



**بهینه‌سازی سبد سرمایه در بازار رمزارزها با استفاده از یادگیری ماشین و روش‌های خوشه‌بندی**

**علیرضا نژادشمسی**

**استاد راهنما**

**دکتر امیرعباس نجفی**

**پایان‌نامه‌ی کارشناسی ارشد**

**رشته‌ی مهندسی صنایع گرایش مهندسی مالی**

**. ۱۴۰۲**

**چکیده**

.

**واژگان کلیدی:** .

**فهرست مطالب**

**فهرست شکل‌ها**

**فهرست جدول‌ها**

فصل اول

# مقدمه و کلیات تحقیق

## ۱-۱- مقدمه

رمزارز به ارزهای دیجیتالی گفته می‌شود که در آن‌ها از تکنولوژی رمزنگاری استفاده شده و معمولاً به صورت نامتمرکز و بدون دخالت یک سازمان مرکزی عمل می‌کند. پس از آغاز به کار رمزارز بیت‌کوین در سال ۲۰۰۹ میلادی، علاقه‌مندان به فناوری‌های جدید به بررسی و استفاده از آن روی آوردند. مزیت‌ها و پتانسیل‌های موجود در این ارز نوظهور موجب شد تا ارزهای دیجیتال متعددی با کاربردهای جدید متولد شوند و توجه سرمایه‌گذاران در سرتاسر جهان را به خود جلب کند. با ایجاد بازار نوظهور رمزارزها، سرمایه‌گذاران باید سیاست‌های مناسبی را اتخاذ کنند تا بتوانند بازدهی بیشتری در سرمایه‌گذاری داشته باشند. در این پژوهش به بهینه‌سازی سبد دارایی برای افزایش بازده و کاهش ریسک در بازار رمزارزها، و با استفاده از روش‌های یادگیری ماشین پرداخته می‌شود.

## ۲-۱- هدف از پژوهش

یکی از مهم‌ترین مسائل در سرمایه‌گذاری، انتخاب یک پرتفوی مناسب از دارایی‌های مورد نیاز برای کسب بیشترین بازده با متحمل شدن کمترین میزان ریسک است. سرمایه‌گذاران بسته به میزان ریسک‌پذیری‌شان، می‌توانند از روش‌های مختلفی برای توزیع دارایی‌ها در سبد سرمایه‌گذاری خود استفاده کنند که منجر به کسب بازدهی‌های متفاوتی می‌شود. هدف از این پژوهش، ارائه‌ی روشی مبتنی بر یادگیری ماشین برای بهینه‌سازی سبد سرمایه‌گذاری و یافتن بهترین رویکرد در بازار رمزارزهاست تا به سرمایه‌گذاران این بازار نوظهور کمک کند تا با توجه به شرایط خاص آن، بهترین سبد از ارزهای دیجیتال موجود را ایجاد کنند.

## ۳-۱- توضیح موضوع پژوهش

سرمایه‌گذاران در بازارهای مختلف در معرض انواع ریسک‌های سیستماتیک و غیرسیستماتیک قرار دارند. با توجه به بالا بودن میزان ریسک سیستماتیک در بازار رمزارزها به علت نوظهور بودن آن و موانع قانونی، سرمایه‌گذاران می‌بایست تا حد امکان اقدام به کاهش ریسک‌های غیرسیستماتیک کنند که به وسیله‌ی تنوع‌بخشی[[1]](#footnote-1) سبد دارایی انجام می‌شود. بنابراین این سرمایه‌گذاران نیاز دارند که پس از تعیین سیاست‌های سرمایه‌گذاری و تجزیه‌وتحلیل رمزارزها، اقدام به تشکیل یک سبد سرمایه‌گذاری مناسب از آن‌ها به وسیله‌ی وزن‌دهی به هر دارایی کنند. برای این کار روش‌های متعددی وجود دارد. استفاده از روش‌های نوین مانند یادگیری ماشین در فرآیند انتخاب سبد سرمایه‌گذاری می‌تواند به بهبود نتایج کمک کند.

## ۴-۱- توجیه، انگیزه و علت انتخاب موضوع

رمزارزها نسبت به ارزهای رایج مزایای بالقوه و باارزشی دارند که آن‌ها را به دارایی‌های ارزشمندی برای سرمایه‌گذاری تبدیل می‌کند. بیشتر رمزارزها نامتمرکز و فاقد یک ساختار مرکزی هستند. این مسئله به این معناست که بر خلاف ارزهای رایج که توسط دولت‌ها و بانک‌های مرکزی کنترل می‌شوند، رمزارزها قابل کنترل توسط هیچ سازمان یا نهادی نیستند و سیاست‌گذاری آن‌ها تنها در پروتکلی انجام می‌شود که در ابتدا توسط سازنده یا سازندگان آن تعریف و برنامه‌نویسی می‌شود. بنابراین در رمزارزها بر خلاف ارزهای رایج، امکان دستکاری میزان عرضه‌ی پول توسط یک نهاد خاص وجود ندارد.

مزیت دیگر رمزارزها شفافیت آن‌هاست که فناوری بلاک‌چین[[2]](#footnote-2) آن را ممکن کرده است. در این فناوری که در اکثر رمزارزهای موجود در بازار از آن استفاده می‌شود، تراکنش‌های همه‌ی افراد در بلاک‌های متصل به یکدیگر ثبت می‌شود که قابل مشاهده برای همه‌ی افراد است. این میزان از شفافیت در هیچ یک از سیستم‌های مالی کنونی یافت نمی‌شود. همچنین متن‌باز بودن این ارزهای دیجیتال علاوه بر بالا بردن میزان شفافیت آن‌ها، امکان استفاده‌ی افراد به صورت ناشناس از شبکه را فراهم می‌کند.

با توجه به مزیت‌های متعدد رمزارزها، رشد زیادی برای آن‌ها در آینده پیش‌بینی می‌شود و از این رو،‌ افراد زیادی از جمله سرمایه‌گذاران بازارهای سنتی، ارزهای دیجیتال را به ترکیب سبد دارایی‌های خود افزوده اند. بنابراین لازم است بیش از پیش به انتخاب و بهینه‌سازی سبد دارایی در این بازار نوظهور توجه شود و ویژگی‌های خاص آن مد نظر پژوهشگران قرار گیرد.

## ۵-۱- اهمیت موضوع

رمزارزها در کنار مزایای ذکرشده دارای خطرات بالقوه‌ای هم هستند. با توجه به حذف نقش دولت‌ها در این اکوسیستم، همواره موانع قانونی زیادی در برابر استفاده از این ارزهای نوظهور وجود داشته است. همچنین به دلیل ویژگی ناشناس بودن استفاده‌کنندگان بسیاری از رمزارزها و امکان استفاده برای مقاصد غیرقانونی، چالش‌های زیادی در مسائل قانونی مربوط به دولت‌ها وجود دارد. علاوه بر این، هک شدن یک شبکه‌ی خاص و از دست رفتن دارایی‌های سرمایه‌گذاران از خطرات دیگری است که دارندگان رمزارزها را تهدید می‌کند. با توجه به خطرات این بازار و رشدهای سریع و حباب‌گونه، سقوط‌های شدیدی هم پیش روی سرمایه‌گذاران به وجود آمده است که ضرورتِ داشتن استراتژی‌های معاملاتی و بهینه‌سازی سبد دارایی‌ها را دوچندان می‌کند. بنابراین لازم است علاوه بر شناخت رمزارزها و پیدا کردن پروژه‌های ارزشمند، ترکیب آن‌ها در پرتفوی دارایی‌ها به گونه‌ای چیده شود که کمترین میزان ریسک را متوجه سرمایه‌گذاران کند. این پژوهش تلاش می‌کند با کمک روش‌های یادگیری ماشین همچون خوشه‌بندی و رگرسیون در مراحل مختلف چینش سبد سرمایه‌گذاری، این فرایند را بهبود بخشد.

## ۶-۱- مرور کلی بر ادبیات موضوع

پس از ایجاد بازار رمزارزها، پژوهشگران در بسیاری از مقالات مرتبط با بهینه‌سازی سبد سرمایه‌گذاری، تأثیر افزودن رمزارزها به سبد دارایی سرمایه‌گذاران بازارهایی مثل سهام، ارز، کالا و... را مورد بررسی قرار دادند. بسیاری از این پژوهشگران مانند برییر[[3]](#footnote-3) و همکاران (۲۰۱۵) و گسمی[[4]](#footnote-4) و همکاران (۲۰۱۸) اضافه کردن بیت‌کوین به سبد سرمایه‌گذاران را مورد مطالعه قرار دادند؛ اما تعدادی از آن‌ها مانند دمیرالای[[5]](#footnote-5) و بایراسی[[6]](#footnote-6) (۲۰۲۱)، تنها به بیت‌کوین بسنده نکردند و افزودن رمزارزهای دیگر به سبد دارایی‌های سنتی را نیز بررسی کردند.

با افزایش کاربرد رمزارزها و تکنولوژی آن، بازار بزرگی از ارزهای دیجیتال شکل گرفت و پژوهشگران بسیاری مانند امبا[[7]](#footnote-7) و همکاران (۲۰۱۸)، آقامحمدی و همکاران (۱۳۹۹)، براونیس[[8]](#footnote-8) و مستل[[9]](#footnote-9) (۲۰۱۹) و پلاتاناکیس[[10]](#footnote-10) و همکاران (۲۰۱۸) روش‌های انتخاب سبد رمزارز بدون حضور سایر دارایی‌ها را مورد بررسی قرار دادند. همچنین پژوهشگرانی مانند لیو[[11]](#footnote-11) (۲۰۱۹) و اسچلینجر[[12]](#footnote-12) (۲۰۲۰) نیاز مربوط به مقایسه‌ی روش‌های متنوع انتخاب و بهینه‌سازی سبد سرمایه‌گذاری را برطرف کردند.

## ۷-۱- جنبه‌های جدید بودن موضوع

بازار رمزارزها دارای قدمت بسیار کمی است. نخستین رمزارز نامتمرکز بیت‌کوین است که فعالیت آن از سال ۲۰۰۹ شروع شده است. در سال ۲۰۱۱ رمزارز‌های دیگری مانند لایت‌کوین[[13]](#footnote-13) پا به عرصه‌ی وجود گذاشتند و پس از آن بود که به مرور بازاری از ارزهای رمزنگاری‌شده شامل اتریوم[[14]](#footnote-14)، ریپل[[15]](#footnote-15)، دش[[16]](#footnote-16) و... ایجاد شد. این بازار در طول یک دهه فعالیت خود، فرازونشیب‌های زیادی داشته است.

پس از هر رشد سریع در قیمت رمزارزها، سرمایه‌گذاران جدیدی به فکر سرمایه‌گذاری در آن‌ها می‌افتادند و به این ترتیب ارزش کل بازار رمزارزها به مرور زمان بالاتر می‌رفت. سقوط‌های ناگهانی و خطرات بالقوه باعث شد سرمایه‌گذاران به فکر مدیریت دارایی‌های خود بیفتند و تنها به خرید و نگهداری یک رمزارز خاص اکتفا نکنند. بنابراین نیاز است پژوهش‌های بیشتری در این حوزه انجام شود و به سرمایه‌گذاران کمک کند تا با توجه به نوظهور بودن این بازار، بتوانند با انتخاب و بهینه‌سازی پرتفوی مناسب علاوه بر کسب بازده به خوبی از ریسک‌های آن نیز در امان بمانند.

## ۸-۱- کاربردهای موضوع پژوهش

کابرد اصلی این پژوهش کمک به سرمایه‌گذاران برای تشکیل و بهینه‌سازی یک سبد سرمایه‌گذاری از ارزهای دیجیتال برای افزایش بازده و کاهش ریسک است. این کار به کمک وزن‌دهی به هر رمزارز و تقسیم بودجه‌ی سرمایه‌گذاری بین آن‌ها انجام می‌شود.

## ۸-۱- کاربران نتایج پژوهش

این پژوهش می‌تواند برای سرمایه‌گذاران رمزارزها که قصد نگهداری بلندمدت یا کوتاه‌مدت آن‌ها را دارند مفید باشد. کاربران دیگر این پژوهش، سرمایه‌گذاران سایر بازارها هستند که قصد ورود به بازار ارزهای دیجیتال را دارند و می‌خواهند با صرف کمترین میزان ریسک، از منافع آن استفاده کنند. نتایج این تحقیق همچنین می‌تواند برای محققان اقتصادی، سیاست‌گذاران بازارهای مالی و پژوهشگران مفید باشد.

## ۹-۱- جمع‌بندی

بازار رمزارزها یک بازار نوپاست که در سال‌های اخیر به طور گسترده‌ای مورد توجه سرمایه‌گذاران قرار گرفته است. با توجه به ریسک بالای این بازار، ضروری است که سرمایه‌گذاران اقدام به مدیریت ریسک سرمایه‌گذاری خود، به خصوص به وسیله‌ی تنوع‌بخشی و با تشکیل سبد دارایی‌ها کنند.

فصل دوم

# مبانی نظری و مرور ادبیات

## ۱-۲- مقدمه

با گسترش رمزارزها در دهه‌ی سوم قرن بیست و یکم، گروه‌های مختلفی از ارزهای دیجیتال پیدا شد که هر یک کارکردها و ریسک‌های خاص خود را دارد. این امر موجب می‌شود که سرمایه‌گذاران تنها به سرمایه‌گذاری در بزرگ‌ترین ارز دیجیتال، یعنی بیت‌کوین بسنده نکنند و به فکر تشکیل سبدی از رمزارزها باشند تا هم از مزایای گروه‌های مختلف این رمزارزها استفاده کنند و هم از ریسک‌هایشان در امان بمانند. بنابراین این سرمایه‌گذاران باید روش‌های بهینه‌سازی سبد سرمایه‌گذاری را بررسی کنند و با توجه به ویژگی‌های این بازار بهترین آن‌ها را برای تشکیل سبد انتخاب کنند. در این فصل به مطالعه‌ی سبد دارایی‌ها و روش‌های انتخاب و بهینه‌سازی آن پرداخته می‌شود.

## ۲-۲- انتخاب سبد سرمایه‌گذاری

سرمایه‌گذاران با هدف کسب سود در آینده، بخشی از نقدینگی خود را به سرمایه‌گذاری در دارایی‌های مختلف اختصاص می‌دهند. هر کدام از این دارایی‌ها دارای ریسک‌های خاص خود هستند و بازده آن‌ها در آینده نامشخص است؛ اما هر سرمایه‌گذار، انتظاری از بازده دارایی‌ها دارد و بر این اساس اقدام به خرید یا فروش آن‌ها می‌کند. تجربه‌ی سرمایه‌گذاران نشان می‌دهد که برای کاهش ریسک سرمایه‌گذاری می‌توانند سبدی از دارایی‌های مختلف تشکیل دهند تا در صورت افت یک یا چند دارایی، متحمل ضرر کمتری شوند. در واقع با این کار می‌توان هم در گروه‌های مختلفی از دارایی‌ها سرمایه‌گذاری کرد و هم ریسک کل سرمایه‌گذاری را کاهش داد.

### ۱-۲-۲- مدل انتخاب ساده[[17]](#footnote-17) (مدل هم‌وزن)

مدل ساده یا هم‌وزن که به آن مدل 1/N یا EW[[18]](#footnote-18) نیز گفته می‌شود، سرمایه‌گذاری برابری را برای تمام دارایی‌ها پیشنهاد می‌کند. این روش نیازی به پیشبینی بازده مورد انتظار یا ماتریس کوواریانس بازده دارایی ها ندارد و با کمترین زمان و هزینه قابل دستیابی است (ابونوری و همکاران، ۱۳۹۷). پژوهش‌های زیادی انجام شده است که بررسی می‌کند که آیا روش ساده می‌تواند نسبت به سایر روش‌های انتخاب سبد برتری داشته باشد یا خیر؛ تا جایی که در برخی از آن‌ها این روش بر بسیاری از روش‌های بهینه‌سازی سبد سرمایه‌گذاری عملکرد بهتری داشته است (پلاتاناکیس و اورکوهارت[[19]](#footnote-19)، ۲۰۱۹).

### ۲-۲-۲- مدل انتخاب سبد بازار

این روش پرتفوی بازار را برای سرمایه‌گذاری پیشنهاد می‌کند. برای انجام این کار، پس از انتخاب دارایی‌های مورد نظر به هر کدام از آن‌ها وزنی متناسب با ارزش کل بازار آن می‌دهیم. بنابراین نسبت هر دارایی از کل سرمایه برابر با رابطه‌ی ۲-۱ خواهد بود:

*(رابطه‌ی ۲-۱)*

این روش هم مانند روش ساده، بیشتر برای مقایسه و سنجش عملکرد سایر روش‌های انتخاب و بهینه‌سازی سبد سرمایه‌گذاری استفاده می‌شود.

## ۳-۲- بهینه‌سازی سبد سرمایه‌گذاری

مسئله‌ی اصلی در بهینه‌سازی سبد سرمایه‌گذاری، انتخاب بهینه‌ی دارایی‌هایی است که با مقدار مشخصی سرمایه می‌توان آن‌ها را تهیه کرد. مارکوویتز (۱۹۵۲) با ایجاد یک نوآوری قابل توجه، مدلی ارائه کرد که به کمک آن می‌توان ریسک سبد سرمایه‌گذاری را به وسیله‌ی تنوع‌بخشی به حداقل رساند. ویلیام شارپ[[20]](#footnote-20) با مشاهده‌ی مشکلات محاسباتی در حل مدل مارکوویتز، شاخص جدیدی به نام بتا[[21]](#footnote-21) را پیشنهاد داد که درجه‌ی حساسیت نرخ بازدهی دارایی به تغییرات شاخص را اندازه‌گیری می‌کند. نتایج تحقیقات شارپ و دو تن از همکارانش منجر به پدید آمدن مدلی به نام مدل قیمت‌گذاری دارایی‌های سرمایه‌ای یا CAPM[[22]](#footnote-22) شد که پارادایمی در حوزه‌ی سرمایه‌گذاری بود (راعی و تلنگی، ۱۳۸۳).

تئوری‌های مالی مثل مدل CAPM دارای پیشفرض‌هایی مثل فرضیه‌ی بازار کارا و عقلایی بودن سرمایه‌گذاری هستند که در چند دهه‌ی اخیر مورد تردید واقع شده است. بر این اساس، بسیاری از سرمایه‌گذاران و محققان این مدل‌ها را دارای اعتبار کافی نمی‌دانند (عباس‌نژاد، ۱۳۸۰). از این رو مدل‌های گسترده و جدیدی برای بهینه‌سازی سبد سرمایه‌گذاری تهیه شده است.

توزیع بازده دارایی‌هایی مثل سهام همواره مورد تحقیقات بسیاری قرار گرفته است و نتایج به‌دست‌آمده حاکی از آن است که این توزیع نرمال نیست. بر این اساس، محققان سنجه‌های ریسک نامطلوب را مطرح کردند که بین نوسان‌های مطلوب و نامطلوب تفاوت قائل می‌شود و تنها نوسان‌های پایین‌تر از بازده انتظاری سرمایه‌گذار را ریسک در نظر می‌گیرد. بنابراین این تئوری‌ها بر اساس رابطه‌ی بازده و ریسک نامطلوب به معیارهای انتخاب سبد بهینه می‌پردازد (استرادا[[23]](#footnote-23)، ۲۰۰۷).

### ۱-۳-۲- مدل مارکوویتز

نظریه‌ی میانگین-واریانس برای اولین بار توسط هری مارکوویتز برای حل مسئله‌ی انتخاب مجموعه‌ی بهینه‌ی دارایی ارائه شد. مارکوویتز این مسئله را به صورت برنامه‌ریزی کوادراتیک با هدف حداقل کردن واریانس پرتفوی دارایی‌ها که سنجه‌ای برای ریسک آن‌ها محسوب می‌شود، با این شرط که بازده انتظاری بیشتر از یک مقدار ثابت باشد مطرح کرد. محدودیت دیگر این مسئله‌ی بهینه‌سازی این است که مجموع متغیرهای تصمیم مسئله یعنی وزن دارایی‌ها برابر با یک باشد و هیچ یک از این اوزان کوچک‌تر از صفر هم نشود (صباحی و همکاران، ۱۳۹۹).

بر طبق مدلی که مارکوویتز (۱۹۵۹) آن را ارائه کرده است، فرض می‌کنیم که دارایی از اوراق بهادار داریم و هر کدام از آن‌ها را با نمایش می‌دهیم. در این صورت بازده ورقه‌ی -ام یعنی ، یک متغیر تصادفی خواهد بود که میانگین آن را و انحراف معیار آن را در نظر می‌گیریم. همچنین کوواریانس بین بازدهی دو دارایی با نمایش داده می‌شود. *در این صورت شکل استاندارد مدل میانگین-واریانس به صورت رابطه‌ی ۲-۲ خواهد بود.*

*(رابطه‌ی ۲-۲)*

در مدل فوق پارامتر حداقل بازده مورد انتظار سرمایه‌گذار است و برای مقادیر مختلف آن، جواب‌های متفاوتی به دست می‌آید که مجموعه‌ی آن‌ها، محدوده‌ای به نام مرز کارا را تشکیل می‌دهد. همه‌ی پرتفوهای واقع در مرز کارا بهینه محسوب می‌شوند و انتخاب هر کدام یک از آن‌ها بستگی به نظر سرمایه‌گذار و روش‌های ارائه‌شده دارد. (کیانی هرچگانی و همکاران، ۱۳۹۳)

با حل این مسئله‌ی تحقیق در عملیات با به حداقل رساندن تابع هدف، وزن بهینه‌ی هر یک از دارایی‌ها در سبد سرمایه‌گذاری به دست می‌آید. در این صورت برای متغیر تصادفی بازده سبد خواهیم داشت:

*(رابطه‌ی ۲-۳)*

### ۲-۳-۲- انتخاب نقطه‌ی بهینه روی مرز کارا

با استفاده از مدل مارکوویتز می‌توان محدوده‌ای شامل پرتفوهای بهینه به دست آورد که به آن مرز کارا گفته می‌شود. مرز کارا در نمودار بازده-ریسک مقعر و صعودی است و انتخاب سبد سرمایه‌گذاری روی آن روش‌های مختلفی دارد که در ادامه ذکر می‌شود.



**شکل ۲-۱:** مرز کارا برای مجموعه‌ای از دارایی‌ها در نمودار بازده-ریسک

#### ۱-۲-۳-۲- روش کمینه کردن واریانس

مارکوویتز (۱۹۵۹) برای انتخاب سبد بهینه روی مرز کارا پیشنهاد داد که حداقل بازده مورد انتظار سهامدار تعیین شود. به این صورت سبدی انتخاب می‌شود که کمترین ریسک را برای بازده تعیین‌شده فراهم می‌کند. اگر بازده مورد نظر در مرز کارا وجود داشته باشد، نقطه‌ی متناظر با این بازده روی مرز کارا به عنوان پرتفوی بهینه تعیین می‌شود و در صورت پایین‌تر بودن آن از مرز کارا، سبد دارای کمترین ریسک روی مرز کارا (پرتفوی MVP[[24]](#footnote-24)) جواب مسئله خواهد بود.

#### ۲-۲-۳-۲- روش بیشینه کردن بازده

در این روش به جای تعیین کف برای بازده مورد نظر سرمایه‌گذار، از یک سقف برای ریسک مورد انتظار استفاده می‌شود. بنابراین اگر ریسک تعیین‌شده در مرزکارا موجود باشد، نقطه‌ی متناظر با آن به عنوان سبد بهینه انتخاب می‌شود و اگر این ریسک از مرز کارا بالاتر باشد، پاسخ مسئله بالاترین نقطه روی مرز کارا خواهد بود که دارای بیشترین بازده و ریسک در این محدوده است (کولجک[[25]](#footnote-25) و همکاران، ۲۰۲۲).

#### ۳-۲-۳-۲- روش بیشینه کردن نسبت شارپ

نسبت شارپ از تقسیم بازده مازاد کسب‌شده‌ی سبد سرمایه‌گذاری در برابر نرخ بازده بدون ریسک، به انحراف معیار آن به دست می‌آید. روی[[26]](#footnote-26) (۱۹۵۲) اشاره می‌کند که پرتفویی که دارای بیشترین نسبت شارپ در مرز کاراست، نقطه ای است که از نرخ بازده بدون ریسک به آن مماس شود. در واقع در این روش حداقل بازده مورد انتظار به عنوان پارامتر در نظر گرفته می‌شود و برای مماس شدن خط به مرز کارا، شیب آن بیشینه می‌شود. (شهرستانی و همکاران، ۱۳۸۹). نسبت شارپ از رابطه‌ی ۲-۴ به دست می‌آید و هر چه مقدار آن برای یک سبد دارایی بیشتر باشد، نشان‌دهنده‌ی عملکرد بهتر آن است.

*(رابطه‌ی ۲-۴)*

#### ۴-۲-۳-۲- روش بیشینه کردن حداقل بازده مورد انتظار

کاتائوکا[[27]](#footnote-27) (۱۹۶۳) توضیح می‌دهد که در این روش همانند روش بیشینه کردن نسبت شارپ، خطی مماس به مرز کارا و از نرخ بازده بدون ریسک رسم می‌کنیم؛ با این تفاوت که شیب آن را پارامتری ثابت در نظر می‌گیریم که باید توسط سرمایه‌گذار تعیین شود و حداقل بازده مورد انتظار را بیشینه می‌کنیم (دینگ[[28]](#footnote-28) و ژانگ[[29]](#footnote-29)، ۲۰۰۹).

#### ۵-۲-۳-۲- روش بیشینه کردن مطلوبیت

یکی دیگر از روش‌های انتخاب سبد از روی مرز کارا، استفاده توابع مطلوبیت است. به این روش، از طریق منحنی‌های بی‌تفاوتی که برای هر سرمایه‌گذار با توجه به میزان ریسک‌پذیری او متفاوت است، می‌توان نقطه‌ی اشتراک تابع مطلوبیت و مرز کارا را به عنوان پرتفوی بهینه در نظر گرفت. به عنوان یک مثال از توابع مطلوبیت، تابع کوادراتیک رابطه‌ی ۲-۵ قابل تعریف است (لیو[[30]](#footnote-30)، ۲۰۱۹):

*(رابطه‌ی ۲-۵)*

### ۳-۳-۲- توسعه‌ی مدل مارکوویتز

توسعه‌های زیادی برای بهبود عملکرد روش میانگین-واریانس صورت گرفته شده است. در ادامه تعدادی از این روش‌های بهبودیافته ذکر خواهد شد.

#### ۱-۳-۳-۲- تغییر سنجه‌ی ریسک

همان طور که در مدل میانگین-واریانس ذکر شد، مارکوویتز (۱۹۵۲) ریسک دارایی‌ها را با سنجه‌ی واریانس اندازه گرفت. به مرور زمان پژوهشگران سنجه‌های دیگری را برای ریسک در نظر گرفتند و از این طریق مدل مارکوویتز را توسعه دادند. سنجه‌های ریسک به طور کلی به سه دسته تقسیم می‌شود. دسته‌ی اول سنجه‌های مبتنی بر تلاطم است که واریانس نیز یکی از این سنجه‌هاست. شاخص‌های پراکندگی در علم آمار مانند دامنه‌ی تغییرات، دامنه‌ی میان‌چارکی، ضریب تغییرات و... نیز از همین دسته هستند. در این سنجه‌ها نوسان‌های داده‌ها مورد توجه قرار می‌گیرد. دسته‌ی دیگر از سنجه‌های ریسک، سنجه‌های مبتنی بر حساسیت است. این سنجه‌ها حساسیت متغیر تصادفی مورد نظر را در مقابل تغییرات یک متغیر تصادفی دیگر بررسی می‌کنند. از جمله این سنجه‌ها می‌توان به دیرش، تحدب و ضریب بتا شاره کرد که هر کدام کارکردهای خاص خود را دارا هستند (زمردیان و همکاران، ۱۳۹۸).

آخرین دسته از سنجه‌های ریسک، سنجه‌های ریسک نامطلوب است که بر خلاف سنجه‌های دیگر که تغییرات مثبت و منفی را به عنوان ریسک در نظر می‌گیرد، تنها بخش مربوط به تغییرات منفی و نامطلوب را محاسبه می‌کند. این سنجه‌ها به دو گروه تقسیم می‌شود. نخستین گروه نیم‌سنجه‌ها هستند که شامل نیم‌واریانس، نیم‌بتا و... است (استرادا، ۲۰۰۷)؛ گروه دوم شامل سنجه‌های مبتنی بر صدک مانند ارزش در معرض ریسک، ریزش مورد انتظار و سنجه‌های طیفی است (یامای[[31]](#footnote-31) و یوشیبا[[32]](#footnote-32)، ۲۰۰۵).

#### ۲-۳-۳-۲- رویکرد فازی

از دیگر روش‌های توسعه‌ی مدل مارکوویتز که مورد توجه پژوهشگران قرار گرفت، استفاده از رویکردهای دیگر برای عدم قطعیت بود که یکی از این رویکردها، روش فازی است. همان طور که اورتی[[33]](#footnote-33) و همکاران (۲۰۰۲) نشان داده اند، با توجه به پیش‌بینی‌ناپذیر بودن بازارهای مالی و پیچیدگی آن‌ها ارائه‌ی یک تخمین دقیق از ریسک و بازده مورد انتظار بسیار دشوار است. بنابراین بسیاری از پژوهشگران به استفاده از روش فازی روی آورده اند. در این روش نرخ بازده، بتا، حداقل بازده مورد انتظار و... به جای اعداد قطعی، با اعداد فازی مدل‌سازی می‌شوند. این اعداد فازی می‌توانند به شکل اعداد مثلثی، ذوزنقه‌ای یا حالت‌های دیگر باشند. (امیری و محبوب قدسی، ۱۳۹۴)

#### ۳-۳-۳-۲- رویکرد بهینه‌سازی استوار

مدل‌های بهینه‌سازی استوار، بازدهی آینده‌ی دارایی‌ها را به صورت ضرایب غیرقطعی در مسئله‌ی بهینه‌سازی در نظر می‌گیرند و درجه‌ی ریسک‌گریزی سرمایه‌گذاری را به درجه‌ی تحمل در مقابل کل خطای حاصل تخمین بازدهی‌ها تصویر می‌کنند (قره‌خانی و همکاران، ۱۳۹۲). بنابراین همان طور که قهطرانی (۱۳۹۱) اشاره می‌کند، در رویکرد بهینه‌سازی استوار، به دنبال جواب‌های نزدیک به بهینه‌ای هستیم که با احتمال بالا موجه باشند. این رویکرد شامل رویکرد استوار سویستر[[34]](#footnote-34) (۱۹۷۳)، رویکرد استوار بن تال[[35]](#footnote-35) و نمیروفسکی[[36]](#footnote-36) (۲۰۰۰)، و رویکرد استوار برتسیماس[[37]](#footnote-37) و سیم[[38]](#footnote-38) (۲۰۰۳) است (پیکانی و روغنیان، ۱۳۹۴).

## ۴-۲- یادگیری ماشین

یادگیری ماشین یکی از روش‌های تشریح مجموعه‌ای از داده‌ها با استفاده از یافتن یک الگوی جامع است. الگوریتم‌های یادگیری ماشین با استفاده از حجم محدودی از داده‌های گرد‌آوری‌شده، الگویی برای ارتباط بین داده‌ها پیدا می‌کنند (ندایی و نجفی، ۲۰۱۶). این داده‌های آموزشی برای بهبود عملکرد و ارائه‌ی پیش‌بینی‌های دقیق‌تر استفاده می‌شود. در ادامه به بررسی انواع مسائل یادگیری ماشین پرداخته می‌شود.

### ۱-۴-۲- یادگیری نظارت‌شده[[39]](#footnote-39)

مسائل یادگیری ماشین ممکن است نظارت‌شده باشند. در این نوع مسائل علاوه بر داده‌های ورودی، داده‌های خروجی نیز توسط مدل‌ساز به مدل داده می‌شود و مدل موظف به یافتن رابطه‌ای بین ورودی‌ها و خروجی‌هاست. به این شکل مدل می‌تواند با یادگیری داده‌های ورودی و داده‌های خروجی که به آن‌ها داده‌های آموزشی[[40]](#footnote-40) گفته می‌شود، تابعی برای نحوه‌ی ارتباط این داده‌ها بسازد و از این تابع برای پیش‌بینی داده‌های آتی استفاده کند. یادگیری ماشین نظارت‌شده را می‌توان بر اساس داده‌های خروجی به دو نوع مسئله تقسیم کرد؛ اگر داده‌های خروجی پیوسته باشد به آن مسئله‌ی رگرسیون گفته می‌شود و در غیر این صورت، آن را یک مسئله‌ی دسته‌بندی می‌نامیم (ندایی و نجفی، ۲۰۱۶).

### ۲-۴-۲- یادگیری بدون نظارت[[41]](#footnote-41)

در مدل‌های یادگیری ماشین بدون نظارت بر خلاف مدل‌های نظارت‌شده، داده‌های خروجی به مدل داده نمی‌شود. هدف اصلی در این مسائل، تقسیم‌بندی داده‌ها بر اساس میزان شباهت آن‌هاست. به این شکل که مدل تلاش می‌کند با یادگیری داده‌های ورودی و یافتن ساختاری برای جدا کردن داده‌های مشابه، آن‌ها را به خوشه‌هایی تقسیم‌بندی کند که دارای بیشترین شباهت درون خوشه باشند و از طرف دیگر، کمترین شباهت بین خوشه‌های مختلف وجود داشته باشد. (ساتیا، ۲۰۱۳) تفاوت مسائل دسته‌بندی با این نوع مسائل که به آن‌ها خوشه‌بندی گفته می‌شود، در آن است که «دسته‌ها» توسط مدل‌ساز از پیش‌تعیین‌شده هستند و مدل تلاش می‌کند ارتباط داده‌های ورودی با این دسته‌ها را بیابد. این در حالی است که «خوشه‌ها» توسط خود مدل و بر مبنای داشتن بیشترین شباهت درون‌گروهی و کمترین شباهت بین‌گروهی ایجاد می‌شوند.

### ۳-۴-۲- روش‌های رگرسیون

#### ۱-۳-۴-۲- روش رگرسیون خطی[[42]](#footnote-42)

رگرسیون خطی یک الگوریتم یادگیری ماشینی پرکاربرد برای وظایف رگرسیون است. این یک روش ساده و موثر است که رابطه بین یک متغیر وابسته و یک یا چند متغیر مستقل را مدل می کند. الگوریتم با نصب یک خط صاف بر روی نقاط داده ای که بهترین نماینده رابطه بین متغیرها هستند، کار می‌کند. این خط به عنوان خط رگرسیون شناخته می‌شود و برای پیش‌بینی مقدار متغیر وابسته بر اساس مقدار متغیر مستقل استفاده می شود. [(داتا و ساهو، ۲۰۲۱)](https://link.springer.com/chapter/10.1007/978-981-16-1919-9_15)

#### ۲-۳-۴-۲- روش رگرسیون بردار پشتیبان[[43]](#footnote-43)

رگرسیون بردار پشتیبان (SVR) یک الگوریتم یادگیری ماشینی است که برای وظایف رگرسیون استفاده می شود. این یک روش قدرتمند است که می تواند با داده های خطی و غیر خطی سروکار داشته باشد. الگوریتم با یافتن یک ابرصفحه که حاشیه بین نقاط داده ای و ابرصفحه را بیشینه می‌کند، کار می‌کند. ابرصفحه برای پیش بینی مقدار متغیر وابسته بر اساس مقدار متغیر مستقل استفاده می‌شود. (کادام و همکاران، ۲۰۲۰)

Kadam, V. S., Kanhere, S. and Mahindrakar. S. 2020. Regression techniques in machine learning & applications: A review. *Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET)* (10):826–830.

#### ۳-۳-۴-۲- روش رگرسیون چندجمله‌ای[[44]](#footnote-44)

رگرسیون چندجمله‌ای یک الگوریتم یادگیری ماشینی است که رابطه بین یک متغیر وابسته و یک یا چند متغیر مستقل را مدل می کند. این یک نوع رگرسیون خطی است که از توابع چندجمله‌ای برای رسم نقاط داده‌ای استفاده می کند. به این شکل که رابطه‌ی بین متغیر مستقل و متغیر وابسته به عنوان یک تابع چندجمله‌ای مدل می‌شود. الگوریتم با یافتن بهترین منحنی مناسب که رابطه بین متغیرها را نشان می‌دهد، کار می‌کند. رگرسیون چند جمله ای می تواند برای مدل کردن داده های غیرخطی استفاده شود. [(داتا و ساهو، ۲۰۲۱)](https://link.springer.com/chapter/10.1007/978-981-16-1919-9_15)

#### ۴-۳-۴-۲- روش رگرسیون درخت تصمیم[[45]](#footnote-45)

رگرسیون درخت تصمیم یک الگوریتم یادگیری ماشینی است که رابطه بین یک متغیر وابسته و یک یا چند متغیر مستقل را مدل می کند. این یک نوع رگرسیون است که از درخت های تصمیم برای رسم نقاط داده‌ای استفاده می‌کند. الگوریتم با تقسیم مجدد داده‌ها به زیر مجموعه هایی بر اساس مقدار متغیر مستقل کار می‌کند. سپس زیرمجموعه ها برای ایجاد یک درخت تصمیم استفاده می‌شوند که رابطه بین متغیرها را نشان می دهد [(سارکر، ۲۰۲۱)](https://link.springer.com/article/10.1007/s42979-021-00592-x).

#### ۴-۳-۴-۲- روش رگرسیون جنگل تصادفی[[46]](#footnote-46)

رگرسیون جنگل تصادفی یک الگوریتم یادگیری ماشینی است که رابطه بین یک متغیر وابسته و یک یا چند متغیر مستقل را مدل می کند. این روش از جنگل تصادفی برای رسم نقاط داده‌ای استفاده می‌کند. به این شکل که چندین درخت تصمیم را ترکیب می کند تا دقت پیش‌بینی‌ها را بهبود بخشد. الگوریتم با ایجاد چندین درخت تصمیم بر روی زیرمجموعه‌های مختلف داده‌ها شروع می‌شود و سپس نتایج را ترکیب می‌کند تا پیش‌بینی نهایی را انجام دهد [(لی و همکاران، ۲۰۲۲)](https://link.springer.com/chapter/10.1007/978-981-16-9229-1_14).

### ۴-۴-۲- روش‌های دسته‌بندی

#### ۱-۴-۴-۲- روش رگرسیون لجستیک[[47]](#footnote-47)

رگرسیون لجستیک یک الگوریتم ساده و موثر است که رابطه‌ی بین یک یا چند متغیر مستقل و یک متغیر وابسته‌ی دوسویی (به معنای این که مربوط به یک واقعه‌ی تصادفی در دو موقعیت ممکنه است) را مدل می‌کند. الگوریتم با رسم یک منحنی لجستیک بر روی نقاط داده‌ای که بهترین نماینده‌ی رابطه‌ی بین متغیرها هستند، کار می‌کند. این منحنی برای پیش‌بینی احتمال متغیر وابسته برای تعلق به یک کلاس خاص بر اساس مقدار متغیر مستقل استفاده می‌شود. [(سارکر، ۲۰۲۱)](https://link.springer.com/article/10.1007/s42979-021-00592-x).

#### ۲-۴-۴-۲- روش درخت تصمیم[[48]](#footnote-48)

درخت تصمیم یک الگوریتم یادگیری ماشینی است که برای مسائل دسته‌بندی استفاده می شود. این روش می‌تواند با داده‌های خطی و غیرخطی سروکار داشته باشد. الگوریتم با تقسیم مجدد داده ها به زیرمجموعه‌هایی بر اساس مقدار متغیر مستقل کار می‌کند. سپس زیرمجموعه‌ها برای ایجاد یک درخت تصمیم استفاده می شوند که رابطه بین متغیرها را نشان می‌دهد [(چودهاری و شوئن، ۲۰۲۰)](https://ieeexplore.ieee.org/document/9249211/).

#### ۳-۴-۴-۲- روش جنگل تصادفی[[49]](#footnote-49)

جنگل تصادفی یک الگوریتم یادگیری ماشینی است که برای وظایف طبقه بندی استفاده می شود. این یک روش مجموعه است که چندین درخت تصمیم را ترکیب می‌کند تا دقت پیش‌بینی‌ها را بهبود بخشد. الگوریتم با ایجاد چندین درخت تصمیم بر روی زیرمجموعه های مختلف داده ها شروع می‌شود و سپس نتایج را ترکیب می‌کند تا پیش‌بینی نهایی را انجام دهد [(لی و همکاران، ۲۰۲۲)](https://link.springer.com/chapter/10.1007/978-981-16-9229-1_14).

#### ۴-۴-۴-۲- روش بیز ساده[[50]](#footnote-50)

بیز ساده یک الگوریتم یادگیری ماشینی است که برای وظایف طبقه بندی استفاده می شود. این یک روش احتمالاتی است که رابطه بین متغیر وابسته و یک یا چند متغیر مستقل را مدل می کند. الگوریتم با محاسبه احتمال متغیر وابسته برای تعلق به یک کلاس خاص بر اساس مقدار متغیر مستقل کار می کند [(چانگ‌گه و همکاران، ۲۰۲۰)](https://link.springer.com/article/10.1007/s11192-020-03614-2).

### ۵-۴-۲- روش‌های خوشه‌بندی

#### ۱-۵-۴-۲- روش ماشین بردار پشتیبانی[[51]](#footnote-51)

ماشین بردار پشتیبانی یک روش خوشه‌بندی است که برای دسته‌بندی داده‌های دو دسته‌ای استفاده می‌شود. این روش با پیدا کردن ابرصفحه‌ای که بهترین جداسازی بین داده‌ها را انجام می‌دهد، کار می‌کند. ابرصفحه به گونه‌ای انتخاب می‌شود که فاصله بین دو دسته را بیشینه می‌کند. فاصله به عنوان فاصله‌ی بین ابرصفحه و نزدیک‌ترین داده‌های هر دسته تعریف می‌شود. SVM یکی از روش‌های خوشه‌بندی محبوب است زیرا در فضاهای بالا بعدی موثر است و می‌تواند با داده‌های غیرخطی نیز استفاده شود. (ردموند و همکاران، ۲۰۱۱)

#### ۲-۵-۴-۲- روش انتشار وابستگی[[52]](#footnote-52)

انتشار وابستگی یک روش خوشه‌بندی است که برای دسته‌بندی داده‌های چنددسته‌ای استفاده می‌شود. این روش با پیدا کردن نمونه‌هایی، یا نقاط نماینده، در مجموعه داده که بهترین خلاصه‌ای از داده‌ها را ارائه می‌دهند، کار می‌کند. هر داده بر اساس شباهت آن با نمونه‌ها به یک نمونه نسبت داده می‌شود. شباهت با استفاده از یک اندازه‌گیری از شباهت داده‌ها با همه داده‌های دیگر در مجموعه داده محاسبه می‌شود. انتشار وابستگی یکی از روش‌های خوشه‌بندی محبوب است زیرا سرعت بالا، دقت بالا و تعداد پارامترهای کمتری نسبت به روش‌های دیگر خوشه‌بندی دارد و نیازی به تعیین تعداد خوشه توسط کاربر ندارد (ردموند و همکاران، ۲۰۱۱).

#### ۳-۵-۴-۲- روش K-Means

روش K-Means یکی از روش‌های خوشه‌بندی است که برای دسته‌بندی داده‌های چند بعدی استفاده می‌شود. در این روش، داده‌ها به تعداد k خوشه تقسیم می‌شوند و هر خوشه به گونه‌ای انتخاب می‌شود که مرکز خوشه بهترین نماینده برای داده‌های آن خوشه باشد. این روش برای داده‌هایی که تعداد بعد آن‌ها کم است، به خوبی عمل می‌کند (ردموند و همکاران، ۲۰۱۱).

#### ۴-۵-۴-۲- روش خوشه‌بندی فضایی مبتنی بر چگالی در کاربردهای دارای نویز[[53]](#footnote-53)

این روش که به اختصار به آن DBSCAN گفته می‌شود، یکی از روش‌های خوشه‌بندی است که برای خوشه‌بندی داده‌هایی با توزیع ناهمگن استفاده می‌شود. در این روش، خوشه‌ها به گونه‌ای تعریف می‌شوند که داده‌هایی که در یک خوشه قرار دارند، از نظر فاصله با یکدیگر نزدیک هستند و داده‌هایی که در خوشه‌ای قرار نمی‌گیرند، از نظر فاصله با داده‌های خوشه‌ها دور هستند. این روش برای داده‌هایی که توزیع آن‌ها ناهمگن است، به خوبی عمل می‌کند (ردموند و همکاران، ۲۰۱۱). این الگوریتم نیازی به تعیین تعداد خوشه توسط کاربر ندارد و خود الگوریتم می‌تواند خوشه‌ها را مبتنی بر غلظت آن‌ها شناسایی کند.

#### ۵-۵-۴-۲- روش خوشه‌بندی سلسله‌مراتبی[[54]](#footnote-54)

خوشه‌بندی سلسله‌مراتبی یکی از روش‌های خوشه‌بندی داده‌های چندبعدی است. در این روش، داده‌ها به صورت یک درخت خوشه‌بندی شده و هر خوشه به گونه‌ای انتخاب می‌شود که فاصله بین داده‌های آن خوشه کمینه شود. این روش برای داده‌هایی که تعداد بعد آن‌ها کم است، به خوبی عمل می‌کند (ردموند و همکاران، ۲۰۱۱).

## ۵-۲- مروری بر ادبیات موضوع

با پیشرفت بازار ارزهای دیجیتال، تحقیقات گسترده‌ای در زمینه‌های مختلف روی رمزارزها انجام شده است. از جمله موضوعات این پژوهش‌ها می‌توان به بررسی مسائل مربوط به تکنولوژی، ماهیت، قانون‌گذاری و سرمایه‌گذاری رمزارزها اشاره کرد. یکی از مهم‌ترین دغدغه‌های سرمایه‌گذاران، تشکیل و انتخاب سبد سرمایه‌گذاری است. این مسئله به خصوص از این جهت اهمیت دارد که هنوز صندوق‌های سرمایه‌گذاری بسیار کمی برای رمزارزها وجود دارد و فرد علاقه‌مند به این بازار می‌بایست خود اقدام به بررسی و انتخاب سبد سرمایه‌گذاری کند. از این جهت مقالات مرتبط با تشکیل و بهینه‌سازی پرتفو برای بازار رمزارزها در سال‌های اخیر توسعه یافته است. در بخشی از این مطالعات از روش‌های یادگیری ماشین مثل رگرسیون و خوشه‌بندی استفاده شده است. در ادامه به مروری بر این مطالعات پرداخته شده است.

### ۱-۵-۲- بررسی مقالات

.

**جدول ۲-۱:** مقایسه‌ی مقالات

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | | | | | |  | | | |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |

**جدول ۲-۱:** مقایسه‌ی مقالات (ادامه)

## ۶-۲- جمع‌بندی

انتخاب یک سبد سرمایه‌گذاری مناسب از دارایی‌های مورد نظر، همواره یکی از دغدغه‌های سرمایه‌گذاران بوده است. همان طور که مشاهده شد، روش‌های گوناگونی برای انتخاب و بهبنه‌سازی سبدهای سرمایه‌گذاری توسعه یافته شده است تا سرمایه‌گذاران بتوانند بازدهی بیشتر و ریسک کمتری را تجربه کنند. پژوهشگران همچنین روش‌های غیربهینه‌ای مثل روش هم‌وزن یا روش انتخاب پرتفوی بازار را در کنار روش‌های بهینه‌سازی مثل مدل مارکوویتز دنبال می‌کنند؛ چرا که هدف اصلی این است که سرمایه‌گذاران بتوانند در عمل از مزایای متنوع‌سازی دارایی‌ها بهره‌مند شوند. همچنین علاوه بر روش‌های ذکرشده، روش‌های جدیدی مانند مدل‌های مبتنی بر هوش مصنوعی به طور فزاینده در حال پیشرفت هستند و مقایسه‌ی این رویکردهای متفاوت با یکدیگر، از مطالعات لازم برای سرمایه‌گذاران است. در فصل آینده، پژوهش‌های انجام‌شده در حیطه‌ی انتخاب و بهینه‌سازی سبد دارایی‌ها در بازار رمزارزها مورد بررسی قرار می‌گیرد.

فصل سوم

# روش‌شناسی تحقیق

## ۱-۳- مقدمه

.

فصل چهارم

# پیاده‌سازی و نتایج پژوهش

## ۱-۴- مقدمه

**.**

فصل پنجم

# نتیجه‌گیری و پیشنهادات برای تحقیقات آتی

## ۱-۵- مقدمه

**.**

## ۲-۵- جمع‌بندی و خلاصه‌ای از پژوهش

.

## ۳-۵- پیشنهادات برای تحقیقات آتی

.

## فهرست مراجع

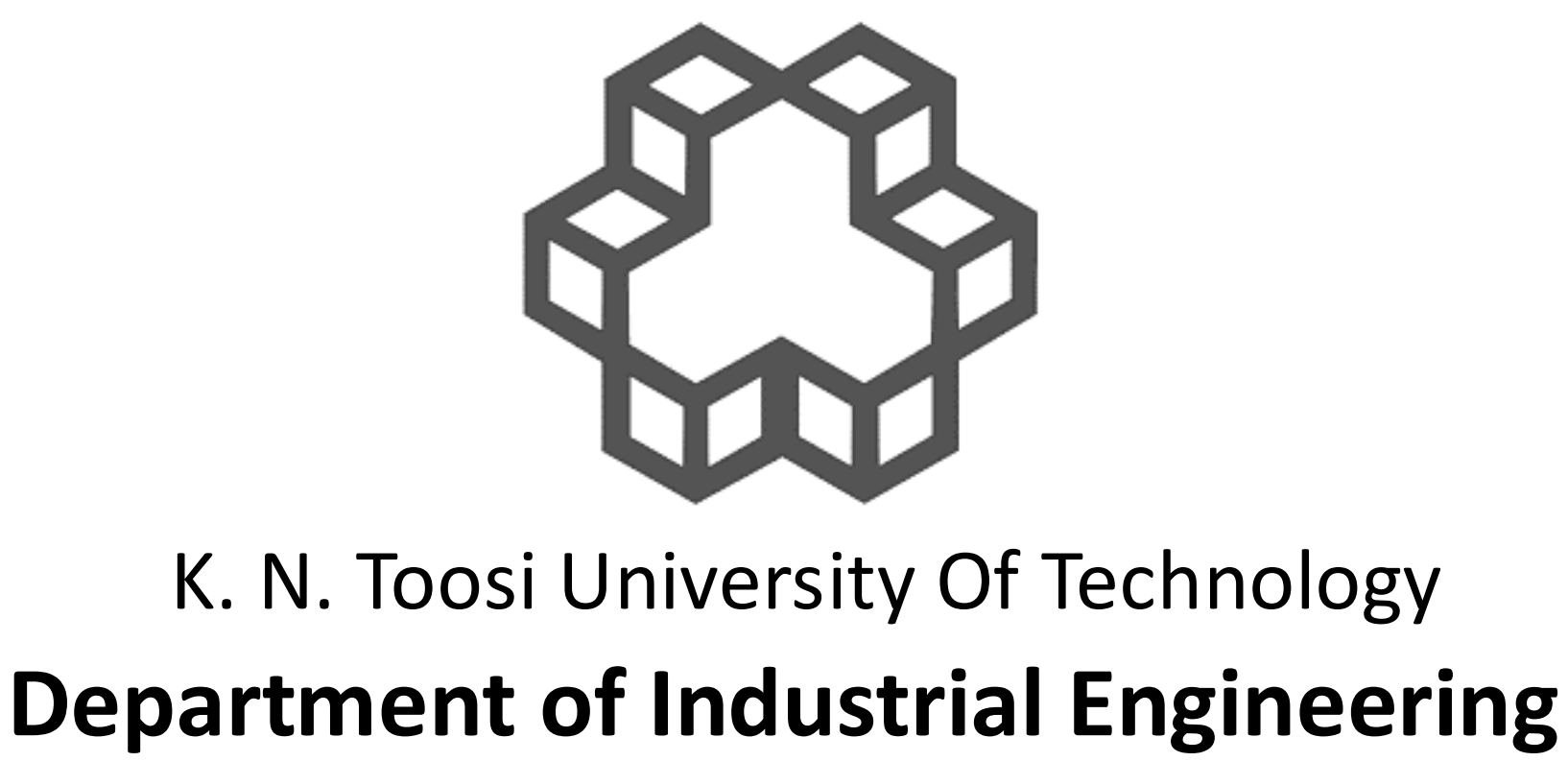
.

**Abstract**

.

**Keywords:**

.



**Portfolio Optimization in the Cryptocurrency Market Using Machine Learning and Clustering**

**Alireza Nezhadshamsi**

**Supervisor**

**Dr. Amir Abbas Najafi**

**Master of Science Thesis in**

**Industrial Engineering Majoring in Financial Engineering**

**. 2023**

1. Diversification [↑](#footnote-ref-1)
2. Blockchain [↑](#footnote-ref-2)
3. Briere [↑](#footnote-ref-3)
4. Guesmi [↑](#footnote-ref-4)
5. Demiralay [↑](#footnote-ref-5)
6. Bayracı [↑](#footnote-ref-6)
7. Mba [↑](#footnote-ref-7)
8. Brauneis [↑](#footnote-ref-8)
9. Mestel [↑](#footnote-ref-9)
10. Platanakis [↑](#footnote-ref-10)
11. Liu [↑](#footnote-ref-11)
12. Schellinger [↑](#footnote-ref-12)
13. Litecoin [↑](#footnote-ref-13)
14. Ethereum [↑](#footnote-ref-14)
15. Ripple [↑](#footnote-ref-15)
16. Dash [↑](#footnote-ref-16)
17. Naïve [↑](#footnote-ref-17)
18. Equally Weighted [↑](#footnote-ref-18)
19. Urquhart [↑](#footnote-ref-19)
20. William Sharpe [↑](#footnote-ref-20)
21. Beta [↑](#footnote-ref-21)
22. Capital Asset Pricing Model [↑](#footnote-ref-22)
23. Estrada [↑](#footnote-ref-23)
24. Minimum Variance Portfolio [↑](#footnote-ref-24)
25. Čuljak [↑](#footnote-ref-25)
26. Roy [↑](#footnote-ref-26)
27. Kataoka [↑](#footnote-ref-27)
28. Ding [↑](#footnote-ref-28)
29. Zhang [↑](#footnote-ref-29)
30. Liu [↑](#footnote-ref-30)
31. Yamai [↑](#footnote-ref-31)
32. Yoshiba [↑](#footnote-ref-32)
33. Orti [↑](#footnote-ref-33)
34. Soyster [↑](#footnote-ref-34)
35. Ben-Tal [↑](#footnote-ref-35)
36. Nemirovski [↑](#footnote-ref-36)
37. Bertsimas [↑](#footnote-ref-37)
38. Sim [↑](#footnote-ref-38)
39. Supervised Learning [↑](#footnote-ref-39)
40. Training data [↑](#footnote-ref-40)
41. Unsupervised Learning [↑](#footnote-ref-41)
42. Linear regression [↑](#footnote-ref-42)
43. Support Vector Regression (SVR) [↑](#footnote-ref-43)
44. Polynomial regression [↑](#footnote-ref-44)
45. Decision Tree Regression [↑](#footnote-ref-45)
46. Random Forest Regression [↑](#footnote-ref-46)
47. Logistic Regression [↑](#footnote-ref-47)
48. Decision Tree [↑](#footnote-ref-48)
49. Random Forest [↑](#footnote-ref-49)
50. Naive Bayes [↑](#footnote-ref-50)
51. Support Vector Machine (SVM) [↑](#footnote-ref-51)
52. Affinity Propagation [↑](#footnote-ref-52)
53. Density Based Spatial Clustering of Applications with Noise (DBSCAN) [↑](#footnote-ref-53)
54. Hierarchical Clustering [↑](#footnote-ref-54)