# 



دانشگاه شهيد بهشتي

دانشکده مهندسی و علوم کامپیوتر

**استفاده از یادگیری تقویتی در صندوق‌های سرمایه‌گذاری رمزارزهای انرژی‌های سبز**

پروژه کارشناسی­

پارسا نوری

دکتر مائده مشرف دهکردی

زمستان ۱۴۰۲

**چکیده**

بازار اعتبارات کربن یکی از بازارهایی است که بعد از توافق پاریس مورد توجه جهان قرار گرفته است. هم‌چنین اعتبارهای کربن که به صورت رمزارزها به نمایش در می‌تواند برای افزایش شفافیت این بازار مورد استفاده قرار بگیرد. یکی از راهکارهایی که سیستم‌های اعتبار کربن باید ارائه دهد، روشی برای تبدیل اعتبار رمزارز به ارزی قابل معامله است. راهکار ما در این جا استفاده از استخرهایی است که از یادگیری تقویتی بهره می‌برند. دلیل استفاده از آن‌های نیز لازمه وجود یک تعادل برای فرکانس تولید رمزارز خروجی است. به دلیل اصل عرضه و تقاضا هرچه میزان تولید ارز کمتر بیشتر شود کمیابی آن بیشتر و ارزش آن نیز بیشتر می‌شود ولی از آن طرف هرچه میزان تولید آن کمتر شود تعداد کاربران کمتری تمایل به سرمایه‌گذاری در آن را دارند. برای حل این مشکل از شبکه یادگیرنده تقویتی مبتنی بر شبکه یادگیرده عمیق Q استفاده کردیم. برای حفظ امنیت قراردادهای هوشموند نیز راهکارهایی اندیشیده شده. برای نحوه استفاده از شبکه یادگیرنده عمیق در قراردادهای هوشمند نیازمند حل چالش‌های بوده که به آن‌ها پرداختیم. در انتها راهکاری برای آماده سازی این یادگیرنده برای حضور در دنیای واقعی و هم‌چنین تا حدودی سنجش عملکرد آن ارايه شد.

**فهرست مطالب**

[فصل اول: کلیّات 1](#_Toc157206123)

[1-1 مقدمه 2](#_Toc157206124)

[1-2 برخی مفاهیم 4](#_Toc157206125)

[1-2-1 مفهوم ERC-20 4](#_Toc157206126)

[1-2-2 مفوهم ERC-721 4](#_Toc157206127)

[1-2-3 مفهوم ERC-1155 5](#_Toc157206128)

[1-3 بیان مسئله 5](#_Toc157206129)

[1-4 کلیات روش پیشنهادی 6](#_Toc157206130)

[1-5 ساختار پروژه 7](#_Toc157206131)

[فصل دوم: سیستم‌های رمزارزی اعتبار کربن موجود 8](#_Toc157206132)

[2-1 مقدمه 9](#_Toc157206133)

[2-2 بررسی سیستم توکان 9](#_Toc157206134)

[2-2-1 بازار انرژی توکان 9](#_Toc157206135)

[2-2-2 سیستم پل در توکان 11](#_Toc157206136)

[2-2-3 بررسی استخرها در سیستم توکان 12](#_Toc157206137)

[2-3 بررسی سیستم انرژی وب اریجین 14](#_Toc157206138)

[2-3-1 بررسی معماری انرژی وب اریجین 15](#_Toc157206139)

[2-4 جمعبندی 18](#_Toc157206140)

[فصل سوم: بررسی روش‌های یادگیری تقویتی 20](#_Toc157206141)

[3-1 مقدمه 21](#_Toc157206142)

[3-2 بررسی روش‌های مختلف یادگیری تقویتی 21](#_Toc157206143)

[3-2-1 تقویت یادگیری بدون مدل 21](#_Toc157206144)

[3-2-2 یادگیری تقویتی با مدل 24](#_Toc157206145)

[3-2-3 یادگیری تقلیدی 26](#_Toc157206146)

[3-2-4 یادگیری تقویتی عمیق 27](#_Toc157206147)

[3-3 استراتژی‌های کاوش 27](#_Toc157206148)

[3-3-1 اپسیلون-حریص 28](#_Toc157206149)

[3-3-2 مرز بالای اعتماد 29](#_Toc157206150)

[3-3-3 نمونه‌برداری تامپسون 29](#_Toc157206151)

[3-4 جمعبندی 30](#_Toc157206152)

[فصل چهارم روش پیشنهادی و روش پیاده‌سازی 31](#_Toc157206153)

[4-1 مقدمه 32](#_Toc157206154)

[4-2 نحوه پیاده‌سازی بخش تبدیل اعتبار کربن به رمزارز 32](#_Toc157206155)

[4-3 34](#_Toc157206156)

[4-3-1 نحوه ارتباط با سرورهای مرکزی 34](#_Toc157206157)

[4-4 راهکار یادگیری تقویتی برای تبدیل ERC-1155 به ERC-20 34](#_Toc157206158)

[4-5 نحوه ارزیابی و پیش‌آموزش برای استفاده در محیط واقعی 37](#_Toc157206159)

[4-6 جمع‌بندی 38](#_Toc157206160)

**فهرست شکل­ها**

[تصویر 1 مکانیزم تبدیل ارز غیر قابل عوض به قابل عوض 7](file:////Users/parsa/PycharmProjects/RLEnergy/BachelorProjectReport.docx#_Toc157206161)

[تصویر 2 شرح پل توکان 11](#_Toc157206162)

[تصویر 3 - معماری انرژی وب اریجین 15](#_Toc157206163)

[تصویر 4 - پروسه تبدیل اعتبار کربن به رمزارز غیرقابل تعویض 32](#_Toc157206164)

[تصویر 5- راهکار ارتباط با مراجع خارج از بلاک‌چین 34](#_Toc157206165)

[تصویر 6 - الگوی سر و شانه 38](#_Toc157206166)

**فهرست جدول­ها**

[جدول 1- مقایسه توکان و انرژی وب اریجین 18](#_Toc157197698)

# فصل اول: کلیّات

## مقدمه

کشورهای توافق پاریس، یا موافقت‌نامه پاریس در خصوص تغییر اقلیم، یک توافق بین المللی است که در ۱۲ دسامبر ۲۰۱۵ در پاریس، فرانسه، امضاء شد. این توافق به عنوان قطبی‌ترین و کامل‌ترین توافق در زمینه تغییر اقلیم تاکنون شناخته می‌شود. اهداف اصلی توافق پاریس عبارتند از کاهش افزایش دما به زیر ۲ درجه سانتی‌گراد نسبت به دوران قبل از صنعت‌گری و تلاش برای محدود کردن این افزایش به حداکثر ۱.۵ درجه سانتی‌گراد. توافق‌نامه پاریس تعهدات مالی، تعهدات کاهش گازهای گلخانه‌ای، و توسعه پایدار را نیز در بر می‌گیرد. این توافق همچنین بر اهمیت همکاری جهانی در مقابله با تغییر اقلیم تأکید دارد و تعهد به ارائه حمایت‌های مالی و فنی به کشورهای در حال توسعه را دارد. [1]

**اعتبار کربن** یک واحد اندازه‌گیری است که کاهش یا حذف یک تن دی‌اکسید کربن یا سایر گازهای گلخانه‌ای از جو را نشان می‌دهد. اعتبارات کربن معمولاً توسط پروژه‌هایی ایجاد می‌شوند که به کاهش انتشار گازهای گلخانه‌ای کمک می‌کنند. **منابع اعتبار کربن** می‌توانند شامل موارد زیر باشند:

* استفاده از انرژی‌های تجدیدپذیر، مانند انرژی خورشیدی و باد
* بهبود بهره‌وری انرژی، مانند عایق‌بندی ساختمان‌ها
* کاشت درخت، که دی‌اکسید کربن را از جو جذب می‌کند
* کاهش انتشار گازهای گلخانه‌ای از صنایع، مانند کاهش استفاده از سوخت‌های فسیلی

**اعتبارات کربن** می‌توانند به روش‌های مختلفی استفاده شوند. یکی از روش‌های استفاده از آنها، خرید و فروش در بازارهای کربن است. در این بازارها، شرکت‌ها و سازمان‌هایی که انتشار گازهای گلخانه‌ای خود را کاهش نمی‌دهند، می‌توانند اعتبارات کربن را از شرکت‌ها و سازمان‌هایی که انتشار گازهای گلخانه‌ای خود را کاهش داده‌اند، خریداری کنند. این امر به شرکت‌ها و سازمان‌ها کمک می‌کند تا به اهداف کاهش انتشار گازهای گلخانه‌ای خود دست یابند.

**اعتبارات کربن** همچنین می‌توانند برای جبران انتشار گازهای گلخانه‌ای استفاده شوند. به عنوان مثال، یک شرکت هواپیمایی ممکن است اعتبارات کربن را خریداری کند تا انتشار گازهای گلخانه‌ای ناشی از پروازهای خود را جبران کند.

**اعتبارات کربن** یک ابزار مهم برای مقابله با تغییر اقلیم هستند. آنها می‌توانند به شرکت‌ها و سازمان‌ها کمک کنند تا به اهداف کاهش انتشار گازهای گلخانه‌ای خود دست یابند و همچنین می‌توانند به جبران انتشار گازهای گلخانه‌ای کمک کنند. [2]

رمزارز اعتبار کربن[[1]](#footnote-1) یک نوع رمزارز است که با اعتبارات کربن پشتیبانی می‌شود. این اعتبارات کربن نماینده یک تن دی‌اکسید کربن هستند که از جو حذف شده است. این اعتبارات از طریق پروژه‌هایی که انتشار گازهای گلخانه‌ای را کاهش می‌دهند، ایجاد می‌شوند. سپس توسعه‌دهندگان پروژه می‌توانند این اعتبارات را به افراد یا شرکت‌هایی که می‌خواهند اثر محیطی خود را جبران کنند، بفروشند. برخی از شرکت‌های فناوری، اعتبارات کربن و توکن‌های[[2]](#footnote-2) رمزارز را با هم ترکیب کرده‌اند. به طور کلی، رمزارزهای اعتبار کربن می‌توانند باعث شوند تا بازار اعتبارات کربن شفاف‌تر و قابل دسترسی‌تر شود. [3]

بازار اعتبار کربن یک بازار جهانی است که در آن شرکت‌ها و افراد می‌توانند برای جبران انتشار کربن خود، اعتبارات کربن خریداری کنند. این اعتبارات از پروژه‌هایی مانند کاشت درختان، توسعه انرژی تجدیدپذیر و کاهش مصرف انرژی به دست می‌آیند. بازار اعتبار کربن سنتی دارای برخی از مشکلات است، از جمله:

* سوءاستفاده: برخی از شرکت‌ها از اعتبارات کربن برای دستیابی به اهداف زیست‌محیطی خود بدون انجام اقدامات واقعی برای کاهش انتشار کربن استفاده می‌کنند.
* عدم شفافیت: ردیابی اعتبارات کربن در بازار سنتی دشوار است.
* هزینه بالای معاملات: معاملات در بازار سنتی می‌تواند گران باشد.

بازار اعتبار کربن در رمز ارزها یک بازار جدید است که در آن اعتبارات کربن به عنوان توکن‌های رمزنگاری شده معامله می‌شوند. این بازار پتانسیل آن را دارد تا برخی از مشکلات بازار سنتی اعتبار کربن را برطرف کند. از مزایای این بازار می‌توان گفت:

* شفافیت: توکن‌های رمزنگاری شده اعتبار کربن را می‌توان به راحتی ردیابی کرد.
* کارایی: معاملات در بازار رمز ارزها معمولاً سریع‌تر و کارآمدتر از معاملات در بازار سنتی است.
* دسترسی: بازار رمز ارزها دسترسی بیشتری به سرمایه‌گذاران و شرکت‌ها دارد.

چالش‌ها

مقررات: مقررات بازار اعتبار کربن در رمز ارزها هنوز در حال توسعه است.

سوءاستفاده: توکن‌های رمزنگاری شده اعتبار کربن نیز می‌توانند مورد سوءاستفاده قرار گیرند.

بازار اعتبار کربن در رمز ارزها هنوز در مراحل اولیه توسعه خود قرار دارد. با این حال، این بازار پتانسیل آن را دارد تا به مقابله با تغییر اقلیم کمک کند.

بازار اعتبار کربن در رمز ارزها یک بازار نوظهور است که پتانسیل آن را دارد تا به مقابله با تغییر اقلیم کمک کند. این بازار هنوز در مراحل اولیه توسعه خود قرار دارد و چالش‌هایی نیز در پیش دارد. با این حال، این بازار پتانسیل آن را دارد تا بازار اعتبار کربن سنتی را متحول کند. [4]

اعتبارات کربن در سیستم‌های رمزارزهای انرژی‌های سبز به طور عمومی به صورت توکن‌های غیرقابل معاوضه در سیستم رمزارزهای تعریف می‌شود. یکی از راه‌کارهایی که سیستم‌های رمزارزی بازار انرژی بایستی ارائه کنند مکانیزم تبدیل این توکن‌ها به توکن‌های قابل معاوضه با استاندارد است تا از آن‌ها بتوان به عنوان یک ارز استفاده کرد..

## برخی مفاهیم

### مفهوم ERC-20

ERC-20 یک استاندارد قرارداد هوش مصنوعی در بلاکچین Ethereum است که برای ایجاد توکن‌های قابل تبادل بر اساس بلاکچین Ethereum استفاده می‌شود.

ERC-20 تعریف مشخصی برای توکن‌های مبادله‌پذیر یا تعویض‌پذیر ایجاد می‌کند و مشخصاتی مانند نام توکن، نماد توکن، تعداد کل توکن‌ها، تعداد اعشار توکن، و قابلیت‌های انتقال توکن را تعریف می‌کند. این استاندارد امکان تبادل و معامله توکن‌ها بین مخاطبان مختلف را فراهم می‌کند و به توکن‌های ERC-20 اجازه می‌دهد تا به راحتی در کیف‌های مختلف Ethereum ذخیره و مدیریت شوند.

به عبارت دیگر، ERC-20به توکن‌های مبادله‌پذیر بر اساس بلاکچین Ethereum یک ساختار و استاندارد مشخص می‌دهد تا امکان انتقال، ذخیره‌سازی و معامله‌ی آنها در اکوسیستم Ethereum را بهبود بخشد و این امکان را به کاربران و توسعه‌دهندگان می‌دهد که به راحتی با توکن‌های مختلف ارتباط برقرار کنند.

### مفوهم ERC-721

ERC-721 یک استاندارد قرارداد هوش مصنوعی برای توکن‌های غیرقابل تعویض [[3]](#footnote-3)در بلاکچین Ethereum است. توکن‌های ERC-721 یک منحصر به فرد هستند و هر یک دارای ویژگی‌ها و مشخصات منحصر به فردی هستند که آنها را از یکدیگر متمایز می‌کنند. به عبارت دیگر، هر توکن ERC-721 نماینده یک دارایی یا عنصر خاص در دنیای واقعی یا مجازی است.

از آنجا که توکن‌های ERC-721 منحصر به فرد هستند و هر یک دارایی یا عنصر خاصی را نمایندگی می‌کنند، این استاندارد به طور گسترده‌ای در برنامه‌ها و بازی‌های آنلاین به کار می‌رود که نیاز به تبادل دارایی‌های منحصر به فرد دارند.

### مفهوم ERC-1155

ERC-1155یک استاندارد قرارداد هوش مصنوعی برای توکن‌های چندپرکنشی [[4]](#footnote-4)در بلاکچین Ethereum است. ERC-1155 به توسعه‌دهندگان امکان ایجاد توکن‌هایی با انواع مختلفی از دارایی‌ها را می‌دهد. به عبارت دیگر با ERC-1155، یک قرارداد هوش مصنوعی می‌تواند توکن‌هایی ایجاد کند که همزمان انواع مختلفی از دارایی‌ها را نمایندگی کنند.

ویژگی‌های مهم ERC-1155 شامل موارد زیر هستند:

1. چندپرکنشی: توکن‌های ERC-1155 می‌توانند انواع مختلفی از دارایی‌ها را نمایندگی کنند. این انواع می‌توانند شامل توکن‌های غیرقابل تعویض (NFTs) و توکن‌های قابل تعویض (FTs) باشند. این امکان به توسعه‌دهندگان اجازه می‌دهد تا توکن‌هایی با ویژگی‌ها و قابلیت‌های مختلف ایجاد کنند.

2. کارایی بهتر: ERC-1155 باعث کاهش هزینه‌ها و پیچیدگی‌ها در مقایسه با استفاده از استانداردهای جداگانه برای NFTs و FTs می‌شود. این به معنای کاهش حجم داده‌های ذخیره‌سازی و تراکنش‌ها در بلاکچین است.

3. انتقال متعدد: توکن‌های ERC-1155 امکان انتقال متعدد را به صورت یک تراکنش در بلاکچین فراهم می‌کنند. این بدین معناست که می‌توانید تعداد زیادی از توکن‌ها را به یک مقصد انتقال دهید.

4. توسعه‌پذیری: ERC-1155 امکان توسعه‌دهندگان را در ایجاد برنامه‌ها و بازی‌هایی که نیاز به تعامل با توکن‌های متنوع دارند، تسهیل می‌کند.

ERC-1155 به عنوان یک استاندارد انعطاف‌پذیر و کارا در برنامه‌نویسی توکن‌های بلاکچین و ایجاد بازی‌های مبتنی بر بلاکچین با انواع مختلفی از دارایی‌ها بسیار مورد توجه قرار گرفته است.

## بیان مسئله

سیستم بازار اعتبارهای کربن یک مکانیزم برای تبدیل اعتبار کربن به ارز قابل معاوضه نیاز دارند. این سیستم تا به حال به این صورت بوده که به ازای هر تن کربن که از کره زمین حذف شده است یا خواهد شد، یک واحد رمزارز قابل عوض به فرد ارائه دهنده اعتبار داده می‌شود.

این کار اگرچه می‌تواند مناسب باشد ولی تنها راهکار نیست. داشتن یک بازار تبدیل توکن‌های بلاعوض به توکن‌های قابل تعویض می‌تواند با تعویض مکانیزم گفته شده با یک استخر سرمایه‌گذاری رمزارزی بسیار تغییر کند. به این صورت که به افراد دارنده اعتبار کربن، به ازای مدت زمان نگه‌داری اعتبار نام‌برده در استخر با فرکانسی توکن قابل تعویض ارائه شود.

حال سوالی که مطرح می‌شود آن است که فرکانس ارائه توکن قابل تعویض در ازای توکن غیرقابل تعویض به چه صورت باشد؟

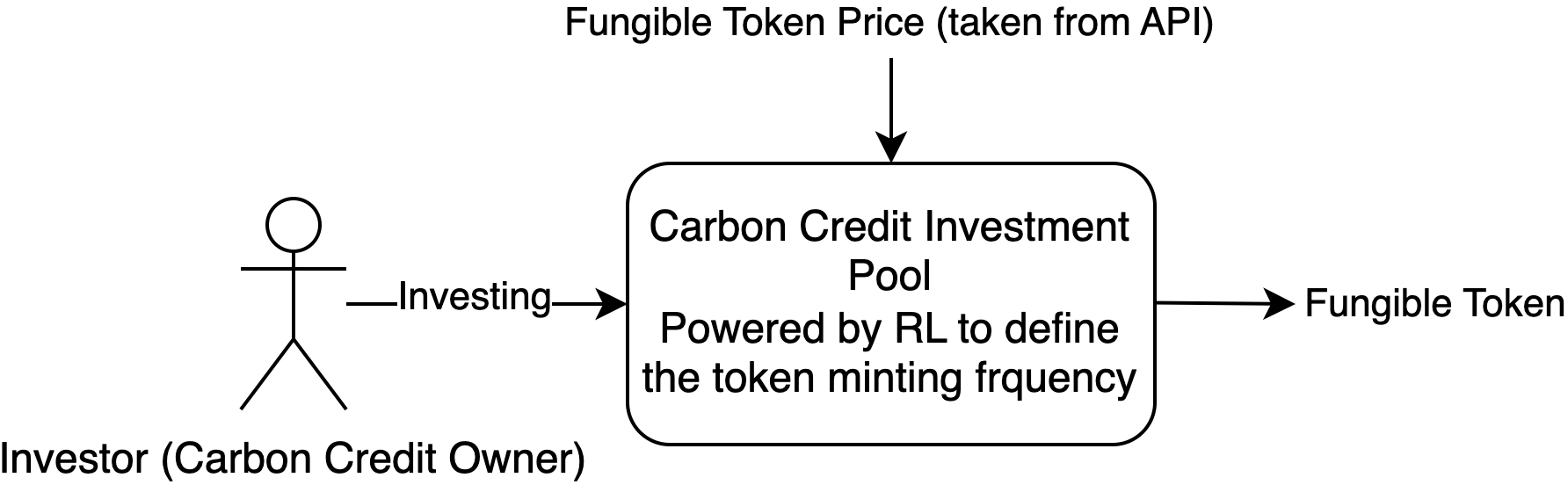
.

## کلیات روش پیشنهادی

در این جا می‌توان از یادگیری تقویتی[[5]](#footnote-5) بهره برد به شرط آن که بتوان مولفه‌های آن را به درستی تعریف نمود. این مولفه‌ها به صورت زیر هستند:

1. **وضعیت:** وضعیت در این مسئله قیمت ارز قابل تعویض و میزان فرکانس سرمایه‌گذاری در استخر یا صندوق سرمایه‌گذاری می‌باشد.
2. **امتیاز و پاداش:** امتیاز در این مسئله افزایش میزان سرمایه‌گذاری یا افزایش نرخ ارز خروجی است. هم‌چنین کاهش این دو نیز می‌تواند عامل منفی در نظر گرفته شود.
3. **عمل عامل:** عامل می‌تواند با تغییر نرخ تولید ارز قابل معاوضه میزان موجودیت آن را در بازار تغییر بدهد.
4. **تاثیر عامل بر محیط و بیان تعادل لازم:** عامل یادگیرنده تقویتی در اینجا با تعیین میزان موجودیت ارز در بازار می‌تواند روی قیمت تاثیر گذار باشد. افزایش قیمت می‌تواند منجر به افزایش میزان سرمایه‌گذاری نیز شود. پس کاهش نرخ توضیه منجر به افزایش قیمت و افزایش قیمت نیز منجر به افزایش میزان سرمایه‌گذاری جهت کسب ارز قابل معاوضه می‌شود. از طرف دیگر هرچه میزان تولید ارز قابل تعویض بیشتر باشد میزان این صندوق سرمایه‌گذاری برای سرمایه‌گذار جذاب‌تر است زیرا میزان بیشتری ارز می‌تواند از آن کسب کند و اگر این میزان کم باشد دیگر کسی مایل به سرمایه‌گذاری در این صندوق نیست زیرا دیگر آورده‌ای برای اون ندارد.

**هدف** از این پروژه پیاده‌سازی سیستم رمزارزهای اعتبارات کربن با استفاده از یادگیری ماشینی به عنوان عامل تبدیل توکن‌های غیر قابل تعویض به قابل تعویض می‌باشد



## ساختار پروژه

در فصل‌هایی که در ادامه می‌آیند در فصل دوم به بیان سیستم‌های فعلی برای بازار رمزارزهای اعتبار کربن پرداخته و سپس در فصل‌های سوم به راه‌کارهای یادگیری تقویتی پرداخته و به نقاط ضعف و قدرت هر یک از آن‌ها برای پروژه خود می‌پردازیم. در فصل چهارم نیز راهکار خود را ارائه نموده و نتیجه‌گیری می‌کنیم.

تصویر 1 مکانیزم تبدیل ارز غیر قابل عوض به قابل عوض

# فصل دوم: سیستم‌های رمزارزی اعتبار کربن موجود

## مقدمه

در حوزه سیستم‌های رمزارزهای اعتبارات کربن دو محصول تجاری به حد مطلوب رسیده‌اند. یکی سیستم توکان [5]و دیگری سیستم انرژی وب اریجین [6] در دو بخش به هر یک از آن‌ها و مزایا و معایب آن‌ها می‌پردازیم.

## بررسی سیستم توکان

### بازار انرژی توکان

توکان به ارائه زيرساخت‌های ديجيتال برای اعتبارهای کربنی توکنيزه شده می‌پردازد. هدف اصلی اين سايت، توسعه داوطلبانه کربن[[6]](#footnote-6) با شفافيت و اعتبار بالا است. در اين راستا، از فناوری بلاکچين برای تضمين شفافيت و صداقت در معاملات اعتبار کربنی استفاده می‌شود.

تمام اطلاعات مربوط به اعتبار کربنی در يک پايگاه داده مبتنی بر بلاکچين ثبت می‌شوند که اين امر به همگان اجازه می‌دهد تا به صورت مستقل از ويژگی‌ها، تاريخچه معاملات و داده‌های قيمتی هر اعتباری مطلع شوند. توکان ارتباط نزديکی با متخصصان، سازمان‌های ثبتی و ديگر نهادهای صنعتی دارد تا از مزايای فناوری‌هاي جديد بهره‌مند شوند، در حالی که به دقت خطرات را مديريت می‌کنند.

زيرساخت‌های توکن تحت قوانين سوئيس فعاليت می‌کنند و برنامه‌هايی برای اجرای تدابير احراز هويت مشتری[[7]](#footnote-7) به منظور حفظ يکپارچگی اعتبارهای کربنی توکنيزه شده دارند. اين زيرساخت ها اعتبار کربنی را با برنامه های کاربردی بلاکچين متصل می کنند که نتيجه آن امکانات نامحدود برای نوآوری در نسل بعدی محصولات و انگيزه های مثبت برای اقليم است تا اقدامات آب و هوايی را در مقياس وسيعی فعال کنند.

از زمان راه اندازی توکان در اکتبر 2021، اتفاقات بزرگی در زمينه بازنشستگی اعتبار کربنی رخ داده است. پلتفرم برای کاربران متنوعی طراحی شده است، از جمله کسانی که به دنبال نظارت بر فعاليت بازار يا استفاده از داده‌های معاملاتی قابل دسترسی عمومی هستند.

در ﺑﺨﺶ ﺗﻮﺻﯿﻔﺎت، ﺳﺎزﻣﺎن‌های ﻣﺨﺘﻠﻔﯽ به ﺗﺎﯾﯿﺪ ﻧﻘﺶ ﺗﻮﮐاﻦ در مقابله ﺑﺎ ﺗﻐﯿﯿﺮات آب و هوایی ﺑﺎ اﺳﺘﻔﺎده از ﻓﻨﺎوری وب ۳ اﺷﺎره ﮐﺮده‌اﻧﺪ. ﻣﺤﺼﻮﻻت اﺻﻠﯽ ﺳﺎﯾﺖ ﺷﺎﻣﻞ ﭘﻞ‌های ﮐﺮﺑﻨﯽ ﺑﺮای ﺗﻮﮐﻨﯿﺰه[[8]](#footnote-8) ﮐﺮدن اعتبارهای ﮐﺮﺑﻨﯽ ﻣﻮﺟﻮد در ثبت‌های ﮐﺮﺑﻨﯽ ﻣﻌﺘﺒﺮ، اﺳﺘﺨﺮھﺎی ﮐﺮﺑﻨﯽ و ﺑﺎزﻧﺸﺴﺘﮕﯽ ﮐﺮﺑﻨﯽ اﺳﺖ ﮐﮫ اﻣﮑﺎن ﺧﺮ ﯾﺪ و ﺑﺎزﻧﺸﺴﺘﮕﯽ اﻋﺘﺒﺎر ﮐﺮﺑﻨﯽ را ﺑﺎ ﺷﻔﺎﻓﯿﺖ فراهم ﻣﯽﮐﻨﺪ.

در نهایت، ﻣﺎﻣﻮرﯾﺖ اﺻﻠﯽ توکان اﯾﺠﺎد زیرساخت‌های ﻣﺎﻟﯽ ﺑﺮای هدایت تریلوین‌ها دﻻر ﺑﮫ ﺳﻤﺖ بهترین ﯾﻦ راه‌حل‌های آب و هوایی اﺳﺖ. ﺳﺎﯾﺖ توکان هم‌چنین از همکاری‌های ﺧﻮد ﺑﺎ سازمان‌ها و نهادهای ﻣﺨﺘﻠﻒ در ﺣﻮزه اﻗﺪاﻣﺎت آب و هوایی و ﻓﻀﺎی ﺑﻼﮐﭽﯿﻦ ﯾﺎد ﻣﯽ ﮐﻨﺪ.

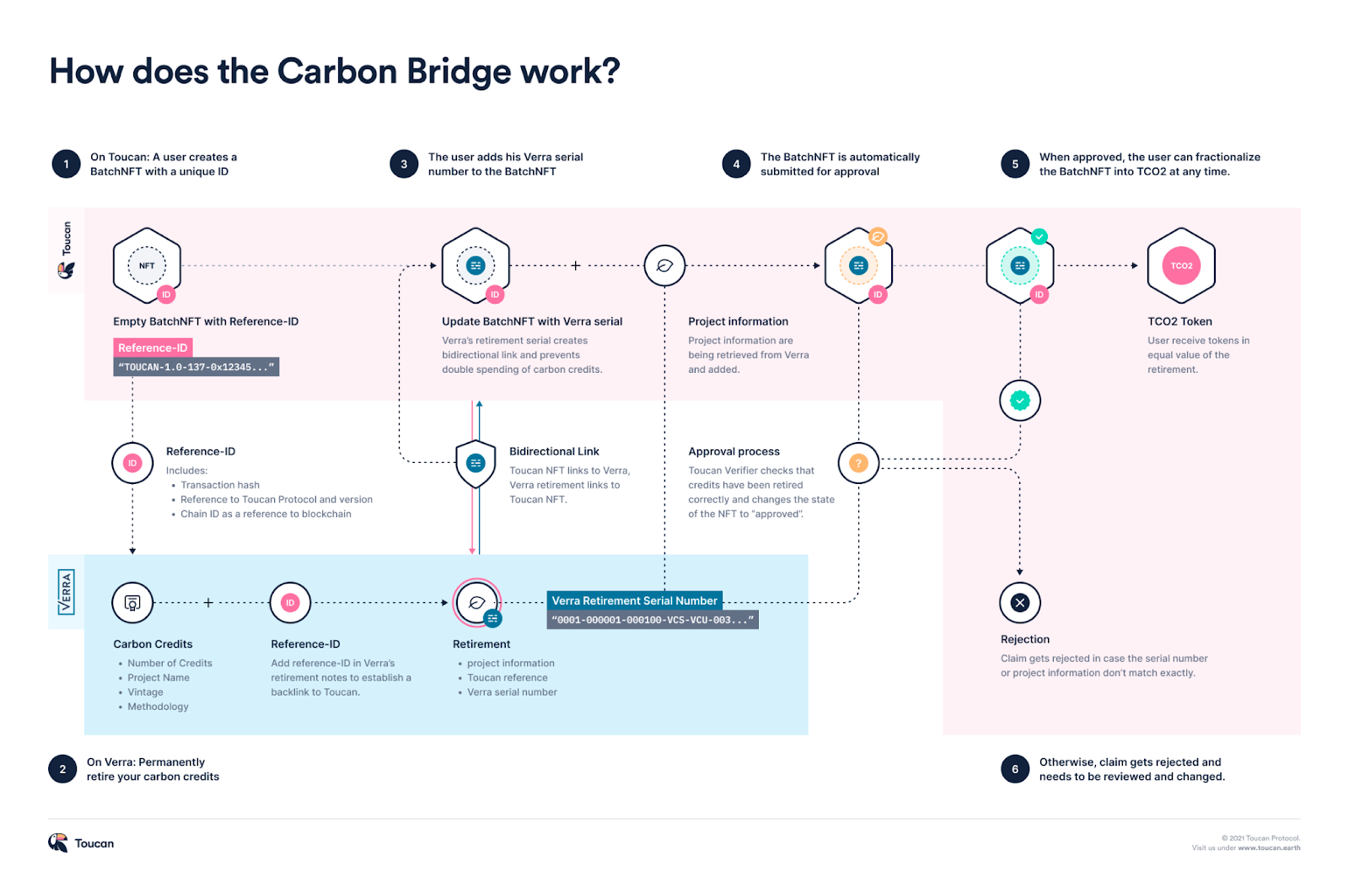
ﭘل ﮐرﺑن ﺗوﮐﺎن ﯾﮏ اﺑﺗﮑﺎر نوآورانه اﺳت که به اﻓراد و سازمان‌ها اﻣﮑﺎن می‌دهد ﺗﺎ اﻋﺗﺑﺎرات ﮐرﺑﻧﯽ ﺧود را به ﺷﮑل توکن‌های دیجیتال در بلاک‌چین ﻗرار دهند. اﯾن فرآیند تبدیل، اعتبارات کربنی سنتی را به دارایی‌های دیجیتالی ﺗﺑدﯾل می‌کند که ﻣﯽﺗواﻧﻧد به راحتی در بازارهای ﻣﺎﻟﯽ دﯾﺟﯾﺗﺎل ﻣورد اﺳﺗﻔﺎده و ﻣﻌﺎﻣﻠﮫ ﻗرار گیرند. اﯾن تحول، فرصت‌های جدیدی را ﺑرای سرمایه‌گذاری و مدیریدت اعتبارات کرنبی فراهم ﻣﯽﮐﻧد.

یکی از مهمترین ویژگی‌های توکن‌های کربنی، شفافیت آن‌ها است. با استفاده از فناوری بلاک‌چین» هر تراکنش و نگهداری این توکن‌ها به طور کامل قابل ردیابی و شفاف است. اين ویژگی به افزایش اعتماد و شفافیت در بازار کربن کمک می‌کند و اطمینان می‌دهد که اعتبارات کربنی به درستی مورد استفاده قرار می‌گیرند. علاوه بر شفافیت» توکن‌سازی اعتبارات کربنی امکان برنامه‌ریزی و تنظیم دقیق این دارایی‌ها را فراهم می‌کند. اين امر به سازمان‌ها و فعالان محیط زیست اجازه می‌دهد تا استراتژی‌های پیچیده‌تری را برای مدیریت اعتبارات کربنی خود طراحی کنند. همچنین» توکن‌های کربنی می‌توانند به راحتی تقسیم شوند» که اين امر دسترسی به بازار کربن را برای سرمایه‌گذاران کوچکتر و فردی ساده‌تر می‌کند.

یکی دیگر از مزایای مهم توکن‌سازی کرین» ترکیب‌پذیری آن با اکوسیستم مالی غیرمتمرکز [[9]](#footnote-9) است. اين ویژگی به توکن‌های کربنی اجازه می‌دهد تا در محصولات و خدمات مالی مختلف مورد استفاده قرار گیرند» از جمله در پروژه‌های سرمایه‌گذاری پایدار و تأمین مالی سبز.

در نهایت، پل کربن توکان بخشی از تلاش‌های گسترده‌تر برای آوردن دارایی‌های محیطی به فضای دیجیتال است. اين ابتکار عمل به دنبال تسهيل دسترسی، مديريت و استفاده از اعتبارات کربنی در مبارزه با تغييرات آب و هوايی است و امکانات جديدی را برای سرمايه گذاری مسئولانه و پايدار در زمينه محيط زيست فراهم می‌کند.

### سیستم پل در توکان



تصویر 2 شرح پل توکان

تصوير نشان‌دهنده نحوه کار "پل کربن" است. مراحل مختلف اين فرآیند به شرح زیر است:

1. در توکان: یک کاربر یک توکن غیر قابل تعویض (دسته‌ای با یک شناسه‌ی منحصر به فرد ایجاد می‌کند. اين توکن خالی با یک "شناسه‌ی مرجع" ارائه می‌شود. شناسه‌ی مرجع شامل اطلاعاتی مانند هش معامله، ارجاع به پروتکل توکان و نسخه» و همچنین شناسه‌ی زنجیره‌ی بلوکی است.

2. در ورا[[10]](#footnote-10) می‌توان اعتبارات کربنی خود را برای هميشه بازنشسته کرد. این شامل تعداد اعتبارات، نام پروژه، و روش است.

3. کاربر شماره‌ی سریال ورا خود را به توکن دسته‌ای اضافه می‌کند. وقتی شماره‌ی سریال ورا به توکن اضافه می‌شود. یک پیوند دوسویه ایجاد می‌شود که جلوی مصرف دو برابر اعتبارات کربن را می‌گیرد.

۴. توکن غیر قابل عوض دسته‌ای به طور خودکار برای تأیید ارسال می‌شود. اطلاعات پروژه از ورا بازیابی شده و اضافه می‌شود. فرآیند تأیید شروع می‌شود. در اینجا سازمان توکان اطمینان حاصل می‌کند که اعتبارات به درستی بازنشسته شده‌اند و وضعیت توکن غیر قابل عوض به "تایید شده" تغییر می‌دهد.

۵.. وقتی تأیید شده است، کاربر می‌تواند توکن غیر قابل عوض دسته‌ای را در هر زمان به توکن‌های قابل عوض با امکان نقدینگی کم با نام TCO2 تجزیه کند. کاربران توکن‌ها را با ارزش معادل بازنشستگی دریافت می‌کنند.

6. در غیر این صورت. اگر اطلاعات دقیقاً مطابقت نداشته باشد» ادعا رد می‌شود و باید مورد بازبینی و تغییر قرار گیرد.

برخی نکات کلیدی عبارتند از:

* توکنیزاسیون اعتبارات کربن: پس از توکنیزه و تجزیه شدن اعتبارات کربن یک پروژه با استفاده از پل کربن» اين

اعتبارات به صورت توکن‌های TCO2نمایش داده می‌شوند.

* ویژگی‌های مخصوص پروژه: هر قرارداد توکن TCO2 تمام ویژگی‌های مخصوص پروژه را حفظ می‌کند×

مانند کشور یا روش‌شناسی پروژه.

* عدم قابلیت تبادل: توکن‌ها از یک پروژه با توکن‌ها از پروژه دیگر قابل تبادل نیستند. حتی اگر بسیار شبیه به هم

باشند.

هدف از این رویکرد، ساده‌سازی فرآیند معامله با اعتبارات کربن و ایجاد یک بازار استاندارد و نقدی‌تر برای آن‌ها است. ویژگی‌های یک توکن کربن می‌تواند به صورت زیر باشد:

standard = Puro — only accepts TCO2s from the Puro standard

vintage = >2021 — only accepts TCO2s from 2021 and later.

country = Colombia — only accepts TCO2s from Colombia.

### بررسی استخرها در سیستم توکان

استخرهای کربن در واقع روشی برای گروه‌بندی چندین تن کربن توکنیزه شده مخصوص پروژه (توکن‌های TCO2) در قالب توکن‌های استخر قابل تبادل هستند. این سیستم به منظور تسهیل در نقدینگی و کشف قیمت برای کلاس‌های مختلف دارایی‌های کربنی طراحی شده است.

استخرهای کربن، چندین تن کربن دی‌اکسید (TCO2) توکنیزه شده خاص پروژه را به توکن‌های استخر قابل تبدیل ترکیب می‌کنند. این امر نقدینگی و کشف قیمت را برای انواع مختلف دارایی‌های کربنی فراهم می‌کند.

**نحوه عملکرد استخرهای کربن**

پس از اینکه اعتبارات کربن یک پروژه با استفاده از پل کربن توکنیزه و خرد شد، این اعتبارات به عنوان TCO2 توکن‌ها (تن‌های کربن توکنیزه شده) نشان داده می‌شوند. قرارداد یک توکن TCO2 همچنان حاوی تمام ویژگی‌های خاص پروژه، مانند کشور یا روش‌شناسی پروژه است. این بدان معناست که توکن‌های یک پروژه با توکن‌های پروژه دیگر قابل تعویض نیستند، حتی اگر بسیار مشابه باشند.

سپس توکن‌های TCO2 می‌توانند در استخرها سپرده شوند که نقدینگی و قابلیت تبدیل را در سراسر انواع پروژه‌های مشابه بهبود می‌بخشند.

**ترجمه فارسی**

استخرهای کربن یک فناوری جدید هستند که برای تسهیل تجارت دارایی‌های کربن طراحی شده‌اند. آنها با ترکیب چندین تن کربن دی‌اکسید (TCO2) توکنیزه شده خاص پروژه به توکن‌های استخر قابل تبدیل، نقدینگی و کشف قیمت را برای انواع مختلف دارایی‌های کربنی فراهم می‌کنند.

عملکرد استخرهای کربن به شرح زیر است:

ابتدا، اعتبارات کربن یک پروژه با استفاده از یک پلتفرم توکن‌سازی مانند Carbon Bridge توکنیزه می‌شوند. این بدان معناست که آنها به عنوان یک دارایی دیجیتالی منحصر به فرد که می‌تواند در بازارهای مالی معامله شود، نمایندگی می‌شوند.

سپس، این اعتبارات کربن توکنیزه شده می‌توانند در یک استخر کربن سپرده شوند. استخر کربن یک قرارداد هوشمند است که مدیریت توکن‌های کربن را در یک گروه خاص کنترل می‌کند.

با سپرده‌گذاری توکن‌های کربن در استخر، آنها به توکن‌های استخر قابل تبدیل تبدیل می‌شوند. توکن‌های استخر قابل تبدیل از نظر ارزش برابر هستند و می‌توانند با یکدیگر معامله شوند.

استخرهای کربن می‌توانند مزایای متعددی برای بازارهای کربن داشته باشند. آنها می‌توانند:

* نقدینگی بازار کربن را افزایش دهند.
* کشف قیمت را بهبود بخشند.
* سرمایه‌گذاری در پروژه‌های کاهش انتشار کربن را آسان‌تر کنند.

استخرهای کربن یک فناوری جدید و نوظهور هستند که پتانسیل تغییر بازارهای کربن را دارند. آنها می‌توانند به کاهش انتشار کربن و مبارزه با تغییرات آب و هوایی کمک کنند.

## بررسی سیستم انرژی وب اریجین

انرژي وب اریجین یک شبکه بلاک‌چینی مخصوص صنعت انرژی است که به منظور ارتقاء انتقال و مدیریت انرژی تجدیدپذیر و پایدار توسعه یافته است. این پروژه به عنوان یک شبکه برای اشیاء [[11]](#footnote-11) انرژی بازارهای جدید را فراهم می‌کند و از تکنولوژی بلاک‌چین برای تضمین امنیت، شفافیت، و کارایی در انتقال اطلاعات و انرژی استفاده می‌کند.

مهمترین اهداف انرژي وب اریجین عبارتند از:

1. ایجاد بازار انرژی دیجیتال: انرژي وب اریجین به تشکیل بازارهای انرژی دیجیتال کمک می‌کند که انرژی تولیدی از منابع تجدیدپذیر مانند باد، خورشید و سایر منابع پایدار را به تولیدکنندگان و مصرف‌کنندگان انرژی متصل کند.
2. افزایش شفافیت: با استفاده از تکنولوژی بلاک‌چین، تمامی تراکنش‌ها و معاملات مربوط به انتقال انرژی به شفافیت بیشتری دست خواهند یافت و اطلاعات دقیقی از میزان تولید و مصرف انرژی در دسترس قرار خواهد گرفت.
3. امنیت: امنیت اطلاعات و انتقال انرژی در انرژي وب اریجین با استفاده از تکنولوژی بلاک‌چین تقویت می‌شود، که از حملات سایبری و تغییرات غیرمجاز جلوگیری می‌کند.
4. ایجاد سیستم‌های انرژی هوش مصنوعی: انرژي وب اریجین به توسعه سیستم‌های هوش مصنوعی که بهینه‌سازی مصرف انرژی و مدیریت آن را بهبود می‌بخشند، کمک می‌کند.
5. حمایت از تجدیدپذیری و پایداری: این پروژه از تولید انرژی از منابع تجدیدپذیر و کاهش انتشار گازهای گلخانه‌ای حمایت می‌کند.

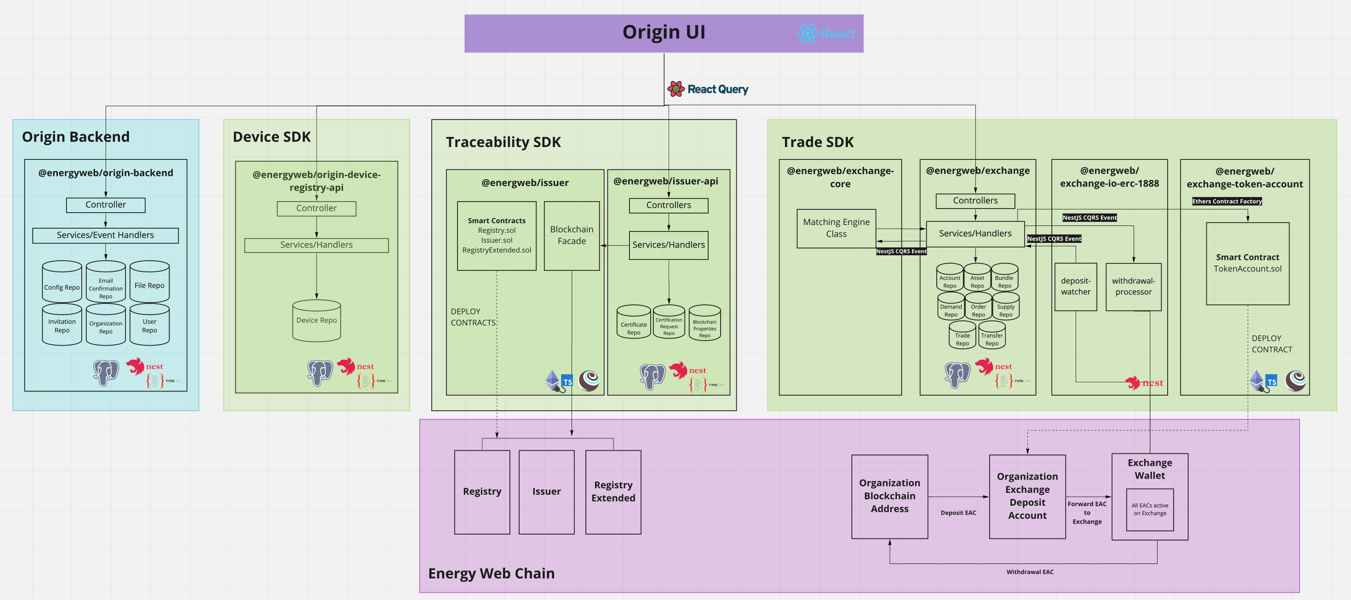
انرژي وب اریجین یکی از پروژه‌های مهم در زمینه بلاک‌چین در صنعت انرژی است که به منظور بهبود مدیریت و انتقال انرژی تجدیدپذیر و پایدار به کار می‌رود.

انرژي وب اریجین یک مجموعه از کیت‌های توسعه نرم‌افزاری[[12]](#footnote-12) و سرویس‌های سمت سرویس‌دهنده[[13]](#footnote-13) است که در مجموع یک پلتفرم برای صدور، مدیریت و معامله گواهینامه‌های ویژگی انرژی[[14]](#footnote-14) ارائه می‌کنند.

گواهینامه ویژگی انرژی یک سند رسمی است که تضمین می‌کند که انرژی تولید شده از منابع تجدیدپذیر می‌آید. استانداردهای مختلفی وجود دارند که نحوه ذخیره و اعتبارسنجی داده‌ها را تنظیم می‌کنند. در اروپا، این سند به نام «گواهی اصالت»[[15]](#footnote-15) شناخته می‌شود، در شمال آمریکا «گواهینامه انرژی تجدیدپذیر»[[16]](#footnote-16) نام دارد و در بخش‌هایی از آسیا، آفریقا، خاورمیانه و آمریکای لاتین استاندارد حاکم بر آن «گواهی انرژی تجدیدپذیر بین‌المللی»[[17]](#footnote-17)است. استانداردها ممکن است متفاوت باشند، اما همه از اصول اساسی مشترکی بهره می‌برند.

قصد اصلی گواهینامه‌های ویژگی انرژی اثبات آن است که انرژی مصرف شده از منابع تجدیدپذیر به دست آمده است. گواهینامه‌های ویژگی انرژی می‌توانند در یک بازار معاملاتی معامله و خریداری شوند.

### بررسی معماری انرژی وب اریجین



تصویر 3 - معماری انرژی وب اریجین

این تصویر یک معماری نرم‌افزاری را نشان می‌دهد که برای پلتفرمی با نام "Energy Web Chain" طراحی شده است. پلتفرم به چند بخش اصلی تقسیم شده است که هر کدام دارای ماژول‌ها و کلاس‌های مختلفی هستند. در اینجا به شرح هر بخش می‌پردازیم:

1. پشتگاه[[18]](#footnote-18) اریجین: این بخش شامل یک کنترلر و سرویس‌ها/رسیدگی‌کننده‌های رویداد است که به پیکربندی، ایمیل، فایل، دعوت‌نامه، سازمان و مخزن کاربران مرتبط می‌شود. این قسمت از نرم‌افزار عمدتا برای مدیریت پشتیبانی و پیکربندی سیستم استفاده می‌شود.

2. وسیله کیت توسعه نرم‌افزار: این بخش نیز شامل یک کنترلر و سرویس‌هایی است که به یک مخزن دستگاه متصل هستند. این کیت توسعه نرم‌افزار ممکن است برای تعامل با دستگاه‌های فیزیکی در شبکه و ثبت اطلاعاتشان در سیستم استفاده شود.

3. کیت توسعه نرم‌افزار ردیابی: این بخش شامل کنترلرها و سرویس‌هایی است که با قراردادهای هوشمند و یک نمای بلاکچین ارتباط دارند. قراردادهای هوشمند شامل ثبت‌نام‌ها و صدور گواهینامه‌ها هستند. این بخش برای ردیابی و اعتبارسنجی تراکنش‌ها و اسناد در شبکه استفاده می‌شود.

4. کیت توسعه نرم‌افزار معامله: این بخش شامل یک موتور مطابقت و کنترلرها و سرویس‌ها/رسیدگی‌کننده‌هایی است که به حساب‌ها، دارایی‌ها، سفارش‌ها و معاملات مرتبط هستند. این بخش به نظر می‌رسد که برای تسهیل معاملات و تبادل‌های تجاری در پلتفرم طراحی شده است.

Energy Web Chain شامل سه بخش اصلی است: ثبت[[19]](#footnote-19)، صادرکننده و ثبت گسترده[[20]](#footnote-20) است.. این قسمت‌ها ممکن است از بلاک‌چین هستند که برای ثبت و صدور گواهینامه‌ها و سایر اسناد مورد استفاده قرار می‌گیرند.

#### کیت توسعه نرم‌افزار ردیابی

کیت توسعه نرم‌افزار ردیابی ردیابی کیت توسعه نرم‌افزار مسئول درخواست، صدور و مبادله گواهینامه‌های ویژگی انرژی در پلتفرم اریجین است. گواهینامه‌ها توسط صاحبان دستگاه درخواست می‌شوند و توسط صادرکنندگان تأیید می‌شوند. صادرکننده نهادی است که مسئول بررسی شواهد تولید انرژی و صدور گواهینامه به گونه‌ای است که با استانداردهای قانونی و صنعتی مطابقت داشته باشد.

کیت توسعه نرم‌افزار ردیابی دارای دو بسته اصلی است:

صادرکننده

بسته صادرکننده مدیریت گواهینامه‌ها را در بلاک چین انجام می‌شود و دارای دو مؤلفه اصلی است:

* قراردادهای هوشمند که چرخه عمر گواهینامه‌ها را در زنجیره بلوک حفظ می‌کنند.
* نمایه‌ها برای تعامل با توابع قرارداد هوشمند. صادرکننده از این نمایه‌ها استفاده می‌کند تا به جای تعامل مستقیم با بلاک چین، از آنها استفاده کند.

API صادرکننده:

گواهینامه‌های ویژگی انرژی توسط نهادهای رسمی به دستگاه‌های تولیدکننده برق صادر می‌شوند تا گواهی کنند که در بازه زمانی مشخص شده، مقدار مشخصی انرژی سبز تولید کرده‌اند. نهادهای صادرکننده رسمی برای مناطق مختلف وجود دارند. گواهینامه‌ها باید اطلاعات دستگاه تولید، حجم کل انرژی تولید شده و بازه زمانی که انرژی در آن تولید شده را در نظر بگیرند. حجم انرژی تولید شده سپس می‌تواند به واحدهای کوچکتر تقسیم شود و به عنوان گواهینامه‌های ویژگی انرژی به خریداران مختلف فروخته شود. توجه داشته باشید که رابط پیاده‌سازی مرجع Origin از Mwh به عنوان واحد استاندارد حجم انرژی استفاده می‌کند، اما هر واحدی می‌تواند بسته به نیازهای پیاده‌سازی استفاده شود. گواهینامه‌هایی که صادر می‌شوند:

* منحصر به فرد هستند.
* قابل ردیابی هستند.
* قابل تأیید هستند.

ردیابی SDK یک ابزار قدرتمند برای مدیریت گواهینامه‌های ویژگی انرژی در بلاک چین است. این SDK اطمینان حاصل می‌کند که گواهینامه‌ها منحصر به فرد، قابل ردیابی و قابل تأیید هستند.

#### بررسی کیت توسعه نرم‌افزار معامله

کیت توسعه نرم‌افزار معالمه برای سازگاری با کیت توسعه نرم‌افزار ردیابی طراحی شده است که یک کیت توسعه نرم‌افزار دیگر برای پلتفرم اریجین است. کیت توسعه نرم‌افزار ردیابی مسئول مدیریت چرخه عمر گواهینامه‌های ویژگی انرژی در بلاک‌چین است. با ادغام با کیت توسعه نرم‌افزار ردیابی، کیت توسعه نرم‌افزار تجارت می‌تواند اطمینان حاصل کند که تنها گواهینامه‌های ویژگی انرژی معتبر در صرافی معامله می‌شوند.

کیت توسعه نرم‌افزار تجارت دارای چهار بسته اصلی است:

**هسته صرافی:** این بسته عملکرد اصلی صرافی را ارائه می‌دهد، مانند مطابقت سفارشات و تسویه حساب.

**صرافی:** این بسته رابط سطح بالاتری را برای ساخت صرافی‌ها ارائه می‌دهد. این پیچیدگی بسته هسته صرافی را خلاصه می‌کند و توسعه صرافی‌های سفارشی را آسان‌تر می‌کند.

**حساب توکن صرافی:** این بسته راهی برای مدیریت گواهینامه‌های ویژگی انرژی در صرافی ارائه می‌دهد. این به توسعه‌دهندگان اجازه می‌دهد تا تعادل گواهینامه ویژگی انرژی را ایجاد، به‌روزرسانی و حذف کنند.

**Exchange IO ERC1888:** این بسته راهی برای تعامل با استاندارد ERC1888 برای مبادلات اتمی ارائه می‌دهد. مبادلات اتمی نوعی تراکنش است که به دو طرف اجازه می‌دهد تا توکن‌ها را در بلاک‌چین‌های مختلف بدون نیاز به واسطه مورد اعتماد مبادله کنند.

کیت توسعه نرم‌افزار تجارت یک ابزار قدرتمند است که می‌توان از آن برای ساخت انواع صرافی‌های سفارشی برای مدیریت گواهینامه‌های ویژگی انرژی استفاده کرد. این برای استفاده آسان و ادغام با سایر کیت‌های توسعه نرم‌افزار طراحی شده است.

**مزایای استفاده از کیت توسعه نرم‌افزار تجارت:**

* **کیت توسعه نرم‌افزار** تجارت یک ابزار قدرتمند و انعطاف‌پذیر است که می‌توان از آن برای ساخت انواع صرافی‌های سفارشی برای گواهینامه‌های ویژگی انرژی استفاده کرد.
* **کیت توسعه نرم‌افزار** تجارت با **کیت توسعه نرم‌افزار** ردیابی برای اطمینان از معامله تنها گواهینامه‌های ویژگی انرژی معتبر ادغام شده است.
* **کیت توسعه نرم‌افزار** تجارت برای سازگاری با سایر سیستم‌های مبتنی بر بلاک‌چین طراحی شده است.

**کاربردهای کیت توسعه نرم‌افزار تجارت:**

* **کیت توسعه نرم‌افزار** تجارت می‌تواند برای طیف وسیعی از کاربردها استفاده شود، از جمله:
* تسهیل مبادلات گواهینامه ویژگی انرژی بین تولیدکنندگان برق تجدیدپذیر و خریداران
* ایجاد بازارهای ثانویه برای گواهینامه‌های ویژگی انرژی
* ردیابی و گزارش تولید برق تجدیدپذیر

**کیت توسعه نرم‌افزار** تجارت یک ابزار ارزشمند برای توسعه‌دهندگانی است که می‌خواهند در زمینه انرژی تجدیدپذیر مشارکت کنند.

## جمع­بندی

در کل می‌توان مزایا و معایب دو سیستم را به صورت جدول زیر خلاصه کرد.

جدول - مقایسه توکان و انرژی وب اریجین

|  |  |
| --- | --- |
| توکان | انرژی وب اریجین |
| محوریت پل کربن برای تبدیل به TCO2. | تنها اعتبار توکنیزه می‌شود. |
| محوریتی برای انتساب اعتبار به دستگاه ندارد. | اعتبار کربن را به دستگاه تولید انرژی سبز منتسب می‌کند. |
| محدود به تولید انرژی سبز نیست. بلکه مقابله با جنگل‌زدایی را نیز در دسته اعتبارات خود می‌داند. | اعتبار کربن را محدود به تولید انرژی سبز کرده. |
| امکان تولید رمزارز بنا به میزان تولید انرژی سبز به صورت آنلاین را ندارد. | بالقوه می‌تواند رمزارز را با توجه به میزان تولید انرژی سبز در دستگاه‌های تولید با کمک IOT و فناوری Oracle تولید کند. |
| پروژه کماکان با توجه به دشواری‌های سر راهش فعال است. | توسعه پروژه متوقف شده و آخرین نسخه آن نیز قابل مستقرسازی[[21]](#footnote-21) و اجرا نیست. |

ساختار پروژه توکان برای پروژه ما مطلوب‌تر بوده با این تفاوت که پروژه ما به جای استخرهای پروژه توکان که توکن‌های TCO2 را یک‌بار فقط به توکن‌های با نقدینگی بیشتر تبدیل کند، آن‌ها را در خود نگه داشته و با توجه به شرایط متفاوت با فرکانس‌های متفاوت توکن‌های با به نقدینگی بالا تولید می‌کند. یعنی استخرها در این پروژّ معادل صندوق سرمایه‌گذاری رمزارزی خواهند بود. خال مسئله اصلی فرکانس تولید توکن با نقدینگی بالاست. این فرکانس باید تعادل بین تمایل مردم به سرمایه‌گذاری بیشتر و نقدینگی و قیمت توکن را حفظ کند.

# فصل سوم: بررسی روش‌های یادگیری تقویتی

## مقدمه

یادگیری تقویتی آن است که چه کاری انجام دهیم؛ چگونه موقعیت ها را به اقدامات نگاشت کنیم؛ تا حداکثر سیگنال پاداش عددی را بدست آورده شود. مانند بسیاری از اشکال یادگیری ماشینی، به یادگیرنده گفته نمی‌شود که کدام اقدامات را انجام دهد، بلکه باید با امتحان آنها کشف کند که کدام اقدامات بیشترین پاداش را می دهند. در مواردی که جالب و چالش برانگیز هستند، اقدامات ممکن است نه تنها پاداش فوری، بلکه موقعیت بعدی و از طریق آن تمام پاداش های بعدی را نیز تحت تأثیر قرار دهند. این دو ویژگی - جستجو و خطا و پاداش تأخیری - دو ویژگی مهم متمایز یادگیری تقویتی هستند.

یادگیری تقویتی با مشخص کردن الگوریتم های یادگیری تعریف نمی‌شود، بلکه با مشخص کردن یک مسئله یادگیری تعریف می شود. هر الگوریتمی که برای حل آن مشکل مناسب باشد را یک الگوریتم یادگیری تقویتی می‌دانیم. ایده اصلی این است که مهمترین جنبه های مشکل واقعی را با استفاده از تعامل با محیط خود که خود یک عامل یادگیری است، برای دستیابی به هدف مشخصی به کار گیریم. چنین عاملی باید تا حدی بتواند وضعیت محیط را حس کند و باید بتواند اقداماتی انجام دهد که بر آن وضعیت تأثیر بگذارد. عامل همچنین باید یک هدف یا اهدافی مربوط به وضعیت محیط داشته باشد. [7]

روش‌های مختلفی از یادگیری ماشینی ارائه شده که بیان هر یک از آن‌ها می‌پردازیم.

## بررسی روش‌های مختلف یادگیری تقویتی

انواع مختلف از الگوریتم‌ها و رویکردها در تقویت یادگیری وجود دارد، از جمله:

### تقویت یادگیری بدون مدل

یادگیری تقویتی بدون مدل[[22]](#footnote-22) یک روش در حوزه یادگیری ماشین است که برای حل مسائل تصمیم‌گیری در محیط‌های تعاملی استفاده می‌شود. در این نوع از یادگیری تقویتی، ما به جای استفاده از یک مدل دقیق از محیط برای پیش‌بینی آینده، تلاش می‌کنیم تا از تجربیات مستقیم با محیط استفاده کنیم تا یک استراتژی بهینه برای تصمیم‌گیری در محیط توسعه دهیم.

در یادگیری تقویتی بدون مدل، یک عامل در یک محیط قرار دارد و در هر مرحله از زمان، عامل اقدامی انجام می‌دهد و از محیط پاداشی دریافت می‌کند. هدف این عامل این است که استراتژی بهینه‌ای را کشف کند که به ماکسیمم کردن جمع ارزش (مجموع پاداش‌ها در طول زمان) منجر می‌شود.

روش‌های یادگیری تقویتی بدون مدل شامل الگوریتم‌هایی مانند Q-Learning، Sarsa و AC3 می‌شوند. این الگوریتم‌ها بر اساس تجربیات جمع‌آوری شده در محیط، توابعی مانند تابع ارزش یا تابع عمل را برای انتخاب عمل بهینه یا تخمین ارزش وضعیت‌ها و عمل‌ها آموزش می‌دهند.

مزیت اصلی یادگیری تقویتی بدون مدل این است که به عامل اجازه می‌دهد تا در محیط‌های پیچیده و غیرقطعی عمل کند و به صورت تجربی به یادگیری بپردازد، بدون نیاز به دانستن مدل دقیقی از محیط. این الگوریتم‌ها معمولاً با استفاده از تکنیک‌هایی مانند تابع‌های مشتق‌گیری و نمونه‌برداری از تجربیات، به عامل اجازه می‌دهند تا به تدریج استراتژی بهینه‌ای را کشف کند و بهبودش دهد.

#### یادگیری Q

یادگیری [[23]](#footnote-23)Qیکی از الگوریتم‌های مهم در حوزه یادگیری تقویتی است که برای یادگیری یک استراتژی بهینه برای تصمیم‌گیری در محیط‌های تعاملی استفاده می‌شود. این الگوریتم به ویژه برای مسائلی که محیط آن‌ها تقریباً قطعی[[24]](#footnote-24) استفاده می‌شود، اما می‌توان آن را به محیط‌های غیرقطعی [[25]](#footnote-25)نیز تعمیم داد.

در یادگیری Q، عامل یک جدول به نام "Q-Table" را تعریف می‌کند. این جدول دارای ابعادی معادل تعداد وضعیت‌ها و تعداد عمل‌ها در محیط است. هر خانه از این جدول مقداری به نام Q-value دارد که نشان‌دهنده ارزش ترکیبی از وضعیت و عمل است.

عامل به صورت تجربی و در طول زمان اقداماتی در محیط انجام می‌دهد و پاداش‌های دریافتی را مشاهده می‌کند. با استفاده از این تجربیات، عامل مقادیر Q-value را به روز می‌کند تا بهبود استراتژی خود برای انتخاب عمل‌ها را دست‌یابی کند.

در این فرمول:

* Q(s, a) : برای وضعیت s و عمل a.
* α (نرخ یادگیری) : یک پارامتر که نشان‌دهنده میزان اهمیت تجربیات جدید است و می‌تواند بین 0 و 1 باشد.
* R(s, a) : پاداش دریافتی برای انجام عمل a در وضعیت s.
* γ (عامل تخفیف): یک پارامتر که نشان‌دهنده تاثیر پاداش‌های آینده است و می‌تواند بین 0 و 1 باشد.
* max(Q(s', a')) : بیشترین Q-value برای وضعیت بعدی s' و همه عمل‌های ممکن در آن وضعیت.
* Q(s, a) : مقدار Q-value قبلی برای وضعیت s و عمل a.

عامل با استفاده از این فرمول مقدار Q-value را به روزرسانی کرده و به تدریج استراتژی بهینه‌ای برای تصمیم‌گیری در محیط را یاد می‌گیرد. این فرآیند ادامه پیدا می‌کند تا مقادیر Q-value به تقریب به مقادیر بهینه برای استراتژی بهینه نزدیک شوند. یادگیری Q معمولاً به عنوان یک روش مدل‌نسبی برای یادگیری تقویتی بدون مدل شناخته می‌شود، زیرا بدون نیاز به دانش دقیق از مدل محیط، عامل می‌تواند به تدریج استراتژی بهینه‌ای را کشف کند. [8]

#### SARSA - State-Action-Reward-State-Action

الگوریتم Sarsa یکی از الگوریتم‌های یادگیری تقویتی است که برای یادگیری یک استراتژی بهینه برای تصمیم‌گیری در محیط‌های تعاملی استفاده می‌شود. نام "Sarsa" از ترکیب اجتماعی اجزای اصلی این الگوریتم گرفته شده است:

S: State

وضعیت فعلی

A: Action

عمل انجام شده

R: Reward

پاداش دریافتی در ازای عمل

S’: State

وضعیت بعدی

A’: Action

عمل بعدی

الگوریتم Sarsa به صورت یک الگوریتم یادگیری Q-value عمل مدل می‌شود. عامل در هر مرحله از زمان، در وضعیت فعلی قرار دارد و یک عمل را انتخاب می‌کند. سپس عمل انتخاب شده را انجام می‌دهد، پاداشی را دریافت می‌کند و به وضعیت بعدی منتقل می‌شود. سپس عامل عمل بعدی را انتخاب می‌کند و فرآیند به این ترتیب ادامه می‌یابد.

روند آموزش الگوریتم Sarsa به این صورت است:

* 1. مقادیر اولیه برای Q-value ها (Q(s, a)) مقداردهی می‌شود.
  2. عامل در وضعیت فعلی (S) قرار دارد و یک عمل (A) را انتخاب می‌کند با استفاده از یک روش معین مانند ε-greedy که امکان انتخاب تصادفی عمل را هم فراهم می‌کند.
  3. عامل عمل انتخاب شده را انجام داده و پاداش (R) را دریافت می‌کند و به وضعیت بعدی (S') منتقل می‌شود.
  4. سپس عامل در وضعیت بعدی (S') عمل بعدی (A') را بر اساس سیاست خود انتخاب می‌کند.
  5. مقدار Q-value برای جفت وضعیت فعلی و عمل فعلی (Q(S, A)) با استفاده از فرمول به روزرسانی زیر به‌روز می‌شود:

که نمادها مانند الگوریتم یادگیریQ می‌باشد.

1. مراحل 2 تا 5 به صورت تکراری ادامه می‌یابد تا عامل یک استراتژی بهینه برای تصمیم‌گیری در محیط یاد بگیرد.

الگوریتم Sarsa به عنوان یک روش مدلی برای یادگیری تقویتی بدون مدل شناخته می‌شود، زیرا به تدریج تخمین بهینه Q-value ها را در محیط انجام می‌دهد و استراتژی بهینه را کشف می‌کند. [9]

#### روش‌های تقویت سیاست:

این روش‌ها به‌صورت مستقیم یک خط مشی که وضعیت‌ها را به اعمال نگاشت می‌کند، یاد می‌گیرند، به‌جای تخمین توابع ارزش. به عنوان مثال می‌توان به روش REINFORCE و روش‌های بازیگر-انتقادی (Actor-Critic) اشاره کرد.

از آن‌جا که این مدل یادگیری بنا به ذات گسسته خود خود با مسئله ما مطابقت ندارد بیش از این به آن نمی‌پردازیم.

### یادگیری تقویتی با مدل

یادگیری تقویتی با مدل[[26]](#footnote-26) یکی از روش‌های یادگیری تقویتی است که در آن یک مدل داخلی از محیط به وسیله عامل ساخته می‌شود و از این مدل برای انجام تصمیم‌گیری‌ها و یادگیری سیاست بهینه استفاده می‌شود. این روش تفاوت‌های مهمی با روش‌های دیگر یادگیری تقویتی دارد. مانند یادگیری تقویتی مبتنی بر خود در یادگیری تقویتی مبتنی بر مدل، مراحل کلی به شرح زیر هستند:

1. تجمیع داده: ابتدا، عامل با تعامل با محیط داده‌های تجربی جمع‌آوری می‌کند. این داده‌ها شامل اطلاعاتی از تجربیات عامل در محیط و نتایج اعمال اقدامات مختلف می‌شوند.

2. ساخت مدل: در مرحله دوم، عامل یک مدل داخلی از محیط ایجاد می‌کند. این مدل معمولاً یک مدل پیش‌بینی[[27]](#footnote-27) نامیده می‌شود. این مدل به عامل اجازه می‌دهد تا پیش‌بینی‌هایی از نتایج اقدامات در محیط انجام دهد.

3. تولید سیاست: با استفاده از مدل داخلی محیط، عامل می‌تواند تعدادی تصمیم را به صورت آزمایشی تولید کند و ارزیابی کند که کدام یک از اقدامات بهینه‌ترین نتیجه را تولید می‌کند. این فرآیند معمولاً با استفاده از الگوریتم‌های بهینه‌سازی مانند الگوریتم‌های بهینه‌سازی تقریباً مطلق انجام می‌شود.

4. آموزش مدل: در مرحله آموزش مدل داخلی، عامل از داده‌های تجربی‌اش استفاده می‌کند تا مدل را بهبود دهد. این به معنای به روزرسانی پارامترهای مدل توسط الگوریتم‌های یادگیری مدل می‌باشد. این به عامل این امکان را می‌دهد که مدل داخلی خود را بهبود داده و تقریب بهتری از محیط ایجاد کند.

مزیت اصلی یادگیری تقویتی مبتنی بر مدل این است که به عامل امکان پیش‌بینی دقیق‌تری از تاثیر اقدامات خود در محیط می‌دهد و این اطلاعات به عامل کمک می‌کند تا به سرعت یادگیری کند. با این حال، این روش نیازمند تولید و نگهداری یک مدل داخلی دقیق و معتبر از محیط است که ممکن است در بعضی موارد دشوار باشد. همچنین، نیاز به جمع‌آوری داده‌های تجربی از محیط نیز ممکن است هزینه و زمان بر باشد.

#### جستجوی درخت مونت کارلو[[28]](#footnote-28)

یک الگوریتم تقویت یادگیری با مدل است که برای انجام تصمیمات در محیط‌هایی که عامل دارای میزان محدودی از اطلاعات است، استفاده می‌شود. این الگوریتم یک درخت از اقدامات ممکن و نتایج آنها را برای انتخاب تصمیمات ساخته و به‌روز می‌کند.

#### کنترل پیش‌بینی مدل[[29]](#footnote-29)

از یک مدل یادگیری شده یا شبیه‌سازی شده از محیط برای پیش‌بینی وضعیت‌های آینده و بهینه‌سازی خط‌مشی کنترلی در یک بازه زمانی محدود استفاده می‌کند.

#### Dyna Q

همزمان از رویکردهای تقویت یادگیری بدون مدل و تقویت یادگیری با مدل استفاده می‌کند، با نگه‌داری از یک مدل از محیط برای شبیه‌سازی تجارب و به‌روزرسانی تابع ارزش.

یادگیری تقویتی با مدل معمولاً به مدل‌سازی محیط و تصمیم‌گیری در محیط‌های نامعین[[30]](#footnote-30) توانمندی دارد. اگرچه لازم به ذکر است که این توانمندی باتوجه به شرایط محیط ممکن است متفاوت باشد. عموما برای محیط‌های نامعین استفاده از این مدل از یادگیرنده‌ها نیازمند تلاش بیشتری است تا نتایج به حد مطلوب برسد.

### یادگیری تقلیدی[[31]](#footnote-31)

یادگیری تقویتی تقلیدی یک رویکرد در یادگیری ماشینی است که در آن عامل یادگیرنده سعی می‌کند عملکردی مشابه با عامل دیگری که به عنوان مرجع متخصص شناخته می‌شود، را تقلید کند. در این روش، عامل یادگیرنده نمونه‌هایی از عملکرد مرجع را مشاهده می‌کند و تلاش می‌کند تا این عملکرد را با تکرار تجربه و تمرین بهبود دهد.

مراحل یادگیری تقویتی تقلیدی عموماً به شرح زیر انجام می‌شود:

1. جمع‌آوری داده‌های آموزشی: در این مرحله، داده‌هایی که توسط مرجع یا اکسپرت جمع‌آوری شده است به عامل یادگیرنده داده می‌شود. این داده‌ها معمولاً شامل توالی‌هایی از وضعیت‌ها و اقدامات (عملکرد مرجع) می‌شوند.

2. مدل سازی: در این مرحله، یک مدل یادگیری عمیق یا شبکه عصبی معمولاً به عنوان شبکه تقلیدی یجاد می‌شود. این شبکه به وضعیت فعلی عامل ورودی داده شده و سعی می‌کند عملی را تولید کند که به عملکرد مرجع نزدیک باشد.

3. آموزش شبکه تقلیدی: شبکه تقلیدی با استفاده از داده‌های آموزشی، بهبود عملکرد خود را ادامه می‌دهد. این بهبود معمولاً توسط الگوریتم‌های یادگیری نظارت می‌شود.

4. ارزیابی عملکرد: پس از آموزش، عامل یادگیرنده باید عملکرد خود را ارزیابی کند. این ارزیابی معمولاً با مقایسه عملکرد عامل یادگیرنده با عملکرد مرجع انجام می‌شود.

5. تنظیم و بهبود: اگر عملکرد عامل یادگیرنده هنوز نسبت به مرجع کافی نباشد، ممکن است مراحل 2 تا 4 چندین بار تکرار شود تا عامل به تقلید بهتری از مرجع برسد.

این روش به عنوان یکی از روش‌های ابتدایی در یادگیری تقویتی به کار می‌رود و به ویژه در مواردی که مشکل یادگیری تقویتی از ابتدا بدون مرجع مشخص و دقیق باشد، مفید است.

در این پروژه متخصصی وجود ندارد که از او نحوه عملکرد مطلوب را جویا بشیم بنابراین این نوع یادگیری مطلوب نیست و حذف می‌شود

### یادگیری تقویتی عمیق

یادگیری تقویتی عمیق[[32]](#footnote-32) یک شاخه از یادگیری ماشین است که به ترکیب دو عنصر اصلی، یعنی یادگیری تقویتی و شبکه‌های عصبی عمیق می‌پردازد. این روش برای حل مسائلی که تصمیم‌های توالی بر اساس عوامل متاثر از تصمیم‌ها نیاز دارند، مورد استفاده قرار می‌گیرد.

شبکه‌های عصبی عمیق[[33]](#footnote-33) نیز برای تقریب تابع ارزش عمل‌ها و تصمیم‌گیری عامل در یادگیری تقویتی استفاده می‌شوند. این شبکه‌ها به عنوان تقریب‌گرهای توابع ارزش عمل‌ها (Q-Values) عمل می‌کنند و عامل با استفاده از آن‌ها می‌تواند بهترین عمل را در هر وضعیت انتخاب کند.

یکی از مطرح‌ترین الگوریتم‌های یادگیری تقویتی عمیق، الگوریتم شبکه Q عمیق است که توسط Google DeepMind معرفی شد و با استفاده از یک شبکه عصبی عمیق، عامل را آموزش می‌دهد تا عمل‌های بهتری انجام دهد. همچنین، الگوریتم‌هایی مانند عملگر-منتقد نیز برای یادگیری تقویتی عمیق مورد استفاده قرار می‌گیرند. [8]

یادگیری تقویتی عمیق در بسیاری از حوزه‌ها مانند بازی‌های ویدئویی، رباتیک، مدیریت منابع، بهینه‌سازی مالی، و زمینه‌های دیگر به کار می‌رود و به عنوان یکی از روش‌های پیشرو در یادگیری ماشین به شمار می‌آید.

## استراتژی‌های کاوش

در یادگیری تقویتی، کاوش[[34]](#footnote-34) و بهره‌برداری[[35]](#footnote-35) دو مفهوم مهمی هستند که در فرآیند تصمیم‌گیری و یادگیری تقویتی نقش دارند. این دو مفهوم به شکل متعادلی با یکدیگر ترکیب می‌شوند تا بهترین عملکرد را به دست آورند. در ادامه به توضیح هر یک از این مفاهیم می‌پردازیم:

1. کاوش

- کاوش به معنای انجام اقداماتی است که به شناخت محیط یادگیری و کشف اطلاعات جدید کمک می‌کند. در یادگیری تقویتی، انجام کاوش به منظور یافتن تجربیات جدید و تخمین ارزش‌های مختلف عملکردها استفاده می‌شود.

- کاوش می‌تواند با انتخاب عملکردهای تصادفی یا آزمون عملکردهای جدید انجام شود. این کار به افزایش دانش مدل در مورد محیط و اطلاعات مربوط به آن کمک می‌کند.

2. بهره‌برداری

- بهره‌برداری به معنای انجام اقداماتی است که بر اساس اطلاعاتی که تا کنون جمع آوری شده است، بهترین عملکرد را انتخاب می‌کند. در واقعیت، این به انجام عملکردهایی که تا کنون به عنوان موثرترین شناخته شده‌اند و باعث کسب مزیت‌های بیشتر می‌شوند، اشاره دارد.

- بهره‌برداری به منظور بهره‌گیری از تجربیات قبلی و بهبود عملکرد با استفاده از اطلاعات به دست آمده از مصالحه انجام می‌شود.

تعادل میان کاوش و بهره‌برداری در یادگیری تقویتی مهم است. اگر تنها مصالحه انجام شود، ممکن است مدل به طور مداوم عملکردهای تصادفی انجام دهد و عملکرد بهبود نیابد. از طرف دیگر، اگر تنها بهره‌برداری صورت گیرد، ممکن است اطلاعات جدید در مورد محیط جمع آوری نشود و عملکرد به حداکثر خود نرسد. بنابراین، تعادل مناسب بین این دو عنصر باید در الگوریتم‌های یادگیری تقویتی حفظ شود تا بهترین عملکرد ممکن به دست آید. [9]

اپسیلون[[36]](#footnote-36) در الگوریتم‌های یادگیری تقویتی به عنوان یک پارامتر مهم و تصمیم‌گیری در روشهای تصادفی مورد استفاده قرار می‌گیرد. اپسیلون معمولاً به عنوان مقداری بین صفر و یک تعریف می‌شود و نمایانگر نسبتی از زمان‌هاست که عامل یادگیری در انتخاب عمل تصادفی عمل می‌کند به جای انتخاب بهترین عمل بر اساس استراتژی یادگیری تصمیم‌گیری یا همان سیاست خود.

به عبارت دیگر، اگر مقدار اپسیلون برابر با صفر باشد، عامل همیشه عمل بهترین تخمین‌گر برای عمل بهره‌برداری را انجام می‌دهد و هیچ عمل تصادفی انجام نمی‌دهد. اما اگر اپسیلون برابر با یک باشد، عامل همیشه تصادفی عمل می‌کند و هیچ‌گاه به عمل بهینه تخمین‌گر نمی‌پردازد و کاوش می‌کند. در حالت عمومی، اپسیلون می‌تواند یک مقدار میانی بین صفر و یک داشته باشد، که عامل به تصادف عمل می‌کند با احتمال اپسیلون و به تمام تصمیمات اشتباه از تصمیمات تخمین‌گر بهینه دوری می‌کند.

استفاده از اپسیلون در الگوریتم‌های یادگیری تقویتی به عامل این امکان را می‌دهد که در طول زمان از تجربیات خود بیاموزد و به بهبود عملکرد خود برسد، در حالی که همچنان از تصمیمات بهینه‌ای که تاکنون یاد گرفته است، استفاده کند. این تعادل بین اکتشاف و بهره‌برداری به عامل اجازه می‌دهد که به دنبال جستجوی بهترین راه حل‌ها باشد و به دست آوردن تجربیات جدید را نیز تشویق کند.

### اپسیلون-حریص

الگوریتم اپسیلون-حریص[[37]](#footnote-37) یکی از روش‌های استفاده شده در مسائل تقسیم منابع برای تجربه کردن و تست کردن چندین گزینه است. این الگوریتم به صورت معمول در مسائلی مورد استفاده قرار می‌گیرد که می‌خواهیم بین چندین گزینه یا عمل مختلف انتخاب کنیم و به طور متوسط بهترین گزینه را پیدا کنیم.

### مرز بالای اعتماد[[38]](#footnote-38)

این مفهوم به تعیین یک سیاست تصمیم‌گیری برای یک عامل یادگیری تقویتی اشاره دارد که برای انتخاب اعمال (actions) مختلف در یک محیط تعاملی استفاده می‌شود.

الگوریتم مرز بالای اعتماد معمولاً در مسائلی به کار می‌رود که به عامل امکان برخورد با شرایط مختلف در محیط را می‌دهند و او نیاز به بیشترین بهره‌برداری از اعمال بازدهی بالا دارد.

مرز بالای اطمینان، برای هر عمل، یک مقدار مرز بالای اطمینان بر اساس داده‌های تاکنون جمع‌آوری شده محاسبه می‌شود. این مقدار نمایانگر اطمینان ما از بازدهی این عمل است و معمولاً با استفاده از اندازه نمونه‌ها و مقدارهای آماری مرتبط محاسبه می‌شود.

با توجه به مقدار مرز بالای اطمینان برای هر عمل، عامل یادگیری تقویتی انتخاب می‌کند که کدام عمل را انجام دهد. این استراتژی کمک می‌کند تا عامل به صورت تدریجی عمل‌های با بازدهی بالاتر را ترجیح دهد و در عین حال اطمینان از صحت تخمین‌هایش را حفظ کند.

الگوریتم مرز بالای اعتماد به طور گسترده در مسائل یادگیری تقویتی مورد استفاده قرار می‌گیرد و بهبود عملکرد عوامل در محیط‌های تعاملی را تسهیل می‌کند.

### نمونه‌برداری تامپسون[[39]](#footnote-39)

نمونه‌برداری تامپسون یک الگوریتم معروف در حوزه یادگیری تقویتی است که برای مسائل تصمیم‌گیری تعاملی با استفاده از اصول بیزی مورد استفاده قرار می‌گیرد. این الگوریتم به عامل یادگیری تقویتی امکان می‌دهد تا بر اساس تخمین توزیع احتمالی مقادیر بازدهی عمل‌ها تصمیم بگیرد.

عملکرد الگوریتم نمونه‌برداری تامپسون به صورت زیر است:

1. هر عمل ممکن در مسئله با یک توزیع احتمالی مرتبط است که نشان‌دهنده توزیع بازدهی آن عمل می‌باشد.
2. الگوریتم نمونه‌برداری تامپسون برای هر عمل یک تخمین از توزیع احتمالی بازدهی آن عمل می‌سازد. این تخمین با استفاده از تحلیل بیزی و استفاده از تاریخچه‌ی بازدهی‌ها به‌روز می‌شود.
3. در هر مرحله، الگوریتم یک نمونه تصادفی از توزیع احتمالی بازدهی هر عمل انتخاب می‌کند. این نمونه‌برداری تصادفی باعث می‌شود که عملی با بیشترین احتمال بازدهی بالا انتخاب شود.
4. عملی که در مرحله 3 انتخاب شده است، انجام می‌شود و بازدهی آن به عامل یادگیری اطلاع داده می‌شود.
5. تخمین‌های بیزی برای توزیع‌های بازدهی عمل‌ها بر اساس داده‌های جدید به‌روزرسانی می‌شود.

این چرخه تکرار می‌شود و الگوریتم تامپسون ادامه می‌دهد تا تاخیر زمانی مشخصی یا تا رسیدن به یک تعداد معین از مراحل. این الگوریتم به عامل این امکان را می‌دهد که به طور آگاهانه و با توجه به تخمین‌های بیزی، تصمیم‌های بهینه‌تری در مورد انتخاب عمل‌ها بگیرد.

الگوریتم تامپسون معمولاً در مسائلی که توزیع بازدهی عمل‌ها تغییر پذیر است یا در مسائلی که نیاز به کاوش و بهره‌برداری همزمان دارید مورد استفاده قرار می‌گیرد و به عنوان یکی از الگوریتم‌های اصولی یادگیری تقویتی به شمار می‌آید.

## جمع­بندی

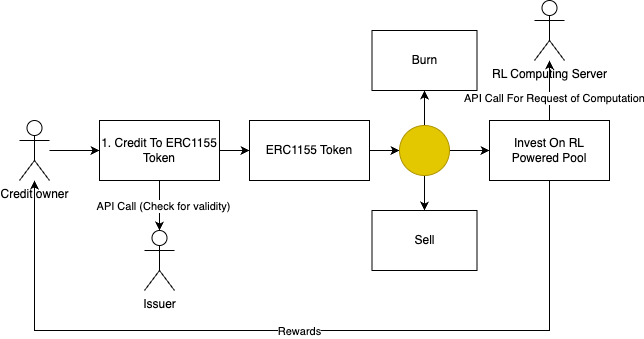
الگوریتم‌های یادگیری تقویتی را می‌توان به طور عمده به دسته‌های یادگیری مدل‌دار و یادگیری بدون مدل تقسیم‌بندی کرد. الگوریتم‌های مدل‌دار سعی در ترسیم یک مدل از محیطی که با آن در تعامل هستند و پویایی آن‌ها بپردازند در حالی که در سیستم استخر سرمایه‌گذاری قیمت ارز خروجی و هم‌چنین نرخ سرمایه‌گذاری کمتر دارای مدل مشخصی هستند. بنابراین استفاده از یک مدل یادگیری بدون مدل راهکار مناسب‌تری است. هم‌چنین در این مسئله بایستی بتوانیم که با مدل مناسبی بتوانیم که مصالحه کاوش و بهره‌برداری را حل نماییم. اگر که راهکارهای موجود نگاهی بیندازیم می‌بینیم که الگوریتم نمونه‌برداری تامپسون بهترین راهکار برای این موضوع است زیرا اگر که از سایر الگوریتم‌ها استفاده کنیم ممکن است که عامل رفتارش بعد از مدتی ثابت شود ولی قیمت و میزان سرمایه‌گذاری این‌چنین نیست. پس نیاز داریم که با راهکاری که هم‌زمان امکان کاوش و بهره‌برداری را ارائه می‌دهد پیش برویم.

# فصل چهارم روش پیشنهادی و روش پیاده‌سازی

## مقدمه

در این فصل به نحوه پیاده‌سازی پل تبدیل اعتبارات کربنی به رمزارزها و هم‌چنین نحوه پیاده‌سازی استخر تبدیل اعتبارات کربنی رمزارزی به دست آمده به رمزارزهای با نقدینگی بالا می‌پردازیم. هم‌چنین بایستی بیان کنیم که چگونه در این سیستم به ایمنی لازم دسترسی پیدا کرده‌ایم. ایمنی در قراردادهای هوشمند در بستر بلاک‌چین دارای اهمیت بالایی است زیرا در صورتی که یک قرارداد در محیط بلاک‌چین مستقر شود، دیگر امکان تغییر آن وجود ندارد و هم‌چنین این قراردادها با پول سر و کار دارند، اگر که امکان حمله به آن‌ها وجود داشته باشد دارایی از دست خواهد رفت. هم‌چنین لازم به ذکر است که یادگیری تقویتی بار پردازشی زیادی لازم دارد و هم‌چنین لازم به حافظه زیادی است که سیستم بلاک‌چین این امکان را به ذاتا در اختیار ما نمی‌گذارد.

## نحوه پیاده‌سازی بخش تبدیل اعتبار کربن به رمزارز



تصویر - فرایند سرمایه گذاری

فرایند سرمایه‌گذاری که در تصویر ۴ به نمایش گذاشته شده است مطابق زیر است:

1. ابتدا صاحب اعتبار کربن باستی در سازمانی که از آن اعتبار کربن خود را گرفته درخواست تبدیل آن به نوع رمزارزی را بدهد.
2. در صورت پذیرش سازمان نام‌برده کاربر از طریق قرارداد هوشمند یک توکن ERC1155 خالی تولید نموده و با استفاده از آن درخواستی به سازمان تولید کننده اعتبار می‌زند و به این صورت از طریق یک تماس API از طریق یک Oracle می‌توان اعتبار را به صورت رمزارزی با قالب توکن ERC1155 در آورد.
3. صاحب اعتبار می‌تواند با استفاده از توکن ERC1155 که معادل فیزیکی اعتبار کربن است، اقدامات متنوعی بکند از جمله:
   1. فروش
   2. سوزاندن
   3. سرمایه‌گذاری
4. در صورتی که راه آخر یعنی سرمایه‌گذاری را انتخاب کند. با استفاده از یک تماس Oracle به سروری مجهز به یادگیرنده تقویتی سرور از سرمایه‌گذاری شخص سرمایه‌گذار در استخر مربوطه مطلع می‌شود. پس در نتیجه شروع به محاسبه تدریجی نرخ سود او می‌کند.
5. کاربر هر موقع که بخواهد می‌تواند با درخواست از طریق قرارداد هوشمند توکن‌های قابل تعویض ERC20 به ازای مدت سرمایه‌گذاری در صندوق دریافت کند.

## توکن‌های ERC1155 با یک‌دیگر متفاوت هستند

یکی از مزایای پروتکل ERC1155 این است که هم‌زمان می‌تواند هم قابل تعویض و غیرقابل تعویض باشد. در این پروژه توکن‌های ERC1155 از ۴ نوع متفاوت هستند که به صورت زیر می‌باشد.

1. نوع استاندارد: سازمان صادر کننده استاندارد
2. نوع جبران کربن: نحوه مقابله با گرمایش زمین مثلا: تولید انرژی پاک، ممانعت از جنگل‌زدایی، جنگل‌زایی
3. حجم جبران کربن: میزان کربن جبران شده از سطح زمین
4. سال صدور

## راهکار یادگیری تقویتی برای تبدیل ERC-1155 به ERC-20

تقاضا و عرضه دو مفهوم کلیدی در اقتصاد هستند که نقش مهمی در تعیین قیمت‌ها و تعاملات اقتصادی ایفا می‌کنند. این دو عنصر به طور مستقیم با یکدیگر مرتبط هستند و تأثیرات متقابلی بر روی اقتصاد دارند. در زیر به توضیح هر یک از این مفاهیم می‌پردازم:

1. تقاضا: تقاضا به میزان کالا یا خدماتی اشاره دارد که افراد یا سازمان‌ها در بازار خریداری می‌کنند. تقاضا معمولاً به عواملی مانند قیمت، درآمد، ترجیحات مصرف‌کنندگان، توقعات و شرایط اقتصادی بستگی دارد. افزایش تقاضا ممکن است به افزایش قیمت‌ها منجر شود، به عنوان مثال وقتی که تقاضا بیشتر از عرضه است. تقاضا می‌تواند بر اساس تغییرات فصلی، ترتیب محصولات جدید و ابتکارات صنعتی نیز تغییر کند.

2. عرضه: عرضه به تعداد کالا یا خدماتی اشاره دارد که توسط تولیدکنندگان یا فروشندگان در بازار قرار داده می‌شود. عرضه نیز به عواملی مانند هزینه تولید، تکنولوژی، تغییرات در منابع، واکنش به تغییرات قیمت و سایر عوامل اقتصادی بستگی دارد. افزایش عرضه ممکن است منجر به کاهش قیمت‌ها شود، به عنوان مثال وقتی که عرضه بیشتر از تقاضا است.

تعادل بین تقاضا و عرضه در بازار می‌تواند به تنظیم قیمت‌ها و مقدار تولید مناسب منجر شود. اگر تقاضا بیشتر از عرضه باشد، قیمت‌ها ممکن است افزایش پیدا کنند، و اگر عرضه بیشتر از تقاضا باشد، قیمت‌ها ممکن است کاهش یابد. این فرآیند تعیین قیمت‌ها را به بازار و نیروهای اقتصادی ترک می‌کند. تغییرات در تقاضا و عرضه می‌توانند به دلایل مختلفی از جمله تغییرات اقتصادی، تغییرات در سیاست‌های دولتی و تغییرات در تکنولوژی رخ دهند و تأثیرات گسترده‌ای بر اقتصاد داشته باشند.

یکی از عاومل موثر بر عرضه یک رمزارز، میزان تولید آن است. به طور مثال بیت‌کوین تقریبا هر چهار سال یک بار، یا به عبارت دیگر هر ۲۱۰ هزار بلاک یک‌بار میزان عرضه بیت‌کوین‌ها را نصف می‌کند. این منجر به ایجاد یک تصاعد هندسی در میزان عرضه این رمزارز شده که منجر به این می‌شود که میزان محدودی از این ارز موجود باشد و دچار تورم نشود.

در این‌جا نیز ارزش رمزارز ERC-20 بسته به میزان تولید آن می‌تواند متفاوت باشد. هرچه بیشتر این ارز تولید شود کمیابی آن کمتر شده و از ارزش آن کاسته می‌شود. هرچه که این ارز کمتر تولید شود نیز برعکس.

حال اگر که جای اینکه میزان معینی از رمز ارز ERC-20 خلق کنیم اگر که آن‌ها را در یک استخر سرمایه‌گذاری قرار دهیم که با فرکانس مشخصی به دارنده ERC-1155 معادل اعتبار کربن به آن‌ها ERC-20 می‌دهد، آیا می‌توان هرچقدر که لازم است فرکانس مذکور را کمتر کرد؟‌ به طور حتم خیر زیرا در این صورت دیگر ارز ERC-20 ای تولید نمی‌شود که بخواهد از ارزش آن کم شود. و به طبع اگر که میزان عرضه آن نیز کمتر شود میزان ارز که به دست سرمایه‌گذار می‌رسد ممکن است که مورد پسند او نباشد پس لازم است که در این جا یک مصالحه صورت بگیرد.

یادگیری تقویتی می‌تواند در اینجا مورد استفاده قرار بگیرد. حال به بیان نحوه استفاده از یادگیری تقویتی در این مسئله می‌پردازیم.

یادگیری تقویتی در این جا به صورت زیر تعریف می‌شود:

1. وضعیت: میزان قیمت ارز ERC-20 خروجی و هم‌چنین میزان سرمایه‌گذاری کاربران
2. پاداش: افزایش ارز ERC-20 معادل پاداش مثبت و کاهش آن نیز معادل پاداش منفی است. افزایش میزان سرمایه‌گذاری در استخر نیز پاداش مثبت و کاهش آن نیز پاداش منفی است.
3. عمل: عامل می‌تواند با افزایش یا کاهش فرکانس تولید ERC-20 با محیط تعامل کند.
4. عامل: موجودیتی که میزان تولید ERC-20 را تعیین و با توجه به وضعیت موجود بهینه‌ترین عمل را تشخیص می‌دهد.

## چرا یادگیری تقویتی؟

در بخش قبل گقتیم که عرضه و تقاضا یکی از عوامل با ارزش شدن یک کالا می‌باشد. اما میزان موجودی تنها عامل تاثیرگذار بر قیمت نیست. به طور کمی دقیق‌تر میزان موجودی نقدی است که میزان قیمت را تعیین می‌کند. این عامل تنها در دست میزان تولید کالا نیست. بلکه عملکر مردم در قبال آن کالا نیز بر میزان قیمت آن تاثیرگذار است. اگر که میزان قیمت یک کالا را به فرض محدود به میزان موجودی آن به طور کلی نکنیم، هر چه قیمت آن‌ بالاتر رود تمایل به کسب آن کالا نیز بیشتر می‌شود. در مسئله ما اگر که به دلیلی غیر از عرضه و تقاضی ناشی از میزان تولید، میزان سرمایه‌گذاری جهت کسب رمزارز خروجی استخر سرمایه‌گذاری بیشتر شود، این برای صاحب استخر منفعت دارد. افزایش قیمت نیز می‌تواند ناشی از کاهش موجودیت ناشی از کمتر تولید کردن رمزارز قابل تعویض باشد که خود می‌تواند عاملی برای افزایش سرمایه‌گذاری شود. پس به طور کلی میزان قیمت و میزان سرمایه‌گذاری می‌توانند به طور مستقل عاملی برای کاهش یا افزایش میزان تولید رمزارز خروجی استخر باشند. پس می‌توان به طور خلاصه در جدول زیر گفت:

جدول - عوامل موثر بر میزان تولید رمزارز خروجی استخر

|  |  |
| --- | --- |
| عوامل افزایش میزان تولید | عوامل کاهش میزان تولید |
| افزایش تقاضا به دلیل افزایش قیمت مستل از میزان تولید | افزایش قیمت به دلیل کاهش میزان تولید |
|  | افزایش تقاضا به دلیل افزایش قیمت ناشی از کمتر شدن تولید |

لازم به ذکر است که در اینجا تغییر میزان تولید یعنی عمل عامل بر محیط مسئله یعنی میزان قیمت و سرمایه‌گذاری تاثیر گذار است. پس نیازمند نوعی هوشمندی داریم که بتواند از تاثیر اعمال خود بر محیط آگاه باشد.

## انتخاب مدل یادگیرنده تقویتی

حال با توجه به تمامی این موارد می‌توان گفت که با انتخاب نوع مطلوب مدل یادگیری تقویتی می‌توان مسئله را حل نمود. با توجه به فصل قبل یادگیری تقویتی را می‌توان به دو دسته تقسیم کرد: یادگیری با مدل و یادگیری بدون مدل. در یادگیری با مدل، الگوریتم‌ها سعی می‌کنند یک مدل از محیطی که با آن تعامل دارند بسازند و با پویایی محیط مدل را به‌روز کنند. اما در یادگیری بدون مدل، سیستم‌ها در تصمیم‌گیری‌های خود از مدل خاصی استفاده نمی‌کنند و به جای آن، به صورت مستقیم با محیط تعامل می‌کنند. از این رو، در مواجهه با وظیفه‌هایی مانند پیش‌بینی قیمت ارز و تصمیم‌گیری در مورد فرکانس استخر سرمایه‌گذاری، یادگیری بدون مدل رویکرد بهتری به نظر می‌آید. در این پروژه به دو نوع یادگیرنده بدون مدل می‌پردازیم و بیان می‌کنیم که هر یک چگونه عمل می‌کند و ما چگونه از یادگیرنده‌ها بهره می‌بریم.

### یادگیرنده تقویتی شبکه یادگیری عمیق Q

این یک نوع یادگیرنده بدون مدل است که سعی در کسب یک سیاست برای نحوه عملکرد در شرایط مختلف بدهد. این یادگیرنده این یادگیرنده از ترکیب یادگیرنده Q و شبکه‌های عصبی عمیق به دست می‌آید.

#### یادگیرنده ْQ

این الگوریتم، عامل از یک جدول به نام "Q-Table" استفاده می‌کند که ابعاد آن برابر با تعداد وضعیت‌ها و تعداد عمل‌ها در محیط است. هر خانه در این جدول دارای مقداری به نام Q-value است که نمایانگر ارزش ترکیب وضعیت و عمل است. عامل با انجام عملی در محیط و مشاهده پاداش‌های دریافتی، تجربی اقدامات خود را بهبود می‌دهد. از این تجربیات به منظور بهبود استراتژی انتخاب عمل‌ها استفاده می‌کند و مقادیر Q-value را در جدول به روز می‌کند. این الگوریتم به ویژه برای محیط‌های قطعی مناسب است، اما می‌توان آن را به محیط‌های غیرقطعی نیز اعمال کرد.

این عامل از رابطه زیر برای سیاست‌گذاری‌های خود استفاده می‌کند.

عامل با استفاده از این فرمول مقدار Q-value را به روزرسانی کرده و به تدریج استراتژی بهینه‌ای برای تصمیم‌گیری در محیط را یاد می‌گیرد. این فرآیند ادامه پیدا می‌کند تا به مرور بهینه شود. [8]

#### شبکه یادگیری عمیق

از این این شبکه‌ها برای یادگیری جدول Q در شبکه نورون‌های عصبی استفاده می‌شود. یعنی به طور تدریجی مقادیر Q-Value کسب شده و یادگیرنده عمیق آن‌ها را به مرور یاد می‌گیرد.

### یادگیرنده عمیق SARSA

این یادگیرنده تماما معادل یادگیرنده تقویتی Q است با این تفاوت که به به جای پیدا کردن بهترین عمل ممکن بعد از عمل a از سیاست تا کنون یادگرفته شده برای تعیین آن استفاده می‌کند. یعنی رابطه توصیف‌گر آن به صورت زیر می‌شود.

## تعیین تابع پاداش

همان‌طور که قبل‌تر بیان شد، پاداش و هم‌چنین عِقاب یک یادگیرنده تقویتی با استفاده از میزان افزایش و کاهش میزان قیمت و سرمایه‌گذاری تعریف می‌شود.

حال برای هر یک نیاز به دو تابع داریم تا معیاری عددی برای هریک ارائه کند.

### تابع قیمت

برای تعیین تاثیر قیمت در پاداش از مشتق زمانی آن یعنی آخرین قیمت منهای یک مانده به آخرین قیمت یا به عبارت دیگر از عبارت زیر استفاده کردیم.

### تابع سرمایه‌گذاری

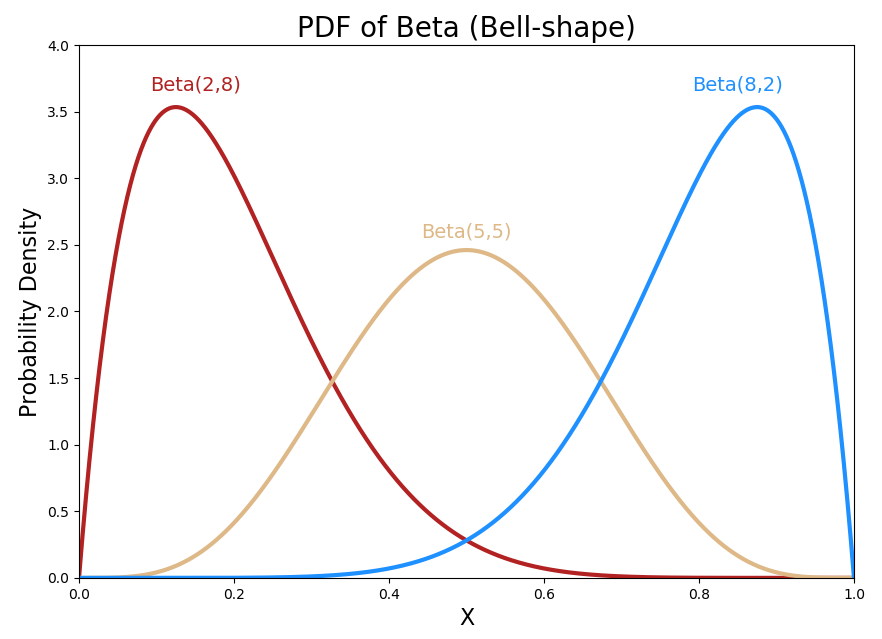
برای تعیین میزان تاثیر سرمایه‌گذاری کمی متفاوت عمل کردیم. به این صورت که مشروط به آنکه دو سرمایه‌گذاری صورت گرفته اخیر فاصله‌ای کمتر از واحد زمانی داشته باشند، پاداش معادل اختلاف آن‌ها باشد. یعنی:

*در نهایت میزان پاداش نهایی جمع وزن‌داری از آن‌ها است. که وزن هر یک در صورت تغییر می‌تواند منجر به تغییر مسئله شود. پس با توجه به نیازمندی‌های خاص مد نظر سرمایه‌گذار می‌توان آن‌ها تعیین کرد.*

## مصالحه کاوش و بهره‌برداری

در این مسئله گاه ممکن است، رفتار مردم بعد مدتی عوض شود. پس انجام عمل مکاشفه بعد از مدتی بهره‌برداری که به خوبی جواب نداده ممکن است لازم باشد. در اینجا یک الگوریتم هیبریدی از نمونه‌برداری تامپسون و الگوریتم اپسیلون-گریدی ارائه می‌دهیم که منجر به عملکرد بهینه شود.

1. الگوریتم اپسیلون-حریص تا جایی که دیگر تغییر نکند ادامه میابد. یعنی به حد پایینی از پیش تعیین شده برسد اجرا می‌شود.
2. سپس الگوریتم نمونه برداری تامپسون اجرا می‌شود. به این صورت که علامت آخرین پاداش منهای پاداش قبلی حاکی از عملکرد خوب یا بد است. میزان عملکرد خوب و بد یک جا ذخیره می‌شود. سپس با استفاده از آن‌ها هنگام تصمیم در مورد کاوش یا بهره‌برداری یک متغیر تصادفی از توزیع بتا تولید می‌شود. این توزیع دارای فرمول چگالی زیر و شکل زیر نیز برخی از مقادیر آلفا و بتا را برای آن نشان می‌دهد.



تصویر - توزیع بتا منبع وبسایت Towards Data Science

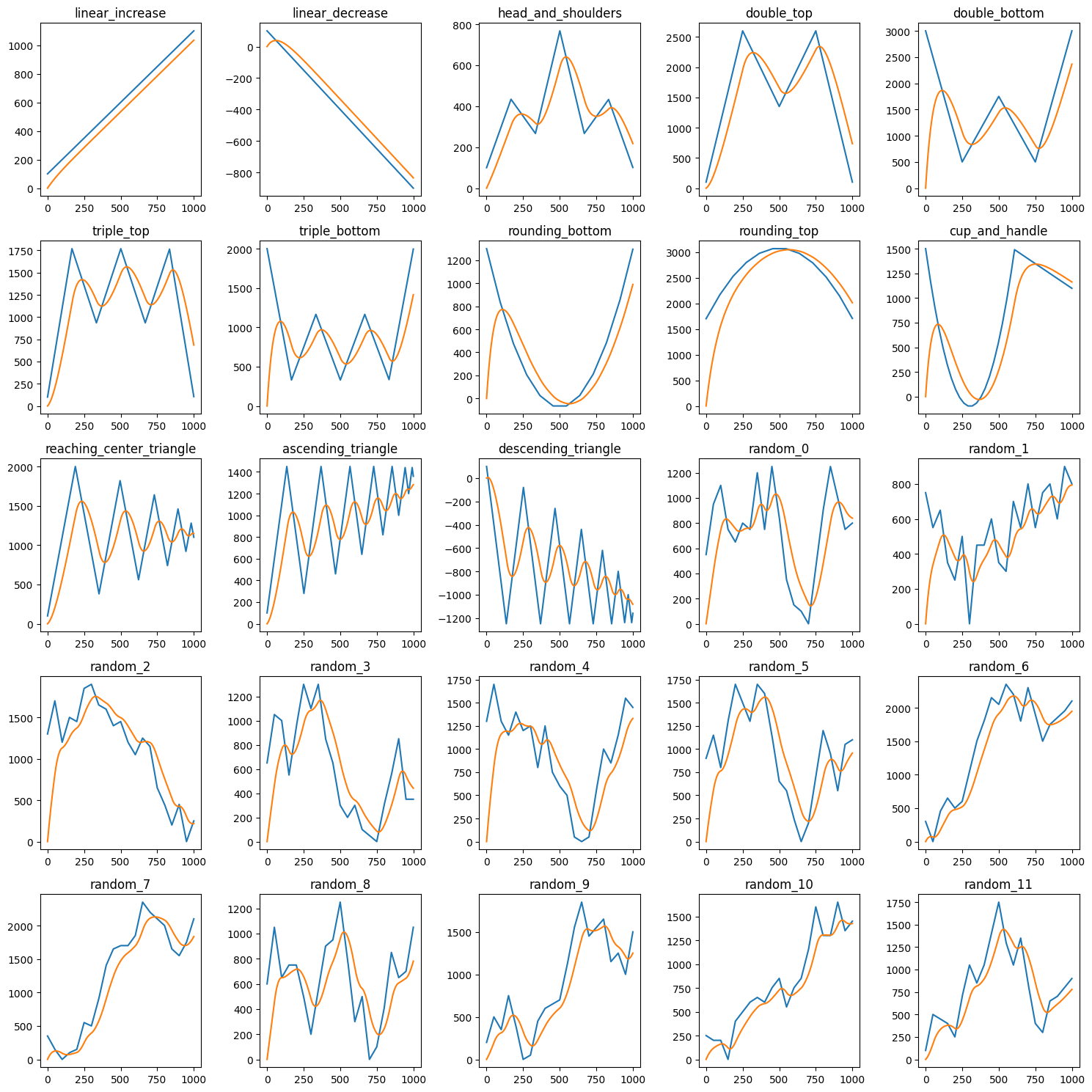
## شبیه‌سازی

از آن‌جا که اصل کار یادگیرنده‌های تقویتی در محیط تعریف می‌شود استفاده از داده‌های از پیش دست یافته شده برای آن شدنی نیست. هم‌چنین تولید داده‌هایی برای قیمت و میزان سرمایه‌گذاری، تاثیر عمل خود عامل را مد نظر نمی‌گیرد.

راهکار چیست؟ پیاده‌سازی یک محیط شبیه‌سازی که در آن تا حدودی عملکرد مردم با توجه به عملکرد عامل تغییر می‌کند.

برای این منظور ابتدا چند نمودار قیمت که ۱۳ تا از آن‌ها از قواعد الگوهای کلاسیک تحلیل تکنیکال پیروی کرده و ۱۲ تا از آن‌ها از قواعد خاصی پیروی نمی‌کنند تولید می‌کنیم. نمودار میزان فرکانس سرمایه‌گذاری را رشد حرکتی نمایی قیمت در ابتدا فرض می‌کنیم.

پس نمودارها به صورت زیر می‌شوند که در آن‌ها آبی قیمت و زرد میزان سرمایه‌گذاری است.



این‌ها البته داده‌های خام اولیه می‌باشند. سپس در ادامه با یک احتمال و با توجه به میزان تغییرات میزان تولید استخر میزان قیمت به صورت معکوس به تناسب تغییر کاهش یافته، میزان سرمایه‌گذاری نیز به کاهش می‌بیاد. هم‌چنین با افزایش قیمت نیز با احتمالی میزان سرمایه‌گذاری به تناسب تغییر افزایش می‌یابد. به طور خلاصه احتمالات گفته شده در جدول زیر آمده است.

|  |  |
| --- | --- |
| احتمال تغییرقیمت با توجه به میزان تولید (عکس) | 0.02 |
| احتمال تغییر میزان سرمایه‌گذاری با توجه به میزان تولید (مستقیم) | ۰.۰۳ |
| احتمال تغییر میزان میزان سرمایه‌گذاری با توجه به تغییرات قیمت (مستقیم) | ۰.۰۲ |

انتخاب مناسب هر یک از احتمالا گفته شده به نزدیکی بیشتر شبیه‌سازی به واقعیت کمک می‌کند. احتمالا گفته شده با توجه به میزان پایداری تغییرات با استفاده از آزمون و خطاهی متوالی به دست آمده است.

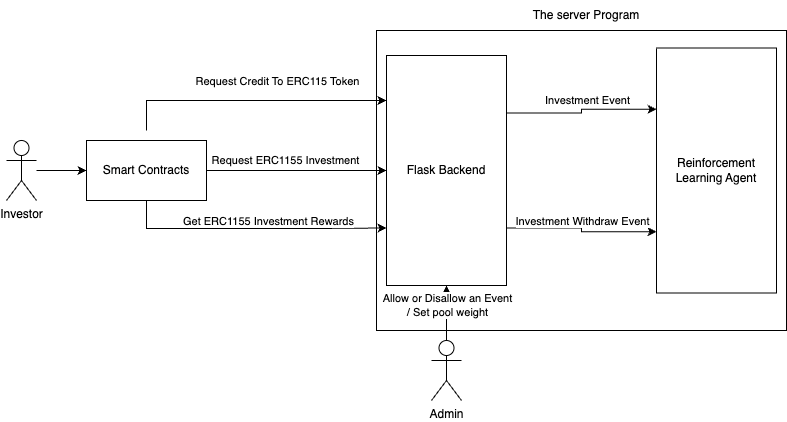
## تست

در نهایت با استفاده از مدل‌های یادگیرنده تقویتی بیان شده و اجرای شبیه‌سازی بر روی آن‌ها برای یادگیری آن‌ها و سپس جمع زدن میزان پاداش‌ها می‌توان مدل‌ها را تست نمود. سه عملگر مورد تست قرار گرفته‌اند که به صورت زیر نتایج تست آن‌ها آمده است.

|  |  |  |
| --- | --- | --- |
| عملگر رندوم | عملگر DQN | عملگر Sarsa |
| 10981.15778705092 | 28328.062903721406 | 25753.4734386721 |

## نحوه پیاده‌سازی

معماری نهایی از حیث پیاده‌سازی به صورت زیر است:



مراحل به صورت زیر است:

1. ابتدا سرمایه‌گذار با استفاده از قرارداد‌های هوشمند درخواست تبدیل اعتبار خود به ERC1155 را صادر می‌کند.
2. این درخواست به سرور Backend می‌رسد. در صورتی که کاربر مدیر در سیستم عمل تبدیل را طبق سیاست‌های خود تایید نموده باشد تبدیل اعتبار به توکن ERC1155 صورت می‌گیرد.
3. کاربر با توکن ERC1155 خود اعمال مختلفی از جمله سرمایه‌گذازی، سوزاندن، فروختن می‌تواند انجام دهد.
4. کاربر با استفاده از توکن خود می‌تواند در سرور درخواست سرمایه‌گذاری به سرور Flask بدهد.

### پیاده‌سازی قراردهای هوشمند

چالش اصلی پیاده‌سازی قراردادهای هوشمند در مصرف کم حافظه و هم‌چنین انجام تماس با بیرون است. برای چالش اول قواعد زیر برای ذخیره‌سازی داده‌ها ارائه شده است.

1. 4 bits -> standard
2. 4 bits -> carbon type
3. 2 byte -> carbon offset
4. 1 byte -> year – 1980
5. با اینکار همه داده‌های مورد نیاز در یک متغیر ۳۲ بیتی جا می‌شوند.
6. برای تماس با بیرون از تماس oracle و محیط تست kovan استفاده شده است.

### پیاده‌سازی سرور backend

این سرور با استفاده از فناوری Flask طراحی شده تا به نیاز‌های قراردهای هوشمند برای اخذ داده و هم‌چنین ارتباط آن‌ها با یادگیرنده تقویتی تعبیه شده است.

### پیاده‌سازی یادگیری تقویتی

در هر یک از مدل‌ها از فناوری PyTorch و numpy و مجموعه Deque از کتاب‌خانه پایتون استفاده شده است.

## جمع‌بندی

می‌توان با استفاده از مدل یادگیری تقویتی به حل مسئله استخر سرمایه‌گذاری رمزارزی پرداخت. شبکه‌های یادگیرنده تقویتی عمیق Q می‌توانند گزینه مناسبی برای حل این دسته مسائل که پیچیدگی‌های فراوانی دارند مورد استفاده قرار بگیرند. هم‌چنین استفاده از الگوریتم پیشنهادی هیبریدی منجر به عملکرد بهینه‌تر یادگیرنده می‌‌شود. برای ارزیابی این مسئله نیاز به یک محیط پویا و تعامل‌دار با یادگیرنده بود. که این محیط ساخته شد. در نهایت با استفاده از تماس‌های Oracle و سرور Flask قراردادهای هوشمند به یادگیرنده و سرویس‌های لازم متصل شدند.

**منابع**

|  |  |
| --- | --- |
| [1] | United Nations، "United Nations Treaties،" 2015. [متصل]. Available: https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg\_no=XXVII-7-d&chapter=27&clang=\_en. |
| [2] | M. G. Kelley Hamrick، "Forest Trends،" 2017. [متصل]. Available: https://www.forest-trends.org/wp-content/uploads/2017/07/doc\_5591.pdf. |
| [3] | Popular Science، 2022. [متصل]. Available: https://www.popsci.com/environment/crypto-carbon-credit-tokens/. |
| [4] | A. R. CHOW، "Time،" 2022. [متصل]. Available: https://time.com/6181907/crypto-carbon-credits/. |
| [5] | Toucan، "Toucan | Carbon Market Infrastructure for climate action،" [متصل]. Available: https://toucan.earth/. |
| [6] | Energy Web Origin، "EW-Origin،" [متصل]. Available: https://energy-web-foundation-origin.readthedocs-hosted.com/en/latest/. |
| [7] | R. Sutton و A. Barto، Reinforcement Learning: An Introduction، MIT Press، 1998. |
| [8] | K. K. D. S. A. G. I. A. D. W. M. R. Volodymyr Mnih، "Playing Atari with Deep Reinforcement Learning،" Arxiv، 2013. |
| [9] | A. G. ‌. Richard S. Sutto، Reinforcement Learning: An Introduction، Bradford Books ، 1992. |

1. Carbon Credit Cryptocurrency [↑](#footnote-ref-1)
2. Token [↑](#footnote-ref-2)
3. Non-Fungible Tokens [↑](#footnote-ref-3)
4. Multi-Fungible Tokens (MFT) [↑](#footnote-ref-4)
5. Reinforcement Leraning [↑](#footnote-ref-5)
6. Voluntary Carbon Market (VCM) [↑](#footnote-ref-6)
7. Know Your Costumer (KYC) [↑](#footnote-ref-7)
8. Tokenization [↑](#footnote-ref-8)
9. Decentralized Finance (Defi) [↑](#footnote-ref-9)
10. Verra [↑](#footnote-ref-10)
11. Internet of Things (IOT) [↑](#footnote-ref-11)
12. Software Development Kit (SDK) [↑](#footnote-ref-12)
13. Backend [↑](#footnote-ref-13)
14. Energy Attribute Certificate (EAC) [↑](#footnote-ref-14)
15. Guarantee of Origin (GO) [↑](#footnote-ref-15)
16. Renewable Energy Certificate (REC) [↑](#footnote-ref-16)
17. International-Renewable Energy Certificate (I-REC) [↑](#footnote-ref-17)
18. Backend [↑](#footnote-ref-18)
19. Registry [↑](#footnote-ref-19)
20. Registry Extended [↑](#footnote-ref-20)
21. Deployment [↑](#footnote-ref-21)
22. Model Free Learning [↑](#footnote-ref-22)
23. Q Learning [↑](#footnote-ref-23)
24. Deterministic [↑](#footnote-ref-24)
25. Stochastic [↑](#footnote-ref-25)
26. Model-Based Reinforcement Learning [↑](#footnote-ref-26)
27. Prediction Model [↑](#footnote-ref-27)
28. Monte Carlo Tree Search - MCTS [↑](#footnote-ref-28)
29. Model Predictive Control – MPC [↑](#footnote-ref-29)
30. Stochastic [↑](#footnote-ref-30)
31. Imitation Learning [↑](#footnote-ref-31)
32. Deep Reinforcement Leraning [↑](#footnote-ref-32)
33. Deep Neural Networks [↑](#footnote-ref-33)
34. Exploration [↑](#footnote-ref-34)
35. Exploitation [↑](#footnote-ref-35)
36. Epsilon [↑](#footnote-ref-36)
37. Epsilon Greedy [↑](#footnote-ref-37)
38. Upper Confidence Bound (UCB) [↑](#footnote-ref-38)
39. Thompson Sampling [↑](#footnote-ref-39)