 (Luca 2020).

## 平均值-方差优化

1. **简介**

均值-方差优化模型，也被称为现代投资组合理论（MPT），是由哈里-马科维茨在1952年提出。这个理论认为，仅仅看某只股票的预期风险和收益是不够的。 ("现代投资组合理论中的投资组合优化" n.d.).通过投资于一只以上的股票，投资者可以获得多样化的好处--其中最主要的是降低投资组合的风险性。它显示了规避风险的投资者如何在给定的市场风险水平的基础上构建投资组合，使预期收益最大化。这个模型需要投资组合中考虑的不同资产的历史回报数据。掌握可能影响资产收益的不同经济条件和变量的数据也是有益的。

1. **关键公式**

在均值-方差优化模型中，投资组合的预期收益是单个资产预期收益的加权总和，其中权重是投资于每种资产的组合总价值的比例。它的计算方法如下：

其中 是投资组合的预期收益、 是资产n在组合中的权重，和 是资产n的预期收益。

投资组合的方差（可通过平方根得到标准差，是衡量风险的标准）计算如下：

其中 是投资组合的方差、 和 是投资组合中第i项和第j项资产的权重，和 是第i项和第j项资产的收益之间的相关性。

夏普比率 衡量风险调整后的回报，然后可以计算如下：

其中 是投资组合的预期收益、 是无风险利率，和 是投资组合的标准差。

1. **优势和劣势**

均值-方差优化模型提供了一种系统的、定量的分散化和风险管理的方法。它非常灵活，因为它可以适用于各种规模的投资组合，并能适应广泛的投资策略和金融资产类型。另外，它还可以进行明确的风险控制。

相反，该模型假设资产收益是正态分布的，并且相关性是固定的和稳定的，而在现实世界的市场中可能不是这样的。它很容易受到关于未来表现的输入假设的影响。输入的微小变化会导致最佳投资组合的重大变化。它假定所有的投资者都是规避风险的，并以较高的回报和较低的风险来实现效用最大化，这可能并不适合所有的投资者。

1. **优化**

在使用财务模型之前，我们对每只股票的价格进行预测。我们使用三种不同的方法（随机森林、决策树和SVM）来预测未来20天的价格变化，并将预测的数据和历史数据作为金融模型的输入。值得一提的是，我们将过去一年的历史价格和预测价格的权重增加了1.5倍（如下图代码块所示），这使得模型在做决策时对近期和未来的数据有更大的偏向。这一部分的优化将用于大多数金融模型，所以我们在后续章节中不会默认重复这一优化。

*# 添加权重*recent\_year = end\_date - pd.DateOffset(year=1)  
recent\_data = prices[prices.index >= recent\_year]  
weighted\_prices = pd.concat([price, recent\_data, recent\_data \* 0.5])

在使用PyPortfolioOpt库时，我们在循环中试验了不同的风险因素伽玛。在每个循环中，通过创建一个EfficientFrontier对象，然后添加一系列约束条件，找到最佳投资组合。我们使用add\_sector\_constraints方法来设置不同股票类型的最大和最小投资限制 ("Mean-Variance Optimization - PyPortfolioOpt 1.5.4 Documentation" n.d.).这是对标准均值方差模型的优化，它不考虑股票类型的比例。然后，我们使用add\_constraint方法，按照客户的要求，为每只股票设置最大（10%）和最小（5%）的投资限制。接下来，我们使用max\_sharpe方法来最大化夏普比率。夏普比率是一个投资组合的预期收益和波动率的比率，它可以成为评估投资组合业绩的一个重要指标。虽然标准的均值-方差模型也考虑了收益和波动率，但它并不直接优化夏普比率。最后，我们通过计算投资组合的基本性能指标来使用超参数调整，并在一个循环中不断选择最佳投资组合，其标准是最大夏普比率。该功能模块如下图1所示。

文本

描述已自动生成

图：投资组合的优化

## 指数模型

1. **简介**

指数模型，也被称为单一指数模型（SIM），是资本资产定价模型（CAPM）的简化版本。该模型假定任何证券的收益与市场指数的收益呈线性关系。它有助于降低与投资组合方差计算相关的复杂性。 (*投资*2023)

这个模型需要的主要数据是所选资产和所选市场指数的历史回报数据。此外，有关无风险利率和市场整体风险水平的细节可以补充对单个股票表现的理解。在这种情况下，我们选择标普500指数（^SPX）作为市场指数。

1. **关键公式**

证券的预期收益率 指数模型中的证券的预期收益率计算如下：

其中 是证券的预期收益 , 是市场回报率为零时证券的预期回报率、 是证券的收益率对市场收益率的敏感性，和 是市场的预期收益率。

一个证券的总风险 (方差）是由系统风险（市场风险）和特定风险（非系统风险）组成：

其中 是安全的总风险 , 是证券的收益率对市场收益率的敏感性、 是市场方差，和 是证券的具体风险 .

夏普比率的计算方法与均值-方差优化模型类似。

1. **优势和劣势**

指数模型大大简化了投资组合风险的计算，将其从一个潜在的庞大协方差矩阵减少到一套简单得多的计算。它很容易理解和实施，系统性风险和特定风险被清楚地分开。

另一方面，指数模型假设证券的回报与市场回报之间存在线性关系，这在现实生活中可能并不总是成立。它可能不像其他模型那样全面，因为它只考虑单一因素（市场指数回报）。

1. **优化**

在这里，我们使用ML算法来预测每只股票的未来价格，就像我们在均值-方差优化模型中所做的一样。此外，我们还实施ML算法来预测指数价格。

之后，投资组合的优化与均值-方差优化模型相同。值得注意的是，我们在计算β值时做了一些优化。我们输入了股票和指数价格的历史和预测数据。然后我们特别注意了数据的排列和去重，以确保计算的准确性。之后，我们用数据来计算协方差矩阵和方差来计算贝塔。

## 资本资产定价模型

1. **简介**

资本资产定价模型（CAPM）确定了一项投资的预期收益，考虑到其系统风险。该模型宣称，一项资产的预期收益等于无风险利率加上风险溢价，而风险溢价是基于资产的β系数，即衡量其市场风险的指标。

CAPM所需的数据主要包括指数模型所需的所有数据，加上无风险利率，可以用长期政府债券的收益率来近似计算。同时，资产和市场收益可以从历史价格数据中获得。在这种情况下，我们从美国政府的10年期国债利率中获取无风险利率，它将是3.71%。 (CNBC n.d.).

1. **关键公式**

根据CAPM，一项资产的预期收益率计算如下：

其中 是资产的预期收益、 是无风险利率、 是该资产的贝塔系数，和 是市场的预期收益。

投资组合的标准差（风险）仍将根据收益的权重和协方差来计算，这可以从历史数据中估算出来。

夏普比率的计算方法与之前的模型相同。

1. **优势和劣势**

CAPM提供了一个有用的风险-收益权衡的措施，并协助评估风险证券和确定适当的所需回报率。它的理解和使用相对简单，只需要几个输入参数。 ("CAPM: 假设和限制 | 证券 | 金融经济学" n.d.)

然而，CAPM假设投资者是理性的，市场是有效的，这在现实世界的情况下可能并不总是正确的。它假定一项资产的市场风险（β）完全解释了其预期收益，而不考虑其他潜在因素。

1. **优化**

实现机器学习算法和优化应该与索引模型相同。

## 套利理论和多因素模型

1. **简介**

套利定价理论（APT）是一种资产定价的一般理论，认为金融资产的预期收益可以被模拟为各种宏观经济因素的线性函数。该理论认为，可以利用一项资产与许多因素之间的关系来预测该资产的收益。

如同CAPM一样，多因素模型也是对这一思想的扩展，即使用多个因素而不是只有一个因素（市场回报）。这些因素可以是影响证券收益的特定公司、特定行业或宏观经济变量。

APT和多因素模型所需的数据包括资产的历史回报，无风险利率，以及所考虑的每个因素的值（这些可能是宏观经济变量，如通货膨胀，GDP增长率，利率等，或金融市场变量，如市场回报）。

1. **关键公式**

根据多因素模型，一项资产的预期收益率计算为：：

其中 是资产的预期收益、 是无风险利率、 是资产收益率对因素的敏感性。 .

投资组合的标准差（风险）仍将根据收益的权重和协方差来计算，这可以从历史数据中估算出来。夏普比率的计算方法与之前的模型相同。

1. **优势和劣势**

APT和多因素模型提供了一个比CAPM更灵活的框架，允许包含可能影响资产回报的多种因素。这些模型比CAPM更能适应实际的市场条件。

然而，APT和多因素模型需要选择一些假定影响资产收益的因素，这些因素的确定可能具有挑战性，而且可能随时间变化。对于包括哪些因素没有具体的准则，使得模型的标准化程度低于CAPM。

1. **优化**

在实践APT多因素模型的过程中，我们加入了岭回归算法，这是一种专门用于处理协方差数据的线性回归方法，其计算方法是带有惩罚项的最小二乘函数，可以用来防止过拟合。协方差问题是指预测变量之间存在高度的线性关系，这会导致线性回归模型的稳定性和准确性下降。在数学表达中，协方差可能将预测矩阵的行列式关闭为零，导致其反矩阵的不稳定，使线性回归系数估计的方差广泛而不准确。岭回归通过在成本函数中加入一个L2参数的惩罚项（系数的平方和）来解决协方差问题。其主要目标是最小化预测误差和系数大小的平方之和。在这个过程中，岭回归保留了所有的特征变量。然而，它减少了特征变量的系数，使模型对单个特征的噪声不那么敏感，从而防止过度拟合。 (Zou and Hastie 2005)

另外，我们还增加了几个因素。除了国内生产总值和标准普尔500指数等主要因素外，我们还增加了纳斯达克、道琼斯50指数和罗素200指数，以帮助我们更准确地预测资产价格。

## 股票估值模型

1. **简介**

股票估值模型是用来计算一个公司股票内在价值的方法。然后，这一价值可以与当前的市场价格进行比较，以确定股票是否被高估或低估。最流行的股票估值模型包括股息折现模型（DDM），折现现金流（DCF），以及价格/收益（P/E）模型。对于这个项目，我们将主要关注贴现现金流（DCF）。

对于DDM，所需的数据包括当前和历史的股息支付，股息的预期增长率，以及所需的回报率。对于这个模型，需要财务报表和历史市场数据。

1. **关键公式**

在DDM中，公司股票的内在价值（P）的计算方法是：：

其中 是下一年的预期股息、 是要求的回报率，和 是红利的增长率。

在股票估值模型的背景下，预期回报率是所需的回报率（k），而股票价格的波动代表风险。

然后，夏普比率将用投资组合相对于无风险利率的超额回报除以超额回报的标准差来计算。

1. **优势和劣势**

这些模型提供了一种简单而直观的方式来估计一只股票的内在价值，可以帮助识别潜在的投资机会。该模型结合了关键的财务指标，并提供了对驱动公司价值的基本因素的理解。

然而，股权评估模型，特别是DDM，对于不支付股息或不定期支付股息的公司来说，可能不太有利。这些模型在很大程度上依赖于输入的准确性，包括所需的回报率和股息的增长率，而要准确估计这些数据是有难度的。

1. **优化**

作为我们分析的一部分，我们注意到有几只选定的股票没有分配股息。这给将股息折现模型（DDM）应用于整个投资组合带来了挑战。作为一种变通方法，我们用每年的股价变化取代这些公司不存在的股息数据，以代表公司的增长率。

亚马逊、特斯拉和谷歌等公司都没有派发股息。我们的分析表明，这其中的原因可能包括：

**再投资：**许多高科技公司正处于快速增长阶段。因此，他们更愿意将所有的利润和现金流再投资到公司，用于进一步发展和扩张。这可能涉及研究新产品，扩大生产能力，增加市场份额，或投资于新的商业机会。对这类公司来说，再投资可能会带来更高的长期回报，对股东来说，这可能比直接获得股息更有吸引力。

**成长阶段：**对于许多高科技公司来说，他们可能还没有实现稳定的、可预测的利润。在这种情况下，公司可能会优先考虑保留现金，以减轻积极的风险，处理潜在的经济衰退，并普遍经受金融风暴。

**信号传递：**从公司财务的角度来看，市场可以将支付股息解释为公司增长放缓的信号。高科技公司通常想让投资者了解其高增长潜力和投资吸引力。因此，他们可能选择不派发股息。

因此，尽管股息折扣模型对许多股票提供了有价值的见解，但其应用必须加以调整，以适应投资组合中每家公司的具体情况和财务战略。通过采取这种细致入微的方法，我们确保我们的投资策略保持稳健，并很好地适应股票市场的动态性质。

我们首先获取所有公司的股息数据。然后，过滤掉那些不分红的公司（图2），取其过去五年的股价，并计算其年度股价变化，以填补数据集（图2和图3）。

同时，我们需要预测这些公司在下一个时期的增长率。在这种情况下，我们还实现了三种ML算法来帮助模型预测增长率（图3）。

在预测了增长率之后，我们还将这些数据输入到Black Litterman模型中，以便得到更精确的结果。对于这个模型的优化，我们将在下一节讨论。

文本

描述已自动生成

图 确定哪些公司不分配股利。

文本

描述已自动生成

图 取出红利或股票价格来分析年增长率。

## 黑衣人模式

1. **简介**

布莱克-利特曼模型是用于投资组合分配的，它采用贝叶斯方法，将投资者对一种或多种资产预期收益的主观看法与市场均衡向量（市场投资组合权重）结合起来，得出一个新的、混合的预期收益估计。 (*投资*2023)

布莱克-李特曼模型利用了市场均衡收益，这是从资产市场价值中得出的，同时还有投资者对资产收益的独特看法。因此，需要的数据包括投资组合中每种资产的市值（用于市场均衡收益），资产收益的协方差矩阵，以及投资者对收益的主观看法（信心值）。

1. **关键公式**

Black-Litterman模型中的预期收益率计算如下：

在哪里？

E(R)是综合收益、

τ是一个标量，代表先验的不确定性、

Σ是资产收益的协方差矩阵、

P代表确定投资者观点中涉及的资产的矩阵、

Ω是投资者观点的误差项的协方差矩阵、

π是均衡收益的向量、

Q是投资者观点的一个向量。

然后可以使用投资组合的协方差矩阵和从Black-Litterman模型得出的权重来计算风险或标准差。

1. **优势和劣势**

该模型提供了一个正式的机制，将投资者的观点纳入投资组合的优化。它允许根据市场预期或投资者的观点来调整投资组合。另外，它还有助于克服马科维茨均值-方差优化模型中经常出现的极端权重问题。

1. **优化**

在优化Black Litterman模型时，我们同样引入了一个预测股票价格的机器学习模型。同时，我们根据预测的价格变化百分比，产生了用户的主观看法。正如我们之前所说，我们的目标是使机器学习算法更多更好地融入金融模型。我们根据算法预测的变化百分比加上之前量化的客户风险承受能力的数值，得到了每只股票的信心值并将其归一化。

文本

描述已自动生成

图 使用机器学习算法来提高置信度的计算。

## 算法交易

1. **简介**

算法交易，也被称为algo-trading或黑箱交易，是使用程序和算法在市场上以人类交易者不可能达到的速度进行交易的过程。这些计算机程序建立在复杂的数学模型和算法上，可以进行高速、高频的交易，并被用来做出购买或出售证券的决定。

算法交易使用高频数据，包括实时或接近实时的价格反馈、交易量信息、订单簿数据和其他市场信号。它还可以纳入更广泛的数据集，如经济指标、新闻馈送，甚至用于情绪分析的社交媒体数据。

1. **优势和劣势**

算法交易可以比人类交易者更快地执行交易，消除了交易中的人为错误风险。算法交易策略可以在部署前对历史数据进行回测，以评估其性能。

然而，存在过度优化的风险，即策略对过去的数据调整得过于精细，在新的数据上表现不佳。另外，高频率的交易可能会导致市场的突然波动。算法中的任何错误或系统故障都可能导致重大损失。

1. **优化**

在实践算法交易的过程中，我们使用了与前面类似的优化策略。我们使用三种不同的机器学习算法分类器来预测未来价格，并对预测的行动进行分类：买入、卖出或观望。在这种情况下，我们根据预测未来股票价格走势，根据波动带（布林线）的位置，产生买入或卖出信号。

布林线是一种技术分析工具，它对价格信息进行统计处理，以产生对价格变动的预测。它的要素构成是一个中轨（均值）和两个上下轨，与此均值的标准偏差。 (Zheng, Li, and Feng 2022)这个策略的主要思想是，价格在大部分时间都会在这个区间内波动。当价格接近上轨时，意味着股票可能处于超买状态，可能会出现下跌；反之，当价格接近下轨时，意味着股票可能处于超买状态，可能出现反弹。而当价格接近中轨时，市场可能处于平衡状态。 (Williams 2006)

在代码中，首先用公式（价格-下限）/（上限-下限）计算每一天的价格位置在波动带中的比例，然后将其作为特征输入机器学习模型。接下来，模型的目标变量（y）是预测的交易信号，它是根据未来真实价格与上波段的关系设定的：如果真实价格超过了上波段，那么就被认为是卖出信号（标记为-1），否则就是买入信号（标记为1）。

通过这种方式，模型学会了价格在波动区间的位置与未来价格走势之间的关系，可以用来对未来进行预测。当模型预测到一个买入信号时，它就进行买入操作；当预测到一个卖出信号时，它就进行卖出操作。

算法学习的一个重要部分是与股票账户的连接以及程序自动获取信息和完成投资策略。在这里，我们使用TWS的API来连接到互动经纪公司的账户。在实现的模型中，我们将从互动经纪公司的账户中获得每只股票的头寸和当前股价（图7）。这是通过调用'reqPositions'方法完成的（图5），该方法从服务器上请求当前账户的所有头寸。服务器返回这些信息并通过调用position方法进行处理。最后，'positionEnd'方法在收到所有仓位后被调用，然后设置一个事件，告诉其他部分所有仓位都已收到（图5）。 ("TWS API v9.72+: Positions" n.d.)

在收到相关数据后，模型被送入机器学习算法，以预测投资信号（图7）。然后程序根据投资信号做出买入或卖出的决定，并通过TWS API下单。在调用这个方法之前，创建一个合同对象来指定你要交易的股票，并创建一个订单对象来指定交易细节（如买/卖、数量、价格等）。然后，通过调用'placeOrder'方法将该订单发送到服务器（图6）。 ("TWS API v9.72+: Placing Orders" n.d.)

文本

描述已自动生成

图 TWS API实现和获取持有股票值

文本

描述已自动生成

图 TWS API订立合同和订单

文本

描述已自动生成

图 使用TWS API计算持股价值并下单

# 股票选择

选择股票进行投资是投资组合管理的一个关键部分，需要对各种数量和质量方面进行仔细评估。

投资股票的选择是投资组合管理的一个组成部分，需要对各种数量和质量方面进行仔细评估。鉴于客户的高风险承受能力，可能会选择一些高波动性的行业。然而，我们的目标不仅是在短期内为客户提供尽可能高的回报率，而且我们还期望与客户建立长期的关系。因此，我们决定选择一些高科技公司的股票，如特斯拉、谷歌、亚马逊等。此外，我们将剩余的资金分配给其他各种行业，包括医疗保健（如辉瑞），线下产业（如好市多），航空航天（如洛克希德-马丁），银行（如汇丰）等。我们选择各种行业，并尽量确保这些行业不存在关联性。这些股票降低了投资组合中股票的相关性。这在股市震荡的情况下提供了最大的风险对冲。然而，在提供坚实的风险对冲的同时，可能会有一个较低的增长率来应对股票的牛市。我们的投资目标是稳健并在一段时间内实现有效的回报。因此，风险对冲是我们所需要的。

考虑到客户的风险承受能力和投资目标，我们选择了以下17只股票：

**AAPL（苹果公司）：**作为一家拥有强大品牌和稳定消费群体的领先科技公司，苹果公司表现出稳定的增长和盈利能力。苹果公司强大的生态系统和创新的产品组合使其成为有弹性的选择。

**AMZN（Amazon.com Inc.）：**亚马逊在其电子商务和云计算领域持续增长。凭借其广泛的影响力和多元化的商业模式，亚马逊已准备好持续增长，并被认为是投资组合中一个可靠的补充。

**BSX（波士顿科学公司）：**波士顿科学公司是一家领先的医疗设备公司，由于在不断增长的医疗市场中不断创新和强大的产品组合，显示出强大的增长潜力。

**CI（Cigna公司）：**Cigna是一家全球公认的健康服务机构。鉴于医疗保健行业的稳定需求，它为投资组合提供了稳定性和多样性。

**COST（Costco Wholesale Corporation）：**好市多有一个稳定的商业模式，展示了持续的增长和坚实的会员续费率。这使其成为缓解风险的有力候选者。

**CVX（雪佛龙公司）：**雪佛龙公司是一家跨国能源公司，提供能源领域的风险，并提供潜在的增长和可观的股息收益。

**GOOGL（Alphabet公司）：**Alphabet在全球数字广告市场的主导地位，以及对云计算和人工智能的涉足，使其成为一项增长型投资。

**高盛集团（Goldman Sachs Group）：**高盛作为全球领先的投资银行、证券和投资管理公司，为金融服务业提供了机会，而金融服务业往往与整体经济增长呈正相关。

**汇丰银行（HSBC Holdings plc）：**汇丰提供国际风险，特别是亚洲市场的风险，增加了投资组合的多样化。

**LMT（洛克希德-马丁公司）：**洛克希德-马丁公司是一家全球航空航天和国防公司，有稳定的政府合同流入，为收入流提供了稳定性。

**MCD（麦当劳公司）：**麦当劳提供稳定的回报和分红，使其成为风险承受能力较低的投资者比较安全的赌注。

**默克（Merck & Co., Inc：）**默克公司是一家全球医疗保健公司，由于其强大的医药产品阵容和潜在药物管道，提供了稳定的增长前景。

**MSFT（微软公司）：**微软公司多样化的产品组合和对基于云的服务的日益重视，使其成为具有吸引力的增长型投资。

**PFE（辉瑞公司）：**辉瑞公司在医疗保健领域提供了一个坚实的存在。其多元化的药物组合和强大的管道，再加上其在COVID-19疫苗接种中的作用，使其具有强大的潜在回报。

**TSLA（特斯拉公司）：**特斯拉是电动汽车市场和可再生能源领域的领先者。它的高增长潜力，虽然伴随着高波动性，但提供了可观的回报前景。

**NVDA（英伟达公司）：**英伟达是图形处理单元（GPU）市场的领导者，处于有利地位，可以从游戏、数据中心和人工智能的趋势中获益，提供强大的增长潜力。

# 投资组合评估

我们的投资组合构建和管理过程从6月12日开始，于6月30日结束。在此，我们讨论在此期间我们的投资组合的演变，对其进行修改的原因，以及这些修改对投资组合的表现的影响。

**最初的投资（6月12日）：**我们最初的投资组合是使用均值-方差优化模型构建的。这是因为，在我们的初步分析中，这个模型相对于其他已实施的模型，如单一指数模型和资本资产定价模型（CAPM），表现出卓越的夏普比率性能。这为我们提供了一个分散的投资组合，最佳地平衡了风险和收益。

**第三天（6月15日）：**在投资组合管理的第三天，我们将英伟达（NVDA）加入我们的投资组合。英伟达公司充满希望的增长前景推动了这一决定。该公司最近发布了一系列基于其最新架构的商用CPU和GPU，标志着计算能力的大幅提升。我们相信这一新产品的发布将推动该公司的增长，使其成为我们投资组合中具有吸引力的投资。

**第11天（6月27日）：**我们的投资组合一直保持稳定，直到第11天，我们从均值-方差优化模型切换到Black-Litterman模型。在比较了Black-Litterman、股票估值和多因素模型的表现后，我们选择做出这一改变。Black-Litterman模型不仅表现出最高的夏普比率，而且还表现出预测价格的波动较小，表明预测更稳定。这一转换旨在优化我们的投资组合在不断变化的市场条件下的风险和收益之间的平衡。

在整个投资组合管理过程中，我们不断监测与我们所选股票相关的业绩和市场条件。只有当我们的模型显示有可能提高业绩或降低风险时，才会进行重新平衡，确保我们与客户的风险承受能力和投资目标保持一致。

我们的投资组合的演变是一个反复的过程，它在很大程度上依赖于我们选择的金融模型及其准确预测市场动向的能力。通过不断地重新评估我们的投资组合和指导其构建的模型，我们可以适应市场条件的变化并优化我们的投资战略。

# 投资组合表现

## 投资组合和基准统计

1. **数据统计**

在本节中，我们将深入分析和讨论我们的投资组合与基准相比的表现。我们将研究各种业绩指标，如平均收益、累积收益、波动率、夏普比率、阿尔法、贝塔、主动收益、主动风险、信息比率，以及投资组合与基准的相关性。表1显示了我们将在本节中用来分析投资组合和基准业绩的性能指标。

表 投资组合和基准统计数据

|  |  |  |
| --- | --- | --- |
|  | **证券公司** | **基准** |
| **算术平均数** | 0.153% | 0.198% |
| **几何平均数** | 0.151% | 0.195% |
| **累计总回报** | 1.976% | 2.569% |
| **年化累计收益率** | 46.116% | 63.497% |
| **总波动率** | 0.00775 | 0.00710 |
| **年化波动率** | 0.12296 | 0.11264 |
| **夏普比率** | 3.44880 |  |
| **阿尔法** | -0.00190 |  |
| **β** | 0.95940 |  |
| **主动风险** | 0.00353 |  |
| **信息比率** | -0.53780 |  |
| **与基准的相关性** | 0.89029 |  |

该投资组合的算术平均收益率为0.153%，几何平均收益率为0.151%，而基准的算术和几何平均收益率分别为0.198%和0.195%。这些数字表明，平均而言，该投资组合的表现略逊于基准。

就总累积回报而言，我们的投资组合回报率为1.976%，年化回报率为46.116%。相比之下，基准的总累积回报率为2.569%，年化回报率为63.497%。这再次表明，我们的投资组合在此期间的表现低于基准。

该投资组合的总波动率为0.00775，比基准的波动率0.00710略高一些。这意味着投资组合的年化波动率为0.12296，而基准为0.11264，表明投资组合的风险略高。

尽管风险较高，但我们的投资组合取得了令人印象深刻的年化夏普比率3.44880，显示了其相对于所承担的风险而言产生回报的效率。值得注意的是，在所提供的示例表中，用于计算夏普比率的公式是：

这个公式中使用的无风险回报是年化的。然而，项目的持续时间都采用了算术平均收益率和波动率（风险），导致三个数据的时间尺度不同，因此可能无法计算出准确的夏普比率。**在此，我们将夏普比率的计算公式修改为：：**

这确保了所有数据都是按年计算的。

该投资组合的β值为0.95940，表明它通常会与基准值同向移动，但波动性较小。投资组合的α值为-0.00190，意味着在此期间投资组合经理的策略没有比基准增加价值。

该组合的主动风险为0.00353，即主动收益的标准差。信息比率，即主动收益除以主动风险，为-0.53780，表明投资组合在风险调整的基础上相对于基准没有表现良好。

最后，我们的投资组合表现出与基准的高度相关性（0.89029），这意味着投资组合的回报与基准的回报紧密相连。

总之，我们在实现该计划提供长期高回报的目标方面进展顺利，同时也注重短期回报。与基准相比，该投资组合显示出略高的风险水平，表现略低于基准。然而，该投资组合的高夏普比率表明，它在补偿所承担的风险方面做得非常好，这也是我们在股票选择方面的决定发挥作用的地方。虽然我们的回报率没有超过基准，但我们有效地管理了风险，这对长期投资成功至关重要。我们未来的努力将集中在改进我们的选择和时机策略，以提供更高的阿尔法，同时保持一个可接受的风险水平。

1. **回归、方差分析和β分析**

在本节中，我们将讨论我们的投资组合的回归、方差分析和β分析结果。表2显示了我们需要讨论和分析的必要数据。

表 回归、方差分析和β分析

表格

描述已自动生成

**回归：**倍数R值为0.88。这意味着投资组合收益率的大约88%的变异性可以由基准收益率的变异性来解释。其他回归方法，包括R平方和调整后的R平方，都保持类似的数值。这个合理的高值表明，投资组合的业绩与基准业绩紧密相连。

**方差分析：**F统计量为42.49，相当高，相关的P值非常小（0.000029），小于0.05的显著性水平。这意味着整体回归模型具有统计学意义，并且具有解释力。

**β：**截距（α）为-0.0019，没有统计学意义（P值为0.734，大于0.05）。这表明没有明显的异常收益是独立于市场变动的。斜率（β）为0.9594，意味着基准收益率每增加1%，投资组合的收益率就会增加0.9594%。这个贝塔值小于1，说明投资组合的波动性比基准要小。

总之，回归分析表明，我们的投资组合紧跟市场表现，正如高倍数和显著的F统计量所显示的那样。小于1的β值表明投资组合的波动性低于市场，而不明显的α值表明在对市场变动进行调整后没有异常的回报。监测这些数字是至关重要的，因为它们可以为投资组合的市场敏感性和业绩提供有价值的见解。

## 绩效归属

本节将讨论业绩归因，分析投资组合管理期间的各种活动和决定，以及它们对整个投资组合业绩的贡献。我们将重点讨论资产配置和股票选择在投资组合业绩中的作用。表3显示了我们将在本节中分析和讨论的业绩指标，包括投资组合和基准业绩、比重、回报和贡献。

表 资产配置和股票选择的业绩归属

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **资产配置** | |  |  |
|  | **投资组合重量** | **基准重量** | **超重** | **基准收益** | **贡献** |
| **股票** | 98.13% | 100% | -1.87% | -1.47% | 0.03% |
| **现金** | 1.87% | 0% | 1.87% | 0 | 0 |
| **资产配置的贡献** | |  |  |  | 0.03% |
|  |  | **股票选择** | |  |  |
|  | **投资组合表现** | **基准性能** | **性能过剩** | **投资组合重量** | **贡献** |
| **选择股票的贡献** | 1.98% | 2.57% | -0.59% | 98.13% | -0.58% |
| **投资组合的总超额收益** | |  |  |  | -0.55% |

1. **资产配置**

资产配置是指在不同的资产类型（如股票、债券和现金）之间划分投资组合。我们的投资组合几乎全部（98.13%）都分配给了股票，剩下的（1.87%）是现金。相比之下，基准是完全投资于股票。

资产配置促成了0.03%的收益增长。发生这种情况是因为投资组合的一部分被分配到了现金，这期间没有受到股票收益下降的影响。这里要注意的是，在投资组合中持有部分现金可以提供稳定性和流动性，这在市场低迷时是很有好处的。

1. **股票选择**

选股是指为投资组合选择个别股票的过程。我们的投资组合在选股方面的表现（回报）为1.98%，而基准表现为2.57%，导致超额表现为-0.59%。这个数字乘以投资组合的股票权重（98.13%），对投资组合的总业绩贡献为-0.58%。

1. **总超额收益**

考虑到资产配置和股票选择，投资组合的总超额收益为-0.55%。这个数字表明，投资组合的表现略逊于基准，主要是由于选股。然而，资产配置的决定通过对投资组合回报的积极贡献提供了一些缓冲。

总之，虽然我们的投资组合的表现与基准相比略显逊色，但从这个归因分析中可以得到明显的启示。改进我们的选股过程可以大大提升投资组合的表现。同时，保持部分投资组合的现金可以在市场低迷时提供急需的稳定性，并能对整体业绩做出积极贡献。因此，这项分析为加强未来的投资组合管理战略提供了可操作的启示。

## 因素风格分析--Fama French三因素模型

1. **数据统计**

在这一节中，我们将使用Fama French三因素模型对投资组合的因素进行分析。Fama French三因素模型通过三个因素来解释投资组合的回报：市场风险（Mkt-RF），规模效应（SMB为Small Minus Big），以及价值效应（HML为High Minus Low）。该模型将帮助我们了解投资组合的表现有多少可以归因于这些系统性风险因素。表4显示了Fama French三因素模型的数据和投资组合的每日回报。请注意，我们在达特茅斯MBA学院检查了发布法玛-弗伦奇数据的网站 ("Kenneth R. French - Data Library" n.d.)并只能获得5月份的数据。然而，我们经常检查是否有6月份的数据发布。这个项目的时间是在6月，所以本节中的分析和结论可能与实际不同。我们相信，当6月份的Fama French数据发布后，我们将能得到一个更准确的判断。

表 Fama French三因素模型分析

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 日期 | 投资组合日收益率 | Mkt-RF | SMB | HML | 射频 | RP-RF |
| 2023/6/12 |  | 0.43 | 0.44 | -1.31 | 0.016 | 0.000% |
| 2023/6/13 | 0.528656686 | -0.21 | -0.5 | -0.1 | 0.016 | -1.071% |
| 2023/6/14 | 0.20366113 | -0.22 | -0.19 | 0.35 | 0.016 | -1.396% |
| 2023/6/15 | 0.912082219 | 0.47 | 0.85 | 0.02 | 0.016 | -0.688% |
| 2023/6/16 | -0.181413065 | -0.72 | -0.33 | -0.48 | 0.016 | -1.781% |
| 2023/6/20 | 0.019912124 | 1.32 | 0.88 | 1.2 | 0.016 | -1.580% |
| 2023/6/21 | -0.659584528 | 0.99 | -0.07 | -0.38 | 0.016 | -2.260% |
| 2023/6/22 | 0.674210621 | -0.26 | -0.46 | -0.32 | 0.016 | -0.926% |
| 2023/6/23 | -0.880466226 | 0.17 | 1.27 | -0.21 | 0.016 | -2.480% |
| 2023/6/26 | -1.393295017 | -1.09 | 0.4 | 1.06 | 0.016 | -2.993% |
| 2023/6/27 | 0.8957358 | -0.72 | -0.32 | -0.2 | 0.016 | -0.704% |
| 2023/6/28 | 0.126491159 | 0.65 | -0.71 | -1.08 | 0.016 | -1.474% |
| 2023/6/29 | 0.445528396 | 1.32 | 0.04 | -0.86 | 0.016 | -1.154% |
| 2023/6/30 | 1.302331261 | -0.01 | -0.15 | -0.42 | 0.016 | -0.298% |

**市场风险（Mkt-RF）：**观察投资组合的回报与Mkt-RF因素的关系，有的时候，投资组合的表现比市场好，有的时候则不然。例如，在6月13日，投资组合回报率为0.528%，而市场回报率为-0.21%，表明投资组合在那一天的表现好于市场。然而，在6月20日，投资组合的回报率为0.02%，而市场回报率为1.32%，表明投资组合的表现低于市场。这种波动是典型的，表明我们的投资组合在一定程度上受到市场风险的影响，这对大多数投资组合来说是典型的。

**规模效应（SMB）：**规模效应是由小盘股和大盘股之间的回报差异（SMB）衡量的。所提供数据中的SMB值随时间变化，与投资组合收益似乎没有一致的关系，这表明我们投资组合的表现可能不会受到规模风险的明显影响。

**价值效应（HML）：**所提供的数据中的HML值也有所不同，这表明投资组合的回报并不总是与高和低账面价值股票之间的回报差异一致。这可能表明，价值风险并没有严重影响我们投资组合的表现。

1. **回归、方差分析和β分析**

在这一节中，我们使用回归模型来实现Fama French三因素模型，以评估这些因素如何定量地影响投资组合的回报。该模型的参数可以通过线性回归进行估计，提供代表投资组合对这些风险因素敏感度的β系数。表5显示了我们将在本节中讨论和分析的数据。再次注意，我们使用的Fama French的数据是5月的，而不是6月的。因此，投资组合数据和Fama French数据的不相关是合理的。一旦有了Fama French 6月份的数据，我们会有更精确和可解释的数据和分析。

表 Fama French三因素模型的回归、方差分析和Beta分析

表格

描述已自动生成

**回归：**代表观察到的和预测的投资组合收益之间的相关性的多重R是0.42018。这个相对较低的数值表明自变量和因变量之间的线性关系较弱。R-Square是0.1765，意味着Fama-French三因素模型可以解释组合收益的大约17.65%。这个数值相对较低，表明模型中没有包括的其他因素也可能影响投资组合的收益。

**方差分析：**F统计量是2.143406，显著性F（P值）非常高，为1.666537。通常情况下，我们正在寻找一个小于0.05的p值来表明一个有统计学意义的模型。鉴于这种情况下的高p值，它表明观察到的R-Square可能没有统计学意义，这意味着模型可能不能很好地适应数据。

**β值：**Mkt-RF的贝塔系数为0.0017，但高P值（0.585）表明该系数在统计上并不显著。这表明该投资组合可能对整体市场回报不敏感。SMB的β系数是-0.0043。负的系数表明，投资组合在规模因素方面的表现往往是相反的。然而，0.31的P值表明这在统计学上并不显著。HML的β系数是-0.0011。同样，负系数表明该组合可能倾向于与价值因素表现相反。0.742的高P值表明这一结果在统计学上并不显著。

总之，回归分析表明，Fama French三因素模型（5月版）并不能显著解释特定时期的投资组合收益。其他因素，可能是特定行业或特定公司，可能在这里起作用。这个模型可能不是最适合投资组合的，应该考虑其他因素模型或专有模型。

# 分析和结论

## 超出期望值 超出期望值

在投资期间，科技公司表现出不同的股价趋势。其中，苹果、微软和特斯拉的股价持续上涨。这可能与这些公司在此期间的行动有关。

苹果在投资开始前的周末发布了一个新产品 ("介绍苹果Vision Pro：苹果的第一台空间计算机--苹果 "n.d.)，尽管股市开盘后第一天的股价反映出人们对这个新产品普遍不乐观。但之后，可能是随着消息的揭开和苹果在这个产品上的最新动作，股价在投资周期内持续上涨，很少有下跌的情况出现 ("苹果股价收盘创下历史新高" 2023年).

微软的股价上涨可能是中长期的，其LLM生成性人工智能的发布 ("Introducing ChatGPT" n.d.)，作为大股东在今年早些时候在全球掀起了巨大的波澜。用户数量在一个月内就突破了1亿人。该公司引发的巨大人工智能热潮推动了其主要投资者微软的股价上涨。 (Movement n.d.).

特斯拉最近一段时间没有发布新产品，但其产品出色的成本控制使其汽车产品在多个国家降价，可以说是成本下降，吸引了不少用户 (Merano 2023).而其老板马斯克最近也是动作频频，在5月底访问了该公司最大的出口国并与政府高官会面后，股价也是一路上涨 (Zhu 2023).这可能意味着马斯克与该国政府的会晤对特斯拉有积极意义--或者从更高的层面上说，对两国之间的贸易有积极意义。

## 表现低于预期

谷歌、亚马逊最近的股价表现不尽如人意。谷歌最近发布了新产品，但市场认可度并不高。另一方面，亚马逊也没有什么作为。这两家公司在某些方面是苹果和微软的竞争对手。当对手公司的股价上涨时，它们的股价下跌是有道理的。

英伟达本月采取了许多行动，首先发布了新的CPU+GPU "超级芯片"，为人工智能提供了更多的算力，并将成本降低到了一个可观的水平 (Edwards 2023).然后，它发布了其入门级家用GPU系列，并解禁了其评论，从中可以看出其性能迭代是出乎意料的 (Hollister 2023).这两款新产品可能是Nvidia的股价最近出现震荡的原因。

## 主要学习内容

总的来说，从投资项目中可以学到很多重要的经验和技巧。在代理投资者的资产之前，最关键的一步是分析投资者的风险承受能力和其他相关指标。大多数投资者没有高风险承受能力，所以保险的做法是通过分散投资于几个行业来对冲风险。

在确定一个财务模型之前，我们应该确定有哪些数据可用，哪些数据是我们感兴趣的，然后根据我们的投资目标选择合适的模型。我们还需要考虑每个金融模型的优势和劣势。机器学习算法可以预测广泛的数据，如股票价格的未来走势。虽然专业投资者会对特定的股票或股份有独特的见解，并能感知到投资的时机，但不可能做到面面俱到，尤其是在处理多个行业的股票时。机器学习算法可以识别和总结历史价格变动中隐藏的模式，这对我们很有帮助。在投资过程中，应不断比较几个金融模型的表现。股票市场瞬息万变，为了获得更高的回报，通常需要持续监测和更新模型。

在进行股票选择时，通常根据指数中包含的公司来选择投资。对大多数人来说，这是明智之举，因为列入指数的公司往往在其行业中占有重要地位。

在投资之后，对仓位数据和指数进行分析是非常必要的。我们通常可以从分析中获得具体的重要数据，如投资组合是否有效地对冲了风险，能否取得与指数相似或更高的收益，以及在面对市场风险、规模效应和价值效应时是否受到严重影响。这对投资组合的风险和收益至关重要。这样的分析必须经常进行，以优化投资组合在下一个交易日的特定方向。

最后，股市是波动的，没有一个或多个模型可以准确预测股市的变化，即使是最有力的模型也面临数据噪音和信息差异问题。然而，我们在投资时可以做的是缩小信息差距，用学到的金融理论来分析股市的变化，用强大的金融模型和机器学习模型来帮助我们做出选择，而不是依赖模型。

## 未来的影响

这个投资项目让我们实践了每个金融模型，并尝试使用机器学习算法优化金融模型。此外，对投资组合和股票市场的分析使我们能够了解投资组合的相关性能指标。我们相信，这个项目将大大影响我们在未来投资时的决策。

在未来，一些模型可以帮助我们与现有的模型相结合，在投资过程中获得更有说服力的预测。例如，在这个项目中，我们只限于分析股票市场数据，而没有分析基本面，如新闻和社交媒体。我们可以通过引入NLP模型来分析主流新闻和社交媒体对该公司未来的讨论，以预测公众对该股票的预期。这在有重大新闻（如新产品发布、政策变化）时尤为重要，可以帮助投资者快速做出决策。同时，在使用上述模型后，算法交易可以用于快速交易，避免了人工交易的延迟和错误概率。

# 参考文献

"苹果股价收盘时创下历史新高"。2023.AppleInsider.2023年6月12日。https://appleinsider.com/articles/23/06/12/apple-share-price-closes-at-an-all-time-record-high-thanks-to-vision-pro-speculation。

Bekena, Sisay Menji.2017."使用决策树分类器来预测收入水平"。

Breiman, Leo.2001."随机森林"。*机器学习* 45 (1)：5–32. https://doi.org/10.1023/A:1010933404324.

"CAPM：Assumptions and Limitations | Securities | Financial Economics." n.d. Accessed July 1st, 2023. https://www.economicsdiscussion.net/portfolio-management/capm/capm-assumptions-and-limitations-securities-financial-economics/29904.

CNBC. n.d. "US TREASURY-CURRENT 10 YEAR's Stock Price (US10Y) in Real Time"。CNBC。2023年7月1日访问。https://www.cnbc.com/quotes/US10Y。

爱德华兹，本杰明，2023年。"Nvidia的新怪物CPU+GPU芯片可能为下一代AI聊天机器人提供动力。"Ars Technica.2023年6月8日。https://arstechnica.com/information-technology/2023/06/nvidias-new-ai-superchip-combines-cpu-and-gpu-to-train-monster-ai-systems/。

"个人投资者的投资政策声明的要素"， n.d. 2023年7月1日访问。https://www.cfainstitute.org/en/advocacy/policy-positions/elements-of-an-investment-policy-statement-for-individual-investors。

霍利斯特，肖恩。2023."Nvidia RTX 4060评论综述：你应该得到更好的。"The Verge.2023年6月29日。https://www.theverge.com/23777753/nvidia-rtx-4060-review-roundup。

"介绍苹果Vision Pro：苹果公司的第一台空间计算机--苹果。" n.d. 2023年7月3日访问。https://www.apple.com/newsroom/2023/06/introducing-apple-vision-pro/。

"Introducing ChatGPT." n.d. Accessed July 3rd, 2023. https://openai.com/blog/chatgpt.

*投资*。2023年。https://www.mheducation.com/highered/product/investments-bodie-kane/M9781264412662.html。

"Kenneth R. French - Data Library." n.d. Accessed July 3rd, 2023. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\_library.html.

Luca, Gabriele De.2020."SVM Vs Neural Network | Baeldung on Computer Science."2020年8月6日。https://www.baeldung.com/cs/svm-vs-neural-network。

"Mean-Variance Optimization - PyPortfolioOpt 1.5.4 Documentation." n.d. Accessed July 1st, 2023. https://pyportfolioopt.readthedocs.io/en/latest/MeanVariance.html。

梅拉诺，玛丽亚。2023."特斯拉对Model Y和Model 3库存单位进行降价。"*TESLARATI*（博客）。2023年6月1日。https://www.teslarati.com/tesla-model-y-model-3-inventory-price-decrease/。

Movement, Q. ai-Powering a Personal Wealth. n.d. "微软股票随着下一代ChatGPT-4的发布而反弹。"福布斯》。2023年7月3日访问。https://www.forbes.com/sites/qai/2023/03/16/microsoft-stock-rallies-as-next-generation-chatgpt-4-is-released/。

"现代投资组合理论中的投资组合优化。" n.d. 2023年7月1日访问。https://developers.refinitiv.com/content/devportal/en\_us/article-catalog/article/portfolio-optimization-modern-portfolio-theory.html。

"TWS API v9.72+：Placing Orders." n.d. Accessed July 2nd, 2023. https://interactivebrokers.github.io/tws-api/order\_submission.html.

"TWS API v9.72+：Positions." n.d. Accessed July 2nd, 2023. https://interactivebrokers.github.io/tws-api/positions.html.

威廉姆斯，奥利弗。2006."布林线盈利能力的实证优化"。SSRN学术论文。罗切斯特，纽约。https://doi.org/10.2139/ssrn.2321140。

郑宇浩，李欣怡，和冯元军。2022."基于布林带策略和多项式回归模型的量化交易策略研究"。In *2022 IEEE 2nd International Conference on Data Science and Computer Application (ICDSCA)*, 1255-60. https://doi.org/10.1109/ICDSCA56264.2022.9988398.

Zhu, Julie.2023."埃隆-马斯克在北京会见了中国副总理丁列明。"*路透社*，2023年6月1日，秒。业务。https://www.reuters.com/business/elon-musk-met-with-chinese-vice-premier-ding-beijing-source-2023-06-01/。

Zou, Hui, and Trevor Hastie.2005."通过弹性网的正则化和变量选择"。*皇家统计学会杂志。B系列（统计方法学）*67（2）：301-20.