



SEABRIDGE FINTECH

SEABRIDGE AI METHODOLOGY WHITEPAPER



90.95

28.87

72.22

52.82

目录

03 机器学习 - 大数据最优解

04 方法论

05 历史数据回测

06 实盘数据展示

07 结论

机器学习-大数据最优解

Seabridge AI 如何利用机器学习进行更好的投资

技术分析、价值投资、因子模型、均值回归、趋势跟踪、统计套利等，许多流派和技术被用在金融市场。但由于市场的动态变化、非平稳以及混沌状态导致现代投资和资产管理在市场的竞争中保持领先变得越来越具有挑战性。

近年来，随着大数据的出现和新算法的进步，探索数据集之间复杂、细微、动态非线性的关系变得更加容易。这个问题的复杂性和变化性超出了传统的统计分析的解决范畴，但新的解决思路已经在计算机算力大幅提升和大数据人工智能的时代背景下蓬勃发展。新人工智能在这个时代背景下的惊人成就为我们带来了投机和投资组合管理的新视角。

Seabridge Fintech 的创始团队成员曾就职于加拿大某投资机构工作，设计并构建了关键任务机器学习系统，其处理数亿个数据点并以生成最佳结果以达成核心分析预测和排名。系统通过最先进的 AI 和机器学习方法吸收大量复杂数据以生成捕获金融市场中动态和非线性模式的 alpha 信号，结合技术和动量分析从而创造了核心性优势。

我们称之为大数据最优解投资法：具有自适应算法的数据为投资提供了一种新的解决方案来帮助广大的投资者。

虽然金融市场每时每刻都在变化，所记录和观察到的数据也是如此。但与传统的统计或定量模型不同的是我们的算法系统设计用于处理这种不断变化发展的数据集。大家都知道线性模型的局限性，尤其是在做出决定之前需要考虑和权衡成百上千的因数。但机器学习，尤其是深度学习便可以通过在不同的数据集中学习数据特征，并发现极其复杂的非线性模式来完成这种任务。这也便是例如谷歌、Facebook 和亚马逊能够在如此短的时间内成倍地扩展业务并颠覆传统巨头的核心秘密。

一些人可能会问，像 AlphaGo 或自动驾驶汽车，大数据最优解投资会打败还是取代当下的人类专家以及散户投资人？对于这个问题，我们倾向于采取更加混合的观点：机器算法更像是副驾驶、投资顾问或者经理，而人类投资者仍处于主导地位。为了阐述我们的观点，我们从构建 Seabridge AI 核心算法的经验中学到的以下三条原则将有所帮助：

- 1，投资交易是这个过程的核心，而不是机器，人类起着至关重要的作用；
- 2，算法投资只有在数据表达正确和人类洞察力指导编码后内才能变成所谓的“高级智慧”，而其“智慧”便是建立在大数据的支持和计算机的算力上；
- 3，在正确的应用环境中时千万不要低估高级人工智能算法的优势。

Seabridge AI 方法论

由于近年来许多定量因子模型的盈利能力下降，市场从业者正在寻找动态和因子时序模型进行选股。很少有人传统的定量方法中取得成功，线性回归模型很难捕捉预测变量和股票收益之间的动态关系。

与当下流行的计量不同的是，Seabridge AI 是机器学习和交易系统驱动的股票选择，预测模型。简单来说就是价格行为与机器学习，深度学习的完美配合。

Seabridge AI 利用大数据和强大的学习算法结合配合着许多价格行为分析方法来寻找最佳信号。例如多因子选股，图形识别，主成分回归、弹性网络回归、随机森林、梯度提升树和神经网络模型等。将许多弱信息源组合成一个强复合信号来识别隐藏的非线性关系。并采用的这种集成方法在这种嘈杂金融数据的低信噪比环境中取得惊人的成功。

为了在数据中找到不仅在过去有效而且从样本中泛化的模式，我们将核心算法应用于更多数据源，有些甚至通常被人工分析师视为极弱信息或噪音，这些无法在人的审查下做出任何的判断，却可以在机器算法分析中取得重要的信号，其 AI 系统引擎可以在处理大量数据的同时识别出超越传统定量方法的复杂信息。我们致力于对 AI 的不断开发，训练和完善，采用先进的投资分析理论，建立完整的风险控制体系和交易系统，利用大数据和算力为用户提升投资交易体验以及获得更高的回报。

当下所使用的一些分析策略组成部分包括：

财务报表信息，包括销售额，市值，净收入，股息，EPS等，

比率包括资本与长期债务、ROA 和 ROE等，

自主研发的公司基本面及盈利能力分数体系，

机器学习自主优化的技术指标包括 最佳均线，倾斜资金流，成交量，相关性相对强弱等

根据价格行为与成交量等构建出动量分析系统

价格趋势逻辑，包括图形识别和蜡烛图组合信息

上述因素的时间序列以及与价格相关性

市场和投资者情绪信息来自替代数据的信息，例如 SEC 报告、股民社区、访问量、电子商务网站的销售等

据数据统计，股票的价格走势只有30%-40%处于趋势运动，而60%-70%则处于震荡。根据此基本原则可推出交易分析里绝不存在“一招鲜吃遍天”，单一策略掌控全市场行情则显为悖论。所以 Seabridge AI 的策略设计根据价格趋势原则进行了区分。我们在基础的趋势和震荡之外还额外添加了动量策略。

数据来自多个数据供应商以提高完整性和质量，在全部基本数据财务报表信息和比率粗筛后获取由 2000 多只基本数据符合要求的个股组成的第一股池，再通过自主研发的 Cashbridge 基本面评分体系，过滤出 700 以内数量的第二股池。Cashbridge 结合了动态学习因子与未来回报之间的变化关系针对给定时间范围内的当前影响项去提取预测和分析数千因子，我们的算法将这一长长的信号列表从数千个减少到几十个。我们通过关注具有高信噪比的因子并避免过度拟合（机器学习中最具挑战性的任务之一）来构建具有更符合当下市场动态的评分系统。每只股票的评级在 1 到 9 之间，分配给股票的得分越高 (9) 表示未来几个月表现优异的可能性越大，而得分 (1) 越低表示表现优异的可能性越低。

在第二股池开始以三条主策略运行对其个股进行分析和价格节点计算及回测对比，并配合我们当下所监控统计得到的实盘目标价触发率以搭配合适的风险回报比例（关键）以当下最优解去对个股做出交易决策。（每日运算）

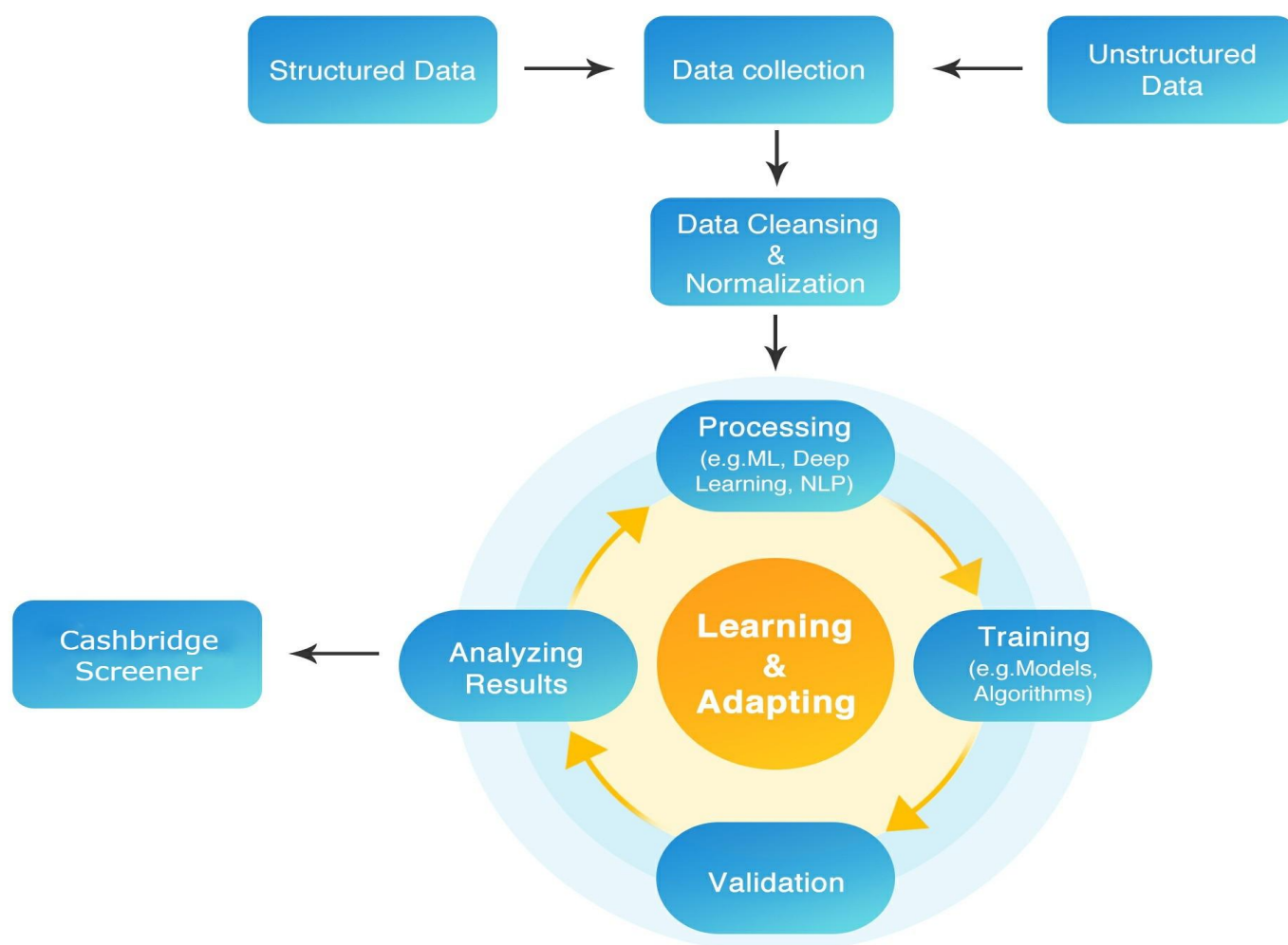
紫色Power Trend策略：以特征识别价格反转所构成的不同规则运动的节点与自生成的多均线系统结合。对股池中提取的公司进行识别聚类后为每个聚类创建随机其森林模型推演，使用最优解对应集群预测，再同理套用在其他自研自生的指标系统进行层叠筛选。同时在自研均线角度和量价的关系排列考虑支撑阻力择优过滤出强势趋势概率最大的股票。

绿色策略：通过使用动量和波动为主要因数，结合紫色power trend, 但去掉对趋势强度的定义。此策略具有4种微调算法(包括Squeeze) 来训练最佳策略模型以选出最具有交易价值的个股。

蓝色策略：弱化紫色策略中的趋势性，并且强化了对价格节点的重要性以及支撑阻力的更长时间窗口的参数，结合上下趋势构成的支撑阻力来寻找高概率突破或反弹的价格范围，配合量能的变化找出当下机会。

将三大策略结果与图形机器视觉识别匹配，并过滤出满足其他附加条件的股票。并利用过去10年数据在本个股上回测策略。

在更高层次上，Seabridge AI引擎设定可以从结构化和非结构化数据（如财务比率、内幕交易、分析师意见、筹码、财务报告、新闻和社交媒体等）中获取信号和因子，并利用机器学习算法来提取其定量信号。这高度模拟了人类的思维方式并在打开类人智慧的前提下附加了几乎无限的算力和大数据库的支持。

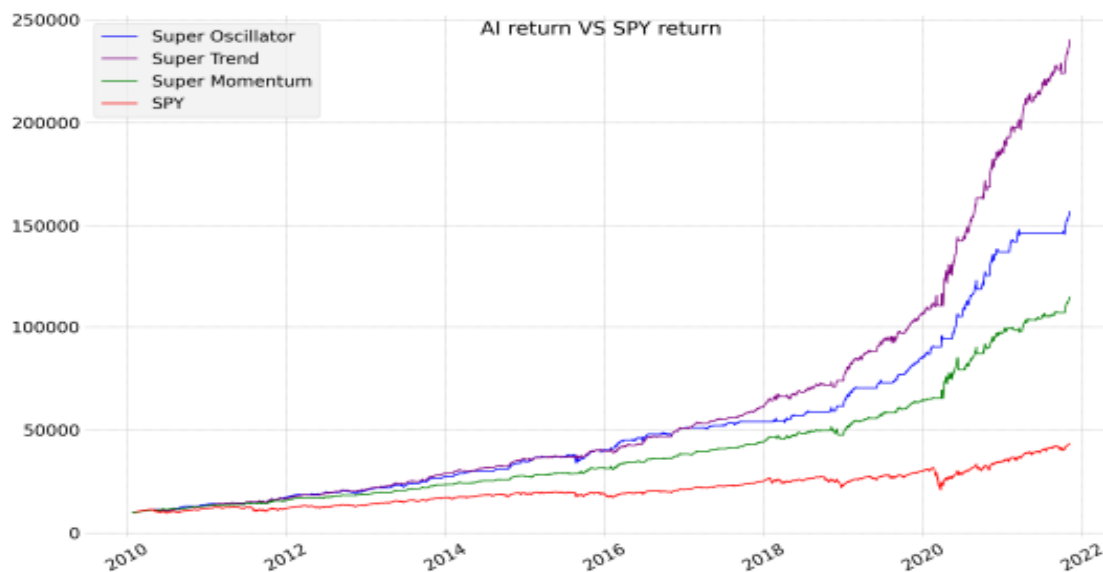


回测结果和投资组合展示

这是策略自构建成功后的10年回测结果及实盘操作中的投资组合结果展示。

回溯测试结果: 将三大策略分别作了测试于所有股票, 过去11年历史数据检测。并在app上完整展示。下面是标普500ETF SPY的三大策略回测结果。

图1. 标普500 VS 三大策略累计收益率(2010/1/4-2021/11/4)



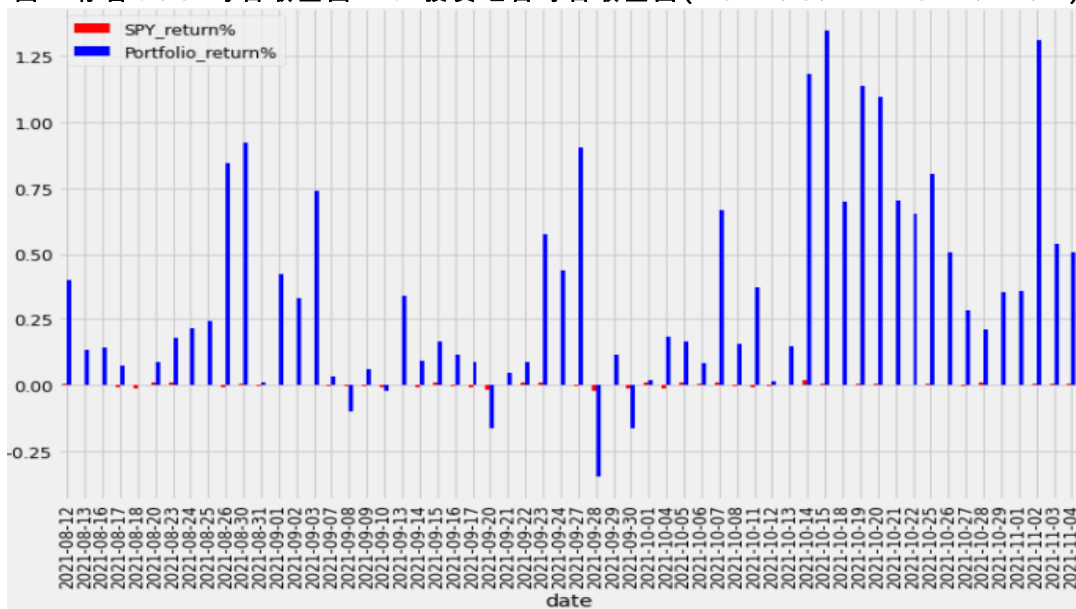
回测结果参数对比

?	SPY	Purple Line	Blue Line	Green Line
CAGR	18.94%	46.73%	39.85%	34.14%
Vol.	20.36%	12.48%	11.04%	11.52%
Sharpe Ratio	0.81	3.55	3.38	3.55
Max. DrawDown	34.10%	5.92%	11.07%	7.70%
Sortino Ratio	34.10%	7.83	7.33	5.62

回测结果和投资组合展示

下面是标普实盘中投资组合的表现。起初因未买入，没有资金量的限制，所以设置等权来计算组合买入量。

图2. 标普500 每日收益图 VS 投资组合每日收益图(2021/8/12-2021/11/4)



总结

Seabridge AI 利用和开发最先进的人工智能和机器学习 (ML) 引擎，将大量复杂数据转化为有用的投资情报。我们的 AI/ML 方法提供了一个框架，该框架不仅有条不紊且功能强大，而且具有适应性和可扩展性。我们将不断优化算法，帮助我们客户在投资方面取得长期的成功。