

سرمایه گذاری در سهام

دریافت سود مناسب با توجه به ریسک



حسین مستان

پوريا حیدریان

امین ذوالفقاریان

استاد درس: استاد باباعلی

درس: حسابگری زیستی

ترم بهار 1402

مسئله‌ی سرمایه گذاری در سهام و بررسی سود یکی از مسائل پر اهمیت در زمینه‌ی اقتصاد است.

حل این مسئله به دلیل توابع غیرخطی، عوامل تاثیرگذار فراوان و محدودیت‌های مختلف در توابع با

روش‌های سنتی امکان پذیر نیست.

به همین دلیل می‌خواهیم به کمک الگوریتم‌های مختلف زیستی تلاش بر حل این مسئله داشته باشیم و عملکرد الگوریتم‌های مختلف (SA, PSO, GA) را در مقایسه با یکدیگر بررسی کنیم.

صورت مسئله

فرض کنید اطلاعات قیمت چندین سهام در یک سال اخیر را در دست داشته باشیم. حال می‌خواهیم با استفاده از این اطلاعات تشخیص بدهیم که چگونه سرمایه‌ی خود را در بین این سهام‌ها پخش کنیم و در آنها سرمایه‌گذاری کنیم تا بیشترین مقدار سود به همراه کمترین مقدار ریسک را داشته باشیم.

تعاریف، ارائه‌ی روش حل

یکی از اصول پایه‌ای سرمایه‌گذاری، پخش کردن سرمایه بین سهام‌های مختلف است. به این ترتیب میزان ریسک کاهش میابد زیرا دیگر فقط وابسته به یک سهام نیستیم.

یک راه حل p را یک آرایه به اندازه‌ی تعداد سهام‌های بررسی شده در نظر می‌گیریم که جمع اعضای آرایه یک باشد. در این صورت می‌توانیم نسبت سرمایه‌گذاری در سهام‌های مختلف را متناظر با p در نظر بگیریم.

همچنین برای بررسی یک سرمایه‌گذاری از مدل Sharpe Ratio(SR) به عنوان Fitness Function استفاده می‌کنیم. که به شرح زیر است:

$$SR = \frac{R_p - R_f}{StdDev(p)}$$

که در آن R_p میانگین سود در سرمایه‌گذاری مورد نظر در p ، R_f حداقل میزان سود مورد نظر و $StdDev(p)$ میزان انحراف از معیار در p است. و p هم که یک جواب برای مسئله است.

حال مشخص است که این مدل در جهت حل مسئله است زیرا R_p میزان میانگین سود است که حداکثر کردن آن از اهداف مسئله است. و میزان انحراف از معیار هم همان نشان دهنده‌ی ریسک است زیرا همانطور که قبلاً اشاره کردیم باید سرمایه‌ی خود را در سهام‌های مختلف پخش کنیم تا ریسک را حداقل کنیم.

نحوه‌ی سرمایه‌گذاری ما در این سهام‌ها نیز میتواند به دو روش باشد. فرض کنیم که از هیچکدام از سهام‌ها نداریم و فقط میخواهیم سرمایه‌گذاری کنیم که در این صورت ضرایب داخل p باید همگی بزرگتر از صفر باشند. یا اینکه فرض کنیم از هرکدام از سهام‌ها مقداری داریم و برنامه به ما علاوه بر برنامه‌ی خرید چندین سهام، فروش آنها را نیز با ضرایب منفی توصیه میکند. این کار دقیقاً همان فروختن سهام‌های کم سود یا حتی مضر است.

الگوریتم‌ها:

الگوریتم Particle Swarm Optimization

کافیست ابتدا particle ها را تعیین کنیم که هرکدام یک آرایه به اندازه‌ی تعداد سهام‌ها هستند که نماینده‌ی یک جواب به ازای ضرایب سهام‌ها هستند.

این الگوریتم یک الگوریتم هوش جمعی است و هرمرحله جواب‌ها را به سمت بهترین جواب تا به حال دست‌یافته سوق میدهد.

برای particle های ابتدایی از مقادیر رندوم استفاده میکنیم. و البته برای اینکه جمع مقادیر هر particle یک بماند بعد از اضافه شدن velocities به جواب‌ها آنها اگر خارج از محدوده بودند یا جمع کمتر یا بیشتر از یک داشتند آنها را با حفظ نسبت به ناحیه مناسب میبریم به طوری که جمع یک داشته باشند.

الگوریتم Simulated Annealing

در این الگوریتم هم هر پاسخ یک آرایه در نظر گرفته میشود که آرایه‌ها همان وزن‌ها برای خرید سهام هستند.

الگوریتم Genetic

در این الگوریتم هر کروموزم یک آرایه از اعداد است که ضرایب سهام‌ها را نشان می‌دهد. که در ابتدا به صورت رندوم تولید شدند.

در مرحله‌ی انتخاب والد 30 درصد برتر در هر نسل به عنوان والد انتخاب میشوند.

برای جهش دو خانه از آرایه‌ی کروموزم را به صورت تصادفی با یکدیگر جابه‌جا میکنیم

برای ترکیب از blend crossover استفاده شده تا جواب‌ها بیشتر نزدیک به والد بهتر باشند.

همینطور از ترکیب arithmetic crossover نیز استفاده شده و نتیجه‌ی هردو را بررسی میکنیم.

برای ساخت نسل بعدی فقط از جمعیت انتخاب شده به عنوان والد استفاده میشود و تمامی اعضای نسل قبلی از بین می‌روند.

اعضای نسل بعدی یا از طریق ترکیب یا از طریق جهش و یا هردو تولید شده‌اند.

یک نکته که در تمامی الگوریتم‌ها یکسان است fitness function است.

همانطور که گفته شد از Sharpe Ratio استفاده شده با این تفاوت که برای محدود کردن تاثیر ریسک در خروجی نهایی، برای واریانس سود حد پایین قائل شدیم تا جواب‌های با ریسک بسیار پایین اما با سود نه چندان قابل قبول برازش بالایی نداشته باشند.

حال به سراغ دیتای مسئله می‌رویم. در نوت‌بوک الگوریتم Genetic بصورت مفصل و کامل نحوه بررسی دیتا و مدل کردن مسئله در کد توضیح داده شده و به همین دلیل در کدهای SA و PSO از این توضیحات صرف نظر شده. برای بررسی پارامترهای مختلف، ما مقادیر زیر را بررسی کردیم:

- R_f : برای مقدار حداقل سود مورد نظر، دو مقدار 0.667 را که بصورت تجربی دیدیم تعادل خوبی از ریسک و سود می‌دهد و 0.2 را که خود مقاله نیز بررسی کرده و یک هدف‌گذاری برای سود بالا است را در نظر گرفتیم.

- بازه وزن‌ها: بازه وزن‌ها را در دو حالت در نظر گرفتیم، حالتی که وزن‌ها بین 0 و 1 باشند که بدان معناست که ما تنها داریم خرید سهام انجام می‌دهیم و حالتی که بین -1 و 1 باشند که بدان

معناست که ما هم خرید انجام می‌دهیم و هم فروش داریم (مجموع وزن‌های هر پاسخ برابر 1 است).

- حد پایین واریانس سود: برای تعادل جوابها و جلوگیری از آنکه جواب‌های با سود پایین بخاطر ریسک بسیار پایین (مانند جوابی با سود 10 درصد اما ریسک 10^{-7} که از برآزش بسیار بالایی بهره‌مند است)، دو حد پایین برای واریانس در نظر گرفتیم. 10^{-3} و 10^{-4} که به ترتیب یعنی تاثیر ریسک تا یک‌دهم درصد و یک هزارم درصد است.

از ویژگی‌های الگوریتم‌ها نیز به این نتیجه رسیدیم که الگوریتم PSO چون یک الگوریتم جمعیت محور است که در آن ذرات به سوی یکدیگر حرکت می‌کنند، باعث کاهش چشمگیر ریسک و بهم نزدیک شدن وزن‌ها می‌شد و در حالتی که بدون محدودیت حد پایین واریانس سود بررسی می‌کردیم، PSO در مقایسه با SA و GA، جواب‌هایی با ریسک بسیار پایین‌تری و برآزش بسیار بالاتری تولید می‌کرد که البته در ازای سود پایین‌تر نیز بود.

در PSO جواب‌هایی همچون سود 10 درصدی با ریسک 10^{-7} و سود 7 درصدی با ریسک 10^{-9} هم مشاهده شد. ولی چنین اعدادی برای ریسک بی‌معنی است چرا که این اعداد عملاً به معنای ریسک صفر هستند در حالی که در دنیای واقعی هیچ معامله‌ای ریسک صفر ندارد و به همین دلیل تصمیم بر محدود کردن حد پایین واریانس سود (ریسک) گرفتیم.

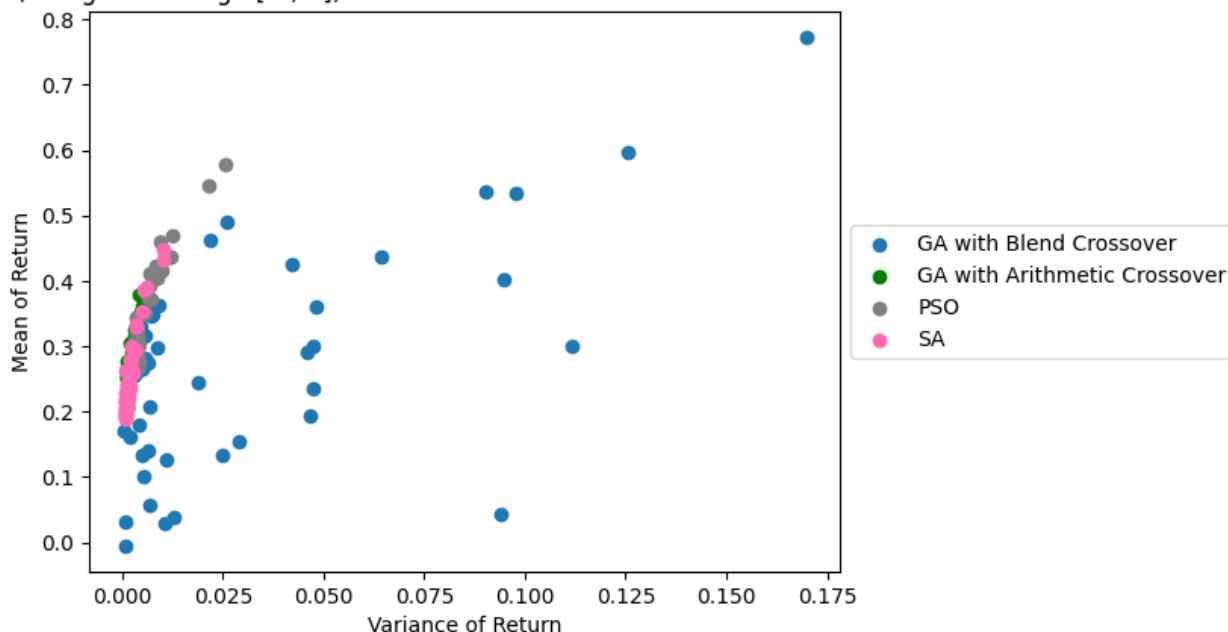
همچنین در میان SA و GA متوجه شدیم که بخاطر ماهیت الگوریتم‌ها، SA تمرکز بیشتری بر روی افزایش سود دارد و توانایی کمتری در کاهش ریسک دارد چرا که تمامی جواب‌های SA بصورت تصادفی ساخته می‌شوند و لزوماً به سمت مشابه هم شدن حرکت نمی‌کنند.

از طرفی در GA بخاطر جهش و باز ترکیب، دست بازتری برای آن داشتیم که هم تا حدودی بر روی کاهش ریسک متمرکز شویم و همچنین عامل تصادفی را در کد نگه داریم تا سود را بیشینه نگه داریم.

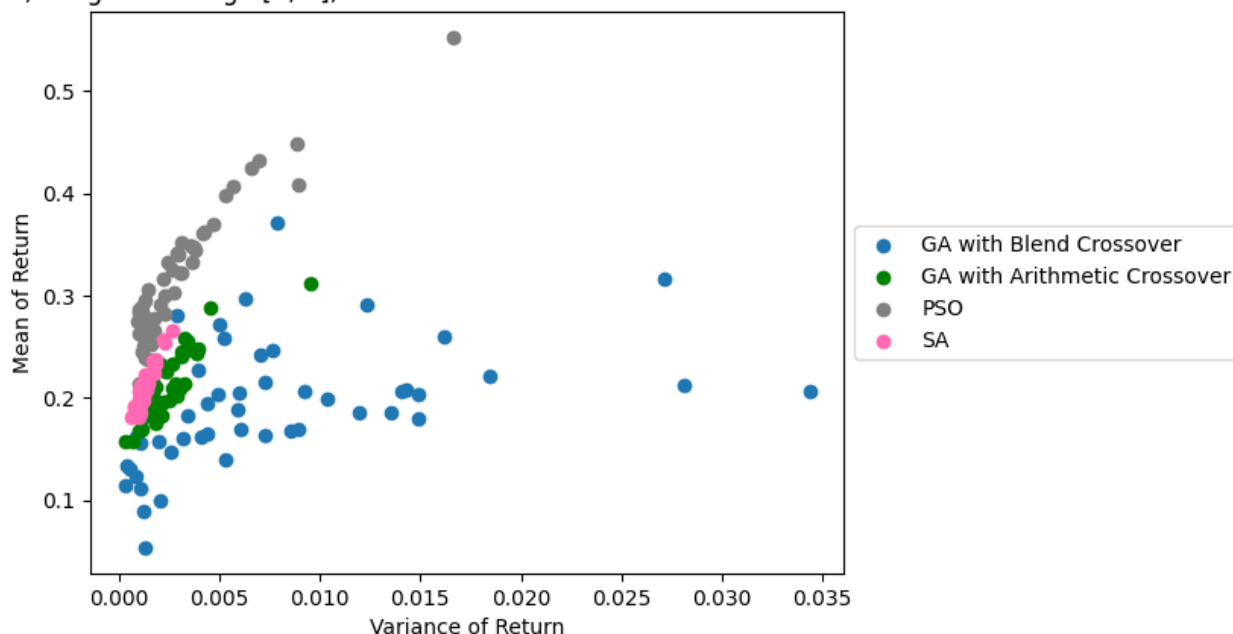
بطور کلی می‌توان گفت که SA بیشترین سود و بیشترین ریسک را داشت، PSO کمترین سود و در عین حال کمترین ریسک را داشت و GA تعادلی میان این دو معیار برقرار کرد.

در ادامه، نمودارهایی از مقایسه عملکرد این کدها بر روی داده‌های مسئله آورده شده است.

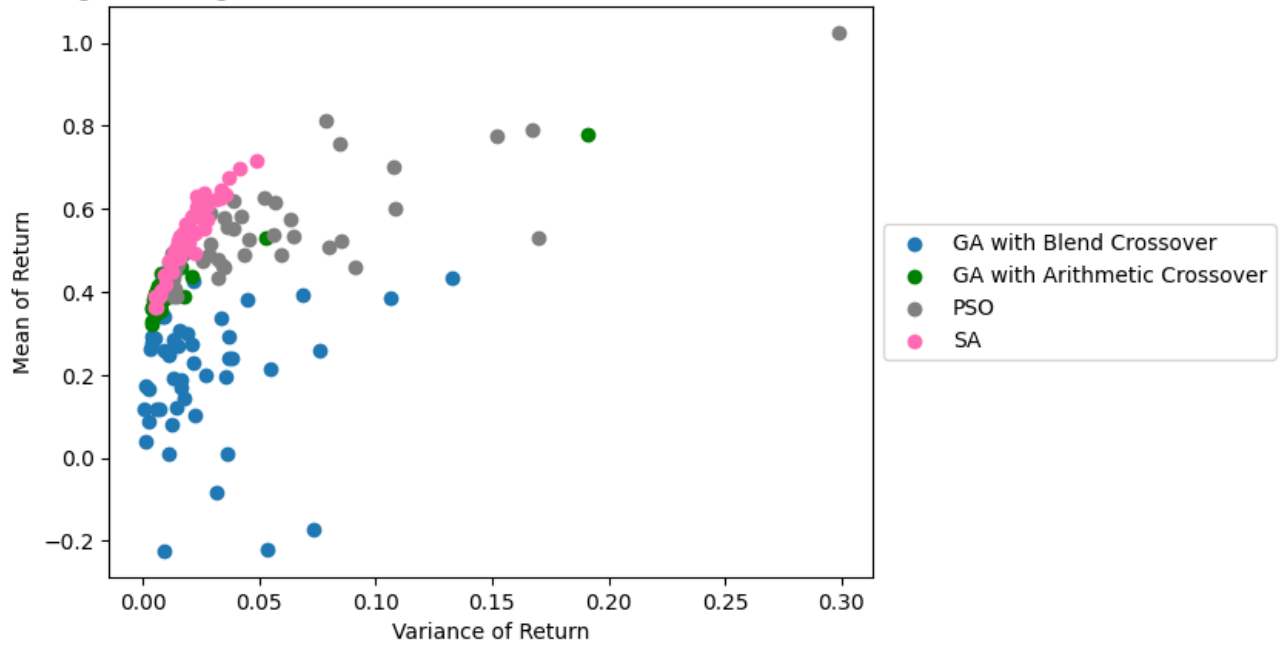
$R_f = 0.0667$, Weights in range $[-1, 1]$, Variance of Return lower bound of effectiveness = 0.001



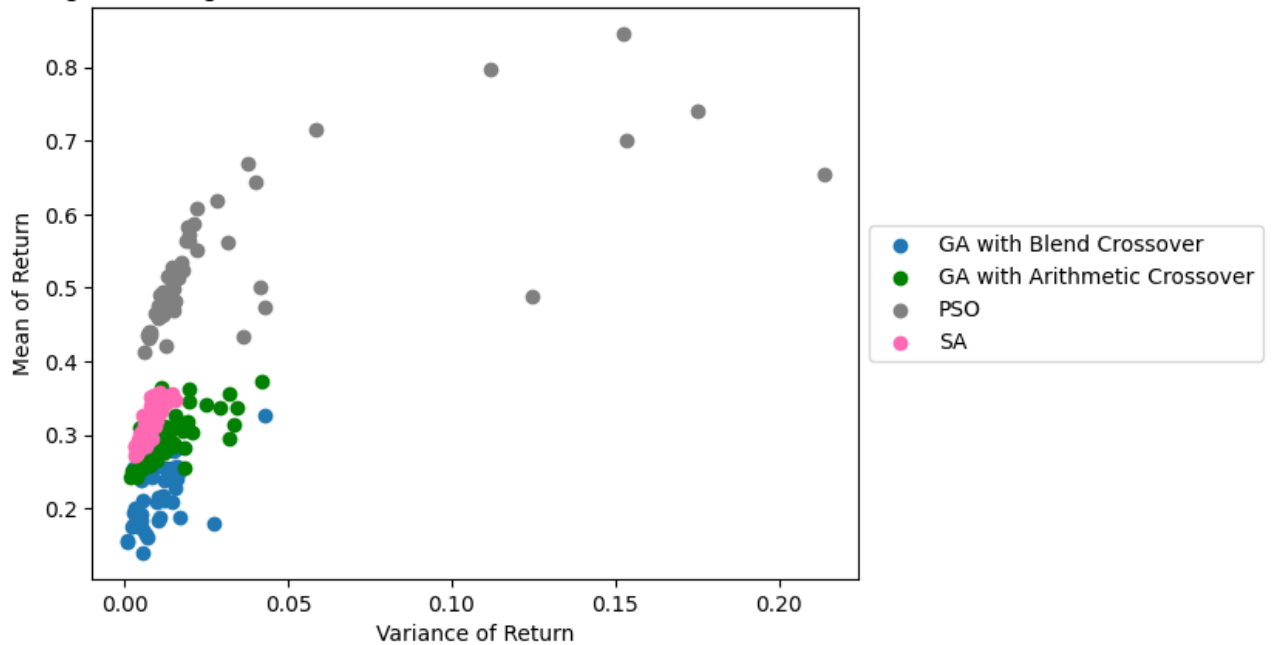
$R_f = 0.0667$, Weights in range $[0, 1]$, Variance of Return lower bound of effectiveness = 0.001



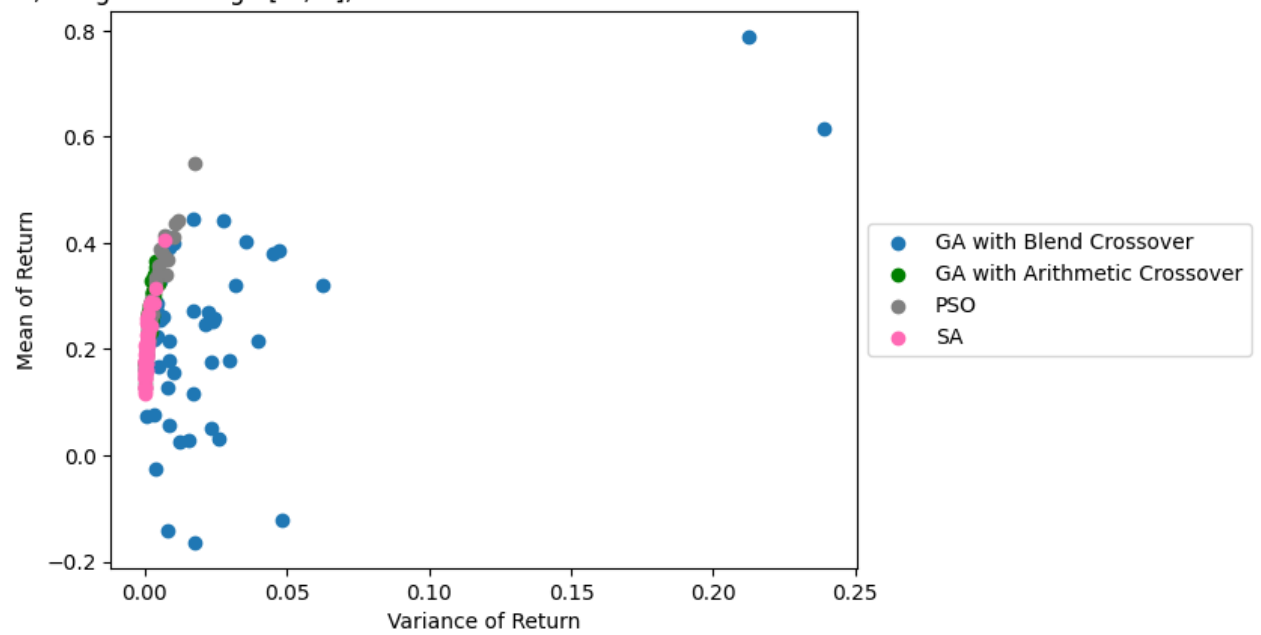
Rf = 0.2, Weights in range [-1, 1], Variance of Return lower bound of effectiveness = 0.001



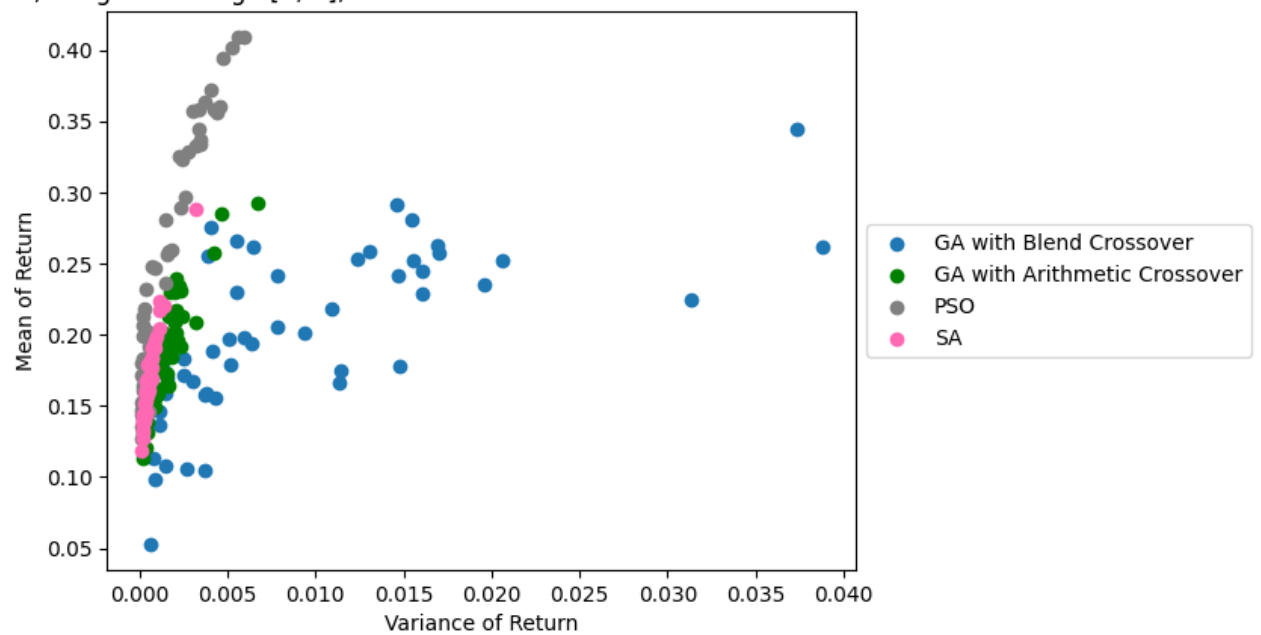
Rf = 0.2, Weights in range [0, 1], Variance of Return lower bound of effectiveness = 0.001



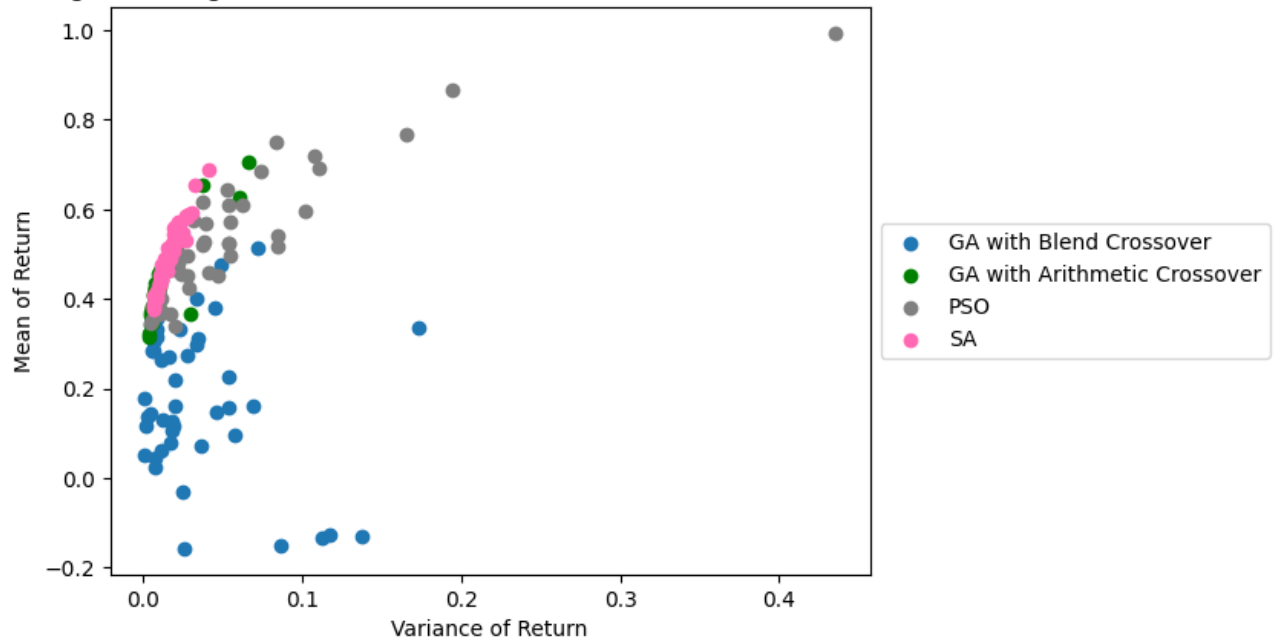
Rf = 0.0667, Weights in range [-1, 1], Variance of Return lower bound of effectiveness = 0.0001



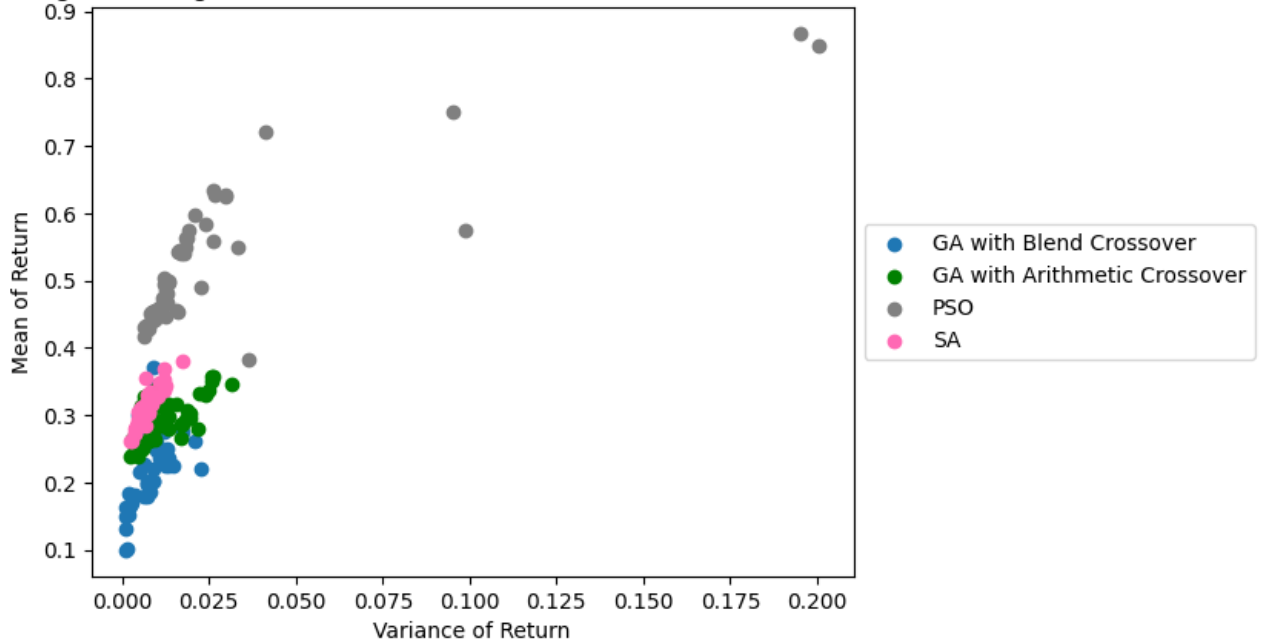
Rf = 0.0667, Weights in range [0, 1], Variance of Return lower bound of effectiveness = 0.0001



Rf = 0.2, Weights in range [-1, 1], Variance of Return lower bound of effectiveness = 0.0001

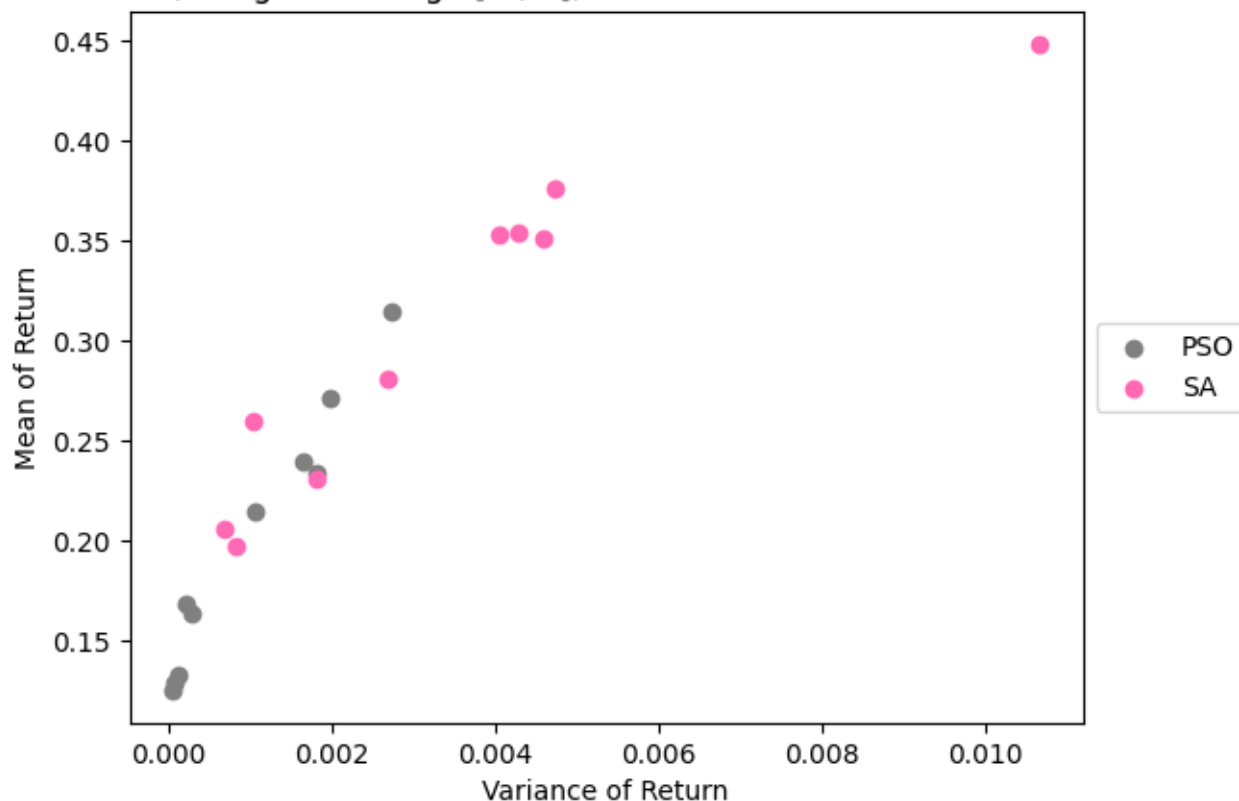


Rf = 0.2, Weights in range [0, 1], Variance of Return lower bound of effectiveness = 0.0001

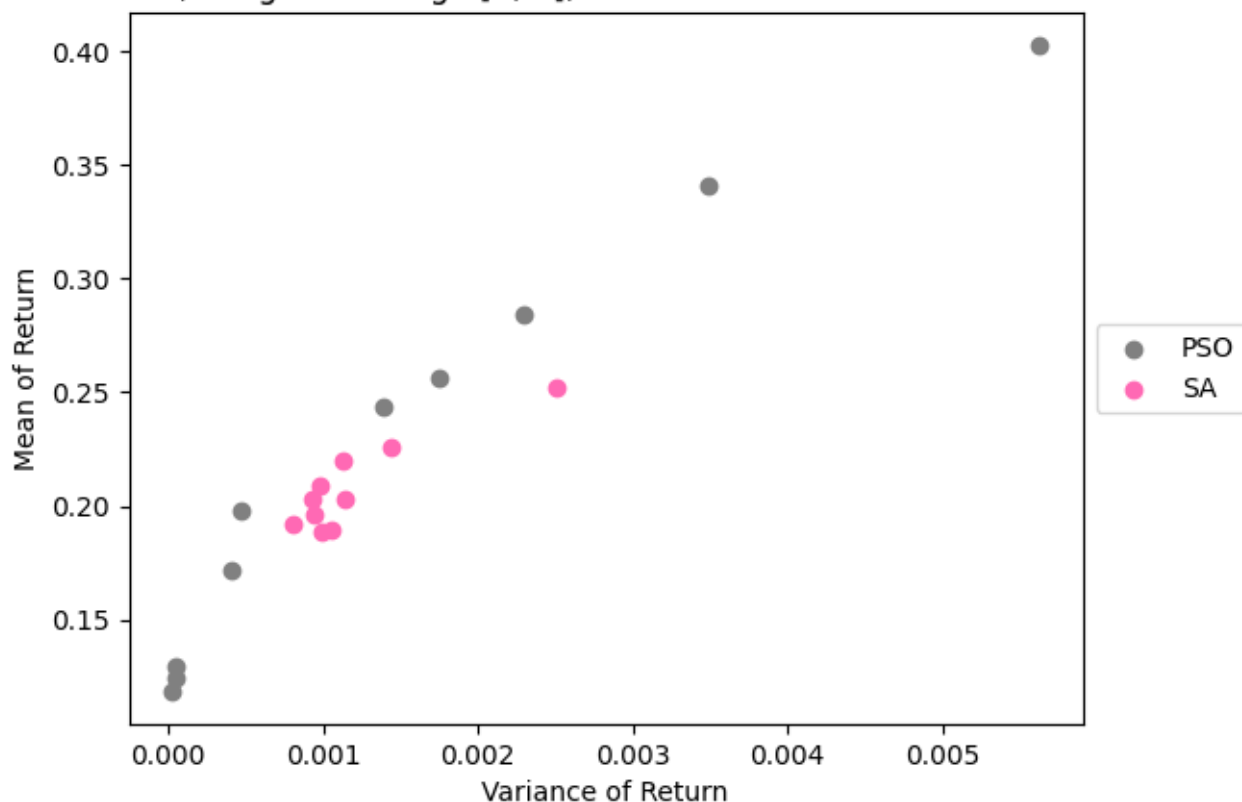


چهار نمودار زیر نشان می‌دهند که درحالی که واریانس را محدود نکنیم، به جز در حالت ضرایب مثبت و $R_f = 0.2$ ، SA بطور میانگین از PSO سود بیشتری دارد ولی ریسک بیشتری نیز دارد.

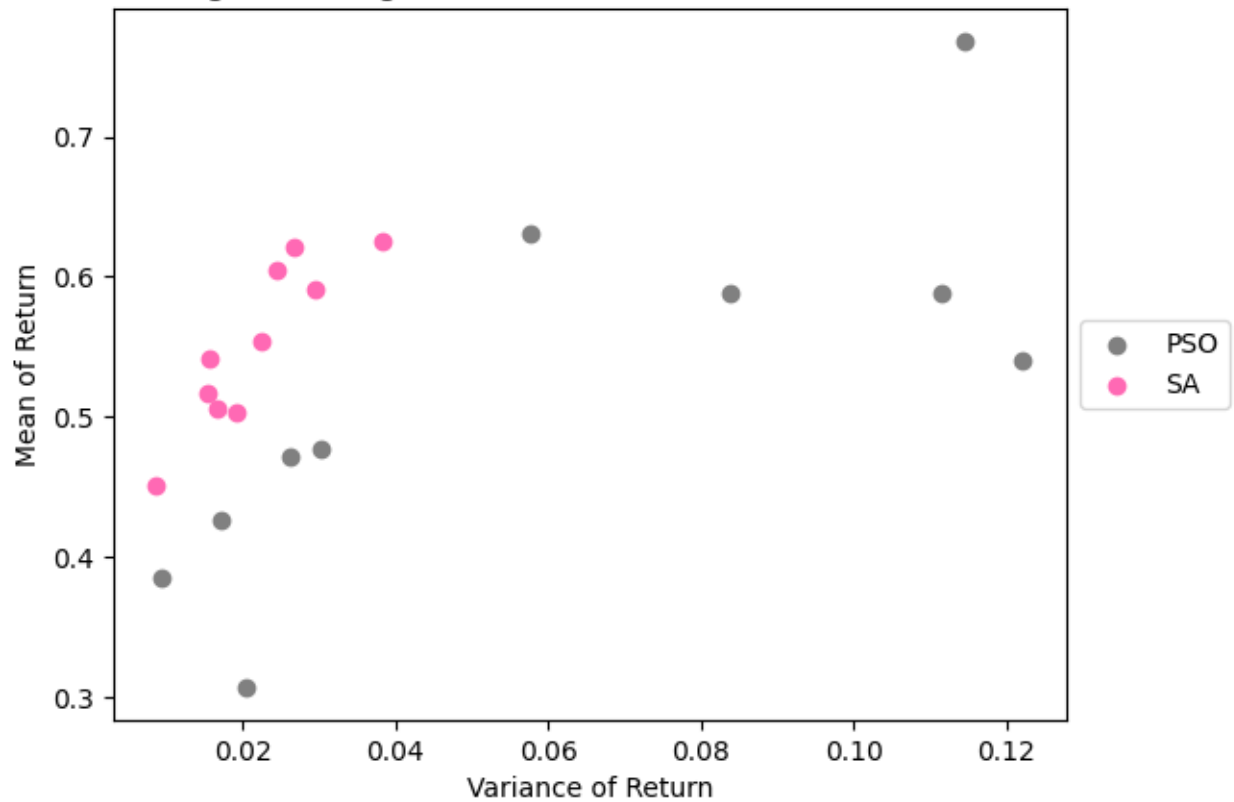
$R_f = 0.0667$, Weights in range $[-1, 1]$, Without Variance of Return lower bound



$R_f = 0.0667$, Weights in range $[0, 1]$, Without Variance of Return lower bound



$R_f = 0.2$, Weights in range $[-1, 1]$, Without Variance of Return lower bound



$R_f = 0.2$, Weights in range $[0, 1]$, Without Variance of Return lower bound

